Clinical Radiology xxx (xxxx) xxx

Contents lists available at ScienceDirect

Clinical Radiology

journal homepage: www.clinicalradiologyonline.net

Deep learning for screening of interstitial lung disease patterns in high-resolution CT images

S. Agarwala [a], M. Kale [b], D. Kumar [a], R. Swaroop [a], A. Kumar [c], A. Kumar Dhara [d,*], S. Basu Thakur [e], A. Sadhu [f], D. Nandi [a]

[a] *Department of Computer Science and Engineering, National Institute of Technology Durgapur, Durgapur, 713209, India*
[b] *Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology Kharagpur, Kharagpur, 721302, India*
[c] *School of Computer and Information Science, University of Hyderabad, Hyderabad, 500046, India*
[d] *Department of Electrical Engineering, National Institute of Technology Durgapur, Durgapur, 713209, India*
[e] *Department of Chest Medicine, Medical College Kolkata, 700073, India*
[f] *Department of Radiology, Medical College Kolkata, 700073, India*

ARTICLE INFORMATION

Article history:
Received 21 August 2019
Accepted 16 January 2020

AIM: To develop a screening tool for the detection of interstitial lung disease (ILD) patterns using a deep-learning method.

MATERIALS AND METHODS: A fully convolutional network was used for semantic segmentation of several ILD patterns. Improved segmentation of ILD patterns was achieved using multi-scale feature extraction. Dilated convolution was used to maintain the resolution of feature maps and to enlarge the receptive field. The proposed method was evaluated on a publicly available ILD database (MedGIFT) and a private clinical research database. Several metrics, such as success rate, sensitivity, and false positives per section were used for quantitative evaluation of the proposed method.

RESULTS: Sections with fibrosis and emphysema were detected with a similar success rate and sensitivity for both databases but the performance of detection was lower for consolidation compared to fibrosis and emphysema.

CONCLUSION: Automatic identification of ILD patterns in a high-resolution computed tomography (CT) image was implemented using a deep-learning framework. Creation of a pre-trained model with natural images and subsequent transfer learning using a particular database gives acceptable results.

* Guarantor and correspondent. A. Kumar Dhara, Department of Electrical Engineering, National Institute of Technology Durgapur, Durgapur, 713209, India. Tel.: +91-9432518150.

E-mail address: ashis.dhara@ee.nitdgp.ac.inAim (A. Kumar Dhara).

https://doi.org/10.1016/j.crad.2020.01.010
0009-9260/

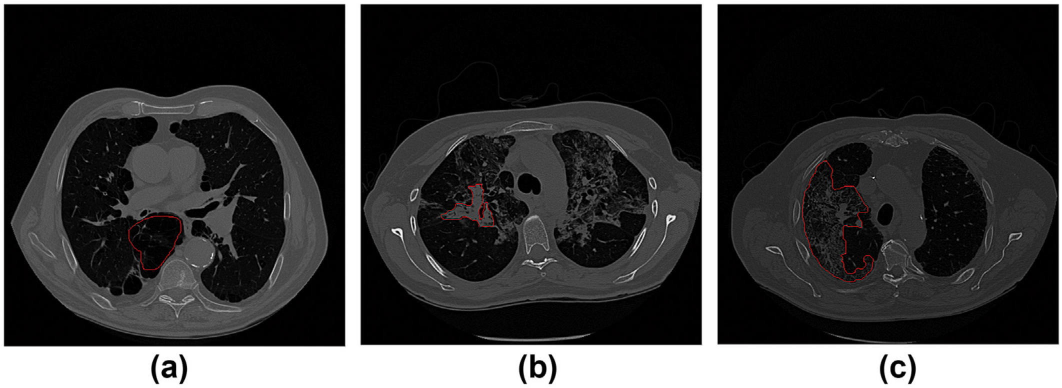

Figure 1 Representation of various types of ILD patterns: (a) emphysema, (b) consolidation, and (c) fibrosis.

Introduction

The term interstitial lung disease (ILD) is used to describe a large group of pulmonary disorders. The majority of these disorders cause progressive scarring of the lung tissues, which eventually leads to the thickening of the interstitium. ILD constitutes diverse group of different lung disorders that tend to exhibit the same clinical manifestations with ambiguous patterns. High-resolution computed tomography (HRCT) images of the lung field regions are used to detect the presence of ILD patterns. HRCT produces high-resolution images of lung alveoli, airways, interstitium, and pulmonary vasculature, which eases the process of pattern recognition for radiologists. ILD tissues are categorised into several patterns based on their appearances. The present work focuses on three common patterns: emphysema, consolidation, and fibrosis (Fig 1). It is an exhausting task to examine all the HRCT sections with the same accuracy and efficiency even for the experienced radiologists. Therefore, a screening tool would be useful to assist and provide another perspective to the radiologist to achieve a differential diagnosis.

The majority of prior research has focused on the improvement of the quality of the hand-crafted features for identification of pathological regions. Several classifiers (k-nearest neighbours (kNN),[1–4] artificial neural network,[5] and support vector machines (SVM)[1,6–12]) have been used to classify ILD patterns. Local binary patterns,[5,13] histogram-based, and scale-invariant feature transform[14,15] have also been used to classify ILD patterns. The limitation of image classification using classical techniques is the manual extraction of hand-crafted features.

Deep-learning approaches have been investigated to overcome the limitation of hand-crafted features in several medical imaging applications. Coupled with advancing technology and the use of new-generation graphics cards, convolution neural networks (CNNs) provide excellent results in the classification and segmentation of natural images. The majority of the work using deep learning focuses on patch-based classification.[16,17] In the study of Van *et al.*,[16] the weight sharing between hidden layers was investigated. The network used was a fully connected neural network, and supervised training was performed with the help of gradient descent. Anthimopoulos *et al.*,[17] classified six ILD patterns by extracting patch sizes of 32×32 and classifying them with the aid of a CNN with five convolutional layers and a filter size of 2×2. The network attained an accuracy of nearly 85.5%.

Patch-based classification is performed with the help of annotated regions of interest (ROIs) provided by radiologists. The process itself is very time-consuming and clinically less preferable. Identification of patterns at section level will be more useful for radiologists. In the study of Gao *et al.*,[18] a pre-trained model of AlexNet was used for fine-tuning purposes. The classification task was completed on whole lung section rather than patch-based classification. The input images were rescaled to fit the AlexNet architectural design and to exploit the advantage of colour images that have been artificially created using different attenuation windows.

Shin *et al.*[19] performed ILD classification on both patch and section levels. At section level, similar to[18] the main objective was to classify the presence of any ILD pattern in the HRCT lung section.

The published work using CNN has focused on the classification of ILD patterns at the patch level. Deep-learning algorithms for determining the presence of ILD patterns from HRCT sections is very important for assisting clinicians.

Table 1
The number of image sections in the MedGIFT ILD database.

ILD pattern	No. Of image sections
Consolidation	116
Emphysema	71
Fibrosis	293
Total	480

Table 2
The number of image sections in the private ILD database.

ILD pattern	No. Of Image sections
Consolidation	10
Emphysema	25
Fibrosis	25
Total	60

Please cite this article as: Agarwala S et al., Deep learning for screening of interstitial lung disease patterns in high-resolution CT images, Clinical Radiology, https://doi.org/10.1016/j.crad.2020.01.010

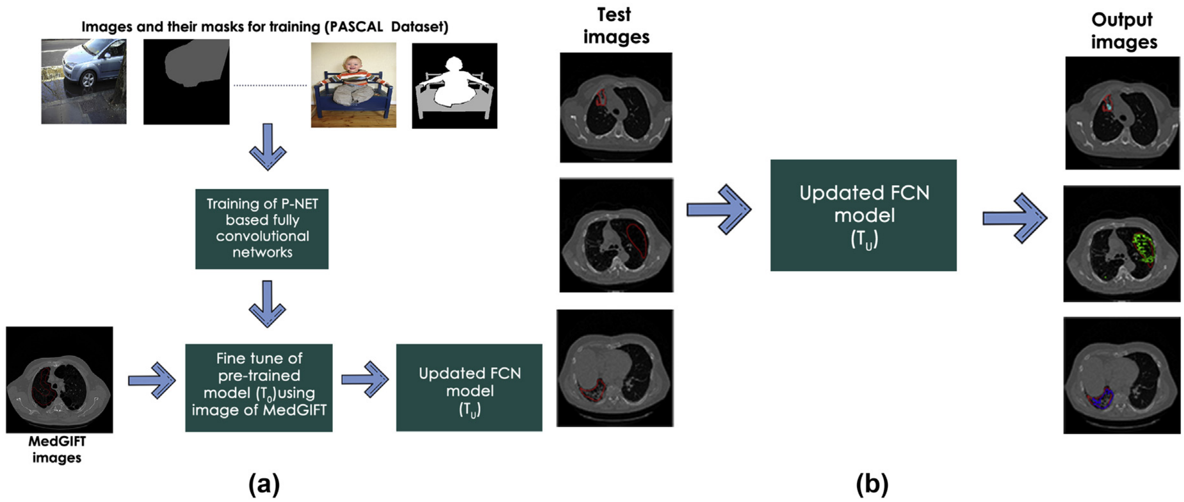

Figure 2 (a) Transfer learning of the pre-trained model, created using PASCAL VOC and (b) screening of ILD patterns (red indicates the ground truth and cyan, green, and blue are contour of consolidation, emphysema, and fibrosis, respectively).

In the present study, a fully convolutional network[20] was used to determine the presence of ILD patterns in a chest HRCT section. The proposed method does not require lung field segmentation as a pre-processing step. The pre-trained model was created using the well-known PASCAL VOC[21] database and fine-tuned using the MedGIFT ILD database[22] in order to learn database-specific features. The pre-trained model was used to overcome the learning limitation due to the insufficient images in the MedGIFT ILD database. This work will help radiologists screening ILD patterns at the section level.

Material and methods

Description of the MedGIFT ILD database

The database contains 108 annotated HRCT image series of different ILD patterns. A total of 1,946 ROIs are provided from 108 HRCT image series, which results in a total volume of 41.65 l of annotated tissue. In a single image section having more than one pattern, for every pattern a different label has been assigned. It is common practice to include more than one experienced radiologist in the annotation process. In one such database for lung nodules, the Lung Imaging Database Consortium (LIDC/IDRI), the nodules are delineated with the agreement of five radiologists. In the case of the MedGIFT ILD database, two radiologists were involved in the process of creating coarse boundaries to highlight pathology in the HRCT section to avoid the inclusion of noisy data.[22] A detailed description and the process of preparing the database can be found in22. The total number of image sections used in the present study was 480. Table 1 provides the statistical details for each pattern.

Description of the private ILD database

The database consists of 60 HRCT sections of 20 patients where the ROIs were annotated by two expert radiologists. The database contains only three patterns: consolidation, emphysema, and fibrosis. Table 2 provides the statistical details for each pattern.

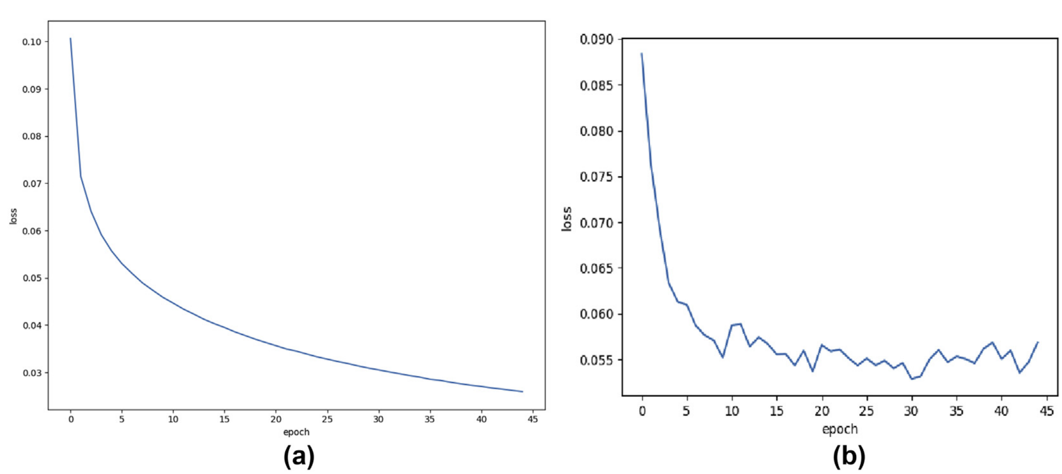

Figure 3 Representation of (a) training loss and (b) validation loss.

Table 3
Results of the proposed method for the MedGIFT ILD database.

Pattern	Success rate (%)	Sensitivity (%)	False positives per section
Consolidation	77	47	0.30
Emphysema	89	71	0.11
Fibrosis	90	78	0.12

Network description

A fully convolutional network named P-Net[20] was used for multi-scale feature extraction. Dilated convolution[23,24] was used to maintain the resolution of feature maps and to enlarge the receptive field for incorporation of larger contextual information. The architecture combined the first 13 convolution layers into five blocks. The first two blocks consist of two convolutional layers and the remaining blocks have three convolutional layers, respectively. In P-Net, there is no block for down-sampling and max-pooling. Therefore, features are extracted from a large receptive field. In this way, the resolution of the feature maps is maintained. In all the convolutional layers, the size of the convolution filter is 3×3. Features extracted from several blocks are concatenated to obtain the multi-scale feature map.

Creation of pre-trained model

The deep-learning models require a huge amount of data in the training process to get a pre-trained model. The majority of the research work on segmentation was evaluated on several natural image databases, such as PASCAL VOC[21] or COCO.[25] The benefit of working with natural colour images is that a huge amount of data can be obtained easily. It is very difficult to get a large number of annotated medical images because annotation by expert radiologists is

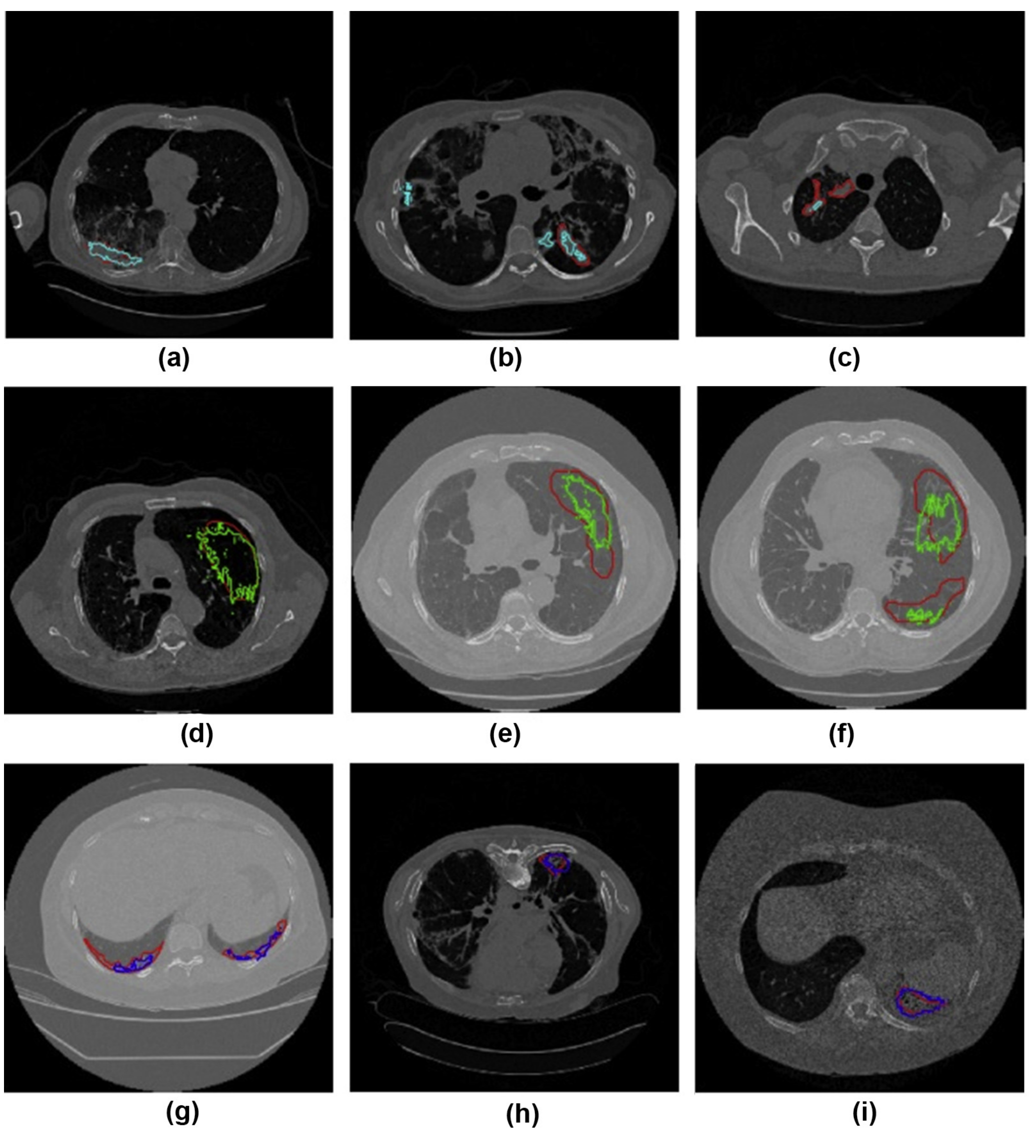

Figure 4 Segmentation of different ILD patterns using P-Net. (a–c) Segmentation for consolidation, (d–f) segmentation of emphysema, and (g–i) segmentation of fibrosis. The ground truth is shown in red. Cyan, green, and blue have been used to represent the contours for consolidation, emphysema, and fibrosis, respectively.

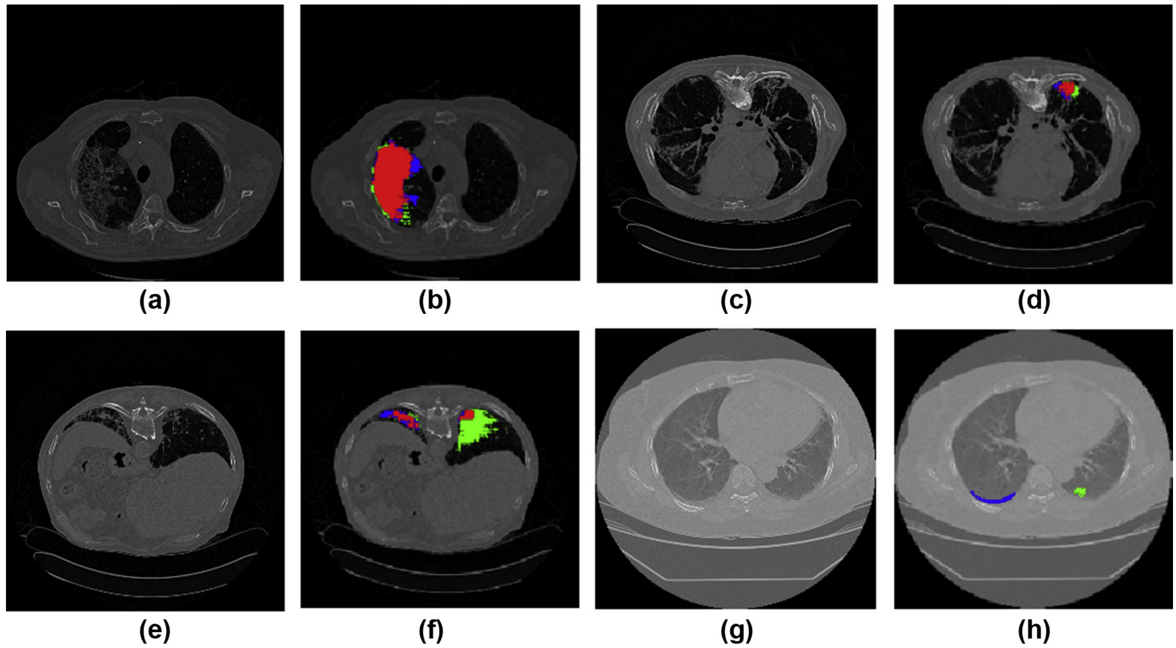

Figure 5 True positive, false positive and false negative areas in different patterns. (a, c, e, g) Test images and (b, d, f, h) corresponding output images. The true positive is shown in red whereas green and blue are used to represent false negative and false positive, respectively.

expensive. Transfer learning has been used to overcome this limitation. In the proposed work, a pre-trained network was created on a very popular natural colour image data set, the PASCAL VOC database, where a mask was provided for different objects. This database contains 21 classes. Three ILD patterns were considered in the present study. Therefore, in order to create a pre-trained model using the PASCAL VOC database (Fig 2), three classes (person, chair, and car) having the maximum number of images available, were considered. In total, 13,028 images were taken from PASCAL VOC to train the network. The images of the PASCAL VOC database were resized to 224×224. Data augmentation techniques (rotation, flip, scaling, shifting, addition of Gaussian noise, and change of contrast) were used to increase the size of the database. The increased amount of existing training data controlled for over-fitting. In the study of Adam,[26] a stochastic gradient-based optimisation technique was used to minimise the cross-entropy loss function.

Transfer learning using images of MedGIFT

Transfer learning is a process to train a network on a dataset and utilise the learned model for another dataset for learning database-specific features. The training of the model mainly updates the weight of the network. During the fine-tuning of the network, the transferred weight of few layers freezes. In the medical field, obtaining a large amount of annotated data is a challenging task. Therefore, the concept of transfer learning is very popular. Sometimes, fine-tuning of the model leads to over-fitting of the network when the dataset is limited. Therefore, there are two possibilities: either fine-tune all layers or only the last layer. It has been observed that the last layer of the network is task-specific and all the other layers are modality dependent. Transferring the weight to and fine-tuning all layers in the network is the best process.[27] In this paper, fine-tuning of the network layers was undertaken. Fig. 3 shows the plot for training loss and validation loss. It reveals that the network is not over-fitted. The learning rate for the Adam optimizer is set to 0.00001. Fine-tuning of the pre-trained network was performed in up to 45 epochs. For the MedGIFT ILD database, 50% of the sections were used for fine-tuning, 25% for validation, and the remaining 25% for testing.

Implementation

The experiment was performed in the Linux environment using NVIDIA GTX 1070 8 GB GPU having Intel core i5 (7^{th} generation) with a 3.5 GHz processor and 32 GB of RAM. The network was implemented in Python with pytorch library.

Results

The proposed method was evaluated on the publicly available MedGIFT ILD database as well as a private ILD

Table 4
Results of the proposed method for private ILD database.

Pattern	Success rate (%)	Sensitivity (%)	False positives per section
Consolidation	67	44	0.58
Emphysema	83	68	0.50
Fibrosis	86	74	0.57

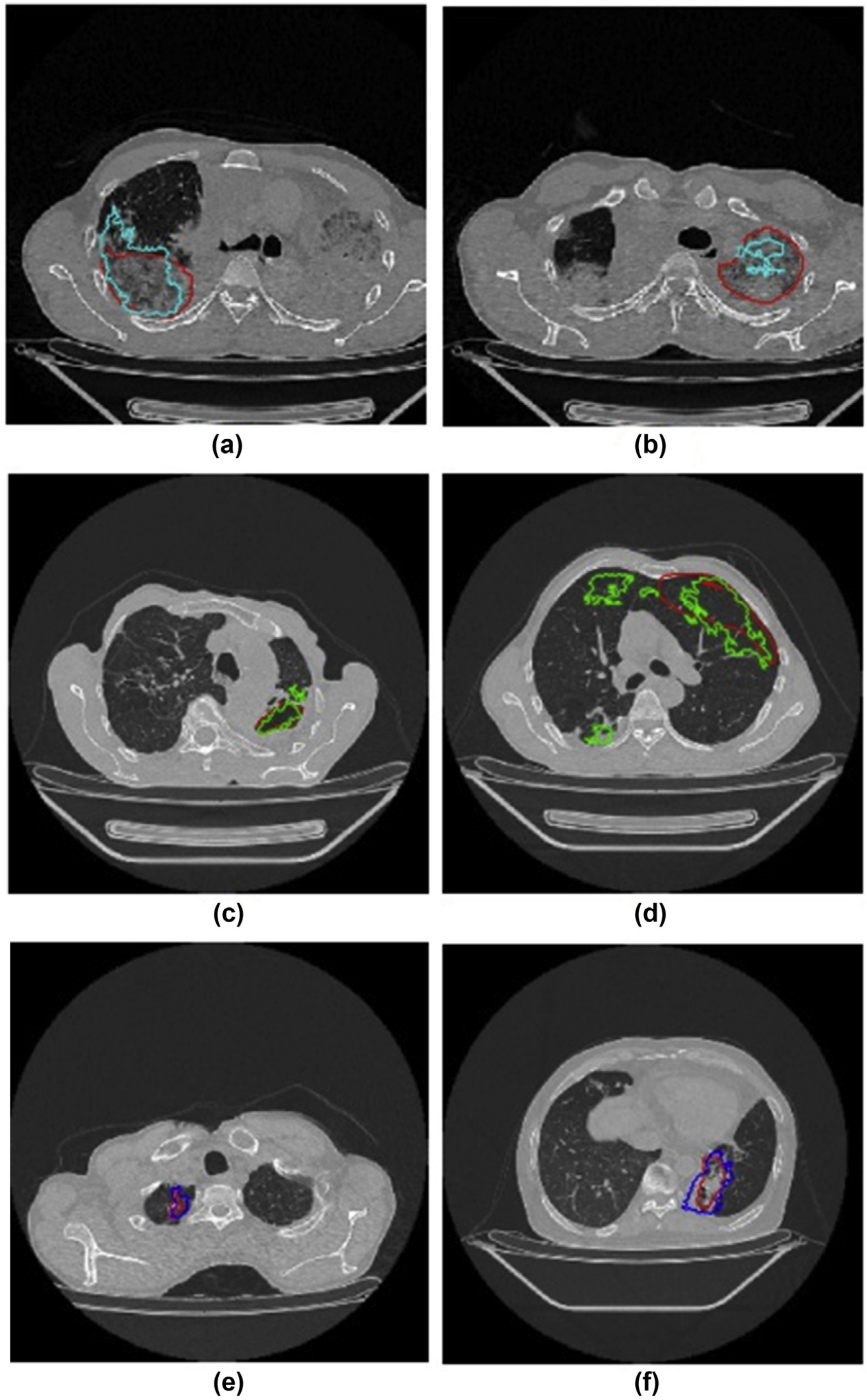

Figure 6 Segmentation of the ILD pattern result using the private database. (a) The segmentation result for consolidation, (b) the segmentation result of emphysema, and (c) fibrosis result, respectively. The ground truth is shown in red. Cyan, green, and blue have been used to represent contours for consolidation, emphysema, and fibrosis, respectively.

database. The performance of the proposed method is reported in terms of success rate, sensitivity, and number of false positives per section. The success rate was computed based on the presence of a particular pathology with respect to the ground truth for that pathology. It represents the percentage of HRCT sections in a dataset where the pathology has been detected successfully. Sensitivity is defined as

$$Sensitivity = \frac{X_{tp}}{X_{tp} + X_{fn}}$$

Where, 'X_{tp}' is true positive and 'X_{fn}' false negative.

Result for the MedGIFT ILD database

Table 3 shows the sensitivity, success rate, and number of false positives per section for different ILD patterns considered in this work. Sections with fibrosis and emphysema were detected with a similar success rate and sensitivity. The sensitivity is low because of some patterns of very small sise were missing. The performance of detection is lower for consolidation than for fibrosis and emphysema. The similarity of contrast between consolidation and the pleura leads to false negatives and affects the success rate and sensitivity.

Fig 4a–c depicts the detection of pixels with consolidation patterns. The consolidation is detected correctly in Fig 4a. In case of Fig 4b, two consolidation patterns are detected correctly, but only one region was marked as consolidation. One ROI with consolidation was missed due to its similarity with the pleura in Fig 4c. The emphysema patterns have been detected properly in Fig 4d–f; however, few pixels with emphysema were missed due to poor contrast. Fibrosis is detected efficiently in Fig 4h and i despite the low contrast between fibrosis and the image background. The performance of detection for three ILD patterns is better for those images, where there is a substantial difference between the foreground and background as shown in Fig 5g and Fig 5h.

Result for private ILD database

Database-specific fine-tuning was performed using 75% of the HRCT sections from the private database and the results were reported on the remaining 25% HRCT sections. Table 4 shows the quantitative results for the private database. Fibrosis and emphysema patterns achieved a better success rate and sensitivity compare with consolidation. The results of the private ILD database are consistent with the MedGIFT database. Fig. 6 shows the qualitative results for the private ILD database. Consolidation was detected correctly in Fig 6a, whereas a portion of the ROI of consolidation was detected in Fig 6b. Emphysema was detected correctly in Fig 6c as compared to Fig 6d. Fibrosis patches were detected efficiently as shown in Fig 6e and f.

Discussion

In the present study, a deep-learning framework for the diagnosis of ILD was implemented for automatic identification of the presence of consolidation, emphysema, and fibrosis from a HRCT section. The pre-trained model was created using natural images and fine-tuned using the ILD database. Creation of a pre-trained model with natural images and subsequent transfer learning using a particular database provided acceptable results, where annotated images are scarce. In future, active learning will be explored to improve the accuracy of screening with minimal use of training images.

Acknowlegments

The authors are thankful to Kolkata Medical College and EKO DIAGNOSTICS Kolkata for receiving useful suggestions for this research work. The research received regular grants from Visvesvaraya PhD scheme of DeitY (Department of Electronics & Information Technology), Govt. of India.

References

1. Gangeh MJ, Sørensen L, Shaker SB, *et al.* A texton-based approach for the classification of lung parenchyma in CT images. In: *International conference on medical image computing and computer-assisted intervention*. Berlin, Heidelberg: Springer; 2010. p. 595–602.
2. Park SC, Tan J, Wang X, *et al.* Computer aided detection of early interstitial lung diseases using low-dose CT images. *Phys Med Biol* 2011;**56**(4):1139–53.
3. Koratis PD, Karahaliou AN, Kazantzi AD, *et al.* Texture-based identication and characterization of interstitial pneumonia patterns in lung multidetector ct. *IEEE Trans Inform Technol Biomed* 2010;**14**(3):675–80.
4. El-Baz A, Soliman A, McClure P, *et al.* Early assessment of malignant lung nodules based on the spatial analysis of detected lung nodules. In: *2012 9th IEEE international symposium on biomedical imaging (ISBI)*. Piscataway: IEEE; 2012. p. 1463–6.
5. Jacobs C, Sánchez CI, Saur SC, *et al.* Computer aided detection of ground glass nodules in thoracic ct images using shape, intensity and context features. In: *International conference on medical image computing and computer-assisted intervention*. Berlin, Heidelberg: Springer; 2011. p. 207–14.
6. Depeursinge A, Foncubierta-Rodriguez A, Van de Ville D, *et al.* Lung texture classification using locally-oriented riesz components. *Med Image Comput Comput Assist Interv* 2011;**14**(Pt 3):231–8.
7. Depeursinge A1, Van de Ville D, Platon A, *et al.* Near-affine-invariant texture learning for lung tissue analysis using isotropic wavelet frames. *IEEE Trans Inf Technol Biomed* 2012 Jul;**16**(4):665–75.
8. Song Y, Cai W, Kim J, *et al.* A multistage discriminative model for tumor and lymph node detection in thoracic images. *IEEE Trans Med Imaging* 2012 May;**31**(5):1061–75.
9. Xu R, Hirano Y, Tachibana R, *et al.* Classification of diffuse lung disease patterns on high-resolution computed tomography by a bag of words approach. *Med Image Comput Comput Assist Interv* 2011;**14**(Pt 3):183–90.
10. Yao J, Dwyer A, Summers RM, *et al.* Computer-aided diagnosis of pulmonary infections using texture analysis and support vector machine classification. *Acad Radiol* 2011 Mar;**18**(3):306–14.
11. Bagci U, Yao J, Wu A, *et al.* Automatic detection and quantification of tree-in-bud (TIB) opacities from CT scans. *IEEE Trans Biomed Eng* 2012 Jun;**59**(6):1620–32.
12. Song Y1, Cai W, Eberl S, *et al.* Discriminative pathological context detection in thoracic images based on multi-level inference. *Med Image Comput Comput Assist Interv* 2011;**14**(Pt 3):191–8.
13. Sørensen L, Shaker SB, de Bruijne M. Quantitative analysis of pulmonary emphysema using local binary patterns. *IEEE Trans Med Imaging* 2010 Feb;**29**(2):559–69.
14. Sørensen L1, Nielsen M, Lo P, *et al.* Texture-based analysis of COPD: a data-driven approach. *IEEE Trans Med Imaging* 2012 Jan;**31**(1):70–8.
15. Song Yang, Cai Weidong, Wang Yue, *et al.* Location classification of lung nodules with optimized graph construction. In: *9th IEEE international symposium on biomedical imaging (ISBI)*. Piscataway: IEEE; 2012. p. 1439–42. 2012.
16. van Tulder Gijs, de Bruijne Marleen. Learning features for tissue classification with the classification restricted Boltzmann machine. In: *International MICCAI workshop on medical computer vision*. Berlin, Heidelberg: Springer; 2014. p. 47–58.
17. 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 May;**35**(5):1207–16.
18. Gao M, Bagci U, Lu L, *et al.* Holistic classication of ct attenuation patterns for interstitial lung diseases via deep convolutional neural networks. *Comput Methods Biomech Biomed Eng Imaging Vis* 2018;**6**(1):1–6.
19. Shin HC, Roth HR, Gao M, *et al.* Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging* 2016 May;**35**(5):1285–98.
20. Wang G, Zuluaga MA, Li W, *et al.* DeepIGeoS: a deep interactive geodesic framework for medical image segmentation. *IEEE Trans Pattern Anal Mach Intell* 2019 Jul;**41**(7):1559–72.

21. Everingham M, Ali Eslami SM, Van Gool L, *et al*. The pascal visual object classes challenge: a retrospective. *Int J Comput Vis* 2015;**111**(1):98–136.
22. Depeursinge A, Vargas A, Platon A, *et al*. Building a reference multimedia database for interstitial lung diseases. *Comput Med Imaging Graph* 2012 Apr;**36**(3):227–38.
23. Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. *arXiv Preprint arXiv:1511.07122* 2015.
24. Chen LC, Papandreou G, Kokkinos I, *et al*. Semantic image segmentation with deep convolutional nets and fully connected crfs. *arXiv Preprint arXiv:1412.7062* 2014.
25. Lin TY, Maire M, Belongie S, *et al*. Microsoft coco: Common objects in context. In: *Of European conference on computer vision*. Cham: Springer; 2014 Sep 6. p. 740–55.
26. Kingma DP, Ba J. Adam: a method for stochastic optimization. *arXiv Preprint arXiv:1412.6980* 2014.
27. Yosinski J, Clune J, Bengio Y, *et al*. How transferable are features in deep neural networks?. In: *In Advances in neural information processing systems*; 2014. p. 3320–8.

Please cite this article as: Agarwala S et al., Deep learning for screening of interstitial lung disease patterns in high-resolution CT images, Clinical Radiology, https://doi.org/10.1016/j.crad.2020.01.010