Interstitial Lung Disease Classification Using Transfer Learning Pre-trained Deep Neural Network

ARUNAVA KIRTAN

Objective

Interstitial lung disease (ILD) is an umbrella term used for a large group of diseases that cause scarring (fibrosis) of the lungs. The scarring causes stiffness in the lungs which makes it difficult to breathe and get oxygen to the bloodstream.

The Problem: Due to the trouble of classifying ILD, analyzing the type of ILD from CT images by radiologists is not accurate and also very time consuming.

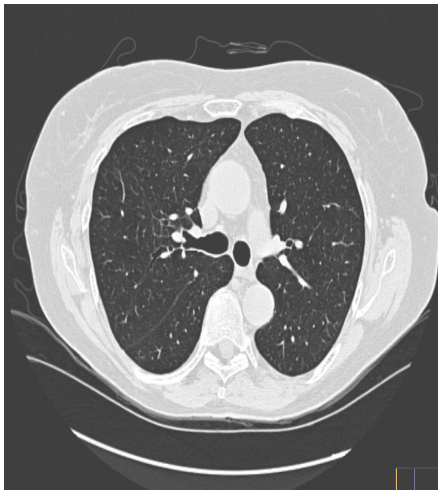
To solve this problem we used a pre-trained deep neural network for detection of various common interstitial lung diseases.

We have selected three common ILDs along with healthy
Which covers approx 75% of total ILDs reported every year

a)Hypersensitivity Pneumonitis

b) Pulmonary Fibrosis

c)Tuberculosis

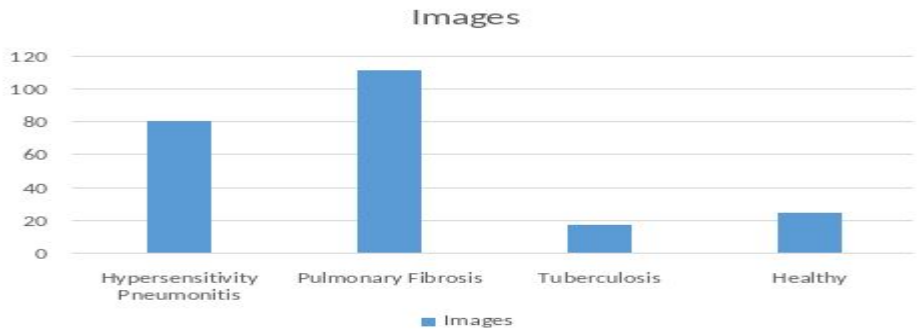


d) Healthy

Input images for this project mainly HRCT lungs images. All of these images are taken from the MedGift data-set.

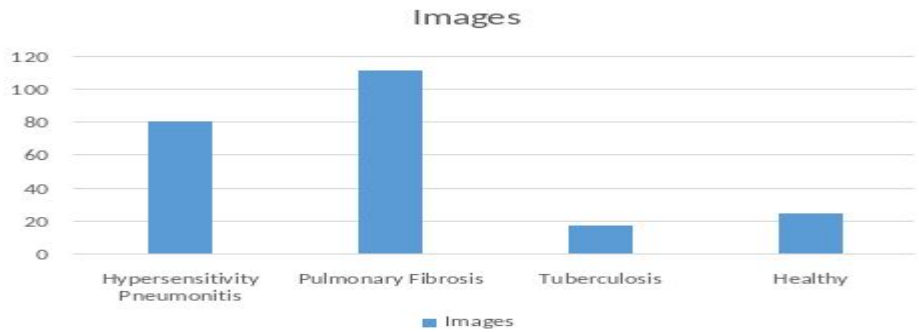
The ultimate dataset was created from the CT images of 4 categories with 35 patients and a total 236 images.

No.	Category of ILD	images
a	Hypersensitivity Pneumonitis	81
b	Pulmonary Fibrosis	112
c	Tuberculosis	18
d	Healthy	25



The limitation of image classification in the medical field is that there is not a very good scope to collect adequate images to train our model so we opt for a method which is called transfer-learning.

Also the differences of the quantities among the images is evident. So it is always a better option to train a pre-train model with a limited dataset.



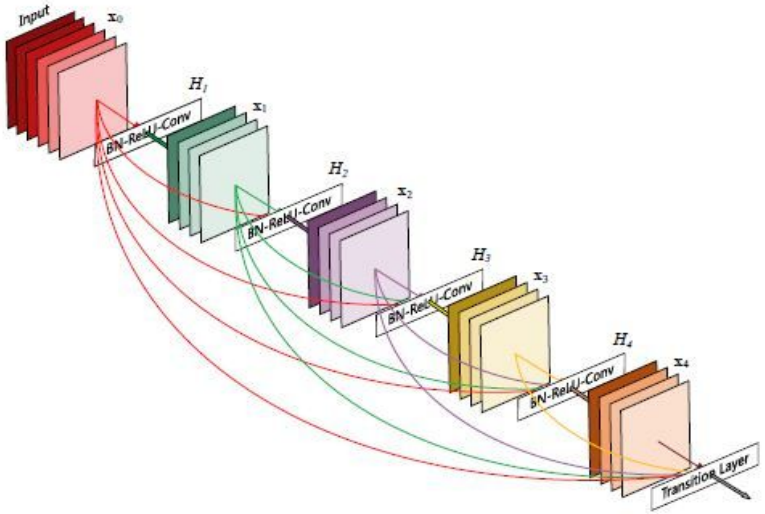
The pre-trained models used:

- **DenseNet121 Model**
- VGG 16 Model
- InceptionV3 Model
- Xception Model
- ResNet50 Model

This pathway is useful because pre-trained models are trained on a huge collection of images and train the model predict on a comparatively different big number of classes.

Because we have only 236 images there is not many adequate images to properly train our models.

DenseNet121 Architecture



Therefore, DenseNet-121 has the following layers:

- 1 7x7 Convolution
- 58 3x3 Convolution
- 61 1x1 Convolution
- 4 AvgPool
- 1 Fully Connected Layer

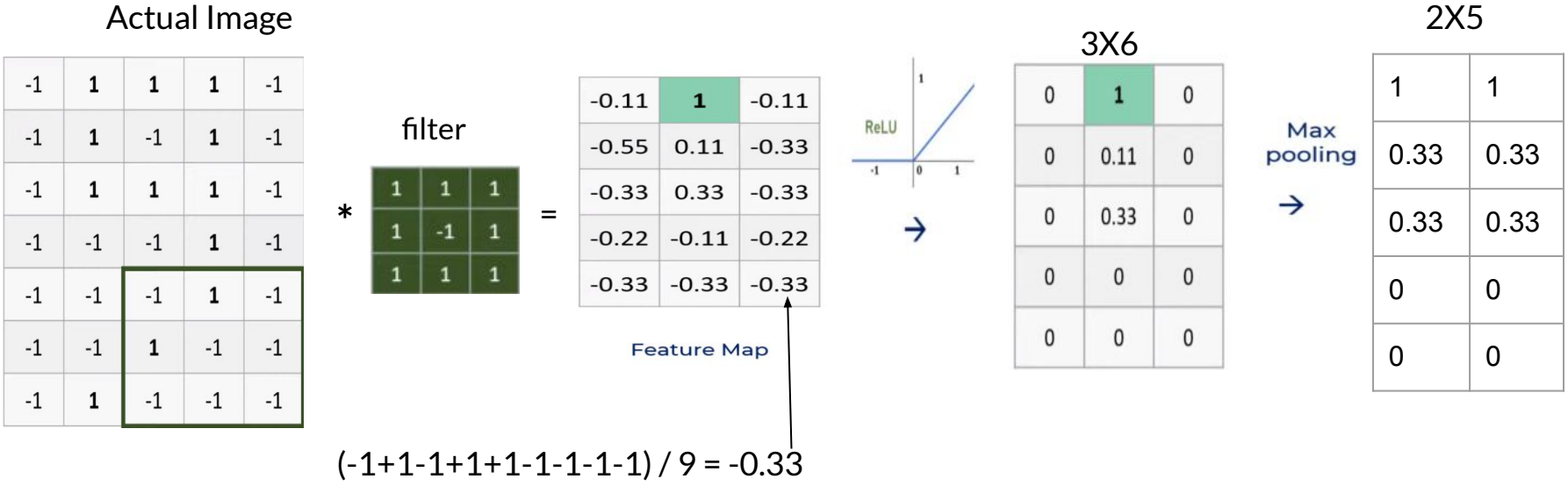
In short,by default DenseNet-121 has 120 Convolutions and 4 AvgPool.

We used **hyper tuning** to get optimal **learning rate** and optimal **neuron number** in fully connected layer.

new weight = w1 - learning rate * dw1

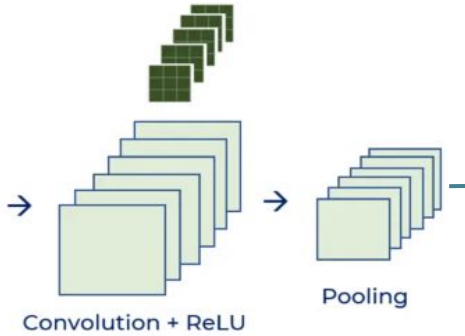
= 1-0.2=0.8

After all the the convolution operation and pooling operation all the output goes to fully connected neural network as a input. The network consists of 1024 neurons in fully connected layer.



0=healthy, 1=hypersensitivity pneumonitis, 2=pulmonary fibrosis,3=tuberculosis

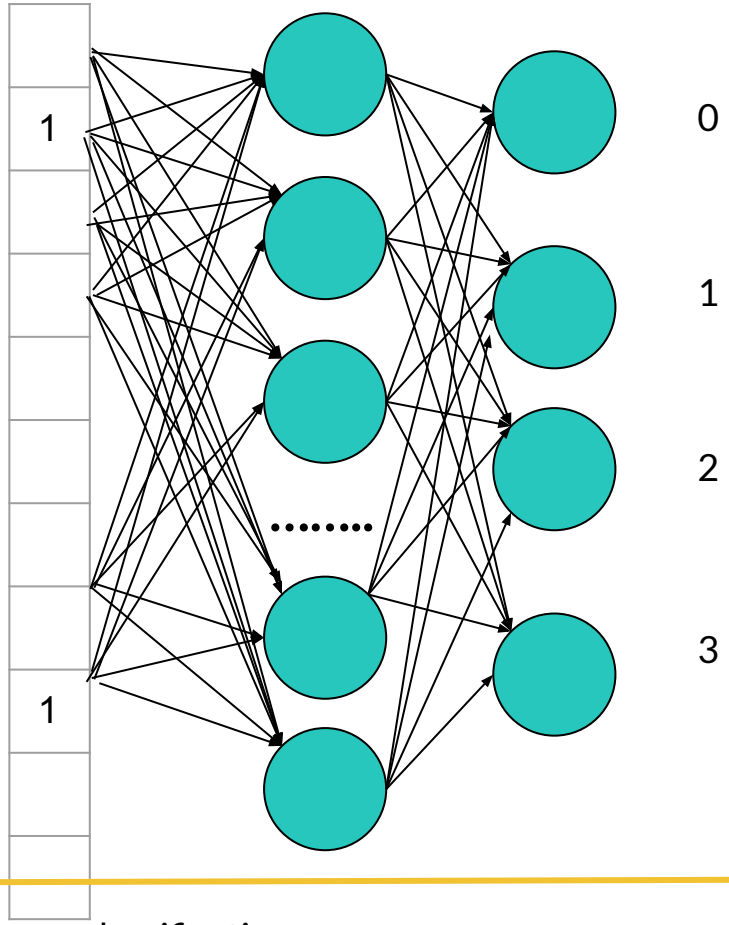
Convolution operation with different filters



120 times

4 times

flatten



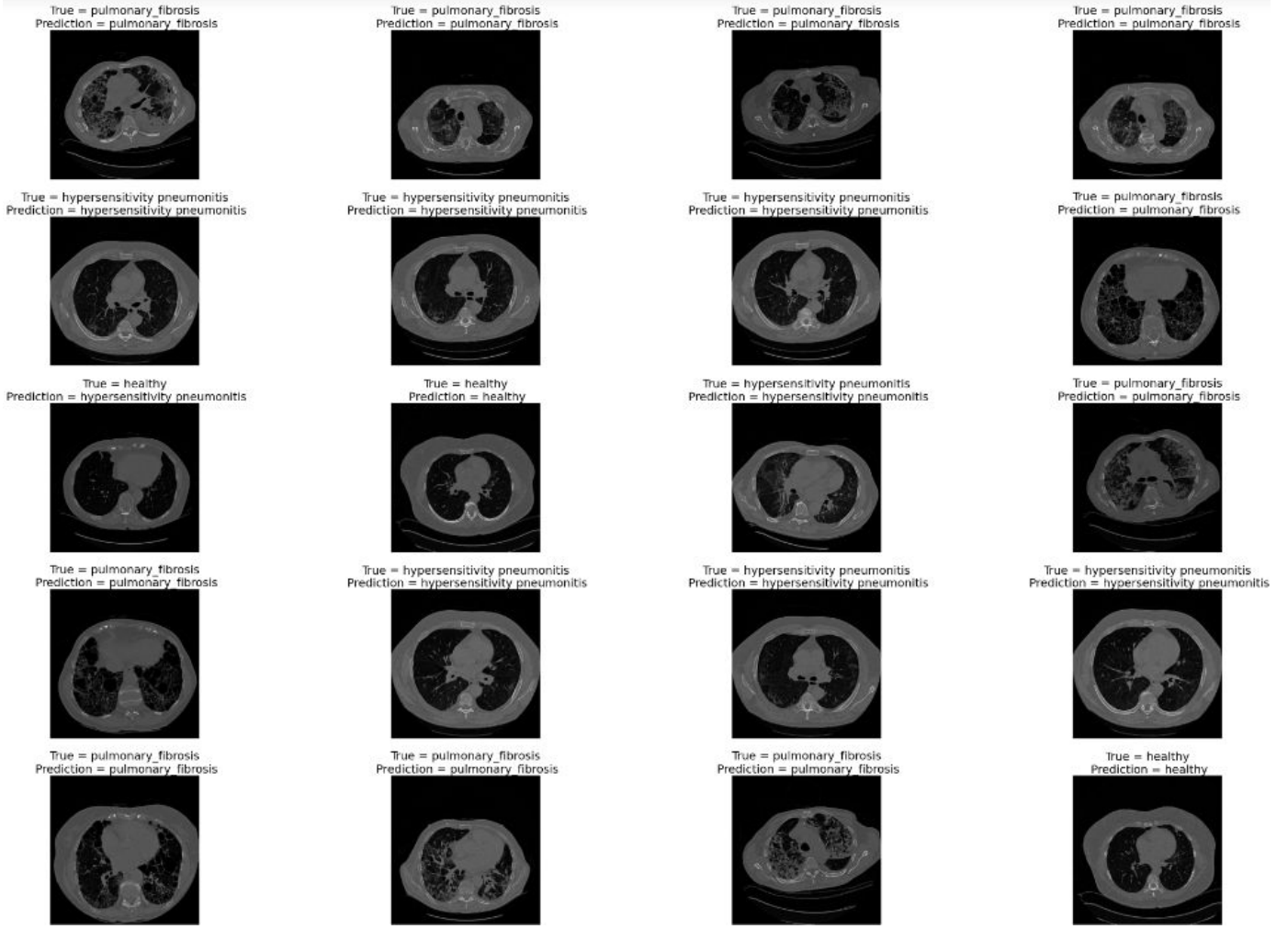
feature extractions

classification

	VALIDATION ACCURACY				
	DENSENET121	VGG16	INCEPTION V3	XCEPTION	RESNET50
FOLD 1	0.83	0.83	0.79	0.86	0.49
FOLD 2	0.88	0.77	0.79	0.79	0.49
FOLD 3	0.86	0.81	0.83	0.79	0.49
FOLD 4	0.84	0.81	0.81	0.79	0.49
FOLD 5	0.88	0.81	0.79	0.75	0.49
AVERAGE	**0.85**	0.80	0.80	0.79	0.49

Test Accuracy

	ACCURACY	SENSITIVITY	SPECIFICITY
DENSE NET	**0.82**	**0.71**	**1.0**
VGG16	0.75	0.57	1.0
INCEPTION V3	0.68	0.42	0.70
XCEPTION	0.65	0.0	1.0
RESNET50	0.34	nan	nan


True = pulmonary_fibrosis
Prediction = pulmonary_fibrosis
True = pulmonary_fibrosis
Prediction = pulmonary_fibrosis
True = pulmonary_fibrosis
Prediction = pulmonary_fibrosis
True = pulmonary_fibrosis
Prediction = pulmonary_fibrosis
True = hypersensitivity pneumonitis
Prediction = hypersensitivity pneumonitis
True = hypersensitivity pneumonitis
Prediction = hypersensitivity pneumonitis
True = hypersensitivity pneumonitis
Prediction = hypersensitivity pneumonitis
True = pulmonary_fibrosis
Prediction = pulmonary_fibrosis
True = healthy
Prediction = hypersensitivity pneumonitis
True = healthy
Prediction = healthy
True = hypersensitivity pneumonitis
Prediction = hypersensitivity pneumonitis
True = pulmonary_fibrosis
Prediction = pulmonary_fibrosis
True = pulmonary_fibrosis
Prediction = pulmonary_fibrosis
True = hypersensitivity pneumonitis
Prediction = hypersensitivity pneumonitis
True = hypersensitivity pneumonitis
Prediction = hypersensitivity pneumonitis
True = hypersensitivity pneumonitis
Prediction = hypersensitivity pneumonitis
True = pulmonary_fibrosis
Prediction = pulmonary_fibrosis
True = pulmonary_fibrosis
Prediction = pulmonary_fibrosis
True = pulmonary_fibrosis
Prediction = pulmonary_fibrosis
True = healthy
Prediction = healthy

The introduced method gave outstanding results, well performed on a very challenging dataset of 236 CT scan images from different hospitals and scanners. All the 5 models have fluctuation of the results, for the same input. In future times, we plan to improve our models accuracy.

	ACCURACY	SENSITIVITY	SPECIFICITY
DENSE NET	**0.82**	**0.71**	**1.0**
VGG16	0.75	0.57	1.0
INCEPTION V3	0.68	0.42	0.70
XCEPTION	0.65	0.0	1.0
RESNET50	0.34	nan	nan