Title

High Throughput Image Labeling and Lung/Lobar Segmentation on Chest CT using Deep Learning

Permalink

https://escholarship.org/uc/item/4p06c8qq

Author

Wang, Xiaoyong N/A

Publication Date

2019

UNIVERSITY OF CALIFORNIA

Los Angeles

A dissertation submitted in partial satisfaction

of the requirements for the degree Doctor of Philosophy

in Bioengineering

by

Xiaoyong Wang

2019

ABSTRACT OF THE DISSERTATION

High Throughput Image Labeling and Lung/Lobar Segmentation on Chest CT using Deep Learning

by

Xiaoyong Wang
Doctor of Philosophy in Bioengineering
University of California, Los Angeles, 2019
Professor Alex Anh-Tuan Bui, Co-Chair
Professor Matthew Sherman Brown, Co-Chair

Chest CT is the most common modality in thoracic imaging, especially for diagnosis of diffuse lung disease and lung cancer screening. When mining image data from PACS or clinical trials or processing large volumes of data without curation, the relevant scans must be identified among irrelevant or redundant data. Only images acquired with appropriate technical factors, patient positioning, and physiological conditions may be applicable to a particular image processing or machine learning task. Following identification of appropriate images for processing, accurate lung and lobar segmentation is a pre-requisite for the subsequent quantitative image analysis, e.g. air trapping measurement, emphysema scoring, fibrosis scoring, nodule detection, etc. Fully automated segmentation on a diverse spectrum of pathological lungs is still a challenge in clinical practice. Both the labeling and segmentation steps currently require significant manual intervention by image analysts and are prohibitive for large scale processing of big data. The goal of this dissertation is to fully automate the labeling and segmentation tasks in chest CT with high accuracy.

In this dissertation, an image based high throughput labeling pipeline using deep learning was proposed, it aimed to identify anatomical coverage, scan direction, scan posture, lung coverage completeness, contrast usage and breath-hold types. They were posed as different classification problems and some of them required further segmentation and identification of anatomical landmarks. Images of different view planes were used depending on the specific classification problem. All of our models achieved an accuracy > 99% on their respective test sets across different tasks using a research database from multi-center clinical trials. Based on the comprehensive labels from deep learning models, an optimal image series at each time point for a given patient was selected prior to lung and lobar segmentation.

Two fully convolutional networks were proposed to sequentially achieve accurate lung and lobar segmentation. Firstly, a 2D ResNet-101 based network was used for lung segmentation and 575 chest CT scans from multi-center clinical trials were used with radiologist approved lung segmentation. Secondly, a 3D DenseNet based network is applied to segment the 5 lobes and a total of 1280 different CT scans were used with radiologist approved lobar segmentation as ground truth. The dataset included various pathological lung diseases and stratified sampling was used to form the training and test sets with a ratio of 4:1 to ensure a balanced number and type of abnormality were present. Using 5-fold cross validation a mean Dice coefficient of 0.988 ± 0.012 and Average Surface Distance of 0.562 ± 0.49 mm were achieved by the proposed 2D CNN

on lung segmentation. The 3D DenseNet on lobar segmentation achieved a Dice score of 0.959 ± 0.087 and an Average surface distance of 0.873 ± 0.61mm.

The dissertation of Xiaoyong Wang is approved.

Jonathan G Goldin

Denise R Aberle

Dan Ruan

William Hsu

Alex Anh-Tuan Bui, Co-Chair

Matthew S Brown, Co-Chair

University of California, Los Angeles

2019

Dedicated to my beloved family and friends!

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEGEMENTS

I would like to express the utmost appreciation to my advisor, Professor Matthew Brown, for his persistent guidance, inspiration and support throughout my PhD career. I am very grateful that he offered me the opportunity to join the Center for Computer Vision and Imaging Biomarkers (CVIB) as graduate student of bioengineering and I have learned so much from him in this journey of research and discovery in the past 5 years. I am so fortunate to have Dr. Brown as my mentor. I would also like to offer my thanks to my committee chair, Dr. Alex Bui, for his guidance of this dissertation work and other committee members, Dr. Dan Ruan, Dr. William Hsu, Dr. Denise Aberle, and Dr. Jonathan Goldin, for offering insights and dedication for my dissertation.

I would like to thank all my lab members in CVIB, whether past or present, for providing productive, cooperative and friendly environment. I am extremely thankful to Dr. Pechin Lo, who supervised me in study and research in my first two years at UCLA. Dr. Lo has shown me, by his example, what desirable skill sets a computer vision scientist should have. Thanks to Dr. Pangyu Teng and Dr. Mahesh Nagarajan for troubleshooting various problems and providing insightful suggestions in terms of data collection/preparation, algorithm development, and software engineering. Thanks to Mitchell Murphy and his team for supporting IT and infrastructure in the lab. I cannot finish my projects without your help. Thanks to Dr. Jonathan Goldin and Dr. Fereidoun Abtin for continuously providing medical knowledge and quality dataset for my projects. Thanks to Eloisa Rodriguez-Mena for incredible support in administration of the whole lab, so I can fully focus on my research. Thanks to my fellow graduate students Muhammad Wahi-Anwar, Yu Shi, Nastaran Emaminejad, John Hoffman, Anthony

Hardy, Gregory Chu, and Daniel Chong for their help in research and friendship. Thanks to Irene da Costa, Koon-Pong Wong, Judy Ho, Bharath Ramakrishna, Liza Shrestha, and Umamaheswarn Anavarathan for efficiently managing all the projects in CVIB. Thanks to the image analyst team and statistic team for data curation and feedback for my research work.

I would like to thank other PIs of CVIB, Dr. Michael McNitt-Gray, Dr. Grace Kim, and Dr. Scott Hsieh for offering valuable strategies in conducting clinical research.

I would like to thank all my friends who have been with me through my doctoral journey at UCLA. Thanks to UCLA medical center for saving my life before my defense, especially surgeons of trauma team.

Finally, I am eternally grateful to my parents, grandparents and all my family members for your unconditional love, caring, support and understanding. I know you are always there for me through both good and difficult times.

VITA

EDUCATION

2014-2019 Ph.D candidate, Bioengineering, University of California, Los Angeles

2011-2013 M.S. Biomedical Engineering, University of California, Irvine

2007-2011 B.S. Biomedical Engineering, Nanjing University of Aeronautics and Astronautics, China

EXPERIENCE

2014-2019 Graduate Student Researcher, Center for Computer Vision and Imaging Biomarkers, UCLA

2018 *Image Data Scientist Intern (summer), Merck*

2011-2013 Research Assistant, Center for Functional Onco-Imaging, UC Irvine

PUBLICATIONS

- **Xiaoyong Wang**, Pangyu Teng, Pechin Lo, Ashley Banola, Grace Kim, Fereidoun Abtin, Jonathan Goldin, Matthew Brown, “High throughput lung and lobar segmentation by 2D and 3D CNN on chest CT with diffuse lung disease”, MICCAI, 2018
- **Xiaoyong Wang**, Caroline Chou, John Hoffman, Michael McNitt-Gray, Matthew Brown, “Robustness of lung segmentation on ultra-low-dose chest CT using Convolutional Neural Network”, AAPM, 2018
- **Xiaoyong Wang**, Pechin Lo, Matthew Brown. “A machine learning approach for classification of anatomical coverage in CT”. Proc. SPIE 9785, Medical Imaging 2016
- **Xiaoyong Wang**, Pangyu Teng, Ashley Banola, Jonathan Goldin, Matthew Brown, “High throughput image labeling on chest CT by deep learning”, Journal of Medical Imaging (submitted)

Chapter 1: Introduction

1.1 Overview

High resolution computed tomography (HRCT) has been used extensively as the gold standard to diagnose various pulmonary diseases, such as ground glass, emphysema, chronic obstructive pulmonary disease (COPD), scleroderma, pulmonary nodules, etc. This non-invasive modality is ideal for quantitative analysis since it provides volumetric visualization, characterization and quantification of different anatomical structures and lung pathologies. Computer-aided systems are necessary to assist radiologists in performing such quantitative analysis. Firstly the number of radiologists is far less than the increasing number of volumetric datasets produced every day. Manual reading and annotating these high resolution scans is extremely labor intensive and inter-reader variability is inevitable. More importantly, it is challenging for radiologists to provide accurate quantitative measurements due to image complexity and heterogeneity. A large number of quantitative research studies have been performed involving chest CT. Kauczor et al.[1] applied multiple neural networks to automatically detect and quantify ground glass opacities. Compared to density mask based methods, they achieved an improvement of 1.3% in sensitivity and 17.3% in specificity. Wang et al.[2] investigated the optimal threshold to quantify emphysema, ranging from -850 HU to -1000 HU and they concluded that -950 HU is still the most appropriate threshold for density based emphysema quantification although the optimal cut-off for individuals varies. Paired inspiratory and expiratory CT scans were used to determine the attenuation threshold value for the detection and quantification of air trapping and they observed highest correlation with pulmonary function tests (PFTs) using upper threshold of -860 HU[3]. Humphries et al.[4] introduced a CT histogram and texture feature based method to detect and measure the amount of fibrosis and their automatic lung fibrosis quantification correlated well with visual assessment of radiologists. Nakagawa et al.[5] quantified honeycombing on patients diagnosed with idiopathic pulmonary fibrosis (IPF) using a multi-threshold based method. Their findings correlated with radiologist annotations as well as parameters of PFTs. Accurate lung and lobar image segmentation is a precursor to perform almost all of the quantitative analyses described above. Another critical factor that is often overlooked is the identification of appropriate image series on which to perform quantitative image analysis. Image acquisition parameters that are inconsistent with those used during algorithm development can lead to erroneous measurements and/or classifications.

1.2 Motivation

Image data from clinical trials from different sites or from hospital PACS are heterogeneous in terms of acquisition parameters. Even for a single patient, there may be multiple anatomical regions imaged and each may have images reconstructed with different parameters. Since many of them can be redundant, irrelevant, or unusable, it is crucial to select the optimal series for processing that meets standardization requirements. Conventionally, the series for processing is manually identified by an image analyst who opens and reviews every series. This process is extremely time-consuming, taking up to 1 hour for patients with multiple time points and is prone to error when dealing with a large dataset. Therefore, automatic image labeling would be beneficial for large scale data mining in clinical research. DICOM is the standard format in medical imaging, and its headers include many acquisition parameters. However, DICOM header tags are often insufficient or unreliable to provide all the necessary labels[6–8]. Image based identification is thus essential to solve this problem. In this work, we aim to identify the following characteristics of chest CT scans: anatomical coverage, scan direction, scan posture, lung coverage completeness, contrast usage and patient breath-hold. Identifying anatomical coverage of scans is critical to decide the appropriate quantitative analysis task afterwards. For example, lung and lobar segmentation are appropriate for scans that cover the chest. Patients can be scanned either head-first or feet-first and most chest CT analysis algorithms are developed expecting head-first. Also, patients can be scanned either prone or supine. Supine imaging is most commonly performed for diagnosis, while prone imaging sometimes helps to resolve dependent opacities that resemble pathology on supine images [9–11]. Full lung coverage is a basic requirement in thoracic imaging, otherwise the resultant measurement, e.g. lung volume, is inaccurate. Contrast CT is often used to provide physicians with better delineation of structures with soft tissue attenuation. Intensity differences between contrast and non-contrast scans can lead to different quantitative results and so labeling in terms of contrast vs non-contrast is important. Finally, patients are imaged at suspended full inspiration or expiration and identification of breath-hold types is necessary for lung volume and attenuation measurements to be interpreted correctly. For example, total lung capacity (TLC) scans are preferred to evaluate most pulmonary diseases, including, emphysema, fibrosis and, nodules, while residual volume (RV) is ideal for air trapping assessment. Collectively, identifications of all of these image labels is a pre-requisite to perform quantitative research on lung CT.

Automated lung segmentation has been studied extensively for over 20 years and many algorithms have achieved good performance on normal lungs[12–14]. However, clinical practice presents a variety of pulmonary diseases, and methods for pathological lung segmentation have also been investigated in many studies. For example, non-rigid registering normal to pathologic lung to include abnormal regions [15]. Anatomical knowledge, such as the curvature of ribs, was also applied to assist the inclusion of pathologies in a semantic model[16]. Machine learning approaches based on texture features have been introduced to detect abnormal lung regions[17–19]. Those methods were only evaluated on a small dataset with limited pathological categories and their generalizability and robustness are uncertain when applied to large and other cohorts. They are also computational expensive which is not ideal for high throughput processing. Lobar segmentation involves detecting or segmenting the fissures and then using them to isolate different lobes. The most common approach has been to exploit the contrast and physical appearance of fissure by applying ridge detection techniques, such as the hessian filter[20–22]. Adjacent anatomical structures, e.g. airway trees and vessels, were used to assist the track of fissures in 3D volume[23,24]. A derivative of stick (DoS) was proposed to improve the visibility of fissures especially when the slice thickness is thin[25]. It defines non-linear derivatives with a stick kernel in multiple directions. Nevertheless, these methods tend to be unreliable in the presence of a large amount of lung pathology near the fissure, especially if the fissure is incomplete.

To handle big data for research and clinical deployment of quantitative chest CT analysis systems, there is still significant manual intervention required by image analysts in both image labeling and lung/lobe segmentation. As such, large scale processing is currently not feasible. The goal of this dissertation is to fully automate the labeling and segmentation tasks in chest CT with high accuracy. We will develop the necessary algorithms and test their accuracy. We expect that since the systems are fully automated the image analyst intervention required will be reduced or eliminated, but direct confirmation of this is outside the scope of the thesis.

1.3 Deep learning in medical imaging

Deep learning has been applied extensively in medical imaging and various modalities[26–28], including CT[29–33], MRI[34–38], ultrasound[39–43], PET[44–48], histopathology[49–53], etc. Specifically for medical image analysis, convolutional neural networks (CNN)[54,55] have been successfully and widely used in a variety of areas, including classification, detection, segmentation, registration and denoising, and it has become the state-of-the-art method due to its powerful feature extraction capability. Instead of using hand-crafted features as traditional machine learning methods, the extraction of representative features is fused into the subsequent learning stage in a CNN, meaning useful features will be learned directly from the data by the network itself.

1.3.1 Classification

Medical image classification, binary or multiple classes, has improved substantially since CNNs were initially developed and applied to differentiate images of 1000 categories for ImageNet Large Scale Visual Recognition Competition

(ILSVRC)[56]. A 5-layer CNN was used to classify anatomical coverages in CT including brain, neck, chest, abdomen and pelvis and high accuracy was achieved on a test set[57]. A similar study was conducted by comparing the classification performance, as well as training efficiency, using AlexNet[58] and GoogleNet[59]. A deep convolutional neural network was used to automatically classify tooth types on dental cone-beam CT[60]. A transfer learning method[61] was introduced to identify diffuse lung diseases, including consolidation, ground-grass opacity (GGO), honeycomb, reticular, emphysema and nodular patterns. CNNs trained from scratch were also applied to recognize various lung patterns[62] and interstitial lung disease[63] on CT images. A CNN based deep learning method was presented to classify different liver masses into 5 categories on contrast enhanced CT[64]. Both 2D and 3D CNNs were proposed to differentiate Alzheimer's disease (AD), lesion and normal tissue on brain CT images[32]. Neural networks have been used to differentiate pulmonary nodules from non-nodules, e.g. vessels, from a large candidate pool [65] and a "massive training artificial neural network" (MTANN) was presented to perform the task on low-dose CT[66]. Fu et al. used a CNN as a feature extractor and the output was combined with hand-crafted features for classification since these manually designed features worked well on false positive reduction[67]. A comparison study investigated performance differences between a CNN and regular deep neural network (DNN) and the DNN surpassed CNN by 0.2% in their experiment[68]. 2D CNNs[69] or multi-views (2.5D) CNNs[70] have been proposed to accomplish this task before 3D CNN[71–74] came to dominate this research. Not surprisingly, 3D CNN based methods tend to outperform earlier 2D and 2.5D CNNs. To overcome the cumbersome manual design of the neural network architecture, a particle swarm optimization (PSO) algorithm[75] was applied to optimize CNN's hyper-parameters automatically for better performance. The most commonly used public data set for machine learning involving lung nodules on CT is LUNA16[76], a subset of LIDC/IDRI database[77]. This data set is frequently used as benchmark to test the superiority of proposed methods. Nodule classification to distinguish benign from malignant is another research area where CNNs have been widely adopted. Early methods used 2D and 2.5D[31,78] CNNs and then multi-scale[79,80] and 3D CNN[81–84] based methods were proposed with better performance. Some groups trained and tested their model on public datasets from the National Cancer Institute (NCI) and Lung Image Database Consortium (LIDC) and some used their own in-house datasets that were relatively small. Similar to false positives reduction, the performance improvement are commonly from augmenting dimensions of input data with more contextual information, deeper and wider network, for example, incorporation of residual block learning. Many CNN based classification methods have showed promising results, demonstrating comparable or even better performance than radiologist in many tasks.

1.3.2 Detection

CNN has also been used to detect or localize anatomical organs, landmarks and lesions. A deep transfer learning scheme was applied to detect polyps in CT colonography, with the transfer of model weights being used to mitigate the lack of a large annotated data set, and their preliminary results showed promising sensitivity. 2D CNNs using images from orthogonal (axial, coronal and sagittal) planes were proposed to detect 3D region of interests (ROIs), including the heart, aortic arch, and descending aorta, with a 3D bounding box and achieved an AUC of 0.988[85]. A similar approach was employed to detect lymph nodes in non-contrast CT[86] and localize the landmarks on the distal femur surface[87]. Such methods simplified the 3D scan as composition of 2D images in orthogonal directions which helped to circumvent expensive parsing in 3D volume. Regression based localization has also been presented and the authors proposed a SpatialConfiguration-Net[88] considering anatomical variations in local appearance. They compared their method with both 2D and 3D fully convolutional neural networks and showed state-of-the-art detection performance. A few 3D CNN based methods were also introduced subsequently. A 3D CNN[89] was trained end-to-end to detect and localize vertebra on pathological CT scans and the additional contextual information helped to improve their detection rate to 96% and computational time was less than 3 seconds. An very similar approach was presented that incorporated a novel joint learning to enhance detection accuracy, and this J-CNN[90] took into account pairwise dependencies between adjacent vertebrae. A two-step approach was used to detect the carotid artery bifurcation on a head-neck CT[91]. Candidates were first extracted by a shallow network and then were accurately classified by a 3D deep network. Separable filter decomposition and network sparsification was introduced to prevent over-fitting. A recurrent attention DenseNet (RADnet) was applied to detect hemorrhage on CT[92] and demonstrated 81.82% prediction accuracy that was better than 3 radiologists. The architecture was based on 2D CNN but a recurrent neural network was added to provide 3D context from neighboring slices. Applications to pulmonary nodule detection have also been investigated using 2D and 3D CNNs. A 2D CNN was trained using raw images, as well as pre-processed images by a Frangi filter, as input to reduce the rate of misidentifying vascular structures as nodules[93]. A 2D CNN using images from orthogonal planes was designed, hoping that coronal and sagittal slice would assist by providing more spatial information[70]. Ginneken et al. exploited the popular detection network OverFeat[94] to extract deep features also based on multi-planner views and a support vector machine (SVM) was ultimately used to detect nodules[95]. Ding et al. first applied Faster Region-based CNN (Faster R-CNN) to detect candidate nodules on 2D axial slices that were then fed into a 3D CNN for false-positive reduction[96]. 3D fully convolutional networks (3D FCNs) were also proposed. Investigators applied them to a sub-volume via a sliding window given the large volume of an entire scan[97]. A similar 3D FCN structure using multiple GPUs and a larger splitting size was proposed to speed up the computation[98]. A multi-scale 3D CNN[99] was presented to mitigate variations in nodule size and that scheme helped to better detect extremely small nodules. In further work, a hybrid-loss function that emphasizes nodule location and size was incorporated in training the network[100]. Also, provided only with location of one voxel of nodule, a weakly supervised CNN was proposed to alleviate the challenge of annotating the entire nodule[71].

1.3.3 Segmentation

Segmentation is another area where deep learning dominates in medical imaging. Almost all approaches have used fully convolutional networks (FCN)[101]. U-Net[102], one of the earliest and most successful architectures applied to biomedical image segmentation, has been widely used in many different applications. For example, brain tumor segmentation in MRI[103], cartilage and meniscus segmentation on knee

MR[104], and iris segmentation[105] .The symmetric architecture with same number of convolutional layers and deconvolutional layers, as well as the skip connections between down-sampling and up-sampling stages, are the 2 main contributions of this approach. To deal with target size variability or a relatively small object in multi-class segmentation, multi-scale U-nets and derivates[106,107] were proposed to segment brain tumors[108,109] and different glands on histological images[110]. The 3D U-net[111] and V-net[112] extended original 2D to 3D image segmentation and they were first applied to segment the prostate and kidneys in MRI. Due to the success of ResNet[113], residual blocks were incorporated into many architectures to improve the network's feature extraction capability[114]. For example, VoxResNet[115] for brain and prostate[116] segmentation on MR, FusionNet[117] for cell membrane segmentation on electron microscopy image, and retinal layer pathology on OCT images[118]. Similarly, DenseNet[119] was converted from its original classification architecture to a segmentation network[120]. Graphical model Markov Random field (MRF)[121] and conditional random field (CRF)[122] have also been used for post-processing in a few studies to improve CNN segmentation and groups directly embedded the CRF in their end-to-end training for 2D retinal vessel segmentation[123,124], 2D cervical nuclei segmentation[125] and 3D brain lesion segmentation on MR[126,127]. Similar methods using 2D[128] and 3D FCNs[129–131] for pulmonary nodule segmentation were also reported. Recurrent neural networks (RNNs) have also been studied for segmentation. Xie et al. proposed using a RNN on perimysium segmentation in histopathology images. Prior information for a given row and column were preserved when classifying the current patch as reference[132]. A multi-dimensional GRU based RNN was presented to segment white and gray matters in brain MR[133]. Overall, CNN based methods have tended to achieve better performance than RNNs.

Generative adversarial network (GAN)[134] based methods have been presented in medical image segmentation as well, and may help the model by introducing extra adversarial loss from discriminators. SegAn[135] was proposed for brain lesion segmentation. The generator has an FCN responsible for generating mask image and a discriminator with a multi-scale L1 loss was used to challenge the mask to provide feedback. So the generator and discriminator were trained against one another in a min-max game. A combination of ACGAN[136] and pix2pix[137] was applied to segment different cells on fluorescent microscopy images[138]. Qin et al. developed a conditional GAN to synthesize CT images with nodules to augment the size of training sets and diversify the variations for the second step of nodule segmentation by 3D FCN.

1.4 Contributions

To solve the problems discussed in Section 1.2, deep learning based methods have been proposed in this dissertation to accelerate and leverage quantitative lung CT analysis in clinical practice. Two major contributions of this work include:

1. *Development of a fully automated image labeling pipeline to replace current manual curation for large chest CT data sets*. CNNs using orthogonal image planes were built to detect anatomical coverages, scan direction, scan postures, lung coverage completeness, contrast usage and patient breath-holds. The method was able to achieve high automated accuracy (> 95%) and is thereby expected to improve the efficiency of curating large data sets.

2. *Development of a fully automated lung and lobar segmentation method on a broad spectrum of pathological lungs*. A 2D ResNet-101 based fully convolutional network was first developed to accurately segment lung. Then, a 3D DenseNet based network was applied to achieve lobar segmentation based on this lung segmentation. The 3D CNN was able to exploit spatial context to separate individual lobes, even when fissures are incomplete or absent. The collective lung/lobar segmentation takes less than 10 seconds for a typical high resolution scan.

Overall, these labeling and segmentation models showed reliable performance and are applicable to a variety of chest CTs in clinical practice, including adult and pediatric scans, high and low resolutions scans, and diagnostic and low-dose scans.

1.5 Organization of this dissertation

Chapter 2 introduces the development of deep learning models for automated image labeling, including identification of anatomical coverage, scan direction, scan posture, lung coverage completeness, contrast usage and patient breath-holds.

Chapter 3 describes the development of automated lung and lobar segmentation using 2D ResNet and 3D DenseNet based fully convolutional networks, including whole lung segmentation, left and right lung separation, and lobar segmentation.

Chapter 4 summarizes the overall research, limitations of the work and discusses potential future extensions.

Chapter 2: Automated image labeling on chest CT

Image data curation, involving identification and labeling of relevant types of scans with consistent acquisition parameters, is important for machine learning with big data. Labeling of images meeting standardization requirements is vital in obtaining reliable research findings from multi-center clinical trials [139–142] since images collected from different sites are often heterogeneous and variable in terms of the types of scan acquired. For example, each patient may include image series from multiple time points and each time point could have multiple series that can be redundant, irrelevant or unusable. The traditional approach is to manually assign the labels by reviewing each image series and selecting the best series to process that meets standardization requirements. It is very time-consuming and prone to human errors. As such, an efficient automatic image labeling method would be beneficial for large scale clinical research. DICOM is the standard format in medical imaging and it contains a variety of scan parameters and other metadata. DICOM headers, e.g. series description, are extensively used to extract information for labeling. Nevertheless, DICOM tags are often insufficient or unreliable due to manual entry or vendor discrepancies[6–8,143,144]. To overcome this limitation, image based identification is imperative. In this work, we aimed to identify anatomical coverage, scan direction, scan posture, lung coverage completeness, contrast usage, and patient breath-hold level. These labeling tasks were posed as classification problems. To our knowledge, there is no related work on this specific problem of classification for comprehensive scan labeling. Although these labels seem basic for chest CT, they are labor intensive to assign for large data sets, and efficiently and correctly extracting them is crucial for data curation in clinical research.

In this chapter, we present a fully automated high throughput labeling method using deep learning to create classification models for each label. Our hypothesis was that each classifier can achieve >95% accuracy individually in assigning its label.

2.1 Anatomical coverage

2.1.1 Introduction

Computed Tomography (CT) has been one of the most popular imaging modality since 1970s and widely used for imaging various anatomies, such as brain, chest, abdomen, pelvis, legs, etc. Image data collected from hospitals and clinical trials from different sites sometimes contains series of different anatomic coverages and it is a pre-requisite to identify the anatomy included before they can be used for image processing and machine learning. Clinically, the four most common anatomical coverages are brain, chest, abdomen-pelvis and chest-abdomen-pelvis (CAP). Identifying scans containing appropriate anatomy is critical to trigger a specific image analysis pipeline. For example, an identified chest or CAP scan can be processed by lung and lobar segmentation, as we will discuss in Chapter 3. Brain scans can be used for detection of aneurysm, stroke and tumors[145,146]. Abdomen-pelvis scans can be used to detect diseases of internal organs, such as lesion in the liver or spleen, and polyps in colons[147–149]. DICOM tags, e.g. ‘SeriesDescription’ may or may not include a statement of the imaged anatomy. Another tag ‘BodyPartExamined’ is very unreliable and may also be blank. Some anatomic classification methods have been introduced.

Fenchel et al.[150] and Shimizu et al.[151] proposed atlas based methods in which a single atlas from a representative training set was built. Unseen (test) scans were non-rigidly registered to the atlas and anatomical labels from the atlas were propagated to them. The accuracy of this approach was limited by intra- and inter-subject variability in organ location and size, such as topological changes in organ boundary due to disease. A multi-atlas approach[152] was introduced to lessen the influence of data variability. However, non-rigid registration to multiple atlases leads to a significant increase of computation time. A regression forest method[153] was presented to detect anatomical structures, e.g. lung, heart, liver on CT scans. All voxels in the CT volume contribute with varying confidence to estimate the position of bounding box of organs and they aimed to learn the non-linear mapping from voxels directly to organ position and size with training focusing on maximizing the confidence of output predictions. As a result, landmarks for a specific organ were identified for localization. A support vector machine (SVM) based machine learning method[154] was proposed to identify anatomic coverage in 3D. Each scan was split into $k \times k \times k$ ($k = 3, 5, 7, 9, 11$) non-overlapping blocks and they were represented by their individual mean intensities. The resultant $1 \times k^3$ feature vector of each scan was used to train a SVM classifier. Although it achieved high accuracy, the down-sampling processing is not efficient to form the final feature vector for classification. Roth et al. proposed a 5-layer CNN to classify the anatomical region in CT images, including brain, neck, chest, abdomen and pelvis[57]. A similar approach, but using AlexNet and GoogleNet, was then introduced with varied training time and accuracies[155]. Yan et al. used a multi-stage deep learning framework to recognize 12 different body parts by approximating a bounding box of local patches[156]. A CNN was first used to extract the most discriminative local patches and these patches were then used to identify body parts. Images from orthogonal planes (axial, sagittal and coronal) were used to build 3 independent CNNs and results were combined to produce a 3D bounding box for each organ, such as heart, lung, kidney, liver[85,157].

In this work, we proposed a CNN using coronal slices to identify brain, chest, abdomen-pelvis, and CAP scans.

2.1.2 Methods

The conventional approach is to identify the anatomy on each axial slice sequentially and then determine the whole coverage depending on the anatomy present in each slice. One big challenge is the annotation of a large number of axial images for training. Alternatively, coronal images visualize the coverage completely and only a single scan-level annotation is required. Additionally, the identification is more efficient since it only requires one coronal slice instead of recognizing every axial slice.

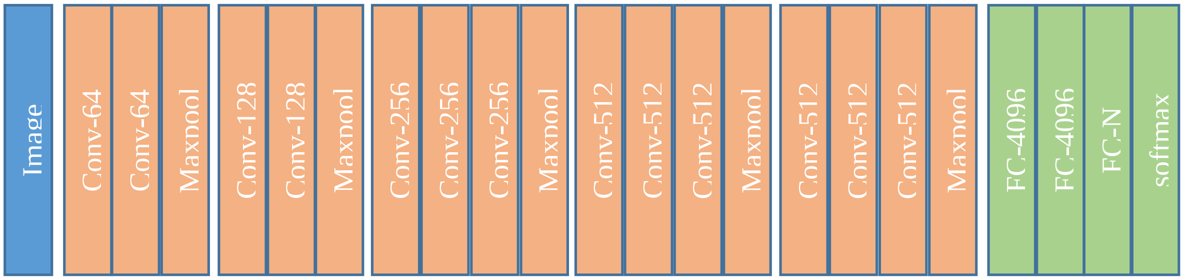


Figure 2.1: Architecture of VGG-16 used for classification.

VGGNet[158] won the localization task of the ILSVRC (Imagenet Large Scale Visual Recognition Competition) and 2rd place in the classification task in 2014. It is still widely used in many areas. In our work, VGG-16 was used as the classification architecture to identify 4 different anatomic regions. It consists of 13 convolution layers plus 2 fully connected layers with 4 classes in last classification layer using ‘softmax’ as activation since it is a multi-class classification problem. The architecture is shown in Figure 2.1

Real-time data augmentation was applied, meaning the CPU generated transformed images with specified augmentation parameters while the GPU was responsible for training the model. Augmentation included horizontal and vertical flipping, zoom in/out of [0.8, 1.2], horizontal shift of [0.1, 0.1] and vertical shift of [0.05, 0.05]. At each iteration, an image will be randomly flipped by x-axis and y-axis (50% by chance), rescaled by a factor between 0.8 and 1.2 of original image size, randomly shifted to left or right by up to 10% of image width, randomly shifted up or down by up to 5% of image height. Vertical flipping helped to make the model insensitive to scan directions (head-first or feet-first). Zoom in/out helped consolidate the mode facing different body size (adult vs. pediatric). The relatively small vertical shift was applied to ensure the transformed images still preserved the original anatomic coverage, while still allowing some tolerance of variation in coverage. For example, it may lead to a cutoff at lung apex/base on the original chest scans and thus diversify the training set by incorporating cases with incomplete lung coverage.

As a result, the generated image at each iteration should be unique since it is a combination of different transformations each and every time. Traditional augmentation pre-generates those images before training and the augmented data will be used over and over through a number of iterations. Real-time augmentation contributes to enhancing independence of individual training samples and greatly saving computational time and storage space compared to conventional approaches.

The same augmentation technique was used in all the following problems but transformation options varied depending on the specific application. Instead of transfer learning, we trained our network from scratch using single channel images. The model was trained on a NVIDIA TITAN X with GPU memory of 12 GB, using the Adam optimizer, and a learning rate of 0.001 with a decay rate of 1e-4.

2.1.3 Materials

Example coronal images of 4 anatomical coverages are shown in Figure 2.2.

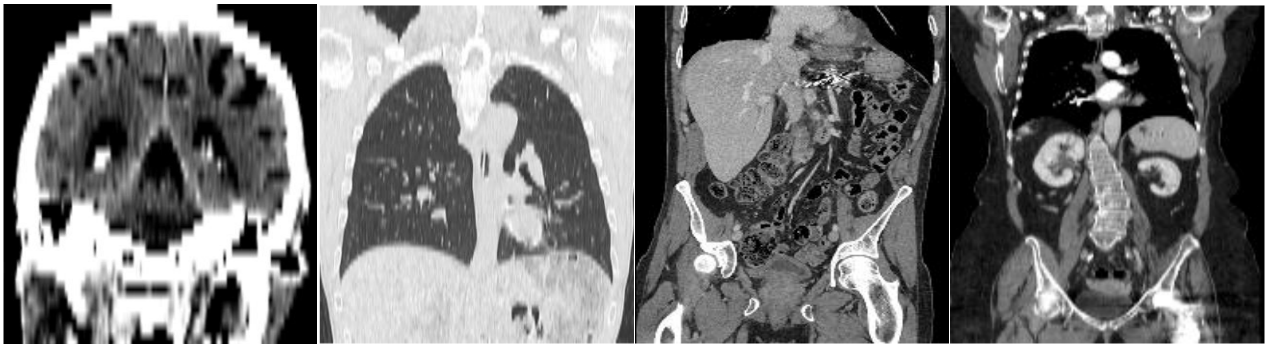


Figure 2.2: Example anatomical coverage in this task. Left to right: brain, chest, abdomen-pelvis and chest-abdomen-pelvis (CAP)

1,000 scans each of the brain, chest, abdomen-pelvis and CAP were used. They were from a collection of 17 clinical trials, including 2 renal cell carcinoma trials, 1 hepatocellular carcinoma trial, 1 prostate cancer trial, 2 lymphoma trials, 1 advanced solid tumor trial, 4 IPF trials, 3 COPD trials, 1 tuberculosis (TB) trial, 1 non-small cell lung cancer (NSCLC) trial, and 1 lung cancer screening trial. Scanners from 4 manufacturers were included, from SIEMENS, GE, PHILIPS and TOSHIBA. Images were reconstructed with smooth, medium and sharp kernels with slice thickness and spacing range of [0.6 mm, 3.0 mm]. From each category, 800 scans were used as a training set and 200 as a test set and the 4 classes were balanced in both sets. The label for each scan was from a trained image analyst. For each scan, the middle coronal slice, plus 2 adjacent slices 10 pixels apart were extracted as representative images. For generalizability, chest scans in the training set included cases with various lung pathologies and different lung coverages. Figure 2.3 shows two examples with incomplete coverage at the lung apex and lung base, respectively.

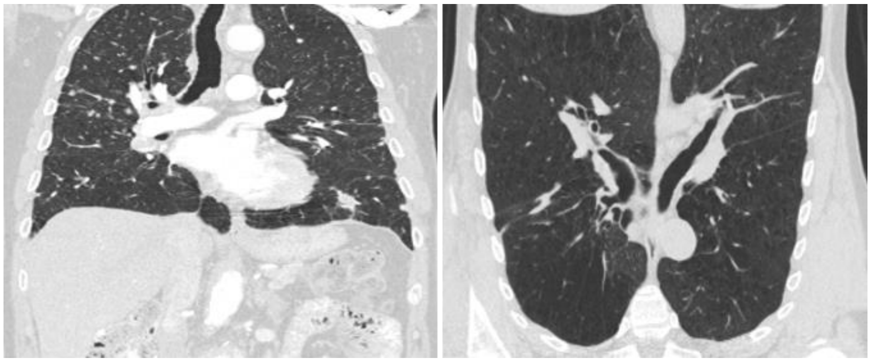


Figure 2.3: Incomplete lung examples included in training set

Figure 2.4: CAP scan variations in the training set

The most significant difference between abdomen-pelvis and CAP scan is the inclusion of the chest anatomy. In clinical practice, the abdomen-pelvis scan does also include a very small portion of the chest to ensure the entire abdomen is covered. For a CAP scan, it should contain complete chest, abdomen and pelvis and sometimes may also cover more superior anatomy, i.e. brain, neck or more inferior coverage, i.e. legs/feet. Figure 2.4 shows two CAP scans that include the brain and legs, respectively. Pre-processing including image down-sampling and normalization were applied and the same method was used in all sub-tasks in this work in both classification and segmentation. For the 2D image based CNN, the image was resized to 256 × 256. For the 3D image based CNN, the image was resized to 128 × 128 × 128.

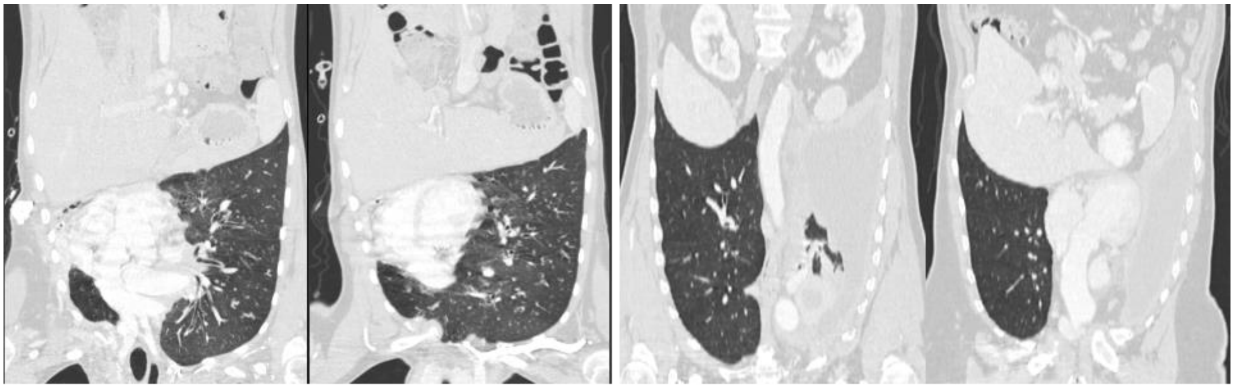
Table 2.1. Summary of Hounsfield units (HU) of different tissue on CT.

Hounsfield units	Tissue
>1000	Bone, calcium, metal
100 to 600	Iodinated CT contrast
30 to 500	Punctate calcification
60 to 100	Intracranial hemorrhage
35	Gray matter
25	White matter
20 to 40	Muscle, soft tissue
0	Water
-30 to-70	Fat
< -1000	Air

As shown in Table 2.1 above, the general intensity range of different tissues on a CT scan are: [-3000HU, 3000HU] and clinically relevant tissues in thoracic imaging are usually within [-1000HU, 1000HU]. Therefore, prior to input to the CNN the image intensity was linearly remapped to the range [-1000 HU, 1000 HU] and then normalized to [0, 1.0].

2.1.4 Results

Both training and validation accuracy were over 99.99% after 7 epochs and the inference time was < 10 ms per case using the machine described previously. 5-fold cross validation was used during evaluation and the model achieved an accuracy of 100%. Figure 2.5 shows two chest scans from the public LObe and Lung Analysis 2011 (LOLA 11) dataset[159]: one with only a partial right lung and the other has only the right lung. Although no similar cases were included in the training set, the trained model still successfully identified them as chest scans.

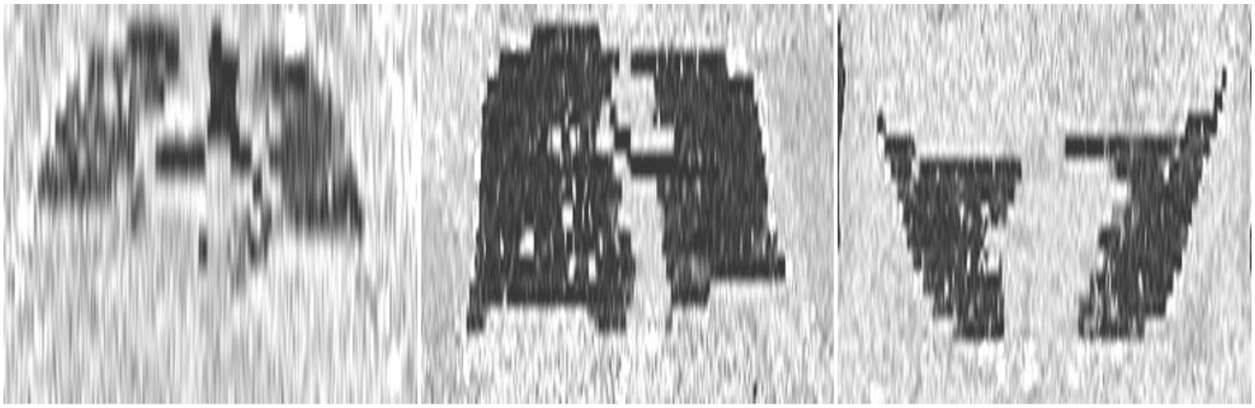


(a) CNN: Chest

(b) CNN: chest

Figure 2.5: Example predictions by the CNN on two cases from LOLA11.

Some chest CT scans are acquired “incrementally”, i.e., with thin slice thickness (e.g., 1 mm), but large slice spacing > 10 mm. As such, an additional 100 incremental chest scans were used in testing and 100% accuracy was achieved by the model on those cases. Figure 2.6 contains three examples with spacing of 20 mm. They were all correctly identified as chest scans even though the model was trained only using scans with slice spacing less than 3 mm. It should be noted that resized coronal images (256 × 256) were fed into the model.



# of slice = 14
CNN: Chest

# of slice = 32
CNN: Chest

# of slice = 36
CNN: Chest

Figure 2.6: Example predictions by CNN on three cases with large spacing of 20 mm.

2.1.5 Discussion

Rather than conventional identification of anatomic regions using axial images, our method demonstrated the feasibility of employing coronal slices to accomplish this task. Large scale manual annotation has been a big challenge, especially in medical imaging, and the proposed method only requires a scan level label. Furthermore, it is computationally more efficient. For example, a scan with 500 axial slices will take 500 times longer using axial images than using a coronal image.

The sensitivity or susceptibility of deep learning models to seemingly minor variations is a common weakness even if there are state-of-the-art in their fields[67,160]. For example, Figure 2.7 is well-known picture[161] showing that adding imperceptible perturbation by human eyes to the image can sometimes change the label predicted by the model.

Figure 2.7: Vulnerability of CNN with perturbation of noise.

In our testing we also applied our model to much noisier low and ultra-low-dose scans to investigate its robustness. A total of 30 chest scans with raw projection data from lung cancer screening were used with an original (100%) dose level around 2 mGy. A CT reconstruction pipeline[162] available in our group was used to simulate images with 1 mm slice thickness and spacing, a medium kernel and dose levels at 10%, 25%, 50% and 100% of the original dose. Weighted filtered back projection (wFBP) was used as the reconstruction algorithm. Our model achieved 100% accuracy on identifying them as chest scans across different dose levels.

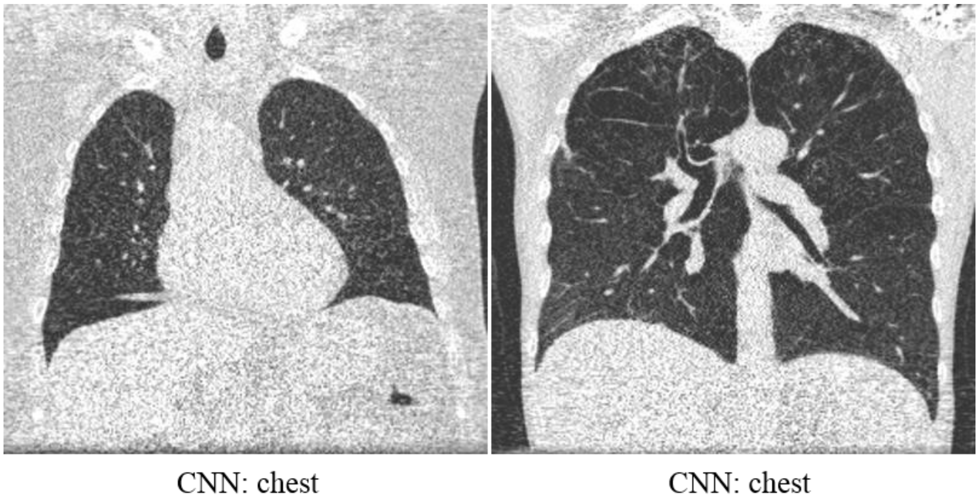


Figure 2.8: Anatomical model inference on ultra-low-dose scans.

Figure 2.8 shows two examples with dose level of 0.2 mGy (~1/75 of training set dose). Although the model was trained using diagnostic scans only (~15 mGy), it remained accurate when applied to low and ultra-low-dose scans without any additional training or post-processing. In addition, the model that was trained only on adult scans was also applied to pediatric patients as shown in Figure 2.9 and they were successfully identified as chest scans.

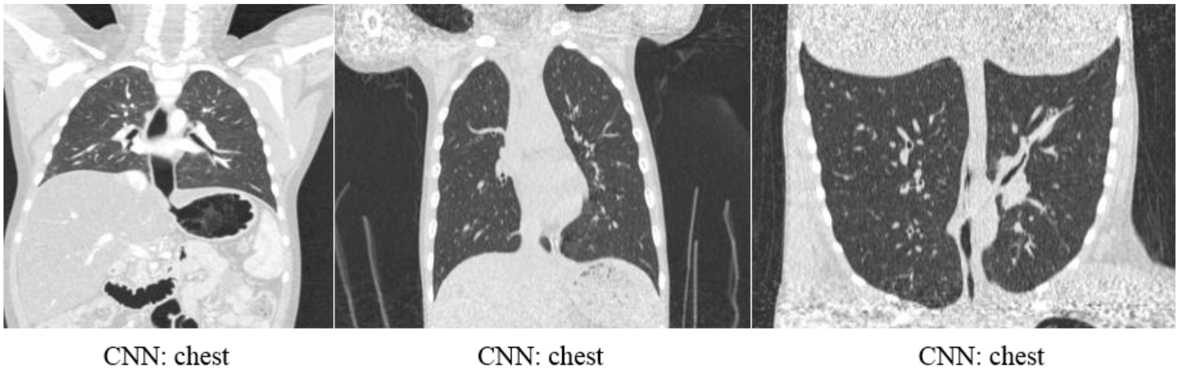


Figure 2.9: Anatomical model inference on pediatric scans.

One concern with deep learning methods is that they obscure the decision method and are considered “black box”. Although a deep learning model may give an accurate prediction, humans are unable to fully understand what kind of features are used by the network for classification and the process of how the decision is being made. One popular approach in attempting to understand CNNs is to visualize various feature maps at different layers and generating artificial representative images of different classes by computing the gradient with respect to them[163]. A successful neural net usually includes hundreds of thousands of filters and thus one drawback of this approach is the difficulty of selecting the most representative filters from an enormous pool. On the other hand, image specific saliency maps help interpret the network using a single back-propagation pass through entire classification process[164].

Saliency refers to the most visually important elements or patterns to the network when it is exposed to an image[165] and the saliency map is a collection reflecting their topographical representation.

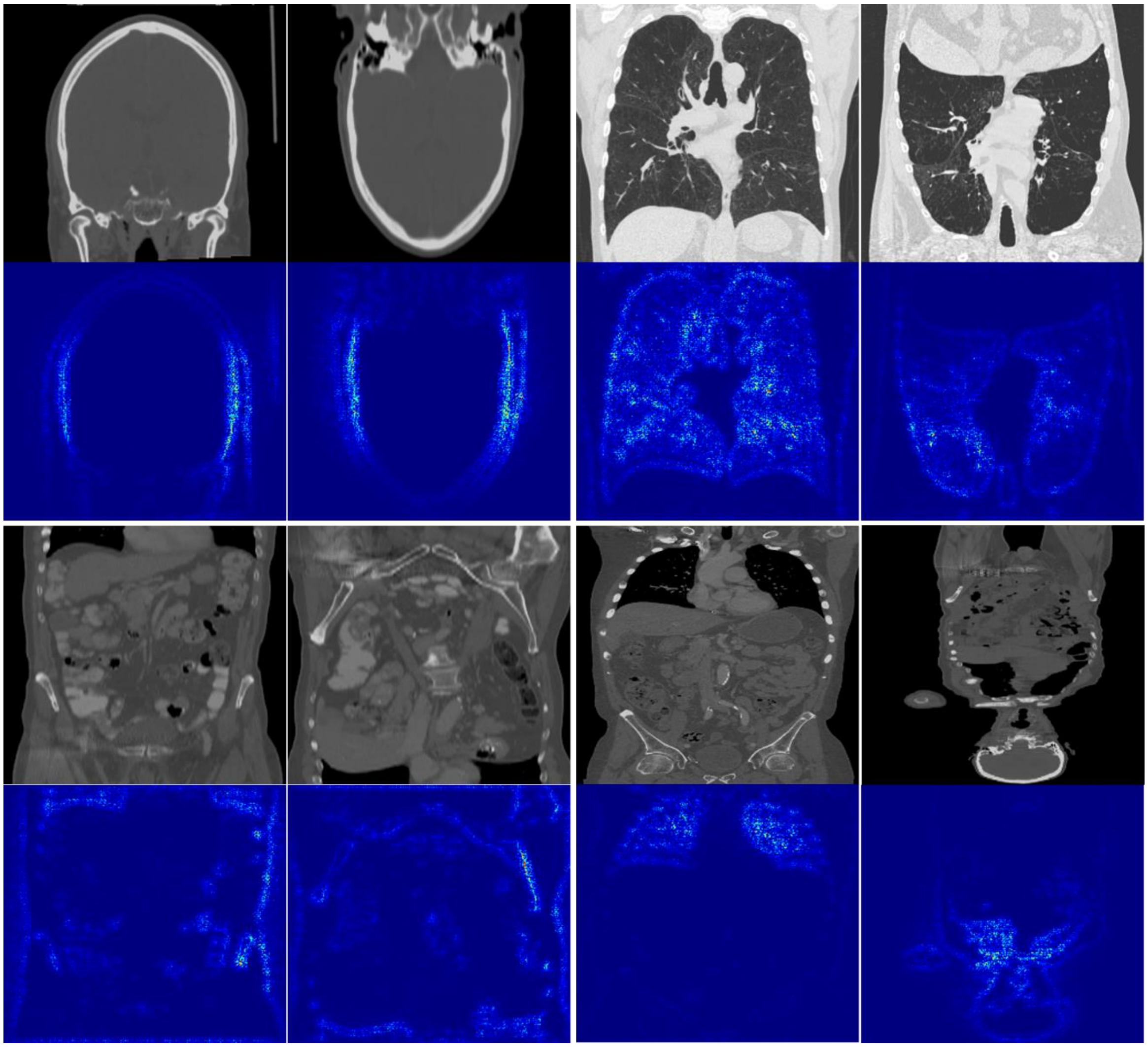


Figure 2.10: Saliency maps of four classes at the last classification layer of the anatomical coverage identification model.

Saliency maps of four classes at the last classification layer are shown in Figure 2.10. The first and third row are the input images to the network and second and fourth row are the corresponding saliency maps of that class.

For brain scans, the skull with a semi-circular shape is used by the network as a characteristic of a brain scan. Similarly, the network exploits the boundaries and content within the lung to identify it as a chest scan. In the case of an abdomen-pelvis scan, the lung as well as bones and other issue on the two lateral sides are extracted as deterministic features. Lastly, pixels within lung are highly activated on the CAP scan and the skull is another significant feature used by the CNN. In comparison with the chest scan, the lung size is much smaller and the location is near either the top or bottom of the image. To sum up, these features extracted by the network to differentiate the four classes are human interpretable and probably consistent with our own thought process. To understand deep learning models better in the future will require more multidisciplinary collaboration including neuro science, computer science, mathematics, etc.

2.1.6 Conclusion

The presented model demonstrated high performance in identifying four different anatomical coverages based coronal images and was shown to be generalizable to scans of high and low resolutions, diagnostic and low CT doses, and adult and pediatric patients in clinical practice.

2.2 Scan direction

2.2.1 Introduction

Patients can be scanned head-first or feet-first. Most chest CT analysis algorithms are developed expecting head-first ordering of images. Incorrect image orientation could cause wrong-side surgery[166,167] or treatment[168,169], e.g., in radiation therapy. For incremental chest scans (with large slice spacing, e.g. >=10mm), they are not suitable for precise lobar segmentation and one commonly used approach is to divide the whole lung into 3 zones: upper, middle, lower zones[170–172] for subsequent analysis. Obviously, correctly identifying the scan direction is critical to determine the proper ordering of the 3 zones. There is a DICOM tag named 'PatientPosition' with values of 'HFX' or 'FFX' where the first 2 characters indicate whether the scan is head-first or feet-first. However, this tag is very unreliable. A survey of 300 chest scans in our database reported an error rate > 50% using this tag. This could be due to manual entry by the CT technologist or flipping of the image during image reconstruction. Therefore, image based scan direction labeling was considered preferable and deep learning the deep learning method described was employed.

2.2.2 Materials and Methods

Similar to the problem of anatomical coverage detection, previous studies have used axial slices to identify different anatomies, including brain, shoulder, chest, abdomen and pelvis and thus scan direction was inferred by the aggregated anatomical order. To alleviate the workload of annotating every axial slice, representative coronal slices were again used to recognize the orientation. 1,000 chest scans of different subjects were used to build the classification model. They were a collection from a research database comprising 6 clinical trials (2 IPF, 2 COPD, 1 scleroderma, and 1 lung cancer screening) using scanners from SIEMENS, GE, PHILIPS and TOSHIBA. Slice thickness and spacing fell within [0.6 mm, 3.0 mm] and smooth to sharp reconstruction kernels were included. For each scan, a middle coronal slice plus 2 adjacent slices 10 pixels apart were selected as representative images. They were all head-first scans confirmed manually and corresponding feet-first scans were generated by vertically flipping the original scans to form balanced classes. As such, each scan generated 6 samples with 3 coronal images for each class. In total, 6,000 slices were used and they were split into training and test sets with a ratio of 4:1. Examples of head-first and feet-first coronal images are shown in Figure 2.11.

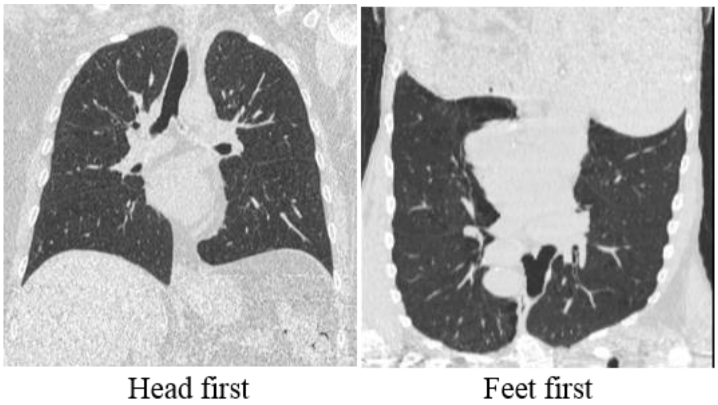


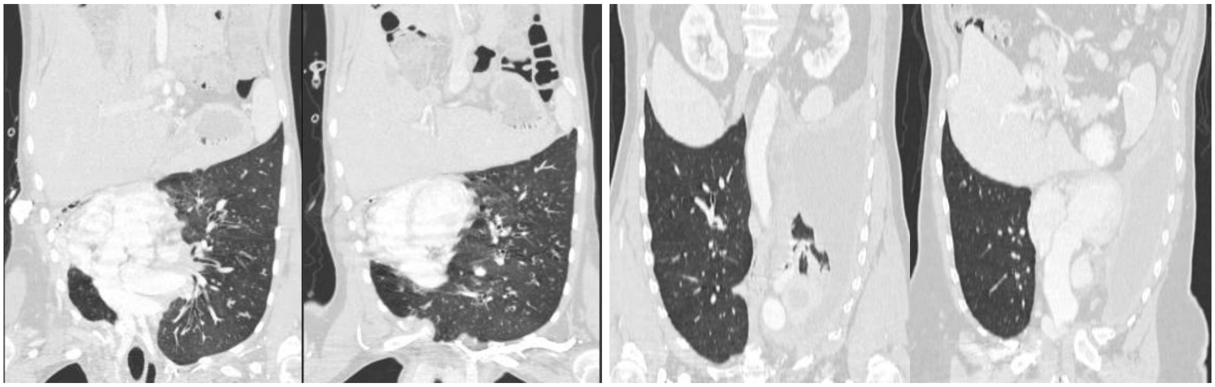
Figure 2.11: Examples of head first and feet first scans in coronal view.

VGG-16 was again employed in this classification task and the architecture is same shown for anatomical coverage identification except that the last classification layer is 2 classes.

Data augmentation was used, including horizontal flipping, zoom in/out of [0.8, 1.2], horizontal shift of [0.1, 0.1] and vertical shift of [0.2, 0.2]. Vertical flipping was dropped in this case. The model was trained from scratch on a NVIDIA TITAN X with GPU memory of 12 GB, using the Adam optimizer, and a learning rate of 0.001 with decay rate of 1e-4.

2.2.3 Results

Both training and validation accuracy were over 99.99% after 5 epochs and inference time was less than 10 ms per case using the GPU mentioned above. 5-fold cross validation was applied during evaluation and the model achieved an accuracy of 100% on the test set. The model worked even when the lung coverage was incomplete, and Figure 2.12 shows two examples from LOLA 11: one with partial right lung and the other has only the right lung and they were successfully identified as feet-first by the model.



(a) CNN: feet-first

(b) CNN: feet-first

Figure 2.12: CNN predictions of scan direction on two cases from LOLA11.

# of slice = 14
CNN: head-first

# of slice = 32
CNN: head-first

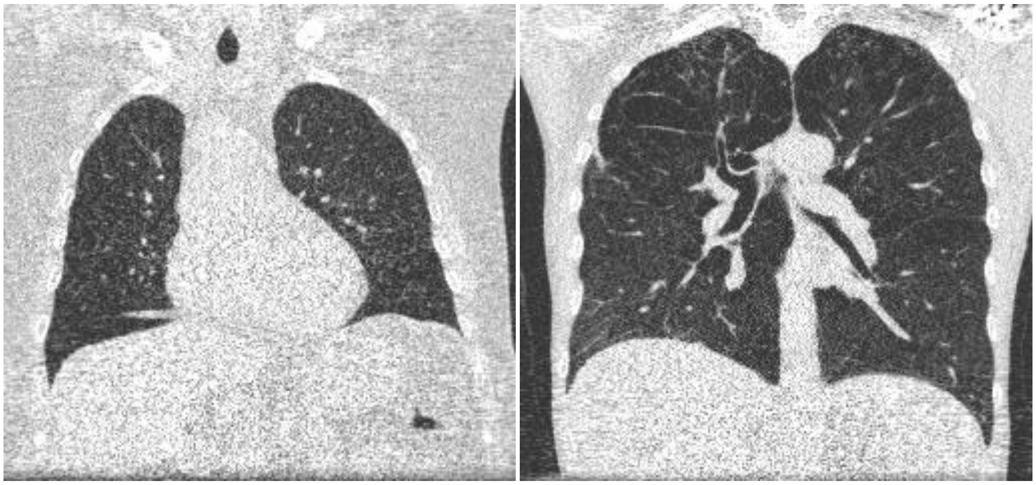
# of slice = 36
CNN: feet-first

Figure 2.13: CNN predictions of scan direction on three low resolution scans.

The model was also assessed on the extremely low resolution dataset with large slice spacing (5-20 mm) and 100% accuracy was achieved on a total of 100 scans. Figure 2.13 shows three examples with varying numbers of slices with slice spacing of 20 mm. These coronal images displayed were resized to 256 x 256 before being fed into the model for scan direction inference.

2.2.4 Discussion

Using a coronal slice instead of axial images helped to alleviate the burden of annotating a large amount of axial slices and allowed very efficient inferencing. Although the model was trained using only high resolution scans, it demonstrated reliable performance when applied to incremental scans.



CNN: head-first   CNN: head-first

Figure 2.14: Scan direction model inference on ultra-low-dose scans.

We also applied the scan direction model to the same low dose data set from lung cancer screening and achieved 100% accuracy. Two examples are shown in Figure 2.14. The model again showed robustness in handling various levels of noise.

Moreover, applications to pediatric scans are also present in Figure 2.15 with the first two cases correctly classified as head-first and last one as feet-first.

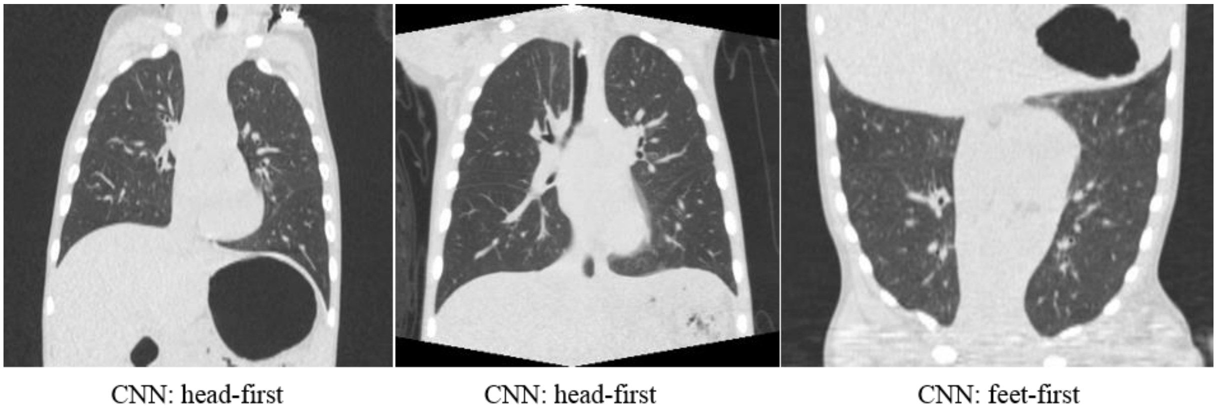


Figure 2.15: Scan direction model inference on pediatric scans.

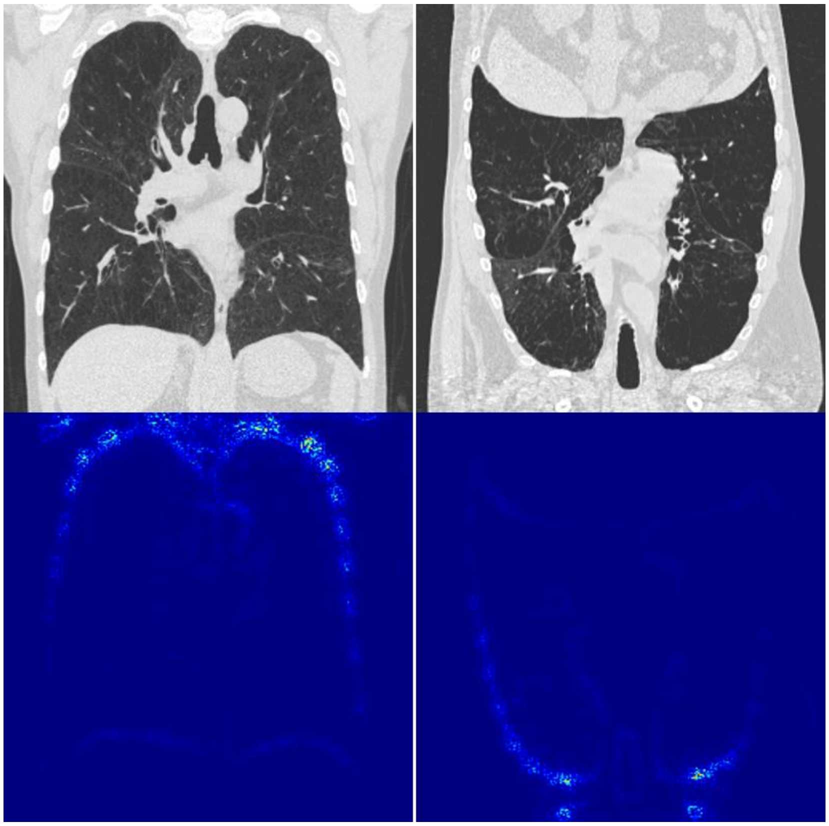


Figure 2.16: Saliency maps at last classification layer of scan direction model with respect to head-first and feet-first images.

Figure 2.16 shows the saliency maps of the last classification layer with respect to the input images. The saliency maps show that the boundaries of the lung parenchyma, especially areas near the apices are used by the scan direction model.

2.2.5 Conclusion

The scan direction model was successfully applied to identify the image orientation and was robust to a broad range of chest CT scans, including high or low resolution, diagnostic or ultra-low-dose, and adult and pediatric patients.

2.3 Scan posture

2.3.1 Introduction

Scan posture, face up (supine) or face down (prone), is dependent on placement of the patient in the scanner. Supine imaging is most commonly performed, for example in lung cancer screening, however, prone imaging is often performed in evaluating interstitial lung disease (ILD)[173–175]. On supine images, there are sometimes opacities due to lung collapse or atelectasis observed in the dependent portions of the lungs[9–11,176]. These dependent opacities may resolve during prone imaging while true opacities resulting from lung disease will not[177][178]. As these opacities may be mistaken as ILD, it is recommended in practice to also perform imaging in the prone position. Also, the posture may be adjusted for patient comfort or in cases of disability. Frequently the DICOM tag “SeriesDescription” does not include any information regarding the position. Another tag “PatientPosition” formatted as ‘XXP’ or ‘XXS’ if a scan is prone or supine, respectively. Although this tag is relatively reliable, it is not always available and it is occasionally inaccurate. A sagittal image based deep learning method was proposed to detect the scan position.

2.3.2 Materials and Methods

Both axial and sagittal images can be used to determine if a scan is prone or supine by checking the location of the table relative to the subject's body. Sagittal images were preferred since they clearly depict the relationship between the entire spine and the table. Example prone and supine scans are shown in Figure 2.17 with the vertical lines representing the scanner table. The middle sagittal slice gives an overall representation of the relevant image features and two other sagittal slices at ¼ and ¾ way through the body in the sagittal direction were also used in case the table was not present in the image center. Figure 2.18 is an example showing the absence of table at the middle sagittal slice and presence on slices at ¼ and ¾.

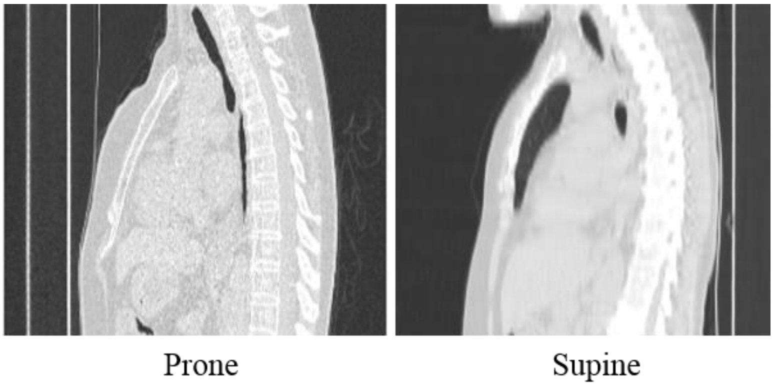


Figure 2.17: Example prone and supine scans.

In total, 5,000 scans (15,000 slices) from 5,000 different subjects were used to form training and testing sets with a ratio of 4:1. They were collected from a research database comprising of 16 clinical trials (5 IPF, 4 scleroderma, 3 COPD, 2 lung cancer screening, 1 tuberculosis, 1 lymphangioleiomyomatosis) using scanners from SIEMENS, GE, PHILIPS and TOSHIBA. Slice thickness and spacing were within [0.5 mm, 3.0 mm]. The number of prone and supine scans were balanced in both sets and the reference was from manual labeling.

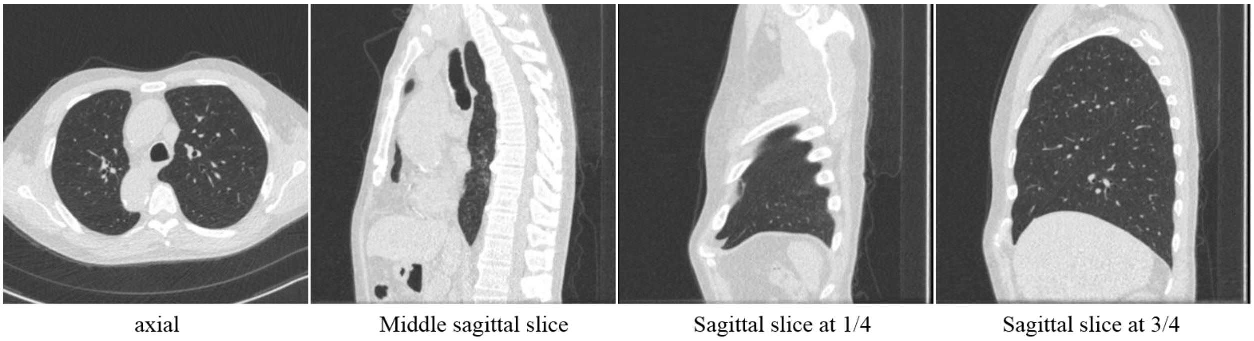


Figure 2.18: Example scans with presence of table at ¼ and ¾ but not middle sagittal slice.

The training set also included a relatively small number of cases without the table present, likely caused by an inappropriate field of view (FOV) during image reconstruction. Figure 2.19 shows two such cases that were labeled as supine by the image analyst. Although it is challenging to confirm the scan posture, the assigned label is likely correct since the spine is much closer to image boundary and the thin curved streak in the image are probably clothing.

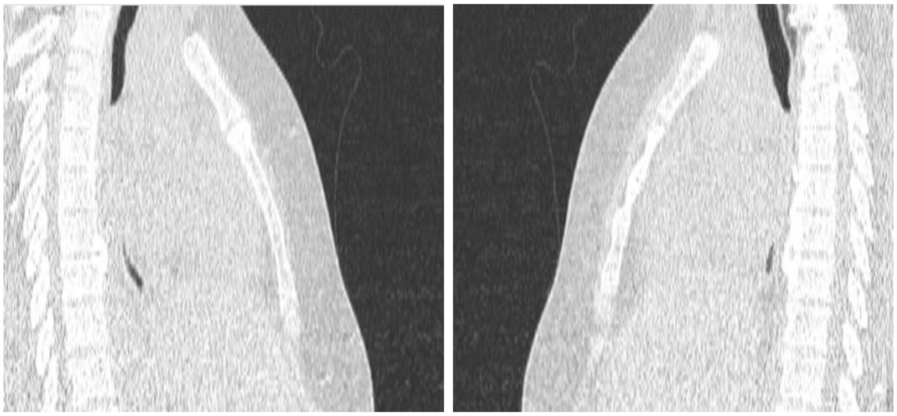


Figure 2.19: Example training samples (supine) with absence of table.

VGG-16 was also employed in this classification task and the architecture was the same as used for anatomical coverage classification except that the last layer has 2 classes. Data augmentation was used, including horizontal and vertical flipping, zoom in/out of [0.8, 1.2], and vertical shift of [0.2, 0.2]. Both horizontal and vertical flipping were applied since the spatial relationship between the spine and table is preserved. Similarly, a vertical shift up to 20% of image height still preserves the original label correctness. However, the position of table varies across different scans in terms of closeness to image boundary. Accordingly, horizontal shift was not applied to avoid table disappearance after translation. The model was trained from scratch on a NVIDIA TITAN X with GPU memory of 12 GB, using the Adam optimizer, with a learning rate of 0.001 and decay rate of 1e-4.

2.3.3 Results

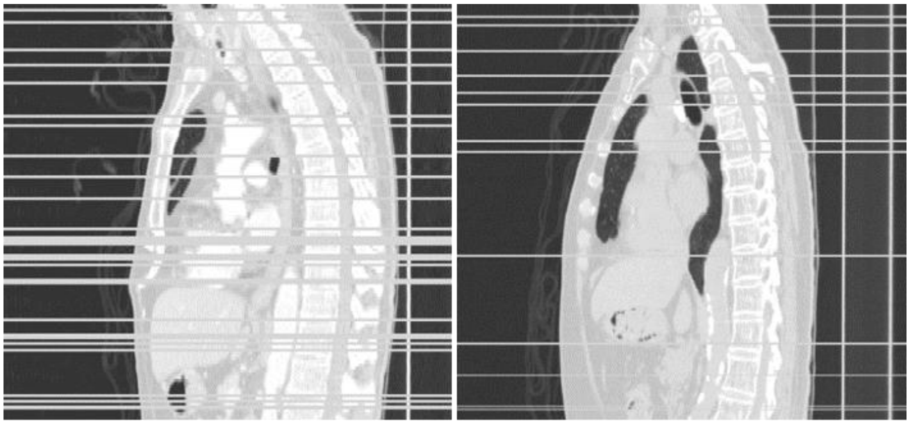
Both training and validation accuracy were over 99.99% after 12 epochs and inference time was less than 10 ms per case. 5-fold cross validation was applied during evaluation and the scan posture model achieved 99.3% accuracy on the test set. Figure 2.20 below shows two example cases where the CNN correctly identified the posture even though the information in the DICOM header tag was incorrect. The DICOM tag “PatientPosition” labeled the left image as prone and the right image as supine which is wrong when the images are flipped. Since horizontal and vertical flipping were applied in data augmentation while training the network, the model could consistently recognize the posture despite changing the table location within the image.

DICOM: prone
CNN: supine

DICOM: supine
CNN: prone

Figure 2.20: Example cases showing conflicting label from DICOM and our model.

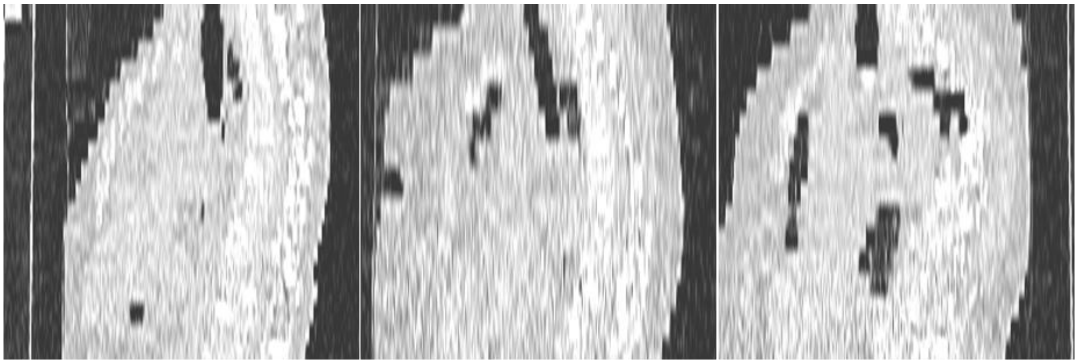
Figure 2.21 shows two unusual cases with intermittent blank slices. The model is still able to correctly identify the posture since these missing slices do not obscure the overall spine-table relationship for the model in the sagittal plane. If axial images were used in building the model errors would be more likely, this represents the advantage of using sagittal images over axial images to identify the scan posture.



CNN: supine

CNN: supine

Figure 2.21: Example cases with intermediate blank slices.



CNN: prone CNN: prone CNN: supine

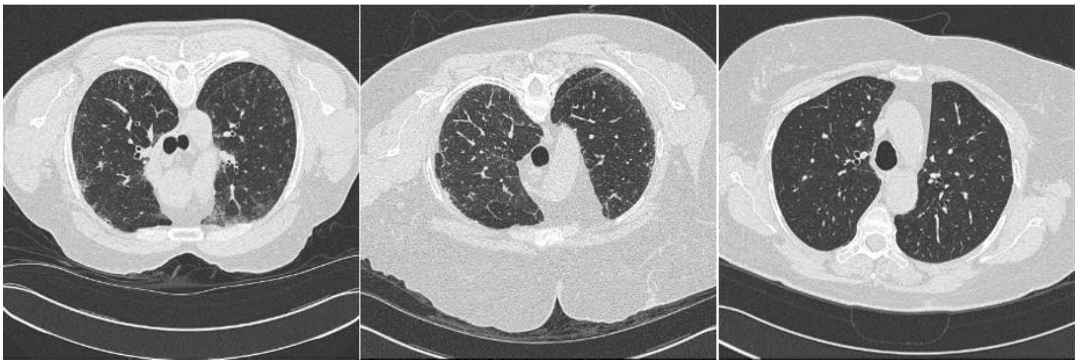


Figure 2.22: Application of scan posture model to low resolution scans.

The scan posture model was also evaluated on the dataset of low resolution chest CT and achieved 100% accuracy on the test set of 100 scans (5-20 mm slice spacing). Figure 2.22 shows the application of the model to three scans with slice spacing of 20 mm. Although those sagittal images are not as visually clear as high resolutions scans, their positions were still successfully detected by the model. Corresponding axial images are also provided in the second row to confirm the CNN inference.

2.3.4 Discussion

The scan posture model made use of sagittal images instead of axial to improve the inference efficacy and efficiency. The model was also assessed on the same low dose dataset from lung cancer screening and achieved 100% accuracy. Example predictions on ultra-low-dose scans by the CNN are shown in Figure 2.23.

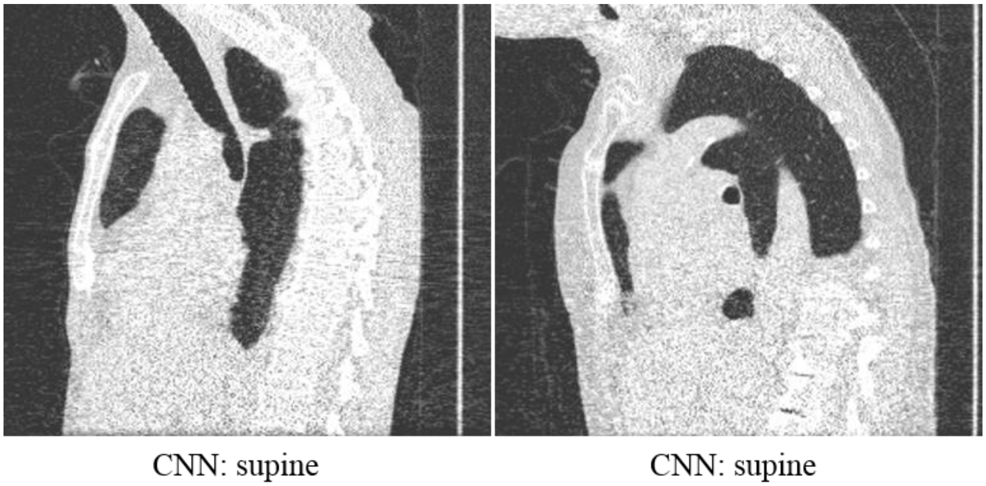


Figure 2.23: Scan posture inference on ultra-low-dose scans (~0.2 mGy).

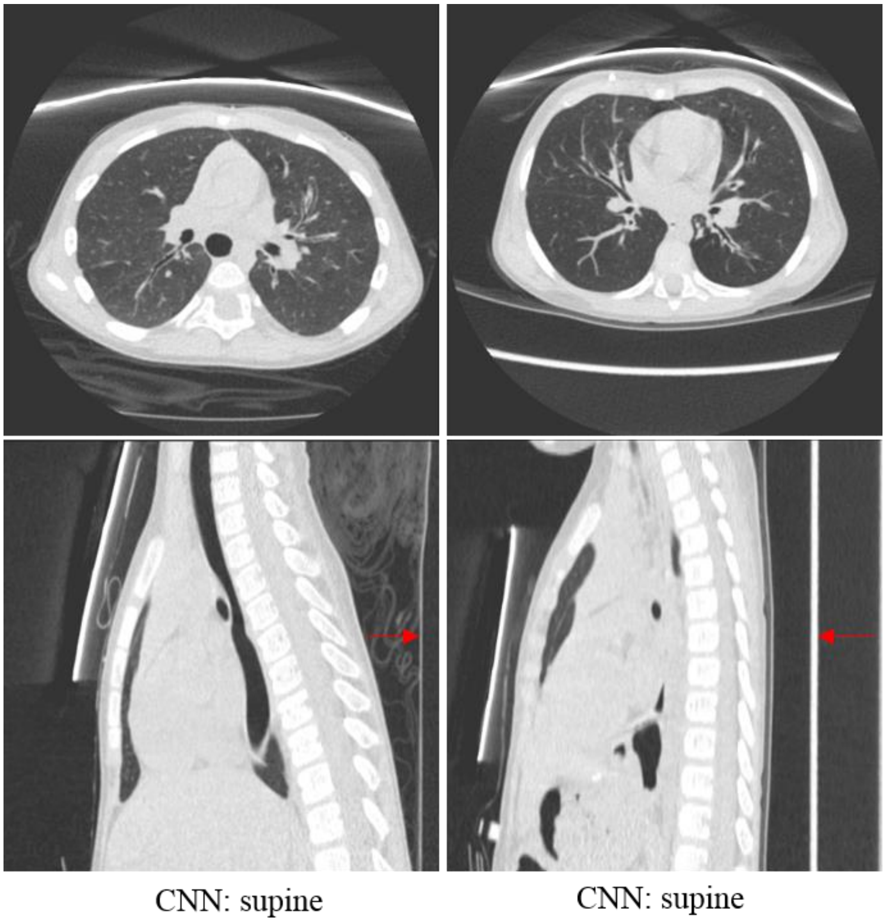


Figure 2.24: Scan posture inference on pediatric scans.

Figure 2.24 shows application of the model to pediatric scans. The first row shows axial images and the second row contains the corresponding sagittal images. They could easily be mistaken for prone scans since there is a table like object above the body, especially from the axial images. Nevertheless, the location of the actual table was confirmed to be behind the body as indicated by the red arrow in the sagittal images. This is another scenario demonstrating the superiority of sagittal images in this application.

Saliency maps from the final classification layer are shown in Figure 2.25. The first two columns are prone scans with the table on different sides. The last two columns correspond to supine scans with table on different sides. Regardless of the location of the table and subject, the most activated pixels are primarily from table and back of the body, including the spine (spinous process more specifically) and muscle/skin. Therefore, the network successfully extracted discriminative features as we anticipated in all possible scenarios to identify the scan posture.

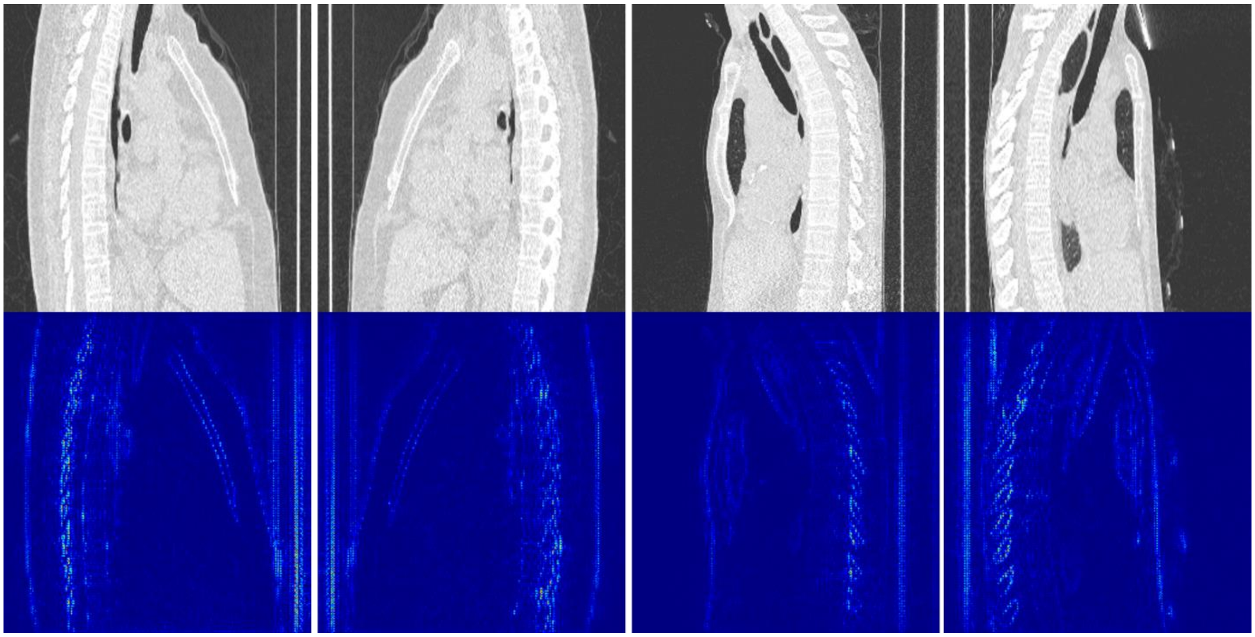


Figure 2.25: Saliency maps of prone and supine scan at last classification layer.

2.3.5 Conclusion

The proposed sagittal image based CNN achieved high accuracy on the test set and demonstrated robustness on both high and low resolutions scans, and regular and ultra-low-dose scans.

2.4 Lung coverage completeness

2.4.1 Introduction

A chest scan can be defined as full lung coverage if it includes the lung apex, intermediate lung and lung base. Although full lung coverage was requested in the National Lung Screening Trial (NLST) protocol[179,180], there are reported imaging defects due to incomplete lung coverage[181,182]. Real world data, e.g., images from PACS or clinical trials, are often a mix of various types, scout scans, biopsy scans, diagnostic scans, etc. The incomplete coverage may result from a specific diagnostic purpose or errors by the CT technologist. As such, it is necessary to screen out partial lung coverage scans since they are not appropriate for many subsequent quantitative analysis, such as lobar segmentation. No DICOM tags, such as “SeriesDescription”, include information regarding lung coverage. A CNN based on axial images was developed to accomplish this task.

2.4.2 Materials and methods

A representative coronal image, such as the middle coronal slice, could be used to check lung coverage. However, for scans where the longitudinal coverage is only slightly incomplete a few coronal slices may appear complete. Sometimes, the lung cutoff is not observed until the coronal slice is very close to the anterior or posterior of the scan. As such, we chose a more reliable approach using axial images. Firstly, we trained a deep learning model to identify whether each axial slice in a scan contained lung or not. After that, the method aggregates individual slice labels to infer whole scan label. A complete lung coverage scan should begin and end with “no-lung” slices and have “lung” slices in the middle, whereas the most superior and/or inferior slices are classified as “lung” for an incomplete lung coverage scan regardless of scan direction (head or feet first).

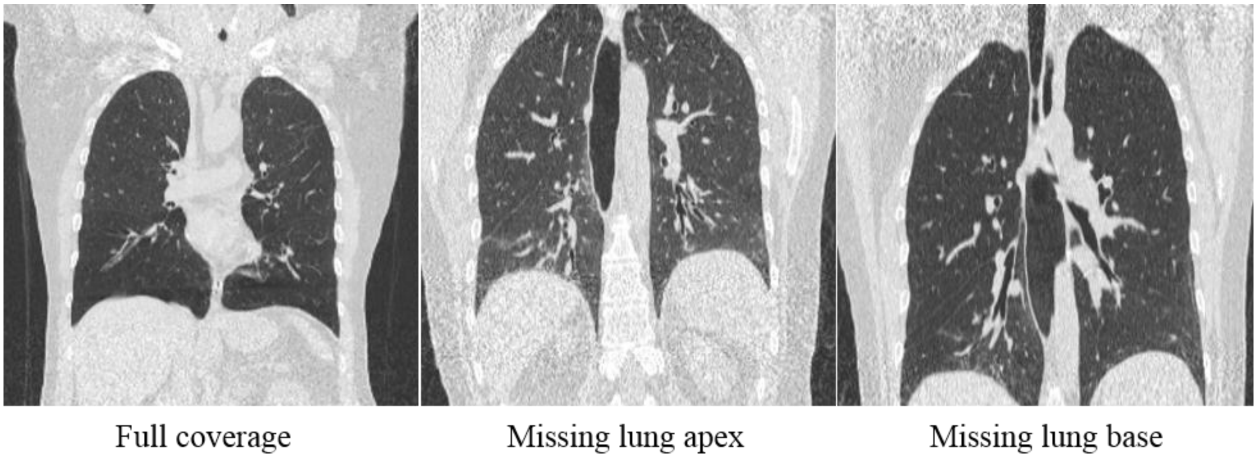


Figure 2.26: Example scans with full lung coverage and missing lung apex or base.

To build the model recognizing lung presence in the axial image, 20,000 axial slices from 210 different scans/patients were used with equal number of lung and non-lung slices. They were collected from 4 clinical trials (2 IPF, 2 COPD) using scanners from SIEMENS, GE, PHILIPS and TOSHIBA. Slice thickness and spacing ranged between [0.6 mm, 3.0 mm] and reconstruction kernels included smooth to sharp. All scans had radiologist approved lung segmentations that were used to determine the individual slice label. Training and test sets were formed with a ratio of 4:1. Figure 2.26

shows examples of complete and incomplete lung scans. The 2nd image lacks the lung apex and 3rd image lacks the lung base.

Figure 2.27 contains example slices with and without lung from superior to inferior. The first row shows typical slices with lung and the second row shows slices without lung at different positions. The lungs on axial images differ in shape and size through the chest. When approaching the lung base, bowel gas that resemble lung may appear and it is critical to differentiate them from the lungs. For example, both images in the last column have bowel gas present but the one in the first row has a small portion of posterior lung near the spine.

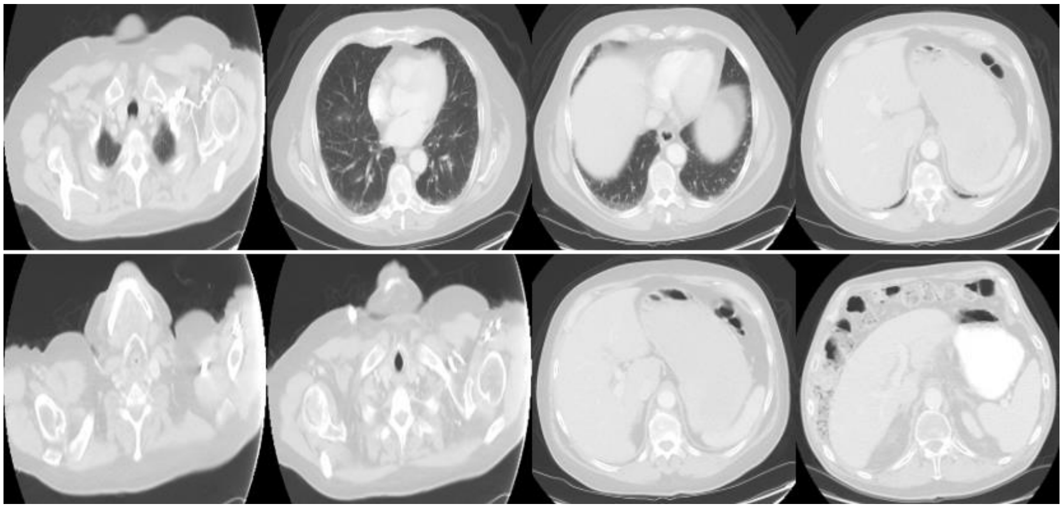


Figure 2.27: Axial slices with (1st row) and without lung (2nd row).

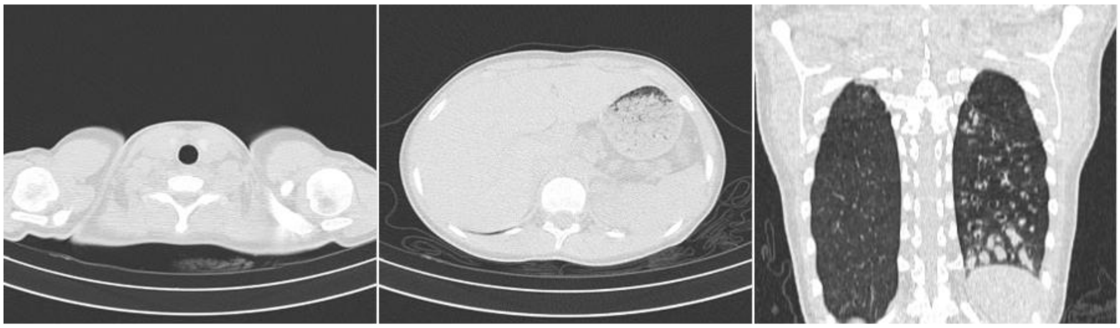
The VGG-16 CNN architecture was also employed in this classification task and is the same as for anatomical coverage identification except that the last layer has two classes. Data augmentation was used, including rotation of [45°, 45°], horizontal flipping, vertical flipping, zoom in/out of [0.8, 1.2], horizontal and vertical shift of [0.2, 0.2]. Random rotation was used in this case since the presence of lung will not be affected by it. At each iteration, the image was be rotated clockwise or counter clockwise by a random degree between 0 and 45°. Similarly, limited translation and flipping of the image does not change the original label. The model was trained from scratch on a NVIDIA TITAN X with GPU memory of 12 GB, using the Adam optimizer, and learning rate of 0.001 with a decay rate of 1e-4.

2.4.3 Results

Both training and validation accuracies were over 99.5% after 15 epochs and about the computation time is around 3 seconds for a scan of 300 slices. The lung coverage detection model achieved 99.1% accuracy on identifying lung presence per axial slice and 100% accuracy on lung completeness recognition at the case level. Figure 2.28 shows two cases with incomplete lung coverage. The first row is an example with an incomplete lung apex (the most superior axial slice contains lung) and the second row corresponds to a case with an incomplete lung base (the most inferior axial slice contains lung). The model was able to correctly classify slices as “no lung” containing only bowel gas rather than lung (see the central image of the first row). Additionally, it was able to correctly identify slices as “lung” with only a small amount of basal lung (see the central image of the second row). The last column shows relevant coronal slices that demonstrating the lung cutoff. In particular, for the case in the second row, the missing right lung base cannot be observed at the very back of body. As indicated by the coronal view in the last column, using representative coronal slices, e.g. middle coronal slice, could miss a scan that is slightly incomplete.

First slice with lung
CNN: lung

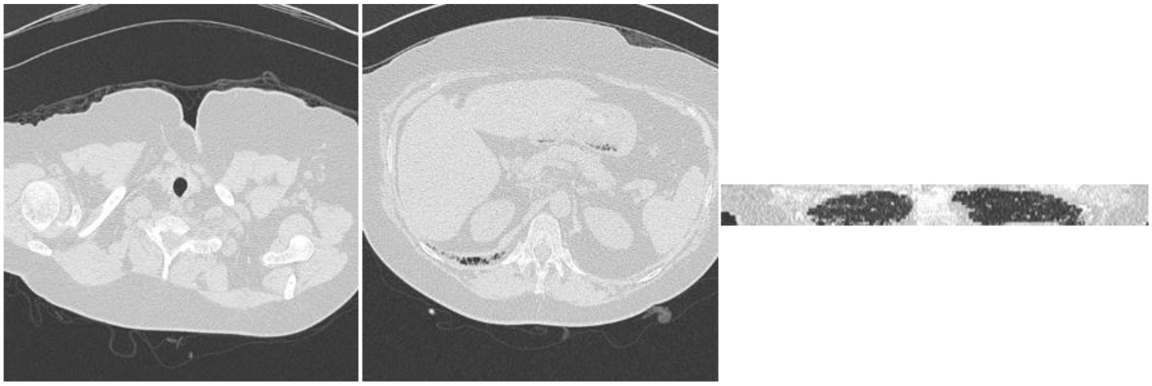
Last slice without lung
CNN: no lung



First slice without lung
CNN: no lung

Last slice with lung
CNN: lung

Figure 2.28: Typical incomplete lung cases with missing apex or base.

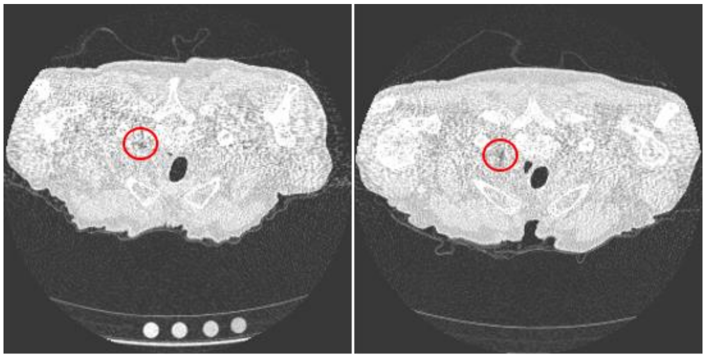


First slice: no-lung by CNN

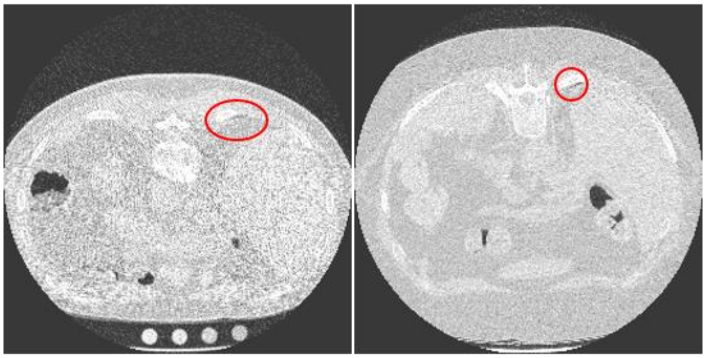
Last slice: lung by CNN

Figure 2.29: An incremental scan identified as incomplete lung coverage by the proposed method.

The model was also evaluated on the incremental dataset (100 scans) to identify lung coverage completeness and 100% accuracy was achieved. Figure 2.29 is an example with slice spacing of 20 mm and 14 slices in total. The last slice was caught by the model with lung presence and the coronal view is displayed for reference.



CNN: no lung CNN: no lung



CNN: no lung CNN: no lung

Figure 2.30: Examples of mis-classified slices from lung apex and base by CNN.

The lung presence axial image classification did not achieve 100% accuracy and some mis-classified slices from the lung apex and lung base are shown in Figure 2.30. All of them were identified as being without lung presence by the model, which is incorrect as indicated by the red circle in each case. Their slice thickness and spacing are 0.625 mm.

2.4.4 Discussion

To avoid the risks inherent in using representative coronal slices, every axial slice of a scan was used to identify the presence of lung. For cases with incomplete lung coverage, the scan direction model was used to further ascertain whether the missing lung is from the apex or base. For the axial lung presence detection model, misclassification sometimes occurs near lung apex or base on high resolution scans. As shown in Figure 2.30, it is very challenging on slices containing just a few pixels of lung. The inference of the overall lung coverage label from the stack of individual slice labels is rarely affected by one or two slice labels unless it happens to start or end with that very slice at lung apex or base. In practice, even such cases of very minor incompleteness were classified as complete lung coverage, it will not affect the following quantitative image analyses, e.g. lung segmentation, emphysema quantification. One potential way to improve performance is by augmenting the training set with more slices from the lung apex and base.

The model was also assessed on the same ultra-low-dose dataset (dose ~0.2 mGy) from lung cancer screening and achieved 100% accuracy in lung coverage completeness. Figure 2.31 is an example of full lung coverage identified by the model. Similar to what was observed in our previous models, the lung presence detection model demonstrated robustness to noise as well.

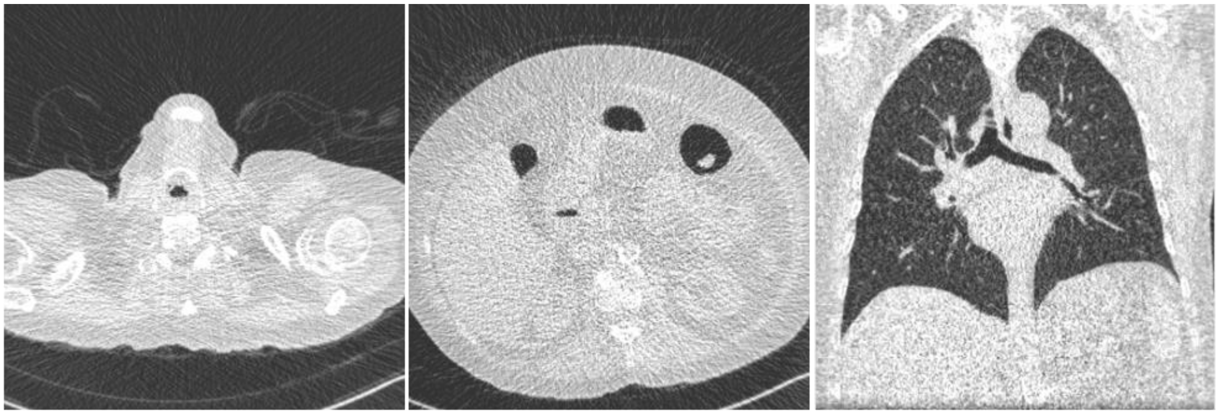

First slice: no-lung by CNN
Last slice: no-lung by CNN
Middle coronal slice

Figure 2.31: Complete lung coverage of an ultra-low-dose scan successfully detected by the model.

The model was also applied to pediatric patients and an example missing the lung base is shown in Figure 2.32, the model successfully detected the last slice as having lung presence. The corresponding coronal image, shown rightmost, demonstrates the incompleteness of the left lung.

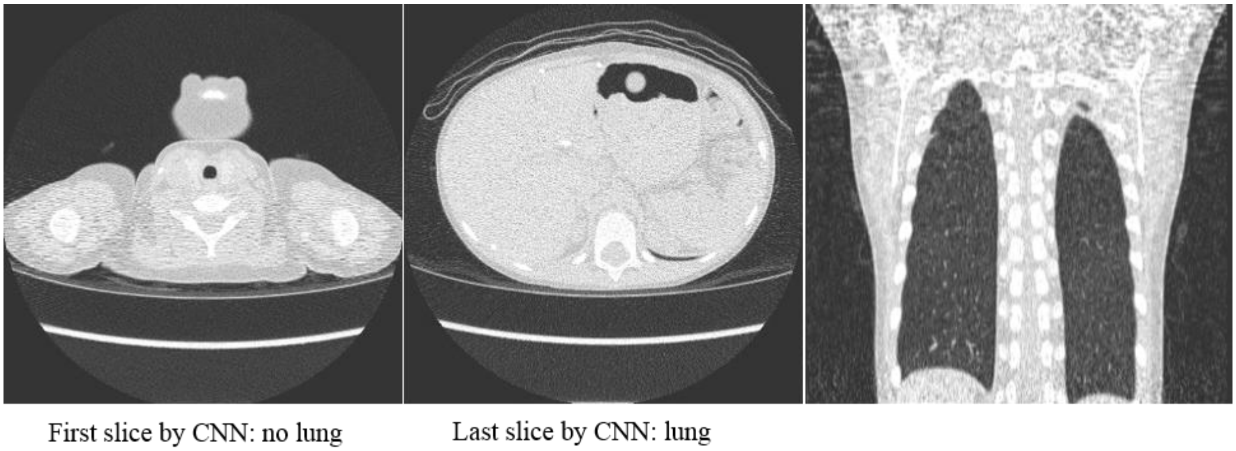


Figure 2.32: A pediatric scan detected with missing lung base by the model.

Saliency maps of lung and non-lung slices at the last classification layer are shown in Figure 2.33 with respect to the input images from superior to inferior. The first row is the input raw slice and the second row is the corresponding maps.

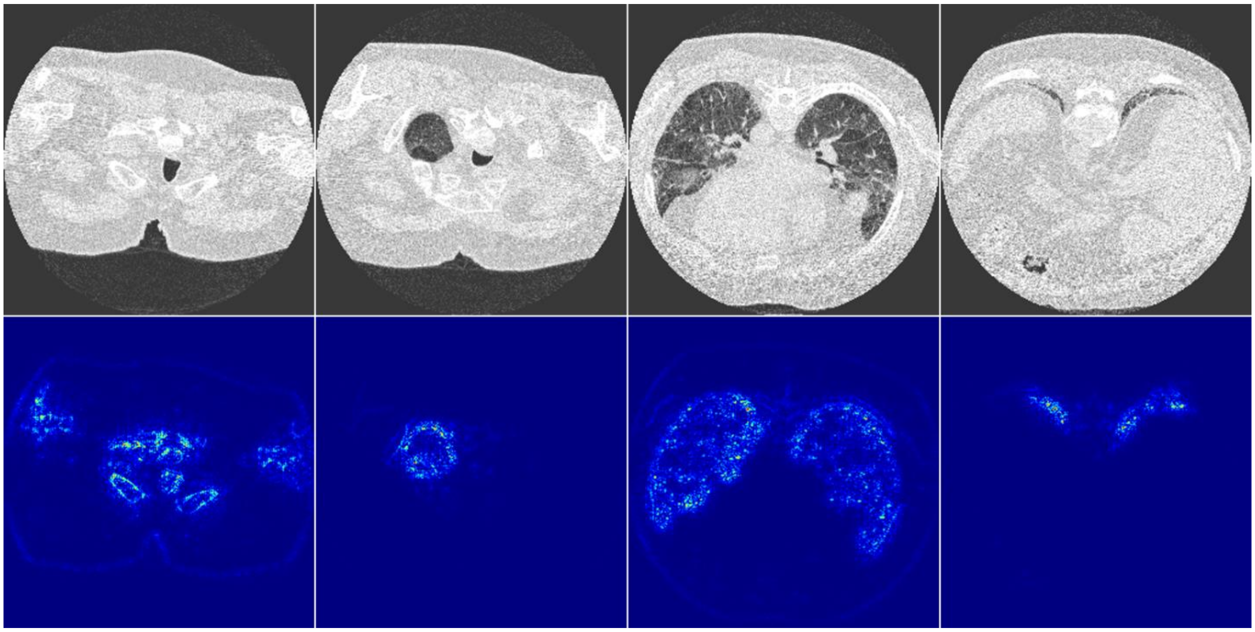


Figure 2.33: Saliency maps of slices with various degrees of lung presence.

For the first column, there is no lung in the images and pixels in high intensity regions, i.e. bone, are highly activated. The model may be using those features to identify ‘no lung’ slices. For the second, third and fourth columns, lung regions of different sizes are present and only pixels within lung are activated by the network. It is interesting that regions with similar intensity of lung (on the 2nd and 4th images), such as trachea, are not activated by the feature extraction. As indicated by these results, the model is able to differentiate real lung from misleading regions, e.g., bowel gas.

2.4.5 Conclusion

Lung coverage completeness classification was achieved by aggregating individual axial slice labels sequentially, and the recognition of lung presence on each slice was accomplished by a deep convolutional neural network. This strategy has been successfully applied to scans of high and low resolution, diagnostic and low-dose, and adult and pediatric patients.

2.5 Contrast usage

2.5.1 Introduction

CT can be ordered either with or without contrast. Contrast CT is able to enhance the intensity of target tissue, e.g. blood vessel, tumor, etc. and assist physicians by providing better structural and functional information. In routine practice, contrast CT is usually applied to evaluate suspicious masses in neck, chest and abdomen[183–186]. Specifically, both malignant masses could potentially develop into adenopathy, lymphoma, sarcoidosis, etc.[187–190]. Pulmonary embolism is one of most common cause of acute cardiovascular disease[191–195] and contrast CT, especially CT Angiography (CTA), is the favored method of diagnosis [196–199]. Intravenous contrast is used to assess vascular related disease, such as aneurysm, clots, and vasculitis[200–202]. Lastly, contrast CT is also applied to lung nodules [203–206].

The presence of enhancement has a considerable impact on the subsequent image analysis results and different strategies can be used to take this into account. For example, case dependent normalization may be required for generalizable analysis. As such, it is crucial to differentiate between contrast and non-contrast scans. Unfortunately, DICOM header tags, such as “SeriesDescription”, often have no relevant information. Accordingly, an image-based deep learning method was proposed for this classification problem.

2.5.2 Methods

In theory, the CNN could be directly applied to raw image assuming the network was able to automatically extract discriminative features from areas with enhancement to differentiate between contrast and non-contrast images. However, there are also other tissues with high HU intensity (bone, calcification, implanted devices, etc.) that may confuse the network. Furthermore, given the network is a “black-box” determining whether this erroneous learning has occurred may be difficult to determine. However, it is observed that enhancement in the aorta is consistently present in contrast-enhanced chest CT scans since intravenous (IV) contrast is most commonly used. Figure 2.34 shows examples of both contrast and non-contrast scans.

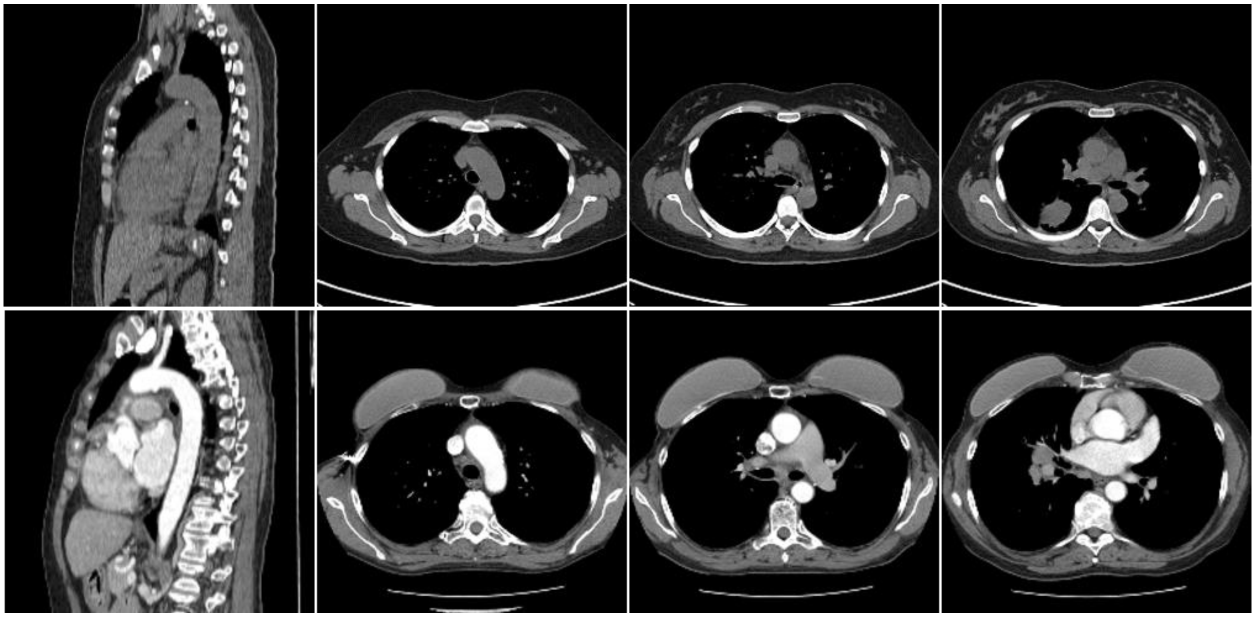


Figure 2.34: Comparison of non-contrast (1st row) and contrast scans (2nd row).

For contrast scans, the intensity within the aorta is over 90 Hounsfield Units (HU), whereas it is less than 50 HU on non-contrast scans[207,208]. Therefore, the aorta is an ideal anatomic landmark in this scenario and the contrast detection task can be reformatted as aorta detection and segmentation. Overall, the contrast detection algorithm can be broken down into: 1. Aorta segmentation; 2. Computing the average intensity within aorta; 3. Determining the presence or absence of contrast using a threshold, e.g., 80 HU. Provided the aorta is segmented accurately, the identification of contrast usage is straightforward. The aorta consists of the aortic arch, ascending and descending aorta. The ascending aorta starts at the aortic root and extends up until the aortic arch. The descending aorta starts at the aortic arch, goes through chest and continues down into abdomen[209,210]. The descending aorta preserves a relatively consistent circular shape and thus it is preferred as our segmentation target. Traditional aorta segmentation methods on chest CT usually apply the Hough transform to detect and track the circular shape because of the tubular structure and treat the whole 3D aorta as cylinders. Usually a seed point needs to be set, either manually[211] or automatically[212], and the segmentation is challenging especially on non-contrast scans using this method. A similar approach using Kalman filtering has also been proposed[213]. 3D level set[214] and graph cut[215] algorithm were applied to further refine the tubular surface based on the previously detected centerline of aorta. Computation efficiency is a big issue when using such algorithms. Dynamic programming was also used to localize aorta and search for the aortic boundaries. Atlas and deformable atlas-based methods have also been presented[216,217] and the aortic boundaries are determined by relying on high contrast edges in these methods. Deformable mesh adaptations was used to match the tubular structure of the aorta[218]. Results of multiple atlases were fused by registration to improve compatibility with aortas of varying sizes[219]. To summarize, conventional methods are not generalizable to images with high variability in the aorta. Further, the segmentation is time-consuming. Most importantly, they are mostly evaluated on contrast enhanced scans and performance markedly declines when applied to non-contrast scans. Recently, a fully convolutional network based on a dilated CNN[220] was applied to segment the aorta on low-dose CT[221] and it is same approach as Google’s deeplab[222]. Three CNNs using images at axial, coronal, sagittal planes were trained individually and the final 3D segmentation results derived by averaging the three probability maps. Only 10 chest CT scans were used in training and testing respectively and achieved a Dice score of 0.88±0.05 on descending aorta segmentation.

To segment the aorta accurately and efficiently, we applied a fully convolutional neural network and we extended the previous VGG-16 classification architecture to a segmentation network by adding a decoder section to recover the original image resolution. As shown in Figure 2.35, the left encoder is exactly same as original classification network but without fully connected layers.

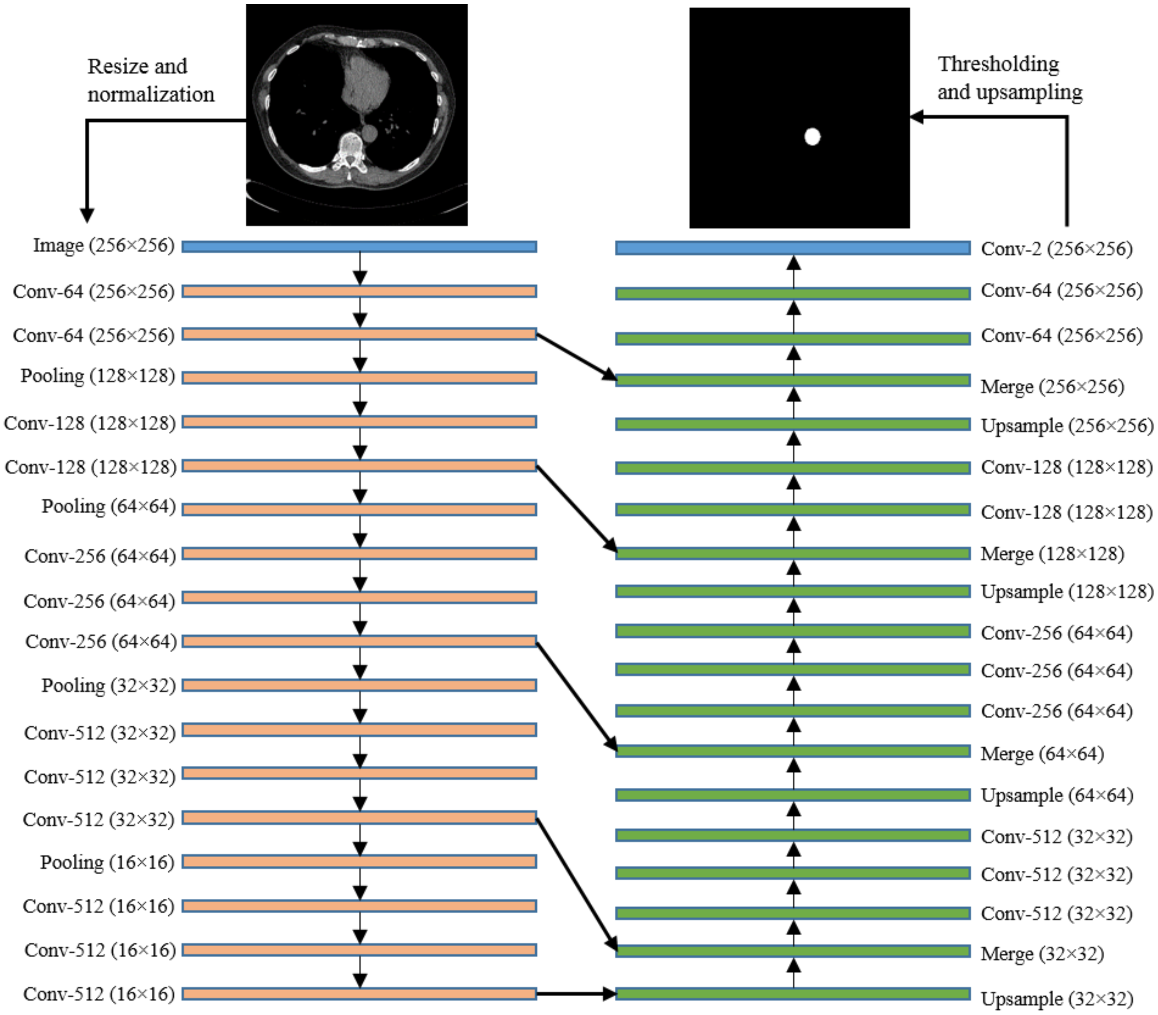


Figure 2.35: VGG-16 based architecture for aorta segmentation.

All the convolutional kernel sizes are 3 × 3 with padding size of 1 × 1 to keep the size intact. Down-sampling was performed by max-pooling with a size of 2 × 2. Overall, the original input image was progressively scaled down 4 times (256 × 256 to 16 × 16).

The right decoder has a corresponding 4 times up-sampling (size of 2 × 2) using transposed convolution[223] (16 × 16 back to 256 × 256). Similar to the strategy of the U-Net, the number of feature maps in the de-convolutional layer matches the number in the corresponding convolutional layer to achieve a symmetric architecture. Also, skip-connections were used in the up-sampling process to incorporate finer details from the lower layers as well as abstract and sematic information from higher layers. After achieving the target resolution, a final 3 by 3 instead of 1 by 1 convolution was applied on the aggregated hyper-columns to achieve pixel-wise classification. This allows incorporation of neighboring information in the class inference for a specific pixel. The final output of the network is a probability map with the same resolution as the input that gives the probability of each pixel belonging to the aorta. The last step is to generate a binary mask by thresholding the probability map at 0.5 and up-sample it to the original image size.

Since the size of the aorta is much smaller than the background, i.e. the number of positive pixels and negative pixels are extremely unbalanced, the regular loss function using cross-entropy would lead to biased model. To solve this class unbalance problem, weighted cross-entropy was used as the cost function during training with a ratio of 150:1 to compensate for the aorta size. Data augmentation was applied in training the segmentation model, including rotation with the range of [90°, 90°], horizontal and vertical shift of [0.2, 0.2], zoom in/out of [0.8, 1.2], horizontal and vertical flipping. The model was trained from scratch on a NVIDIA TITAN X with GPU memory of 12 GB, using the Adam optimizer, and learning rate of 0.001 with a decay rate of 1e-

4. The augmentation was very helpful since the dataset used was not very big as described in materials section below.

2.5.3 Materials

116 chest CT scans from 70 different subjects from a lung cancer trial were used in this study. 70 of them were non-contrast scans and 46 were contrast scans. Because the aorta segmentation is more difficult on scans without enhancement, slightly more non-contrast scans were used. Scanners from two manufacturers, SIEMENS and GE were used and they were reconstructed with smooth to sharp kernels with slice thickness and spacing ranging between [1 mm, 3.0 mm]. In total, 8,447 slices were used having manual segmentation of the aorta. Since it is a relatively small data set and our goal is contrast usage detection, 500 slices (half from contrast scans and half from non-contrast scans) were saved as a test set for the aorta segmentation evaluation and the rest were used in training the segmentation model. An additional independent set of 500 scans with only a scan level label (contrast or not) were used for contrast detection evaluation, and contrast/non-contrast scans accounted for 250 each in this set. They were collected from 5 clinical trials (2 IPF, 2 COPD, 1 scleroderma).

Because our goal is contrast detection only, the reference segmentation of the training set done by image analysts did not cover the whole thoracic descending aorta. An example of aorta segmentation in training set is shown in Figure 2.36.

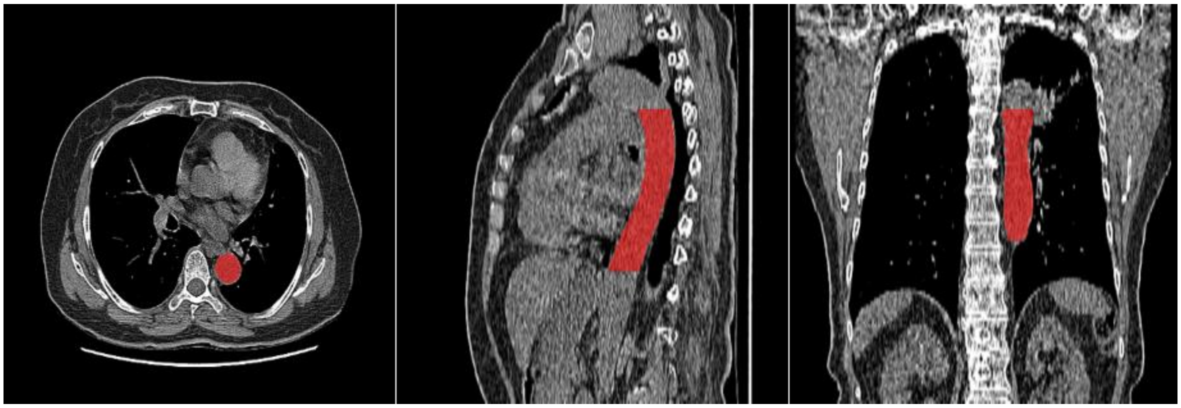


Figure 2.36: Reference segmentation of aorta from training set.

2.5.4 Results

The model was trained for 15 epochs, and after that there was no significant performance improvement on the training or validation sets. It takes about 5 seconds to segment the aorta on a chest scan with 300 slices using the machine mentioned previously. Since our ground truth does not cover the whole aorta, we evaluated the aorta segmentation based on 500 selected slices and achieved a dice coefficient of 0.932±0.14. In terms of contrast detection, the system achieved an accuracy of 100% on the test set of 500 scans using 80 HU as a threshold. Examples of aorta segmentation on contrast and non-contrast scans are shown in Figure 2.37 and Figure 2.38. Figure 2.39 is a non-contrast scan from the public LObe and Lung Analysis 2011 (LOLA11)[159] data set with a warped aorta that was successfully segmented by the model.

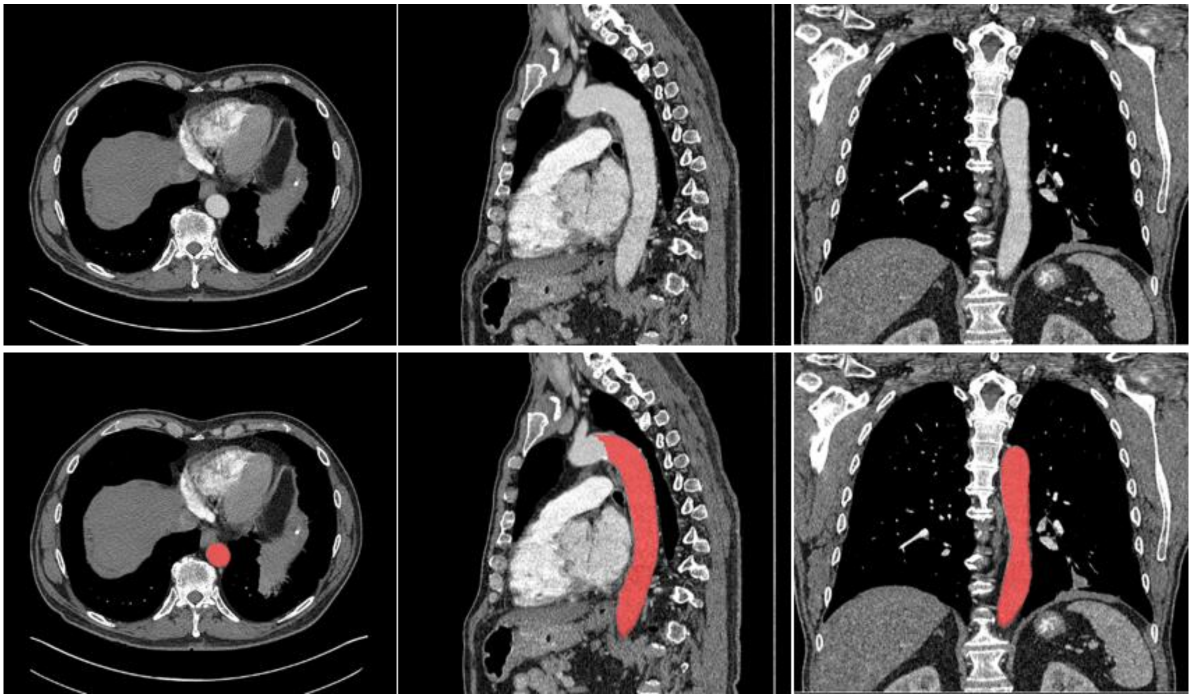


Figure 2.37: Example segmentation on a contrast scan by the model.

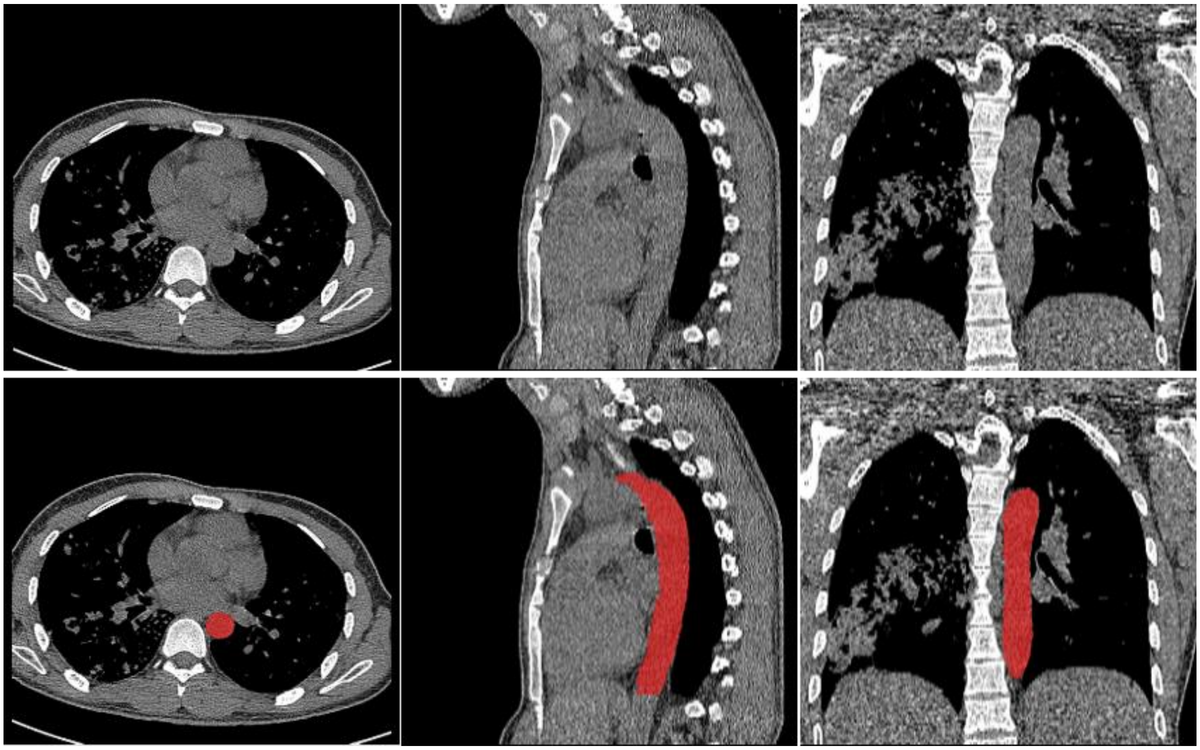


Figure 2.38: Example segmentation on a non-contrast scan by the model.

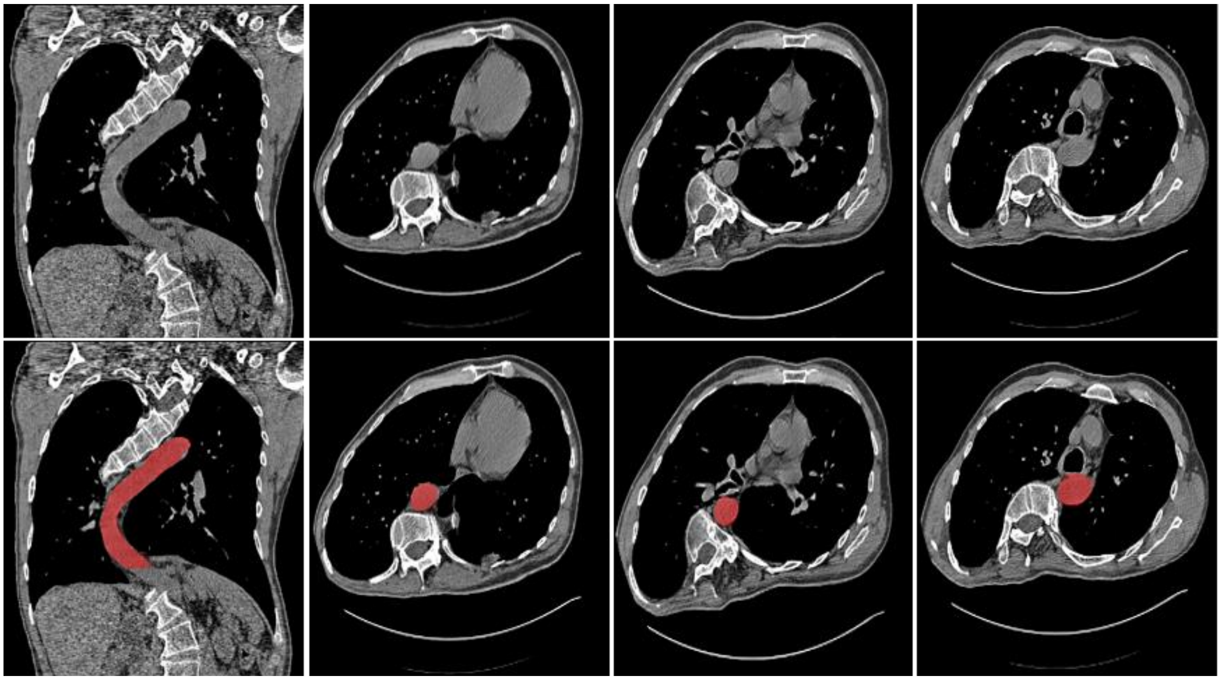


Figure 2.39: Segmentation on a non-contrast scan from LOLA11 with a warped aorta by the model.

Since we do not have a reference segmentation for this LOLA11 data set, no quantitative assessment regarding the aorta segmentation was performed.

2.5.5 Discussion

The aorta segmentation is an intermediate step to detect contrast usage and under/over-segmentation of a few pixels is unlikely to affect the final result. Therefore, it is not necessary to achieve perfect aorta segmentation in this problem. However, the model did achieve high performance compared to other published methods. Furthermore, there are some existing chest CT analysis algorithms dependent on image denoising using standard deviation within aorta and thus our method could automate that process without manual intervention.

Because the proposed CNN is 2D axial slice based, it is applicable to incremental scans as well. In the example from LOLA 11 in the results section, the 2D

CNN would probably outperform 3D CNN since the target descending aorta is not the typical straight tubular structure. Compared to traditional approaches, our method works on both contrast and non-contrast scans and the computation speed is good for high throughput processing in clinical practice. The segmentation model was built only using diagnostic scans but was also evaluated on ultra-low-dose scans from lung cancer screening. Figure 2.40 is an example scan reconstructed with a medium kernel and slice thickness/spacing of 1 mm and the dose level is only 10% of the original screening scan and 1/75 of a diagnostic scan. Since we do not have ground truth segmentation on this ultra-low-dose scan, a quantitative evaluation such as dice score was not performed.

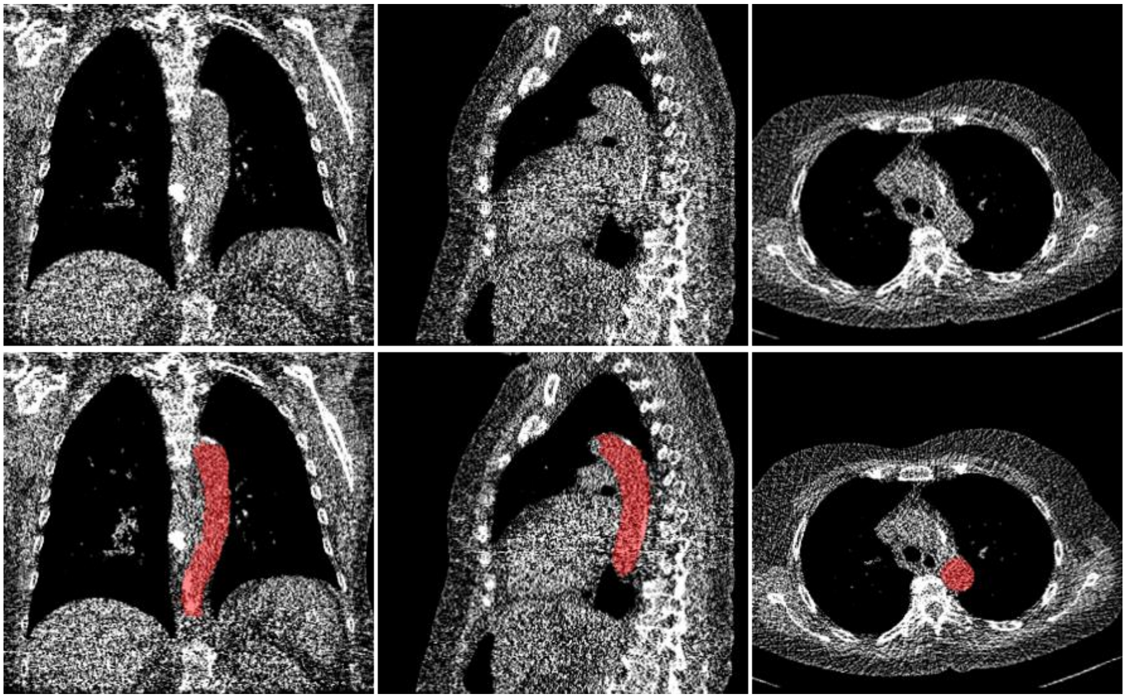


Figure 2.40: Aorta segmentation on an ultra-low-dose scan by the model.

Application of the aorta segmentation model to pediatric scans was performed as well, and an example is shown in Figure 2.41 with slice thickness of 0.48 mm and slice spacing of 1.25 mm.

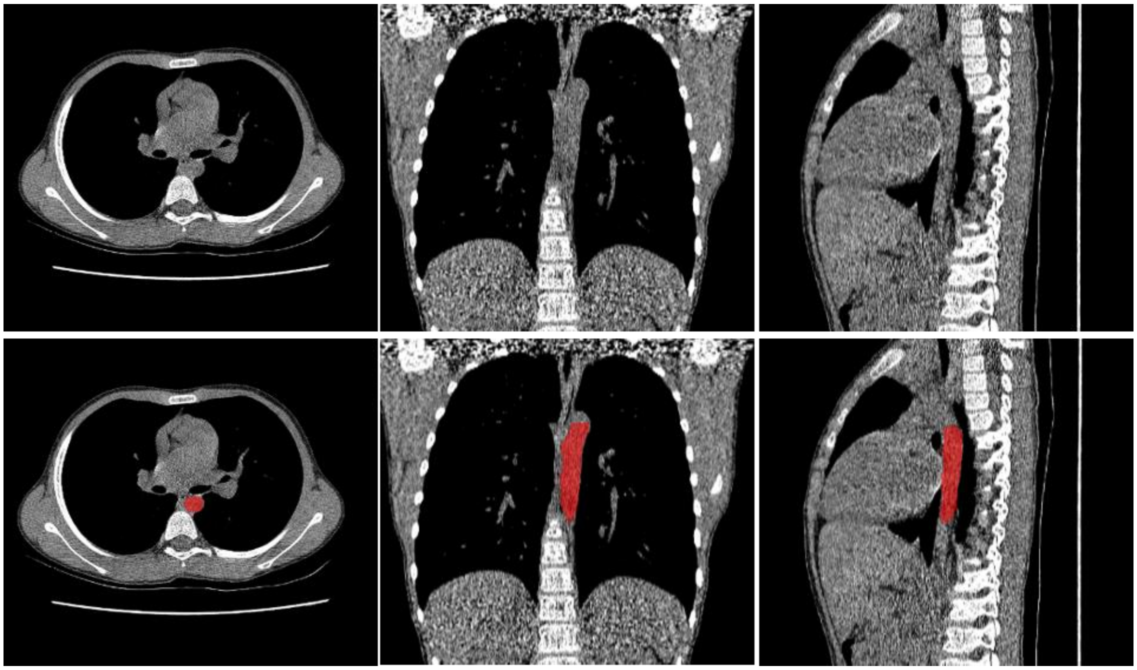


Figure 2.41: Example of aorta segmentation on a pediatric scan.

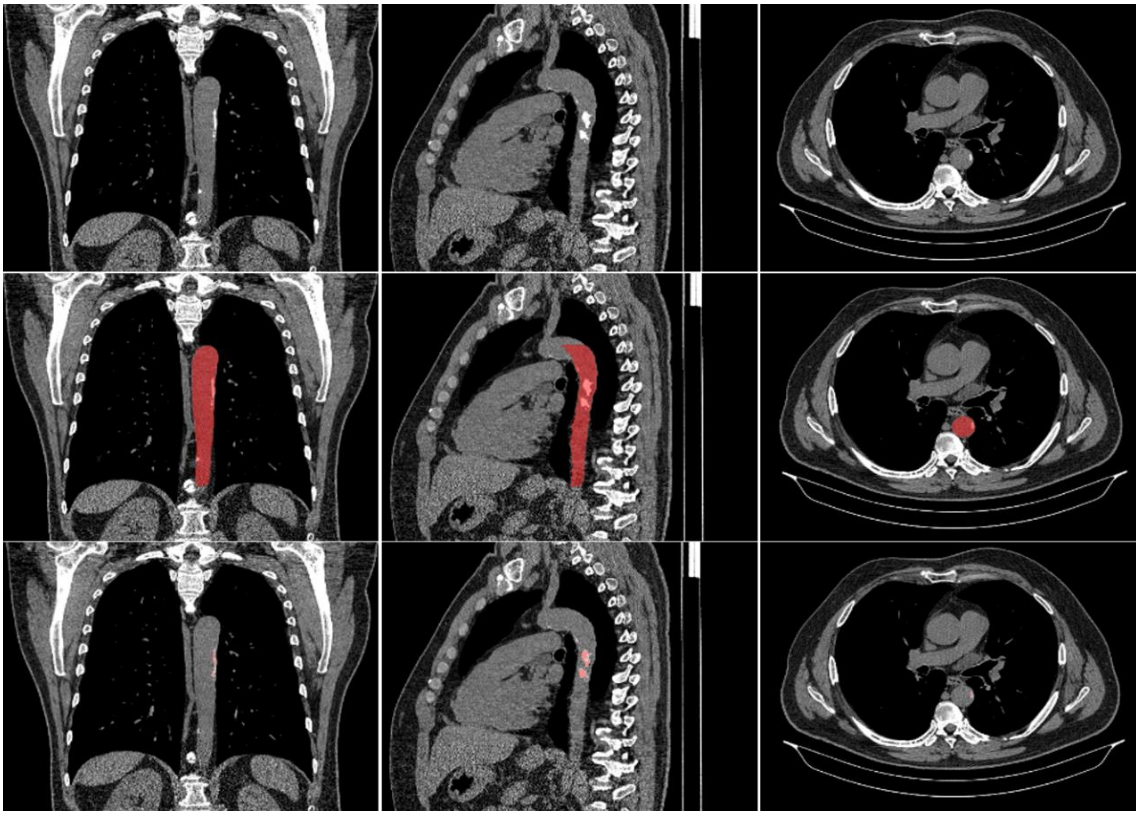


Figure 2.42: Detection of calcification within aorta using HU=400 as threshold.

Since the model segments the aorta accurately, other vessel related disease analysis could be performed. For example, one application is measurement of calcification within the vessel. Calcification has an intensity range from 200 to 800 HU and we can compute the amount of calcification within aorta using a simple threshold. Figure 2.42 is an example using 400 HU as threshold to detect calcification. Moreover, the descending aorta that is closer to the left lung, could be utilized as anatomical landmarks for automatic quality control (QC) to verify the location of left and right lungs.

Compared to the dilated CNN presented on low-dose CT[224], we differ in the following aspects: 1. the goal of the study is slightly different and our model may not segment the entire descending aorta, especially the aortic arch and the abdominal aorta; 2. their network is actually 2.5D using images from 3 planes whereas we only used 2D axial images and our model has higher complexity in terms of depth and width; 3. we used a much bigger training and test set with more diversity including both contrast and non-contrast scans; 4. their network was trained using low-dose scans and our model was trained using diagnostic scans, but it demonstrated robustness when applied to low-dose and even ultra-low-dose scans.

Saliency maps at the final convolutional layer of the network with respect to contrast and non-contrast images are shown in Figure 2.43. Pixels at the circumference of descending aorta are highly activated in both cases and the ascending aorta is also mildly visible on the maps. Since the network is able to differentiate ascending and descending aorta on a 2D slice, those lateral bright regions near the boundaries or corner of the image are possibly used to determine their locations.

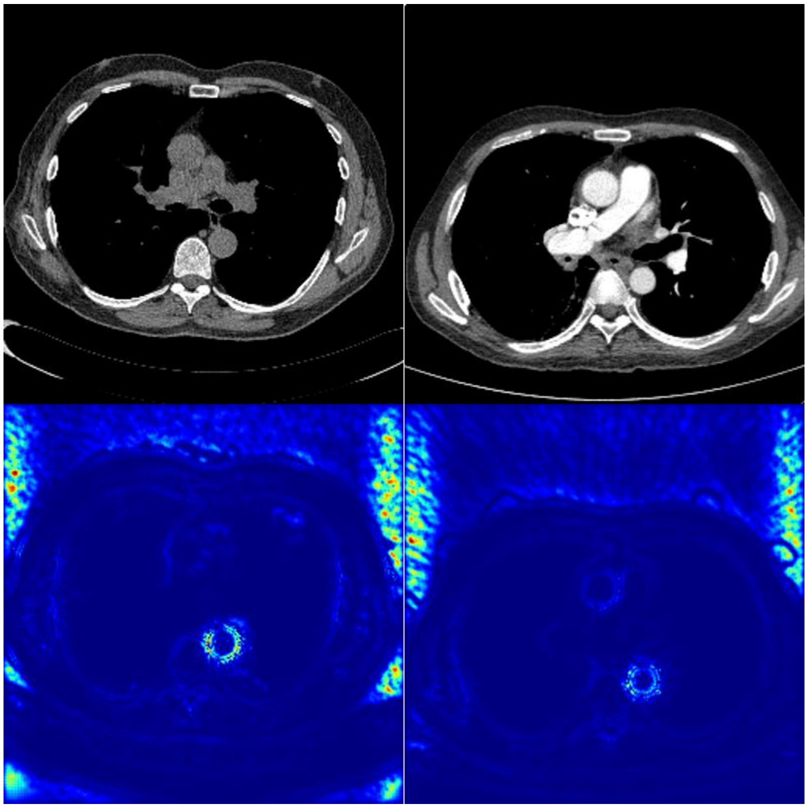


Figure 2.43: Saliency maps of non-contrast and contrast scan from the proposed aorta segmentation model.

The current model was only trained to segment the descending aorta and the next step is to expand the model to include the ascending aorta and potentially employ a 3D CNN if we are able to collect better training data in the future.

2.5.6 Conclusion

We formulated the contrast detection problem as one involving aorta segmentation. The proposed CNN was able to accurately segment the descending aorta on both contrast and non-contrast scans and thus enable reliable detection of contrast. It demonstrated strong performance on low dose scans and pediatric scans.

2.6 Breath-hold types

2.6.1 Introduction

Patients undergoing CT scans are usually asked to hold their breath during scanning at full inspiration, i.e., at total lung capacity (TLC), or at full expiration, i.e., at residual volume (RV). Typically, inspiratory chest CT scans are the preferred imaging method in patients with pulmonary disease, whereas expiratory scans are sometimes used to demonstrate pathophysiological alterations such as air trapping. For example, TLC scans are commonly used in evaluation of most diffuse lung diseases, including fibrosis, chronic obstructive pulmonary disease (COPD), etc.[225–228]. RV scan has been used to reveal air trapping in patients with airway related diseases, such as bronchiectasis, sarcoidosis, and asthma[229–231]. In some studies, the subject is imaged at both breath-holds [232–234]. For example, measurement of lung density and air trapping on patients with cystic fibrosis[235]. Due to the large air volume difference within the lung after full exhalation or inhalation, resulting computations, e.g. lung density, volume, will be quite different. It is essential to correctly identify and consider the patient breath-hold when drawing conclusions based on quantitative lung analyses. The DICOM tag “SeriesDescription” often fails to provide any information regarding breath-hold and may be incorrectly entered if there is. As such, it is necessary to identify the breath-hold types using the images and a deep learning method was proposed.

2.6.2 Methods

Lung volume alone is not sufficient to differentiate between TLC and RV scans since it is impacted by a variety of factors, such as gender, age, pulmonary disease, etc.

Clinically, the compression of the trachea is often used by radiologists to identify RV scans. It is characterized by the collapse of carina and the posterior wall of trachea bows forward[62–66]. In contrast, the shape of the carina is quite circular on TLC scans. Some examples of TLC and RV axial images are shown in Figure 2.44 and the carina shapes are noticeably different. Therefore, the breath-hold identification can be divided into three steps: 1. segmentation of the trachea, 2. Identification of the carina, 3. Identification of RV or TLC based on the shape of the carina.

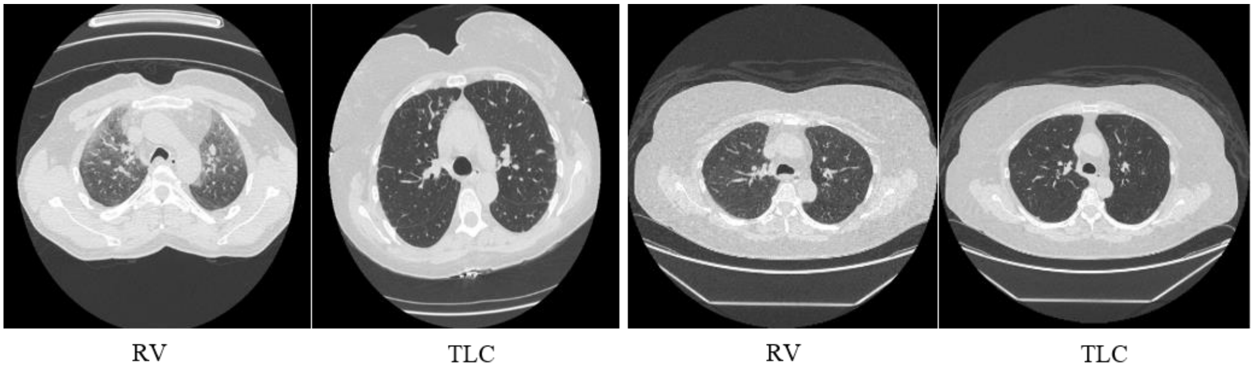


Figure 2.44: Example RV and TLC scans demonstrating differences in carina shape.

Threshold-based region growing methods are most commonly used for airway segmentation on chest CT[241–243]. A seed point is placed manually or automatically[244] and the algorithm starts growing by adding neighboring voxels within a pre-defined intensity range. The biggest issue with these approaches is leakage out of the airways during the growing phase. For example, noise pixels or lung can be mistaken as airway if they share similar intensity in the image. Several methods have been proposed to solve the leakage problem. Morphological operators for 3D filtering and leakage removal were applied to improve the segmentation of peripheral branches[245–247]. A fuzzy connectivity region growing[248] was proposed to prevent leakage by taking advantage of the cylindrical shape of different airway branches. The centerline of the trachea was extracted to track and guide the propagation path of growing[249–251]. A volume of interest (VOI) was introduced by dividing a lung segment into sub-regions and apply topology of airway tree based on the previous VOI to avoid leakage[252,253]. Machine learning methods were also proposed to classify voxels into air and surrounding tissues[254]. Lo et al.[255] presented an appearance model for airway pixel classification and adjacent vessels were segmented to measure their orientation similarity to the candidate airway. To help minimize the cost of fast marching, a multi-scale wellness measure to detect the bronchial wall was used to compute a probability model of the airway[256]. A ConvNet[257] was proposed to detect and remove leakage and prune the segmentation after applying a similar approach as Lo et al.[255]. A 2.5D convolutional neural network[258] using axial, coronal and sagittal slices was applied to classify pixels at branches after initial airway segmentation by region growing. A 3D U-Net like network was trained to segment the whole airway tree[259]. A sliding window with a stride of 104 pixels along z-axis was applied to extract a small path or chunk from the original scan to achieve consistent size input to the network. Probability maps were aggregated to reconstruct the full size airway segmentation. It was trained with 12 chest scans and tested on 6 images and achieved an average Dice coefficient of 0.8.

The VGG-16 based segmentation architecture used in the aorta segmentation was also applied in this task. Because the size of the trachea is much smaller in comparison with background, the same weighted cross-entropy loss function (with ratio of 150:1) introduced in the aorta segmentation was again used. The same data augmentation techniques were applied. A connected component analysis was performed to extract the largest mask and eliminate small isolated artifacts.

After trachea segmentation, the carina was identified by searching slice by slice until bifurcation. Using the previous model for scan direction detection, the search direction was corrected if a scan was feet-first. To train the carina-based breath-hold classification, multiple slices above the bifurcation were selected as carina slices and majority voting was used during the inference stage to generate the final output.

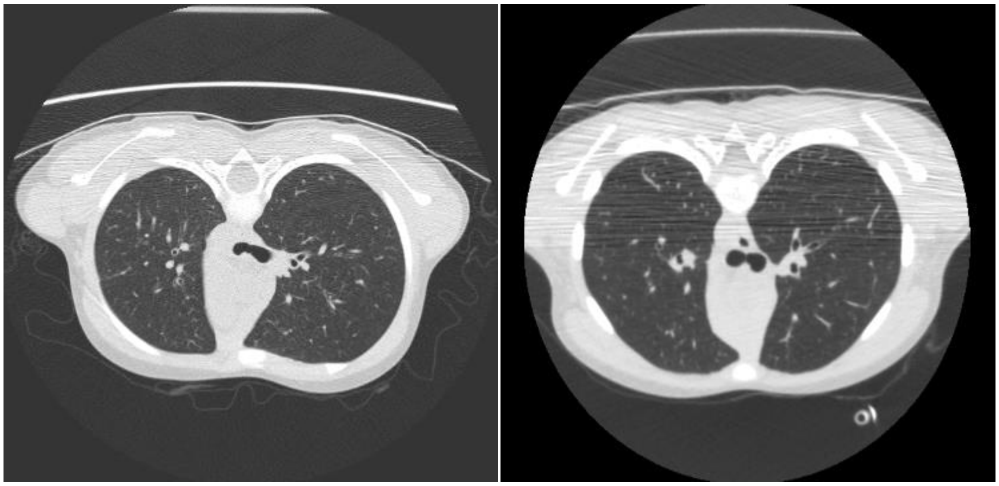


Figure 2.45: TLC scans with the anterior membrane of the carina collapse backward.

Figure 2.45 shows two special TLC scans where the anterior membrane collapse is backward and the spine is up. Such a carina shape could potentially confuse the classification model compared to real RV scans as shown in Figure 2.46. The first row is the raw slice with the related carina segmentation of a TLC scan where the spine is up. The second row is the raw slice with carina segmentation of a RV scan where the spine is down.

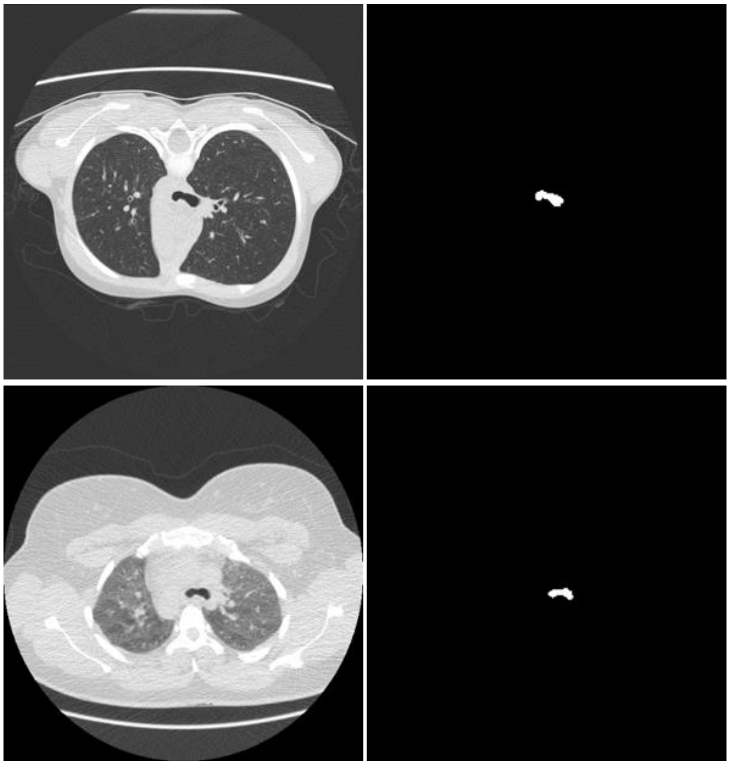


Figure 2.46: Potential confusion caused by uncertainty of spine location.

In order to exploit the shape feature of the carina, it is vital to identify the spine location to confirm the posterior wall is down in the image before employing them in breath-hold classification.

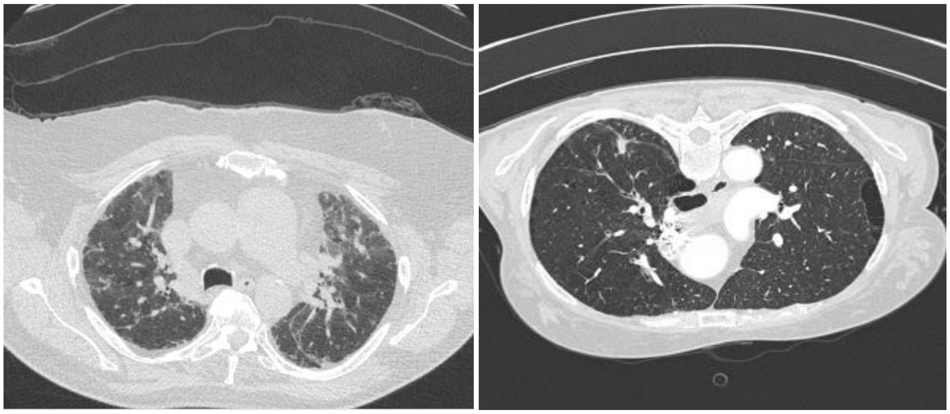


Figure 2.47: Inconsistence between scan posture and spine location.

Since we have built a scan posture model, we could infer an image is spine down (if it is supine) or spine up (if it is prone). Unfortunately, the assumption of correlation between scan posture and spine location is not always true. Figure 2.47 is an example showing exceptions in both cases. The left one is a prone scan but the spine is down and the right one is a supine scan but the spine is up.

Therefore, the previous scan posture classifier cannot be used to determine the spine location and an independent CNN was proposed to accomplish this using axial images. Different from the prone vs. supine classification, this model is supposed to target only on the position of spine instead of table location.

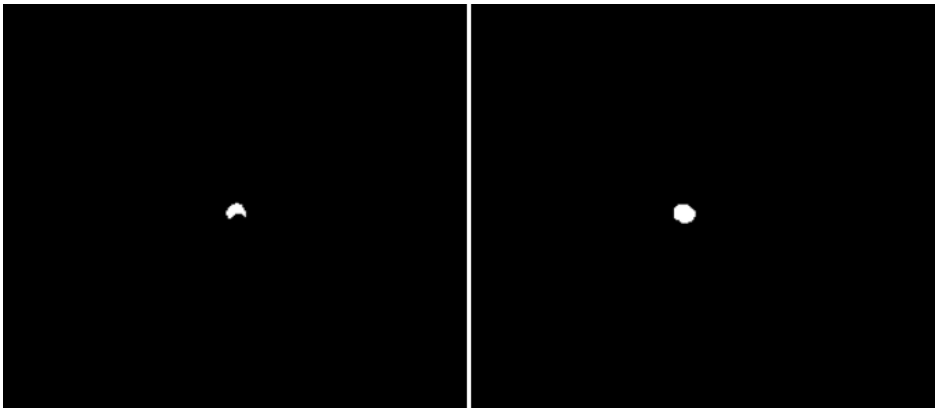


Figure 2.48: Example masks of carina on RV and TLC scan.

The last step is to build the breath-hold recognition model based on the shape of the carina. Since we have performed trachea segmentation, mask images of the carina will be used to simplify the complexity of network. Figure 2.48 shows two mask images of the carina from RV and TLC scans, respectively.

To build classifiers differentiating between spine up and down as well as RV and TLC, the VGG-16 network was used again. Data augmentation was applied in training the segmentation model, including rotation range of [90°, 90°], horizontal and vertical shift of [0.2, 0.2], zoom in/out of [0.8, 1.2], horizontal flipping and vertical flipping for the spine detection model. The models were trained from scratch on a NVIDIA TITAN X with GPU memory of 12 GB, using the Adam optimizer, and a learning rate of 0.001 with a decay rate of 1e-4. Similar data augmentation was used in training the spine detection model and the breath-hold recognition model.

2.6.3 Materials

356 scans (32,298 slices) of 92 patients from multi-center clinical trials were used to build the trachea segmentation model. 176 were RV scans and 180 were TLC scans. Scanners from two manufacturers, SIEMENS and GE, were used and images were reconstructed with smooth to sharp kernels with slice thickness and spacing ranging from [0.6 mm, 1.5 mm]. An independent 1,000 slices were used for trachea segmentation evaluation. The reference segmentation of the training set was done by thresholding plus manual editing, covering the trachea and left/right bronchi beyond the bifurcation. Because the goal of the trachea segmentation was to find the carina, it was not necessary to segment the whole airway tree which alleviated the annotation workload for the image analysts. An example of the manual segmentation is shown in Figure 2.49.

Figure 2.49: Reference segmentation of trachea on training set.

To build the model for detecting the spine location, 40 scans from different patients were used from 1 IPF trial and 1 COPD trial. In total, 21,804 axial slices were used and they were split into training and test sets in a 4:1 ratio. The number of spine up and down slices were balanced in both sets and labels were from image analysts.

To build the model differentiating RV and TLC breath-holds, 1,633 scans from 495 patients were used, including 11,948 RV slices and 11,953 TLC slices and the test set accounts for 20% of the cases. They were collected from 11 clinical trials (4 IPF, 3 COPD, 1 scleroderma, 1 NSCLC, 1 lung cancer screening, 1 tuberculosis). The two classes were balanced in both sets. Their labels were manually confirmed. Lastly, the training and test sets were independent of the dataset used in scan direction classification and spine localization.

2.6.4 Results

The trachea segmentation model was trained for 21 epochs, beyond that there was no significant improvement in either the training or validation sets. The algorithm takes about 5 seconds to segment a scan with 300 slices. 1,000 slices were used in the trachea segmentation evaluation and the dice coefficient was 0.941±0.06. Figure 2.50 shows an example of the trachea segmentation on a scan with slice thickness of 1 mm. Figure 2.51 shows the result from the model on a scan with an unusual chest shape from LOLA11.

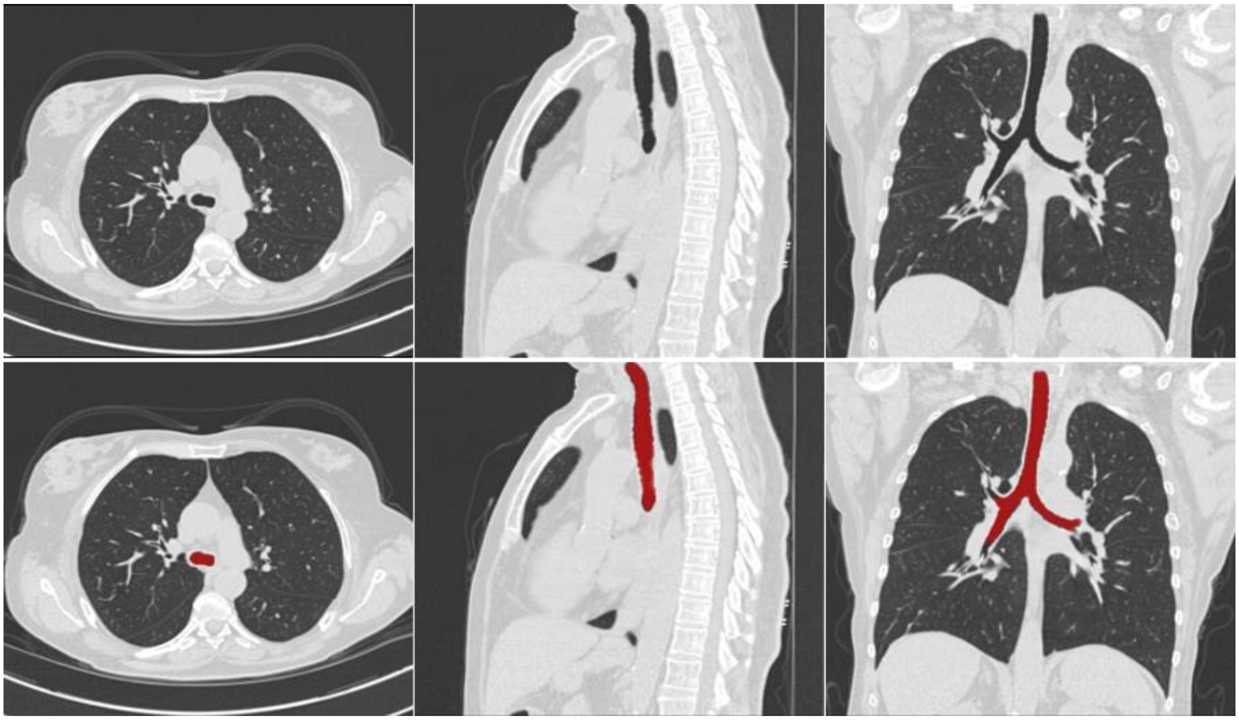


Figure 2.50: Example trachea segmentation by CNN.

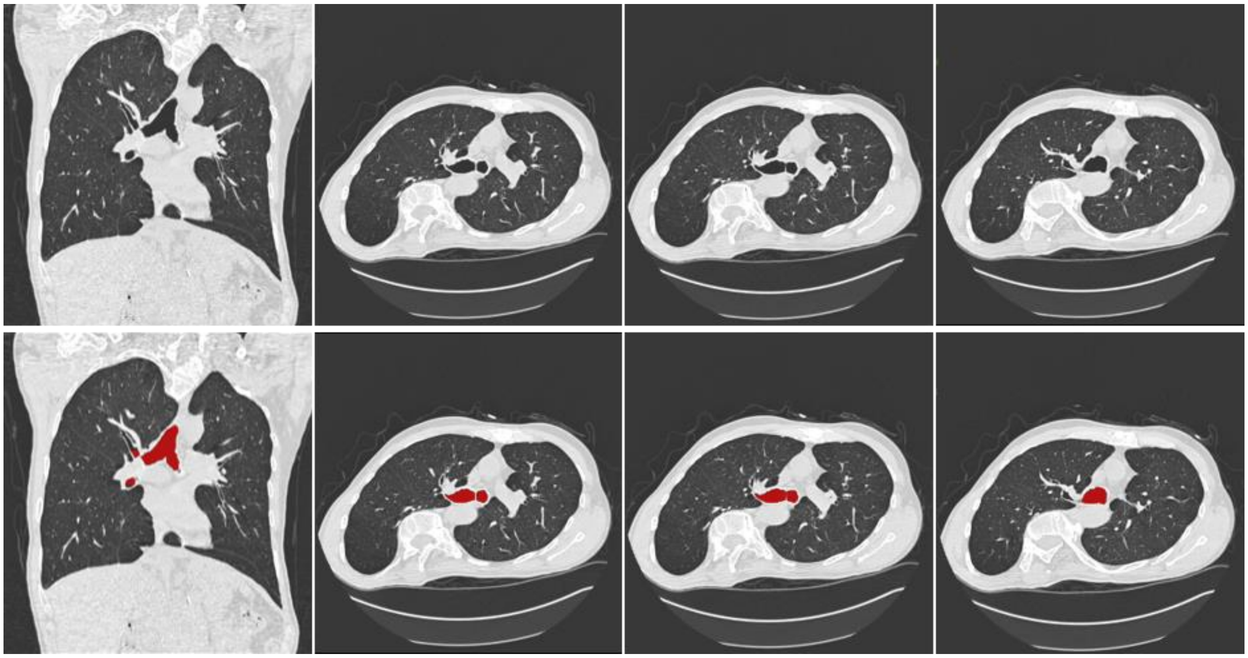


Figure 2.51: Trachea segmentation on a scan with tilted chest by the model.

After trachea segmentation, the second step was to localize the carina. Figure 2.52 demonstrates the search process for the carina slice based on the 3D trachea segmentation on a scan with a slice spacing of 1.5 mm.

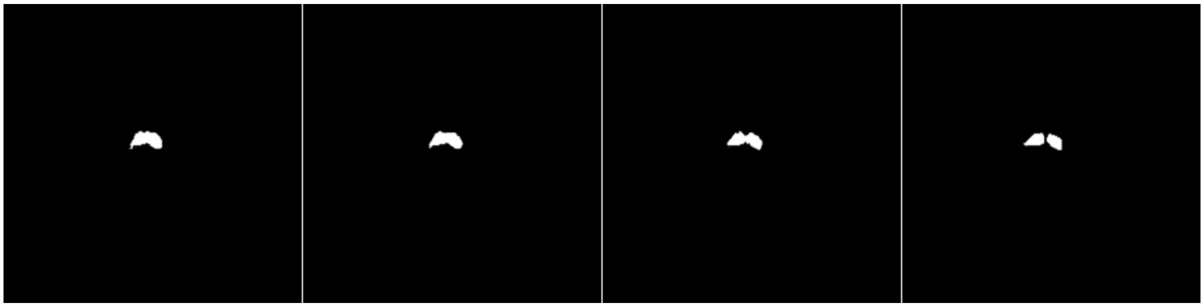
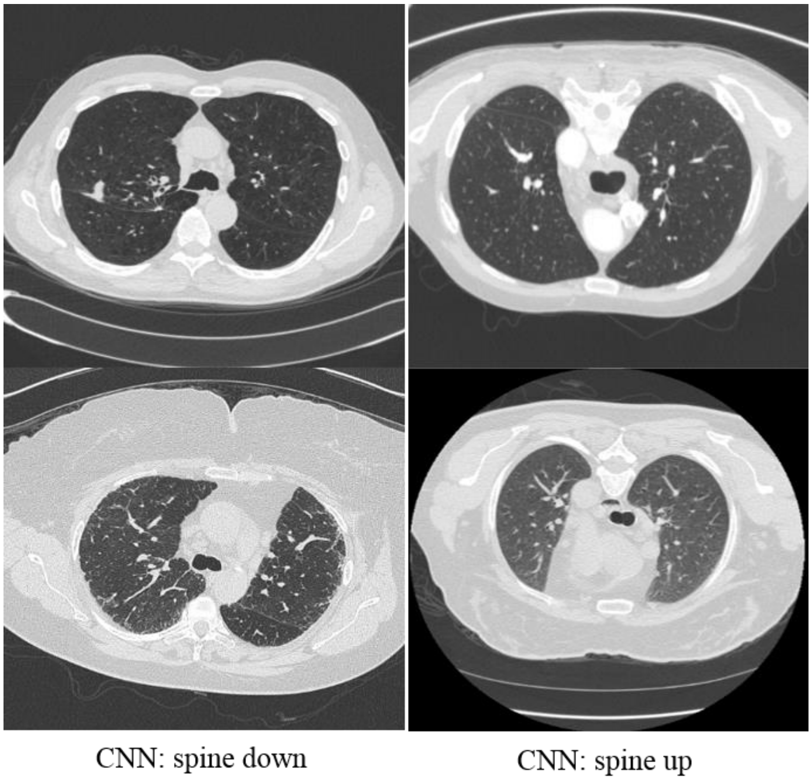


Figure 2.52: Identification of carina slice by searching bifurcation.

For the spine localization model, both the training and validation accuracy were over 99.99% after 3 epochs, and the system achieved 100% accuracy on test set under 5-fold cross validation.



CNN: spine down

CNN: spine up

Figure 2.53: Example of spine up and down scans detected by the CNN.

Some examples of spine up and spine down images detected by the CNN are shown in Figure 2.53. First column has 2 spine down cases with the top one being supine and bottom one prone. The second column has 2 spine up cases, with top one being supine and the bottom one prone.

In practice, multiple carina slices were used in breath-hold classification to make the inference more robust, and the model achieved an average accuracy of 99.3% on the test set under 5-fold cross validation. An example mask image of a carina classified as RV and one classified as TLC are shown in Figure 2.54. The first row are 3 RV cases with variations in the curvature of the posterior wall, and the second row are 3 TLC cases with differing degrees of circularity of carina shape.

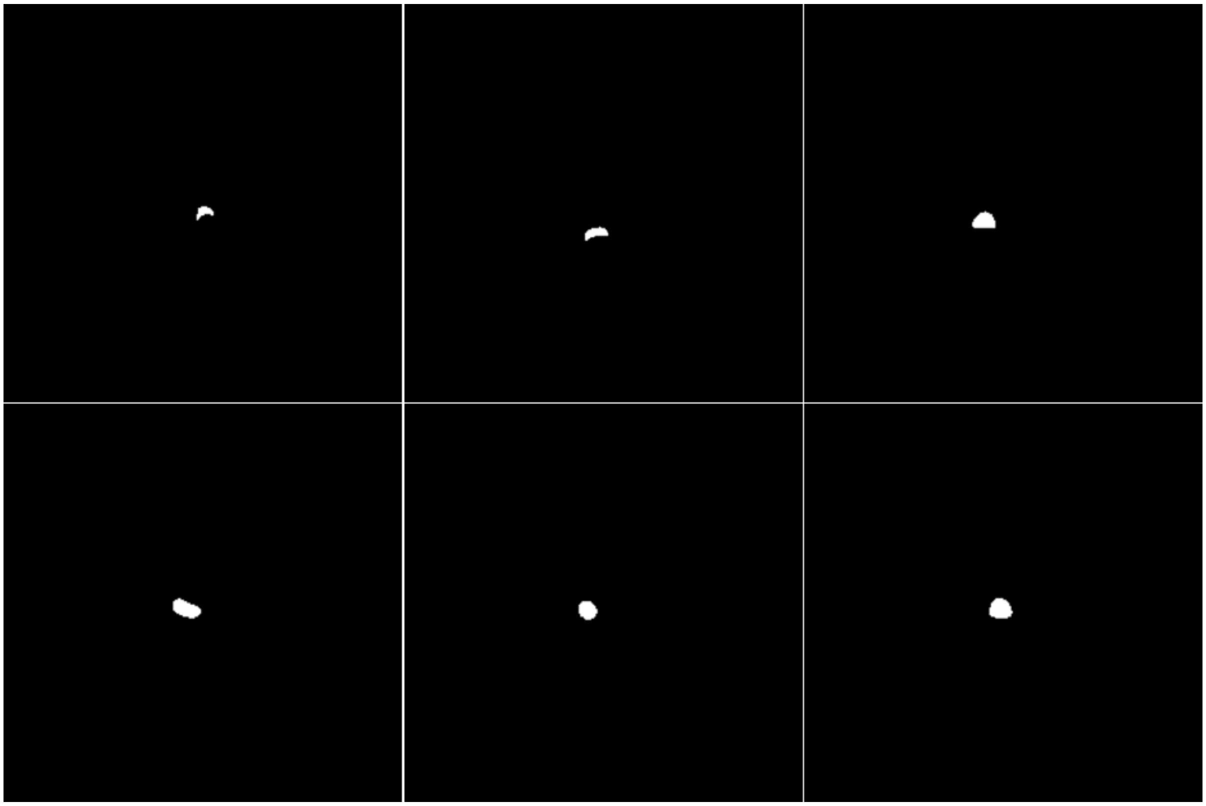
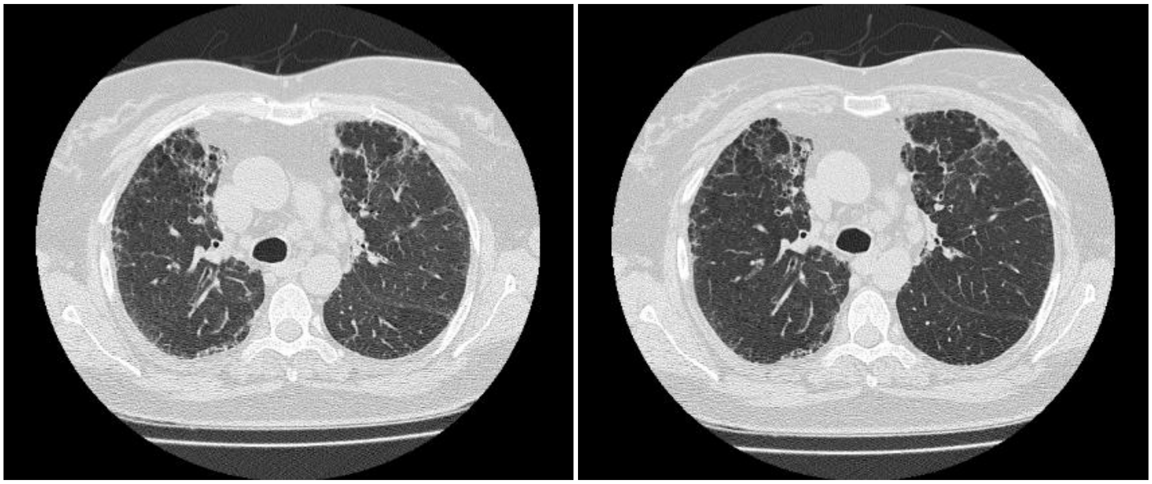


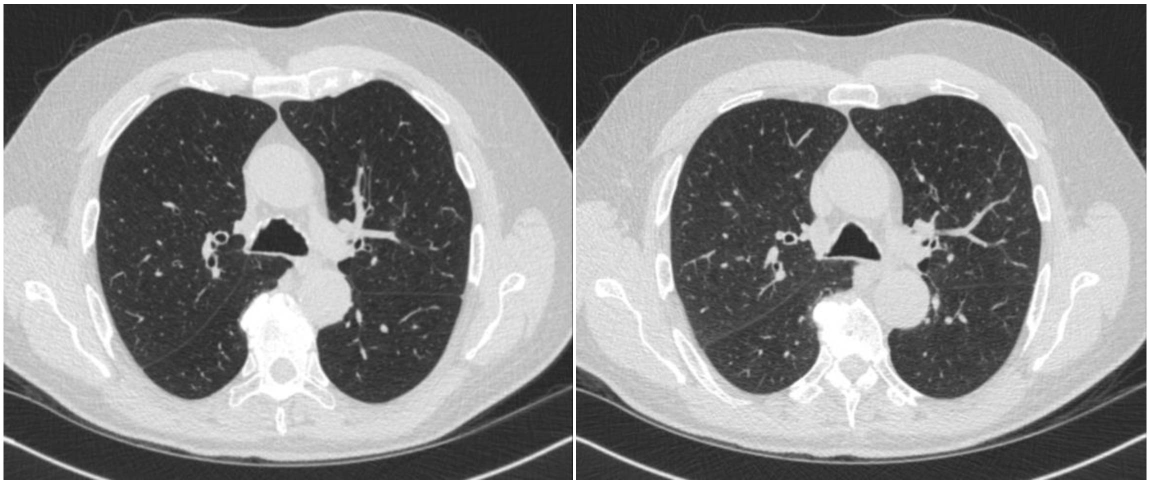
Figure 2.54: Mask images classified as RV (first row) and TLC (second row) by the model.

2.6.5 Discussion

In clinical practice, we noticed that the shape of carina alone sometimes failed to identify the breath-hold. Two unsuccessful cases are shown in Figure 2.55.



Truth: RV vs. CNN: TLC　　Truth & CNN: TLC



Truth: TLC vs. CNN: RV　　Truth & CNN: RV

Figure 2.55: Difficult RV and TLC cases without the usual characteristic carina shapes.

The first row is a pair of RV and TLC scans from the same patient. The shape of the carina on both the RV and TLC scans appear circular and the CNN identified both as TLC. We can probably determine which one is more likely to be RV or TLC by comparing the lung volumes or attenuation within the lungs if both scans were present side by side. However, there is no guarantee that both breath-holds of a given patient will be available and our goal is to identify the breath-hold on a single image rather than an image pair. The second row is an opposite example where the CNN mis-classified the TLC scan as RV. The posterior membrane of the carina is flat (like a letter ‘D’) on both the RV and TLC scans. Unless both breath-holds of the same patient are provided, it is very challenging to make a correct or appropriate decision on a single image. On the other hand, it is possible that the patients did not take full breaths in/out during scanning as reflected by their carina shapes. This could be caused by a patient’s inability to completely exhale or inhale or because they did not effectively follow the instructions given by the CT technologist. We tried to augment the input to the network by adding one axial slice and coronal slice at the center of carina, hoping that the CNN could extract additional features, such as lung volume, density, position of diaphragm, etc. as supplements. However, the network still struggled when confronted with these challenging cases.

Our goal in this project was to identify the breath-hold, which is slightly different from most studies aimed at accurate airway segmentation. We still endeavored to accurately segment the trachea since the classification outcome relies on the shape of carina. Because of this limited application, we only segmented the main bronchi and left and right bronchus beyond bifurcation instead of the whole airway tree. Segmentation of the entire airway tree on high resolution CT remains very challenging.

The segmentation model was also applied to the same low-dose and ultra-low-dose dataset used in robustness assessment of previous classification models. Figure

2.56 is an example of the trachea segmentation on an ultra-low-dose scan from lung cancer screening with a dose level around 0.2 mGy (1/75 of regular diagnostic).

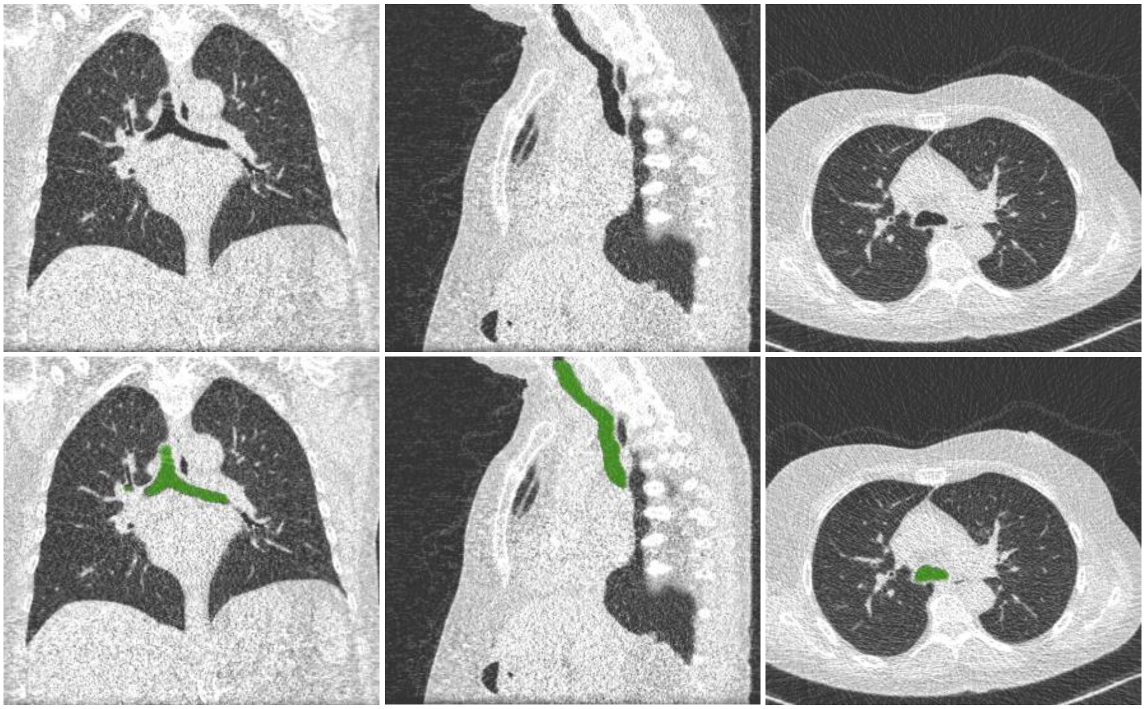


Figure 2.56: Trachea segmentation on an ultra-low-dose scan by the model.

Figure 2.57: Additional use of trachea segmentation.

The shape of the trachea could be a biomarker of airway related diseases. The left of Figure 2.57 is an example with posterior and lateral irregularities of the carina that could imply stenosis with a lesion protruding into the tracheal lumen[260–262]. The image on the right of Figure 2.57 is an example showing a decreased tracheal index (TI), i.e. the ratio between transverse diameter and anteroposterior diameter, a characteristic of COPD patients[263,264].

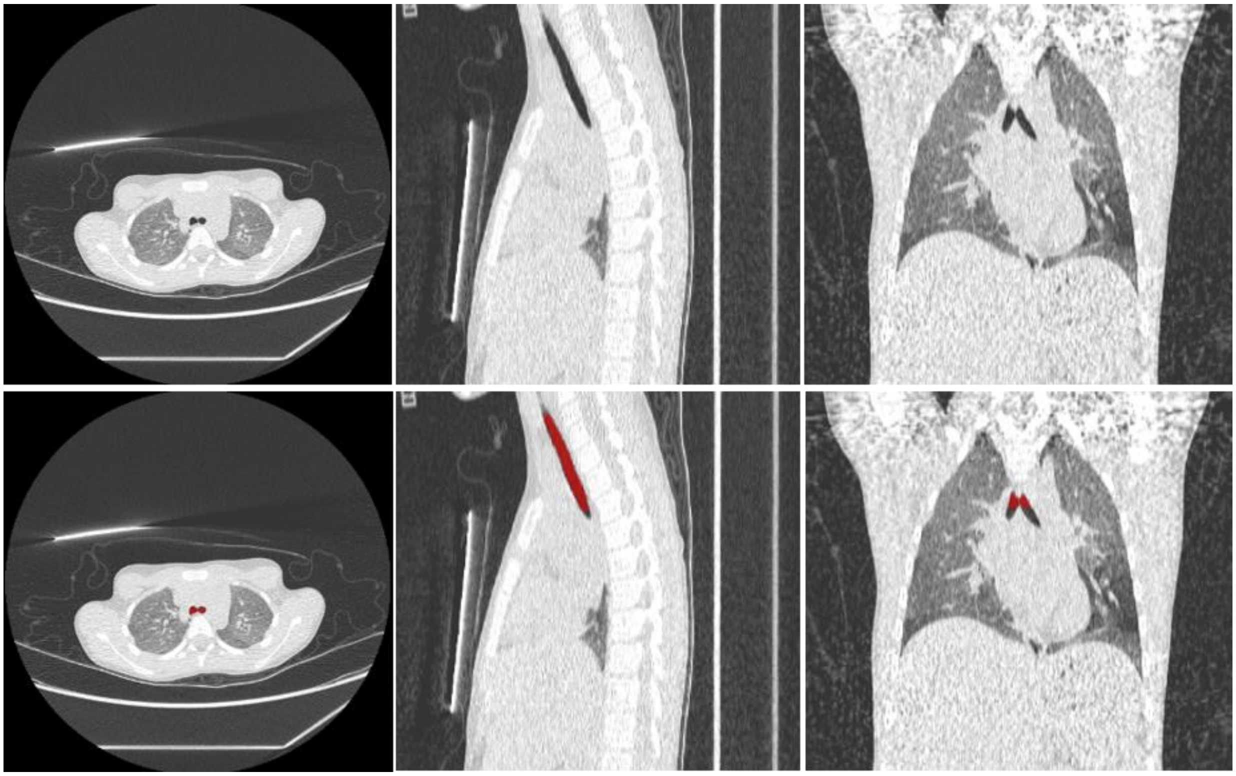


Figure 2.58: Trachea segmentation on a pediatric scan.

The model was applied to pediatric patients as well and an example is shown in Figure 2.58. Although the trachea of a child is much smaller and the model was trained using only adult scans, it was still able to segment it. The diversity in the training set and the image rescaling applied in data augmentation probably contributed to the model reliability in both adult and pediatric patients.

Saliency maps at the last convolutional layer of the trachea segmentation model is shown in Figure 2.59. The boundaries of the trachea were most activated by the network and all other pixels are suppressed.

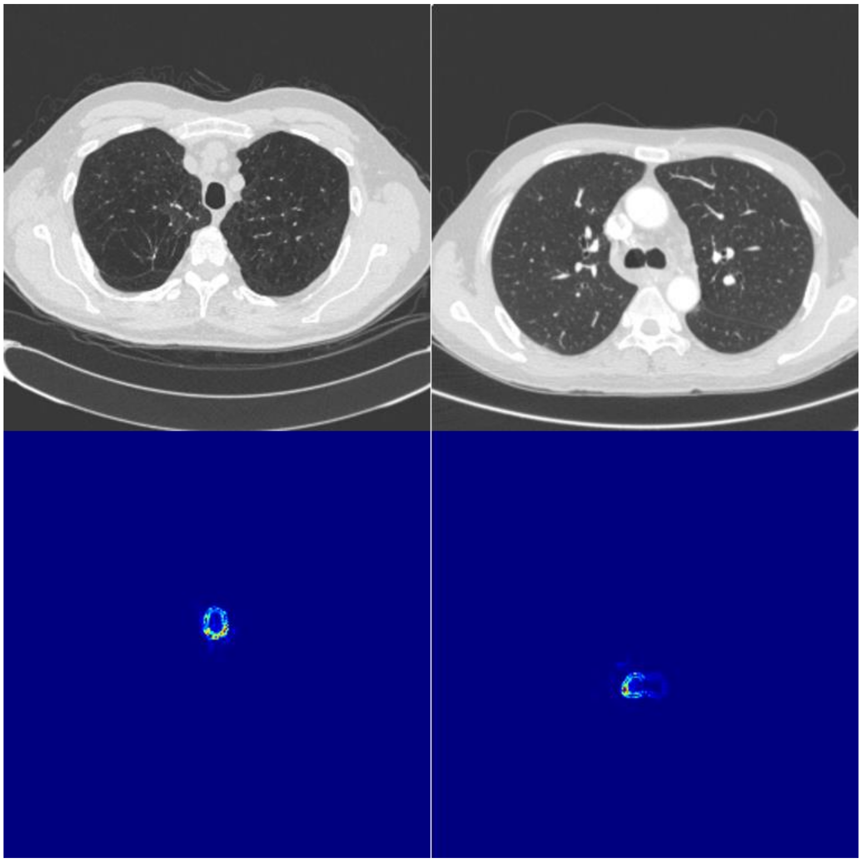


Figure 2.59: Saliency maps from the trachea segmentation model.

Saliency maps from the last classification layer of the spine detection model are shown in Figure 2.60. As anticipated, features exploited by the network for spine up/down classification are primarily from the spine and pixels belonging to the table are not activated. Image pairs with an absence of table in the second column are more intuitive regarding this observation. This is in contrast to the activation maps from scan posture (prone or supine) classification model where pixels from both spine and table were activated.

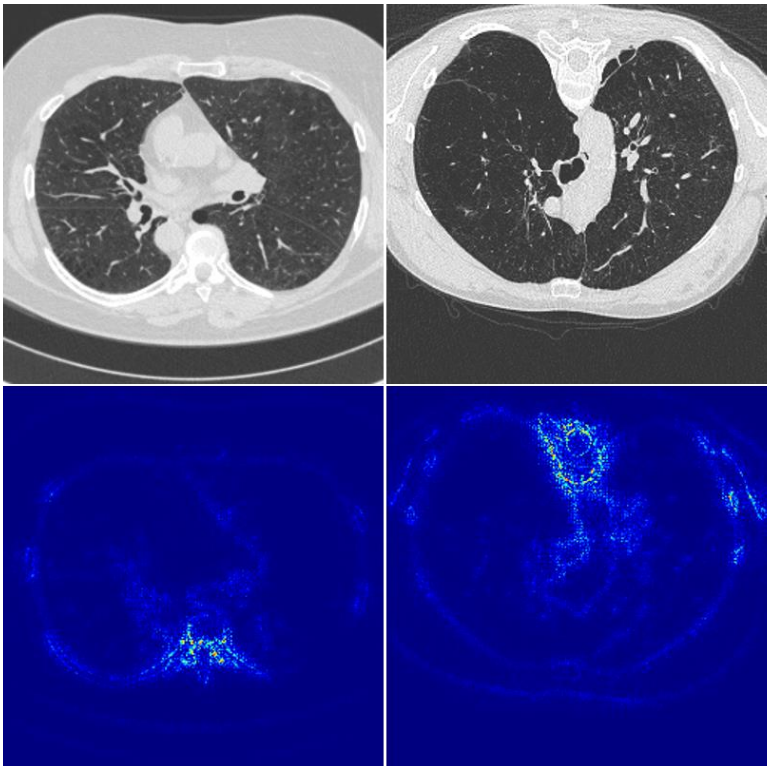


Figure 2.60: Saliency maps from the spine detection model.

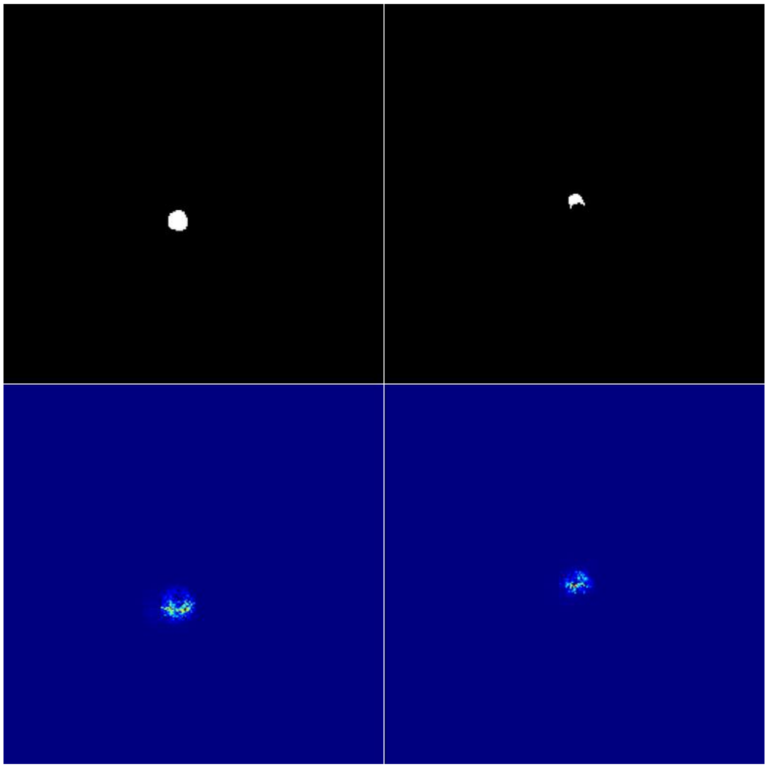


Figure 2.61: Saliency maps of RV and TLC classification from the breath-hold detection model.

The last saliency maps from the breath-hold detection model are shown in Figure 2.61, including cases of RV and TLC respectively. Pixels near the posterior wall of the carina are highly activated by the classification model, which is what we expect the network to focus on.

2.6.6 Conclusion

We proposed a method to accomplish breath-hold classification by three steps: 1. trachea segmentation; 2. carina localization; 3. breath-hold classification between RV and TLC based on the shape of the carina. Our deep learning models demonstrated high performance and robustness on trachea segmentation and subsequent breath-hold classification.

Chapter 3: Lung and lobar segmentation on chest CT

3.1 Lung segmentation

3.1.1 Introduction

Lung segmentation is often performed in chest Computed Tomography (CT) as a precursor to quantitative image analysis of the lung parenchyma[265], nodule detection[266], emphysema and fibrosis scoring[267–270], and particularly lobar functional analysis[23]. Extensive research has been devoted to automated lung segmentation on chest CT[271–274], with many methods demonstrating accurate segmentation in normal lungs[12,14,275,276]. However, traditional segmentation methods, including threshold-based and shape-model-based methods, often fail to accurately segment lungs with dense pathologies[15,17,277]. More specifically, in lungs with severe Idiopathic Pulmonary Fibrosis (IPF)[278,279], peripheral regions with increased CT attenuation tend to be excluded due to inadequate contrast between the lung and surrounding tissues. Given that the majority of chest CT scans acquired in a clinical setting will contain some abnormalities, improving segmentation accuracy of lungs with pathology will be crucial in increasing throughput and translating quantitative lung imaging to routine clinical practice.

A number of studies have investigated segmentation of pathologic lungs. A segmentation-by-registration scheme was introduced in which a normal lung was elastically registered to a pathologic lung followed by k-nearest neighbor voxel classification[15]. Voxel classification using local texture features has also been applied to segment abnormal lung regions[17–19]. Applying anatomical model constraints was another strategy to address the challenge of pathological lung segmentation[280,281]. For example, curvature of the ribs[16] was used to assist the selection of optimal thresholds to segment the lung. A graph search algorithm[282,283] was proposed to detect the lung surface by solving a maximum flow problem. Similarly, a graph cut based method modeled the foreground and background objects in the image as a Gaussian mixture model and optimized the probability of foreground using expectation maximization[284]. A modified convex hull algorithm[285] was introduced to extract the coarse lung region containing diffuse lung disease followed by morphological analysis. Most of these were evaluated on relatively small datasets (<50 scans). Since these unsupervised learning methods usually involve empirical determination of algorithm parameters, it is very challenging to make them generalizable and applicable to a wide spectrum of abnormal lungs. Additionally, segmentation efficiency is required for high throughput processing of chest CT images.

Deep learning methods have been successfully applied to various image classification and recognition tasks in medical imaging. In the past few years, the methods have shown superiority in the image segmentation field. A fully convolutional neural network[101] was proposed for segmentation including precise pixel-wise prediction. It outperformed many traditional state-of-the-art segmentation methods. With a similar idea, the hypercolumn at each pixel was defined as the vector of activations of all CNN units for object segmentation and localization[286]. The U-Net[102] architecture demonstrated success in many biomedical image segmentation problems, including cell segmentation in microscopic images, vertebral level localization for lumbar surgery and glandular structure segmentation on colon histopathology images. SegNet[287,288] which is a deep Encoder-Decoder neural network showed promising results for lung segmentation in chest radiographs. Volumetric image based methods were then introduced, including 3D U-Net[111] and V-Net[112] to segment the prostate and kidneys in MR images. For deep learning application to lung segmentation on chest CT, a progressive and multi-path holistically nested network (P-HNN)[289] was proposed and claimed to merge more outputs from different network stages to generate masks with finer details. SegCaps[290] was proposed more recently and is an extension of the original capsule network[291] by adding de-convolutional capsules for segmentation. Generative models were also used in lung segmentation by introducing adversarial loss. Based on a fully convolutional network, Zhao et al. proposed to add multi-instance and conditional adversary loss to tackle moderate and severe pathological conditions[292]. An improved Wasserstein GAN[293] was applied to enhance learning stability by employing Earth Mover (EM)[294] as the loss function[295].

Residual Network (ResNet)[113] has been the state-of-the-art image recognition architecture and won first place in the ILSVRC-2015 classification task. Due to its much deeper network with residual learning unit, it demonstrated exceptional ability in feature extraction, surpassing previous models such as VGG[158] and GoogleNet[296]. In our work, a ResNet-101 based 2D fully convolutional neural network was proposed for pathological lung segmentation on chest.

3.1.2 Methods

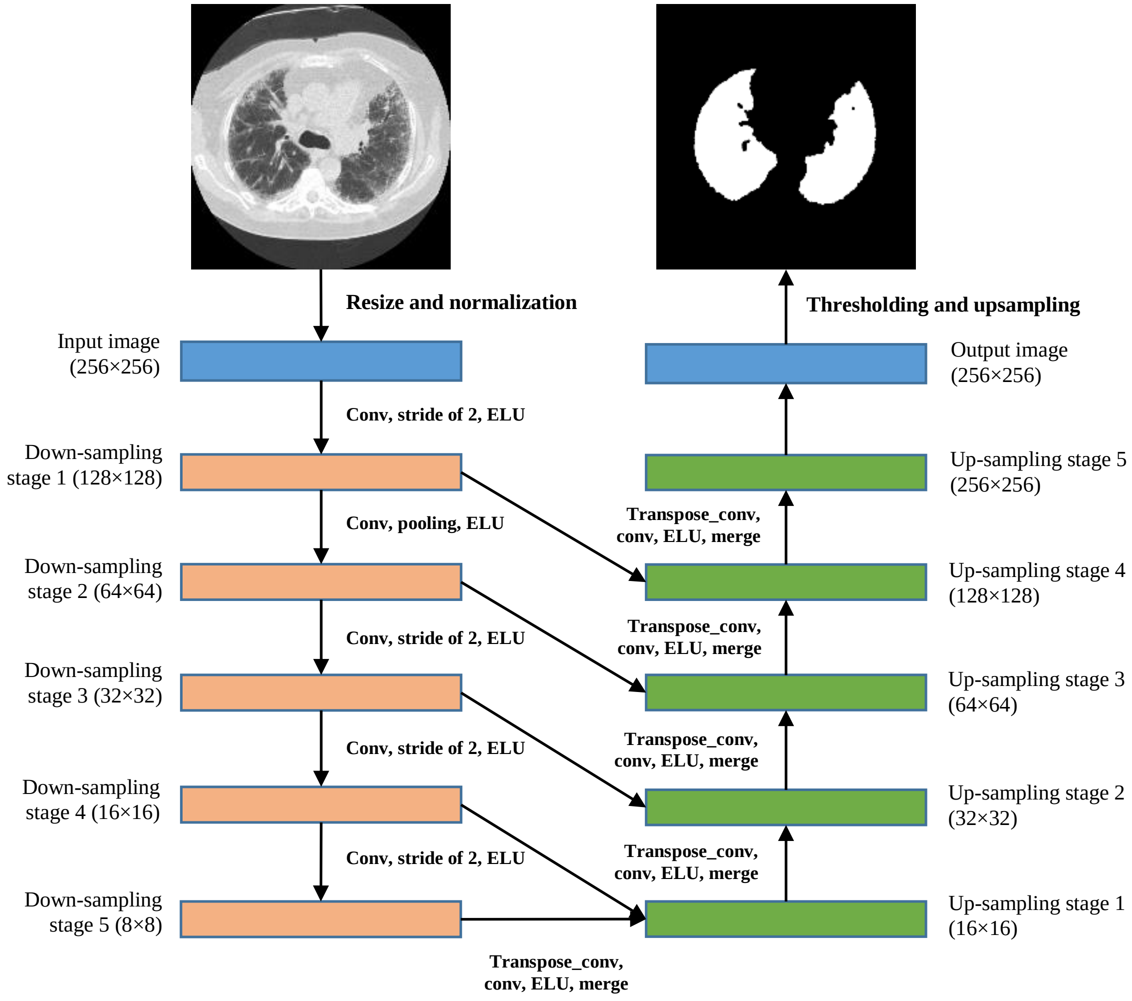


Figure 3.1: Resnet-101 based lung segmentation architecture.

As shown in Figure 3.1, the convolution section or feature extraction is the same as the original ResNet-101 but without fully connected layers and a final classification layer. All of the convolutional filter sizes were 3 × 3 with padding of 1 × 1 and stride of 1 x 1 to keep the image size intact. Downsizing was performed by convolution with a size of 2 × 2 at each stage. Overall, the original input image was progressively scaled down 5 times (from 256 × 256 to 8 × 8). In the deconvolution section, a corresponding 5 times up-sampling (size of 2 × 2) with transposed convolution was used to recover the resolution of the input image (from 8 × 8 back to 256 × 256). Following the same strategy as suggested by Szegedy et al.[59], the number of feature maps in the de-convolutional layer was cut by half every time the image size doubled. In addition, skip-connections were used in the up-sampling process to incorporate finer details from the lower layers as well as abstract and sematic information from higher layers. After achieving the target resolution, a final 3 by 3 instead of 1 by 1 convolution was applied on the aggregated hyper-columns to achieve pixel-wise classification.

One concern of using a 2D slice based CNN is the lesser spatial context available compared to a 3D CNN. To investigate the necessity of spatial information, a 3D CNN was also built for performance comparison. Due to computer hardware limitations, it is not feasible to build a comparable Resnet-101 based 3D CNN. Instead, a 3D U-Net architecture was used with an input image size of 128 × 128 x 128. However, it only involved 3 times down-sampling (from 128 to 16) in the convolution section and 3 times up-sampling (from 16 to 128) in the deconvolution section. Moreover, the number of feature maps used in each layer was also much smaller compared to the 2D model aforementioned. It is evident that the ResNet-101 based 2D model is much more complex in terms of the depth and width of the network.

Data augmentation was applied in training, including rotation of [45°, 45°], horizontal and vertical shift of [0.2, 0.2], horizontal and vertical flipping, zoom in/out of [0.8, 1.2]. Dice loss (1-Dice coefficient) was used as the cost function since it has been shown to be robust in segmentation tasks[297]. It also explicitly reflects segmentation accuracy at a specific iteration instead of just a declining trend of training/validation loss.

The model was trained from scratch on a NVIDIA TITAN X with GPU memory of 12 GB, using the Adam optimizer, and learning rate of 0.01 with a decay rate of 1e-4.

To compare the CNN segmentation against the ground truth the Dice Similarity Coefficient (DSC) and Average Surface Distance (ASD) were used. The Dice Coefficient is defined as:

$$D = \frac{2P \cap T}{P + T}$$

Where P is the segmentation predicted by the model and T is the ground truth.

Average Surface Distance is defined as:

$$ASD = \frac{1}{n_s + n_{s'}} \left(\sum_{s=1}^{n_s} d(\mathrm{S}, \mathrm{S'}) + \sum_{s'=1}^{n_{s'}} d(\mathrm{S'}, \mathrm{S}) \right)$$

where d is the average of all point-to-point distances from the two surfaces. S is the surface of the automatic segmentation and S' is reference.

A paired t-test was used to test the significance of the Dice score difference between the CNN based and a previous threshold-based anatomical model method.

3.1.3 Materials

Chest CT images used in this study were collected retrospectively from 6 multi-center clinical trials. Image acquisition at different sites provided variability in CT acquisition parameters, such as slice thickness, reconstruction kernel, etc. The slice thickness range was [0.625 mm, 3 mm], in-plane (x-y) spacing range was [0.467 mm, 1mm], tube current range was [80 mA, 644 mA], and reconstruction kernels ranged from smooth to sharp. In total, 575 chest CT scans from different subjects were used, each with a radiologist-approved lung segmentation. 143 scans were from subjects enrolled in Chronic Obstructive Pulmonary Disease (COPD) clinical trials and 432 were from interstitial pulmonary fibrosis (IPF) trials. These cases contained common lung parenchymal abnormalities including emphysema, ground glass, fibrosis, nodules, and honeycombing. Sample axial images from these disease patterns are shown in Figure 3.2.

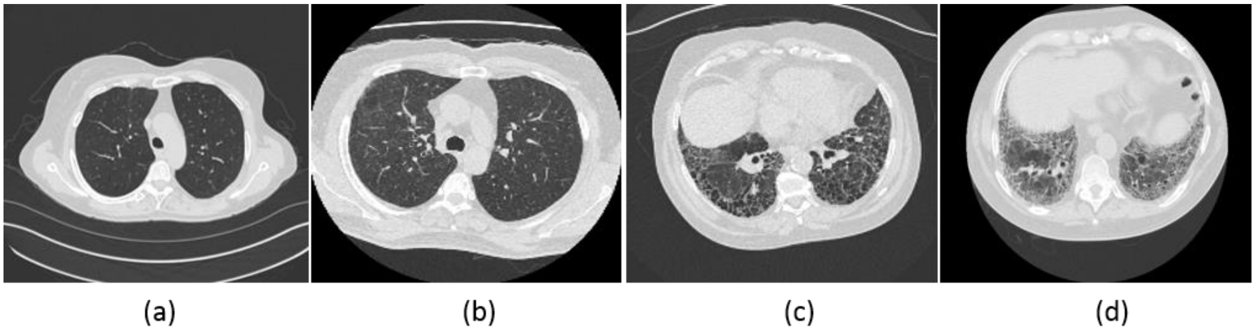


Figure 3.2: Sample axial slice of chest CT of cases with different lung diseases in our dataset, including (a) emphysema, (b) ground glass, (c) honeycomb, (d) fibrosis

The ground truth lung segmentations of these scans were derived using an independent semi-automated segmentation. Specifically, the scans were segmented using a threshold-based anatomical model technique[298,299], followed by manual editing by an image analyst. Finally, radiologists reviewed and edited as needed, and approved the final segmentation. Two thoracic radiologists were involved, and both of them have more than 20 years' experience.

Balanced sampling, rather than random sampling, was used to form training and test sets that were balanced in terms of the number and type of abnormalities present. The whole image set was sorted based on the mean Hounsfield Unit (HU) within the lung in ascending order. Scans with lower and higher mean lung intensity likely correspond to cases with emphysema and IPF, respectively. Every 5 scans in the ordered image set were split into training and test sets with a ratio of 4:1. 5-fold cross validation with balanced distribution was used and each scan was used for testing exactly once. At each fold, the training set consisted of 460 scans, and the test set of 115 scans. Using this approach, similar to stratified cross validation[300], the composition distributions of training and test sets were similar to the original dataset. For a specific scan, only slices (axial images) with lung segmentation (i.e., containing lung) were used.

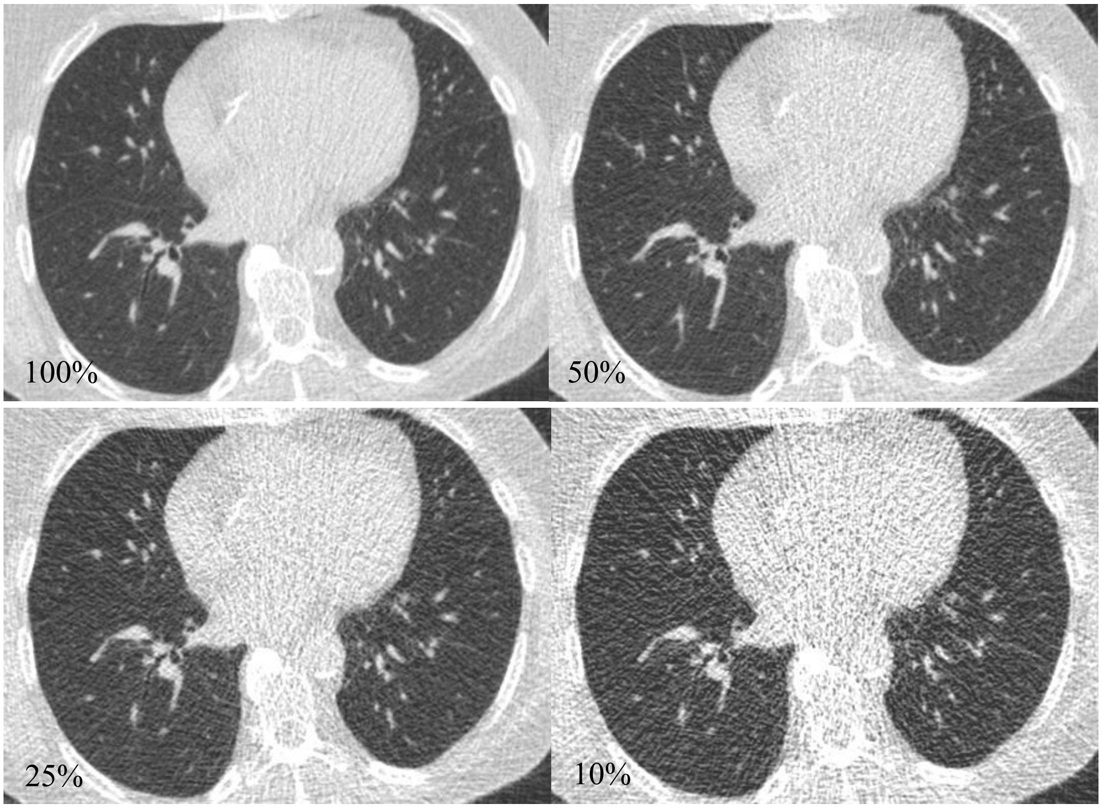


Figure 3.3: Example ultra-low-dose scan from lung cancer screening.

In the evaluation stage, we also applied the model to low and ultra-low-dose scans from lung cancer screening. In total, 40 scans from different subjects with corresponding raw CT projection data were used and they were divided into two groups.

Cohort 1 included 30 scans from lung cancer screening and the original dose level at 100% was around 2 mGy. Cohort 2 comprised 10 scans from diffuse disease trials and the original dose level at 100% was around 15 mGy. A CT reconstruction pipeline in our group was applied to simulate images with 1 mm slice thickness and spacing, medium kernel and dose levels of 10%, 25%, 50% and 100%, and weighted filtered back projection (wFBP)[301] was used as the reconstruction algorithm. Some example images from the two cohorts at different dose levels are shown in Figure 3.3 and Figure 3.4.

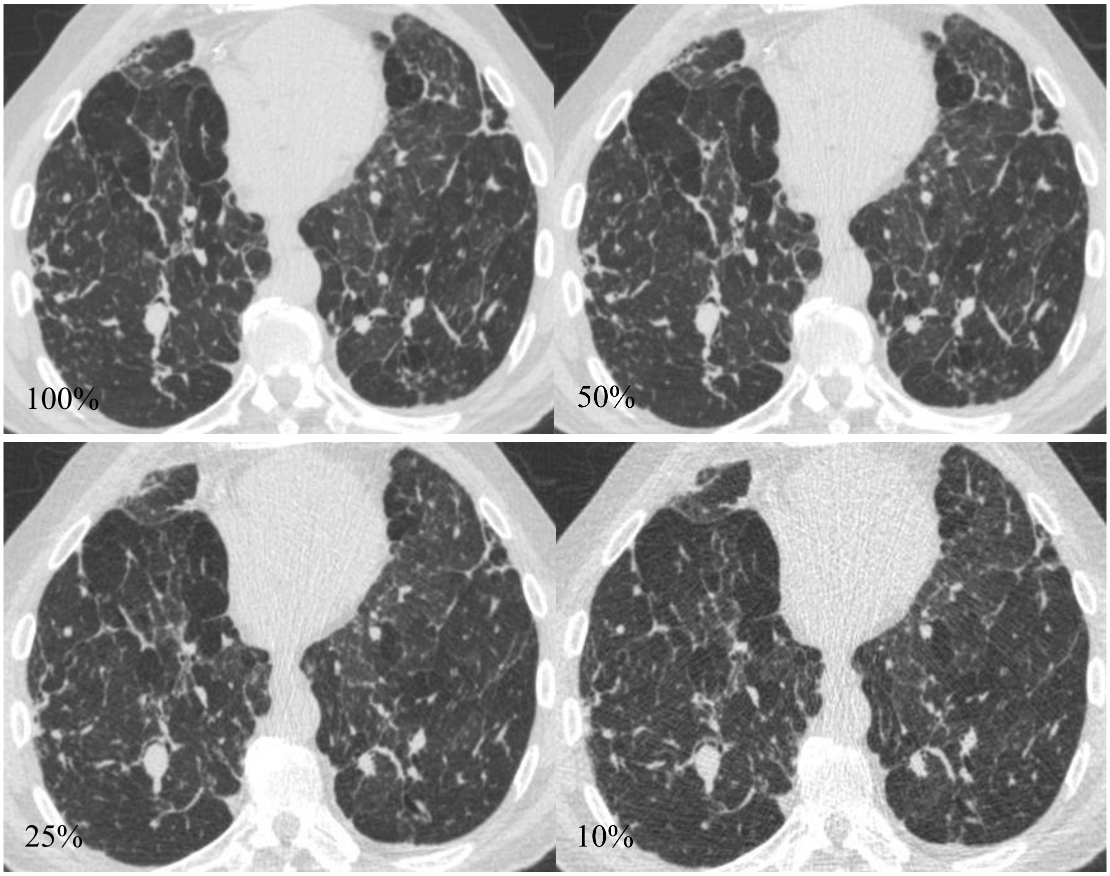


Figure 3.4: Example low-dose scan from diffuse lung disease trials.

3.1.4 Results

5-fold cross validation was performed with 460 training and 115 test scans in each fold. In testing, the whole scan was fed into the segmentation network to generate a volumetric lung segmentation (of all slices) for both the 2D and 3D CNN models. We trained the ResNet-101 based model for 16 epochs and the 3D CNN for 100 epochs, beyond which there was no significant improvement indicated by the loss function. Using the hardware described above, the entire 5-fold cross validation took about 1 week to complete. For a single test scan with 300 slices, it took about 5 seconds to segment the whole lung with the trained model.

Table 3.1 illustrates the segmentation results by the proposed 2D CNN model, 3D CNN model and also the previous threshold based anatomical model method developed by our group. Only 3 out of the total 575 scans from 2D CNN segmentation and 9 scans from 3D CNN segmentation had a Dice coefficient below 0.95. On the contrary, there are 111 cases with Dice score below 0.95 for our previous method. A paired t-test yielded a significant difference in Dice Coefficient with $p < 0.001$ when comparing the previous threshold based method against either the 2D or 3D CNN based methods. The 2D CNN outperformed the 3D CNN model, but not significantly.

Table 3.1 Segmentation results comparison between the CNN based and previous threshold and anatomical model based methods, including mean and standard deviation of Dice coefficient and Average Surface Distance.

	Dice coefficient	ASD (mm)
2D CNN based	0.988 ± 0.012	0.562 ± 0.49
3D CNN based	0.980 ± 0.017	0.581 ± 0.52
Threshold and anatomical model	0.965 ± 0.023	0.599 ± 0.47

Figure 3.5 shows three example segmentations with different amounts of fibrosis present (from mild to severe) by the proposed 2D based CNN, 3D based CNN and previous threshold based anatomical model methods.

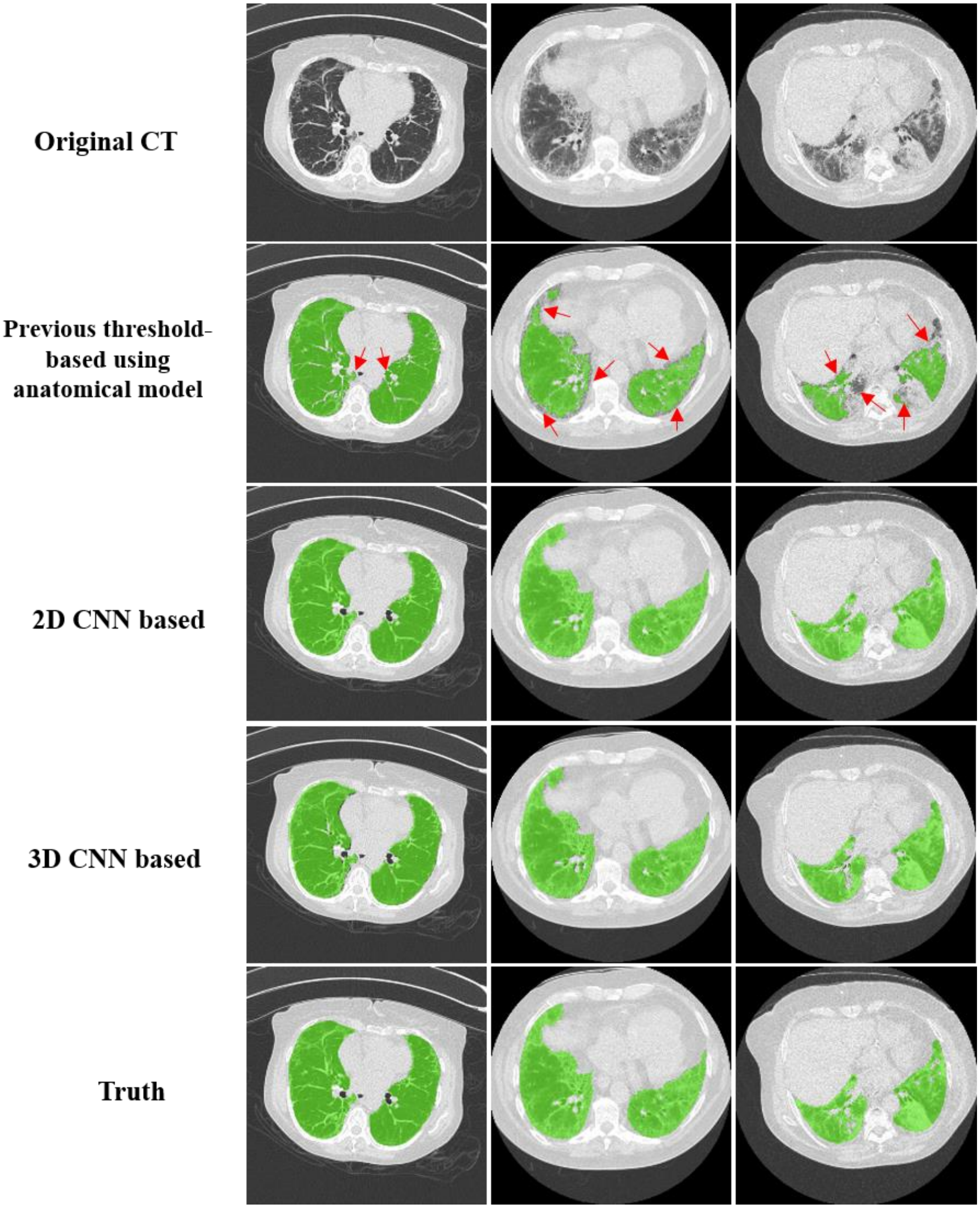


Figure 3.5: Comparison of different segmentation methods on fibrotic lungs.

Second row corresponds to Dice scores of (0.954, 0.931, 0.847) and ASD of (0.575mm, 1.689mm, 4.51mm) by the threshold and anatomical model based method. The 2D CNN achieved Dice scores of (0.989, 0.980, 0.976) and ASD of (0.36 mm, 0.505 mm, 1.18mm). The 3D CNN achieved Dice scores of (0.979, 0.969, 0.970) and ASD of (0.484 mm, 0.677 mm, 1.653 mm). The CNN based methods achieved much better segmentation when the attenuation of the lung was changed significantly due to disease. More specifically, the boundary areas near the chest wall and other soft tissues were under-segmented by the threshold based anatomical method, while CNN based method still identified them as part of the lungs.

Figure 3.6 shows example segmentations of emphysema, ground glass and honeycombing (from 1st to 3rd column) by the proposed the CNN methods as well as our previous threshold-based anatomical model method for comparison. In these examples, there are no major attenuation changes in the lung and all three methods were able to achieve good segmentation relative to the ground truth. The Dice scores of these three cases (from left to right) are (0.992, 0.973, 0.974) by the threshold based anatomical model method, (0.991, 0.984, 0.976) by 2D CNN, and (0.989, 0.981, 0.972) by 3D CNN. Comparable ASDs were achieved by the three segmentation methods: (0.355mm, 0.55mm, 0.503mm) vs. (0.352mm, 0.327mm, 0.479) vs. (0.356 mm, 0.374 mm, 0.488 mm). One further difference is the successful exclusion of central airways by the CNN in this honeycomb case.

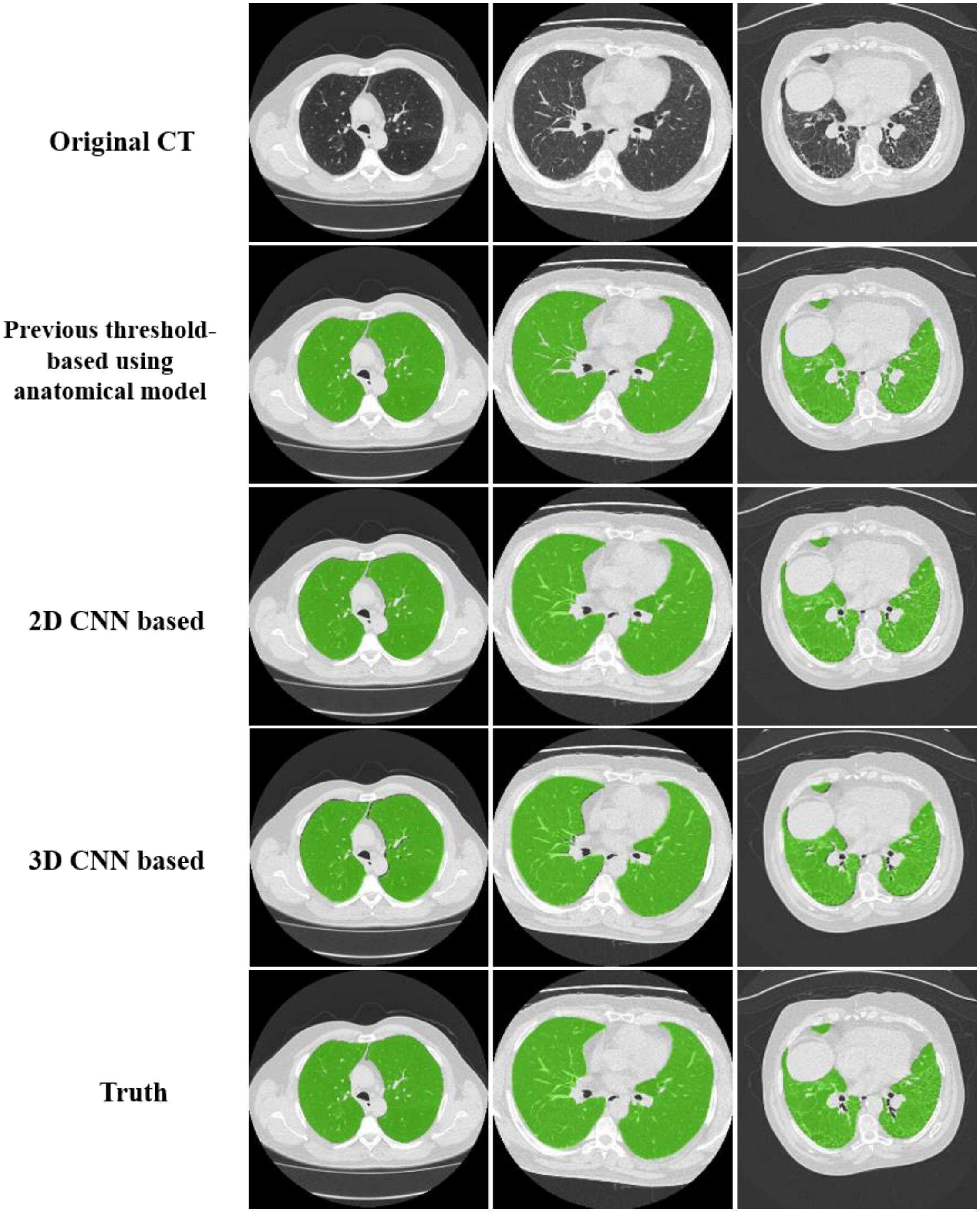


Figure 3.6: Comparison of different segmentation methods on cases with emphysema, ground glass and honeycombing.

Most pulmonary nodule detection algorithms require accurate lung segmentation to constrain the search space. Since our lung segmentation model was built using mostly diffuse lung disease data, blood vessels and large nodules are excluded. To compensate CNN segmentation for nodule detection, 3D binary closing can be applied to incorporate possible missed nodules. Figure 3.7 is example of applying binary closing with an ellipsoid of size (3, 3, 3) after lung segmentation by the CNN.

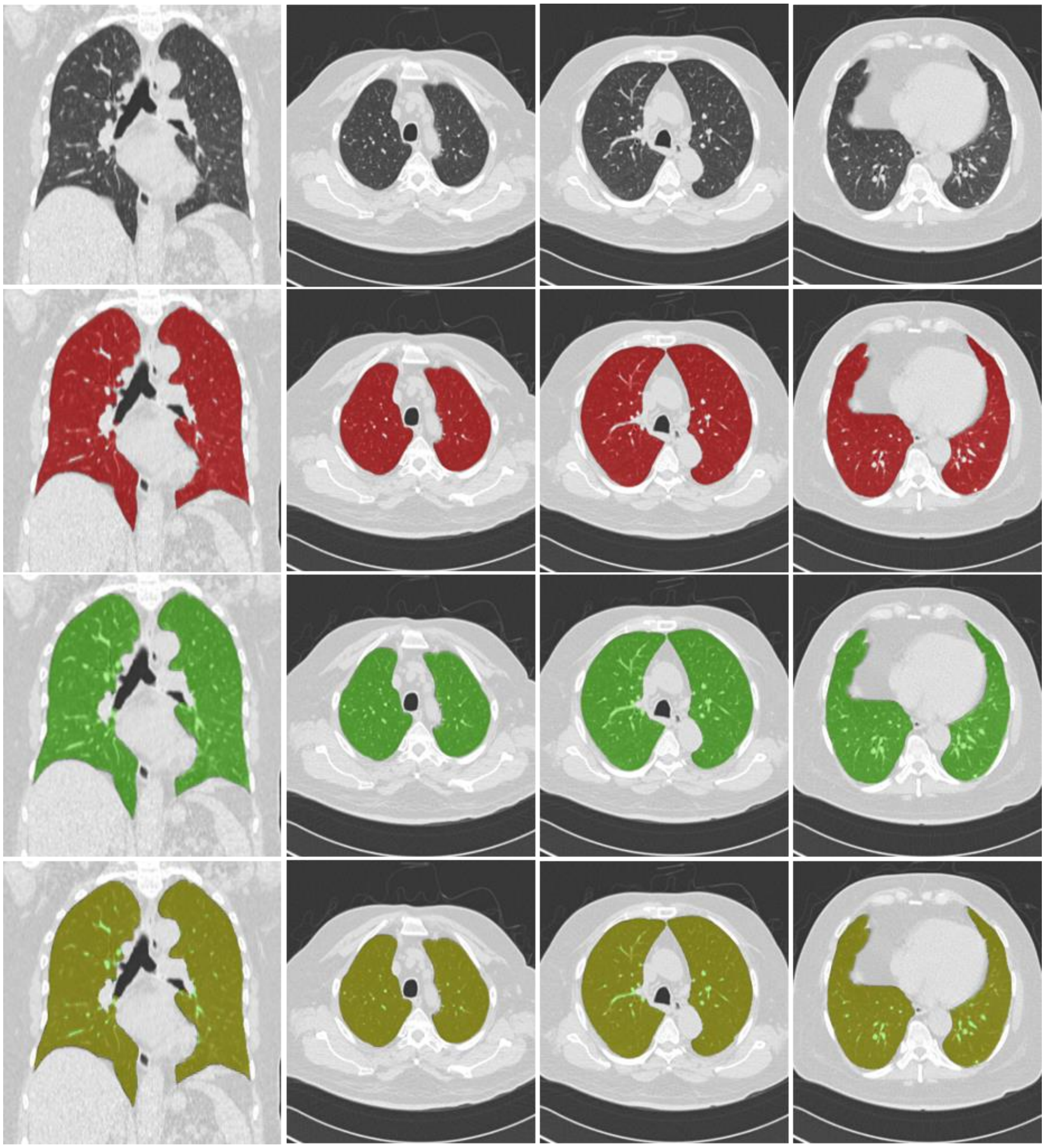


Figure 3.7: Comparison of CNN lung segmentations before and after applying 3D binary closing.

The second row (overlay shown in red) shows the excluded vessels and some nodules by the CNN and the third row (overlay shown in green) is the result after 3D hole filling. The last row shows an overlay of two masks to compare the difference indicated by light green.

The CNN lung segmentation model was also assessed on low and ultra-low-dose scans. Table 3.2 shows the average Dice score and ASD compared to the reference segmentation. The reference segmentation was obtained using a threshold based anatomical model method followed by manual editing. Table 3.3 gives a comparison of the CNN segmentation on the 100% dose scan as reference. In fact, the latter scenario is more straightforward to investigate the impact of dose variations since they were segmented using the same method.

Table 3.2 CNN segmentation results on cases simulated at 4 dose levels from lung cancer screening.

	DSC	ASD (mm)
100%	0.987 ± 0.009	0.826 ± 0.671
50%	0.985 ± 0.013	0.917 ± 0.884
25%	0.982 ± 0.012	1.059 ± 0.947
10%	0.978 ± 0.017	1.218 ± 1.145

Table 3.3 CNN segmentation results on cases simulated at 4 dose levels from lung cancer screening using CNN segmentation on 100% dose scan as reference.

	DSC	ASD (mm)
50%	0.994 ± 0.0022	0.262 ± 0.115
25%	0.992 ± 0.0028	0.394 ± 0.158
10%	0.989 ± 0.0037	0.534 ± 0.195

Figure 3.8 shows lung segmentation by the model on a case from lung cancer screening at dose levels of 100%, 50%, 25% and 10%.

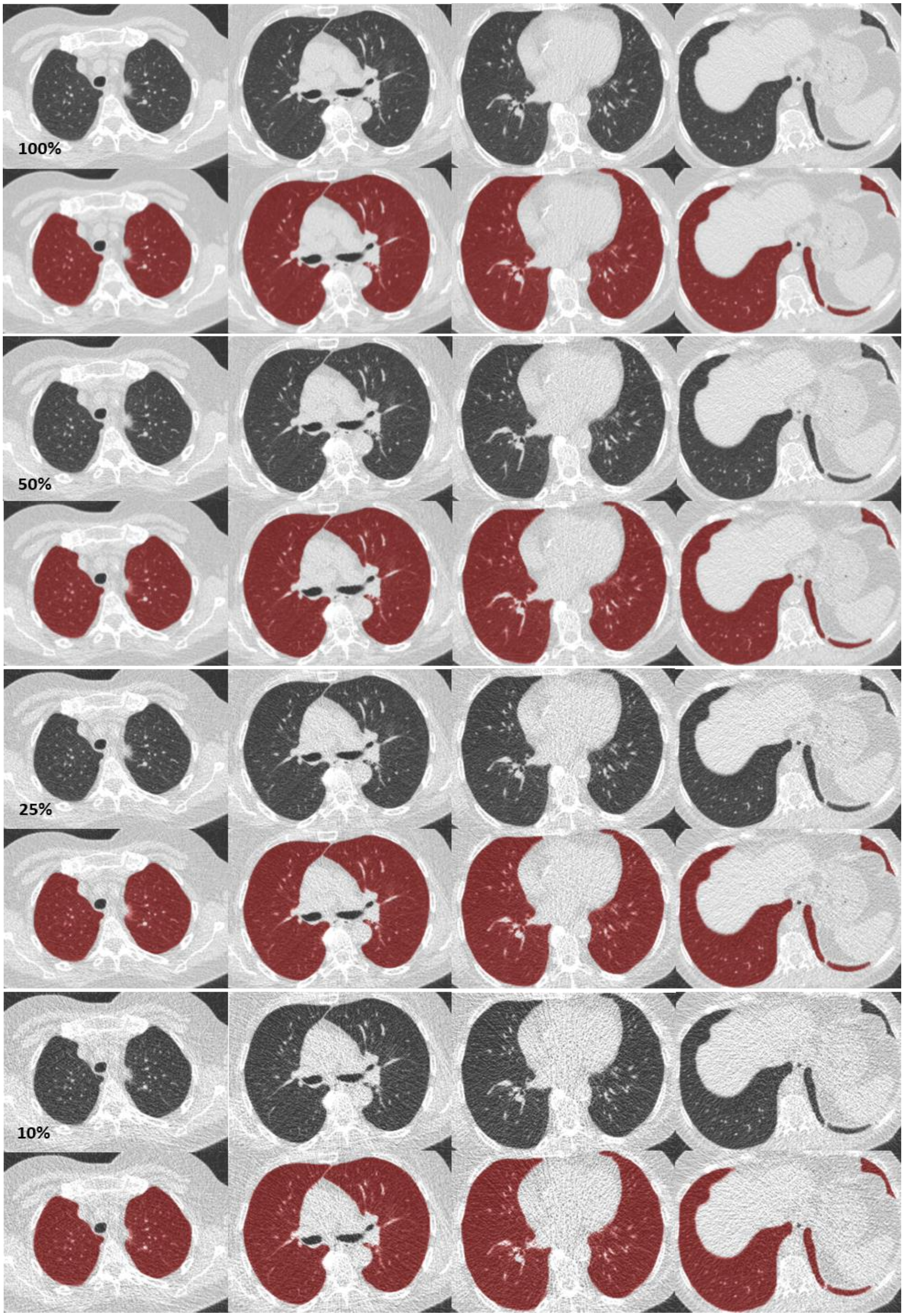


Figure 3.8: CNN segmentation on a cohort from lung cancer screening at different dose levels.

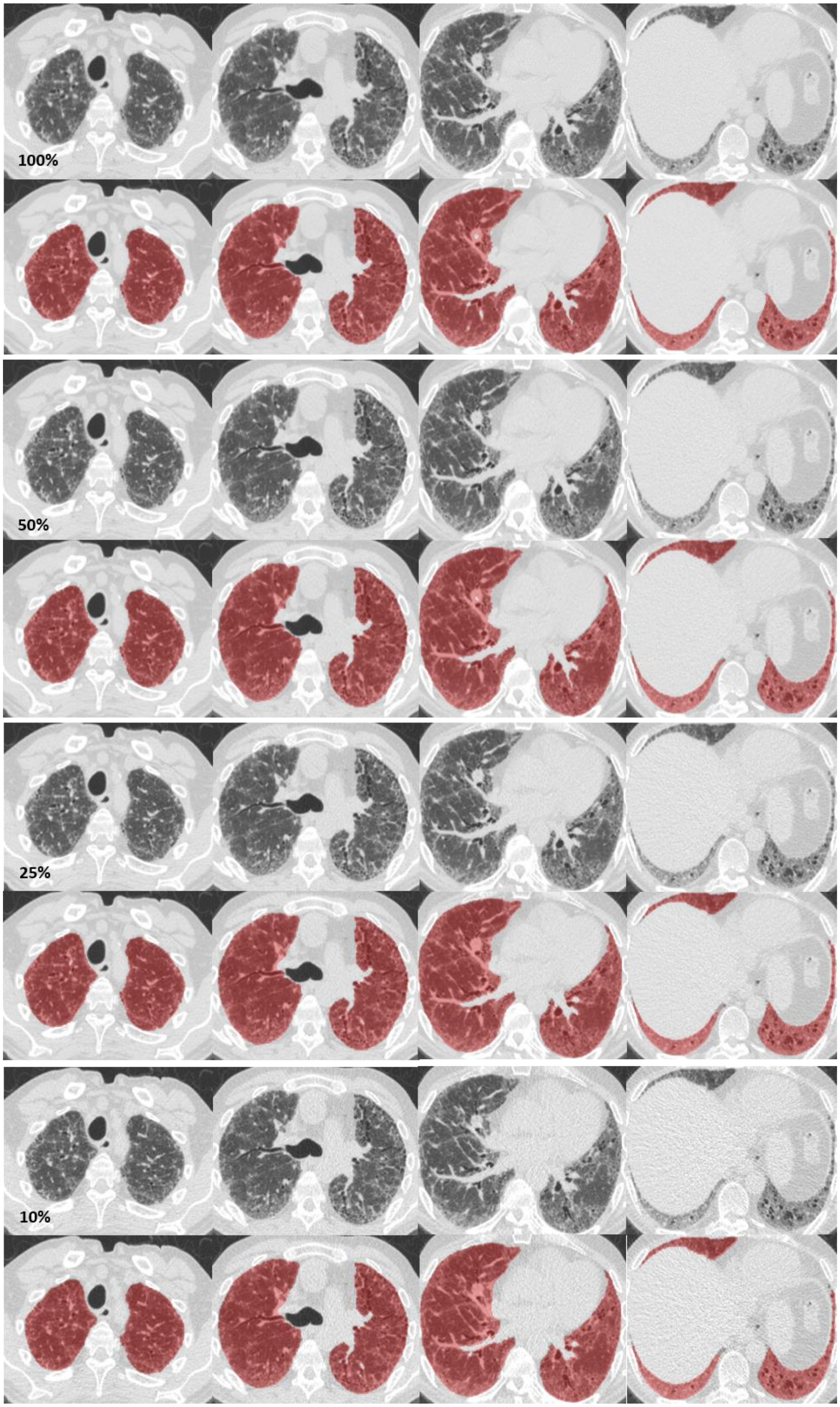


Figure 3.9: CNN segmentation on a cohort from diffuse lung disease trials at different dose levels.

Figure 3.9 shows CNN segmentation results on a case from a diffuse lung disease trial at dose levels of 100%, 50%, 25% and 10%. Table 3.4 shows the average Dice score and ASD compared to the reference segmentation. Table 3.5 is a comparison using the CNN segmentation on the 100% dose scan as reference.

Table 3.4 CNN segmentation results on cases simulated at 4 dose levels from diffuse lung disease trials.

	DSC	ASD (mm)
100%	0.978 ± 0.011	0.842 ± 0.624
50%	0.973 ± 0.015	0.926 ± 0.597
25%	0.969 ± 0.012	1.083 ± 0.739
10%	0.961 ± 0.017	1.242 ± 1.075

Table 3.5 CNN segmentation results on cases simulated at 4 dose levels from diffuse lung disease trials using CNN segmentation on 100% dose scan as reference.

	DSC	ASD (mm)
50%	0.994 ± 0.0014	0.162 ± 0.030
25%	0.990 ± 0.0025	0.275 ± 0.063
10%	0.985 ± 0.0038	0.416 ± 0.138

3.1.5 Discussion

The CNN based method achieved highly accurate lung segmentation based on Dice score and Average surface distance. The 2D CNN model has been successfully applied to segment over 5,000 chest CT scans in clinical practice. In cases of emphysema lung, which have been major applications for lung quantitation, conventional methods are able to segment the lung accurately and our 2D or 3D based CNN segmentation method achieved comparable performance. However, the method showed substantial advantages when extensive lung fibrosis was present, especially in peripheral areas close to other soft tissue. Fibrosis increases the attenuation of the lung and previous segmentation techniques based on thresholding and region growing do not perform as well. Another strength of the CNN based method was its ability to consistently exclude airway trees regardless of the disease pattern and other methods occasionally require post-processing or manual editing. Smoothness constraints on the segmentation could be applied between 2D slices but we obtained accurate results on thin section CT scans without them. The segmentation improvements we achieved are important for quantitative scoring of IPF[302–304] and in applying quantitative imaging in routine clinical practice where a variety of lung parenchymal abnormalities are present.

The ground truth segmentation was derived by manual editing of the threshold based anatomical model method that was used in the comparison, whereas the CNN method was completely independent of the ground truth. Therefore, there may be some bias in favor of the threshold-based technique in our evaluation comparison. Despite this, the CNN still significantly improved segmentation performance.

There has been very little published research to date on pathologic lung segmentation using convolutional neural networks, and thus this work represents a significant contribution. One similar CNN based approach is P-HNN[305] by investigators at the NIH, and they also showed promising results on pathological lung segmentation. Our work differs in terms of using a radiologist edited and approved reference segmentation for training and testing. Secondly, they applied different window-levels to form a 3-channel input image for each slice while we used a simpler single-channel architecture (trained from scratch) without post-processing (such as 3D hole filling) and thus achieved computational efficiency advantages. The most recent SegCaps[290] seems to be a more advanced network over CNN and claimed that a regular CNN based segmentation model was unable to exclude airway completely after comparing with results from U-net. However, it is contradicted by the results in our experiments and our qualitative experience in clinical practice. Besides, they did not prove their model robustness on different kinds of pathological lungs.

Using current datasets and hardware, the 2D slice based CNN slightly outperformed the 3D volume based CNN. Although the 3D model was able to incorporate more contextual information, this is actually not surprising considering the following. Firstly, the variation of slice spacing along the z direction is much larger than in the x-y plane. One strategy would be to resample all the scans to the same resolution but this is computationally expensive. Secondly, the number of training samples is also much smaller when a 3D scan is used instead of individual 2D slices. As such, the 2D model was built with a larger and more diverse dataset. Thirdly, the depth and width of the 3D architecture and number of feature maps used is far less compared to the ResNet-101 based 2D model. Lastly, the image was down-sampled more aggressively in the 3D CNN (512 to 128 vs. 512 to 256). This caused more information loss, especially of fine details near the lung boundary and this was also observed in work by IBM research[306]. However, we believe that 3D volume based segmentation could further improve segmentation if we can build a deeper network with larger input image size and sufficient feature maps at various layers comparable to the 2D counterpart.

The most valuable merit of our model is its generalizability and robustness when applied to diverse chest CT images in clinical practice. The 2D CNN model can be applied not only to high resolution volumetric scans but also incremental scans with large spacing and this is an advantage of the 2D slice model over the 3D volume model.

Additionally, the model is compatible with different lung sizes, e.g. adult or pediatric, RV or TLC. Lastly, the CNN model demonstrated strong performance when applied to low and ultra-low-dose scans.

If necessary, additional types of lung abnormalities or acquisition parameters could be added to the training set to further improve robustness of the lung segmentation in clinical practice. In addition, the RestNet-101 based segmentation architecture is general and could potentially be applied to other object localization and segmentation tasks in medical images.

3.1.6 Conclusion

The proposed 2D ResNet-101 based segmentation method was successfully applied to a variety of pathological lungs and achieved high accuracy without the need for any additional post processing.

3.2 Left and right lung separation

3.2.1 Introduction

The CNN segmentation introduced above results in a whole lung mask and the next step was to separate left and right lungs precisely. It is not as trivial as using a straight line or plane to separate them in 2D or 3D since the two lungs are often closely connected. Figure 3.10 exhibited some example images with very small junctions between left and right lungs at anterior and posterior lung respectively.

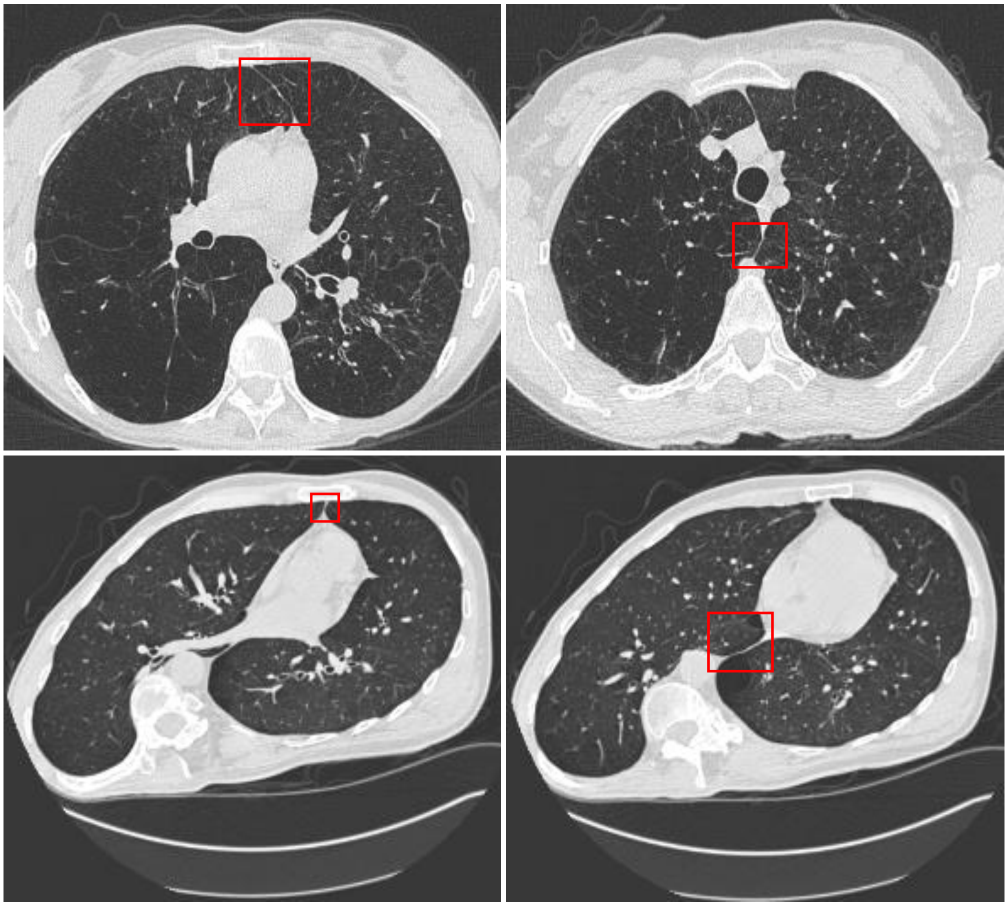


Figure 3.10: Scans with thin junctions at anterior and posterior lungs.

The simplest way is to apply morphological operations[12,307], e.g. erosion followed by dilation to isolate the two lungs. However, it is challenging to find an optimal kernel size and number of iterations for a broad spectrum of images, and excessive erosion could severely deteriorate the original shapes. Lee et al.[308] added Hessian matrix analysis to limit the iterations of 3D erosion and a surface-fitting algorithm was used to optimize the separating planes. A region-growing based method[309] was also used that assumed that the growing from two seeds from left and right lung would meet each other at the correct locations, whereas it was not robust in case with thin junctions or cases with pulmonary diseases. The most common way is first to identify suspicious connecting areas and then find the best separating lines between two lungs using minimum path strategies[275,310]. Brown et al.[298] segmented the mediastinum using an anatomic knowledge in a semantic model and determined the separating curve by dynamic programming. Leader et al.[311] applied a heuristic method to detect the narrowest junction regions in the anteroposterior direction and then locate the separating line by finding the brightest pixels in the junctions. To overcome the uncertainty in separating lungs on a single 2D slice, a volumetric scan based method was introduced incorporating sequential information under the assumption that the change between slices is very small on high resolution CT[308]. Although contextual information from adjacent slices were used, the separating line was still identified in a slice by slice fashion and was unsuccessful if there were two junctions at the anterior and posterior regions of the lung on the same slice. In summary, a few drawbacks included: 1. additional segmentation of other landmarks to find suspicious connecting junctions and dependency of scan orientation to search the best lines; 2. the separating line is identified slice by slice regardless of whether the CT is volumetric or not; 3. most of the methods were computationally expensive, especially the dynamic programming algorithm since it tends to find all potential pathways; 4. they are not really generalizable when confronted with a variety of pathological lungs or incremental scans (large slice spacing) since many of them have empirical parameter settings, e.g. lung volume.

It is challenging to separate the left and right lungs based on 2D axial images since not all context is available. Considering its spatial relationship with surrounding anatomies, such as the aorta, liver, and spleen, it should be feasible to perform the separation in 3D. Moreover, performing separation in 3D space (instead of slice by slice as in previous approaches) would allow resolution of multiple junction slices in parallel and significantly improve the computation efficiency. As such, a 3D CNN for whole lung segmentation before was proposed in this work, with output of separated left and right lung masks.

3.2.2 Methods and materials

The same normalization was applied and the input image was resized to 128 × 128 × 128 to address memory limitations. As mentioned in the discussion of the whole lung segmentation in chapter 3.1.5, the resultant lung segmentation was sometimes under-segmented near the boundaries due to the aggressive down-sampling, especially in the z axis. The down-sized image may lose useful details in junction areas and the resulting separation may not be as good as using the original resolutions. Alternatively, the watershed algorithm[312–314] is widely used to separate adjacent objects in many segmentation problems. Presumably, it should be able to correct the overflow if the 3D left and right lung separation is not perfect from the 3D CNN. As for the under-segmentation problem, watersheds could also contribute by exploiting the previous whole lung segmentation to supplement. To sum up, the watershed algorithm was used to take the left and right lungs from the 3D CNN as seed points or chunks while using previous accurate whole lung segmentation as a mask to fill initially under-segmented pixels. The flood filling of these areas is fast by the watershed algorithm since the majority of pixels (> 99%) have been covered in the original left and right lung mask from the 3D CNN.

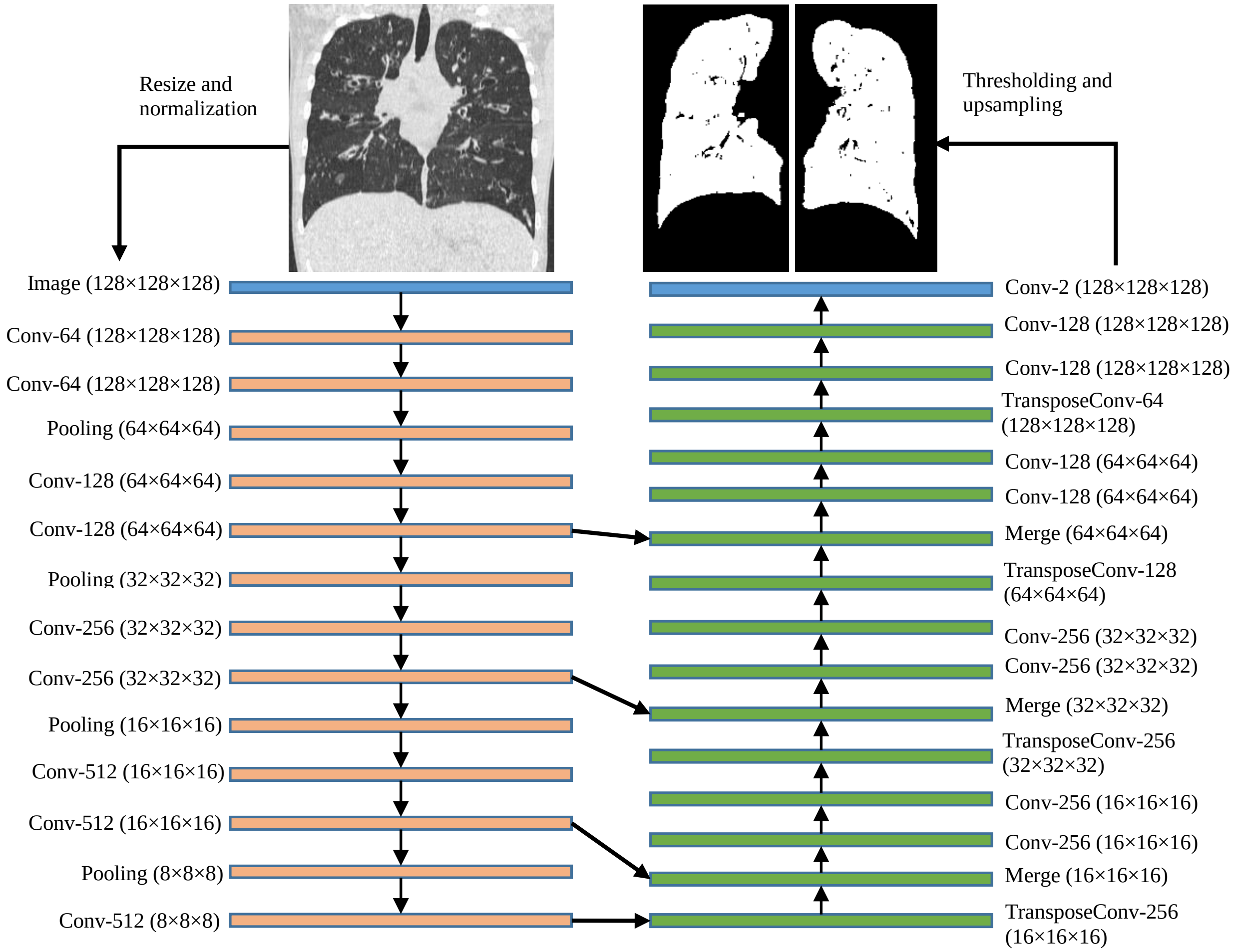


Figure 3.11: 3D CNN architecture to segment the whole lung with output of left and right lung simultaneously.

The architecture to segment and separate the left and right lungs is shown in Figure 3.11. The input s a raw scan and output is an isolated 3D segmentation of the left and right lungs. Dice loss was used as the cost function. Data augmentation was applied in training, including rotation up to 30 degrees within the x-y, y-z and z-x planes, random flipping along each axis. Accordingly, the trained model was designed to be robust to scans of different orientations, e.g. A-P or P-A, L-R or R-L, S-I or I-S. The model was trained from scratch on a NVIDIA TITAN X with GPU memory of 12 GB, using the Adam optimizer, and learning rate of 0.001 with a decay rate of 1e-4.

The same data set used for the whole lung segmentation was used again for the 3D left and right segmentation. The same data sampling and splitting strategy was applied. In the test set, an additional 100 low resolution scans with slice spacing of 5 mm to 20 mm were used to evaluate the left and right lung identification.

3.2.3 Results

Table 3.6 lists results of the left and right lung separation by 3D CNN plus watershed. To exclude the impact of using different lung segmentation methods on the Dice score, watershed was performed based on the whole lung mask from the reference instead of the CNN segmentation introduced in Chapter 2.

Table 3.6 Results of left and right lung separation using 3D CNN and watershed.

	Dice score of left lung	Dice score of right lung
High resolution scans (115 scans)	99.96%	99.97%
Low resolution scans (100 scans)	99.99%	99.99%

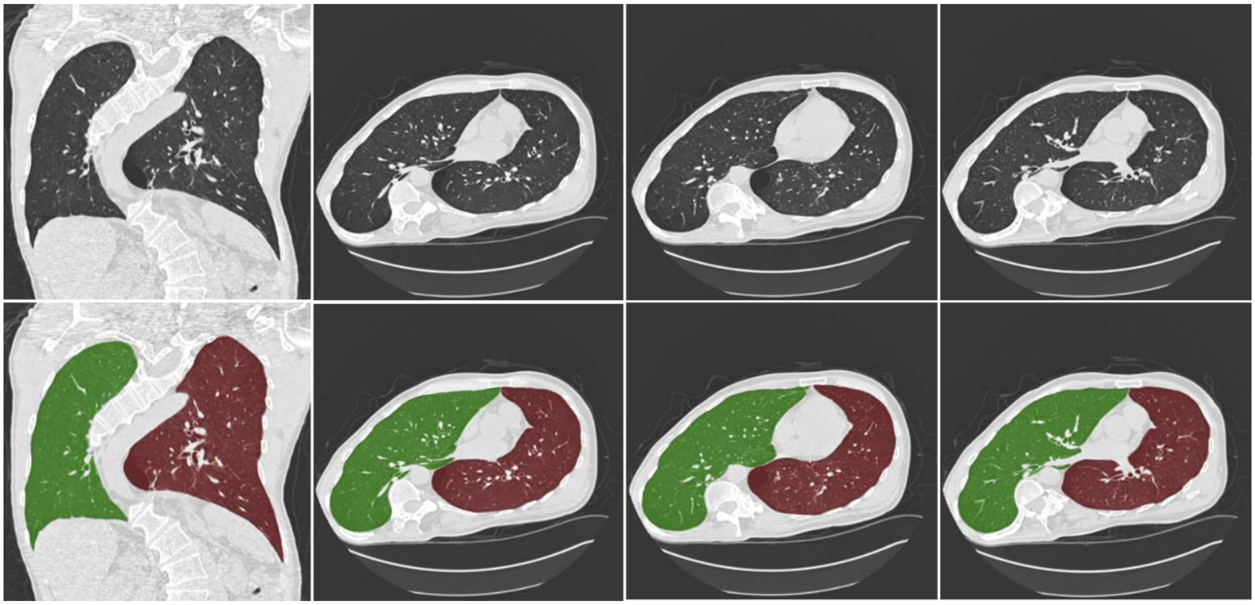


Figure 3.12: An example from LOLA11 with multiple small junctions at both anterior and posterior of lung.

Figure 3.12 shows a case from LOLA 11 with a tilted chest. It is a challenging case with multiple left and right lung junctions both anteriorly and posteriorly. The proposed method was able to separate the left and right lungs precisely.

One concern was the application of the 3D model to incremental scans (slice spacing > 10 mm) since the model was trained using only high resolution scans (0.6-3 mm). Figure 3.13 contains an example scan with slice spacing of 20 mm and 14 slices in total. The proposed method successfully achieved accurate left and right lung separation.

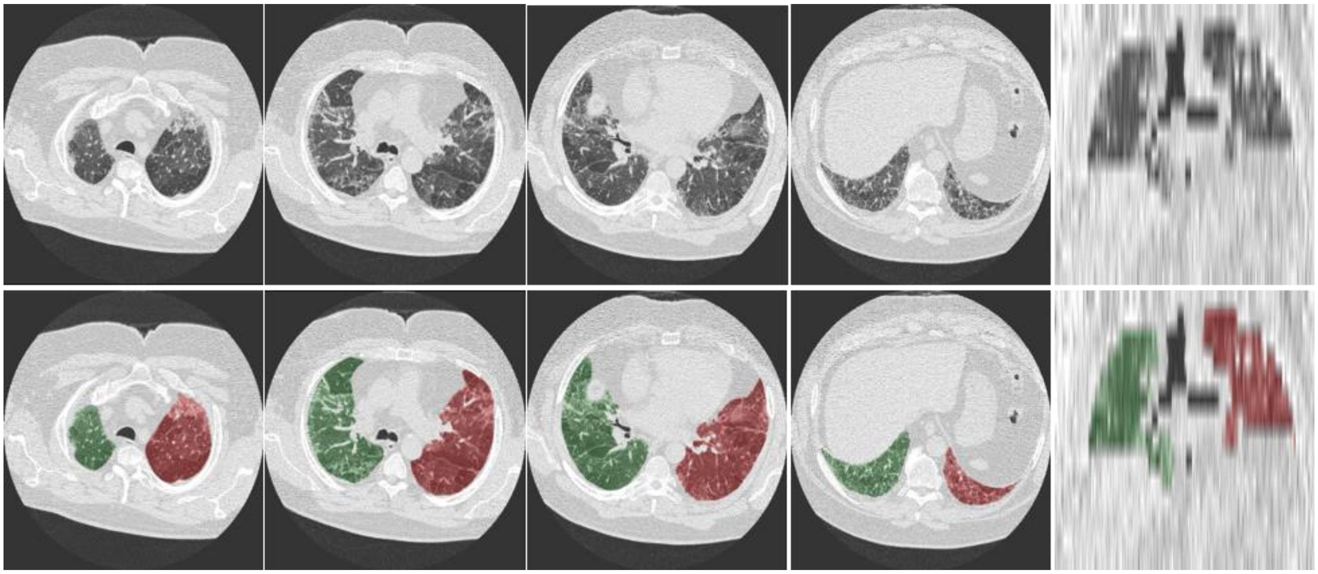


Figure 3.13: Left and right lung separation on an ultra-low resolution scan segmented by the 3D CNN and watershed algorithm.

The robustness was also tested on pediatric patients, and Figure 3.14 is an example with slice thickness and spacing of 1.25 mm. The previous 2D ResNet contributed to accurate whole lung segmentation and the proposed separation approach was able to isolate the two lungs precisely.

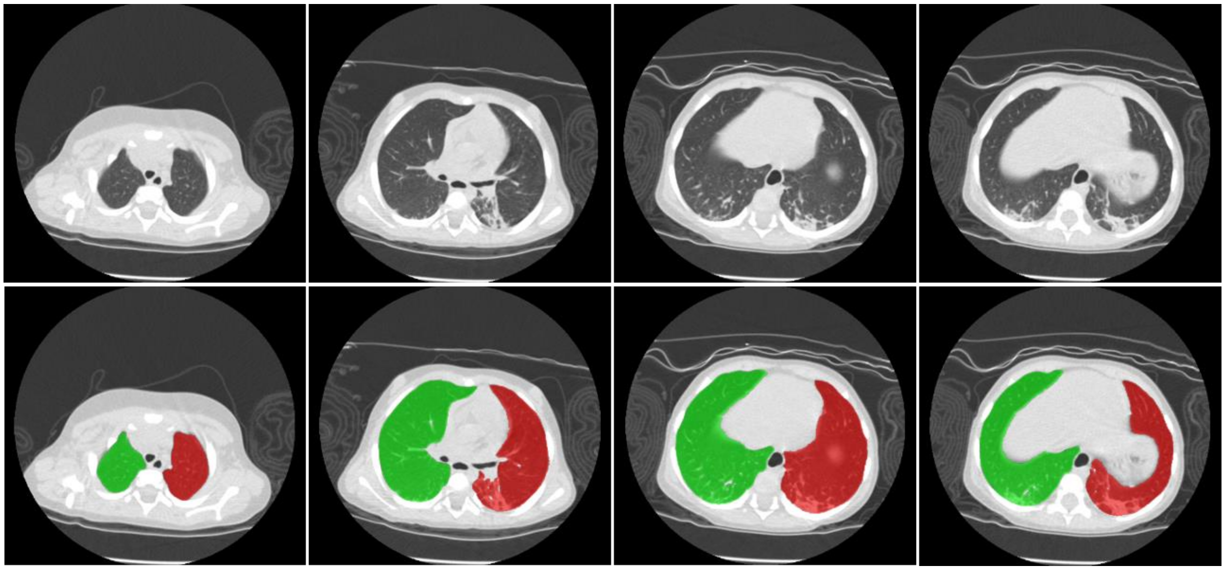


Figure 3.14: Left and right lung segmentation on a pediatric scan by the proposed method.

3.2.4 Conclusion

The proposed 3D CNN plus watershed supplement was successfully applied to separate left and right lung from the whole lung segmentation of chest CT and it is compatible with both high and low resolution scans, and adult and pediatric patients.

3.3 Lobar segmentation

3.3.1 Introduction

The entire human lung usually consists of five pulmonary lobes and they can to some extent be treated as individual functional units since both bronchial and vascular systems are mostly separated without many connections across the different lobes. As shown in Figure 3.15 [315], the left lung is usually divided into the left upper lobe (LUL) and left lower lobe (LLL) by the left major fissure (or left oblique fissure). The right lung contains three lobes, namely the right upper lobe (RUL), right middle lobe (RML), and right lower lobe (RLL), separated by two pulmonary fissures. The right minor fissure, also known as right horizontal fissure, separates the right upper lobe from the right middle lobe, and the right major fissure, a.k.a right oblique fissure, subdivides the right middle lobe and right lower lobe. Lobar segmentation is of great clinical significance to perform local or regional quantitative analysis. For example, idiopathic pulmonary fibrosis is more likely to affect the lower lobes[316], while emphysema, tuberculosis and pulmonary nodule are preferentially related to the upper lobes[317–320]. Furthermore, lobar segmentation on CT can be mapped to other imaging modalities, e.g. MRI[321], where inter-lobar boundaries are not visually apparent.

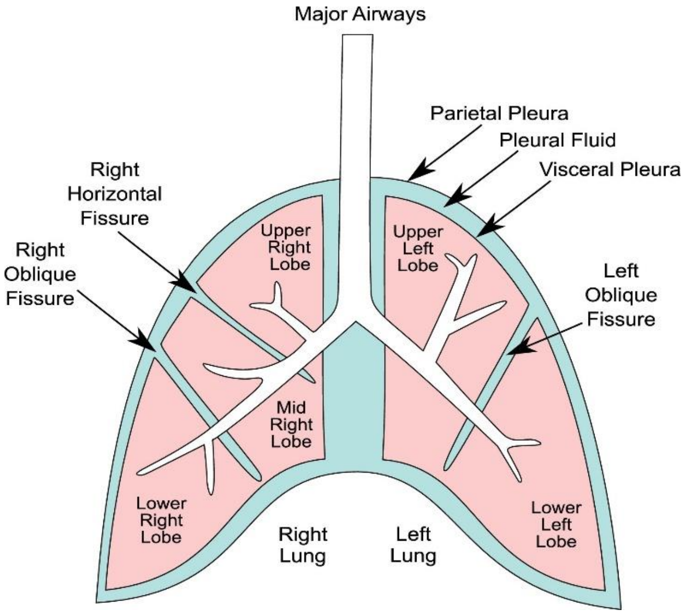


Figure 3.15: Typical anatomy of lung and three fissures to separate left lung into two lobes and right lung into three lobes.

Laborious manual segmentation of the five lobes is not feasible in clinical practice since it is extremely time-consuming (>1 hour per case) to outline the lobar boundaries. As such, automated lobar segmentation is essential for high throughput quantitative analysis on chest CT. Many automated lobar segmentation methods have been proposed and most of them aim to detect or segment the three fissures to separate the lobes. Early approaches focused on the physical appearance of fissures and more recent algorithms adopted further anatomical structure and semantic knowledge as supplements. For example, the spatial relationships of the lungs, airway tree, and vessels at different levels were utilized. A 3D geometry mesh based method[23] was developed using Laplacian smoothing, and marching cubes to boost the surface shape contrast within the lung, and radial basis function based surface fitting was applied to segment the lobes. A similar approach using B-spline fitting was also proposed[24]. This type of method struggles to differentiate planes caused by lung pathologies from real fissures. Hessian matrix based filtering was used to extract candidate fissure points in several studies[20,21]. A similar method[22] by machine learning was also introduced and features from multi-scale Gaussian filters were incorporated besides the Hessian matrix. It achieved better performance at the cost of computational complexity. Zhou et al.[322] proposed to search for fissures by exploiting blood vessels connected within the lobes, and segmented vessels were designated to different lobes following the extension of the airway tree. Another approach also using vessel assistance applied watershed segmentation to a distance transformed search image and marker points generated from the airway were used to segment lobes. A derivative of stick (DoS) filter[25] was introduced to enhance fissure visibility by defining non-linear derivatives with a stick kernel in multiple directions. Post-processing including cascading plane integration with 3D surface shape tuning.

Overall, those traditional methods are not generalizable when applied to a wide range of lungs with various diseases, especially when the fissures are incomplete. Raasch et al.[323] reported 40% and 46% of fissures as incomplete for the upper and lower left major fissures. As for the right lung, the incomplete rate is 70% for the upper major fissures, 47% for the lower major fissures and 94% for the minor fissures. Deep learning methods have also been used for fissure detection. Gerard et al. proposed FissureNet[324], a coarse-to-fine cascade of two convolutional neural networks to extract and classify fissures, and it mitigates the challenge of directly segmenting thin structures in a 3D volume. George et al.[305] presented a progressive and multi-path holistically nested 2D network (P-HNN) method for lobar segmentation using 2D axial slices in which a crude CNN segmentation was followed by a 3D random walker to refine it. The lobar segmentation from the 2D CNN was often inaccurate and they were used as seed points or patches for further post-processing. Presumably, it is challenging to consistently segment different lobes based on 2D slices without spatial context, especially on slices containing misleading pathologies.

In our work, we proposed a 3D DenseNet[119] neural network to directly segment the five lobes, skipping fissure detection as a prerequisite.

3.3.2 Methods

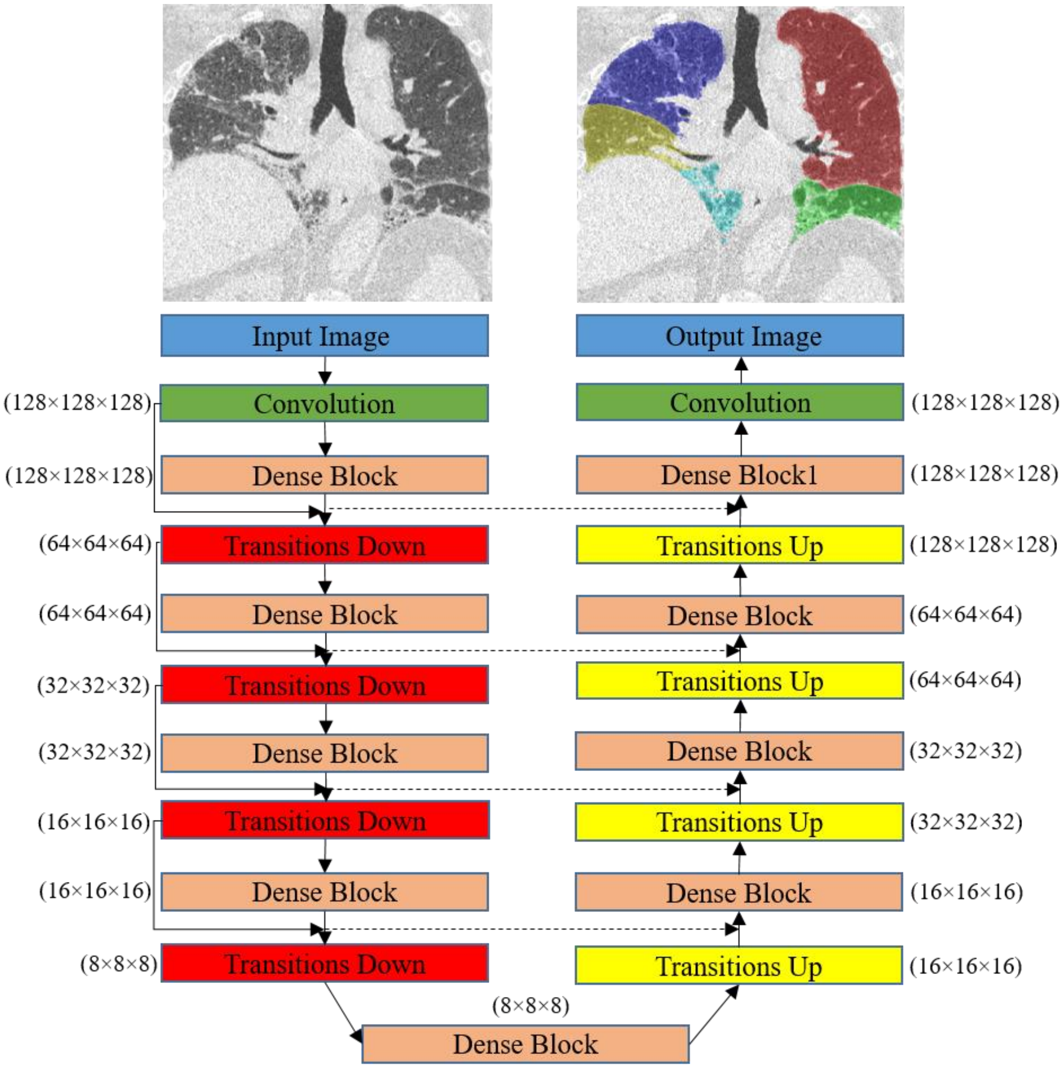


Figure 3.16: 3D DenseNet based lobar segmentation architecture.

The 3D DenseNet is shown in Figure 3.16. The input and output sizes are 128 × 128 × 128. The down-sampling path includes 4 Transitions Down and 4 Dense Blocks. The corresponding up-sampling path includes 4 Transitions Up and 4 Dense Blocks. The transitions Down modules include [Convolution3D, Dropout and Max-pooling]. The transitions Up modules include [Transposed Convolution3D, Convlution3D, and

Dropout]. The Dense Block includes 4 densely connected layers, meaning each layer has access to all previous layers.

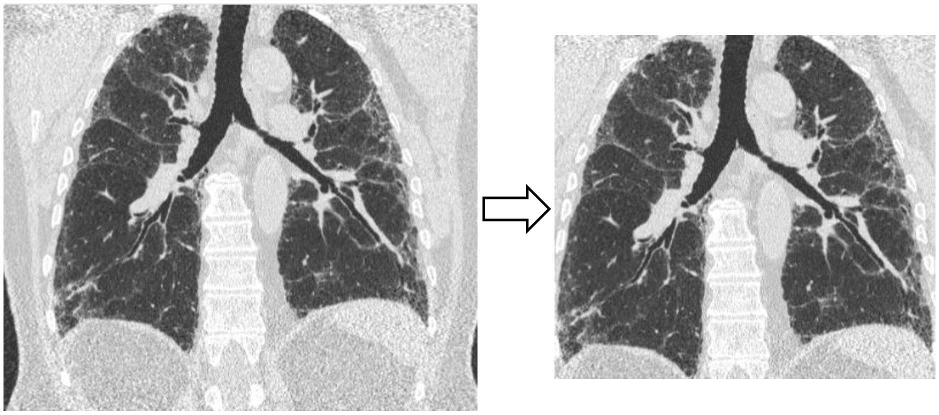


Figure 3.17: Cropping the scan based on lung segmentation.

To mitigate memory constraints, we performed lobar segmentation on the left and right lung separately. A chest CT scan always contains a small portion of the shoulder above the lung, and the abdomen below, that are not necessary for lobar segmentation. To allow the network to focus on the lungs with better receptive fields, each scan is cropped based on the lung segmentation as shown in Figure 3.17. After cropping, each scan along with the corresponding lung segmentation was resized to 128 × 128 × 128. The input to the network is a 2-channel 3D image, composed of the raw image plus the lung mask acquired from CNN lung segmentation introduced before. The lung mask is expected to help save the network from performing lung segmentation again and pay more attention to regional features within the lung.

Data augmentation was applied in training, including rotation up to 15 degrees based on x-y, y-z and z-x planes, random flipping along x, y, z-axis. This helps to make the model robust to scans with a variety of orientations. The model was trained from scratch on a NVIDIA TITAN X with GPU memory of 12 GB, using the Adam optimizer, and a learning rate of 0.001 with a decay rate of 1e-4. Initially, regular dice loss was used as the loss function and it was noticed that the right middle lobe tend to be under-segmented. It was not too surprising since the right middle lobe is relatively small compared to others. As such, a weighted dice loss using [1.20 1.2, 1.5, 4, 1.5] for [LUL, LLL, RUL, RML, RLL] was applied to reward loss on smaller lobes.

3.3.3 Materials

A total number of 1,280 scans from different patients were used in building the model, all with radiologist approved lobar segmentations as ground truth. The data was from 12 multi-center clinical trials, using scanners from GE, SIEMENS, PHILIPS, and TOSHIBA. 515 scans were from COPD trials and 765 scans were from IPF trials. The slice thickness range was [0.6 mm, 3 mm] and reconstruction kernels were from smooth to sharp. The resolution (z spacing) limit was <= 3mm since our method is 3D volume based and larger spacing is not feasible for lobar segmentation.

Similar to the dataset used in lung segmentation, it also included different types of common lung abnormalities, e.g. emphysema, ground glass, fibrosis, etc. The same data sampling and splitting strategy was used to form training and test sets with a 4:1 ratio, and 5-fold cross validation was applied during evaluation. In terms of breath-hold, 1,193 scans were acquired at Total Lung Capacity (TLC) and 87 at Residual Volume (RV). Figure 3.18 shows two emphysema cases with the left lung in the first column and right lung in the second column. The case in the first row has a clear fissure in the presence of emphysema. However, fissures are barely visible on the case in the second row.

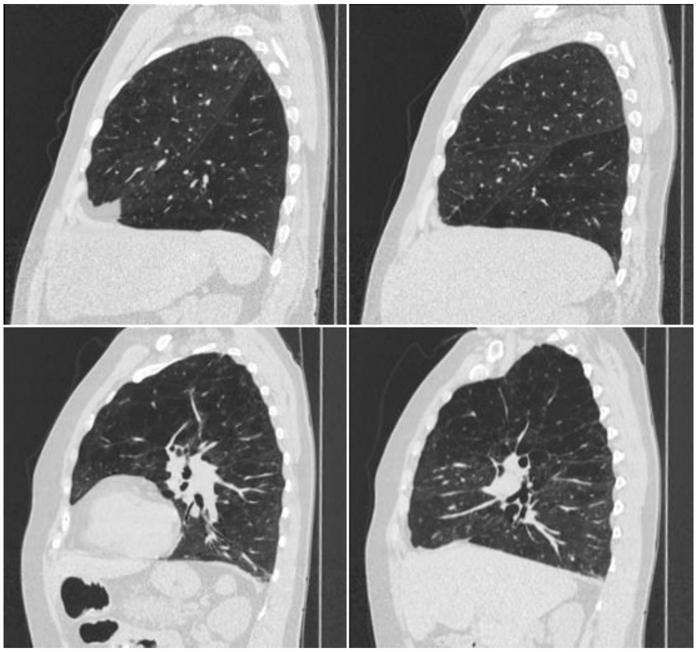


Figure 3.18: Emphysema cases with and without visible fissures from training set.

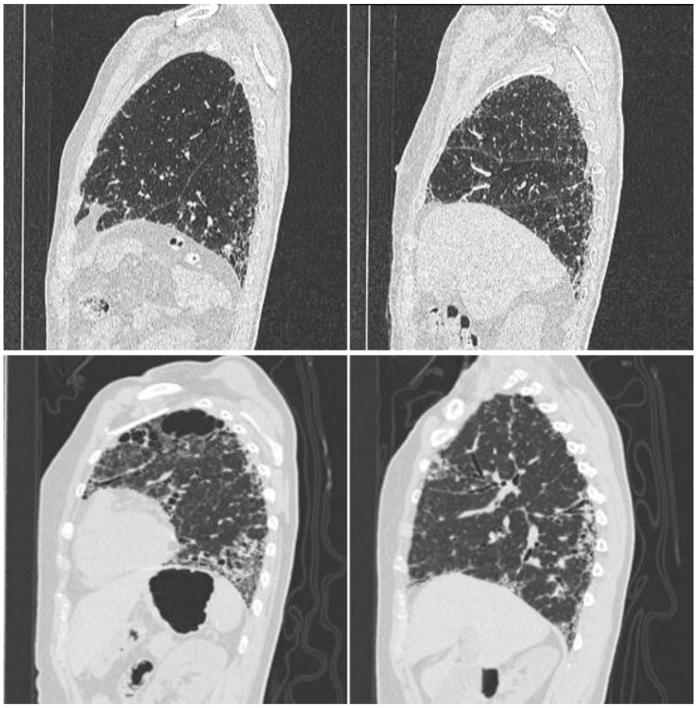


Figure 3.19: Fibrotic cases with and without visible fissures from training set.

Figure 3.19 are two examples of IPF scans. The one in the first row shows discernible fissures on both the left and right lung despite the presence of fibrosis. The case in the second row contains with more intense pathology and it is difficult to distinguish the fissures.

3.3.4 Results

5-fold cross validation was used during evaluation. The overall Dice score across the 5 lobes was 0.959 ± 0.087 and the average surface distance was 0.873 ± 0.61 mm. More specific performance on each lobe is shown in Table 3.7.

Table 3.7 Segmentation results of 5 different lobes by 3D DenseNet.

	Dice coefficient	ASD (mm)
RUL	0.971 ± 0.078	0.699 ± 0.432
RML	0.923 ± 0.114	1.542 ± 1.164
RLL	0.970 ± 0.126	0.783 ± 0.372
LUL	0.972 ± 0.083	0.807 ± 0.594
LLL	0.962 ± 0.105	0.861 ± 0.753

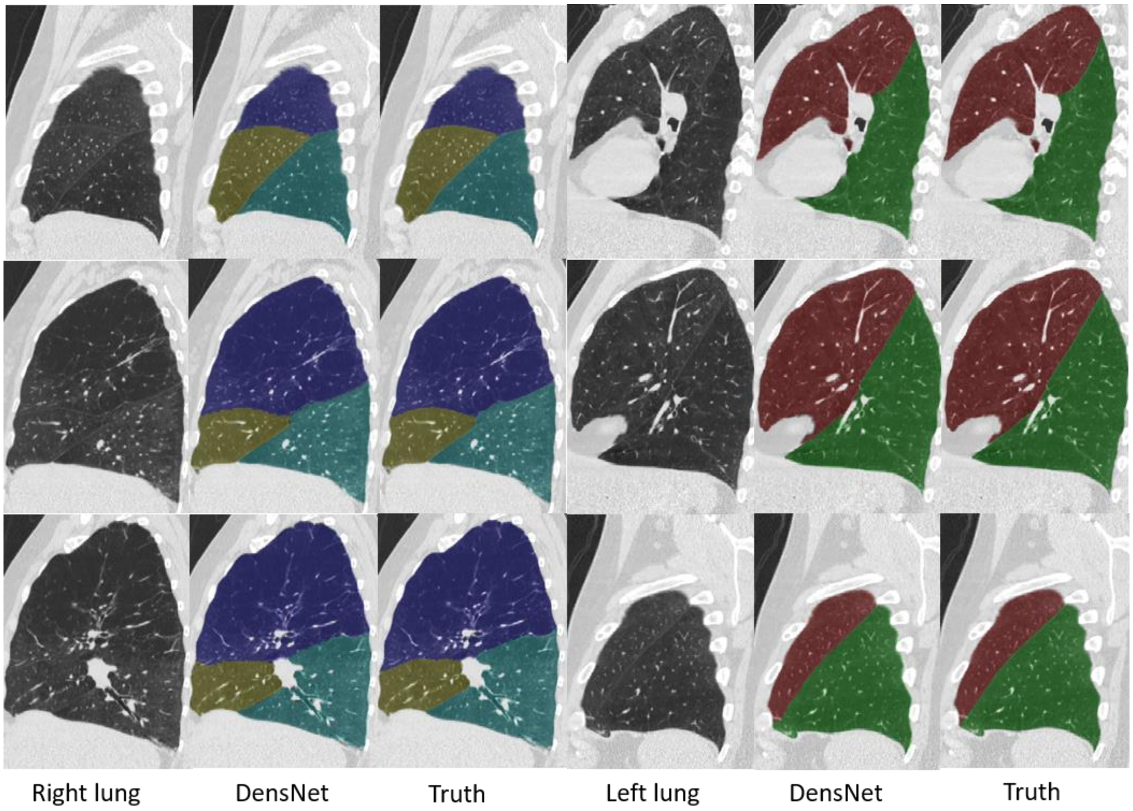


Figure 3.20: Lobar segmentation on an emphysema case at TLC by CNN.

Figure 3.20 shows a lobar segmentation example in the sagittal plane: an emphysema case at TLC and results for each lobe are [LUL: 0.983 and 0.514 mm, LLL: 0.981 and 0.570 mm, RUL: 0.981 and 0.561 mm, RML: 0.970 and 0.673 mm, RLL: 0.977 and 0.635 mm].

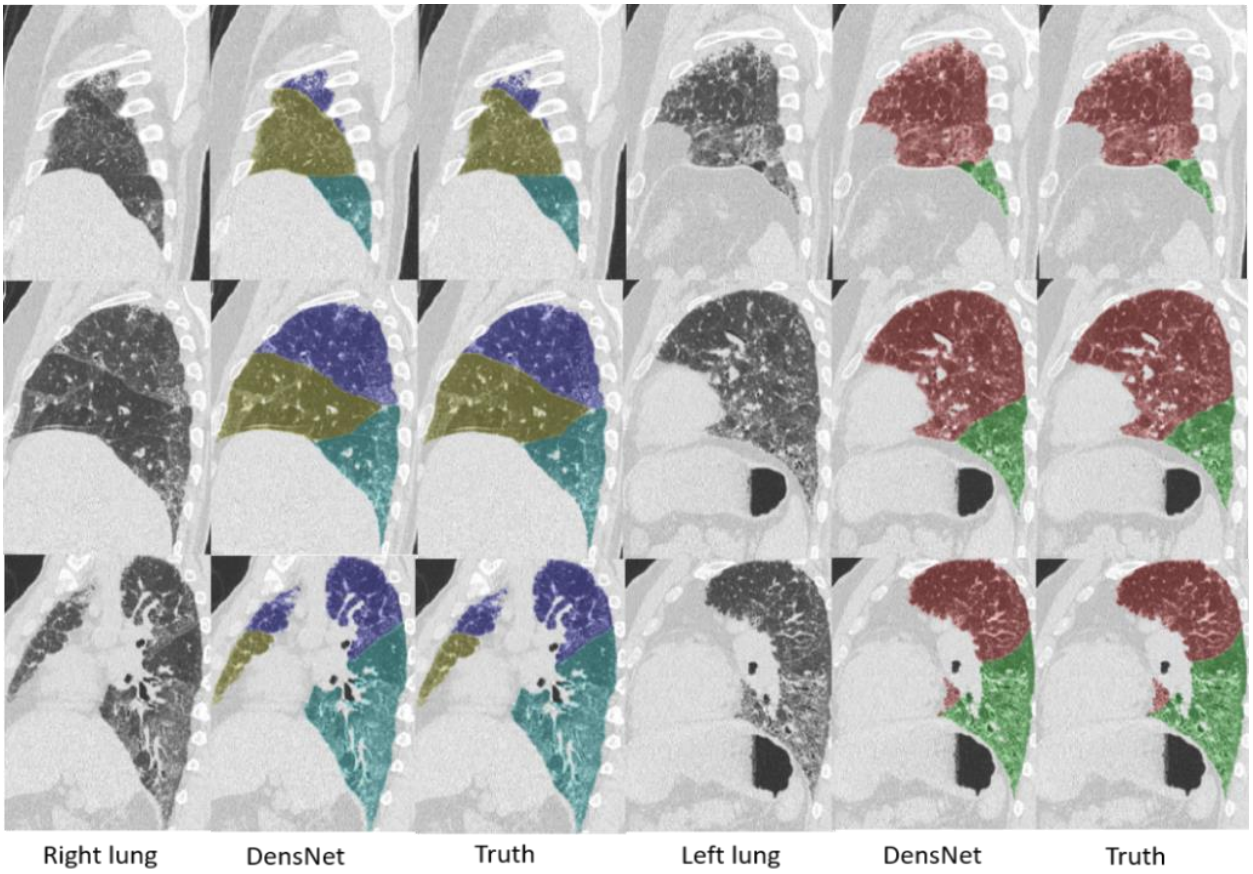


Figure 3.21: Lobar segmentation on a fibrosis case at TLC by CNN.

Figure 3.21 shows lobar segmentation in a fibrosis case at TLC [LUL: 0.978 and 0.427 mm, LLL: 0.960 and 0.538 mm, RUL: 0.975 and 0.496 mm, RML: 0.967 and 0.585 mm, RLL: 0.965 and 0.583 mm]. Figure 3.22 is segmentation on a scleroderma case at RV [LUL: 0.977 and 0.453mm, LLL: 0.973 and 0.544mm, RUL: 0.979 and 0.444mm, RML: 0.947 and 0.584mm, RLL: 0.974 and 0.496mm].

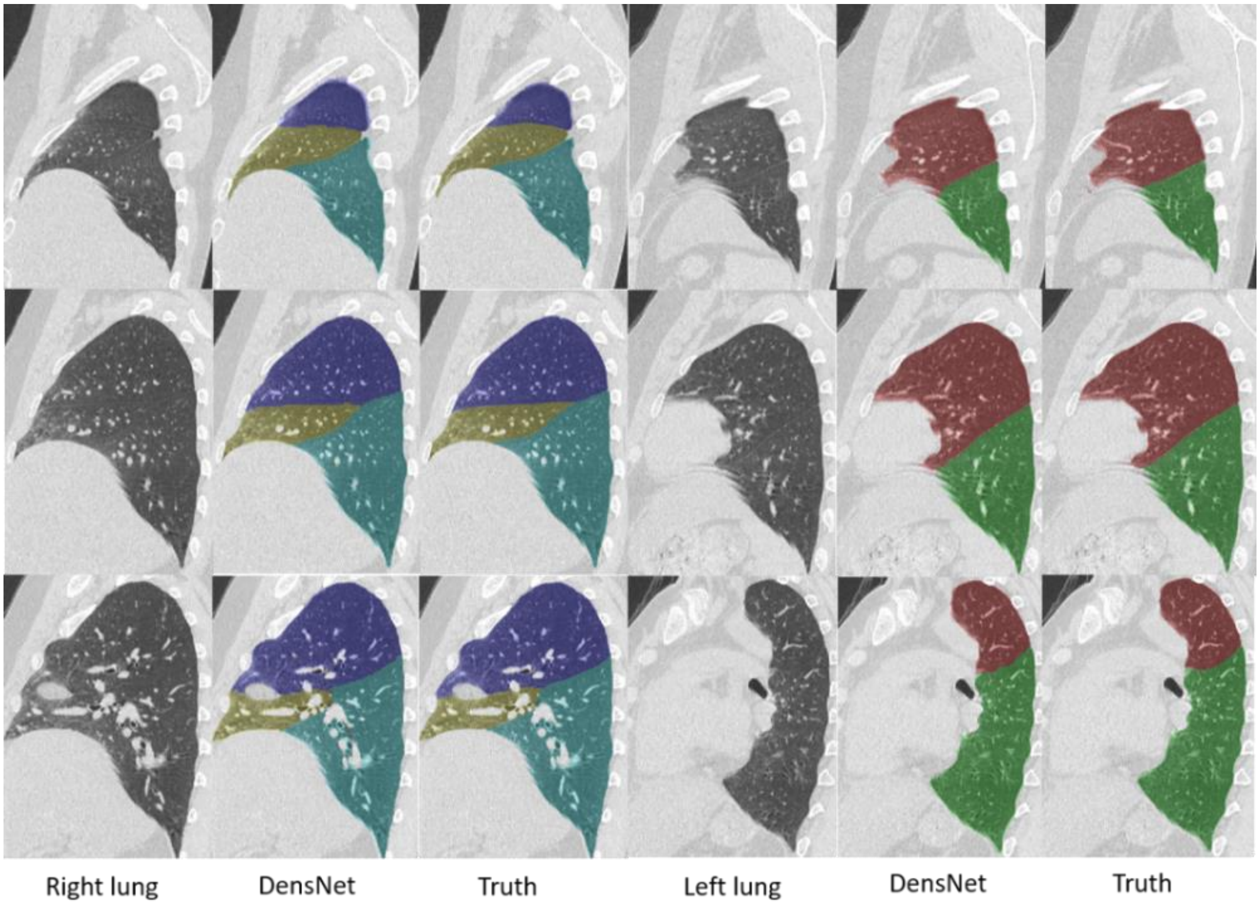


Figure 3.22: Lobar segmentation on a scleroderma case at RV by CNN.

3.3.5 Discussion

Using the initial lung segmentation from the ResNet-101, the 3D DenseNet model achieved good lobar segmentation performance on various pathological lungs. Conventional fissure detection based methods often fail when the fissures are incomplete or impacted by lung abnormalities, such as emphysema and fibrosis. However, our 3D CNN model demonstrated robustness in those challenging cases. One deficiency of the current model is the assumption that five lobes are present, which may not be true when one or more lobes collapse. In cases with atelectasis[325], potential solutions using current models can be applied depending on the situation. For the left lung, lobar segmentation is actually not necessary if there is one collapsed lobe and similarly in the right lung if there are 2 collapsed lobes. Therefore, the most difficult scenario is when there is one collapsed lobe in the right lung. In fact, the task becomes dividing the right lung into two lobes as in the left lung. Figure 3.23 is an example of applying the left lobar segmentation model to the right lung with a collapsed right middle lobe and performance is as follows: [LUL: 0.981 and 0.343mm, LLL: 0.987 and 0.375mm, RUL: 0.977 and 0.344mm, RLL: 0.975 and 0.412mm].

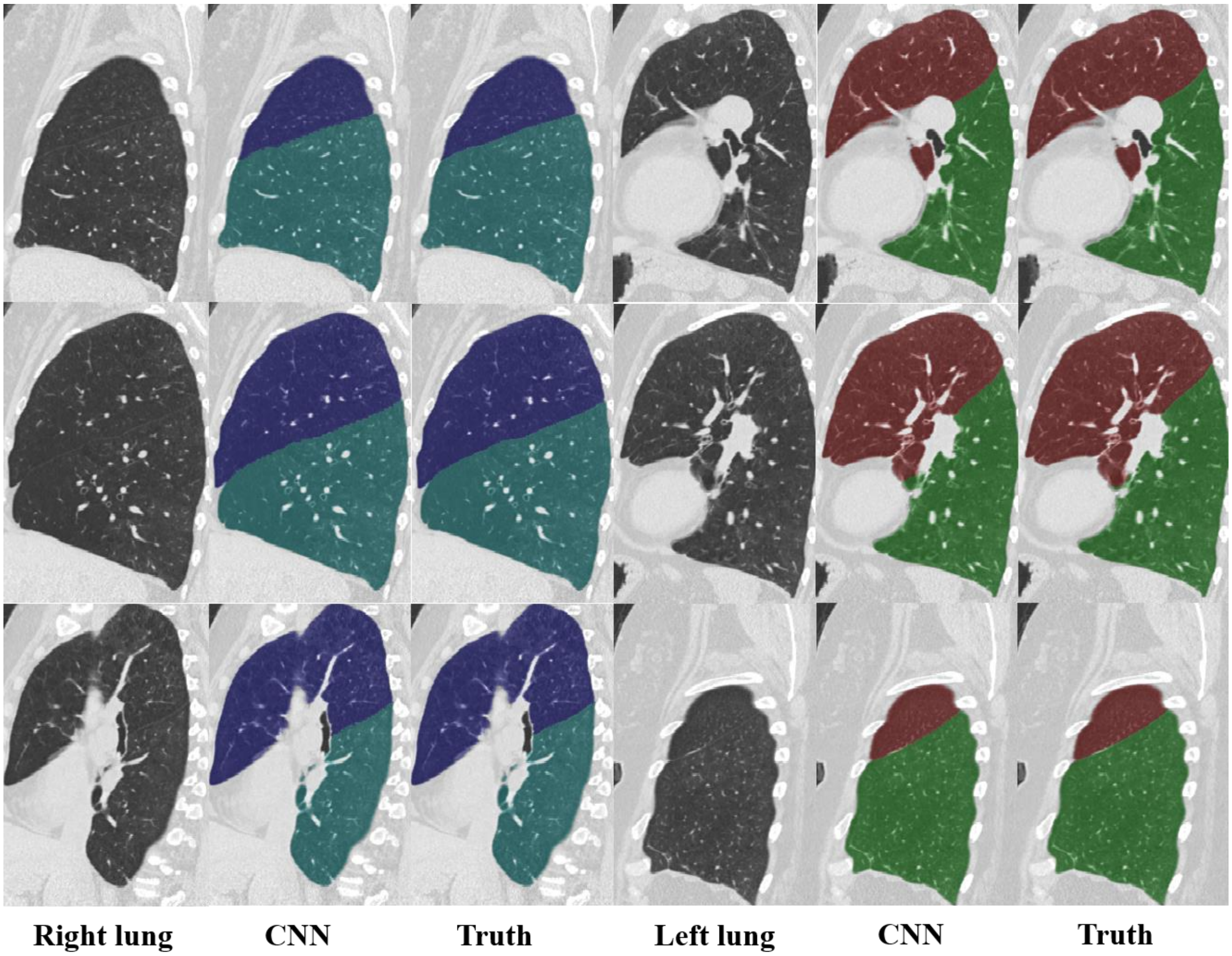


Figure 3.23: Left and right lung (no RML) segmented by the same left lobar segmentation model.

Recently we augmented the training set to 4,000 cases and also diversified it with about 70 cases with atelectasis either on left or right lung. However, there is no significant improvement on cases with collapsed lobes. The atelectasis cohort is probably too small compared to the total number of training samples and one concern is that the network may be confused between cases without visible fissures and those with collapsed lobes.

The lobar segmentation model was also evaluated on low-dose and ultra-low-dose dataset. Figure 3.24 shows an example of lobar segmentation on an ultra-low-dose lung cancer screening scan with a dose level of 0.2 mGy (1/75 of training set dose). The CNN demonstrated impressive robustness in segmenting the lobes as indicated by the previous proposed deep learning based classification and other segmentation methods (aorta, trachea, lung). Quantitative results, such as the Dice score and ASD, are not presented since no reference segmentation is available.

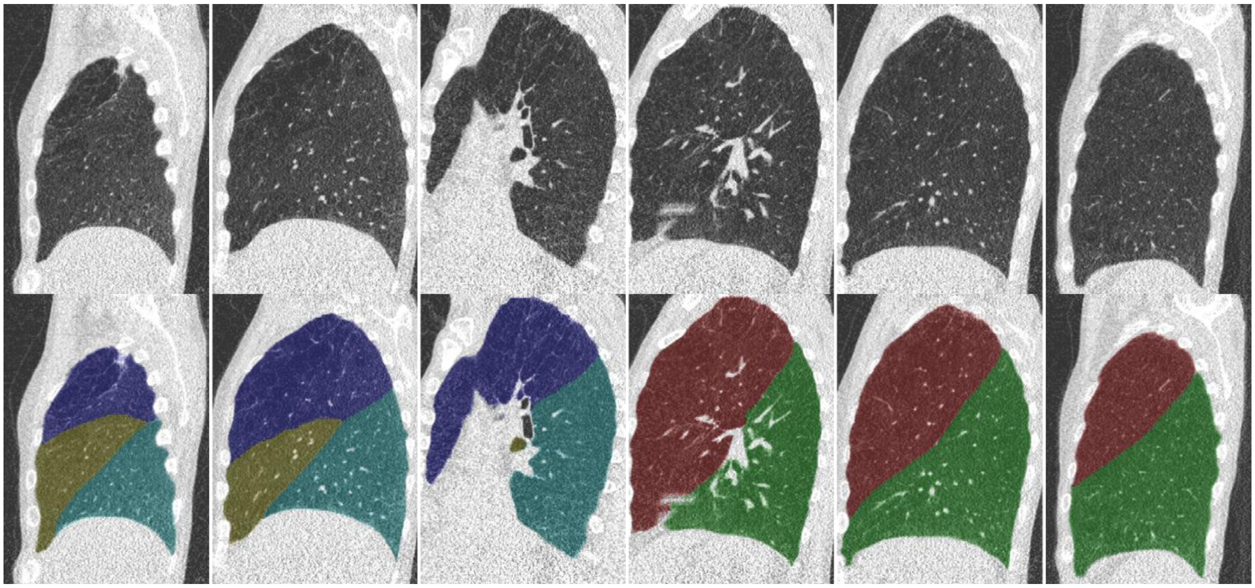


Figure 3.24: Lobar segmentation on an ultra-low-dose scan from lung cancer screening by 3D CNN.

Figure 3.25 is example of segmentation on a low-dose scan from a diffuse lung disease trial. Similarly, no quantitative evaluation was performed since we do not have reference segmentations. This model resilience will be highly beneficial since low-dose scans are used extensively in lung cancer screening.

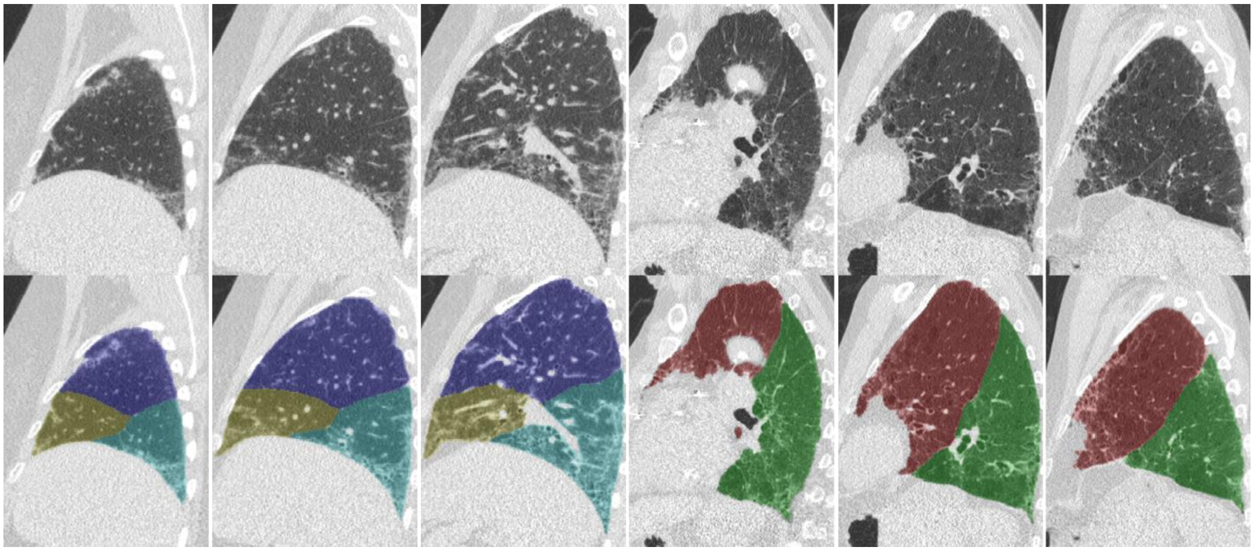


Figure 3.25: Lobar segmentation on a low-dose scan from diffuse lung disease trial by 3D CNN.

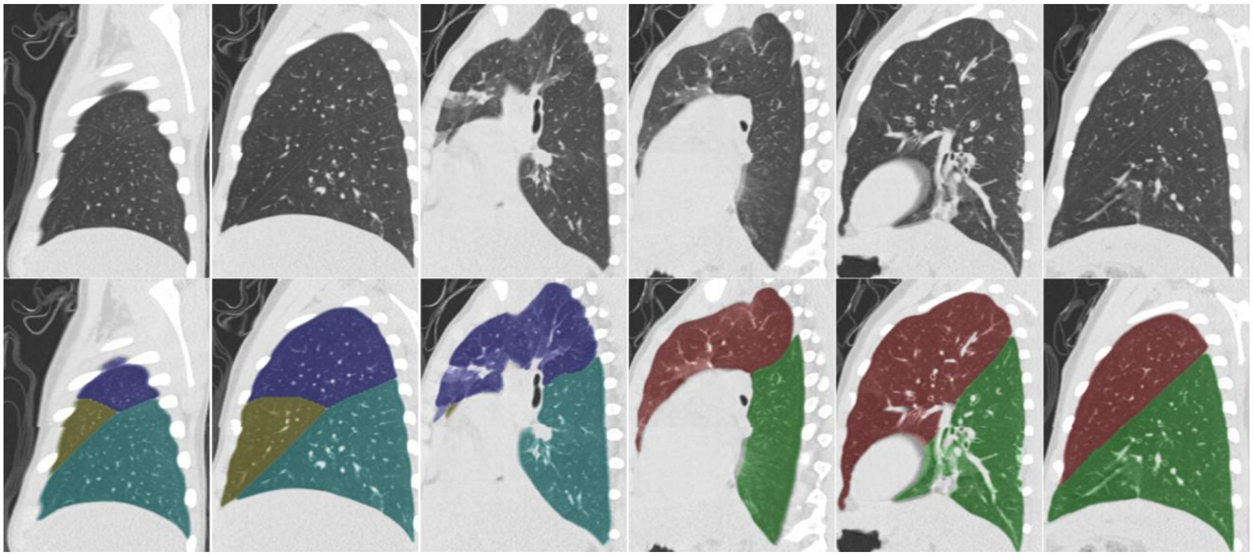


Figure 3.26: Lobar segmentation on a pediatric scan by the model.

The lobar segmentation model was also applied to pediatric patients in clinical practice and Figure 3.26 shows a fibrosis case at TLC.

The current lobar segmentation scheme operates on the left and right lung separately, primarily due to memory constraints. If better machines are available in the future, it will be preferable to segment the five lobes directly from the whole lung segmentation with higher or even original image resolutions. The left and right lung separation could be undertaken simultaneously replacing the sequential workflow described above.

3.3.6 Conclusion

The proposed 3D DenseNet achieved high accuracy in lobar segmentation of various pathological lungs and has been successfully applied to both adult and pediatric subjects, different breath-holds, and diagnostic and low-dose scans.

Chapter 4: Summary and Discussion

4.1 Summary

Image series from clinical trials and PACS are heterogeneous, and it is critical to automatically select the optimal series that meet requirements for processing among other irrelevant or redundant ones. This dissertation addresses the need for techniques to enable quantitative analysis on chest CT for big data mining. Fully automated image labeling and segmentation methods were developed to improve accuracy and efficiency of data curation in the setting of large scale clinical research. The specific contributions of this dissertation are as follows:

1. *A fully automated high throughput image labeling pipeline for chest CT to speed up processing image series of clinical trials.* The comprehensive labeling system was able to detect anatomical coverage, scan direction, scan posture, lung coverage completeness, contrast usage and patient breath-hold.

2. *A fully automated lung and lobar segmentation approach to accurately and efficiently segment a wide spectrum of pathological lungs and individual lobes.*

Both the labeling and segmentation pipelines are currently being used in clinical practice and they substantially reduce the effort and time required from image analysts compared to previous workflows. Each sub-task is accomplished by deep learning using a large volume of data from multi-center clinical trials. Rigorous cross validation was performed during evaluation to ensure generalizability for clinical deployment. Mining with such a large number of images in training and test is not often seen in medical imaging and thus these models demonstrated impressive generalizability and robustness when applied to different types of real world cases: adult and pediatric, high and low resolutions, diagnostic and low dose, and various lung pathologies. For lung segmentation, the proposed model was able to include areas with increased attenuation caused by fibrosis and consistently exclude central airways without post-processing. For lobar segmentation, our methods alleviate the challenge of fissure detection and the separate lobes directly, which has benefits since fissures are sometimes incomplete and the presence of various lung pathologies can be misleading when attempting to detect the fissures. Furthermore, these deep learning based models enable high throughput processing of chest CT that is important for big data mining of clinical trials.

4.2 Limitation and Future work

Multiple CNNs were proposed to solve individual problems and other alternative methods, e.g. traditional machine learning approaches, were also studied. For example, our previous 3D scan down-sampling and SVM based method[154] was applied to identify anatomical coverages and contrast usage. Each scan was evenly split into k × k × k (k = 3, 5, 7, 9, 11) blocks in which average intensity was computed as a feature, i.e. similar to average-pooling used in deep learning, for classification. It achieved an accuracy over 95% for anatomic coverage identification and 80% on contrast detection. Given a relatively small dataset without annotated aorta segmentation, such conventional methods could also be an option with some compromise of accuracy. In addition, it may be feasible to apply a CNN directly to identify enhancement on the image, i.e. skipping aorta segmentation as an intermediate step. If successful, saliency maps can be used to reveal where the distinctive features are extracted.

In terms of comparing different deep learning networks, other representative and up-to-date innovative architectures were also investigated. For example, ResNet, GoogleNet, DenseNet, MobileNet[326], NASNet[327] as well as their variants with squeeze and excitation (SE)[328] modules embedded, have all been applied to the classification tasks mentioned in Chapter 2 and they all achieved comparable performance as VGG-16. One big advantage of those more recent networks is the smaller model size, which would be beneficial for deployment on a CPU or on a mobile device. Similarly, other well-known segmentation networks, e.g. deeplabv3[222], were also tested for lung segmentation and no significant difference was observed. For lobar segmentation, a 3D SE block was integrated with the DenseNet based segmentation network but there was no discernible performance gain. These experiments show that the neural network architecture to build a successful model on a specific problem is not unique and probably they all have some parameter redundancy. A shallower network with just a few convolutional layers may be sufficient to accomplish tasks such as scan direction recognition. The exploration of the best or most economical architecture, while maintaining strong performance, is beyond the scope of this dissertation and is a possible extension in future research.

For model evaluation, accuracy in a test set was used as the primary metric. However, accuracy is not very helpful, or even misleading, in many real world problems when the dataset is highly unbalanced, e.g. classification of pulmonary nodule and non-nodules. In such scenarios, other metrics, such as sensitivity, specificity, or F1 score may be more appropriate to assess the models. Analysis of individual false positive or negative cases could help to uncover deficiencies of the model in specific cohorts. Similarly, Dice coefficient and ASD were used for segmentation evaluation and it is worth exploring other metrics which may reveal more regional results besides overall performance. All models were built with large datasets from multiple clinical trials and 5-fold cross validation was applied for robustness evaluation. In practice, the collection of large datasets is challenging, especially in medical imaging and it is necessary to investigate the best fold to perform cross validation. The aim of this work was development of fully automated image labeling and segmentation to eliminate time-consuming manual intervention. In this work, the computation time was roughly measured on individual models rather than in the setting of a comprehensive pipeline for quantitative chest CT analysis. It would be desirable to measure the exact time saved for the image analysts using these deep learning models.

Although both the classification and segmentation models were built using only diagnostic dose scans, they demonstrated remarkable robustness when applied to low and ultra-low-dose scans. We could incorporate low dose data in training to make the model even more robust using the CT reconstruction pipeline in our group. The current lung segmentation model is trained using data from diffuse lung disease trials and the data set could be expanding by including cohorts with other lung abnormalities, such as nodules, tuberculosis, effusion, etc. Additionally, generative models could be used to produce more diverse images with different combinations of pathologies.

For lobar segmentation, the current model is unable to detect if there is lobar collapse, and this could possibly be resolved by collecting more atelectasis cases. Furthermore, we could perform lobar segmentation using images at the original image resolution if better hardware was available. Similarly, lung and lobar segmentation could potentially be coupled seamlessly without explicit left and right lung separation using higher resolution images. Lastly, all the classification and segmentation tasks introduced in

Chapter 2 and Chapter 3 are based on supervised learning, whereas annotation of large dataset is extremely expensive in medical imaging. As such, it is worth investigating unsupervised (no annotation) or weakly supervised (using little annotation) approaches to further improve and enable deep learning applications in medical imaging.

To sum up, the availability of large high quality datasets plays a central role in developing successful models and collection of such data sets requires multidisciplinary collaboration, including radiologists, computer scientists, image analysts, etc. Another extension of this dissertation could be to apply deep leaning techniques to subsequent quantitative image analyses on chest CT, e.g. pulmonary nodule detection and segmentation, quantification of emphysema, ground glass opacity detection and fibrosis scoring to facilitate the diagnosis and treatment of patients. The ideal pipeline using deep learning techniques would include: 1. image labeling; 2. lung and lobar segmentation; 3. quantitative analysis of pulmonary diseases; and 4. clinical reporting for physicians. This dissertation has laid the foundation for building such a high throughput system for quantitative image analysis on chest CT.

Reference

1. Kauczor H-U, Heitmann K, Heussel CP, Marwede D, Uthmann T, Thelen M. Automatic Detection and Quantification of Ground-Glass Opacities on High-Resolution CT Using Multiple Neural Networks. Am J Roentgenol. American Roentgen Ray Society; 2000;175: 1329–1334. doi:10.2214/ajr.175.5.1751329
2. Wang Z, Gu S, Leader JK, Kundu S, Tedrow JR, Sciurba FC, et al. Optimal threshold in CT quantification of emphysema. Eur Radiol. NIH Public Access; 2013;23: 975–84. doi:10.1007/s00330-012-2683-z
3. Matsuoka S, Kurihara Y, Yagihashi K, Hoshino M, Watanabe N, Nakajima Y. Quantitative Assessment of Air Trapping in Chronic Obstructive Pulmonary Disease Using Inspiratory and Expiratory Volumetric MDCT. Am J Roentgenol. American Roentgen Ray Society; 2008;190: 762–769. doi:10.2214/AJR.07.2820
4. Humphries SM, Yagihashi K, Huckleberry J, Rho B-H, Schroeder JD, Strand M, et al. Idiopathic Pulmonary Fibrosis: Data-driven Textural Analysis of Extent of Fibrosis at Baseline and 15-Month Follow-up. Radiology. Radiological Society of North America ; 2017;285: 270–278. doi:10.1148/radiol.2017161177
5. Nakagawa H, Nagatani Y, Takahashi M, Ogawa E, Tho N Van, Ryujin Y, et al. Quantitative CT analysis of honeycombing area in idiopathic pulmonary fibrosis: Correlations with pulmonary function tests. Eur J Radiol. 2016;85: 125–130. doi:10.1016/j.ejrad.2015.11.011
6. Kuzmak PM, Dayhoff RE. Minimizing Digital Imaging and Communications in Medicine (DICOM) Modality Worklist patient/study selection errors. J Digit Imaging. Springer-Verlag; 2001;14: 153–157. doi:10.1007/BF03190323
7. O. G uld M, Kohnen M, Keysers D, Schubert H, Wein B, M. Lehmann T. Quality of DICOM header information for image categorization. Proc SPIE - Int Soc Opt Eng. 2002;4685. doi:10.1117/12.467017
8. Mongkolwat P, Bhalodia P, Gehl JA, Channin DS. Validating DICOM content in a remote storage model. J Digit Imaging. 2005/01/12. Springer-Verlag; 2005;18: 37–41. doi:10.1007/s10278-004-1034-2
9. Doyle TC, Lawler GA. CT features of rounded atelectasis of the lung. AJR Am J Roentgenol. American Roentgen Ray Society; 1984;143: 225–8. doi:10.2214/ajr.143.2.225
10. Woodring JH. Determining the cause of pulmonary atelectasis: a comparison of plain radiography and CT. AJR Am J Roentgenol. American Roentgen Ray Society; 1988;150: 757–63. doi:10.2214/ajr.150.4.757
11. David M, Karmrodt J, Bletz C, David S, Herweling A, Kauczor H-U, et al. Analysis of Atelectasis, Ventilated, and Hyperinflated Lung During Mechanical Ventilation by Dynamic CT. Chest. Elsevier; 2005;128: 3757–3770. doi:10.1378/CHEST.128.5.3757
12. III SGA, Sensakovic WF. Automated lung segmentation for thoracic CT: Impact on computer-aided diagnosis1. Acad Radiol. 2004;11: 1011–1021. doi:https://doi.org/10.1016/j.acra.2004.06.005

13. Brown MS, McNitt-Gray MF, Mankovich NJ, Goldin JG, Hiller J, Wilson LS, et al. Method for segmenting chest CT image data using an anatomical model: preliminary results. IEEE Trans Med Imaging. 1997;16: 828–839. doi:10.1109/42.650879

14. De Nunzio G, Tommasi E, Agrusti A, Cataldo R, De Mitri I, Favetta M, et al. Automatic Lung Segmentation in CT Images with Accurate Handling of the Hilar Region. J Digit Imaging. 2011;24: 11–27. doi:10.1007/s10278-009-9229-1

15. Sluimer I, Prokop M, Van Ginneken B. Toward automated segmentation of the pathological lung in CT. IEEE Trans Med Imaging. 2005;24: 1025–1038. doi:10.1109/TMI.2005.851757

16. Prasad MN, Brown MS, Ahmad S, Abtin F, Allen J, da Costa I, et al. Automatic Segmentation of Lung Parenchyma in the Presence of Diseases Based on Curvature of Ribs. Acad Radiol. 2008;15: 1173–1180. doi:10.1016/j.acra.2008.02.004

17. Mansoor A, Bagci U, Xu Z, Foster B, Olivier KN, Elinoff JM, et al. A Generic Approach to Pathological Lung Segmentation. IEEE Trans Med Imaging. 2014;33: 2293–2310. doi:10.1109/TMI.2014.2337057

18. Korfiatis P, Kalogeropoulou C, Karahaliou A, Kazantzi A, Skiadopoulos S, Costaridou L. Texture classification-based segmentation of lung affected by interstitial pneumonia in high-resolution CT. Med Phys. 2008;35: 5290–5302. doi:10.1118/1.3003066

19. Wang J, Li F, Li Q. Automated segmentation of lungs with severe interstitial lung disease in CT. Med Phys. 2009;36: 4592–4599. doi:10.1118/1.3222872

20. Lassen B, Kuhnigk J-M, Friman O, Krass S, Peitgen H-O. Automatic segmentation of lung lobes in ct images based on fissures, vessels, and bronchi. 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE; 2010. pp. 560–563. doi:10.1109/ISBI.2010.5490284

21. Ross JC, San José Estépar R, Kindlmann G, Díaz A, Westin C-F, Silverman EK, et al. Automatic Lung Lobe Segmentation Using Particles, Thin Plate Splines, and Maximum a Posteriori Estimation. Springer, Berlin, Heidelberg; 2010. pp. 163–171. doi:10.1007/978-3-642-15711-0_21

22. van Rikxoort EM, van Ginneken B. Automated segmentation of pulmonary structures in thoracic computed tomography scans: a review. Phys Med Biol. 2013;58: R187–R220. doi:10.1088/0031-9155/58/17/R187

23. Jiantao Pu J, Bin Zheng B, Leader JK, Fuhrman C, Knollmann F, Klym A, et al. Pulmonary Lobe Segmentation in CT Examinations Using Implicit Surface Fitting. IEEE Trans Med Imaging. 2009;28: 1986–1996. doi:10.1109/TMI.2009.2027117

24. Gu S, Zheng Q, Siegfried J, Pu J. Robust pulmonary lobe segmentation against incomplete fissures. In: van Ginneken B, Novak CL, editors. International Society for Optics and Photonics; 2012. p. 831535. doi:10.1117/12.911073

25. Xiao C, Stoel BC, Bakker ME, Peng Y, Stolk J, Staring M. Pulmonary Fissure Detection in CT Images Using a Derivative of Stick Filter. IEEE Trans Med Imaging. 2016;35: 1488–1500. doi:10.1109/TMI.2016.2517680

26. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A Survey on Deep Learning in Medical Image Analysis. 2017; Available: http://arxiv.org/abs/1702.05747

27. Shen D, Wu G, Suk H-I. Deep Learning in Medical Image Analysis. Annu Rev Biomed Eng. NIH Public Access; 2017;19: 221–248. doi:10.1146/annurev-bioeng-071516-044442

28. Suzuki K. Overview of deep learning in medical imaging. Radiol Phys Technol. 2017;10: 257–273. doi:10.1007/s12194-017-0406-5

29. Shin H-C, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Trans Med Imaging. 2016;35: 1285–1298. doi:10.1109/TMI.2016.2528162

30. Lakhani P, Sundaram B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. Radiology. Radiological Society of North America; 2017;284: 574–582. doi:10.1148/radiol.2017162326

31. Hua K-L, Hsu C-H, Hidayati SC, Cheng W-H, Chen Y-J. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. Onco Targets Ther. Dove Press; 2015;8: 2015–22. doi:10.2147/OTT.S80733

32. Gao XW, Hui R, Tian Z. Classification of CT brain images based on deep learning networks. Comput Methods Programs Biomed. Elsevier; 2017;138: 49–56. doi:10.1016/J.CMPB.2016.10.007

33. Sun W, Zheng B, Qian W. Computer aided lung cancer diagnosis with deep learning algorithms. In: Tourassi GD, Armato SG, editors. International Society for Optics and Photonics; 2016. p. 97850Z. doi:10.1117/12.2216307

34. Avendi MR, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. Med Image Anal. Elsevier; 2016;30: 108–119. doi:10.1016/J.MEDIA.2016.01.005

35. Li R, Zhang W, Suk H-I, Wang L, Li J, Shen D, et al. Deep Learning Based Imaging Data Completion for Improved Brain Disease Diagnosis. Springer, Cham; 2014. pp. 305–312. doi:10.1007/978-3-319-10443-0_39

36. Prasoon A, Petersen K, Igel C, Lauze F, Dam E, Nielsen M. Deep Feature Learning for Knee Cartilage Segmentation Using a Triplanar Convolutional Neural Network. Springer, Berlin, Heidelberg; 2013. pp. 246–253. doi:10.1007/978-3-642-40763-5_31

37. Yamins DLK, DiCarlo JJ. Using goal-driven deep learning models to understand sensory cortex. Nat Neurosci. Nature Publishing Group; 2016;19: 356–365. doi:10.1038/nn.4244

38. Plis SM, Hjelm DR, Slakhutdinov R, Allen EA, Bockholt HJ, Long JD, et al. Deep learning for neuroimaging: A validation study. Front Neurosci. 2014; doi:10.3389/fnins.2014.00229

39. Cheng JZ, Ni D, Chou YH, Qin J, Tiu CM, Chang YC, et al. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. Sci Rep. nature.com; 2016; Available: https://www.nature.com/articles/srep24454

40. Milletari F, Ahmadi SA, Kroll C, Plate A, ... Hough-CNN: deep learning for segmentation of deep brain regions in MRI and ultrasound. Comput Vis …. Elsevier; 2017; Available: https://www.sciencedirect.com/science/article/pii/S1077314217300620

41. Shi J, Zhou S, Liu X, Zhang Q, Lu M, Wang T. Stacked deep polynomial network based representation learning for tumor classification with small ultrasound image dataset.

Neurocomputing. Elsevier; 2016; Available: https://www.sciencedirect.com/science/article/pii/S0925231216002344

42. Carneiro G, Nascimento JC. Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data. IEEE Trans pattern …. ieeexplore.ieee.org; 2013; Available: https://ieeexplore.ieee.org/abstract/document/6517436/

43. Han S, Kang HK, Jeong JY, Park MH, ... A deep learning framework for supporting the classification of breast lesions in ultrasound images. Phys Med …. iopscience.iop.org; 2017; Available: https://iopscience.iop.org/article/10.1088/1361-6560/aa82ec/meta

44. Suk H Il, Shen D. Deep learning-based feature representation for AD/MCI classification. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2013. pp. 583–590. doi:10.1007/978-3-642-40763-5_72

45. Suk H-I, Lee S-W, Shen D. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. Neuroimage. Academic Press; 2014;101: 569–582. doi:10.1016/J.NEUROIMAGE.2014.06.077

46. Vieira S, Pinaya WHL, Mechelli A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. Neurosci Biobehav Rev. Pergamon; 2017;74: 58–75. doi:10.1016/J.NEUBIOREV.2017.01.002

47. Ortiz A, Munilla J, Górriz JM, Ramírez J. Ensembles of Deep Learning Architectures for the Early Diagnosis of the Alzheimer's Disease. Int J Neural Syst. World Scientific Publishing Company ; 2016;26: 1650025. doi:10.1142/S0129065716500258

48. Liu S, Liu S, Cai W, Pujol S, Kikinis R, Feng D. Early diagnosis of Alzheimer's disease with deep learning. 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI). IEEE; 2014. pp. 1015–1018. doi:10.1109/ISBI.2014.6868045

49. Cruz-Roa A, Gilmore H, Basavanhally A, Feldman M, Ganesan S, Shih NNC, et al. Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent. Sci Rep. Nature Publishing Group; 2017;7: 46450. doi:10.1038/srep46450

50. Bar Y, Diamant I, Wolf L, Greenspan H. Deep learning with non-medical training used for chest pathology identification. In: Hadjiiski LM, Tourassi GD, editors. International Society for Optics and Photonics; 2015. p. 94140V. doi:10.1117/12.2083124

51. Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci Rep. Nature Publishing Group; 2016;6: 26286. doi:10.1038/srep26286

52. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. J Pathol Inform. Wolters Kluwer -- Medknow Publications; 2016;7: 29. doi:10.4103/2153-3539.186902

53. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH. Deep Learning for Identifying Metastatic Breast Cancer. 2016; Available: http://arxiv.org/abs/1606.05718

54. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. Nature Publishing Group; 2015;521: 436–444. doi:10.1038/nature14539

55. Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press; 2016.

56. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. ImageNet Large Scale Visual Recognition Challenge. Int J Comput Vis. 2015;115: 211–252. doi:10.1007/s11263-015-0816-y

57. Roth HR, Lee CT, Shin H, Seff A, Kim L, Yao J, et al. Anatomy-specific classification of medical images using deep convolutional nets. 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI). 2015. pp. 101–104. doi:10.1109/ISBI.2015.7163826

58. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks [Internet]. Available: http://code.google.com/p/cuda-convnet/

59. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the Inception Architecture for Computer Vision. 2015; doi:10.1109/CVPR.2016.308

60. Miki Y, Muramatsu C, Hayashi T, Zhou X, Hara T, Katsumata A, et al. Classification of teeth in cone-beam CT using deep convolutional neural network. Comput Biol Med. Pergamon; 2017;80: 24–29. doi:10.1016/J.COMPBIOMED.2016.11.003

61. Shouno H, Suzuki S, Kido S. A Transfer Learning Method with Deep Convolutional Neural Network for Diffuse Lung Disease Classification. Springer, Cham; 2015. pp. 199–207. doi:10.1007/978-3-319-26532-2_22

62. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. IEEE Trans Med Imaging. 2016;35: 1207–1216. doi:10.1109/TMI.2016.2535865

63. Gao M, Bagci U, Lu L, Wu A, Buty M, Shin H-C, et al. Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks. Comput Methods Biomech Biomed Eng Imaging Vis. Taylor & Francis; 2018;6: 1–6. doi:10.1080/21681163.2015.1124249

64. Yasaka K, Akai H, Abe O, Kiryu S. Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study. Radiology. Radiological Society of North America; 2018;286: 887–896. doi:10.1148/radiol.2017170706

65. Lin JS, Ligomenides PA, Freedman MT, Mun SK. Application of artificial neural networks for reduction of false-positive detections in digital chest radiographs. Proceedings Symp Comput Appl Med Care. American Medical Informatics Association; 1993; 434–8. Available: http://www.ncbi.nlm.nih.gov/pubmed/8130511

66. Suzuki K, Armato SG, Li F, Sone S, Doi K. Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography. Med Phys. John Wiley & Sons, Ltd; 2003;30: 1602–1617. doi:10.1118/1.1580485

67. Wei W, Liu L, Loper M, Truex S, Yu L, Gursoy ME, et al. Adversarial Examples in Deep Learning: Characterization and Divergence. 2018; Available: http://arxiv.org/abs/1807.00051

68. Tan J, Huo Y, Liang Z, Li L. A COMPARISON STUDY ON THE EFFECT OF FALSE POSITIVE REDUCTION IN DEEP LEARNING BASED DETECTION FOR JUXTAPLEURAL LUNG NODULES: CNN VS DNN [Internet]. Available: http://delivery.acm.org/10.1145/3110000/3108768/a8-tan.pdf?ip=149.142.46.151&id=3108768&acc=ACTIVE

SERVICE&key=CA367851C7E3CE77.79535EF926D6BC05.4D4702B0C3E38B35.4D4702B0C3E38B35&__acm__=1551999935_f2c8649798a035b599e37753dd5a0cdc

69. Manivannan S, Li W, Zhang J, Trucco E, McKenna SJ. Structure Prediction for Gland Segmentation With Hand-Crafted and Deep Convolutional Features. IEEE Trans Med Imaging. 2018;37: 210–221. doi:10.1109/TMI.2017.2750210

70. Arindra A, Setio A, Ciompi F, Litjens G, Gerke P, Jacobs C, et al. Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks. IEEE Trans Med Imaging. 2016;35: 1160–1169. doi:10.1109/TMI.2016.2536809

71. Anirudh R, Thiagarajan JJ, Bremer T, Kim H. Lung nodule detection using 3D convolutional neural networks trained on weakly labeled data. In: Tourassi GD, Armato SG, editors. International Society for Optics and Photonics; 2016. p. 978532. doi:10.1117/12.2214876

72. Teramoto A, Fujita H, Yamamuro O, Tamaki T. Automated detection of pulmonary nodules in PET/CT images: Ensemble false-positive reduction using a convolutional neural network technique. Med Phys. John Wiley & Sons, Ltd; 2016;43: 2821–2827. doi:10.1118/1.4948498

73. Dou Q, Chen H, Yu L, Qin J, Heng P-A. Multilevel Contextual 3-D CNNs for False Positive Reduction in Pulmonary Nodule Detection. IEEE Trans Biomed Eng. 2017;64: 1558–1567. doi:10.1109/TBME.2016.2613502

74. Jin H, Li Z, Tong R, Lin L. A deep 3D residual CNN for false-positive reduction in pulmonary nodule detection. Med Phys. John Wiley & Sons, Ltd; 2018;45: 2097–2107. doi:10.1002/mp.12846

75. da Silva GLF, Valente TLA, Silva AC, de Paiva AC, Gattass M. Convolutional neural network-based PSO for lung nodule false positive reduction on CT images. Comput Methods Programs Biomed. Elsevier; 2018;162: 109–118. doi:10.1016/J.CMPB.2018.05.006

76. LUNA16 - Data [Internet]. [cited 7 Mar 2019]. Available: https://luna16.grand-challenge.org/data/

77. Armato SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, 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. 2011;38: 915–931. doi:10.1118/1.3528204

78. Kumar D, Wong A, Clausi DA. Lung Nodule Classification Using Deep Features in CT Images. 2015 12th Conference on Computer and Robot Vision. IEEE; 2015. pp. 133–138. doi:10.1109/CRV.2015.25

79. Shen W, Zhou M, Yang F, Yang C, Tian J. Multi-scale Convolutional Neural Networks for Lung Nodule Classification. Springer, Cham; 2015. pp. 588–599. doi:10.1007/978-3-319-19992-4_46

80. Shen W, Zhou M, Yang F, Yu D, Dong D, Yang C, et al. Multi-crop Convolutional Neural Networks for lung nodule malignancy suspiciousness classification. Pattern Recognit. Pergamon; 2017;61: 663–673. doi:10.1016/J.PATCOG.2016.05.029

81. Yang H, Yu H, Wang G. Deep Learning for the Classification of Lung Nodules. 2016; Available: http://arxiv.org/abs/1611.06651

82. Li W, Cao P, Zhao D, Wang J. Pulmonary Nodule Classification with Deep Convolutional Neural

Networks on Computed Tomography Images. Comput Math Methods Med. Hindawi; 2016;2016: 1–7. doi:10.1155/2016/6215085

83. Song Q, Zhao L, Luo X, Dou X. Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images. J Healthc Eng. Hindawi; 2017;2017: 1–7. doi:10.1155/2017/8314740

84. Zhu W, Liu C, Fan W, Xie X. DeepLung: Deep 3D Dual Path Nets for Automated Pulmonary Nodule Detection and Classification. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE; 2018. pp. 673–681. doi:10.1109/WACV.2018.00079

85. de Vos BD, Wolterink JM, de Jong PA, Viergever MA, Išgum I. 2D image classification for 3D anatomy localization: employing deep convolutional neural networks. In: Styner MA, Angelini ED, editors. International Society for Optics and Photonics; 2016. p. 97841Y. doi:10.1117/12.2216971

86. Roth HR, Lu L, Seff A, Cherry KM, Hoffman J, Wang S, et al. A New 2.5D Representation for Lymph Node Detection Using Random Sets of Deep Convolutional Neural Network Observations. Springer, Cham; 2014. pp. 520–527. doi:10.1007/978-3-319-10404-1_65

87. Yang D, Zhang S, Yan Z, Tan C, Li K, Metaxas D. Automated anatomical landmark detection ondistal femur surface using convolutional neural network. 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI). IEEE; 2015. pp. 17–21. doi:10.1109/ISBI.2015.7163806

88. Payer C, Štern D, Bischof H, Urschler M. Regressing Heatmaps for Multiple Landmark Localization Using CNNs. Springer, Cham; 2016. pp. 230–238. doi:10.1007/978-3-319-46723-8_27

89. Suzani A, Seitel A, Liu Y, Fels S, Rohling RN, Abolmaesumi P. Fast Automatic Vertebrae Detection and Localization in Pathological CT Scans - A Deep Learning Approach. Springer, Cham; 2015. pp. 678–686. doi:10.1007/978-3-319-24574-4_81

90. Chen H, Shen C, Qin J, Ni D, Shi L, Cheng JCY, et al. Automatic Localization and Identification of Vertebrae in Spine CT via a Joint Learning Model with Deep Neural Networks. Springer, Cham; 2015. pp. 515–522. doi:10.1007/978-3-319-24553-9_63

91. Zheng Y, Liu D, Georgescu B, Nguyen H, Comaniciu D. 3D Deep Learning for Efficient and Robust Landmark Detection in Volumetric Data. Springer, Cham; 2015. pp. 565–572. doi:10.1007/978-3-319-24553-9_69

92. Grewal M, Srivastava MM, Kumar P, Varadarajan S. RADnet: Radiologist level accuracy using deep learning for hemorrhage detection in CT scans. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE; 2018. pp. 281–284. doi:10.1109/ISBI.2018.8363574

93. Jiang H, Ma H, Qian W, Gao M, Li Y. An Automatic Detection System of Lung Nodule Based on Multigroup Patch-Based Deep Learning Network. IEEE J Biomed Heal Informatics. 2018;22: 1227–1237. doi:10.1109/JBHI.2017.2725903

94. Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, LeCun Y. OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks. 2013; Available: http://arxiv.org/abs/1312.6229

95. van Ginneken B, Setio AAA, Jacobs C, Ciompi F. Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans. 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI). IEEE; 2015. pp. 286–289.

doi:10.1109/ISBI.2015.7163869

96. Ding J, Li A, Hu Z, Wang L. Accurate Pulmonary Nodule Detection in Computed Tomography Images Using Deep Convolutional Neural Networks. Springer, Cham; 2017. pp. 559–567. doi:10.1007/978-3-319-66179-7_64

97. Golan R, Jacob C, Denzinger J. Lung nodule detection in CT images using deep convolutional neural networks. 2016 International Joint Conference on Neural Networks (IJCNN). IEEE; 2016. pp. 243–250. doi:10.1109/IJCNN.2016.7727205

98. Hamidian S, Sahiner B, Petrick N, Pezeshk A. 3D convolutional neural network for automatic detection of lung nodules in chest CT. In: Armato SG, Petrick NA, editors. International Society for Optics and Photonics; 2017. p. 1013409. doi:10.1117/12.2255795

99. Gu Y, Lu X, Yang L, Zhang B, Yu D, Zhao Y, et al. Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs. Comput Biol Med. 2018;103: 220–231. doi:10.1016/j.compbiomed.2018.10.011

100. Dou Q, Chen H, Jin Y, Lin H, Qin J, Heng P-A. Automated Pulmonary Nodule Detection via 3D ConvNets with Online Sample Filtering and Hybrid-Loss Residual Learning. Springer, Cham; 2017. pp. 630–638. doi:10.1007/978-3-319-66179-7_72

101. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2015. pp. 3431–3440. doi:10.1109/CVPR.2015.7298965

102. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Med Image Comput Comput Interv -- MICCAI 2015. 2015; 234–241. doi:10.1007/978-3-319-24574-4_28

103. Dong H, Yang G, Liu F, Mo Y, Guo Y. Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks. Springer, Cham; 2017. pp. 506–517. doi:10.1007/978-3-319-60964-5_44

104. Norman B, Pedoia V, Majumdar S. Use of 2D U-Net Convolutional Neural Networks for Automated Cartilage and Meniscus Segmentation of Knee MR Imaging Data to Determine Relaxometry and Morphometry. Radiology. Radiological Society of North America ; 2018;288: 177–185. doi:10.1148/radiol.2018172322

105. Lozej J, Meden B, Struc V, Peer P. End-to-End Iris Segmentation Using U-Net. 2018 IEEE International Work Conference on Bioinspired Intelligence (IWOBI). IEEE; 2018. pp. 1–6. doi:10.1109/IWOBI.2018.8464213

106. He D, Cohen S, Price B, Kifer D, Giles CL. Multi-scale Multi-task FCN for Semantic Page Segmentation and Table Detection. 2017; doi:10.1109/ICDAR.2017.50

107. Zhou Q, Yang W, Gao G, Ou W, Lu H, Chen J, et al. Multi-scale deep context convolutional neural networks for semantic segmentation. doi:10.1007/s11280-018-0556-3

108. Kamnitsas K, Chen L, Ledig C, Rueckert D, Glocker B. Multi-scale 3D convolutional neural networks for lesion segmentation in brain MRI [Internet]. 2015. Available: https://www.semanticscholar.org/paper/Multi-scale-3D-convolutional-neural-networks-for-in-Kamnitsas-Chen/465dedb91349eda9a33b3b3b96ff3036b337695e

109. Bao S, Chung ACS. Multi-scale structured CNN with label consistency for brain MR image segmentation. Comput Methods Biomech Biomed Eng Imaging Vis. Taylor & Francis; 2018;6: 113–117. doi:10.1080/21681163.2016.1182072

110. Li J, Sarma K V, Chung Ho K, Gertych A, Knudsen BS, Arnold CW. A Multi-scale U-Net for Semantic Segmentation of Histological Images from Radical Prostatectomies. AMIA . Annu Symp proceedings AMIA Symp. American Medical Informatics Association; 2017;2017: 1140–1148. Available: http://www.ncbi.nlm.nih.gov/pubmed/29854182

111. Çiçek Ö, Abdulkadir A, Lienkamp SS, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation ¨. arXiv Prepr arXiv160606650. 2016; doi:10.1007/978-3-319-46723-8_49

112. Milletari F, Navab N, Ahmadi S-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. arXiv Prepr arXiv160604797. 2016; 1–11. doi:10.1109/3DV.2016.79

113. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. CoRR. 2015;abs/1512.0. Available: http://arxiv.org/abs/1512.03385

114. Drozdzal M, Vorontsov E, Chartrand G, Kadoury S, Pal C. The Importance of Skip Connections in Biomedical Image Segmentation. Springer, Cham; 2016. pp. 179–187. doi:10.1007/978-3-319-46976-8_19

115. Chen H, Dou Q, Yu L, Heng P-A. VoxResNet: Deep Voxelwise Residual Networks for Volumetric Brain Segmentation. 2016; Available: http://arxiv.org/abs/1608.05895

116. Yu L, Yang X, Chen H, Qin J, Heng PA. Volumetric ConvNets with Mixed Residual Connections for Automated Prostate Segmentation from 3D MR Images. Thirty-First AAAI Conf Artif Intell. 2017; Available: https://www.aaai.org/ocs/index.php/AAAI/AAAI17/paper/viewPaper/14719

117. Quan TM, Hildebrand DGC, Jeong W-K. FusionNet: A deep fully residual convolutional neural network for image segmentation in connectomics. 2016; Available: http://arxiv.org/abs/1612.05360

118. Apostolopoulos S, De Zanet S, Ciller C, Wolf S, Sznitman R. Pathological OCT Retinal Layer Segmentation Using Branch Residual U-Shape Networks. Springer, Cham; 2017. pp. 294–301. doi:10.1007/978-3-319-66179-7_34

119. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. Proc - 30th IEEE Conf Comput Vis Pattern Recognition, CVPR 2017. 2017;2017–Janua: 2261–2269. doi:10.1109/CVPR.2017.243

120. Jegou S, Drozdzal M, Vazquez D, Romero A, Bengio Y. The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation [Internet]. 2017. pp. 11–19. Available: http://openaccess.thecvf.com/content_cvpr_2017_workshops/w13/html/Jegou_The_One_Hundred_CVPR_2017_paper.html

121. Shakeri M, Tsogkas S, Ferrante E, Lippe S, Kadoury S, Paragios N, et al. Sub-cortical brain structure segmentation using F-CNN'S. 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). IEEE; 2016. pp. 269–272. doi:10.1109/ISBI.2016.7493261

122. Alansary A, Kamnitsas K, Davidson A, Khlebnikov R, Rajchl M, Malamateniou C, et al. Fast Fully Automatic Segmentation of the Human Placenta from Motion Corrupted MRI. Springer, Cham;

2016. pp. 589–597. doi:10.1007/978-3-319-46723-8_68

123. Luo Y, Yang L, Wang L, Cheng H. Efficient CNN-CRF Network for Retinal Image Segmentation. Springer, Singapore; 2017. pp. 157–165. doi:10.1007/978-981-10-5230-9_17

124. Fu H, Xu Y, Lin S, Kee Wong DW, Liu J. DeepVessel: Retinal Vessel Segmentation via Deep Learning and Conditional Random Field. Springer, Cham; 2016. pp. 132–139. doi:10.1007/978-3-319-46723-8_16

125. Liu Y, Zhang P, Song Q, Li A, Zhang P, Gui Z. Automatic Segmentation of Cervical Nuclei Based on Deep Learning and a Conditional Random Field. IEEE Access. 2018;6: 53709–53721. doi:10.1109/ACCESS.2018.2871153

126. Shen H, Zhang J. Fully connected CRF with data-driven prior for multi-class brain tumor segmentation. 2017 IEEE International Conference on Image Processing (ICIP). IEEE; 2017. pp. 1727–1731. doi:10.1109/ICIP.2017.8296577

127. Kamnitsas K, Ledig C, Newcombe VFJ, Simpson JP, Kane AD, Menon DK, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med Image Anal. Elsevier; 2017;36: 61–78. doi:10.1016/J.MEDIA.2016.10.004

128. Wang S, Zhou M, Gevaert O, Tang Z, Dong D, Liu Z, et al. A multi-view deep convolutional neural networks for lung nodule segmentation. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2017. pp. 1752–1755. doi:10.1109/EMBC.2017.8037182

129. Bronmans B, Supervisor EH. Lung Nodule Segmentation Using 3D Convolutional Neural Networks [Internet]. 2018. Available: https://beta.vu.nl/nl/Images/werkstuk-bronmans_tcm235-875352.pdf

130. Nie Y, Zhuo D, Song G, Wen S. Pulmonary Nodule Segmentation Method of CT Images Based on 3D-FCN. Springer, Cham; 2018. pp. 134–141. doi:10.1007/978-3-030-01298-4_13

131. Liu M, Dong J, Dong X, Yu H, Qi L. Segmentation of Lung Nodule in CT Images Based on Mask R-CNN. 2018 9th International Conference on Awareness Science and Technology (iCAST). IEEE; 2018. pp. 1–6. doi:10.1109/ICAwST.2018.8517248

132. Xie Y, Zhang Z, Sapkota M, Yang L. Spatial Clockwork Recurrent Neural Network for Muscle Perimysium Segmentation [Internet]. Available: https://www.cise.ufl.edu/~zizhao/paper_list/paper_miccai_1144.pdf

133. Andermatt S, Pezold S, Cattin P. Multi-dimensional Gated Recurrent Units for the Segmentation of Biomedical 3D-Data. Springer, Cham; 2016. pp. 142–151. doi:10.1007/978-3-319-46976-8_15

134. Goodfellow I, Pouget-Abadie J, Mirza M. Generative Adversarial Networks. arXiv Prepr arXiv …. 2014; 1–9. Available: http://arxiv.org/abs/1406.2661

135. Xue Y, Xu T, Zhang H, Long R, Huang X. SegAN: Adversarial Network with Multi-scale L_1 Loss for Medical Image Segmentation. 2017; doi:10.1007/s12021-018-9377-x

136. Odena A, Olah C, Shlens J. Conditional Image Synthesis With Auxiliary Classifier GANs. 2016; 1–16. Available: http://arxiv.org/abs/1610.09585

137. Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-Image Translation with Conditional Adversarial

Networks. 2016; Available: http://arxiv.org/abs/1611.07004

138. Li Y, Shen L. cC-GAN: A Robust Transfer-Learning Framework for HEp-2 Specimen Image Segmentation. IEEE Access. 2018;6: 14048–14058. doi:10.1109/ACCESS.2018.2808938

139. Krishnankutty B, Bellary S, Kumar NBR, Moodahadu LS. Data management in clinical research: An overview. Indian J Pharmacol. Medknow Publications & Media Pvt Ltd; 2012;44: 168–172. doi:10.4103/0253-7613.93842

140. Gaddale JR. Clinical Data Acquisition Standards Harmonization importance and benefits in clinical data management. Perspect Clin Res. Medknow Publications & Media Pvt Ltd; 2015;6: 179–183. doi:10.4103/2229-3485.167101

141. Houston L, Probst Y, Yu P, Martin A. Exploring Data Quality Management within Clinical Trials. Appl Clin Inform. 2018/01/31. Schattauer GmbH; 2018;9: 72–81. doi:10.1055/s-0037-1621702

142. Ducar C, Smith D, Pinzon C, Stirewalt M, Cooper C, McElrath MJ, et al. Benefits of a comprehensive quality program for cryopreserved PBMC covering 28 clinical trials sites utilizing an integrated, analytical web-based portal. J Immunol Methods. 2014/04/04. 2014;409: 9–20. doi:10.1016/j.jim.2014.03.024

143. Selvarani AG, Annadurai S. Medical Image Retrieval by Combining Low Level Features and DICOM Features. International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007). 2007. pp. 587–589. doi:10.1109/ICCIMA.2007.336

144. Lituiev DS, Trivedi H, Panahiazar M, Norgeot B, Seo Y, Franc B, et al. Automatic Labeling of Special Diagnostic Mammography Views from Images and DICOM Headers. Journal of Digital Imaging; 2018;

145. Gong T, Liu R, Tan CL, Farzad N, Lee CK, Pang BC, et al. Classification of CT Brain Images of Head Trauma. Pattern Recognition in Bioinformatics. Berlin, Heidelberg: Springer Berlin Heidelberg; 2007. pp. 401–408. doi:10.1007/978-3-540-75286-8_38

146. Jeena RS, Kumar S. A comparative analysis of MRI and CT brain images for stroke diagnosis. 2013 Annual International Conference on Emerging Research Areas and 2013 International Conference on Microelectronics, Communications and Renewable Energy. IEEE; 2013. pp. 1–5. doi:10.1109/AICERA-ICMiCR.2013.6575935

147. Gore RM, Pickhardt PJ, Mortele KJ, Fishman EK, Horowitz JM, Fimmel CJ, et al. Management of Incidental Liver Lesions on CT: A White Paper of the ACR Incidental Findings Committee. J Am Coll Radiol. Elsevier; 2017;14: 1429–1437. doi:10.1016/j.jacr.2017.07.018

148. Scalise P, Mantarro A, Pancrazi F, Neri E. Computed tomography colonography for the practicing radiologist: A review of current recommendations on methodology and clinical indications. World J Radiol. Baishideng Publishing Group Inc; 2016;8: 472–83. doi:10.4329/wjr.v8.i5.472

149. Vancauwenberghe T, Snoeckx A, Vanbeckevoort D, Dymarkowski S, Vanhoenacker FM. Imaging of the spleen: what the clinician needs to know. Singapore Med J. Singapore Medical Association; 2015;56: 133–44. doi:10.11622/SMEDJ.2015040

150. Fenchel M, Thesen S, Schilling A. Automatic Labeling of Anatomical Structures in MR FastView Images Using a Statistical Atlas. Springer, Berlin, Heidelberg; 2008. pp. 576–584. doi:10.1007/978-3-540-85988-8_69

151. SHIMIZU, A. Multi-organ Segmentation in Three Dimensional Abdominal CT Images. Proc Comput Assist Radiol Surgery, 2006. 2006; Available: https://ci.nii.ac.jp/naid/10019160403/

152. Bartels-Rutten A, Prokop M, Išgum I, Staring M, Rutten A, Viergever MA, et al. Multi-Atlas-Based Segmentation With Local Decision Fusion-Application to Cardiac and Aortic Segmentation in CT Scans. IEEE Trans Med Imaging. 2009;28. doi:10.1109/TMI.2008.2011480

153. Criminisi A, Robertson D, Konukoglu E, Shotton J, Pathak S, White S, et al. Regression forests for efficient anatomy detection and localization in computed tomography scans. Med Image Anal. Elsevier; 2013;17: 1293–1303. doi:10.1016/J.MEDIA.2013.01.001

154. Wang X, Lo P, Ramakrishna B, Goldin J, Brown M. A machine learning approach for classification of anatomical coverage in CT. In: Tourassi GD, Armato SG, editors. International Society for Optics and Photonics; 2016. p. 97853P. doi:10.1117/12.2216259

155. Sugimori H. Classification of Computed Tomography Images in Different Slice Positions Using Deep Learning. J Healthc Eng. Hindawi; 2018;2018: 1–9. doi:10.1155/2018/1753480

156. Yan Z, Zhan Y, Peng Z, Liao S, Shinagawa Y, Metaxas DN, et al. Bodypart Recognition Using Multi-stage Deep Learning. Inf Process Med Imaging. 2015;24: 449–61. Available: http://www.ncbi.nlm.nih.gov/pubmed/26221694

157. Humpire-Mamani GE, Setio AAA, van Ginneken B, Jacobs C. Efficient organ localization using multi-label convolutional neural networks in thorax-abdomen CT scans. Phys Med Biol. IOP Publishing; 2018;63: 085003. doi:10.1088/1361-6560/aab4b3

158. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. ImageNet Chall. 2014; 1–10. doi:10.1016/j.infsof.2008.09.005

159. LOLA11 - Home [Internet]. [cited 5 Feb 2019]. Available: https://lola11.grand-challenge.org/

160. Dodge S, Karam L. Understanding How Image Quality Affects Deep Neural Networks [Internet]. Available: https://arxiv.org/pdf/1604.04004.pdf

161. Goodfellow IJ, Shlens J, Szegedy C. EXPLAINING AND HARNESSING ADVERSARIAL EXAMPLES [Internet]. Available: https://github.com/lisa-lab/pylearn2/tree/master/pylearn2/scripts/

162. Hoffman J, Young S, Noo F, McNitt-Gray M. Technical Note: FreeCT_wFBP: A robust, efficient, open-source implementation of weighted filtered backprojection for helical, fan-beam CT. Med Phys. 2016;43: 1411–1420. doi:10.1118/1.4941953

163. Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks [Internet]. Available: https://arxiv.org/pdf/1311.2901.pdf

164. Simonyan K, Vedaldi A, Zisserman A. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps [Internet]. Available: http://code.google.com/p/cuda-convnet/

165. He S, Pugeault N. Deep saliency: What is learnt by a deep network about saliency? 2018; Available: http://arxiv.org/abs/1801.04261

166. Schmidt D, Odland R. Mirror-Image Reversal of Coronal Computed Tomography Scans. Laryngoscope. John Wiley & Sons, Ltd; 2004;114: 1562–1565. doi:10.1097/00005537-200409000-00011

167. Beyea SC. Ensuring correct site surgery. AORN J. John Wiley & Sons, Ltd; 2002;76: 880–882. doi:10.1016/S0001-2092(06)61044-2

168. Christensen JD, Hutchins GC, McDonald CJ. Computer automated detection of head orientation for prevention of wrong-side treatment errors. AMIA . Annu Symp proceedings AMIA Symp. American Medical Informatics Association; 2006;2006: 136–140. Available: https://www.ncbi.nlm.nih.gov/pubmed/17238318

169. Lee YH, Yang J, Suh J-S. Detection and Correction of Laterality Errors in Radiology Reports. J Digit Imaging. 2015/02/25. Springer US; 2015;28: 412–416. doi:10.1007/s10278-015-9772-x

170. Kozo Morimoto AK. A Method for Visual Scoring of Pulmonary Mycobacterium Avium Complex Disease: ?NICE Scoring System? Mycobact Dis. OMICS International; 2013;03: 1–5. doi:10.4172/2161-1068.1000127

171. Goldin JG, McNitt-Gray MF, Sorenson SM, Johnson TD, Dauphinee B, Kleerup EC, et al. Airway hyperreactivity: assessment with helical thin-section CT. Radiology. 1998;208: 321–329. doi:10.1148/radiology.208.2.9680554

172. Kim HJ, Tashkin DP, Clements P, Li G, Brown MS, Elashoff R, et al. A computer-aided diagnosis system for quantitative scoring of extent of lung fibrosis in scleroderma patients. Clin Exp Rheumatol. 2010;28: S26–S35. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3177564/

173. Verschakelen JA, Van fraeyenhoven L, Laureys G, Demedts M, Baert AL. Differences in CT density between dependent and nondependent portions of the lung: influence of lung volume. AJR Am J Roentgenol. American Public Health Association; 1993;161: 713–7. doi:10.2214/ajr.161.4.8372744

174. Mayo JR. CT Evaluation of Diffuse Infiltrative Lung Disease. J Thorac Imaging. 2009;24: 252–259. doi:10.1097/RTI.0b013e3181c227b2

175. Verschakelen JA. The role of high-resolution computed tomography in the work-up of interstitial lung disease. Curr Opin Pulm Med. 2010;16: 503–510. doi:10.1097/MCP.0b013e32833cc997

176. Gotway MB, Freemer MM, King Jr TE. Challenges in pulmonary fibrosis. 1: Use of high resolution CT scanning of the lung for the evaluation of patients with idiopathic interstitial pneumonias. Thorax. BMJ Group; 2007;62: 546–553. doi:10.1136/thx.2004.040022

177. Lee K-N, Yoon SK, Sohn CH, Choi PJ, Webb WR. Dependent Lung Opacity at Thin-Section CT: Evaluation by Spirometrically-Gated CT of the Influence of Lung Volume. Korean J Radiol. 2002;3: 24. doi:10.3348/kjr.2002.3.1.24

178. Lynch DA, Gamsu G, Aberle DR. Conventional and high resolution computed tomography in the diagnosis of asbestos-related diseases. RadioGraphics. 1989;9: 523–551. doi:10.1148/radiographics.9.3.2727359

179. National Lung Screening Trial Research Team, Aberle DR, Berg CD, Black WC, Church TR, Fagerstrom RM, et al. The National Lung Screening Trial: Overview and Study Design. Radiology. 2011;258: 243–253. doi:10.1148/radiol.10091808

180. Moore SM, Gierada DS, Clark KW, Blaine GJ, Group* P-NQAW. Image Quality Assurance in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial Network of the National Lung

Screening Trial. J Digit Imaging. Springer-Verlag; 2005;18: 242–250. doi:10.1007/s10278-005-5153-1

181. Gierada DS, Garg K, Nath H, Strollo DC, Fagerstrom RM, Ford MB. CT Quality Assurance in the Lung Screening Study Component of the National Lung Screening Trial: Implications for Multicenter Imaging Trials. Am J Roentgenol. American Roentgen Ray Society; 2009;193: 419–424. doi:10.2214/AJR.08.1995

182. Clark KW, Gierada DS, Marquez G, Moore SM, Maffitt DR, Moulton JD, et al. Collecting 48,000 CT Exams for the Lung Screening Study of the National Lung Screening Trial. J Digit Imaging. Springer-Verlag; 2009;22: 667–680. doi:10.1007/s10278-008-9145-9

183. Mittal MK, Malik A, Sureka B, Thukral BB. Cystic masses of neck: A pictorial review. Indian J Radiol Imaging. Wolters Kluwer -- Medknow Publications; 2012;22: 334–43. doi:10.4103/0971-3026.111488

184. Karandikar A, Gummalla KM, Loke SC, Goh J, Tan TY. Approach to intensely enhancing neck nodes. Diagn Interv Radiol. Turkish Society of Radiology; 2016;22: 168–72. doi:10.5152/dir.2015.14561

185. Harders SW, Madsen HH, Nellemann HM, Rasmussen TR, Thygesen J, Hager H, et al. Dynamic contrast-enhanced CT in suspected lung cancer: quantitative results. Br J Radiol. British Institute of Radiology; 2013;86: 20130257. doi:10.1259/bjr.20130257

186. García-Garrigós E, Arenas-Jiménez JJ, Sánchez-Payá J. Best Protocol for Combined Contrast-Enhanced Thoracic and Abdominal CT for Lung Cancer: A Single-Institution Randomized Crossover Clinical Trial. Am J Roentgenol. American Roentgen Ray Society ; 2018;210: 1226–1234. doi:10.2214/AJR.17.19185

187. Subhawong TK, Fishman EK, Swart JE, Carrino JA, Attar S, Fayad LM. Soft-tissue masses and masslike conditions: what does CT add to diagnosis and management? AJR Am J Roentgenol. NIH Public Access; 2010;194: 1559–67. doi:10.2214/AJR.09.3736

188. Razek AA, Huang BY. Soft Tissue Tumors of the Head and Neck: Imaging-based Review of the WHO Classification. RadioGraphics. Radiological Society of North America; 2011;31: 1923–1954. doi:10.1148/rg.317115095

189. Brown RE, Harave S. Diagnostic imaging of benign and malignant neck masses in children-a pictorial review. Quant Imaging Med Surg. AME Publications; 2016;6: 591–604. doi:10.21037/qims.2016.10.10

190. Gezer NS, Başara I, Altay C, Harman M, Rocher L, Karabulut N, et al. Abdominal sarcoidosis: cross-sectional imaging findings. Diagn Interv Radiol. Turkish Society of Radiology; 2015;21: 111–7. doi:10.5152/dir.2014.14210

191. Alpert JS, Smith R, Carlson CJ, Ockene IS, Dexter L, Dalen JE. Mortality in Patients Treated for Pulmonary Embolism. JAMA J Am Med Assoc. American Medical Association; 1976;236: 1477. doi:10.1001/jama.1976.03270140029017

192. Stein PD, Terrin ML, Hales CA, Palevsky HI, Saltzman HA, Thompson BT, et al. Clinical, Laboratory, Roentgenographic, and Electrocardiographic Findings in Patients with Acute Pulmonary Embolism and No Pre-Existing Cardiac or Pulmonary Disease. Chest. Elsevier; 1991;100: 598–603. doi:10.1378/CHEST.100.3.598

193. Stein PD, Willis PW, DeMets DL, Bell W, Blackmon JR, Genton E, et al. History and physical examination in acute pulmonary embolism in patients without preexisting cardiac or pulmonary disease. Am J Cardiol. Excerpta Medica; 1981;47: 218–223. doi:10.1016/0002-9149(81)90388-X

194. Curkendall SM, deLuise C, Jones JK, Lanes S, Stang MR, Goehring E, et al. Cardiovascular Disease in Patients with Chronic Obstructive Pulmonary Disease, Saskatchewan Canada: Cardiovascular Disease in COPD Patients. Ann Epidemiol. Elsevier; 2006;16: 63–70. doi:10.1016/J.ANNEPIDEM.2005.04.008

195. Pesavento R, Piovella C, Prandoni P. Heart disease in patients with pulmonary embolism. Curr Opin Pulm Med. 2010;16: 415–418. doi:10.1097/MCP.0b013e32833b6581

196. Remy-Jardin M, Remy J, Wattinne L, Giraud F. Central pulmonary thromboembolism: diagnosis with spiral volumetric CT with the single-breath-hold technique--comparison with pulmonary angiography. Radiology. 1992;185: 381–7. doi:10.1148/radiology.185.2.1410342

197. Teigen CL, Maus TP, Sheedy PF, Stanson AW, Johnson CM, Breen JF, et al. Pulmonary embolism: diagnosis with contrast-enhanced electron-beam CT and comparison with pulmonary angiography. Radiology. 1995;194: 313–9. doi:10.1148/radiology.194.2.7824704

198. Qanadli SD, Hajjam M El, Mesurolle B, Barré O, Bruckert F, Joseph T, et al. Pulmonary Embolism Detection: Prospective Evaluation of Dual-Section Helical CT versus Selective Pulmonary Arteriography in 157 Patients. Radiology. Radiological Society of North America ; 2000;217: 447–455. doi:10.1148/radiology.217.2.r00nv01447

199. Wittram C, Maher MM, Yoo AJ, Kalra MK, Shepard J-AO, McLoud TC. CT Angiography of Pulmonary Embolism: Diagnostic Criteria and Causes of Misdiagnosis. RadioGraphics. Radiological Society of North America ; 2004;24: 1219–1238. doi:10.1148/rg.245045008

200. Hochhegger B, Rottenfusser R, Marchiori E. When is the use of contrast media in chest CT indicated? J Bras Pneumol. Sociedade Brasileira de Pneumologia e Tisiologia (Brazilian Thoracic Society); 2017;43: 400. doi:10.1590/S1806-37562017000000179

201. Bae KT. Intravenous Contrast Medium Administration and Scan Timing at CT: Considerations and Approaches. Radiology. Radiological Society of North America, Inc.; 2010;256: 32–61. doi:10.1148/radiol.10090908

202. Cademartiri F, Nieman K, van der Lugt A, Raaijmakers RH, Mollet N, Pattynama PMT, et al. Intravenous Contrast Material Administration at 16–Detector Row Helical CT Coronary Angiography: Test Bolus versus Bolus-tracking Technique. Radiology. Radiological Society of North America ; 2004;233: 817–823. doi:10.1148/radiol.2333030668

203. Swensen SJ, Brown LR, Colby T V, Weaver AL. Pulmonary nodules: CT evaluation of enhancement with iodinated contrast material. Radiology. 1995;194: 393–8. doi:10.1148/radiology.194.2.7824716

204. Swensen SJ, Viggiano RW, Midthun DE, Müller NL, Sherrick A, Yamashita K, et al. Lung Nodule Enhancement at CT: Multicenter Study. Radiology. Radiological Society of North America ; 2000;214: 73–80. doi:10.1148/radiology.214.1.r00ja1473

205. Lind JSW, Meijerink MR, Dingemans A-MC, van Kuijk C, Öllers MC, de Ruysscher D, et al. Dynamic contrast-enhanced CT in patients treated with sorafenib and erlotinib for non-small cell lung cancer: a new method of monitoring treatment? Eur Radiol. Springer-Verlag; 2010;20: 2890–

2898. doi:10.1007/s00330-010-1869-5

206. Patz EF, Erasmus JJ, Mcadams HP, Connolly JE, Marom EM, Goodman PC, et al. Lung Cancer Staging and Management: Comparison of Contrast-enhanced and Nonenhanced Helical CT of the Thorax 1 [Internet]. Radiology. 1999. Available: https://pubs.rsna.org/doi/pdf/10.1148/radiology.212.1.r99jl1956

207. M.H. Lev RGG. Hounsfield Scale - an overview | ScienceDirect Topics [Internet]. 2002. Available: https://www.sciencedirect.com/topics/medicine-and-dentistry/hounsfield-scale

208. Bibb R, Eggbeer D, Paterson A. Medical modelling : the application of advanced design and rapid prototyping techniques in medicine.

209. Honikman R, Rhee AJ. Normal Anatomy and Flow During the Complete Examination: Epiaortic Imaging. Perioper Transesophageal Echocardiogr. W.B. Saunders; 2014; 47–53. doi:10.1016/B978-1-4557-0761-4.00008-6

210. Korea S. The Dissected Aorta Part III#{149} Anatomy and Radiologic Diagnosis of Branch-Vessel Compromise' From the Departments of Radiology [Internet]. Radiology. 1997. Available: https://pubs.rsna.org/doi/pdf/10.1148/radiology.203.1.9122414

211. Behrens T, Rohr K, Stiehl HS. Robust segmentation of tubular structures in 3-D medical images by parametric object detection and tracking. IEEE Trans Syst Man Cybern Part B. 2003;33: 554–561. doi:10.1109/TSMCB.2003.814305

212. Adame IM, van der Geest RJ, Bluemke DA, Lima JAC, Reiber JHC, Lelieveldt BPF. Automatic vessel wall contour detection and quantification of wall thickness in in-vivo MR images of the human aorta. J Magn Reson Imaging. 2006;24: 595–602. doi:10.1002/jmri.20662

213. Gao Y, Zhu L, Cates J, MacLeod RS, Bouix S, Tannenbaum A. A Kalman Filtering Perspective for Multiatlas Segmentation. SIAM J Imaging Sci. NIH Public Access; 2015;8: 1007–1029. doi:10.1137/130933423

214. Kurugol S, San Jose Estepar R, Ross J, Washko GR. Aorta segmentation with a 3D level set approach and quantification of aortic calcifications in non-contrast chest CT. 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2012. pp. 2343–2346. doi:10.1109/EMBC.2012.6346433

215. Cha J won, Henn A, Stoddard M, Amini A. 3D segmentation of the ascending and descending aorta from CT data via graph-cuts. In: Gimi B, Krol A, editors. Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging. SPIE; 2018. p. 62. doi:10.1117/12.2295967

216. Kovács T, Cattin P, Alkadhi H, Wildermuth S, Székely G. Automatic Segmentation of the Vessel Lumen from 3D CTA Images of Aortic Dissection [Internet]. Available: https://link.springer.com/content/pdf/10.1007%2F3-540-32137-3_33.pdf

217. Lorigo LM, Faugeras OD, Grimson WE, Keriven R, Kikinis R, Nabavi A, et al. CURVES: curve evolution for vessel segmentation. Med Image Anal. 2001;5: 195–206. Available: http://www.ncbi.nlm.nih.gov/pubmed/11524226

218. Biesdorf A, Wörz S, Müller T, Weber TF, Heye T, Hosch W, et al. Model-based segmentation and motion analysis of the thoracic aorta from 4D ECG-gated CTA images. Med Image Comput

Comput Assist Interv. 2011;14: 589–96. Available: http://www.ncbi.nlm.nih.gov/pubmed/22003666

219. Isgum I, Staring M, Rutten A, Prokop M, Viergever MA, van Ginneken B. Multi-Atlas-Based Segmentation With Local Decision Fusion—Application to Cardiac and Aortic Segmentation in CT Scans. IEEE Trans Med Imaging. 2009;28: 1000–1010. doi:10.1109/TMI.2008.2011480

220. Chen L-C, Papandreou G, Schroff F, Adam H. Rethinking Atrous Convolution for Semantic Image Segmentation [Internet]. Available: https://arxiv.org/pdf/1706.05587.pdf

221. Noothout JMH, De Vos BD, Wolterink JM, Išgum I. Automatic Segmentation of Thoracic Aorta Segments in Low-Dose Chest CT [Internet]. Available: https://arxiv.org/pdf/1810.05727.pdf

222. Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Trans Pattern Anal Mach Intell. 2018;40: 834–848. doi:10.1109/TPAMI.2017.2699184

223. Dumoulin V, Visin F, Box GEP. A guide to convolution arithmetic for deep learning [Internet]. 2018. Available: http://ethanschoonover.com/solarized

224. Noothout J, de Vos B, Wolterink J, Isgum I. Automatic segmentation of thoracic aorta segments in low-dose chest CT. In: Angelini ED, Landman BA, editors. Medical Imaging 2018: Image Processing. SPIE; 2018. p. 63. doi:10.1117/12.2293114

225. Grydeland TB, Thorsen E, Dirksen A, Jensen R, Coxson HO, Pillai SG, et al. Quantitative CT measures of emphysema and airway wall thickness are related to DLCO. Respir Med. 2011;105: 343–351. doi:10.1016/j.rmed.2010.10.018

226. Hansell DM, Goldin JG, King TE, Lynch DA, Richeldi L, Wells AU. CT staging and monitoring of fibrotic interstitial lung diseases in clinical practice and treatment trials: a Position Paper from the Fleischner society. Lancet Respir Med. 2015;3: 483–496. doi:10.1016/S2213-2600(15)00096-X

227. Lynch DA, Sverzellati N, Travis WD, Brown KK, Colby T V, Galvin JR, et al. Diagnostic criteria for idiopathic pulmonary fibrosis: a Fleischner Society White Paper. Lancet Respir Med. 2018;6: 138–153. doi:10.1016/S2213-2600(17)30433-2

228. Goldin JG. Quantitative CT of emphysema and the airways. J Thorac Imaging. 2004;19: 235–40. Available: http://www.ncbi.nlm.nih.gov/pubmed/15502610

229. Newell JD, Sieren J, Hoffman EA, Hoffman EA. Development of quantitative computed tomography lung protocols. J Thorac Imaging. NIH Public Access; 2013;28: 266–71. doi:10.1097/RTI.0b013e31829f6796

230. Miller WT, Chatzkel J, Hewitt MG. Expiratory Air Trapping on Thoracic Computed Tomography. A Diagnostic Subclassification. Ann Am Thorac Soc. American Thoracic Society; 2014;11: 874–881. doi:10.1513/AnnalsATS.201311-390OC

231. Tashkin DP, Kim HJ, Zeidler M, Kleerup E, Goldin J, Angeles L. Evaluating small-airways disease in asthmatic patients: The utility of quantitative computed tomography. J Allergy Clin Immunol. 2017;139: 49–51.e2. doi:10.1016/j.jaci.2016.11.010

232. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and

management. Am J Respir Crit Care Med. American Thoracic Society; 2011;183: 788–824. doi:10.1164/rccm.2009-040GL

233. Arakawa H, Niimi H, Kurihara Y, Nakajima Y, Webb WR. Expiratory High-Resolution CT. Am J Roentgenol. American Roentgen Ray Society; 2000;175: 1537–1543. doi:10.2214/ajr.175.6.1751537

234. Hersh CP, Washko GR, Estépar RSJ, Lutz S, Friedman PJ, Han MK, et al. Paired inspiratory-expiratory chest CT scans to assess for small airways disease in COPD. Respir Res. BioMed Central; 2013;14: 42. doi:10.1186/1465-9921-14-42

235. Kongstad T, Buchvald FF, Green K, Lindblad A, Robinson TE, Nielsen KG. Improved air trapping evaluation in chest computed tomography in children with cystic fibrosis using real-time spirometric monitoring and biofeedback. J Cyst Fibros. Elsevier; 2013;12: 559–566. doi:10.1016/J.JCF.2013.05.012

236. Loeve M, Lequin MH, de Bruijne M, Hartmann IJC, Gerbrands K, van Straten M, et al. Cystic Fibrosis: Are Volumetric Ultra-Low-Dose Expiratory CT Scans Sufficient for Monitoring Related Lung Disease? Radiology. Radiological Society of North America, Inc.; 2009;253: 223–229. doi:10.1148/radiol.2532090306

237. Mets OM, Buckens CFM, Zanen P, Isgum I, van Ginneken B, Prokop M, et al. Identification of Chronic Obstructive Pulmonary Disease in Lung Cancer Screening Computed Tomographic Scans. JAMA. American Medical Association; 2011;306: 1775–1781. doi:10.1001/jama.2011.1531

238. Exhibit E, Argota NB, Vaamonde AG, Cervantes S, Ferrero AO, Applications-detection C. What every radiologist should know about expiratory and in the prone position chest computed tomography. 2015;

239. Gaeta M, Minutoli F, Girbino G, Murabito A, Benedetto C, Contiguglia R, et al. Expiratory CT scan in patients with normal inspiratory CT scan: a finding of obliterative bronchiolitis and other causes of bronchiolar obstruction. Multidiscip Respir Med. BioMed Central; 2013;8: 44. doi:10.1186/2049-6958-8-44

240. Park J, Jung J, Yoon SH, Goo JM, Hong H, Yoon J-H. Inspiratory Lung Expansion in Patients with Interstitial Lung Disease: CT Histogram Analyses. Sci Rep. Nature Publishing Group; 2018;8: 15265. doi:10.1038/s41598-018-33638-x

241. Schlathoelter T, Lorenz C, Carlsen IC, Renisch S, Deschamps T. Simultaneous segmentation and tree reconstruction of the airways for virtual bronchoscopy. In: Sonka M, Fitzpatrick JM, editors. International Society for Optics and Photonics; 2002. p. 103. doi:10.1117/12.467061

242. Tschirren J, Hoffman EA, McLennan G, Sonka M. Intrathoracic airway trees: segmentation and airway morphology analysis from low-dose CT scans. IEEE Trans Med Imaging. 2005;24: 1529–1539. doi:10.1109/TMI.2005.857654

243. Kitasaka T, Mori K, Suenaga Y, Hasegawa J, Toriwaki J. A Method for Segmenting Bronchial Trees from 3D Chest X-ray CT Images. Springer, Berlin, Heidelberg; 2003. pp. 603–610. doi:10.1007/978-3-540-39903-2_74

244. Mori K, Hasegawa J, Toriwaki J, Anno H, Katada K. Recognition of bronchus in three-dimensional X-ray CT images with applications to virtualized bronchoscopy system. Proceedings of 13th International Conference on Pattern Recognition. IEEE; 1996. pp. 528–532 vol.3.

doi:10.1109/ICPR.1996.547003

245. Kiraly AP, Higgins WE, McLennan G, Hoffman EA, Reinhardt JM. Three-dimensional human airway segmentation methods for clinical virtual bronchoscopy. Acad Radiol. 2002;9: 1153–68. Available: http://www.ncbi.nlm.nih.gov/pubmed/12385510

246. Gao D, Gao X, Ni C, Zhang T. MGRG-morphological gradient based 3D region growing algorithm for airway tree segmentation in image guided intervention therapy. International Symposium on Bioelectronics and Bioinformations 2011. IEEE; 2011. pp. 76–79. doi:10.1109/ISBB.2011.6107649

247. Zhu C, Qi S, van Triest H, Wang S, Kang Y, Yue Y. Automatic 3D segmentation of human airway tree in CT image. 2010 3rd International Conference on Biomedical Engineering and Informatics. IEEE; 2010. pp. 132–136. doi:10.1109/BMEI.2010.5639658

248. Rizi FY, Ahmadian A, Rezaie N, Iranmanesh SA. Leakage suppression in human airway tree segmentation using shape optimization based on fuzzy connectivity method. Int J Imaging Syst Technol. John Wiley & Sons, Ltd; 2013;23: 71–84. doi:10.1002/ima.22040

249. Singh H, Crawford M, Curtin J, Zwiggelaar R. Automated 3D Segmentation of the Lung Airway Tree Using Gain-Based Region Growing Approach. Springer, Berlin, Heidelberg; 2004. pp. 975–982. doi:10.1007/978-3-540-30136-3_118

250. Bauer C, Pock T, Bischof H, Beichel R. Airway Tree Reconstruction Based on Tube Detection [Internet]. Available: http://www.lungworkshop.org/2009/proc2009/203.pdf

251. Hofstad EF, Sorger H, Leira HO, Amundsen T, Langø T. Automatic registration of CT images to patient during the initial phase of bronchoscopy: A clinical pilot study. Med Phys. John Wiley & Sons, Ltd; 2014;41: 041903. doi:10.1118/1.4866884

252. Lai K, Zhao P, Huang Y, Liu J, Wang C, Feng H, et al. Automatic 3D Segmentation of Lung Airway Tree: A Novel Adaptive Region Growing Approach. 2009 3rd International Conference on Bioinformatics and Biomedical Engineering. IEEE; 2009. pp. 1–4. doi:10.1109/ICBBE.2009.5162427

253. Tschirren J, Hoffman EA, McLennan G, Sonka M. Intrathoracic airway trees: segmentation and airway morphology analysis from low-dose CT scans. IEEE Trans Med Imaging. NIH Public Access; 2005;24: 1529–39. doi:10.1109/TMI.2005.857654

254. Nardelli P, Khan KA, Corvò A, Moore N, Murphy MJ, Twomey M, et al. Optimizing parameters of an open-source airway segmentation algorithm using different CT images. Biomed Eng Online. BioMed Central; 2015;14: 62. doi:10.1186/s12938-015-0060-2

255. Lo P, Sporring J, Ashraf H, Pedersen JJH, de Bruijne M. Vessel-guided airway tree segmentation: A voxel classification approach. Med Image Anal. Elsevier; 2010;14: 527–538. doi:10.1016/J.MEDIA.2010.03.004

256. Rudyanto RD, Munoz-Barrutia A, Diaz AA, Ross J, Washko GR, Ortiz-de-Solorzano C, et al. Modeling airway probability. 2013 IEEE 10th International Symposium on Biomedical Imaging. IEEE; 2013. pp. 378–381. doi:10.1109/ISBI.2013.6556491

257. Charbonnier JP, Rikxoort EM va., Setio AAA, Schaefer-Prokop CM, Ginneken B van, Ciompi F. Improving airway segmentation in computed tomography using leak detection with convolutional networks. Med Image Anal. Elsevier B.V.; 2017;36: 52–60.

doi:10.1016/j.media.2016.11.001

258. Yun J, Park J, Yu D, Yi J, Lee M, Park HJ, et al. Improvement of fully automated airway segmentation on volumetric computed tomographic images using a 2.5 dimensional convolutional neural net. Med Image Anal. Elsevier; 2019;51: 13–20. doi:10.1016/J.MEDIA.2018.10.006

259. Juarez AG-U, Tiddens HAWM, de Bruijne M. Automatic Airway Segmentation in chest CT using Convolutional Neural Networks. 2018; Available: http://arxiv.org/abs/1808.04576

260. Shepard J-AO, Flores EJ, Abbott GF. Imaging of the trachea. Ann Cardiothorac Surg. AME Publications; 2018;7: 197–209. doi:10.21037/acs.2018.03.09

261. Laroia AT, Thompson BH, Laroia ST, van Beek E. Modern imaging of the tracheo-bronchial tree. World J Radiol. Baishideng Publishing Group Inc; 2010;2: 237–48. doi:10.4329/wjr.v2.i7.237

262. Marchiori E, Pozes AS, Souza Junior AS, Escuissato DL, Irion KL, Araujo Neto C de, et al. Alterações difusas da traquéia: aspectos na tomografia computadorizada. J Bras Pneumol. Sociedade Brasileira de Pneumologia e Tisiologia; 2008;34: 47–54. doi:10.1590/S1806-37132008000100009

263. Singh A, Kumar S, Mishra AK, Kumar M, Kant S, Verma SK, et al. Correlation between clinical characteristics, spirometric indices and high resolution computed tomography findings in patients of chronic obstructive pulmonary disease. Lung India. Wolters Kluwer -- Medknow Publications; 2016;33: 42–8. doi:10.4103/0970-2113.173064

264. Gallardo Estrella L, Pompe E, Kuhnigk J-M, Lynch DA, Bhatt SP, van Ginneken B, et al. Computed tomography quantification of tracheal abnormalities in COPD and their influence on airflow limitation. Med Phys. NIH Public Access; 2017;44: 3594–3603. doi:10.1002/mp.12274

265. Antonelli M, Lazzerini B, Marcelloni F. Segmentation and reconstruction of the lung volume in CT images. Proceedings of the 2005 ACM symposium on Applied computing - SAC '05. New York, New York, USA: ACM Press; 2005. p. 255. doi:10.1145/1066677.1066738

266. Armato SG, Giger ML, Moran CJ, Blackburn JT, Doi K, MacMahon H. Computerized Detection of Pulmonary Nodules on CT Scans. RadioGraphics. 1999;19: 1303–1311. doi:10.1148/radiographics.19.5.g99se181303

267. Müller NL, Staples CA, Miller RR, Abboud RT. "Density Mask": An Objective Method to Quantitate Emphysema Using Computed Tomography. Chest. 1988;94: 782–787. doi:https://doi.org/10.1378/chest.94.4.782

268. Bartholmai BJ, Raghunath S, Karwoski RA, Moua T, Rajagopalan S, Maldonado F, et al. Quantitative CT Imaging of Interstitial Lung Diseases. J Thorac Imaging. 2013;28: 10.1097/RTI.0b013e3182a21969. doi:10.1097/RTI.0b013e3182a21969

269. Mansoor A, Bagci U, Foster B, Xu Z, Papadakis GZ, Folio LR, et al. Segmentation and Image Analysis of Abnormal Lungs at CT: Current Approaches, Challenges, and Future Trends. Radiographics. 2015;35: 1056–76. doi:10.1148/rg.2015140232

270. Kauczor HU, Heussel CP, Fischer B, Klamm R, Mildenberger P, Thelen M. Assessment of lung volumes using helical CT at inspiration and expiration: comparison with pulmonary function tests. AJR Am J Roentgenol. American Public Health Association; 1998;171: 1091–5. doi:10.2214/ajr.171.4.9763003

271. Sluimer I, Schilham A, Prokop M, Van Ginneken B. Computer analysis of computed tomography scans of the lung: A survey. IEEE Trans Med Imaging. 2006;25: 385–405. doi:10.1109/TMI.2005.862753

272. JishaKrishnan, RejimolRobinson. A SURVEY ON LUNG SEGMENTATION TECHNIQUES [Internet]. Available: http://www.ijcset.com/docs/IJCSET13-04-05-083.pdf

273. Amanda AR, Widita R. Comparison of image segmentation of lungs using methods: connected threshold, neighborhood connected, and threshold level set segmentation. J Phys Conf Ser. IOP Publishing; 2016;694: 012048. doi:10.1088/1742-6596/694/1/012048

274. Elnakib A, Gimel'farb G, Suri J, El-Baz A. Chapter 1 2 Medical Image Segmentation: A Brief Survey Deformable models. doi:10.1007/978-1-4419-8204-9_1

275. Brown MS, McNitt-Gray MF, Mankovich NJ, Goldin JG, Hiller J, Wilson LS, et al. Method for segmenting chest CT image data using an anatomical model: preliminary results. IEEE Trans Med Imaging. 1997;16: 828–839. doi:10.1109/42.650879

276. Soliman A, Khalifa F, Elnakib A, El-Ghar MA, Dunlap N, Wang B, et al. Accurate Lungs Segmentation on CT Chest Images by Adaptive Appearance-Guided Shape Modeling. IEEE Trans Med Imaging. 2017;36: 263–276. doi:10.1109/TMI.2016.2606370

277. Gill G, Beichel RR. Segmentation of Lungs with Interstitial Lung Disease in CT Scans: A TV-L1 Based Texture Analysis Approach. Springer, Cham; 2014. pp. 511–520. doi:10.1007/978-3-319-14249-4_48

278. Vancheri C, Failla M, Crimi N, Raghu G. Idiopathic pulmonary fibrosis: a disease with similarities and links to cancer biology. The European respiratory journal. England; 2010. pp. 496–504. doi:10.1183/09031936.00077309

279. Lynch DA, Godwin JD, Safrin S, Starko KM, Hormel P, Brown KK, et al. High-resolution computed tomography in idiopathic pulmonary fibrosis: Diagnosis and prognosis. Am J Respir Crit Care Med. 2005;172: 488–493. doi:10.1164/rccm.200412-1756OC

280. Birkbeck N, Kohlberger T, Zhang J, Sofka M, Kaftan J, Comaniciu D, et al. Lung segmentation from CT with severe pathologies using anatomical constraints. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 2014;8673 LNCS: 804–811. doi:10.1007/978-3-319-10404-1_100

281. Sun S, Mclennan G, Hoffman E a. Model-Based Segmentation of Pathological Lungs in Volumetric CT Data. Miccai 2010. 2010;1: 31–40.

282. Song Q, Sonka M, Hoffman E a, Reinhardt JM. SEGMENTATION OF PATHOLOGICAL AND DISEASED LUNG TISSUE IN CT IMAGES USING A GRAPH-SEARCH ALGORITHM Department of Electrical and Computer Engineering , University of Iowa , Iowa City , IA , 52242 Department of Radiology , University of Iowa , Iowa City , IA. Iee. 2011; 2072–2075. doi:10.1109/ISBI.2011.5872820

283. Massoptier L, Misra A, Sowmya A. Automatic lung segmentation in HRCT images with diffuse parenchymal lung disease using graph-cut. 2009 24th International Conference Image and Vision Computing New Zealand, IVCNZ 2009 - Conference Proceedings. 2009. pp. 266–270. doi:10.1109/IVCNZ.2009.5378398

284. Dai S, Lu K, Dong J, Zhang Y, Chen Y. A Novel Approach of Lung Segmentation on Chest CT Images

Using Graph Cuts. Neurocomput. Amsterdam, The Netherlands, The Netherlands: Elsevier Science Publishers B. V.; 2015;168: 799–807. doi:10.1016/j.neucom.2015.05.044

285. Pulagam AR, Kande GB, Ede VKR, Inampudi RB. Automated Lung Segmentation from HRCT Scans with Diffuse Parenchymal Lung Diseases. J Digit Imaging. 2016;29: 507–519. doi:10.1007/s10278-016-9875-z

286. Hariharan B, Arbeláez P, Girshick R, Malik J. Hypercolumns for object segmentation and fine-grained localization. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2015. pp. 447–456. doi:10.1109/CVPR.2015.7298642

287. Badrinarayanan V, Kendall A, Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. Cvpr 2015. 2015; 5. doi:10.1103/PhysRevX.5.041024

288. Kalinovsky A, Kalinovsky A, Kovalev V. Lung Image Segmentation Using Deep Learning Methods and Convolutional Neural Networks Lung Image Segmentation Using Deep Learning Methods and Convolutional Neural Networks. 2016;

289. Harrison AP, Xu Z, George K, Lu L, Summers RM, Mollura DJ. Progressive and Multi-Path Holistically Nested Neural Networks for Pathological Lung Segmentation from CT Images. CoRR. 2017;abs/1706.0. Available: http://arxiv.org/abs/1706.03702

290. LaLonde R, Bagci U. Capsules for Object Segmentation. 2018; Available: http://arxiv.org/abs/1804.04241

291. Sabour S, Frosst N, Hinton GE. Dynamic Routing Between Capsules. 2017; Available: http://arxiv.org/abs/1710.09829

292. Zhao T, Gao D, Wang J, Tin Z. Lung segmentation in CT images using a fully convolutional neural network with multi-instance and conditional adversary loss. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE; 2018. pp. 505–509. doi:10.1109/ISBI.2018.8363626

293. Arjovsky M, Chintala S, Bottou L. Wasserstein GAN. 2017; Available: http://arxiv.org/abs/1701.07875

294. Rubner Y, Tomasi C, Guibas LJ. The Earth Mover's Distance as a Metric for Image Retrieval [Internet]. Available: https://www.cs.cmu.edu/~efros/courses/AP06/Papers/rubner-jcviu-00.pdf

295. Tan J, Jing L, Huo Y, Tian Y, Akin O. LGAN: Lung Segmentation in CT Scans Using Generative Adversarial Network [Internet]. Available: https://arxiv.org/pdf/1901.03473.pdf

296. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going Deeper with Convolutions. Computer Vision and Pattern Recognition (CVPR). 2015. Available: http://arxiv.org/abs/1409.4842

297. Sudre CH, Li W, Vercauteren T, Ourselin S, Jorge Cardoso M. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 2017;10553 LNCS: 240–248. doi:10.1007/978-3-319-67558-9_28

298. Brown MS, Goldin JG, McNitt-Gray MF, Greaser LE, Sapra A, Li K-T, et al. Knowledge-based segmentation of thoracic computed tomography images for assessment of split lung function.

Med Phys. John Wiley & Sons, Ltd; 2000;27: 592–598. doi:10.1118/1.598898

299. Brown MS, Kim HJ, Abtin F, Costa I Da, Pais R, Ahmad S, et al. Reproducibility of Lung and Lobar Volume Measurements Using Computed Tomography. Acad Radiol. 2010;17: 316–322. doi:http://doi.org/10.1016/j.acra.2009.10.005

300. Kohavi R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. Morgan Kaufmann; 1995. pp. 1137–1143.

301. Devaney AJ. A filtered backpropagation algorithm for diffraction tomography. Ultrason Imaging. No longer published by Elsevier; 1982;4: 336–350. doi:10.1016/0161-7346(82)90017-7

302. Oda K, Ishimoto H, Yatera K, Naito K, Ogoshi T, Yamasaki K, et al. High-resolution CT scoring system-based grading scale predicts the clinical outcomes in patients with idiopathic pulmonary fibrosis. Respir Res. 2014;15: 1–9. doi:10.1186/1465-9921-15-10

303. Fujimoto H, Kobayashi T, Azuma A. Idiopathic Pulmonary Fibrosis : Treatment and Prognosis. 2015;9: 179–185. doi:10.4137/CCRPM.S23321.TYPE

304. Robbie H, Daccord C, Chua F, Devaraj A. Evaluating disease severity in idiopathic pulmonary fibrosis. Eur Respir Rev. 2017;26: 1–12. doi:10.1183/16000617.0051-2017

305. George K, Harrison A ~P., Jin D, Xu Z, Mollura D ~J. Pathological Pulmonary Lobe Segmentation from CT Images using Progressive Holistically Nested Neural Networks and Random Walker. ArXiv e-prints. 2017;

306. Negahdar M, Beymer D. Automated volumetric lung segmentation of thoracic CT images using fully convolutional neural network. 2018; doi:10.1117/12.2293723

307. Ukil S, Reinhardt JM. Smoothing lung segmentation surfaces in 3D x-ray CT images using anatomic guidance. In: Fitzpatrick JM, Sonka M, editors. International Society for Optics and Photonics; 2004. p. 1066. doi:10.1117/12.536891

308. Park SC, Leader JK, Tan J, Lee GS, Kim SH, Na IS, et al. Separation of left and right lungs using 3-dimensional information of sequential computed tomography images and a guided dynamic programming algorithm. J Comput Assist Tomogr. NIH Public Access; 2011;35: 280–9. doi:10.1097/RCT.0b013e31820e4389

309. Shojaii R, Alirezaie J, Babyn P. Automatic lung segmentation in CT images using watershed transform. IEEE International Conference on Image Processing 2005. IEEE; 2005. p. II-1270. doi:10.1109/ICIP.2005.1530294

310. Hu S, Hoffman EA, Reinhardt JM. Automatic Lung Segmentation for Accurate Quantitation of Volumetric X-Ray CT Images [Internet]. IEEE TRANSACTIONS ON MEDICAL IMAGING. 2001. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.18.7030&rep=rep1&type=pdf

311. Leader JK, Zheng B, Rogers RM, Sciurba FC, Perez A, Chapman BE, et al. Automated Lung Segmentation in X-Ray Computed Tomography: Development and Evaluation of a Heuristic Threshold-Based Scheme. Acad Radiol. 2003;10: 1224–1236. doi:10.1016/S1076-6332(03)00380-5

312. Vincent L, Soille P. Watersheds in digital spaces: an efficient algorithm based on immersion

simulations. IEEE Trans Pattern Anal Mach Intell. 1991;13: 583–598. doi:10.1109/34.87344

313. Beucher S, Meyer F. The morphological approach to segmentation: The watershed transformation. Mathematical Morphology in Image Processing. 1993. pp. 433–481.

314. Roerdink J, Meijster A. The Watershed Transform: Definitions, Algorithms and Parallelization Strategies. Fundam Inf. 2003.

315. Doel T, Gavaghan DJ, Grau V. Review of automatic pulmonary lobe segmentation methods from CT. Comput Med Imaging Graph. Pergamon; 2015;40: 13–29. doi:10.1016/J.COMPMEDIMAG.2014.10.008

316. Chung JH, Lynch DA. The Value of a Multidisciplinary Approach to the Diagnosis of Usual Interstitial Pneumonitis and Idiopathic Pulmonary Fibrosis: Radiology, Pathology, and Clinical Correlation. Am J Roentgenol. American Roentgen Ray Society ; 2016;206: 463–471. doi:10.2214/AJR.15.15627

317. Leung AN. Pulmonary Tuberculosis: The Essentials. Radiology. Radiological Society of North America ; 1999;210: 307–322. doi:10.1148/radiology.210.2.r99ja34307

318. Jeffery PK. Structural and inflammatory changes in COPD: a comparison with asthma. Thorax. BMJ Publishing Group Ltd; 1998;53: 129–36. doi:10.1136/THX.53.2.129

319. Gurcan MN, Sahiner B, Petrick N, Chan H-P, Kazerooni EA, Cascade PN, et al. Lung nodule detection on thoracic computed tomography images: Preliminary evaluation of a computer-aided diagnosis system. Med Phys. John Wiley & Sons, Ltd; 2002;29: 2552–2558. doi:10.1118/1.1515762

320. Aberle DR, Gamsu G, Henschke CI, Naidich DP, Swensen SJ. A consensus statement of the Society of Thoracic Radiology: screening for lung cancer with helical computed tomography. J Thorac Imaging. 2001;16: 65–8. Available: http://www.ncbi.nlm.nih.gov/pubmed/11149694

321. Ireland RH, Woodhouse N, Hoggard N, Swinscoe JA, Foran BH, Hatton MQ, et al. An image acquisition and registration strategy for the fusion of hyperpolarized helium-3 MRI and x-ray CT images of the lung. Phys Med Biol. IOP Publishing; 2008;53: 6055–6063. doi:10.1088/0031-9155/53/21/011

322. Zhou X, Hayashi T, Hara T, Fujita H, Yokoyama R, Kiryu T, et al. Automatic segmentation and recognition of anatomical lung structures from high-resolution chest CT images. Comput Med Imaging Graph. Pergamon; 2006;30: 299–313. doi:10.1016/J.COMPMEDIMAG.2006.06.002

323. Raasch B, Carsky E, Lane E, O'Callaghan J, Heitzman E. Radiographic anatomy of the interlobar fissures: a study of 100 specimens. Am J Roentgenol. 1982;138: 1043–1049. doi:10.2214/ajr.138.6.1043

324. Gerard SE, Patton TJ, Christensen GE, Bayouth JE, Reinhardt JM. FissureNet: A Deep Learning Approach For Pulmonary Fissure Detection in CT Images. IEEE Trans Med Imaging. 2019;38: 156–166. doi:10.1109/TMI.2018.2858202

325. Restrepo RD, Braverman J. Current challenges in the recognition, prevention and treatment of perioperative pulmonary atelectasis. Expert Rev Respir Med. Informa Healthcare; 2015;9: 97–107. doi:10.1586/17476348.2015.996134

326. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. 2017; Available: https://arxiv.org/abs/1704.04861

327. Zoph B, Brain G, Vasudevan V, Shlens J, Le Google Brain Q V. Learning Transferable Architectures for Scalable Image Recognition [Internet]. Available: https://arxiv.org/pdf/1707.07012.pdf

328. Hu J, Shen L, Albanie S, Sun G, Wu E. Squeeze-and-Excitation Networks. 2017; Available: http://arxiv.org/abs/1709.01507