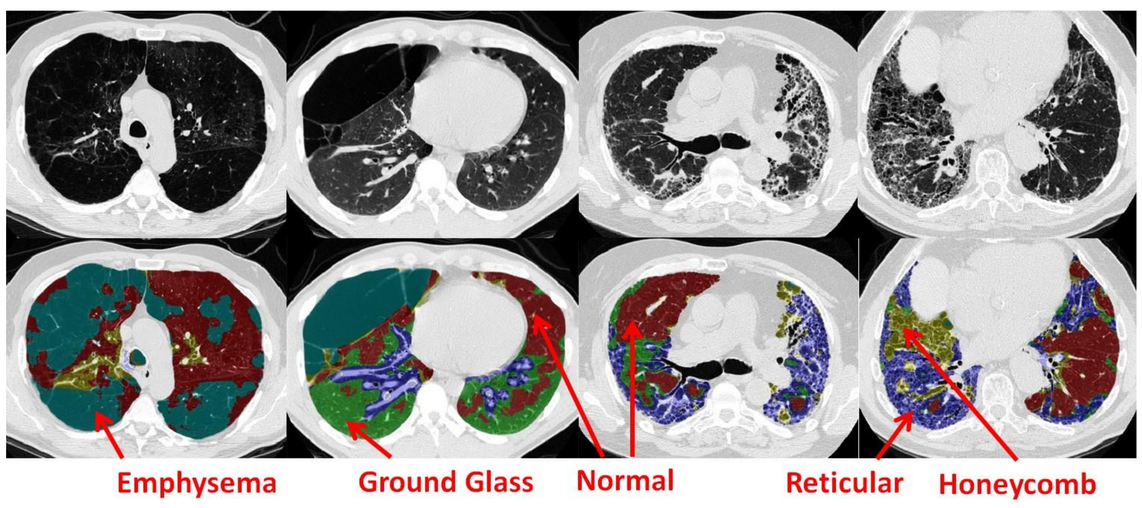
Multi-label Deep Regression and Unordered Pooling for Holistic Interstitial Lung Disease Pattern Detection

Mingchen Gao, Ziyue Xu, Le Lu, Adam P. Harrison, Ronald M. Summers, Daniel J. Mollura
Radiology and Imaging Sciences, National Institutes of Health, Bethesda, MD

Objective

- Holistically detecting interstitial lung disease (ILD) patterns from CT images is challenging yet clinically important.
- Previous solutions rely on manually provided regions of interest, and only predict one ILD label from each slice/patch.
- We propose a multi label deep CNN regression model (**MLCNN-R**) and Fisher Vector CNN (**FV-CNN**) to detect multiple ILD patterns simultaneously.



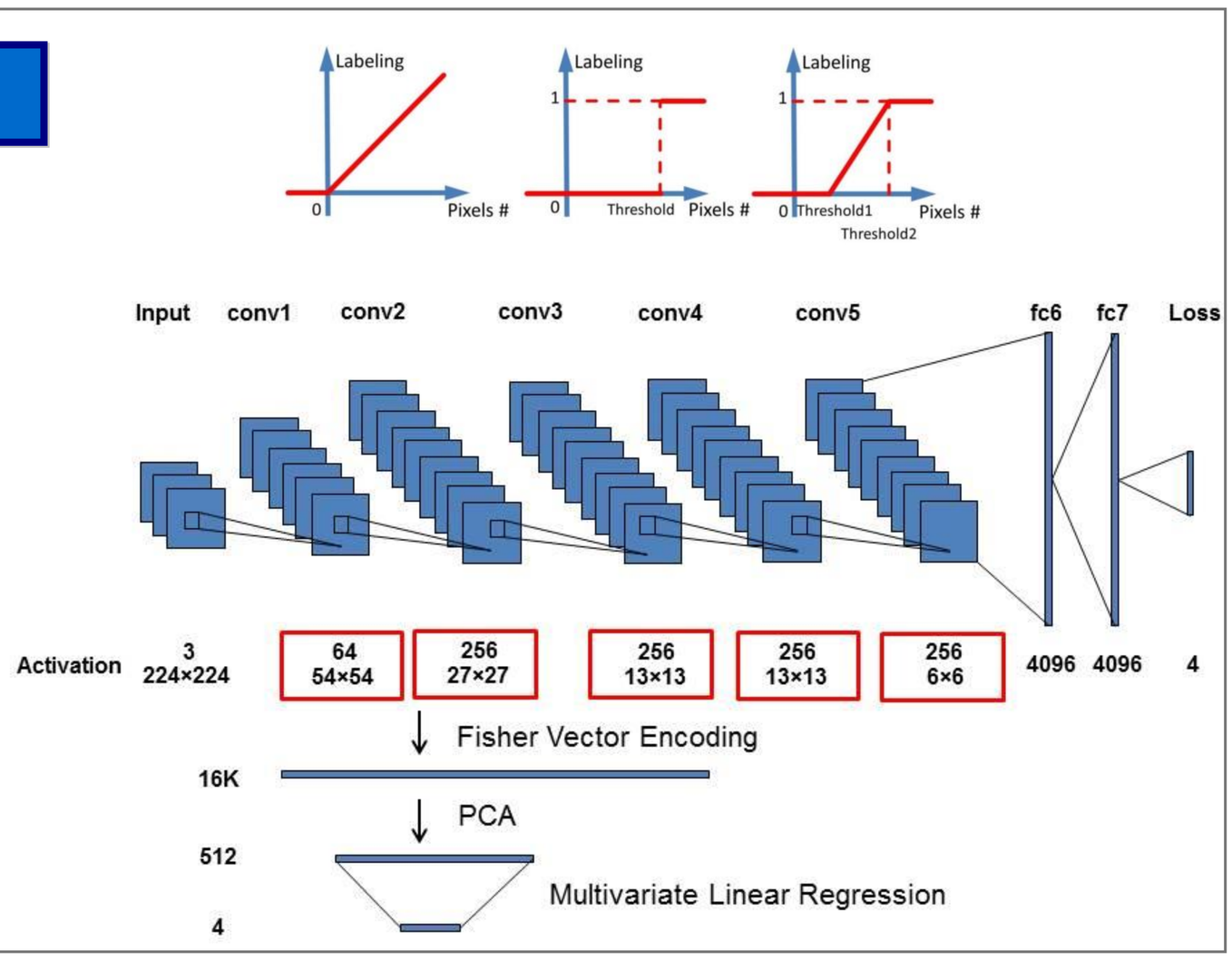
Examples of ILD patterns. Every voxel is labeled as healthy or one of the four ILDs in the LTRC dataset [1].

Method

- MLCNN-R
  - The L2 regression loss functions are minimized in an end-to-end system.

$$L(\boldsymbol{y_i}, \hat{\boldsymbol{y}}_{\boldsymbol{i}}) = \sum_{k=1}^{c} (y_{ik} - \hat{y}_{ik})^2.$$

  - Convert pixel-wise annotations to slice-wise labels, using either pixels numbers occupied per ILD class or the binary [0,1] occurring status.
- FV-CNN
  - While CNNs are powerful tools, their feature learning strategy is not invariant to the spatial locations of objects or textures within a scene.
  - We invoke Fisher Vector (FV) encoding [2] to remove the spatial configurations in activation map.

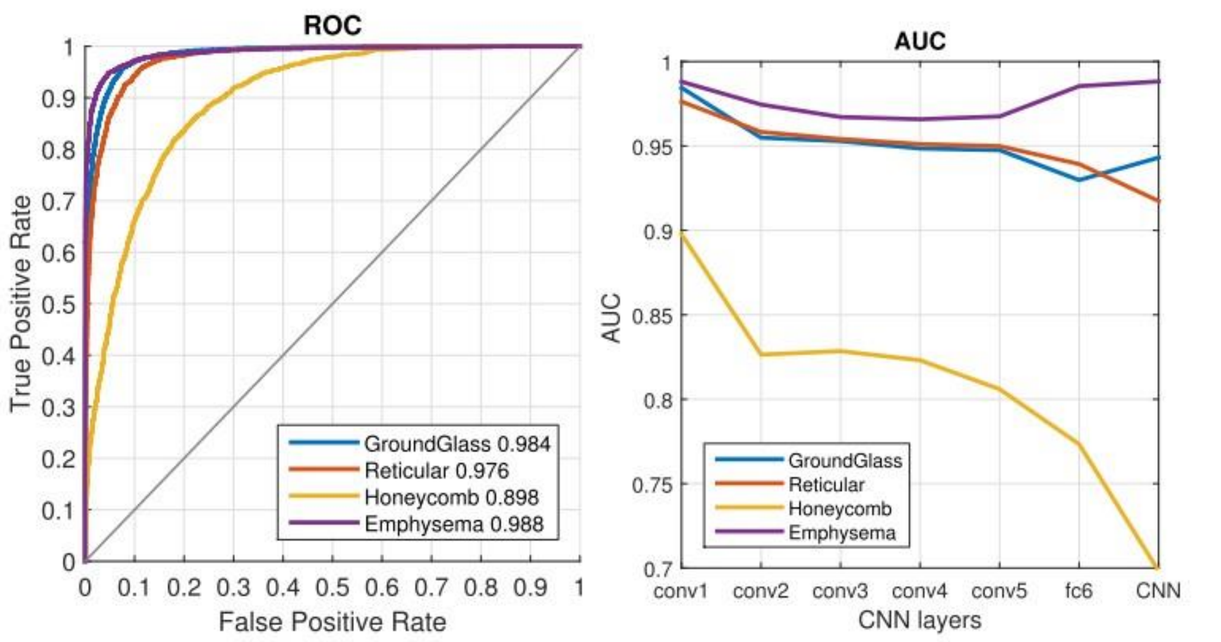


Experiments and Results

- We tested our system on 533 patients, consisting of 18883 slices, for training and testing. Five-fold cross evaluation was performed, where data were separated on patient level.

ILD pattern	Positive	Negative
Healthy	11677	7206
Ground Glass	3368	15515
Reticular	1606	17277
Honeycomb	1247	17636
Emphysema	2639	16244

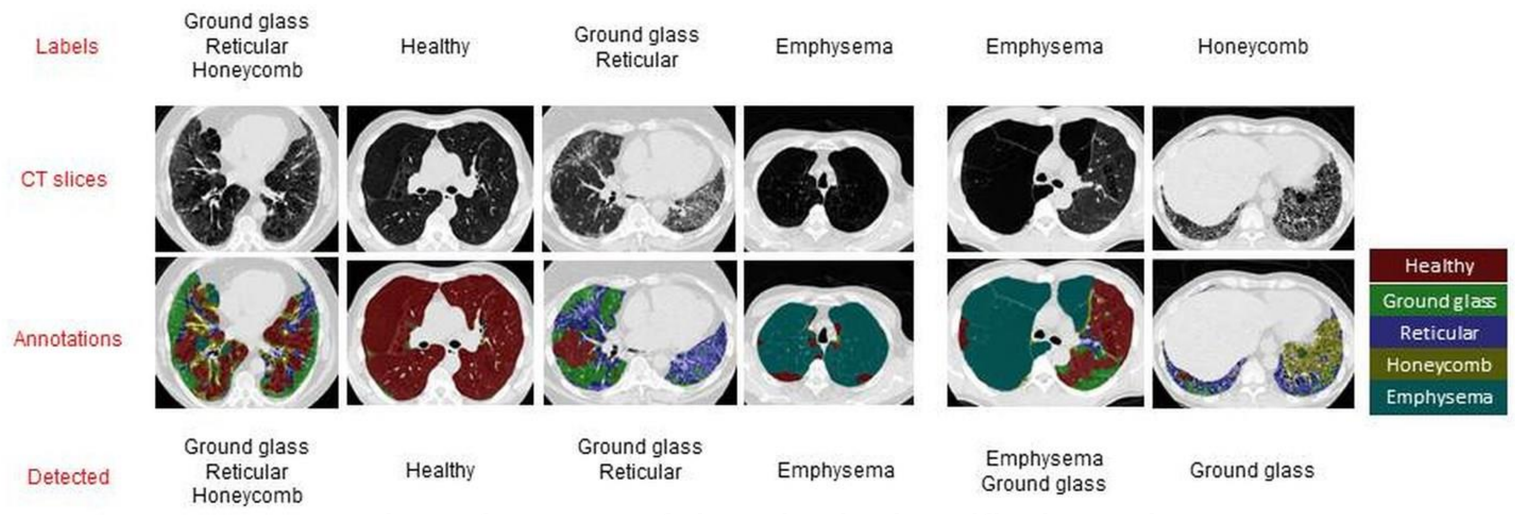
Healthy	One Disease	Two Diseases	Three Diseases	Four Diseases
11677	5675	1410	119	2



Left: Detection results of FV-CNN via the unordered feature pooling using conv1 layer.

Right: AUC of FV-CNN at different convolutional layers and MLCNN-R.

- To construct the FV-encoded features, the local convolutional image descriptors were pooled into 32 Gaussian components, producing 16K dimension features. Those features were further reduced into 512 dimensions using PCA.
- Each ILD pattern was evaluated separately. FV-CNN encoding feature representation has better performances than MLCNN-R.



Examples of successful and misclassified results.

References

[1] Bartholmai, B., Karwoski, R., Zavaletta, V., Robb, R., Holmes, D.: The lung tissue research consortium: An extensive open database containing histological, clinical, and radiological data to study chronic lung disease. (07 2006)

[2] Perronnin, F., S´ anchez, J., Mensink, T.: Improving the fisher kernel for large-scale image classification. In: ECCV. (2010) 143–156