



Development and validation of the artificial intelligence (AI)-based diagnostic model for bronchial lumen identification

Ying Li^{1,2,3}, Xiaoxuan Zheng^{1,2,3}, Fangfang Xie^{1,2,3}, Lin Ye^{1,2,3}, Elena Bignami⁴, Yasmeen K. Tandon⁵, María Rodríguez⁶, Yun Gu^{7,8}, Jiayuan Sun^{1,2,3}

¹Department of Respiratory Endoscopy, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, China; ²Department of Respiratory and Critical Care Medicine, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, China; ³Shanghai Engineering Research Center of Respiratory Endoscopy, Shanghai, China; ⁴Anesthesiology, Critical Care and Pain Medicine Division, Department of Medicine and Surgery, University of Parma, Parma, Italy; ⁵Department of Radiology, Mayo Clinic, Rochester, MN, USA; ⁶Department of Thoracic Surgery, Clínica Universidad de Navarra, Madrid, Spain; ⁷Institute of Medical Robotics, Shanghai Jiao Tong University, Shanghai, China; ⁸Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, Shanghai, China

Contributions: (I) Conception and design: J Sun, Y Gu, Y Li; (II) Administrative support: J Sun; (III) Provision of study materials or patients: J Sun, Y Li, X Zheng, F Xie, L Ye; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: Y Li, X Zheng, F Xie, L Ye, Y Gu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jiayuan Sun. Department of Respiratory Endoscopy, Department of Respiratory and Critical Care Medicine, Shanghai Chest Hospital, Shanghai Jiao Tong University; Shanghai Engineering Research Center of Respiratory Endoscopy, 241 West Huaihai Road, Shanghai 200030, China. Email: xkyjysun@163.com; Yun Gu. Institute of Medical Robotics, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China. Email: yungu@ieee.org.

Background: Bronchoscopy is a key step in the diagnosis and treatment of respiratory diseases. However, the level of expertise varies among different bronchoscopists. Artificial intelligence (AI) may help them identify bronchial lumens. Thus, a bronchoscopy quality-control system based on AI was built to improve the performance of bronchoscopists.

Methods: This single-center observational study consecutively collected bronchoscopy videos from Shanghai Chest Hospital and segmented each video into 31 different anatomical locations to develop an AI-assisted system based on a convolutional neural network (CNN) model. We then designed a single-center trial to compare the accuracy of lumen recognition by bronchoscopists with and without the assistance of the AI system.

Results: A total of 28,441 qualified images of bronchial lumen were used to train the CNNs. In the cross-validation set, the optimal accuracy of the six models was between 91.83% and 96.62%. In the test set, the visual geometry group 16 (VGG-16) achieved optimal performance with an accuracy of 91.88%, and an area under the curve of 0.995. In the clinical evaluation, the accuracy rate of the AI system alone was 54.30% (202/372). For the identification of bronchi except for segmental bronchi, the accuracy was 82.69% (129/156). In group 1, the recognition accuracy rates of doctors A, B, a and b alone were 42.47%, 34.68%, 28.76%, and 29.57%, respectively, but increased to 57.53%, 54.57%, 54.57%, and 46.24% respectively when combined with the AI system. Similarly, in group 2, the recognition accuracy rates of doctors C, D, c, and d were 37.90%, 41.40%, 30.91%, and 33.60% respectively, but increased to 51.61%, 47.85%, 53.49%, and 54.30% respectively, when combined with the AI system. Except for doctor D, the accuracy of doctors in recognizing lumen was significantly higher with AI assistance than without AI assistance, regardless of their experience ($P < 0.001$).

Conclusions: Our AI system could better recognize bronchial lumen and reduce differences in the operation levels of different bronchoscopists. It could be used to improve the quality of everyday bronchoscopies.

Keywords: Artificial intelligence (AI); bronchoscopy; convolutional neural network (CNN); bronchial lumen

Submitted Aug 16, 2022. Accepted for publication Nov 08, 2022.

doi: 10.21037/tlcr-22-761

View this article at: <https://dx.doi.org/10.21037/tlcr-22-761>

Introduction

The diagnosis of endobronchial lesions remains a major challenge in clinical practice. Lung cancer is the most common malignant lesion that causes endobronchial lesions, and it is also the leading cause of cancer-related deaths worldwide (1,2). The overall 5-year relative survival rate of lung cancer patients is only 19% (3), while that of stage IA patients exceeds 90% (4). Unfortunately, most lung cancer patients are diagnosed at an advanced stage resulting in no significant chance of a cure with palliative treatment as the only option. For this reason, the question of how to improve the early diagnosis rate of lung cancer patients has become a hot topic in recent lung cancer research.

Bronchoscopy is a key step in the diagnosis and treatment of respiratory diseases (5). Since it was first introduced in clinical medicine in China in the 1970s, bronchoscopy has become an important tool for respiratory diseases diagnosis, treatment, and emergency rescue. However, there are still a number of unmet needs related to endoscopies in China, such as lack of specialist doctors, an uneven distribution of advanced bronchoscopic units, and poor quality control (6,7). With increasing demands for bronchoscopy and their relevance in the decision-making processes, the workload and psychological pressure placed on doctors also continue to increase. As bronchoscopy is highly dependent on doctor's clinical experience and operating skills, high-load endoscopy units can see the quality or their interventions reduced due to different factors, such as incomplete explorations of the airway, missed lesions or wrong interpretation of the findings observed.

Respiratory physicians early in their career may not be skilled in the technical aspects of bronchoscopy, or familiar with anatomic differences in airways, and physiological variations in patients which can cause challenges. The anatomy of the airway is like a labyrinth, and in performing a bronchoscopy, it is easy to get lost after entering each lumen, which can result in some segmental lumens being missed or observed repeatedly. On the other hand, respiratory physicians later in their career, can become very proficient in the bronchoscopy operation, pay too much attention to a specific lesion and ignore the observation of the rest of the lumens. To improve this variation among different specialists, there is a large number of guidelines based on expert consensus (8-10) on the optimization of bronchoscopy; however, the quality of the intervention remains very variable among doctors. Thus, for daily endoscopies, practical methods need to be developed to

enable and facilitate the implementation of these guidelines.

Artificial intelligence (AI) has attracted attention since it was first proposed in 1956, especially in the last ten years (11,12). With the advent of the big-data era, the dramatic increase in computers' computing power, and breakthroughs in algorithm research, AI has reached a period of rapid development. The combination of AI and medicine has had a profound effect on the medical system (13), especially in medical imaging (14-17), pathological examination (18-20), and endoscopy [e.g., colposcopy (21), gastroscopy (22), and colonoscopy (23)]. The combination of AI and imaging has greatly improved the accuracy and efficiency of clinicians in diagnosing various diseases. However, only limited progress has been made in the application of AI to bronchoscopy. Matava *et al.* (24) used 775 laryngoscopy and bronchoscopy videos as a data set, trained 3 types of CNNs to classify vocal cords and the tracheal airway anatomy in real-time during video laryngoscopy or bronchoscopy. Yoo *et al.* (25) used video bronchoscopy images to train an AI model through deep learning to identify anatomical locations among the carina and both the main bronchi, and the performance of the AI model was comparable to that of the most-experienced human expert.

At present, we are not aware of AI technology for bronchoscopy quality control. This research aimed to develop and validate an AI auxiliary system for bronchial lumen identification to help standardize and improve the quality of bronchoscopies. We present the following article in accordance with the TRIPOD reporting checklist (available at <https://tldr.amegroups.com/article/view/10.21037/tlcr-22-761/rc>).

Methods

This cross-sectional study consecutively collected the operation videos of bronchoscopies from the Respiratory Endoscopy Center of Shanghai Chest Hospital from June 2020 to September 2020. For each video, 31 locations, including the glottis, trachea, carina, left main bronchus (LMB), right main bronchus (RMB), left upper lobe (LUL) bronchus, left lower lobe (LLL) bronchus, right upper lobe (RUL) bronchus, truncus intermedius bronchus (TIB), upper division bronchus (UDB), lower division bronchus (lingular bronchus) (LDB), right middle lobe (RML) bronchus, right lower lobe (RLL) bronchus, left bronchus (LB) 1+2, LB3, LB4, LB5, LB6, LB8, LB9, LB10, right bronchus (RB) 1, RB2, RB3, RB4, RB5, RB6, RB7, RB8, RB9, and RB10, were observed in order.

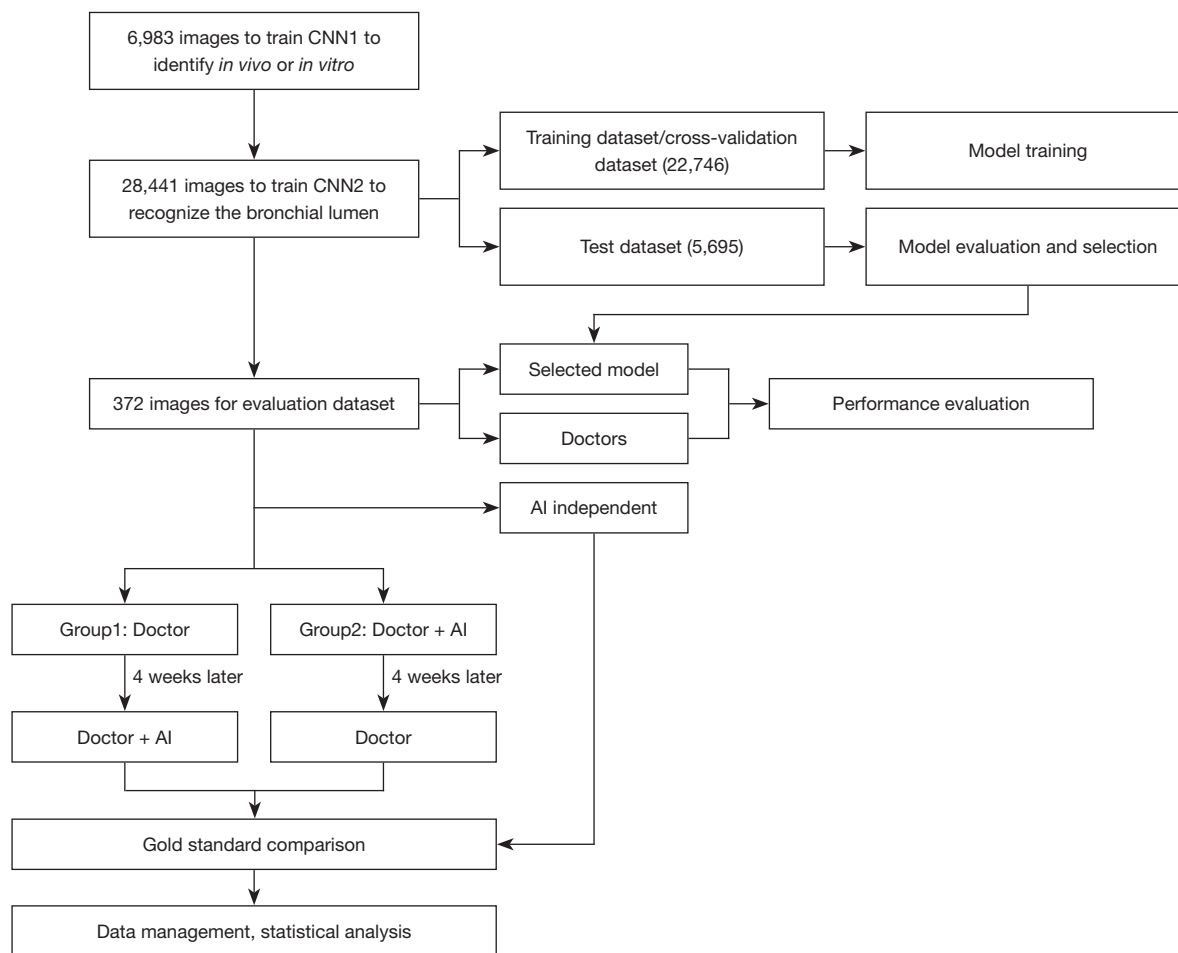


Figure 1 Image pre-processing and flow chart. CNN, convolutional neural network; AI, artificial intelligence.

The inclusion criteria were completed video of bronchoscopy operation and clear lumen images. The exclusion criteria included unclear lumen images, and lumen lesions affecting the identification of bronchial lumen. The primary endpoint was the accuracy of AI in identifying bronchial lumens, defined as all lumens AI correctly identified divided by the total lumens AI identified. The secondary endpoints were the difference in the accuracy of AI in identifying different locations, and the comparison between AI and doctors in identifying lumens.

Data set and pre-processing

A total of 342 bronchoscopy procedure videos were collected and segmented into image frames at a frequency of 3 frames per second. For the experimental images, only a squared area containing the bronchoscopic view was

cropped, and the rest of the frame was removed. All the images were resized to 1,000 by 1,000 pixels. The study design is presented in *Figure 1*.

First, *in-vitro*, *in-vivo*, and unqualified images were selected randomly and marked by 3 respiratory intervention clinicians to train the convolutional neural network (CNN) to identify whether the bronchoscope was *in vivo* or *in vitro*. This model is hereafter referred to as “CNN1.”

Second, the qualified bronchoscopy images were used to train the classification network to recognize the bronchial lumen. This model is hereafter referred to as “CNN2.” Using a ratio of 6:2:2, the data were randomly divided into the training dataset, cross-validation dataset, and test dataset. The respiratory tract was divided into 31 locations, including the glottis, trachea, carina, left and right main bronchi, lobar bronchi, and segmental bronchi (see *Figure 2*), which were independently marked by 3 professional respiratory

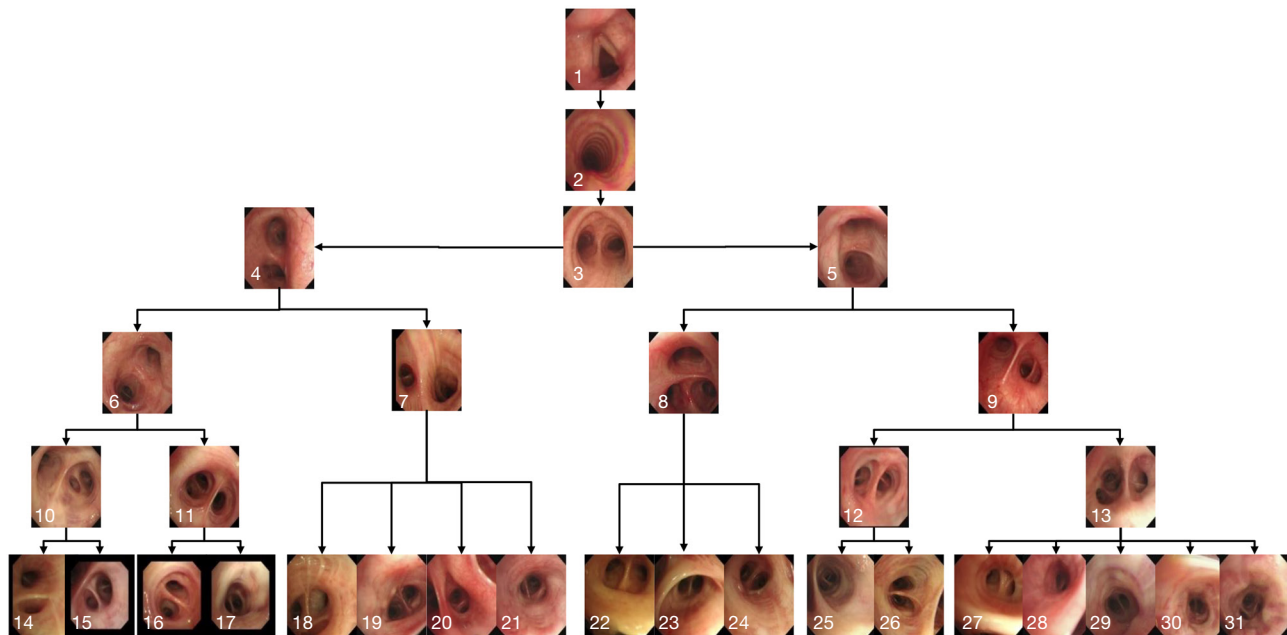


Figure 2 Typical images of 31 locations of the bronchial lumen. 1, glottis; 2, trachea; 3, carina; 4, LMB; 5, RMB; 6, LUL; 7, LLL bronchus; 8, RUL bronchus; 9, TIB; 10, UDB; 11, LDB; 12, RML; 13, RLL; 14, LB1+2; 15, LB3; 16, LB4; 17, LB5; 18, LB6; 19, LB8; 20, LB9; 21, LB10; 22, RB1; 23, RB2; 24, RB3; 25, RB4; 26, RB5; 27, RB6; 28, RB7; 29, RB8; 30, RB9; 31, RB10. LMB, left main bronchus; RMB, right main bronchus; LUL, left upper lobe; LLL, left lower lobe; RUL, right upper lobe; TIB, truncus intermedius bronchus; UDB, upper division bronchus; LDB, lower division bronchus; RML, right middle lobe; RLL, right lower lobe; LB, left bronchus; RB, right bronchus.

intervention doctors. To reduce bias, the images were only included when at least 2 doctors were in agreement. The specific marking results of the 31 locations of the bronchus are shown in *Figure 3*.

Model training and testing

For *in-vivo/in-vitro* recognition, visual geometry group 16 (VGG-16) and ResNet-50 were used to train the model. For the recognition of the 31 locations of the bronchus, InceptionV3, MobileNet, ResNet-50, VGG-16, VGG-19, and Xception were used to train the model. Among the dataset, the training set was used to train the model, the cross-validation dataset was used to determine the network structure and adjust the hyperparameters of the model, and the test dataset was used to explore the generalizability of the model. Data augmentation, including reflection, rotation, shift and early stop, were used to prevent overfitting.

Clinical research design

Evaluation data set

In total, 12 complete bronchoscopy operation videos,

showing 31 locations, were further included in evaluation data set. Again, each video was segmented into image frames. A typical image was selected for each location of each video, and thus a data set of 372 images was formed for clinical evaluation.

Clinical evaluation

The clinicians, who were required to identify the bronchial locations, comprised 4 senior doctors (hereafter referred to as doctors A, B, C, and D, with greater than 5 years of experience in interventional pulmonology) and 4 junior doctors (hereafter referred to as doctors a, b, c, and d, with less than 1 year of experience in interventional pulmonology), who were randomly divided into Group 1 and Group 2 respectively. Thus, Group 1 and Group 2 each comprised 2 senior doctors and 2 junior doctors, respectively. All of clinicians and the AI system were blind to the 372 images that comprised the evaluation dataset. For the evaluation dataset, the doctors in group 1 first identified independently the bronchoscopic images. AI was used 4 weeks later as a diagnosis aid in a similar process. Doctors in group 2 first used AI to assist in the identification of the images, and then independently, repeated the process, without AI help 4 weeks

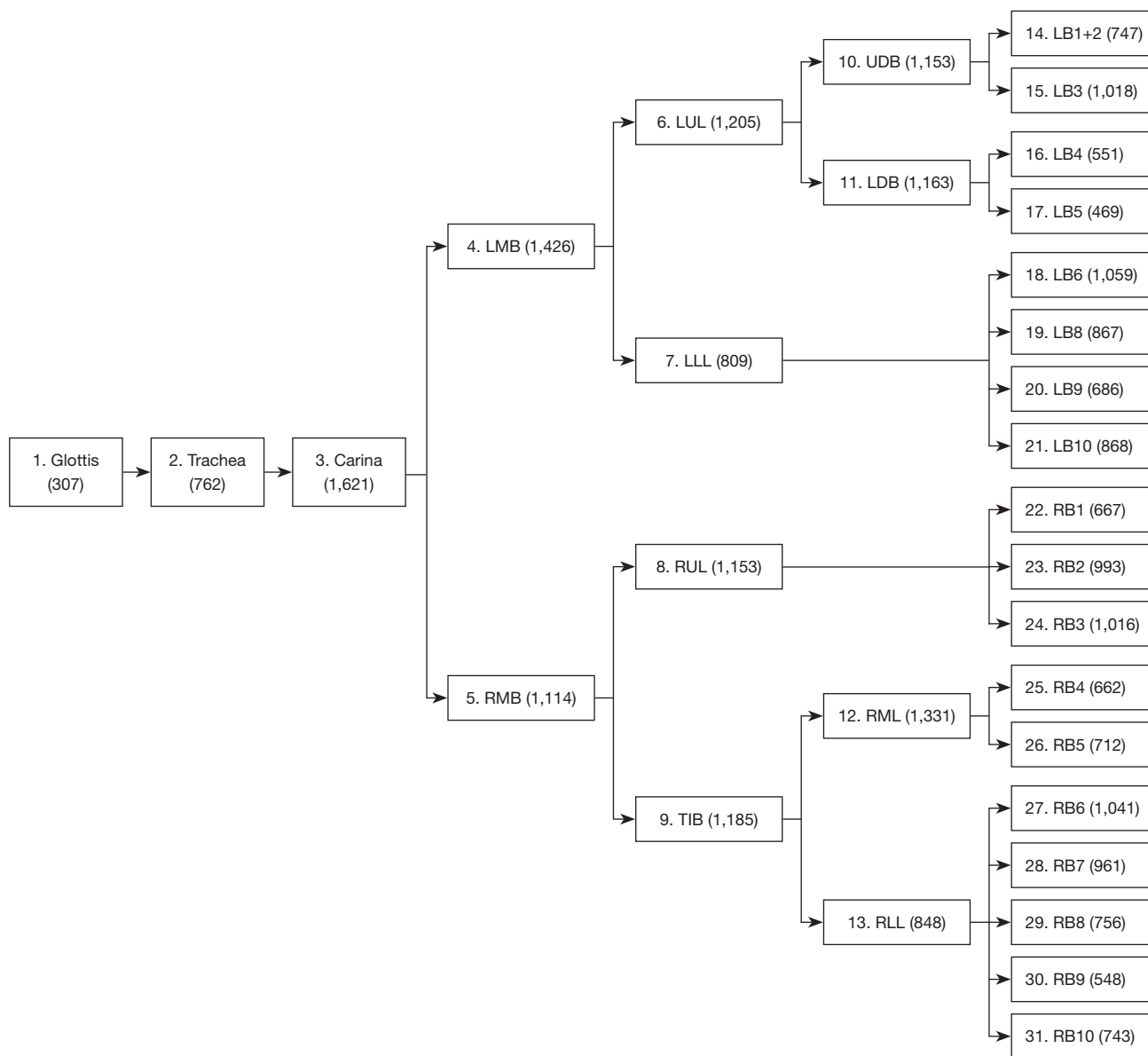


Figure 3 Number of labeled images in 31 locations of the bronchial lumen. LMB, left main bronchus; RMB, right main bronchus; LUL, left upper lobe; LLL, left lower lobe; RUL, right upper lobe; TIB, truncus intermedius bronchus; UDB, upper division bronchus; LDB, lower division bronchus; RML, right middle lobe; RLL, right lower lobe; LB, left bronchus; RB, right bronchus.

later. The difference in accuracy among the senior and junior doctors and the AI system was compared and analyzed.

Statistical analysis

The statistical analyses were performed using SPSS software (version 25.0; IBM Corp., Armonk, NY, USA). We presented standard descriptive statistics, with categorical

variables presented as percentages. The Pearson’s chi-squared test was used to compare the categorical variables. All P values were two-sided. P values <0.05 were considered statistically significant.

Ethical statement

The study was conducted in accordance with the

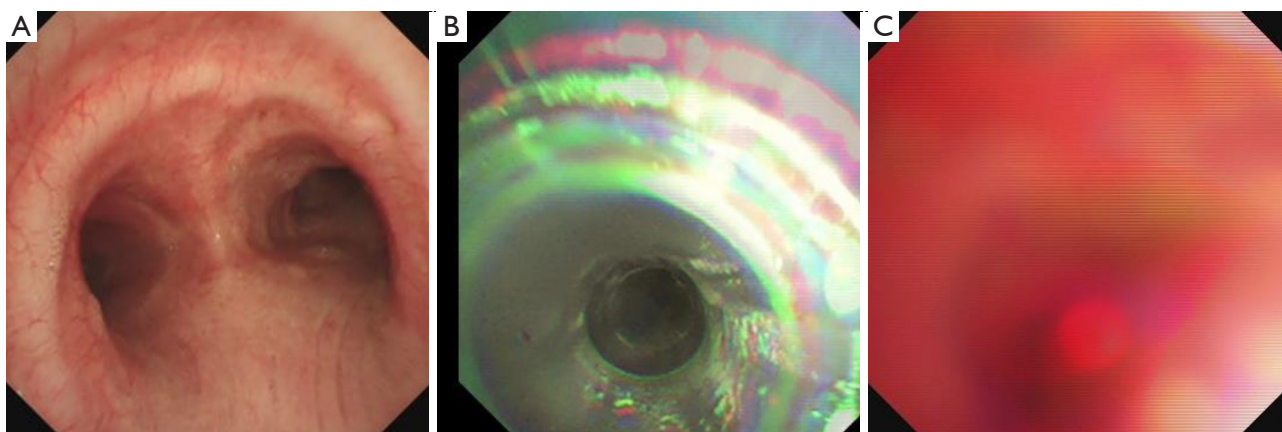


Figure 4 Representative *in-vivo* (A), *in-vitro* (B), and unqualified (C) images.

Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of Shanghai Chest Hospital (No. IS2125). Because of the retrospective nature of the study, the requirement for informed consent was waived.

Results

In-vivo/in-vitro recognition

Of all the images, 1,805 *in-vitro*, 2,708 *in-vivo*, and 2,470 unqualified bronchoscopic images were selected; representative images are shown in *Figure 4*. VGG-16 and ResNet-50 were used to train the recognition *in-vivo/in-vitro* model. The results showed that the recognition accuracy rates of VGG-16 and ResNet-50 did not differ greatly, and had values of 98.02% and 98.39%, respectively. The receiver operating characteristic (ROC) curves of the 2 accuracy rates were drawn, and the areas under the ROC curve (AUCs) of both accuracy rates reached 0.99, indicating that the models' performance was good. Performance metric changes, the ROC curves, and confusion matrixes of the ResNet50 and VGG-16 for *in-vivo/in-vitro/unqualified* image recognition are shown in *Figure 5*.

Identification of 31 locations of the bronchial lumen

A total of 28,441 qualified bronchoscopy images were selected. Multiple CNNs were used to identify 31 locations of the bronchus. The performance metric changes of each model during the training process are shown in *Figure 6*.

The ROC curve of each model is shown in *Figure 7*. The AUC of each model was above 0.96, indicating that these models had good generalizability. The identification results are shown in *Table 1*. The results showed that in the cross-validation set, the optimal accuracy of the six models was between 91.83% and 96.62%. But in the test data, the optimal accuracy rate of the VGG-16 reached 91.88%, which also had the best recognition accuracy rate among the tested models. For this reason, VGG-16 was selected as the AI system for the next test.

We also constructed a confusion matrix based on VGG-16 (see *Figure 8*), which showed the classification performance of the model for each category. The confusion matrixes for InceptionV3, MobileNet, ResNet-50, VGG-19, and Xception are shown in *Figure S1*. According to the confusion matrixes, there was significant difference in the ability of the model to recognize specific categories. Thus, we further measured the recognition accuracy of VGG-16 for each location. As *Figure 9* shows, the recognition accuracy rates of different locations ranged from 75.67% to 99.54%. Among them, the recognition accuracy rates for the glottis, carina, and LMB, RMB, RUL bronchus, and TIB were relatively high, with values >95%. However, the recognition accuracy rates for the LLL bronchus, LDB, LB4, and most basal segment bronchi were <85%, which was significantly lower than that of the other locations. The similarity between these locations may be the reason for the lower recognition accuracy.

Clinical evaluation results

The accuracy rate of the AI system alone in identifying

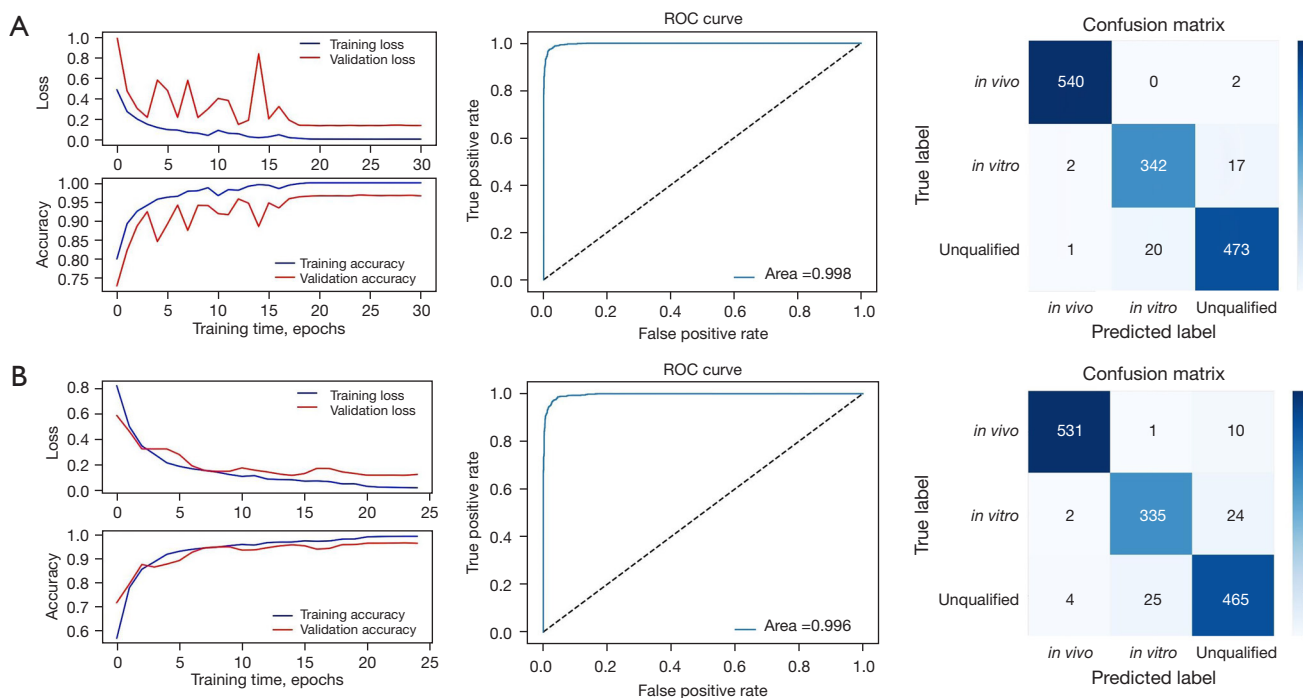


Figure 5 Performance metric changes, the ROC curves, and confusion matrixes of ResNet50 and VGG-16 for *in-vivo/in-vitro/unqualified* image recognition. (A) ResNet50, (B) VGG-16. ROC, receiver operating characteristic; VGG-16, visual geometry group 16.

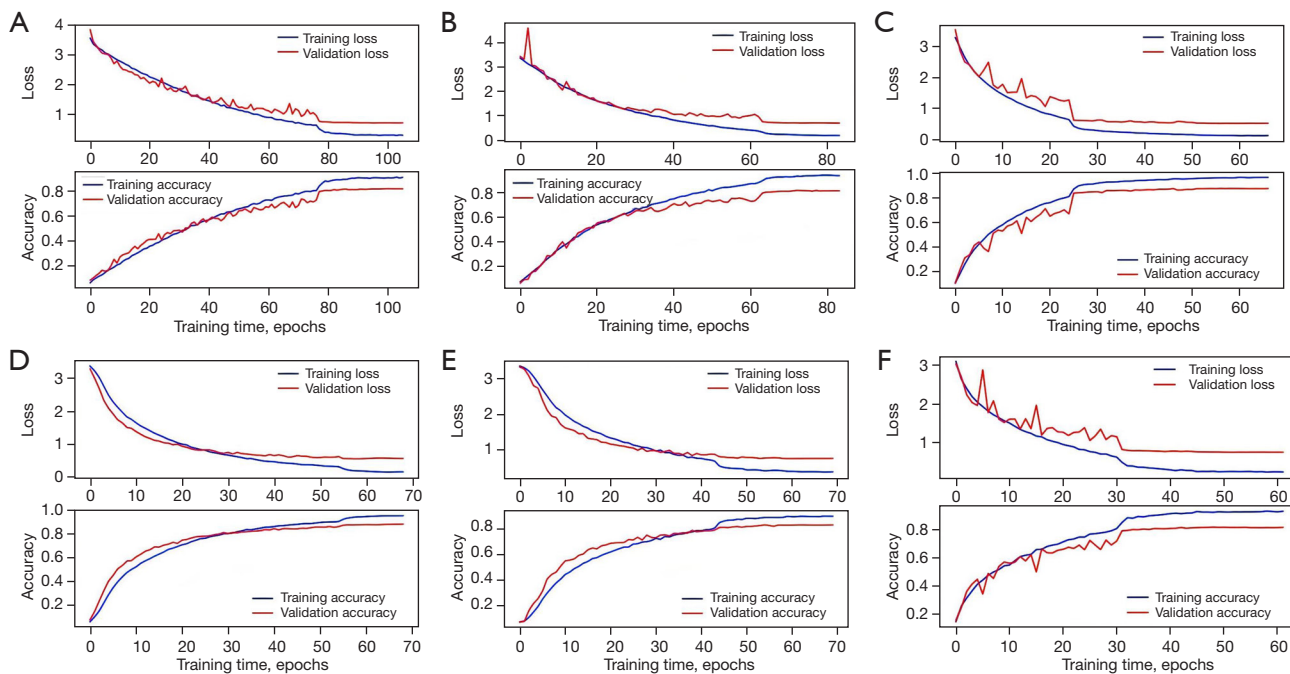


Figure 6 Performance metric changes during the training process for the 31 locations of each model. (A) InceptionV3, (B) MobileNet, (C) ResNet50, (D) VGG-16, (E) VGG-19, (F) Xception. VGG, visual geometry group.

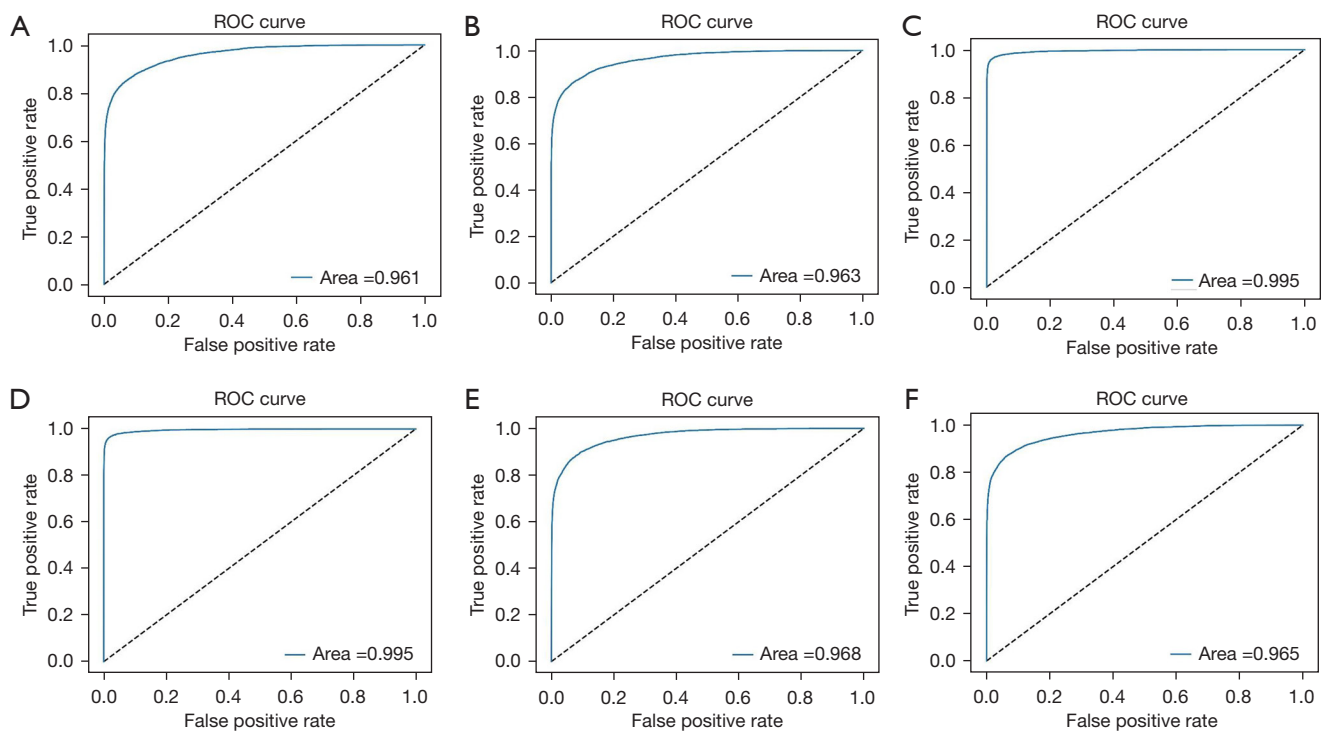


Figure 7 The AUCs of each model in the test data set. (A) InceptionV3; (B) MobileNet; (C) ResNet50; (D) VGG-16; (E) VGG-19; (F) Xception. ROC, receiver operating characteristic; AUC, the area under the ROC curve; VGG, visual geometry group.

Table 1 The optimal recognition accuracy results for the 31 locations of the bronchus.

Model	Cross-validation set accuracy	Test set accuracy
VGG-16	95.62%	91.88%
ResNet50	96.62%	91.76%
Xception	96.26%	73.28%
VGG-19	91.83%	71.77%
MobileNet	95.37%	70.97%
InceptionV3	93.42%	70.34%

VGG, visual geometry group.

the 372 bronchial images was 54.30% (202/372). For the identification of bronchi except for segmental bronchi, the accuracy was 82.69% (129/156). The recognition accuracy of segmental bronchi (B1–B10) was only 33.80% (73/216). The accuracy rates of the doctors are shown in *Table 2*. In Group 1, the accuracy rates of doctors A, B, a, and b increased from 42.47%, 34.68%, 28.76%, and 29.57%, respectively, when conducting the identification alone to 57.53%, 54.57%, 54.57%, and 46.24%, respectively,

when conducting the identification with the assistance of the AI system. In Group 2, the accuracy rates of doctors C, D, c, and d increased from 37.90%, 41.40%, 30.91%, and 33.60%, when conducting the identification alone to 51.61%, 47.85%, 53.49%, and 54.30%, respectively, when conducting the identification with the assistance of the AI system. Except for doctor D, the recognition accuracy rates of the other 7 doctors alone (i.e., without AI assistance) and their recognition accuracy rates with AI assistance differed significantly ($P < 0.001$). Thus, the AI system significantly improved the doctors' recognition of the bronchial lumen, especially among the younger doctors. In relation to the specific results for the identification of the 31 locations, the doctors accurately identified the glottis, trachea, carina, LMB, RMB, RUL bronchus, and TIB. However, in relation to the identification of the segmental bronchus, the recognition accuracy of the doctors was generally low.

Discussion

Millions of people worldwide undergo bronchoscopy every year, and high-quality endoscopy can improve the robustness of respiratory diseases diagnosis (5,26–28).

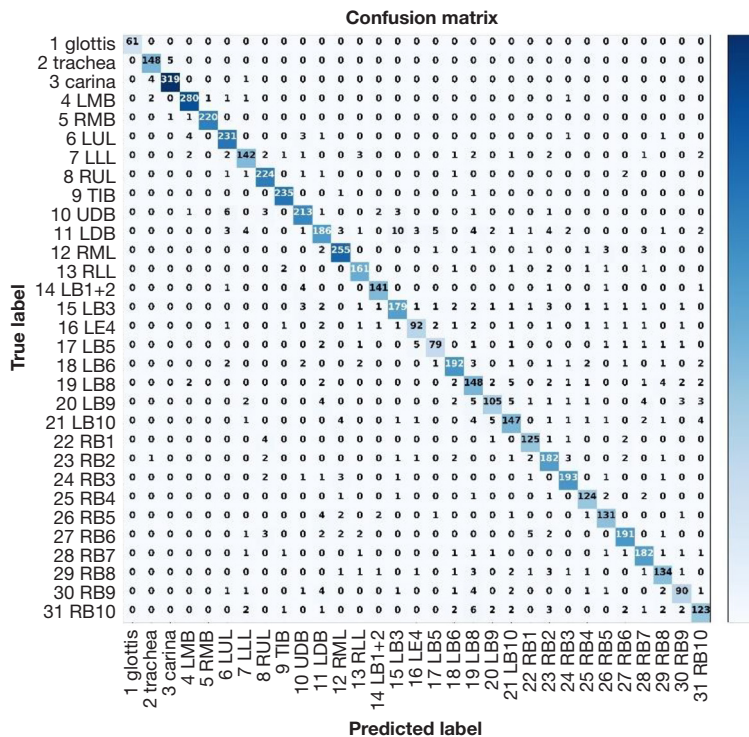


Figure 8 Confusion matrix for the classification of the 31 locations in the test data set by VGG-16. VGG-16, visual geometry group 16.

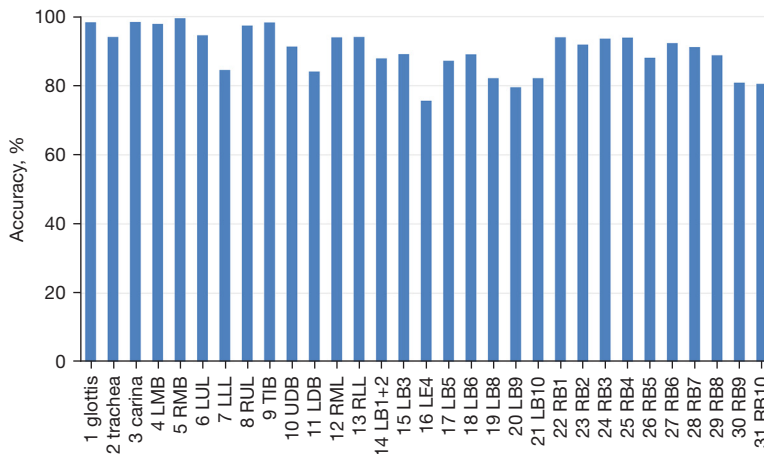


Figure 9 Recognition accuracy rates of VGG-16 for different locations in the test data set. VGG-16, visual geometry group 16.

However, there are significant differences in the level of expertise between endoscopists, which affect the detection rate of lesions. The question of how to correctly identify and judge the bronchial lumen has become an important part of bronchoscopy examinations and treatments. This is especially important when the lesion is located at the distal end of the bronchus and is invisible under a bronchoscope,

as a positioning error may directly lead to examination or treatment failure. A series of guidelines for regulating bronchoscopy operations based on expert consensus have been proposed (8-10); however, due to a lack of supervision and practical tools, these guidelines are often not well implemented, especially in developing countries.

It was previously reported that some scholars applied AI

Table 2 The identification results of the doctors

Identification results	Method		Total	P value
	Doctor	Doctor + AI		
Group 1				
A				<0.001
Identification				
Yes	158	214	372	
No	214	158	372	
Total	372	372	744	
Accuracy	42.47%	57.53%	50.00%	
B				<0.001
Identification				
Yes	129	203	332	
No	243	169	412	
Total	372	372	744	
Accuracy	34.68%	54.57%	44.62%	
a				<0.001
Identification				
Yes	107	203	310	
No	265	169	434	
Total	372	372	744	
Accuracy	28.76%	54.57%	41.67%	
b				<0.001
Identification				
Yes	110	172	282	
No	262	200	462	
Total	372	372	744	
Accuracy	29.57%	46.24%	37.90%	
Group 2				
C				<0.001
Identification				
Yes	141	192	333	
No	231	180	411	
Total	372	372	744	
Accuracy	37.90%	51.61%	44.76%	

Table 2 (continued)

Table 2 (continued)

Identification results	Method		Total	P value
	Doctor	Doctor + AI		
D				0.080
Identification				
Yes	154	178	332	
No	218	194	412	
Total	372	372	744	
Accuracy	41.40%	47.85%	44.62%	
c				<0.001
Identification				
Yes	115	199	314	
No	257	173	430	
Total	372	372	744	
Accuracy	30.91%	53.49%	42.20%	
d				<0.001
Identification				
Yes	125	202	327	
No	247	170	417	
Total	372	372	744	
Accuracy	33.60%	54.30%	43.95%	

AI, artificial intelligence.

to bronchoscopy. Tan *et al.* (29) proposed a new migration learning method on DenseNet for the recognition of lung diseases, including cancer and tuberculosis, under bronchoscopy. Using this method, the authors identified 87% of cancers, 54% of tuberculosis cases, and 91% of normal cases with good detection accuracy. Feng *et al.* (30) used fluorescent bronchoscopy pictures of 12 cases of adenocarcinoma and 11 cases of squamous cell carcinoma as a data set, and linear regression machine learning methods for classification, and ultimately achieved an accuracy of 83% for lung cancer classification. Hotta *et al.* (31) also trained CNN model to predict benign or malignant lesions based on endobronchial ultrasonography findings. These studies mainly focused on AI assisted diagnosis of bronchial lesions. However, the recognition of bronchial lumen was only limited to the recognition of vocal cords and tracheal (24), carina and both the main bronchi (25).

Our AI system was able to accurately identify the main

and lobar bronchi. AI systems could help inexperienced doctors to quickly identify the lumen and prevent missed diagnoses and misdiagnoses. For senior doctors, bronchoscopies can be standardized by AI systems, thereby improving the quality of the interventions performed.

In terms of endoscopy, Wu *et al.* combined a deep CNN with deep reinforcement learning to train a real-time quality improvement system (WINSENSE) for esophageal gastroduodenoscopy. Monitoring blind spots during the examination has become a powerful auxiliary tool to reduce the different skills of endoscopists and improve the quality of daily endoscopies (32); however, in relation to bronchoscopy, most studies have focused on the identification of the pathological specimens of transbronchial biopsies (33,34). To our knowledge, this is the first study to apply CNN training to identify bronchial anatomy.

Due to the hierarchical relationship between the

bronchial segments, the bronchial lumens are usually checked in order during bronchoscopy; thus, there is no need to actually consider the specific shape of each lumen. In addition, the segmental bronchi, especially the lingual segment and the basal segment of the lower lobes, are very similar; making it extremely difficult to identify the bronchial lumen based on images. In this study, neither junior nor senior doctors had very high accuracy rates for recognizing bronchial lumen images. This indicates that image-based bronchial lumen recognition is challenging. In relation to the recognition accuracy of the doctors for each location, the 2 groups of doctors could better identify the glottis, trachea, carina, LMB, RMB, RUL bronchus, and TIB than the remaining bronchial lumens of the 31 different locations. For these locations, the senior doctors had a slightly higher recognition accuracy rate than the junior doctors, which indicates that these locations have more obvious characteristics that facilitate their recognition. However, for the segmental bronchi B1-B10, the recognition accuracy rates of both the senior and junior doctors declined steeply, and the accuracy rates of all the doctors were between 2.31% and 20.37%. These results demonstrate that these locations are difficult to recognize based on images alone. Conversely, the AI system had a better ability to recognize the lumen, especially the segmental bronchus. The recognition accuracy rate of the AI system was significantly higher than that of all the doctors. Thus, the AI system was able to identify lumen that the doctors could not properly identify.

An understanding of the specific shape of each lumen could increase the ability of junior doctors to identify and more quickly familiarize themselves with the bronchial structure, thereby improving their ability to recognize the lumen and reducing the time spent identifying them during bronchoscopy. With the assistance of AI, the recognition accuracy of both junior and senior doctors increased significantly. The improvement in the recognition accuracy of the junior doctors was more obvious than of the senior ones. This showed that the AI system improved doctors' ability to recognize the bronchial lumen, and enabled the junior doctors to achieve a level of accuracy similar to that of the senior doctors.

The introduction of an AI system in bronchoscopy has a number of advantages. First, it can help junior doctors to quickly familiarize themselves with bronchoscopy techniques and thus avoid missing or repeating the viewing of certain lumens during bronchoscopy. Second, it can improve the operation quality by enabling doctors to gain

a comprehensive understanding of the inspection process. Third, it can improve the detection rate of lesions, and enable the automatic identification of possible lesions.

Although the results of our study warrant further investigations, it has some limitations. First, our results were obtained using Olympus endoscopes (Olympus is a mainstream model); however, we hope that the AI system can be applied to endoscopes from other vendors. Theoretically, this could be achieved by using the transfer learning method without additional algorithm tuning (35). Second, this study only focused on the recognition of bronchial lumen based on images, and complete video bronchoscopy training was not carried out. Third, this was a single-center study, and the number of clinicians involved was small. Thus, we intend to train and test the AI system using more videos of bronchoscopy operations from multiple centers in the future.

Conclusions

In conclusion, this AI model is an endoscopy quality improvement system based on CNNs. The AI system was better able to recognize lumen than the junior and senior doctors alone were. Furthermore, our findings suggest that the supplemental application of the AI system could reduce the differences in the endoscopic skills of doctors with different levels of experience, and it could become a powerful auxiliary tool to improve the quality of daily endoscopic examinations. In the future, the AI system is expected to be used to monitor the blind spot rate during bronchoscopy and determine the inspection time.

Acknowledgments

The authors would like to thank Dr. Wu Weijin for his help in training the CNN models. The authors also appreciate the academic support from AME Artificial Intelligence Collaborative Group.

Funding: This work was supported by funding from the SJTU Trans-med Awards Research Grant (No. 20210101), and the Multi-Disciplinary Collaborative Clinical Research Innovation Project of Shanghai Chest Hospital (No. YJXT20190205).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://tclr>.

amegroups.com/article/view/10.21037/tlcr-22-761/rc

Data Sharing Statement: Available at <https://tlcr.amegroups.com/article/view/10.21037/tlcr-22-761/dss>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://tlcr.amegroups.com/article/view/10.21037/tlcr-22-761/coif>). MR reports personal fees from Abex, Intuitive, and AstraZeneca, outside the submitted work. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of Shanghai Chest Hospital (No. IS2125). Because of the retrospective nature of the study, the requirement for informed consent was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. *CA Cancer J Clin* 2016;66:115-32.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* 2020;70:7-30.
- Schabath MB, Cote ML. Cancer Progress and Priorities: Lung Cancer. *Cancer Epidemiol Biomarkers Prev* 2019;28:1563-79.
- Murakami S, Ito H, Tsubokawa N, et al. Prognostic value of the new IASLC/ATS/ERS classification of clinical stage IA lung adenocarcinoma. *Lung Cancer* 2015;90:199-204.
- Gasparini S, Bonifazi M. Management of endobronchial tumors. *Curr Opin Pulm Med* 2016;22:245-51.
- Lin J, Tao X, Xia W, et al. A multicenter survey of pediatric flexible bronchoscopy in western China. *Transl Pediatr* 2021;10:83-91.
- Nie XM, Cai G, Shen X, et al. Bronchoscopy in some tertiary grade A hospitals in China: two years' development. *Chin Med J (Engl)* 2012;125:2115-9.
- Jin F, Li Q, Li S, et al. Interventional Bronchoscopy for the Treatment of Malignant Central Airway Stenosis: An Expert Recommendation for China. *Respiration* 2019;97:484-94.
- Ernst A, Wahidi MM, Read CA, et al. Adult Bronchoscopy Training: Current State and Suggestions for the Future: CHEST Expert Panel Report. *Chest* 2015;148:321-32.
- Du Rand IA, Blaikley J, Booton R, et al. British Thoracic Society guideline for diagnostic flexible bronchoscopy in adults: accredited by NICE. *Thorax* 2013;68 Suppl 1:i1-i44.
- Howard J. Artificial intelligence: Implications for the future of work. *Am J Ind Med* 2019;62:917-26.
- Xiang Y, Zhao L, Liu Z, et al. Implementation of artificial intelligence in medicine: Status analysis and development suggestions. *Artif Intell Med* 2020;102:101780.
- Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;25:44-56.
- Lassau N, Estienne T, de Vomécourt P, et al. Five simultaneous artificial intelligence data challenges on ultrasound, CT, and MRI. *Diagn Interv Imaging* 2019;100:199-209.
- Masood A, Sheng B, Li P, et al. Computer-Assisted Decision Support System in Pulmonary Cancer detection and stage classification on CT images. *J Biomed Inform* 2018;79:117-28.
- Becker AS, Marcon M, Ghafoor S, et al. Deep Learning in Mammography: Diagnostic Accuracy of a Multipurpose Image Analysis Software in the Detection of Breast Cancer. *Invest Radiol* 2017;52:434-40.
- Aramendía-Vidaurreta V, Cabeza R, Villanueva A, et al. Ultrasound Image Discrimination between Benign and Malignant Adnexal Masses Based on a Neural Network Approach. *Ultrasound Med Biol* 2016;42:742-52.
- Bera K, Schalper KA, Rimm DL, et al. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol* 2019;16:703-15.
- Awan R, Sirinukunwattana K, Epstein D, et al. Glandular Morphometrics for Objective Grading of Colorectal Adenocarcinoma Histology Images. *Sci Rep* 2017;7:16852.
- Komura D, Ishikawa S. Machine Learning Methods for Histopathological Image Analysis. *Comput Struct Biotechnol J* 2018;16:34-42.
- Miyagi Y, Takehara K, Miyake T. Application of deep learning to the classification of uterine cervical squamous

- epithelial lesion from colposcopy images. *Mol Clin Oncol* 2019;11:583-9.
22. Itoh T, Kawahira H, Nakashima H, et al. Deep learning analyzes *Helicobacter pylori* infection by upper gastrointestinal endoscopy images. *Endosc Int Open* 2018;6:E139-44.
 23. Urban G, Tripathi P, Alkayali T, et al. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. *Gastroenterology* 2018;155:1069-1078.e8.
 24. Matava C, Pankiv E, Raisbeck S, et al. A Convolutional Neural Network for Real Time Classification, Identification, and Labelling of Vocal Cord and Tracheal Using Laryngoscopy and Bronchoscopy Video. *J Med Syst* 2020;44:44.
 25. Yoo JY, Kang SY, Park JS, et al. Deep learning for anatomical interpretation of video bronchoscopy images. *Sci Rep* 2021;11:23765.
 26. Beaudoin EL, Chee A, Stather DR. Interventional pulmonology: an update for internal medicine physicians. *Minerva Med* 2014;105:197-209.
 27. d'Hooghe J, Alvarez Martinez H, Pietersen PI, et al. ERS International Congress, Madrid, 2019: highlights from the Clinical Techniques, Imaging and Endoscopy Assembly. *ERJ Open Res* 2020;6:e00116-2020.
 28. Rutter MD, Senore C, Bisschops R, et al. The European Society of Gastrointestinal Endoscopy Quality Improvement Initiative: developing performance measures. *Endoscopy* 2016;48:81-9.
 29. Tan T, Li Z, Liu H, et al. Optimize Transfer Learning for Lung Diseases in Bronchoscopy Using a New Concept: Sequential Fine-Tuning. *IEEE J Transl Eng Health Med* 2018;6:1800808.
 30. Feng PH, Lin YT, Lo CM. A machine learning texture model for classifying lung cancer subtypes using preliminary bronchoscopic findings. *Med Phys* 2018;45:5509-14.
 31. Hotta T, Kurimoto N, Shiratsuki Y, et al. Deep learning-based diagnosis from endobronchial ultrasonography images of pulmonary lesions. *Sci Rep* 2022;12:13710.
 32. Wu L, Zhang J, Zhou W, et al. Randomised controlled trial of WISENSE, a real-time quality improving system for monitoring blind spots during esophagogastroduodenoscopy. *Gut* 2019;68:2161-9.
 33. Asfahan S, Elhence P, Dutt N, et al. Digital-Rapid On-site Examination in Endobronchial Ultrasound-Guided Transbronchial Needle Aspiration (DEBUT): a proof of concept study for the application of artificial intelligence in the bronchoscopy suite. *Eur Respir J* 2021;58:2100915.
 34. Zaizen Y, Kanahori Y, Ishijima S, et al. Deep-Learning-Aided Detection of Mycobacteria in Pathology Specimens Increases the Sensitivity in Early Diagnosis of Pulmonary Tuberculosis Compared with Bacteriology Tests. *Diagnostics (Basel)* 2022;12:709.
 35. Chang P, Grinband J, Weinberg BD, et al. Deep-Learning Convolutional Neural Networks Accurately Classify Genetic Mutations in Gliomas. *AJNR Am J Neuroradiol* 2018;39:1201-7.
- (English Language Editor: L. Huleatt)

Cite this article as: Li Y, Zheng X, Xie F, Ye L, Bignami E, Tandon YK, Rodríguez M, Gu Y, Sun J. Development and validation of the artificial intelligence (AI)-based diagnostic model for bronchial lumen identification. *Transl Lung Cancer Res* 2022;11(11):2261-2274. doi: 10.21037/tlcr-22-761

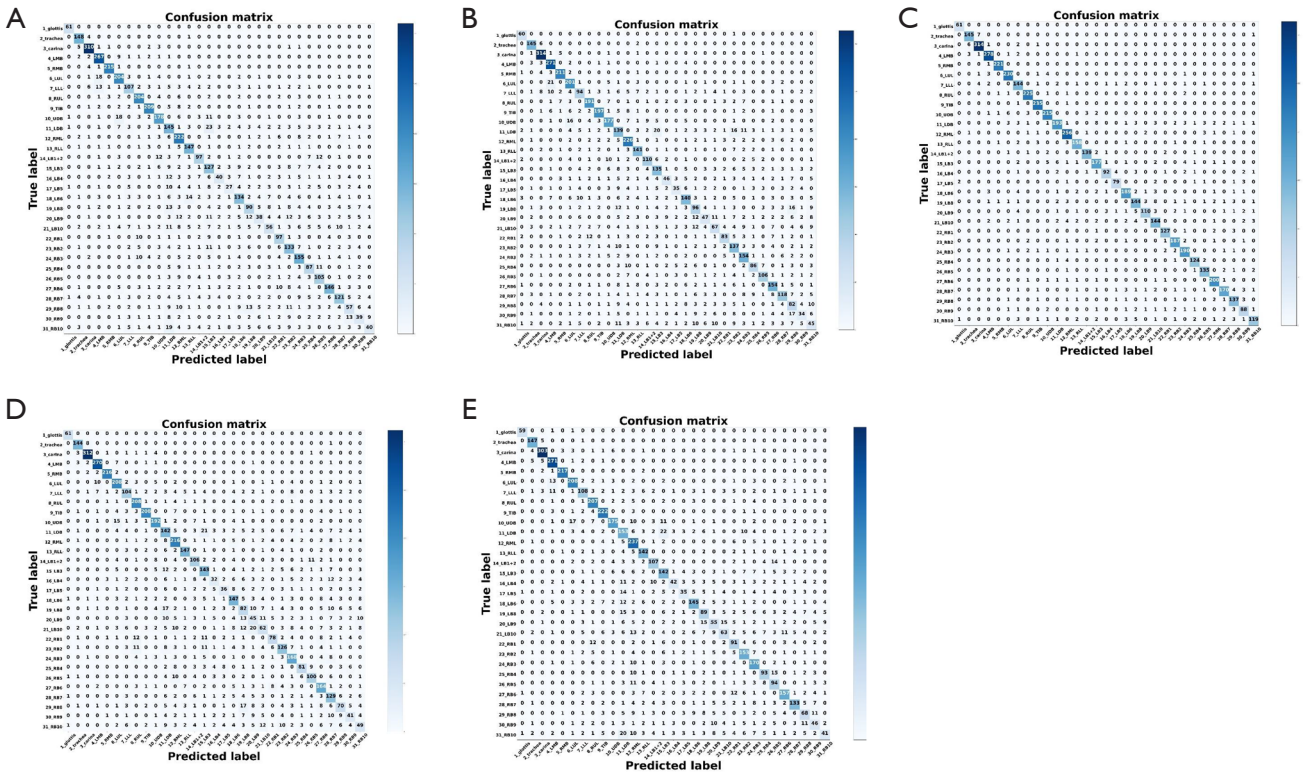


Figure S1 Confusion matrixes for the classification of the 31 locations in the test data set by the models. (A) InceptionV3, (B) MobileNet, (C) ResNet50, (D) VGG-19, (E) Xception. VGG-19, visual geometry group 19.