




# Recognition of Peripheral Lung Cancer and Focal Pneumonia on Chest Computed Tomography Images Based on Convolutional Neural Network

Technology in Cancer Research & Treatment  
 Volume 21: 1-12  
 © The Author(s) 2022  
 Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
 DOI: 10.1177/15330338221085375  
[journals.sagepub.com/home/tct](https://journals.sagepub.com/home/tct)  


Xiaoyue Cheng, MM<sup>1</sup> , He Wen, MD<sup>1</sup>, Hao You, PhD<sup>2,3</sup>, Li Hua, PhD<sup>2</sup> ,  
 Wu Xiaohua, MD<sup>1</sup>, Cao Qiuting, MM<sup>1</sup>, and Liu Jiabao, MD<sup>1</sup>

## Abstract

**Introduction:** Chest computed tomography (CT) is important for the early screening of lung diseases and clinical diagnosis, particularly during the COVID-19 pandemic. We propose a method for classifying peripheral lung cancer and focal pneumonia on chest CT images and undertake 5 window settings to study the effect on the artificial intelligence processing results. **Methods:** A retrospective collection of CT images from 357 patients with peripheral lung cancer having solitary solid nodule or focal pneumonia with a solitary consolidation was applied. We segmented and aligned the lung parenchyma based on some morphological methods and cropped this region of the lung parenchyma with the minimum 3D bounding box. Using these 3D cropped volumes of all cases, we designed a 3D neural network to classify them into 2 categories. We also compared the classification results of the 3 physicians with different experience levels on the same dataset. **Results:** We conducted experiments using 5 window settings. After cropping and alignment based on an automatic preprocessing procedure, our neural network achieved an average classification accuracy of 91.596% under a 5-fold cross-validation in the full window, in which the area under the curve (AUC) was 0.946. The classification accuracy and AUC value were 90.48% and 0.957 for the junior physician, 94.96% and 0.989 for the intermediate physician, and 96.92% and 0.980 for the senior physician, respectively. After removing the error prediction, the accuracy improved significantly, reaching 98.79% in the self-defined window2. **Conclusion:** Using the proposed neural network, in separating peripheral lung cancer and focal pneumonia in chest CT data, we achieved an accuracy competitive to that of a junior physician. Through a data ablation study, the proposed 3D CNN can achieve a slightly higher accuracy compared with senior physicians in the same subset. The self-defined window2 was the best for data training and evaluation.

## Keywords

chest CT, peripheral lung cancer, focal pneumonia, 3D CNN, window settings

## Abbreviations

AUC, area under the curve; Chest CT, chest computed tomography; CNN, convolutional neural network; FW, full window; IID, independent and identically distributed; PLC, peripheral lung cancer; ROC, operating characteristic curves; SW1, self-defined window1; SW2, self-defined window2; WW, window width; WL, window level.

Received: August 8, 2021; Revised: January 27, 2022; Accepted: February 16, 2022.

## Introduction

Chest computed tomography (CT) is an extremely important tool for the early screening of lung diseases and clinical diagnosis, particularly in the current COVID-19 pandemic. Radiologists interpret chest images by recognizing the important radiographic patterns of lung disease.<sup>1</sup> With an increase in the pathogen spectrum, the aging of society, and the wide use of antibiotics, the imaging findings of pneumonia are less

<sup>1</sup> Department of Radiology, Beijing Friendship Hospital, Capital Medical University, Beijing, China

<sup>2</sup> Key laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

<sup>3</sup> University of Chinese Academy of Sciences, Beijing, China

### Corresponding Author:

He Wen, MD, Department of Radiology, Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China.  
 Email: [hewen202003@163.com](mailto:hewen202003@163.com)



typical. An example for this is focal pneumonia, which is caused by bacteria or viruses. Such pneumonia can be acute inflammation of the lung or a manifestation of the absorption process of lobar or segmental pneumonia. It is difficult to differentiate peripheral lung cancer (PLC) from focal pneumonia on chest CT, particularly for junior doctors or less experienced radiologists. Clinically and through imaging, focal pneumonia is often misdiagnosed as PLC, and patients may therefore receive unnecessary surgical treatment, which also has a certain impact on the physical and mental health of the patient. It would be helpful if an automated method to assist physicians in classifying these 2 types of lesions could be made available.

With the rapid development of an intelligent diagnosis of medical images, artificial intelligence (AI) methods have been widely used in medical image processing and have accomplished many achievements.<sup>2-5</sup> AI algorithms have made significant progress in many medical image application scenarios, such as the detection of pulmonary nodules,<sup>6-9</sup> the classification of benign and malignant pulmonary nodules,<sup>10-14</sup> the staging of lung cancer,<sup>15,16</sup> a study on lung cancer prognostication,<sup>17</sup> and the classification of interstitial<sup>18-20</sup> and other<sup>21-23</sup> lung diseases. A convolutional neural network (CNN) is a deep learning technology model. A CNN is designed to make better use of spatial and configuration information.<sup>24</sup> In 2016, Cicek *et al.* extended the U-Net model to a 3D U-net.<sup>25</sup> In addition, Deniz *et al.* compared a 2D CNN with a 3D CNN and found that in some commonly used analysis metrics, the vanilla 3D U-net model showed better results than the vanilla 2D U-net model.<sup>26</sup> Liu *et al.* also showed that a single 3D CNN model and ensemble models with a 3D CNN outperformed their respective counterparts constructed using only traditional models.<sup>27</sup> Moreover, Harmon *et al.* proposed the use of a 3D CNN to identify COVID-19-associated pneumonia and non-COVID-related pneumonia in CT scans.<sup>28</sup> The application of a CNN in the field of computer vision is being transformed into applications in medical image analysis.

In this paper, we propose a novel method for classifying PLC and focal pneumonia on chest CT images and apply 5 window settings to study the effect on the AI processing results. Few previous studies have discussed the influence of window settings on the AI results. Our method consists of the following parts: (1) *Segmentation*: First, we segment the lung parenchyma based on morphological methods. (2) *Orientation alignment*: Considering the different poses and orientations of the subjects during scanning, we propose a method for determining the orientation of the subjects based on lung segmentation. (3) *Training*: After alignment, we crop the 3D region with the minimal bounding box of the segmentation. With the cropped 3D volume as input, we train a neural network to classify the consolidation and nodule subjects. (4) *Window setting*: During the training, we also consider the compact of different window settings of CT values in Hounsfield units (HU). (5) *Evaluation*: To evaluate the effectiveness of the trained neural network, we compared the classification results of neural networks to those of 3 radiologists with different experience levels on the same dataset.

## Methods

### Data Collection

We retrospectively collected patients who underwent chest CT at Beijing Friendship Hospital from March 2013 to March 2018. The CT scanner was a GE Discovery CT 750 HD with scanner parameters of 120 kVp, a 5-mm collimated acquisition with a 0.984–1 pitch, and a reconstructed thickness of 1.25 mm, or a Philips Brilliance iCT with scanner parameters of 120 kVp, a 5-mm collimated acquisition with a 0.915 pitch, and a reconstructed thickness of 1.25 mm. The pixel resolution of each image was 512×512 and the scans were conducted without an intravenous contrast. We have de-identified all patient details. The inclusion and exclusion criteria for subjects used in this study are as follows: (1) CT manifestations were either solitary solid pulmonary nodule or solitary consolidation. Solitary solid pulmonary nodule is single, well-defined, opaque, and straight lesion surrounded by air containing lung tissue, which is roughly spherical. Solitary consolidation is single, usually occupies an anatomical unit in morphology, and is distributed according to lobes or segments of the lungs, and generally, the above structures are wedged. (2) Patients with solitary solid pulmonary nodule were diagnosed as PLC, which was pathology-proven (pathological types including adenocarcinoma, squamous cell carcinoma, and small cell carcinoma). (3) Patients with solitary consolidation were clinically diagnosed as focal pneumonia and the solitary consolidation was reexamined and absorbed after treatment. Finally, (4) Patients with other chest lesions, having artifacts such as breathing movements, or with nodule of less than 1 cm, were excluded. For the measurement, we first measure the long-axis diameter of the lesion, then measure the short-axis perpendicular to the long axis on the same CT slice, and finally calculate the average size.

According to the standard, 357 patients were enrolled in the study group. The patients were divided into 2 groups: a PLC group (n = 172) and a focal pneumonia group (n = 185). The 172 cases in the PLC group were obtained from 94 men and 78 women (age range of 32–82 years; mean age of 63 years). The 185 patients in the focal pneumonia group were obtained from 85 men and 100 women (age range of 8–96 years; mean age of 44 years). The locations of each lesion of the 357 patients are listed in Table 1. Examples of these 2 groups are presented in Figure 1. A scan prior to the diagnosis was selected. Ethical approval for this study was obtained from the Medical ethics committee of our hospital (the approval number 2018-P2-100-01). The need for informed consent from all patients was waived because of the retrospective nature of the data acquisition.

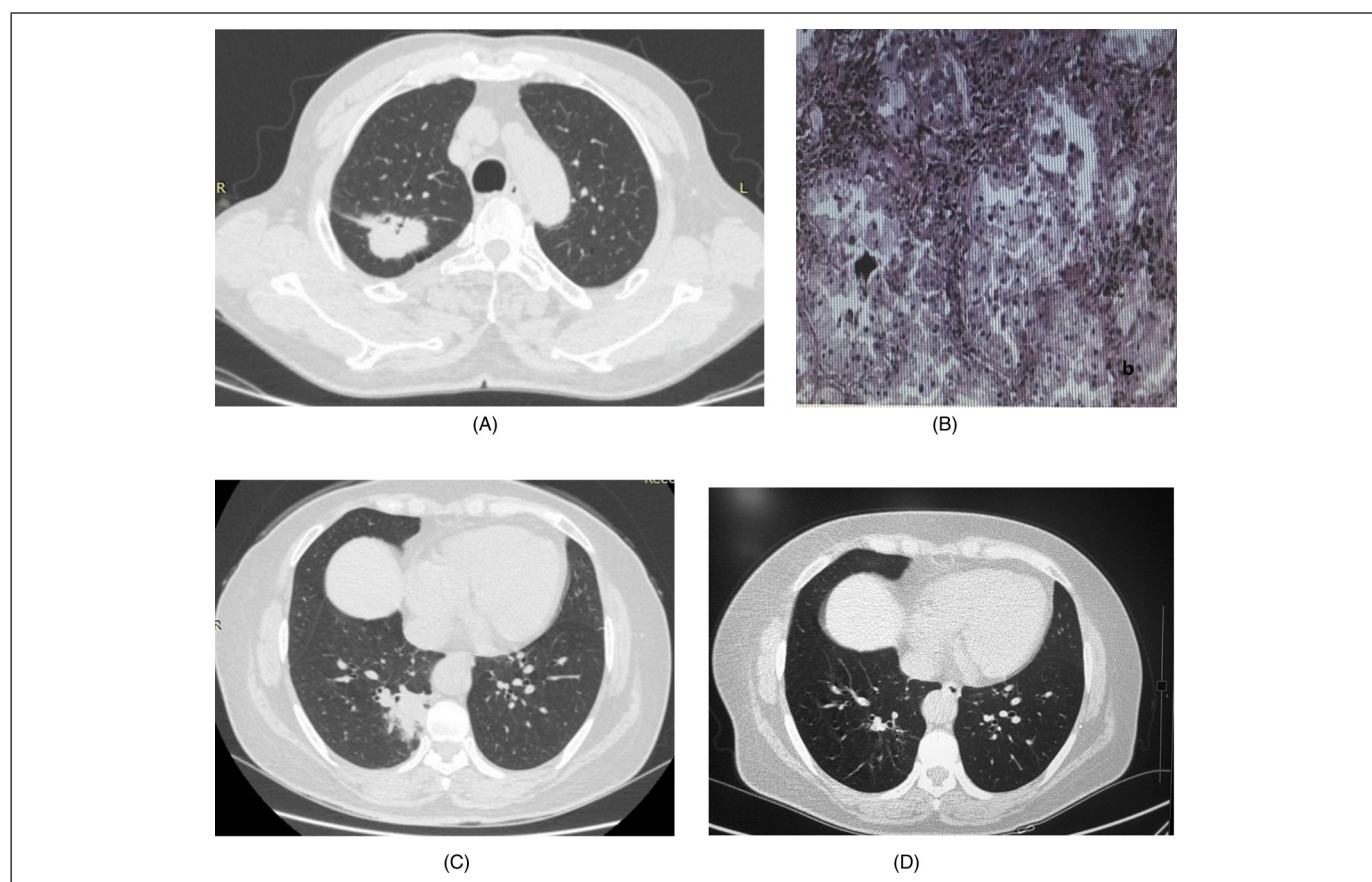
### Window Settings

The application of window width (WW) and window level (WL) technologies in the interpretation of a lung CT is important. WW mainly affects the image contrast, and WL mainly

**Table 1.** The Locations of Each Lesion of the 357 Patients.

The locations of each lesion	The superior lobe of the right lung	The middle lobe of the right lung	The inferior lobe of the right lung	The superior lobe propria of the left lung	The lingual segment of the left lung	The inferior pulmonary lobe of the left lung
The cases of the PLC group	42	12	43	37	13	25
The cases of the focal pneumonia group	17	36	50	13	21	48

Abbreviation: PLC, peripheral lung cancer.

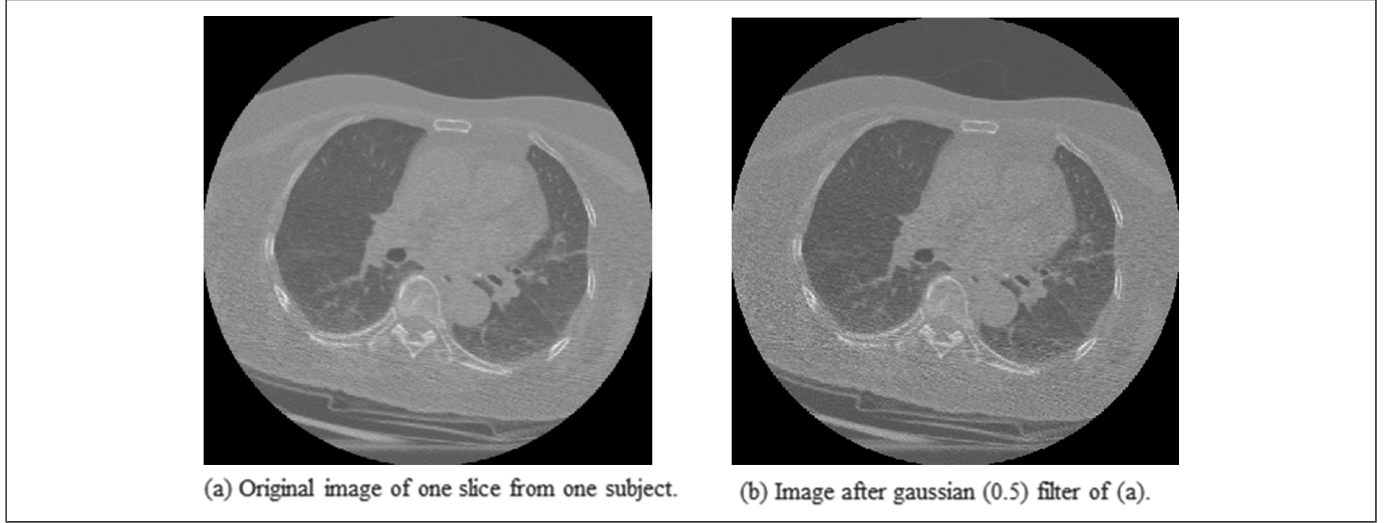


**Figure 1.** Some examples from the collected dataset. Images (A) and (B) are from one patient, the diagnosis of which is peripheral lung cancer (PLC): (A) The nodule in the upper lobe of the right lung is round-like and with spiculations. (B) A pathological examination shows adenocarcinoma. Images (C) and (D) are from another patient, the diagnosis of which is focal pneumonia: (C) The consolidation in the lower lobe of the right lung shows an irregular shape, with the apex pointed toward the hilar side of the lung. (D) The scan shows that the consolidation was absorbed after 1 month.

affects the brightness of the image (black and white). Under different values of WW and WL, the density contrast of the images will change, the edge morphology of the lesions will be affected, and the apparent size of the mass will be altered.

We calculated the following 5 window settings for the data: (1) the lung window (LW) with WW and WL as (1500, -600), (2) the mediastinal window (MW) as (350, 40), (3) the self-defined window 1 (SW1) as (1000, 40), (4) the self-defined

window 2 (SW2) as (1000, -100), and (5) the full window (FW) as (4096, 1024). SW1 is converted from the mediastinal window. We chose the same WL of the MW and expanded the WW to 1000 (in a clinical study, we found that SW1 can consider both the edge of the lesion and the difference in the density of the lesion to a certain extent). SW2 is the shift of SW1. The FW is converted from the CT value range of the Digital Imaging and Communication in Medicine image stored in the CT scan into -1024-3071HU.



**Figure 2.** Example of original image and after applying a Gaussian filter. In the images of some of the subjects, some background noise appeared, affecting the subsequent segmentation processes.

After applied the window on the image, the intensity of the images will be converted following Equation (1):

$$I' = \begin{cases} WL - \frac{WW}{2}, & \text{where } I < \left(WL - \frac{WW}{2}\right) \\ I, & \text{where } \left(WL - \frac{WW}{2}\right) \leq I \leq \left(WL + \frac{WW}{2}\right) \\ WL + \frac{WW}{2}, & \text{where } I > \left(WL + \frac{WW}{2}\right) \end{cases}, \quad (1)$$

where  $I$  is the original image intensity, and  $I'$  is the image intensity after trimmed by the window.

1. Lung Window (LW):  $(WW, WL) = (1500, -600)$ . And the intensity of the image will be trimmed into  $[-1350, 150]$ .
2. Mediastinal Window (MW):  $(WW, WL) = (350, 40)$ . And the intensity of the image will be trimmed into  $[-135, 215]$ .
3. Self-defined Window 1 (SW1):  $(WW, WL) = (1000, 40)$ . And the intensity of the image will be trimmed into  $[-460, 540]$ .
4. Self-defined Window 2 (SW2):  $(WW, WL) = (1000, -100)$ . And the intensity of the image will be trimmed into  $[-600, 400]$ .
5. Full Window (FW):  $(WW, WL) = (4096, 1024)$ . And the intensity of the image will keep the original range  $[-1024, 3071]$ .

### Preprocess

To make training easier, we preprocess training data images before feeding them into the neural network. Because an independent and identically distributed (IID) assumption is the basis of current learning methods, regardless of what kind of model we use in the subsequent learning process, uniformity in the

data distribution will significantly help with the learning process.

Our main preprocessing includes the following steps, all of which are conducted in a 3D image space.

**S1: Gaussian denoising.** In some scanned images, some noise is generated in the CT image during the imaging process, as one slice of an example shows in Figure 2A. The noise will significantly affect the subsequent threshold-based segmentation; therefore, we first apply a Gaussian filter with a sigma of 0.5. The image after denoising is shown in Figure 2B. The corresponding histograms are shown in Figure 3A.

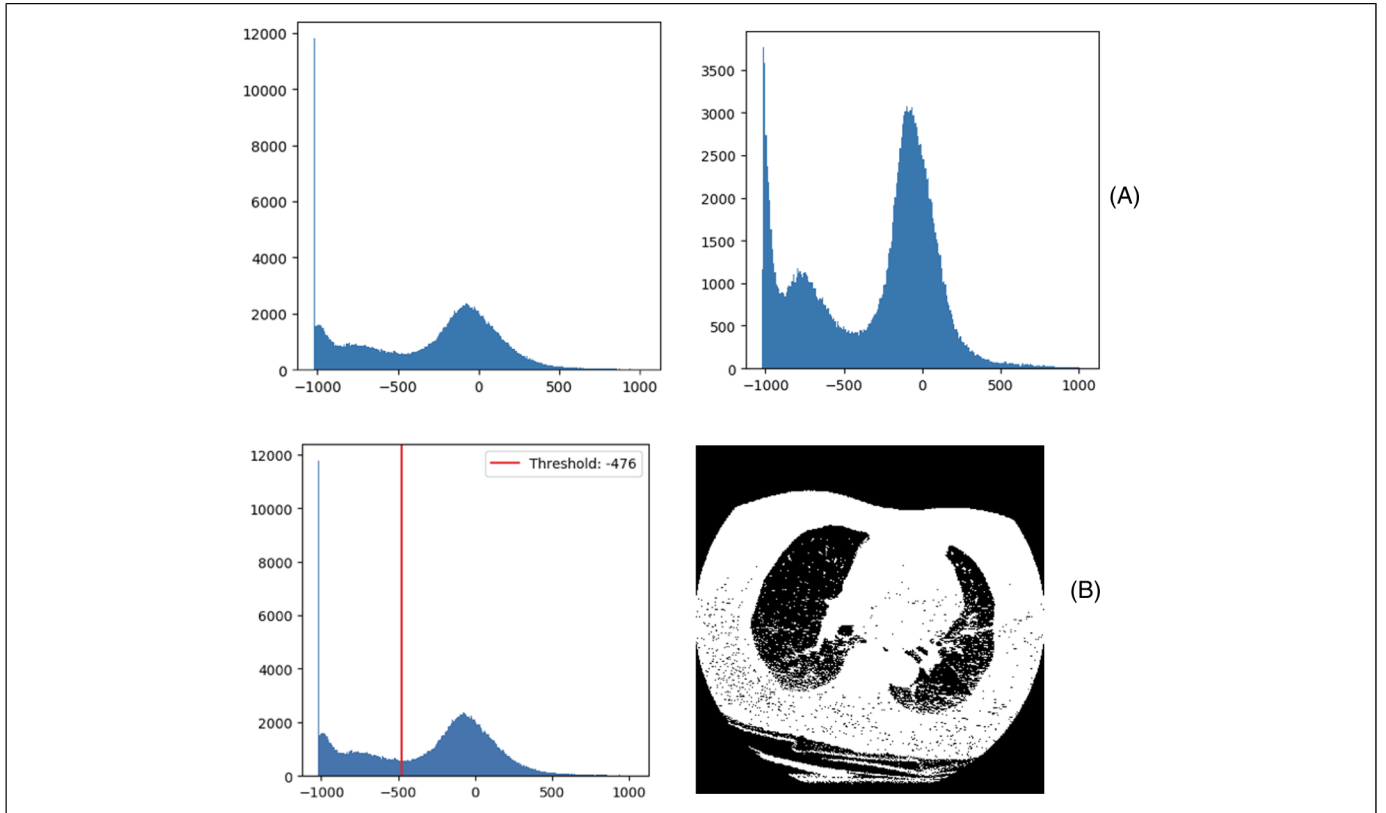
**S2: Lung segmentation.** From the histograms of all subjects in the dataset, we can see that the histogram has obvious bimodal characteristics. Therefore, we can conduct a preliminary segmentation based on the threshold method. An example is shown in Figure 3B.

With some morphology post-processing steps, we can obtain coarse segmentation of the lung in a robust manner. An example of this is shown in Figure 4A. The lung mask after 3D rendering is shown in Figure 4B.

To evaluate the performance of these segmentation algorithms, we conducted an experiment on the LUNA16-51 dataset,<sup>29</sup> which has the ground truth of the lung masks of 51 subjects. We then evaluate the Dice coefficient, as shown in Equation (2), for each subject in the dataset, as depicted in Figure 4C. The Dice values are within the range  $[0, 1]$ . A similarity of 1 means that the 2 segmentations A and B are perfectly matched.

$$Dice = \frac{2 * (A \cap B)}{A + B} \quad (2)$$

From the results, we can see that all Dice values are greater than 0.96. The mean and standard deviation are  $0.9881 \pm 0.0070$ ,



**Figure 3.** (A) The histograms of the original image and the image after applying a Gaussian filter. (B) The threshold from the OTSU method determined automatically and a binary image after segmentation using the threshold.

which proves that our segmentation matches the ground truth extremely well.

**S3: Pose alignment.** From the example shown in Figure 4A, we can see that certain objects may occur in different directions, and not strictly upward, when scanned. The different orientations will definitely affect the subsequent learning process, although we can ease this situation using data augmentation techniques. Therefore, we developed a method for detecting the orientation of the subject based on the lung masks and then align them in a strictly upward direction.

We first selected 5 slices with the largest area of the lung mask. We can then obtain the bottom of both the left and right halves of the lungs. We can then define the orientation for each 2D slice. The average orientation of the 5 slices will be the final orientation of the subject. An example of this is shown in Figure 5.

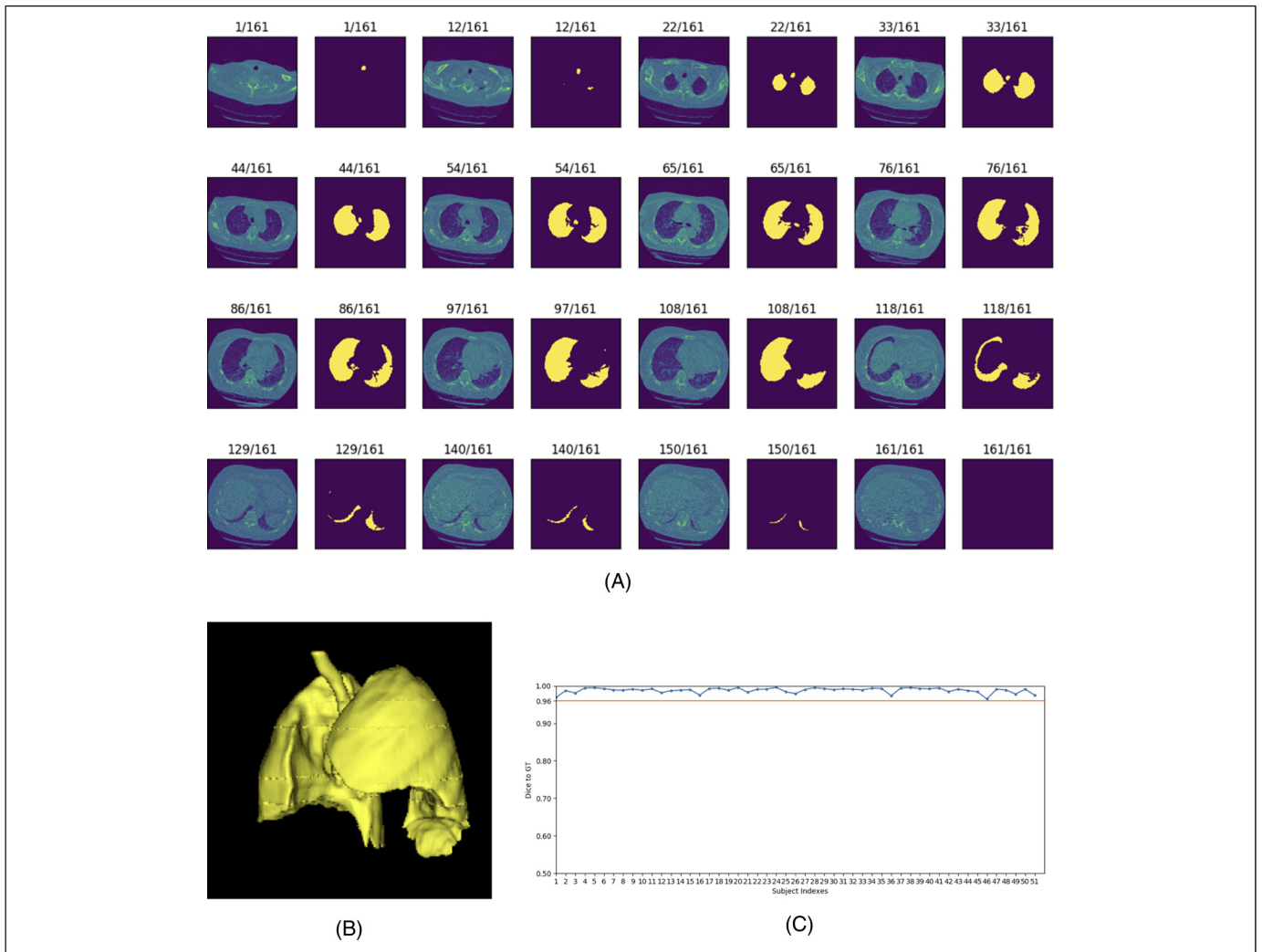
**S4: Coarse localization.** In addition, with the aligned subject and mask, we can obtain the coarse bounding box of the lung region. The bounding box of the region is padded to a certain extent to ensure that the entire lung is included in the region. An example of this is shown in Figure 6.

According to the segmentation results of the lung parenchyma region, we can obtain the bounding box of the lung parenchyma region, which we cut out as a part of our final input into the neural network. Because the size of the

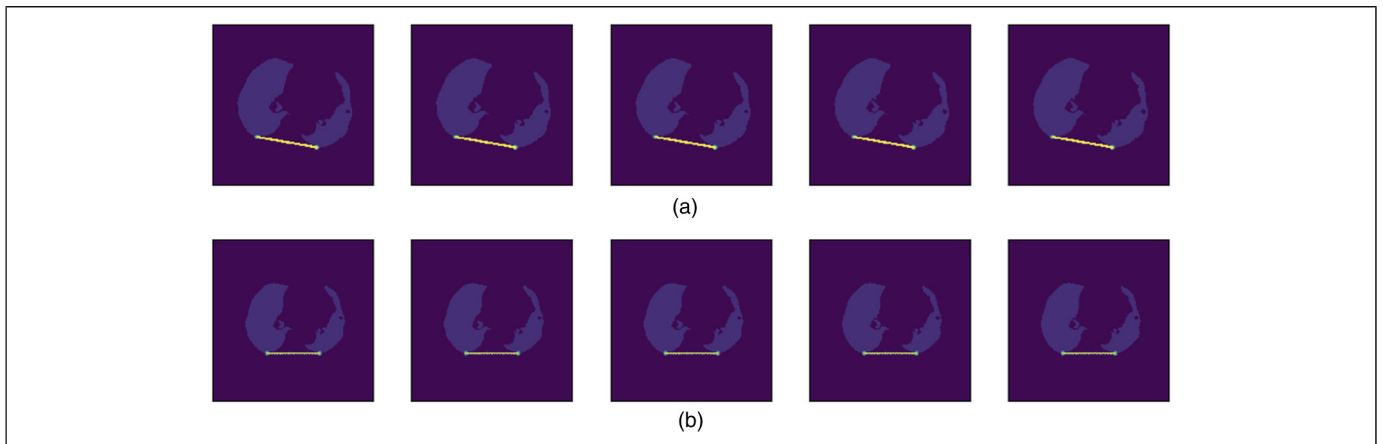
pulmonary parenchymal areas varies from case to case, we need to scale these areas to the same size. Considering the computing complexity and network architecture design, we set the input size as (96, 64, 96).

### Network Architecture

We used the entire 3D scan as the input and a full 3D CNN as the model for classification. The entire classification workflow is shown in Figure 7. We used 3D ResNet with 18 layers as the backbone of our network. The core idea of ResNet is to introduce an identity connection that skips several layers.<sup>30</sup> In this way, the networks can be increased into deeper layers while avoiding the notorious vanishing gradient problem. Therefore, we chose the 3D ResNet with 18 layers as the backbone. The architecture is illustrated in Fig. 8. In each Conv block, there is one convolution layer with the parameters listed in the block, one batch normalization layer, and one ReLu activation layer. The first parameter in the block, “ $3 \times 3 \times 3$ ,” indicates the 3D kernel size, the third parameter, “64, 128, 256, 512,” represents the number of channels, and the last parameter, “/2,” indicates a pooling layer with the stride of 2. Finally, with a global average pool layer and a fully connected layer with 2 outputs, the model can classify the case into 2 different classes: focal pneumonia and PLC. To avoid an overfitting of the training data, we also employed a dropout layer with a ratio of 0.5 before the last fully connected layer (Figure 8).



**Figure 4.** (A) Example of coarse lung segmentation using threshold-based method. The title of each subfigure represents the index/total number of slices in the subject. (B) 3D rendered lung mask. (C) Dice distances between the threshold-based lung segmentation and the ground-truth of all subjects in the LUNA16-51 dataset.



**Figure 5.** Example of the orientation alignment of the subject based on the lung mask.

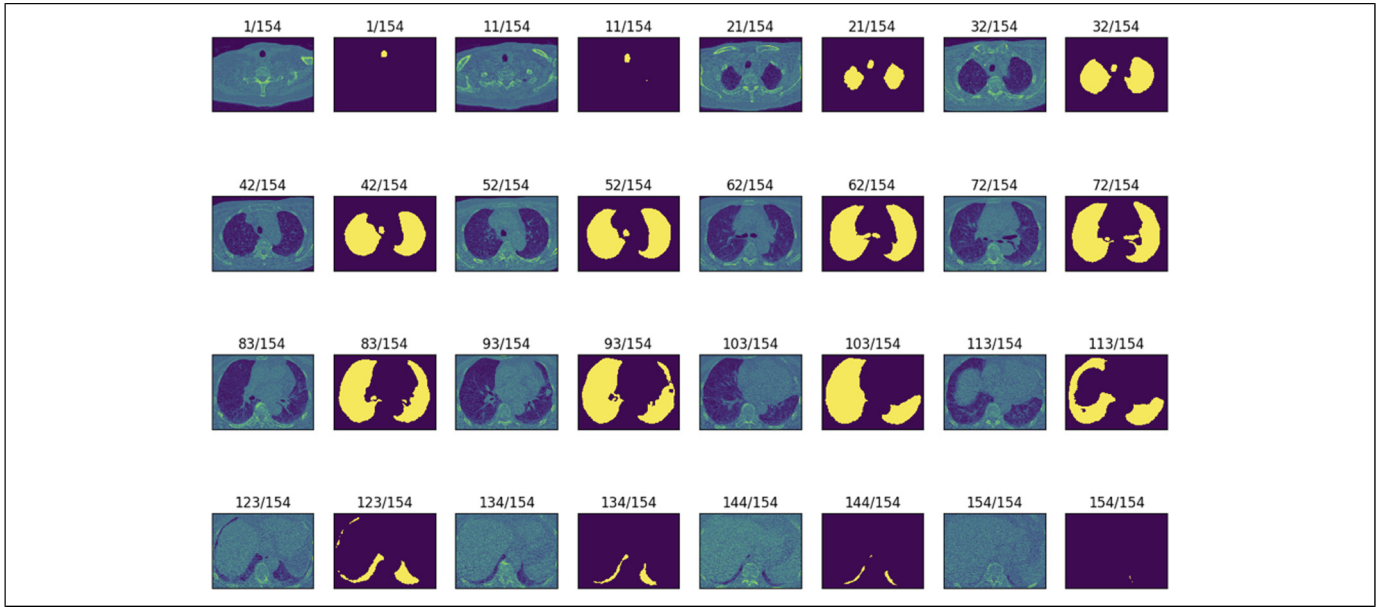


Figure 6. Coarse localization of the lung region.

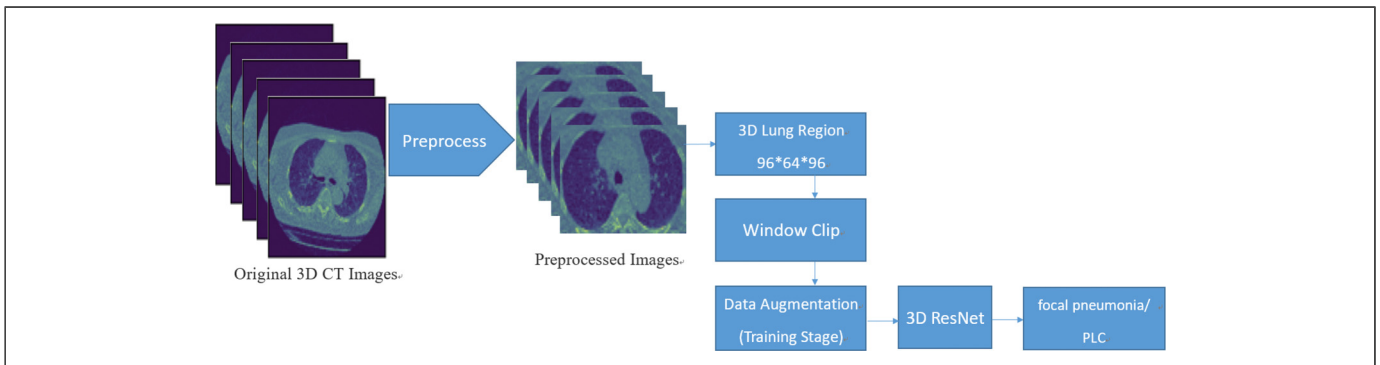


Figure 7. Workflow of the entire classification model.

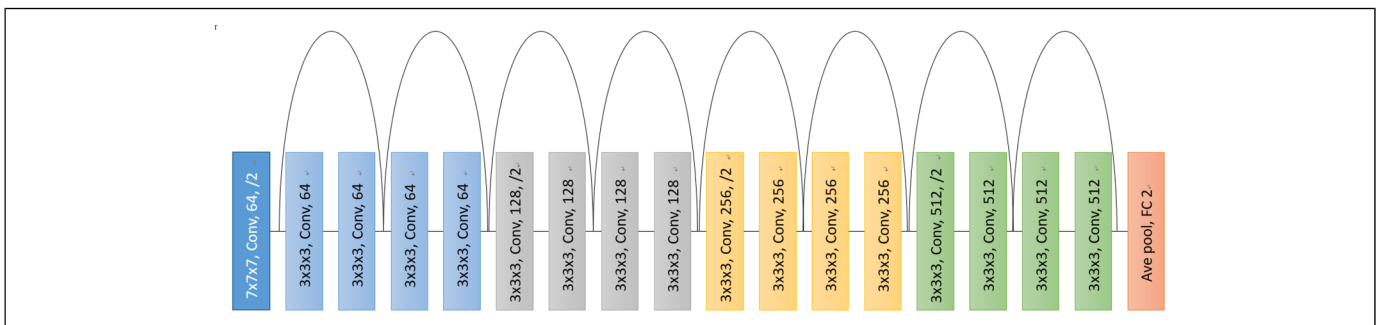
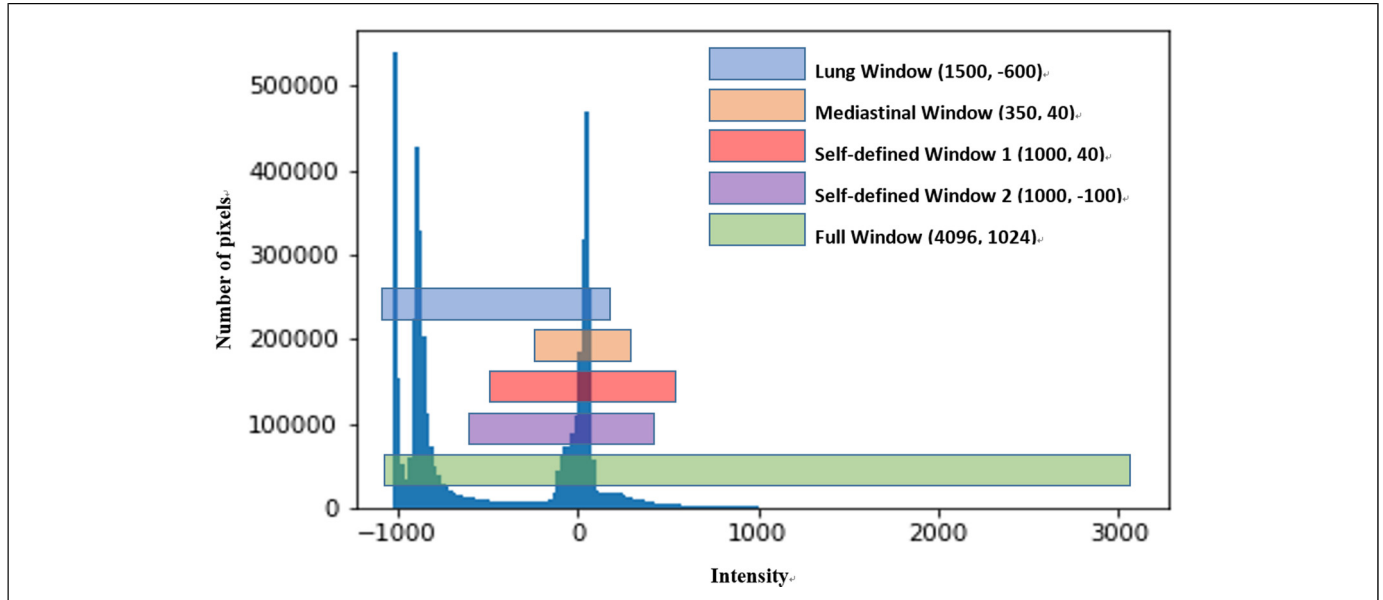


Figure 8. Architecture of 3D ResNet.

### Size Normalization

Because the size of the pulmonary parenchymal areas varies from case to case, before inputting the data into the network, all subjects should be resized to the same size. There are several ways to deal with size differences in the subjects. Other methods normalize

them into the same pixel size and retain the original real size of the lung. This method is reasonable if there are sufficient cases in the training set to cover the variations in the real lung size; thus, the model can obtain the ability to handle different lung sizes. However, in our method, we chose to resize all of the



**Figure 9.** Different window settings on the histogram of a single example.

images into the same shape because our training data are not too large. Therefore, we do not need too much data to cover the variation in the real lung size because we have normalized them into the same size. As another reason, we developed a reliable lung segmentation method, as described in “S2: Lung segmentation,” which can be used to coarsely localize the lung region, which makes available a resizing to the same size.

### Intensity Window Settings

We used 5 different window settings to analyze the impact on the final classification performance. From Figure 9, we can see the difference between the different window settings.

### Data Augmentation

In this experiment, we used 2 different transformations for data augmentation: a random rotation at an angle of  $[-10, 10]$  in the axis plane, and adding random Gaussian noise with an average value of zero and a standard deviation of 0.01 (the input intensity was normalized to  $[-1, 1]$ ).

### Experiment Details

The hyperparameters of the network are as follows. An Adam optimizer was used. The initial learning rate was 0.0001, beta 1 = 0.9, beta 2 = 0.999, and the weight decay was 0.001. All experiments were trained for 200 epochs. We conducted our experiments on a computer equipped with an Intel Core i7-8700 CPU and an NVIDIA GTX1080ti GPU. PyTorch was used as the deep learning platform.

Statistical Analysis: IBM SPSS Statistics (version 19) was used for data analysis. We calculated the classification accuracy, made

receiver operating characteristic (ROC) curve, and the AUC. The reporting of this study conforms to STARD guidelines.<sup>31</sup>

## Results

### Classification Under Different Window Settings

First, to evaluate the method proposed above, we conducted experiments on the dataset with different window settings. First, to conduct a cross-validation, the whole dataset was randomly split into 5 folds. We then conducted a 5-fold cross validation. The details, WW, and WL are listed in Tables 2 to 4.

It can be seen from the table that different window settings have a certain impact on the final classification result. When the WW was wider, the accuracy improved. The 3D ResNet achieved the highest accuracy of 91.596% in the FW setting.

### Comparison to Clinical Physicians

To compare the classification of the 3D ResNet with that of clinical physicians, we involved 3 radiologists with different experiences to interpret the CT images of all cases without knowing the patient information and history. The levels of these 3 radiologists were as follows: (1) a junior doctor, that is, a radiologist with 3 years of subspecialty experience in chest imaging, (2) an intermediate doctor, that is, a radiologist with 5 years of subspecialty experience in chest imaging, and (3) a senior doctor, that is, a radiologist with 10 years of experience in chest imaging. These physicians then scored the images according to the following rules: focal pneumonia, score of 1; likely focal pneumonia, score of 2; borderline, score of 3; likely PLC, score of 4; and PLC, score of 5.



**Table 2.** Classification Results of 5-Fold Cross Validation of 5 Window Settings.

Window setting	(WW, WL)	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Ave.	AUC
Lung	(1500, -600)	87.50	88.73	91.55	91.55	93.06	90.478	0.929
Mediastinal	(350, 40)	80.56	87.32	90.14	85.92	84.72	85.732	0.871
SW1	(1000, 40)	86.11	92.96	88.73	91.55	90.28	89.926	0.918
SW2	(1000, -100)	87.50	92.96	90.14	91.55	90.28	90.486	0.923
Full	(4096, 1024)	87.50	90.14	90.14	94.37	95.83	<b>91.596</b>	<b>0.946</b>

Abbreviations: AUC, area under the curve; WL, window level; WW, window width. Bold is best.

**Table 3.** The Details of Subjects in Different Folds of the Original Dataset (357 Subjects in Total).

	Type		Gender		Age		Location						Size (cm)			
	Group 1	Group 2	M	F	<50	≥50	SLRL	MLRL	ILRL	ILLL	SLPLL	LSSL	<2	[2,3)	[3, 4)	≥4
Fold 1	39	33	41	31	24	48	17	11	19	10	9	6	18	22	16	16
Fold 2	41	30	37	34	28	43	9	10	14	19	10	9	12	24	22	13
Fold 3	37	34	38	33	26	45	5	9	21	15	14	7	12	24	15	20
Fold 4	33	38	31	40	25	46	14	8	15	16	11	7	20	27	6	18
Fold 5	35	37	32	40	26	46	14	10	24	13	6	5	16	25	10	21

**Table 4.** The Details of Subjects in Different Folds of the Subset (328 Subjects in Total).

	Type		Gender		Age		Location						Size (cm)			
	Group 1	Group 2	M	F	<50	≥50	SLRL	MLRL	ILRL	ILLL	SLPLL	LSSL	<2	[2,3)	[3, 4)	≥4
Fold 1	34	31	35	30	22	43	16	11	15	10	9	4	18	18	14	15
Fold 2	37	30	34	33	25	42	8	10	14	17	9	9	10	24	21	12
Fold 3	30	33	34	29	24	39	5	9	16	15	13	5	10	23	11	19
Fold 4	28	37	29	36	22	43	12	7	15	15	11	5	18	24	6	17
Fold 5	33	35	29	39	26	42	13	10	22	13	5	5	14	25	8	21

Abbreviations: Group 1, the focal pneumonia group; Group 2, the PLC group; SLRL, superior lobe of the right lung; MLRL, The middle lobe of the right lung; ILRL, inferior lobe of the right lung; SLPLL, superior lobe propria of the left lung; LSSL, lingual segment of the left lung; ILLL, inferior lobe of the left lung.

The results showed that the accuracy of 3D ResNet (Full) was 91.596%, and the AUC was 0.946. The accuracies of the junior, intermediate, and senior doctors were 90.48%, 94.96%, and 96.92%, respectively. In addition, the AUCs of the junior, intermediate, and senior doctors were 0.957, 0.989, and 0.980, respectively. The ROC curves are shown in Figure 10. We can see that the average accuracy of 3D ResNet with a FW is slightly higher than that of the junior physicians. This proves that our 3D-CNN can effectively distinguish solitary solid pulmonary nodules from solitary consolidation.

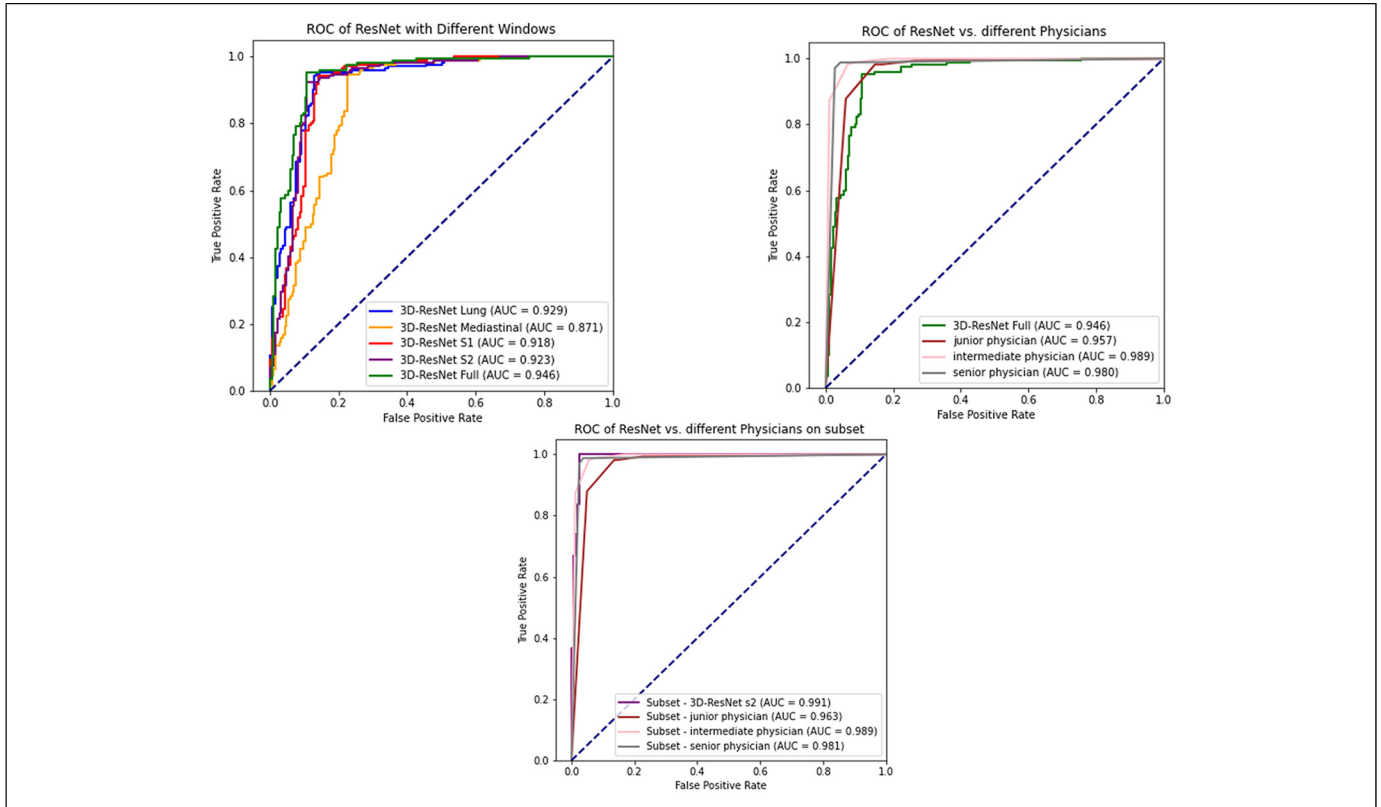
### Data Ablation Study

In this experiment, we removed several subjects that 3D ResNet predicted as wrong in previous experiments and generated a sub-dataset of 328 cases, as listed in Table 5. We conducted this experiment to explore what caused the lower accuracy of the CNN in comparison to that of the intermediate and advanced physicians in the previous experiment.

We can see that, after removing the error prediction cases, the accuracy improved significantly, reaching 98.79% in

SW2. However, the accuracy of all physicians did not change significantly, even for the junior physicians. This means that for the physicians, the removed cases of mispredictions are not difficult for them. There must therefore be other reasons for the incorrect prediction of the CNN. We suppose that this is due to insufficient data. Because a CNN requires sufficient data to train and fit the actual distribution of cases, and our training data are limited (only 357 in total), certain types of cases will not appear in all folds while splitting. Therefore, we assume that if there are more training data, we can obtain better results.

This experiment conducted on the subset also proved this to a certain extent. The results showed that the accuracy of 3D ResNet (SW2) was 98.79%, and the AUC was 0.991. The accuracies of the junior, intermediate, and senior doctors were 91.77%, 95.73%, and 97.26%, respectively. The AUCs of the junior, intermediate, and senior doctors were 0.963, 0.989, and 0.981, respectively. The ROC curves are shown in Figure 10. On the same subset, 3D ResNet can achieve a slightly higher accuracy than the senior physician.



**Figure 10.** Receiver operating characteristic (ROC) curves of 3D ResNet with different window settings and those of 3 physicians, and ROC curves of 3D ResNet with different window settings and 3 physicians for a subset.

**Table 5.** Classification Results of 3D ResNet on the Subset of Different Window Settings.

Window setting	(WW, WL)	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Ave.	AUC
Lung	(1500, -600)	96.92	95.52	98.41	98.46	98.53	97.57	0.985
Mediastinal	(350, 40)	90.77	94.03	95.24	92.31	89.71	92.41	0.939
SW1	(1000, 40)	96.92	100	98.41	98.46	97.06	98.17	0.985
SW2	(1000, -100)	98.46	100	100	98.46	97.06	<b>98.79</b>	<b>0.991</b>
Full	(4096, 1024)	96.92	97.01	100	98.46	97.06	97.89	0.986

Abbreviations: AUC, area under the curve; WL, window level; WW, window width. Bold is best.

## Discussion

In our experiments, we designed a 3D CNN to classify PLC and focal pneumonia on chest CT images and applied 5 window settings to study the effect on the AI processing results. We took each case as a whole, extracted the features layer by layer through a 3D convolution, bypassed the location of the lesion, and classified and judged the entire case. In addition, we selected 3 radiologists to score the CT images, calculated the classification accuracy, created the ROC curve, and calculated the AUC. Our neural network achieved an average accuracy of 91.596% in the 5-fold cross-validation.

In recent years, CNNs have been widely used in lung diseases, particularly lung cancer,<sup>10–17</sup> and a number of studies have begun to use a CNN to classify more lung diseases. For example, some studies have begun using CNNs to classify

interstitial lung diseases.<sup>18–20</sup> Lakhani *et al.* evaluated the efficacy of deep CNNs for detecting tuberculosis on chest radiographs.<sup>21</sup> Taylor *et al.* trained deep convolutional networks to screen for a moderate or large potentially emergent pneumothorax.<sup>22</sup> Zech *et al.* conducted a cross-sectional study on the variable generalization performance of a deep learning model to detect pneumonia on chest radiographs.<sup>23</sup> Our accuracy of 91.596% was similar to or higher than that reported above, demonstrating that our 3D-CNN can effectively distinguish between PLC and focal pneumonia. The classification accuracy of our neural network was 91.596%, which was higher than that of junior physicians (90.48%) and was lower than that of the intermediate and senior physicians. In fact, our classification method can improve the efficiency of the image interpretation and can be used in grassroots hospitals lacking senior doctors. It can

also be used in physical examination centers for disease screening.

From the data ablation study, we can see that after removing a few cases of incorrect predictions, the 3D CNN achieved a higher accuracy than that of the senior physician on the same subset. We believe that if we can collect more data, the current results can be improved to a certain extent.

In this study, we used 5 window settings to study the effects on AI processing results. SW1 is converted from the mediastinal window: We chose the same WL of the MW and expanded the WW to 1000 (in a clinical study, we found that SW1 can consider both the edge of the lesion and the difference in the density of the lesion to a certain extent). SW2 is a shift of SW1. Our results prove that after removing the error prediction cases, the accuracy improved significantly, reaching 98.79% in SW2. In other words, the self-defined window2 (1000, -100) was the best choice for data training and evaluation. However, research on this aspect is relatively limited and has not yet reached a consensus. It is hoped that the factors of the window settings can be considered in future AI studies.

Compared to traditional machine learning, deep learning does not require a manual intervention and can run through the entire process from the input image to the final classification. However, although we have spent a lot of energy collecting a relatively large number of cases, it is still limited by the number of training samples. Our study selected patients scanned by 2 CT scanner. Ideally, CT scans of all patients will be collected with consistent acquisition parameters. However, this is a challenge in practice. In addition, our retrospective study involved only PLC and focal pneumonia. For other lung diseases (such as benign nodules, interstitial lung diseases, etc), further research is needed in the future. In the future, we can further use a 3D-CNN to classify other patterns of lesions on a lung CT. For the size of the data set, more data can be included in the future, too.

## Conclusion

In conclusion, we designed a 3D-CNN for classifying PLC and focal pneumonia on chest CT to simulate the reading process of radiologists and applied 5 window settings to study the effect on the processing results of the AI. We plan to consider new applications for AI in the area of lung diseases.

## Authors' Note

Ethical approval for this study was obtained from the Medical ethics committee of Beijing Friendship Hospital, Capital Medical University (the approval number 2018-P2-100-01). Approval for our study was received from our institutional review board. Because of the retrospective nature of the data acquisition, need for an informed consent from all patients was waived.

## Acknowledgments

This paper is supported by the National Key R&D Program of China (2017YFB1002703), the National Key Basic Research Program of China (2015CB554507), and the National Natural Science

Foundation of China (Grant No. 61379082). The authors also would like to thank Joshua Sheng and Zhao Mengyu for their edits.



## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: My paper is supported by the National Key R&D Program of China (2017YFB1002703), the National Key Basic Research Program of China (2015CB554507), and the National Natural Science Foundation of China (Grant No. 61379082).

## ORCID iD

Xiaoyue Cheng  <https://orcid.org/0000-0002-4143-8987>  
Li Hua  <https://orcid.org/0000-0003-4982-6455>

## References

1. McLoud TC, Boiselle PM. *Thoracic Radiology: The Requisites*. 2nd ed. Mosby Elsevier; 2010:27-57.
2. van Ginneken B. Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. *Radiol Phys Technol*. 2017(1);10:23-32.
3. Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin*. 2019;69(2):127-157.
4. Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*. 2018;172(5):1122-1131.
5. Geng Y, Ren Y, Hou R, et al. 2.5D CNN model for detecting lung disease using weak supervision. *Proc. SPIE 10950, Medical Imaging 2019: Computer-Aided Diagnosis*, 109503O (13 March 2019). DOI:10.1117/12.2513631.
6. Setio AAA, Ciompi F, Litjens G, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE Trans Med Imaging*. 2016;35(5):1160-1169.
7. Winkels M, Cohen TS. 3D G-CNNs for pulmonary nodule detection. 1st Conference on Medical Imaging with Deep Learning, Amsterdam, The Netherlands. 2018:1-11.
8. Dou Q, Chen H, Yu L, et al. Multi-level contextual 3D CNNs for false positive reduction in pulmonary nodule detection. *IEEE Trans Biomed Eng*. 2017;64(7):1558-1567.
9. Kailasam SP, Sathik MM. A novel hybrid feature extraction model for classification on pulmonary nodules. *Asian Pac J Cancer Prev*. 2019;20(2):457-468.
10. Beig N, Khorrani M, Alilou M, et al. Perinodular and intranodular radiomic features on lung CT images distinguish adenocarcinomas from granulomas. *Radiology*. 2019;290(3):783-792.
11. Kang G, Liu K, Hou B, et al. 3D multi-view convolutional neural networks for lung nodule classification. *PLoS One*. 2017;12(11):12-22.

12. Shen W, Zhou M, Yang F, et al. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. *Pattern Recognit.* 2017;61(61):663-673.
13. Wang S, Wang R, Zhang S, et al. 3D Convolutional neural network for differentiating pre-invasive lesions from invasive adenocarcinomas appearing as ground-glass nodules with diameters  $\leq 3$  cm using HRCT. *Quant Imaging Med Surg.* 2018;8(5):491-499.
14. Paul R, Hall L, Goldgof D, et al. Predicting nodule malignancy using a CNN ensemble approach. 2018 International Joint Conference on Neural Networks (IJCNN). DOI:10.1109/ijcnn.2018.8489345.
15. Kirienko M, Sollini M, Silvestri G, et al. Convolutional neural networks promising in lung cancer T-parameter assessment on baseline FDG-PET/CT. *Contrast Media Mol Imaging.* 2018 Oct 30;2018:1382309:1-6. DOI:10.1155/2018/1382309
16. Zhao T, Wang H, Li L, et al. A hybrid CNN feature model for pulmonary nodule differentiation. *J Xray Sci Technol.* 2018;26(2):171-187.
17. Hosny A, Parmar C, Coroller TP, et al. Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. *PLoS Med.* 2018;15(11):e1002711. DOI:10.1371/journal.pmed.1002711
18. Anthimopoulos M, Christodoulidis S, Ebner L, et al. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE Trans Med Imaging.* 2016;35(5):1207-1216.
19. Kim GB, Jung KH, Lee Y, et al. Comparison of shallow and deep learning methods on classifying the regional pattern of diffuse lung disease. *J Digit Imaging.* 2018;31(4):415-424.
20. Gao M, Bagci U, Lu L, et al. Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks. *Comput Methods Biomech Biomed Eng Imaging Vis.* 2016;6(1):1-6.
21. Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology.* 2017;284(2):574-582.
22. Taylor AG, Mielke C, Mongan J. Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: a retrospective study. *PLoS Med.* 2018;15(11):e1002697. DOI:10.1371/journal.pmed.1002697
23. Zech JR, Badgeley MA, Liu M, et al. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. *PLoS Med.* 2018;15(11):e1002683. DOI:10.1371/journal.pmed.1002683
24. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev Biomed Eng.* 2017 Jun 21;19:221-248.
25. Çiçek Ö, Abdulkadir A, Lienkamp SS, et al. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2016. DOI: 10.1007/978-3-319-46723-8\_49.
26. Deniz CM, Xiang S, Hallyburton RS, et al. Segmentation of the proximal femur from MR images using deep convolutional neural networks. *Sci Rep.* 2018;8(1):16485.
27. Liu S, Xie Y, Jirapatnakul A, et al. Pulmonary nodule classification in lung cancer screening with three-dimensional convolutional neural networks. *J Med Imaging.* 2017;4(4):041308. DOI:10.1117/1.JMI.4.4.041308.
28. Harmon SA, Sanford TH, Xu S, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nat Commun.* 2020 Aug 14;11(1):1-7.
29. Tang H, Zhang C, Xie X. Automatic pulmonary lobe segmentation using deep learning. In 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) 2019 Apr 8 (pp. 1225–1228). IEEE.
30. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 770–778).
31. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Br Med J.* 2015;61(12):h5527.