**Pathology detection using Chest X-rays - extended to COVID -19**

Durga Prasad, Akaash, Suresh, Nitish, Karthik

***Abstract*** Across the globe respiratory-related diseases claim most of the lives and there is a wide variety of pathologies that cause respiratory diseases. The Deep learning models trained on natural images are commonly used for different classification tasks in the medical domain. We develop a model that can detect various pathogens from chest X-rays at a level in par with practicing radiologists. In the current model, we are aiming to train the model using deep learning approaches. Convolutional Neural Network (CNN) has gained popularity in learning mid and high-level representations. We explore this capability of CNN deep learning architectures to identify various pathologies in the chest x-ray images. The final trained model will take an Xray input and will be able to classify and predict the presence of a particular pathogen. The model will also being extended to classify the latest virus which is causing chaos over the world the novel coronavirus COVID-19 which affected more than 3.5m people and claiming 270k+ lives

Following are the class of pathologies the model can classify by taking the input chest x-ray image

|  |  |  |  |
| --- | --- | --- | --- |
| **1.** | **Atelectasis** | **9.** | **Consolidation** |
| **2.** | **Cardiomegaly** | **10.** | **Edema** |
| **3.** | **Effusion** | **11.** | **Emphysema** |
| **4.** | **Infiltration** | **12.** | **Fibrosis** |
| **5.** | **Mass** | **13.** | **Pleural\_Thickening** |
| **6.** | **Nodule** | **14.** | **Hernia** |
| **7.** | **Pneumonia** | **15.** | **Coronavirus COVID-19** |
| **8.** | **Pneumothorax** |  |  |

***1. Introduction* :**Chest radiographs are the very common diagnostic imaging tool used across the hospitals in identifying the chest abnormalities. Brief description of each Pathology is provided below with sample representation chest radiograph.

**1.1 Atelectasis :**

Atelectasis (refer Fig 1.1) is another word for lung collapse. The commonest cause is a bronchial obstruction that results in distal gas resorption and a reduction in the volume of gas in the affected lung, lobe, segment or subsegment. As the gas is resorbed, the walls of the alveoli collapse in on themselves and the size of the affected area reduces.

****

*Fig. 1.1*

**1.2. Cardiomegaly :**

Cardiomegaly means enlargement of the heart. The definition is when the transverse diameter of the cardiac silhouette is greater than or equal to 50% of the transverse diameter of the chest (increased cardiothoracic ratio) on a posterior-anterior projection of a chest radiograph or a computed tomography.

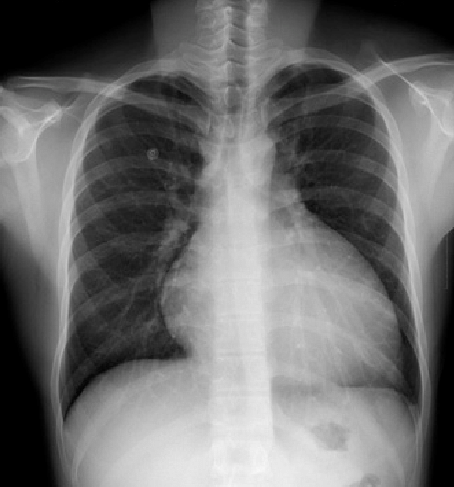


Fig 2.1

**1.3. Effusion :**

A pleural effusion is an abnormal collection of fluid in the pleural space resulting from excess fluid production or decreased absorption or both.



Fig 3.1

**1.4. Infiltration (refers Fig 4.1 & 4.2)**

Persistent pulmonary infiltrate results when a substance denser than air (e.g., pus, oedema, blood, surfactant, protein, or cells) lingers within the lung parenchyma. Non-resolving and slowly resolving pneumonias are the most common broad categories of persistent pulmonary infiltrate.

****

Fig 4.1 Chest X-ray showing infiltrate in the right lung.

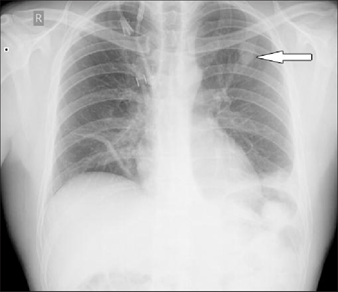
****

Fig4.2 Chest X-ray image. Local infiltration in the upper lobe of the left lung (white arrow)

**1.5. Mass:**

Lung cancer is a mass or growth in the lung made up of cancer cells, but not all masses in the lung are caused by cancer. Lung masses and growths generally fall under two categories:

- Benign growths (non-cancerous) - Malignant growths (cancerous)

Malignant masses are often larger than benign nodules. Malignant masses are also more likely to show up in certain parts of your lung.

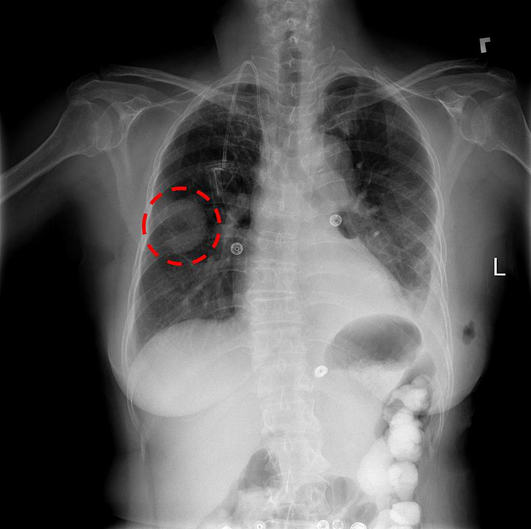
****

Fig. 5.1 Lung Mass

**1.6. Nodule:**

Usually, the first sign that a pulmonary nodule is present is a spot on the lung that shows up on a chest X-ray or a CT scan.

Single pulmonary nodules seen on chest x-rays are generally at least 8 to 10 millimeters in diameter. If they are smaller than that, they are unlikely to be visible on a chest X-ray. The larger the nodule is, and the more irregularly shaped it is, the more likely it is to be cancerous. Those located in the upper portions of the lung are also more likely to be cancerous.

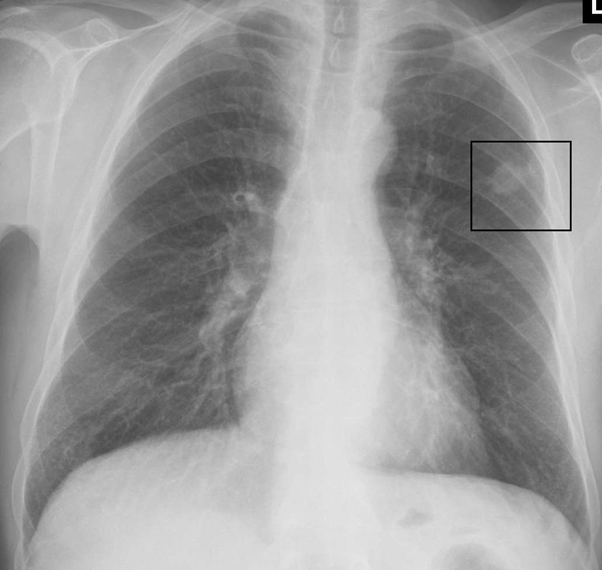
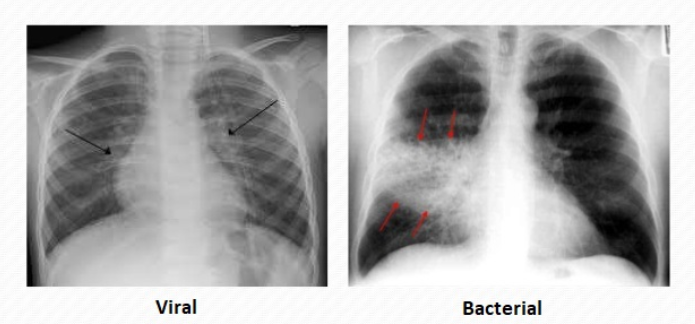


Fig. 6.1 Lung Nodule

**1.7. *Pneumonia*** refers to figure 4. Infection that inflames air sacs in one or both lungs, which may fill with fluid. With pneumonia, the air sacs may fill with fluid or pus. The infection can be life-threatening to anyone, but particularly to infants, children, and people over 65.



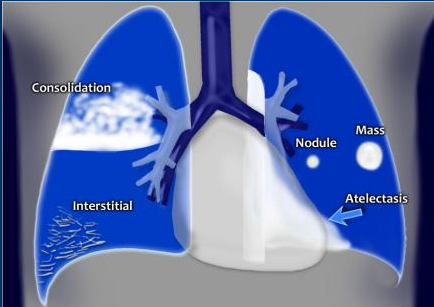
*Fig. 4 Pneumonia*

**1.8. Pneumothorax /collapsed lung** refer to figure 2. It is a condition that occurs when air enters the space between the chest wall and the lung(Pleural space) and causes a collapsed lung. Pneumothorax can be a complete lung collapse or a collapse of only a portion of the lung.



*Fig. 2. Pneumothorax (left lung is collapsed)*

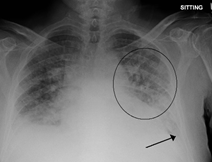
# **1.9. Consolidation** refers to figure 3. A pulmonary consolidation is a region of normally compressible lung tissue that has filled with liquid instead of air. Consolidation occurs through the accumulation of inflammatory cellular exudate in the alveoli and adjoining ducts. The liquid can be pulmonary edema, inflammatory exudate, pus, inhaled water, or blood (from bronchial tree or hemorrhage from a pulmonary artery).



*Fig. 3. Consolidation(right lung)*

**1.10. *Edema***

Pulmonary edema is a condition that is caused by accumulation of excess fluid in the lungs. This condition results in breathing difficulties as the fluid gets accumulated in the lungs air sacs.This Fluid accumulation is caused due to multiple reasons like congestive heart failure, pneumonia, exposure to certain toxins, trauma to the chest wall etc. ( Refer to Figure 6 )



*Fig. 6 Edema*

**1.11. Emphysema**

Pulmonary Emphysema is a condition in which the air sacs of the lungs are damaged and enlarged, causing breathlessness. Emphysema belongs to a group of lung diseases known as chronic obstructive pulmonary disease. Once it develops, emphysema cannot be reversed.When emphysema develops, the alveoli and lung tissue are destroyed & oxygen movement into the bloodstream will be reduced.

****

Fig 11.1 : Emphysema

**1.12. Fibrosis**

Pulmonary fibrosis is a condition that occurs when the lung tissue becomes damaged and scarred. Pulmonary fibrosis scars and thickens the tissue around and between the air sacs. This lung damage makes it difficult for the lungs to function properly. The damage caused to the lungs cannot be repaired and as the condition worsens, lungs become shorter of breath**. ( Refer to Figure 7 )**

****

*Fig. 7 Fibrosis*

**1.13. Pleural\_Thickening :**

Pleural thickening is a common finding on routine chest X-rays. It typically involves the apex of the lung, which is called ‘pulmonary apical cap’. On chest X-rays, the apical cap is an irregular density located at the extreme apex and is less than 5 mm in width. Four areas of the lung (apical, upper, middle, and lower portions) were examined bilaterally for the presence of pleural thickening.  ****

Fig 13.1 : Pleural Thickening

**1.14. Hernia :**

A hernia occurs when an organ pushes through an opening in the muscle or tissue that holds it in place.

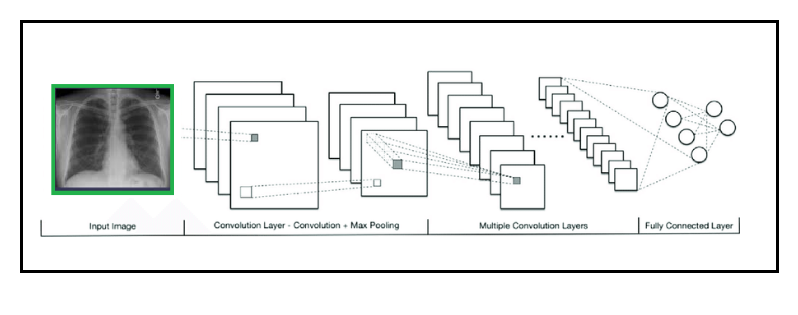
**1.15. Covid-19 :**

COVID-19 is a viral disease also known as SARS-CoV-2 or severe acute respiratory syndrome coronavirus 2.In humans, several coronaviruses are known to cause respiratory infections ranging from the common cold to more severe diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). The most recently discovered coronavirus causes coronavirus disease COVID-19 - World Health Organization.

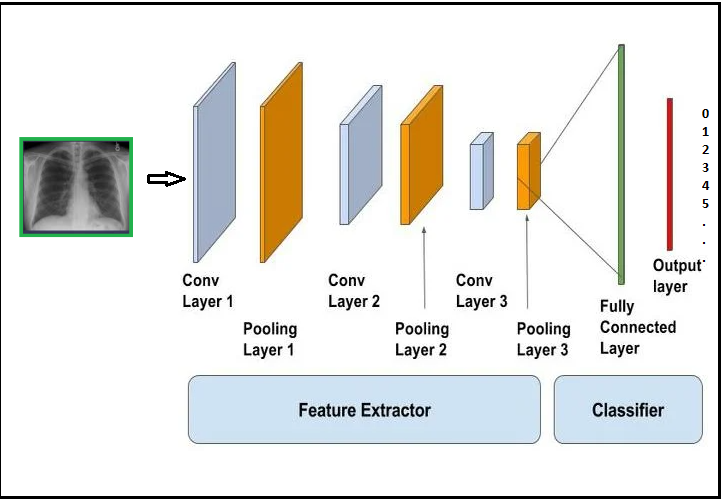


**1.16. NORMAL :**

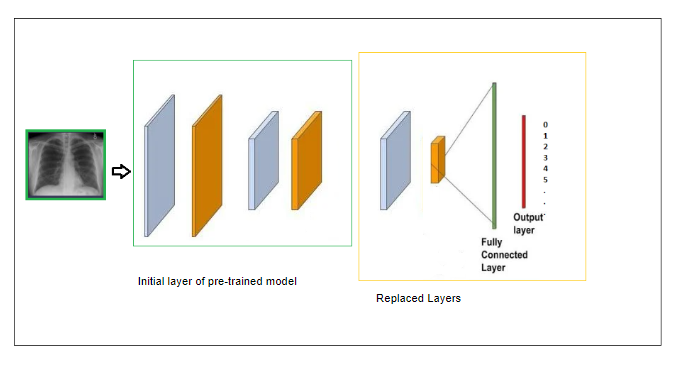
**Model will be trained with chest X-ray images which are classified as normal.**

***2. The Model*** refers to figure 5.1& 5.2. The proposed model will contain multiple convolutional and pooling layers for feature extraction and then a flatten layer that connects to a fully connected layer. Figure 5.1 is only for the depiction purpose of the proposed model and the actual model, number of convolutions, pooling layers, fully connected layers will differ based on the further work and tuning of the model. We are also aiming at experimenting with the transferred learning and use pre-trained models refer to figure 5.3******

*Fig. 5.1 Overview of Convolutional Neural Network (CNN) Architecture for pathology detection*

**

*Fig. 5.2 Overview of Convolutional Neural Network (CNN) Architecture for pathology detection*

**

*Fig. 5.3 Convolutional Neural Network (CNN) with transferred learning*

***3. Sample Data***

ChestX-ray dataset comprises 112,120 frontal-view X-ray images of 30,805 unique patients with the text-mined fourteen disease image labels (where each image can have multi-labels), mined from the associated radiological reports using natural language processing. Fifteen thoracic pathologies include Atelectasis, Consolidation, Infiltration, Pneumothorax, Edema, Emphysema, Fibrosis, Effusion, Pneumonia, Pleural\_thickening, Cardiomegaly, Nodule, Mass, Hernia and Coronavirus COVID-19.

Planning to have the training data into 16 different folders based on the classification of pathogens.

**Contents:**

1. 112,120 frontal-view chest X-ray PNG images in 1024\*1024 resolution (under images folder)

2. Metadata for all images (Data\_Entry\_2017.csv): Image Index, Finding Labels, Follow-up #, Patient ID, Patient Age, Patient Gender, View Position, Original Image Size and Original Image Pixel Spacing.

3. Bounding boxes for ~1000 images (BBox\_List\_2017.csv): Image Index, Finding Label, Bbox[x, y, w, h]. [x y] are coordinates of each box's top left corner. [w h] represent the width and height of each box.

4. Two data split files (train\_val\_list.txt and test\_list.txt) are provided. Images in the chest X-ray dataset are divided into these two sets on the patient level. All studies from the same patient will only appear in either training/validation or testing set.

**References :**

[1] Proto AV, Tocino I. Radiographic manifestations of lobar collapse. Semin Roentgenol. 1980;15 (2): 117-73. - Pubmed citation

[2] Amin H, Siddiqui WJ. Cardiomegaly. [Updated 2019 Jun 4]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan-.

[3] Diaz-Guzman E, Dweik RA. Diagnosis and management of pleural effusions: a practical approach. Compr Ther. 2007 Winter. 33(4):237-46. [Medline].