

Received 27 May 2018; revised 1 August 2018; accepted 3 August 2018. Date of publication 16 August 2018; date of current version 30 August 2018.

Digital Object Identifier 10.1109/JTEHM.2018.2865787

Optimize Transfer Learning for Lung Diseases in Bronchoscopy Using a New Concept: Sequential Fine-Tuning

TAO TAN^{1,2}, ZHANG LI³, (Member, IEEE), HAIXIA LIU⁴, FARHAD G. ZANJANI⁵,
QUCHANG OUYANG⁶, YULING TANG⁷, ZHEYU HU⁶, AND QIANG LI⁸

¹Department of Biomedical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands

²ScreenPoint Medical, 6512 AB Nijmegen, The Netherlands

³College of Aerospace Science and Engineering, National University of Defense Technology, Changsha 410073, China

⁴School Of Computer Science, University of Nottingham Malaysia Campus, 43500 Semenyih, Malaysia

⁵Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands

⁶Hunan Cancer Hospital, The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha 410000, China

⁷First Hospital of Changsha City, Changsha 410000, China

⁸Department of Respiratory Medicine, Shanghai East Hospital, Tongji University School of Medicine, Shanghai 200120, China

(Tao Tan and Zhang Li contributed equally to this work.) CORRESPONDING AUTHORS: Y. TANG (tyl71523@sina.com), Z. HU (huzheyu@hnszly.com), AND Q. LI (liqressh@hotmail.com)

This work was supported in part by the Hunan Provincial Natural Science Foundation of China, under Grant 2018sk2124, in part by the Health and Family Planning Commission of Hunan Province Foundation under Grant B20180393, and Grant C20180386 and in part by the Changsha City Technology Program under Grant kq1701008.

ABSTRACT Bronchoscopy inspection, as a follow-up procedure next to the radiological imaging, plays a key role in the diagnosis and treatment design for lung disease patients. When performing bronchoscopy, doctors have to make a decision immediately whether to perform a biopsy. Because biopsies may cause uncontrollable and life-threatening bleeding of the lung tissue, thus doctors need to be selective with biopsies. In this paper, to help doctors to be more selective on biopsies and provide a second opinion on diagnosis, we propose a computer-aided diagnosis (CAD) system for lung diseases, including cancers and tuberculosis (TB). Based on transfer learning (TL), we propose a novel TL method on the top of DenseNet: sequential fine-tuning (SFT). Compared with traditional fine-tuning (FT) methods, our method achieves the best performance. In a data set of recruited 81 normal cases, 76 TB cases and 277 lung cancer cases, SFT provided an overall accuracy of 82% while other traditional TL methods achieved an accuracy from 70% to 74%. The detection accuracy of SFT for cancers, TB, and normal cases are 87%, 54%, and 91%, respectively. This indicates that the CAD system has the potential to improve lung disease diagnosis accuracy in bronchoscopy and it may be used to be more selective with biopsies.

INDEX TERMS Bronchoscopy, lung cancer, tuberculosis, DenseNet, deep learning, sequential fine-tuning, computer-aided diagnosis, transfer learning.

I. INTRODUCTION

Lung cancer is also called bronchogenic carcinoma, because about 95% of primary pulmonary cancers originate from bronchial mucosa. Lung cancer is the most deadly cancer, with a five-year survival rate of 18.1% (based on 2007-2013 SEER database [1]). In 2014, it was estimated that there were 527,228 people living with bronchogenic carcinoma in the United States.¹ In China, lung cancer is the most common cancer and the leading cause of cancer death, especially for men in urban areas [2]. There were 546,259 tracheal,

bronchus, and lung (TBL) cancer deaths, about one third of the 1,639,646 deaths on a global scale in 2013 [3]. Another serious health problem in developing countries derived from the lung is tuberculosis (TB). China accounts for more than 10% of the global TB burden. Currently, Chinese government aims to suppress the TB prevalence from 390 per 100,000 people to 163 per 100,000 people and stabilize it by 2050 [4].

Chest X-ray is a cheap and fast imaging device which are commonly used for the diagnosis of lung diseases including pneumonia, tuberculosis, emphysema and cancer. It is particularly useful for emergency case. With a very small dose of

¹<https://seer.cancer.gov/statfacts/html/lungb.html>

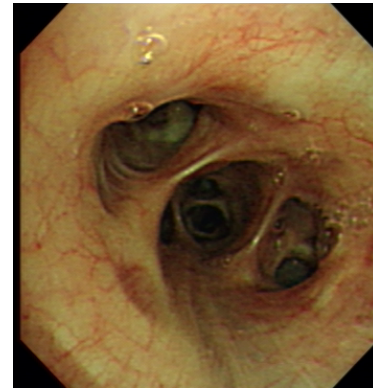
radiation, X-ray generates a 2D projection image of chest and lungs. However, due to its limitation in visualizing lung in 3D, it is gradually replaced by chest CT that offers 3D imaging. The chest CT is widely used for lung nodule detection. The downside of chest CT is its relatively higher radiation. However, in developing countries, chest X-ray is still used as a primary tool for tuberculosis screening or diagnosis. With these radiological imaging tools, radiologists are able to diagnose lung diseases or make a referral in a screening situation.

Once patients are suspected to have lung cancer or TB with X-ray or CT, bronchoscopy can be a followed-up from radiological imaging. Bronchoscopy is used as one invasive tool to directly detect diseases since 1960s [5]. Fig. 1 shows an example of normal tissue, TB and lung cancer in bronchoscopy. We can observe that for invasive TB, the lumen surface suffers from inflammatory change with hyperemia, edema and ulceration. Lung adenocarcinomas grow extraluminally and lead to lumen stenosis without affecting mucosal surface of lumen and therefore, the mucosal surface of lumen is relatively smooth. However, squamous lung cancers always form intruding nodules and are difficult to be differentiated from TB granuloma visually. Computational aid is therefore needed in bronchoscopy, especially for lesion discrimination and detection. Accurately targeting disease areas could significantly reduce the biopsy trauma and increase diagnostic accuracy [6].

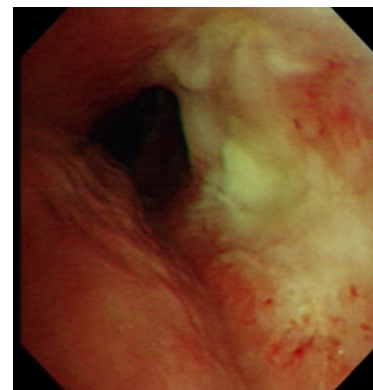
One typical computer-aided diagnosis (CAD) technique for bronchoscopy is so called virtual bronchoscopy (VB) [7], [8]. VB is normally created from CT scans and used for guiding the bronchoscopy to locate lesions [9], [10]. Several techniques, such as segmentation [11], [12], registration [13]–[16] and tracking [17]–[19], has been introduced to the VB to facilitate the guiding process. Despite the guiding, VB also improves diagnostic accuracy for peripheral lesions compared with traditional bronchoscopy [9].

Due to the limitation of traditional bronchoscopy in detecting small lesions (e.g. a few cells thick), autofluorescence bronchoscopy (AFB) [20]–[22] and narrow band imaging (NBI) [23] are introduced. These new imaging techniques improved the sensitivity [24] or specificity [23] for the early stage cancer detection. Although AFB and NBI have their advantages for lung cancer diagnosis, the traditional bronchoscopy is still the most used technique in daily clinical routine practice and bronchoscopic biopsy is a cornerstone for the pathological diagnosis. However, bronchoscopic biopsy may cause life-threatening bleeding during the operation. Therefore, it is very necessary to be selective on the bronchoscopic biopsies. To improve the diagnostic accuracy with less trauma in bronchoscopic biopsies, CADs technique can play a role.

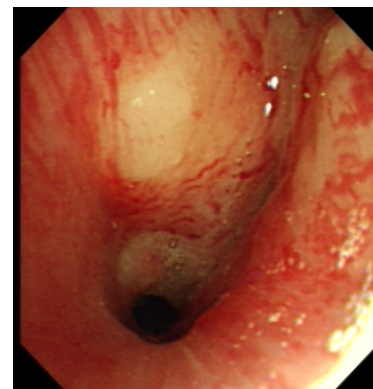
Although CADs were widely studied in other medical imaging area [25]–[31], to the best of our knowledge, CAD system has not properly studied for the traditional bronchoscopy yet. In this study, we are the first group to develop a CAD system to classify cancer, TB and normal tissues in



(a)



(b)



(c)

FIGURE 1. A bronchoscopy image of a normal (a), TB (b) and cancer (c) case.

traditional bronchoscopy. To boost classification performance, we applied the latest deep learning techniques in this study. To deal with the limited number of labeled data, we are the first to propose a novel transfer learning concept: sequential fine-tuning.

II. METHOD

A. CONVOLUTIONAL NEURAL NETWORKS

Convolutional neural networks (CNNs) [32] is a powerful tool for automatically classifying 2D or 3D image patches (input). It usually contains several pairs of a convolution layer

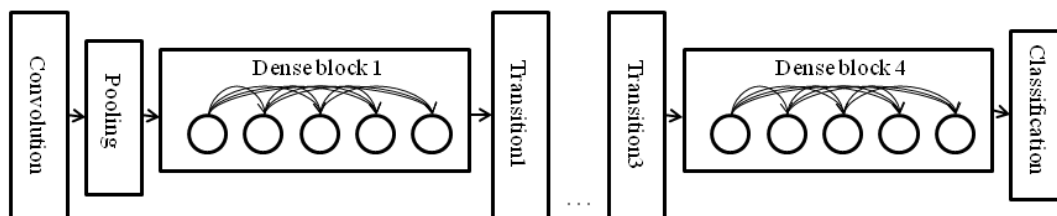


FIGURE 2. Demonstration of DenseNet 121 used in our system.

and a pooling layer. The intermediate outputs of these layers are fully connected to a multi-layer perceptron neural network. Recently, new tricks, including dropout [33], batch normalization [34] and resnet block [35], have been introduced. The purpose of dropout is to solve over-fitting that is caused by co-adaptations during training. Batch normalization helps to accelerate the training of deep networks by normalizing activations. It achieves the same accuracy with 14 times fewer training steps [34], and outperformed the original model. Resnet block is proposed in [35], where the authors find that identity shortcut connections and identity after addition activations are important for smoothing information propagation. They also design 1000-layer deep networks and obtain a better classification accuracy. Their novel deep networks are designed for creating a direct path for propagating information through the entire networks instead of within one residual unit and are trained easily in comparison with the original ResNet architecture [36].

B. USING PRE-TRAINED NETWORKS / TRANSFER LEARNING

A big advantage of CNN is that when the number of training images increases, the performance of the networks improves. Training a deep learning model requires a large amount of labeled data. However, in many medical image classification cases, the number of labeled data is limited for training. Transfer learning has been proposed [37] to effectively tackle this problem. Transfer learning literally means that experience gained from one domain can be transferred to other domains. From CNN perspective, it means that the parameters trained on one type of dataset can be reused for a new type of dataset. Practically, the first n layers of a trained CNN model could be copied to the first n layers of a new model. The remaining layers of the new model are initialized randomly and trained according to the new task [38]. Afterwards, to train the networks from the new dataset, either we only allow the parameters from the fully connected layers of the networks to be tuned, or we fine-tune more layers or even the whole network layers. It is also possible to keep the first convolutional layer fixed, as this layer is often used for edge extraction which is common for all kinds of problems. Since not all parameters are retrained or trained from scratch, the transfer learning is beneficial to problems with a small labeled dataset which is common in medical imaging field.

Transfer-learning have been applied on two specific CAD problems: thoraco-abdominal lymph node (LN) detection and interstitial lung disease (ILD) classification [39]. The state-of-the-art performance has been achieved on the mediastinal LN detection. This CNN model analysis can be extended to design high performance CAD systems for other medical imaging tasks. In ultrasound, transfer learning is used in fetal ultrasound [40], [41] breast cancer classification [42], prostate cancer detection [43], thyroid nodules classification [44], [45], liver fibrosis classification [46] and abdominal classification [47]. Christodoulidis *et al.* [48] pre-train networks on six public texture datasets and further fine-tune the network architecture on the lung tissue data. The resulting conversational features are fused with the original knowledge, which is then compressed back to the network. Compared to the same network without using transfer learning, their proposed method improves the performance by 2%. In [49], experiments are conducted to answer the research question “Can the use of pre-trained deep CNNs with sufficient fine-tuning eliminate the need for training a deep CNN from scratch?”. They conclude that deeply fine-tuned CNNs are useful for analyzing medical images and they perform as well as fully trained CNNs. When the training data is limited, the fine-tuned CNNs even outperformed the fully trained ones.

C. OUR SYSTEM

In this work, we took a pre-trained DenseNet as our pre-trained model. Huang *et al.* [50] propose DenseNet as a network architecture where each layer is directly connected to every other layer in a feed-forward fashion (within each dense block). For each layer, the feature maps of all preceding layers are treated as separate inputs whereas its own feature maps are passed on as inputs to all subsequent layers. In their work, this connectivity pattern yields state-of-the-art accuracies. On the large scale ILSVRC 2012 (ImageNet) dataset, DenseNet achieves a similar accuracy as ResNet by using less than half the amount of parameters and roughly half the number of FLOPs (floating-point operations per second). Fig. 2 shows the architecture of this work. The number 121 corresponds to the number of layers with trainable weights (excluding batch normalization layer). The additional 5 layers include the initial 7×7 convolutional layer, 3 transitional layers and a fully connected layer.

Inspired by previous work on transfer learning, in this paper we proposed a novel way of performing transfer learning. Instead of only fine-tuning the fully connected (FC) layers or fine-tuning the whole networks, we made best use of limited data meanwhile also had the power of all layers of the networks. We sequentially fine-tuned a pre-trained networks from FC layers to layers before. For a following epoch or a following set of epochs, we additionally allowed fine-tuning s layers prior to layers which were already fine-tuned in previous epochs. For example, we used n epochs for training and our networks consisted of m layers. For each sequential step, we performed x epochs. In total, to train the whole networks, we needed to perform n/x steps of SFT and for each step, we allowed extra $m*x/n$ layers to be trained. n , x and m were parameters to set.

Since the percentage of each class was different during training, to eliminate the effect of the unbalanced data, we used weighted-cross-entropy as the cost function to update the parameters of our networks. To obtain the final label or class of each sample, we just assigned the label or class where the corresponding node in the last layer gave the highest likelihood value. For training, we limited the number of epochs (n) to be 150 and set x to be 5. We resized all input images to the size of $224*224$ to fit the pre-trained model.

III. MATERIALS

A total of 434 patients who were suspected to have lung diseases by CT/X-ray images were enrolled in this study. All patients were processed at the Department of Respiration in the First Hospital of Changsha City, China, from January 2016 to November 2017. Inclusion criteria included: 1) image-suspected lung/bronchus diseases; 2) aged between 18 and 70 years old; 3) liver, renal and blood test with neutrophil count > 2.0 g/l, Hb > 9 g/l, platelet count > 100 g/l, AST and ALT > 0.5 ULN, TBIL < 1.5 ULN, and Cr < 1.0 ULN. The exclusion criteria included: 1) patients with immune-deficiency or organ-transplantation history; 2) patients with severe heart disease or heart abnormalities, such as cardiac infarction, or severe cardiac arrhythmia. This study was approved by the Ethics Committee of the First Hospital of Changsha City. Informed consent was obtained from each patient before study. Basic demographic and clinical information, including age, sex, radiological images and treatment history were recorded.

Before bronchoscopy was performed, patients were prohibited to eat and drink for at least 12 hours. 5-10 minutes before performing bronchoscopy, patients received 2 percent lidocaine (by high pressure pump) plus localized infiltrating anesthesia. Some received additional conscious sedation or general anesthesia. During the operation, flexible biferopic bronchoscopy (Olympus BF-260) was inserted to nasal cavity, glottis and bronchus. Computer workstation was configured to receive bronchoscopy images. Once abnormality of suspect was detected visually, the area was captured by

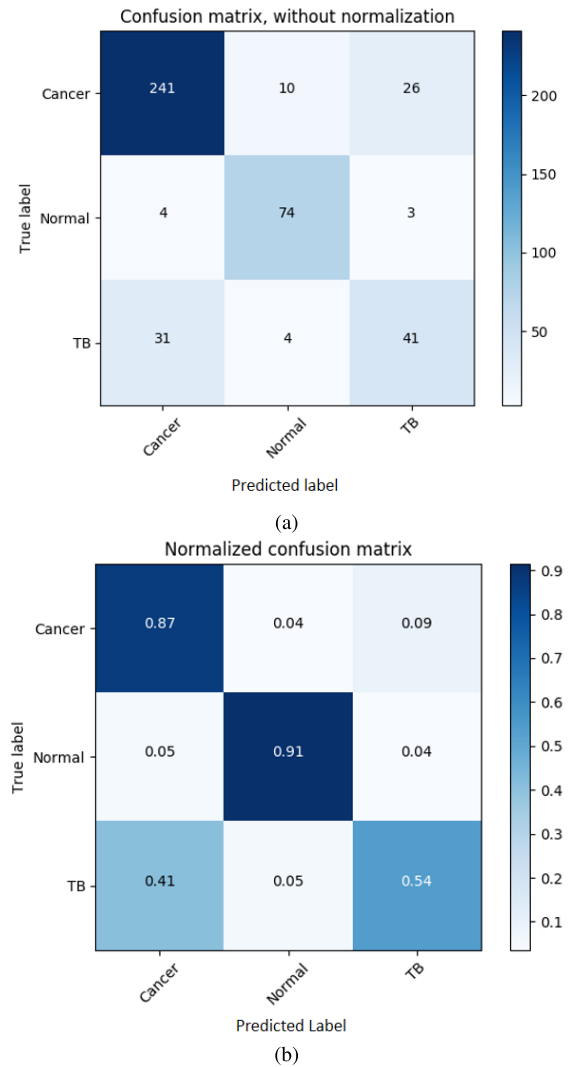


FIGURE 3. The confusion matrix and the normalized confusion matrix of our proposed method using sequential fine-tuning. (a) Confusion matrix. (b) Normalized confusion matrix.

a camera from a high-definition television (HDTV) system and saved as JPG or BMP files (319 by 302 pixels).

Pathological test was the gold standard for the diagnosis of malignant/premalignant airway disease. Therefore, specimens from bronchial biopsy were obtained in all cases of this study. Specimens for pathologic diagnosis were obtained from the following ways: brushing from the lesion or bronchial washings, fine needle aspiration biopsy, and forcep biopsy from visible tumor or TB lesions or suspected regions. Histological diagnosis was made by experienced pathologists. Two independent pathologists firstly made their diagnosis individually. If their diagnosis were inconsistent, another arbitrator pathologist would make the decision. Such histological results were used as ground truth of this study. According to pathological confirmation, among recruited 434 patients, 81 cases were diagnosed as healthy, 76 were diagnosed as TB, and 277 were diagnosed as lung

TABLE 1. Patient number statistics.

Male	Female	Age ≤ 30s	Age = 40s	Age = 50s	Age = 60s	Age ≥ 70s
272	162	35	52	122	160	65

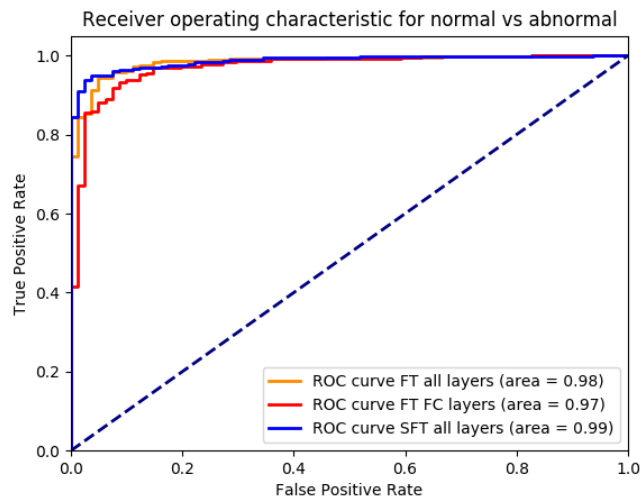


FIGURE 4. ROC curve of binary classification of abnormal (TB+cancer, positives) and normal (negatives) with fine-tuning all layers together (orange), only fine-tuning the fully connected layers (red), our proposed method using sequential fine-tuning (blue), respectively.

cancer patients. Table 1 summarized patient distribution in terms of age and gender.

IV. EXPERIMENTS AND RESULTS

A. EXPERIMENTS

In this work, considering the limited number of samples, in order to obtain an unbiased evaluation of the classification performance, a 2-fold cross-validation was employed to evaluate the performance of our method. Specifically, the input dataset was randomly divided into 2 equal parts, where one part was left for testing, and the other part was split again for training (70%) and validation (30%) to avoid the bias. The best classifier based on the validation set was used for testing. Such a procedure was repeated 2 times with a different part used for testing. We pooled the results from both two parts and evaluated the performance measurements.

Since we aimed to solve a three-class classification problem, the measurements of accuracy (ACC) and confusion matrix were used for the evaluation purpose. We also investigated two-class (binary) classification problems such as abnormal versus normal cases, TB versus cancer cases and non-cancer cases versus cancer cases. The receiver operating characteristic (ROC) analysis and the area under the ROC curve (AUC) were used for the evaluating the two-class classification performance.

B. RESULTS

The overall three-class classification accuracy of the three methods was 73.7%, 70.2% and 82.0% for from fine-tuning

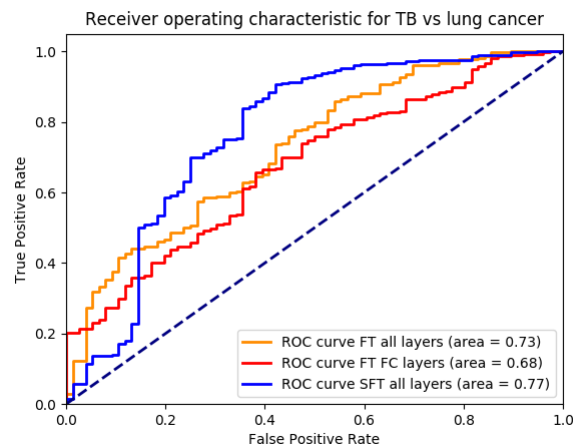


FIGURE 5. ROC curve of binary situations of TB (negatives) and cancer (positives) with fine-tuning all layers together (orange), only fine-tuning the fully connected layers (red), our proposed method using sequential fine-tuning (blue), respectively.

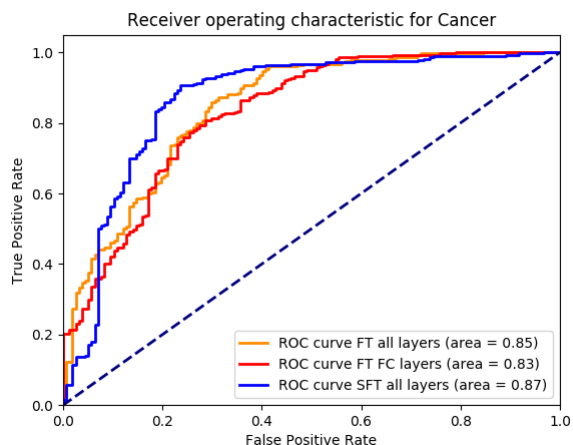
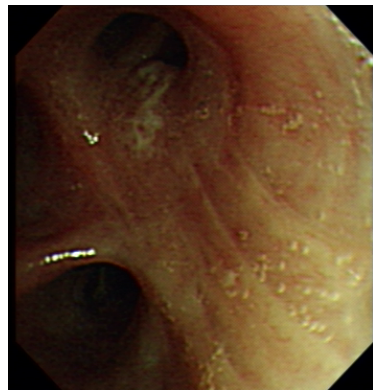


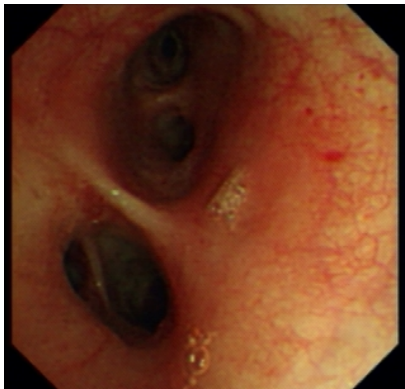
FIGURE 6. ROC curve of binary situations of non-cancer (negatives) and cancer (positives) with fine-tuning all layers together (orange), only fine-tuning the fully connected layers (red), our proposed method using sequential fine-tuning (blue), respectively.

all layer together, fine-tuning the fully connected layers and our proposed method (sequential fine-tuning), respectively. Our proposed method achieved the most accurate result. Fig. 3(a) and Fig. 3(b) showed the confusion matrix and the normalized confusion matrix from our proposed method (sequential fine-tuning).

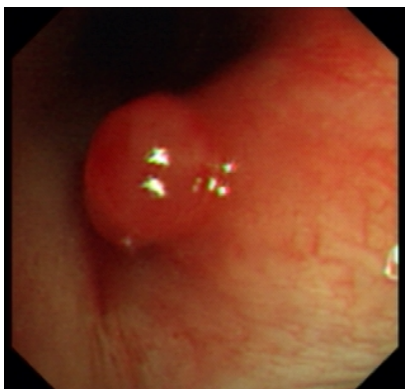
Fig. 4 showed ROC for the binary classification between normal cases and abnormal cases (TB+cancer) from fine-tuning all layer together, fine-tuning the fully connected layers and our proposed method. The AUC was 0.98, 0.97 and 0.99, respectively.



(a)



(b)



(c)

FIGURE 7. A cancer case was misclassified as a normal case (a), a normal case was misclassified as a cancer case (b), a TB case was misclassified as a cancer case(c).

Fig. 5 showed the ROC for the binary classification between TB cases and cancer cases from fine-tuning all layers together, fine-tuning the fully connected layers and our proposed method. The AUC was 0.73, 0.68 and 0.77, respectively.

Fig. 6 showed the ROC for the binary situation between non-cancer cases and cancer cases from fine-tuning all layers together, fine-tuning the fully connected layers and our proposed method. The AUC was 0.85, 0.83 and 0.87, respectively.

TABLE 2. Performance measures included three-class classification accuracy (ACC), AUC for problem 1 (P1) of abnormal versus normal cases, problem 2 (P2) of cancer versus TB cases and problem 3 (P3) of non-cancer versus cancer cases from different methods.

method	ACC	AUC P1	AUC P2	AUC P3
FT all layers	0.74	0.98	0.73	0.85
FT FC layers	0.70	0.97	0.68	0.83
SFT all layers	0.82	0.99	0.77	0.87

Fig. 7 showed examples of misclassified cases. Fig. 7a indicated a cancer case that had pale mucosa and yellow secretion and it was mis-classified as a normal case by CAD, while Fig. 7b indicated a normal case that had smooth red mucosa and was mis-classified as a cancer case by CAD. Fig. 7c. indicated a TB case that had a round nodule with smooth mucosa and was mis-classified as a cancer case by CAD. Mucosa color, secretions and smoothness were important features for cancer and TB discrimination. In cancer cases, the tumor mucosa is pale, rigid and has dirty secretion. The mis-classification in Fig. 7 might be due to a small dataset for training. Larger training set would extract more minor mucosa features to avoid such mis-classification.

Table 2 summarized different performance measures from different methods. Our proposed methods outperformed other compared methods regarding all measures.

V. CONCLUSION AND DISCUSSION

A computer aided diagnosis system was developed for the classification of normal, tuberculosis and lung cancer cases in bronchoscopy. In the system, a deep learning model based on pre-trained DenseNet was applied. Using the sequential fine tuning, our model in combination with 2-fold-cross-validation, obtained a overall accuracy of 82.0% for a dataset of 81 normal cases, 76 tuberculosis cases and 277 lung cancer cases. The detection accuracy for cancers, TB and normal cases were 87%, 54% and 91% respectively. This indicated that the CAD system had potential to improve diagnosis and it also might be used for more selective biopsies. Furthermore, we showed that the performance of the deep-learning model was improved with our proposed sequential fine-tuning.

To our best knowledge, we were the first group to bring up the concept of sequential fine-tuning in deep learning networks and we showed the benefits of using sequential fine-tuning compared to fine-tuning all layers and fine-tuning only fully connected layers. Our explanation was that since the dataset size was relatively small, it was not reasonable to fine-tune a very large set of parameters of the whole networks at the beginning. Therefore, we chose to sequentially and gradually fine-tune more and more layers from a pre-trained model. The other benefits of doing sequential fine-tuning was that instead of fitting data to two sub-models of the DenseNet (a model with no layers fixed and a model with fully connected layers fixed), we fitted our data to more sub-models as sequentially we fixed different sets of layers.

By doing so, we had a better chance of finding a good model for the data.

We also investigated our classification power of different binary classification situations. The area under the ROC curve from the binary classification of abnormal cases and normal cases was very high (0.99). From the ROC curve, we could see that we were able to keep the sensitivity of detecting abnormal cases of our CAD system to be 1 while the specificity was 0.65. It meant that our CAD system could identify 65% normal cases without missing any abnormal cases. It had the potential to reduce the false positive rate of doctors and avoid further with biopsies of these normal patients. The area under the ROC curve from the binary classification of TB cases and cancer cases was 0.77 where there was still space to improve. Although the discrimination power was not very high, we could still triage these abnormal patients and almost 10% of TB patients were correctly identified by our CAD system without missing any cancer patients. Again, these patients would not necessarily go for biopsies. For some cases, the CAD system did not perform well. Fig. 7 showed misclassified cases. The TB case was mistakenly classified as a lung cancer case by our CAD system. This TB nodule looked very like a malignant tumor. However, for doctors, there was still one feature for the discrimination: TB surface was more smooth than cancer surface. With a larger training set, more features would be extracted automatically, and this kind of mistakes would be eliminated or suppressed. In this study, TB cases from the training set was small and thus the trained model was not good enough to capture subtle features of difficult cases.

In this study, we investigated neither the actual diagnostic performance of doctors on bronchoscopy images nor the performance of doctors with the aid of our CAD system. In the future, we would conduct a reader study to evaluate the benefits of using our CAD system. Bronchoscopy as an invasive instrument plays a key role in lung disease diagnosis and determining treatment plans for the patients. With bronchoscopy doctors can directly observe the lung tissue and diagnose the problem to some extent. The doctors needs to make a decision whether to biopsy the patients timely when performing bronchoscopy. Although mortality because of the biopesis is only between 0%-0.04% [51], bleeding during diagnostic bronchoscopy occurs in between 0.26% and 5% of cases. In the case of an massive bleeding or an acute massive iatrogenic haemoptysis after biopsy, a life-threatening situation associated with a high mortality rate can develop. Furthermore, avoiding unnecessary biopsies will also reduce the anxiety of patents. With aid of our computer system, doctors can already correctly eliminate 65%normal patients, 10% of TB patients to avoid unnecessary biopsies/risk for patients which is of great help in clinical operation. To further suppress the number of biopsies, in the future, we will investigate the possibility of boosting this CAD system for identifying specific types of lung cancers. That means more labeled data should be collected in the future. The future work may also extend our CAD system in combining with other imaging

techniques(e.g. AFB) to cover broader a range of diseases and meanwhile combining deep learning networks together with human crafted features from domain knowledge.

REFERENCES

- [1] N. Howlader *et al.*, "SEER cancer statistics review, 1975–2014," Nat. Cancer Inst., Bethesda, MD, USA, Apr. 2017. [Online]. Available: https://seer.cancer.gov/csr/1975_2014/
- [2] R. Zheng *et al.*, "Lung cancer incidence and mortality in China, 2011," *Thoracic Cancer*, vol. 7, no. 1, pp. 94–99, 2016.
- [3] C. Fitzmaurice *et al.*, "The global burden of cancer 2013," *JAMA Oncol.*, vol. 1, no. 4, pp. 505–527, 2015.
- [4] K. Xu *et al.*, "Tuberculosis in China: A longitudinal predictive model of the general population and recommendations for achieving WHO goals," *Respirology*, vol. 22, no. 7, pp. 1423–1429, 2017.
- [5] H. A. Andersen, R. S. Fontana, and E. G. Harrison, Jr., "Transbronchoscopic lung biopsy in diffuse pulmonary disease," *Diseases Chest*, vol. 48, no. 2, pp. 187–192, 1965.
- [6] V. Poletti, G. L. Casoni, C. Gurioli, J. H. Ryu, and S. Tomassetti, "Lung cryobiopsies: A paradigm shift in diagnostic bronchoscopy?" *Respirology*, vol. 19, no. 5, pp. 645–654, 2014.
- [7] D. J. Vining, K. Liu, R. H. Choplin, and E. F. Haponik, "Virtual bronchoscopy: Relationships of virtual reality endobronchial simulations to actual bronchoscopic findings," *Chest*, vol. 109, no. 2, pp. 549–553, 1996.
- [8] R. M. Summers *et al.*, "Computer-assisted detection of endobronchial lesions using virtual bronchoscopy: Application of concepts from differential geometry," in *Proc. Conf. Math. Models Med. Health Sci.*, 1997.
- [9] P. J. Reynisson *et al.*, "Navigated bronchoscopy: A technical review," *J. Bronchol. Interventional Pulmonol.*, vol. 21, no. 3, pp. 242–264, 2014.
- [10] W. E. Higgins, R. Cheirsilp, X. Zang, and P. Byrnes, "Multimodal system for the planning and guidance of bronchoscopy," *Proc. SPIE*, vol. 9415, p. 941508, Mar. 2015.
- [11] R. M. Summers, D. H. Feng, S. M. Holland, M. C. Sneller, and J. H. Shelhamer, "Virtual bronchoscopy: Segmentation method for real-time display," *Radiology*, vol. 200, no. 3, pp. 857–862, 1996.
- [12] R. Eberhardt, N. Kahn, D. Gompelmann, M. Schumann, C. P. Heussel, and F. J. F. Herth, "LungPoint—A new approach to peripheral lesions," *J. Thoracic Oncol.*, vol. 5, no. 10, pp. 1559–1563, 2010.
- [13] K. Mori *et al.*, "Tracking of a bronchoscope using epipolar geometry analysis and intensity-based image registration of real and virtual endoscopic images," *Med. Image Anal.*, vol. 6, no. 3, pp. 321–336, 2002.
- [14] I. Wegner, M. Vetter, M. Schoebinger, I. Wolf, and H.-P. Meinzer, "Development of a navigation system for endoluminal brachytherapy in human lungs," *Proc. SPIE*, vol. 6141, p. 614105, Mar. 2006.
- [15] L. Appelbaum, J. Sosna, Y. Nissenbaum, A. Benshtein, and S. N. Goldberg, "Electromagnetic navigation system for CT-guided biopsy of small lesions," *Amer. J. Roentgenol.*, vol. 196, no. 5, pp. 1194–1200, 2011.
- [16] I. Bricault, G. Ferretti, and P. Cinquin, "Registration of real and CT-derived virtual bronchoscopic images to assist transbronchial biopsy," *IEEE Trans. Med. Imag.*, vol. 17, no. 5, pp. 703–714, Oct. 1998.
- [17] S. B. Solomon, P. White, Jr., C. M. Wiener, J. B. Orens, and K. P. Wang, "Three-dimensional CT-guided bronchoscopy with a real-time electromagnetic position sensor: A comparison of two image registration methods," *Chest*, vol. 118, no. 6, pp. 1783–1787, 2000.
- [18] L. Rai, J. P. Helferty, and W. E. Higgins, "Combined video tracking and image-video registration for continuous bronchoscopic guidance," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 3, nos. 3–4, pp. 315–329, 2008.
- [19] K. Mori *et al.*, "A method for tracking the camera motion of real endoscope by epipolar geometry analysis and virtual endoscopy system," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2001, pp. 1–8.
- [20] T. Gabrecht, T. M. Glanzmann, L. Freitag, B.-C. Weber, H. van den Bergh, and G. A. Wagnières, "Optimized autofluorescence bronchoscopy using additional backscattered red light," *J. Biomed. Opt.*, vol. 12, no. 6, p. 064016, 2007.
- [21] B. Zaric *et al.*, "Advanced bronchoscopic techniques in diagnosis and staging of lung cancer," *J. Thoracic Disease*, vol. 5, no. 4, p. S359, 2013.

- [22] A. Tremblay *et al.*, “Low prevalence of high-grade lesions detected with autofluorescence bronchoscopy in the setting of lung cancer screening in the pan-canadian lung cancer screening study,” *Chest*, vol. 150, no. 5, pp. 1015–1022, 2016.
- [23] F. J. F. Herth, R. Eberhardt, D. Anantham, D. Gompelmann, M. W. Zakaria, and A. Ernst, “Narrow-band imaging bronchoscopy increases the specificity of bronchoscopic early lung cancer detection,” *J. Thoracic Oncol.*, vol. 4, no. 9, pp. 1060–1065, 2009.
- [24] J. Sun *et al.*, “The value of autofluorescence bronchoscopy combined with white light bronchoscopy compared with white light alone in the diagnosis of intraepithelial neoplasia and invasive lung cancer: A meta-analysis,” *J. Thoracic Oncol.*, vol. 6, no. 8, pp. 1336–1344, 2011.
- [25] T. Tan, B. Platel, H. Huisman, C. I. Sánchez, R. Mus, and N. Karssemeijer, “Computer-aided lesion diagnosis in automated 3-D breast ultrasound using coronal spiculation,” *IEEE Trans. Med. Imag.*, vol. 31, no. 5, pp. 1034–1042, May 2012.
- [26] T. Tan, B. Platel, R. Mus, L. Tabar, R. M. Mann, and N. Karssemeijer, “Computer-aided detection of cancer in automated 3-D breast ultrasound,” *IEEE Trans. Med. Imag.*, vol. 32, no. 9, pp. 1698–1706, Sep. 2013.
- [27] T. Tan *et al.*, “Computer-aided detection of breast cancers using Haar-like features in automated 3D breast ultrasound,” *Med. Phys.*, vol. 42, no. 4, pp. 1498–1504, 2015.
- [28] T. Tan *et al.*, “Evaluation of the effect of computer-aided classification of benign and malignant lesions on reader performance in automated three-dimensional breast ultrasound,” *Acad. Radiol.*, vol. 20, no. 11, pp. 1381–1388, 2013.
- [29] H. Liu, T. Tan, J. van Zelst, R. Mann, N. Karssemeijer, and B. Platel, “Incorporating texture features in a computer-aided breast lesion diagnosis system for automated three-dimensional breast ultrasound,” *J. Med. Imag.*, vol. 1, no. 2, p. 024501, 2014.
- [30] K. Drukker, C. A. Sennett, and M. L. Giger, “Computerized detection of breast cancer on automated breast ultrasound imaging of women with dense breasts,” *Med. Phys.*, vol. 41, no. 1, p. 012901, 2014.
- [31] E. Kozegar, M. Soryani, H. Behnam, M. Salamati, and T. Tan, “Breast cancer detection in automated 3D breast ultrasound using iso-contours and cascaded RUSBoosts,” *Ultrasonics*, vol. 79, pp. 68–80, Aug. 2017.
- [32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [33] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [34] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.
- [35] K. He, X. Zhang, S. Ren, and J. Sun, “Identity mappings in deep residual networks,” in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 630–645.
- [36] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
- [37] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
- [38] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks?” in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 3320–3328.
- [39] H.-C. Shin *et al.*, “Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,” *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016.
- [40] H. Chen *et al.*, “Standard plane localization in fetal ultrasound via domain transferred deep neural networks,” *IEEE J. Biomed. Health Inform.*, vol. 19, no. 5, pp. 1627–1636, Sep. 2015.
- [41] H. Chen *et al.*, “Automatic fetal ultrasound standard plane detection using knowledge transferred recurrent neural networks,” in *Proc. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, Oct. 2015, pp. 507–514.
- [42] M. Yap *et al.*, “Automated breast ultrasound lesions detection using convolutional neural networks,” *IEEE J. Biomed. Health Inform.*, vol. 22, no. 4, pp. 1218–1226, Jul. 2018.
- [43] S. Azizi *et al.*, “Detection and grading of prostate cancer using temporal enhanced ultrasound: Combining deep neural networks and tissue mimicking simulations,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 8, pp. 1293–1305, Aug. 2017.
- [44] J.-K. Liu *et al.*, “An assisted diagnosis system for detection of early pulmonary nodule in computed tomography images,” *J. Med. Syst.*, vol. 41, no. 2, p. 30, 2017.
- [45] J. Chi, E. Walia, P. Babyn, J. Wang, G. Groot, and M. Eramian, “Thyroid nodule classification in ultrasound images by fine-tuning deep convolutional neural network,” *J. Digit. Imag.*, vol. 30, no. 4, pp. 477–486, Aug. 2017.
- [46] D. Meng, L. Zhang, G. Cao, W. Cao, G. Zhang, and B. Hu, “Liver fibrosis classification based on transfer learning and FCNet for ultrasound images,” *IEEE Access*, vol. 5, pp. 5804–5810, 2017.
- [47] P. M. Cheng and H. S. Malhi, “Transfer learning with convolutional neural networks for classification of abdominal ultrasound images,” *J. Digit. Imag.*, vol. 30, no. 2, pp. 234–243, Apr. 2017.
- [48] S. Christodoulidis, M. Anthimopoulos, L. Ebner, A. Christe, and S. Mougialakou, “Multi-source transfer learning with convolutional neural networks for lung pattern analysis,” *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 76–84, Jan. 2017.
- [49] N. Tajbakhsh *et al.*, “Convolutional neural networks for medical image analysis: Full training or fine tuning?” *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299–1312, May 2016.
- [50] G. Huang, Z. Liu, K. Q. Weinberger, and L. van der Maaten. (2016). “Densely connected convolutional networks.” [Online]. Available: <https://arxiv.org/abs/1608.06993>
- [51] M. Bernasconi *et al.*, “Iatrogenic bleeding during flexible bronchoscopy: Risk factors, prophylactic measures and management,” *ERJ Open Res.*, vol. 3, no. 2, p. 00084, 2017.