Developing an Early Warning Tool to detect COVID-19 Pneumonia

By Angeline Chandraatmadja

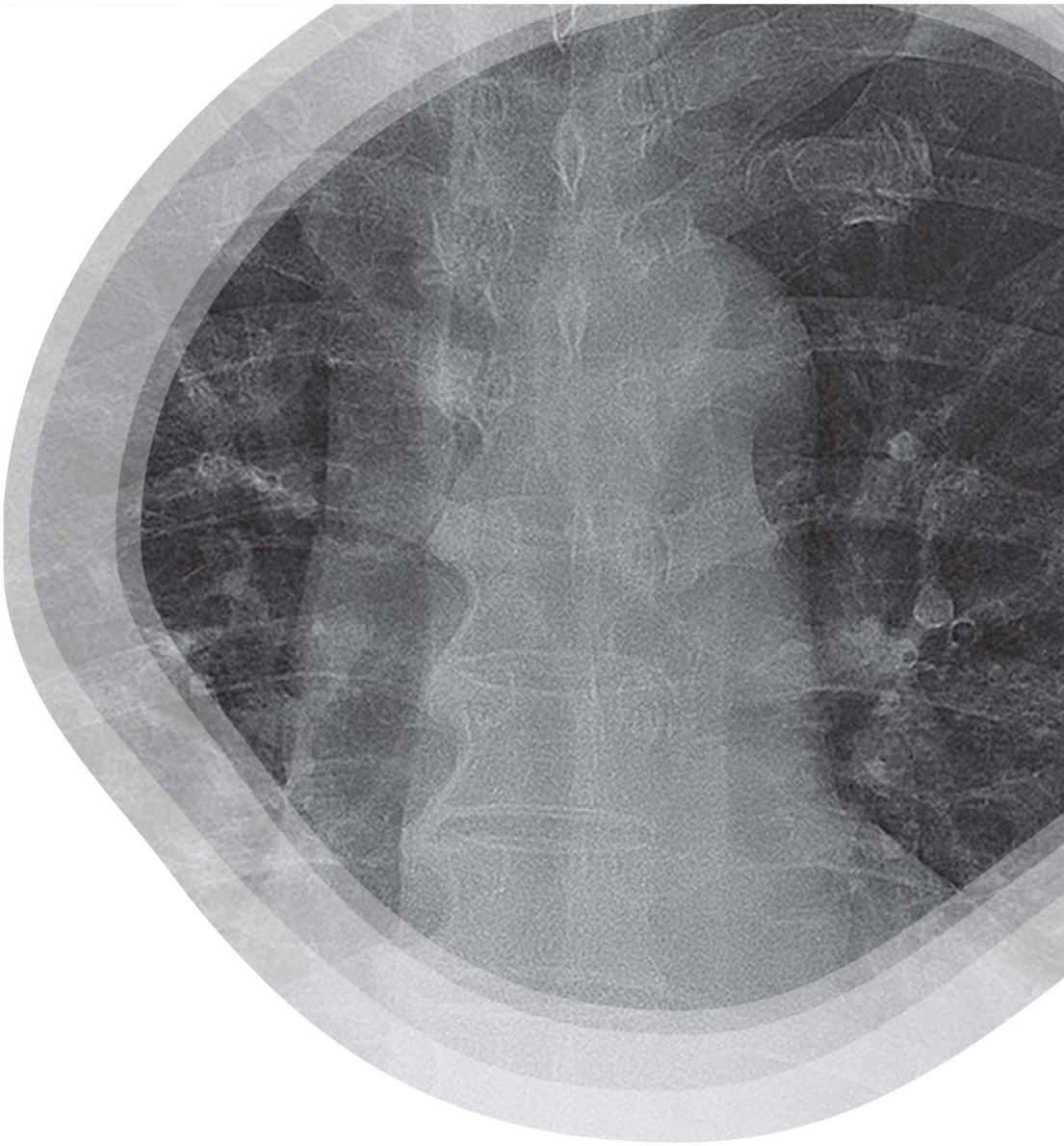
Chest Radiographs in COVID-19 management

- Evaluating severity of illness
  - used for prioritization of patients
  - prescriptions
- Monitoring
  - post COVID-19 chronic lung abnormalities

Machine Learning as an early warning tool for covid-19 pneumonia

- Diagnosing COVID-19 Pneumonia is difficult: Average sensitivity of 80.4% amongst radiologists

- Evaluation Metrics: F1 Score
- Baseline Accuracy: 53%



Dataset

- Format: JPEG, PNG, JPG
- Size: 5 GB
- Dimensions: Varying sizes (1148 x 908 , 2144x 2129, 1845 x 1761, etc)
- Classes:
  - COVID: 363 Images, 6.5%
  - NOFINDING: 1,408 Images, 25.6%
  - THORAXDISEASE: 3,763 Images, 67.9%

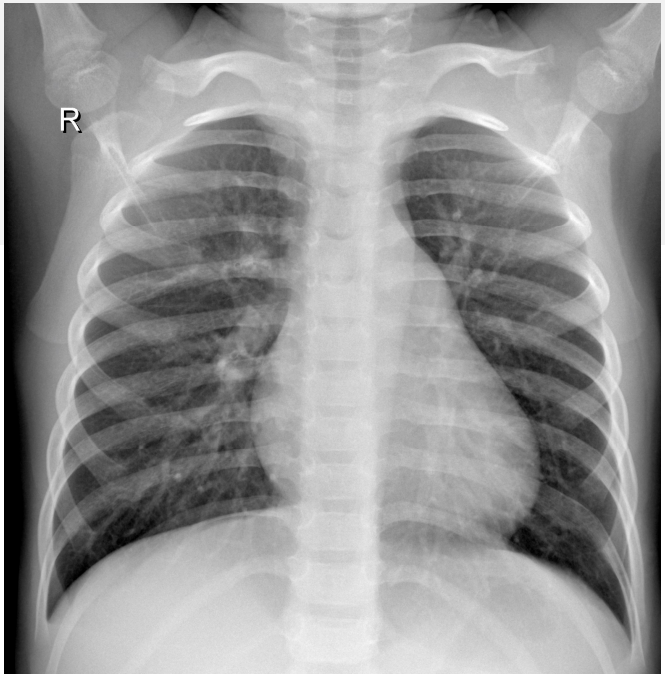
BENCHMARK: X-RAYS (RIPE)

- **Rotation**
- **Inspiration**
- **Projection**
- **Exposure**

Normal Lungs: Bad Quality

Spine is relatively centred in the picture, inspiration effort is poor resulting in an enlarged cardiac shadow, poor visibility of the vertebrae

Normal Lungs: Good Quality



Spine is relatively centred in the picture, inspiration is deep showing a clear view of the mediastinum, exposure is clear

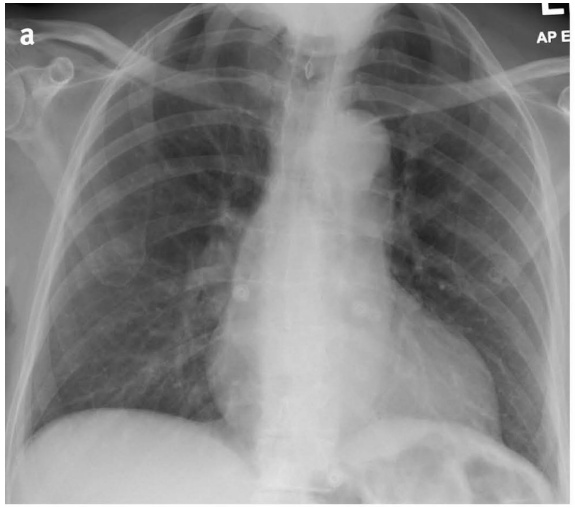
Image transformation

50% of the trained image is flipped horizontally

- Increase variability

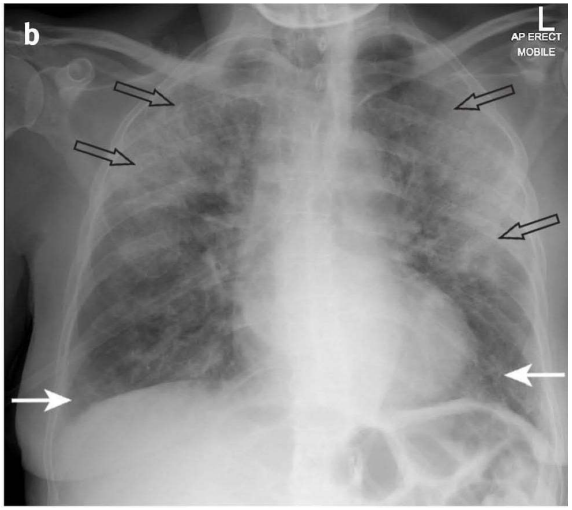
CLASS
CHARACTERISTICS

NOFINDING



- Central mediastinum and heart appear normal
- Lungs predominantly contain air
- Lung markings are present representing blood vessels

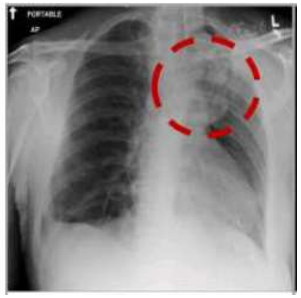
COVID-19 Pneumonia



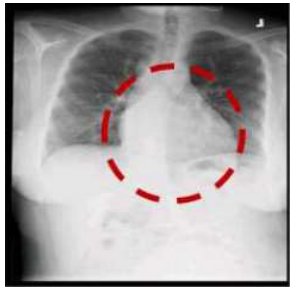
- Ground glass opacity
- Consolidation
- Findings are predominantly in the middle lower zone

Class Characteristics: Thorax Disease

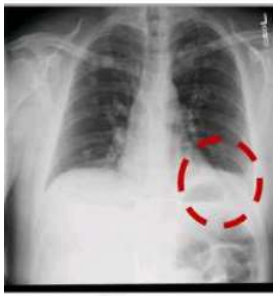
Eight common thoracic diseases observed in our dataset



Atelectasis
Partial lung collapse



Cardiomegaly
Heart enlargement



Effusion
Water between the lung and the sac

Infiltration
Water inside the lungs

Mass
Cancer/Tumour
Bigger than nodule

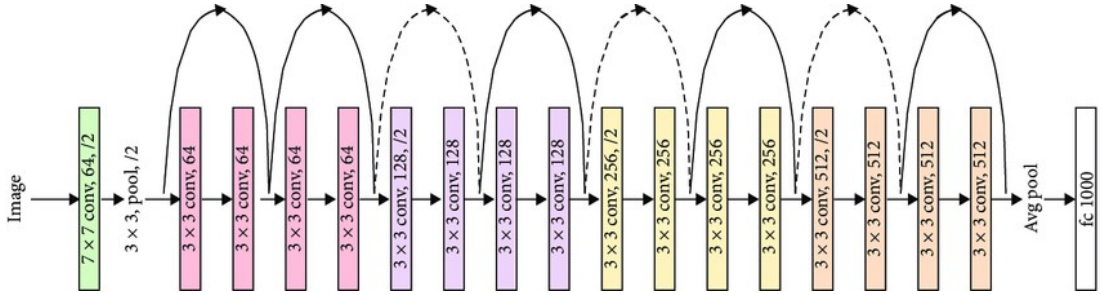
Nodule
Cancer/Tumour

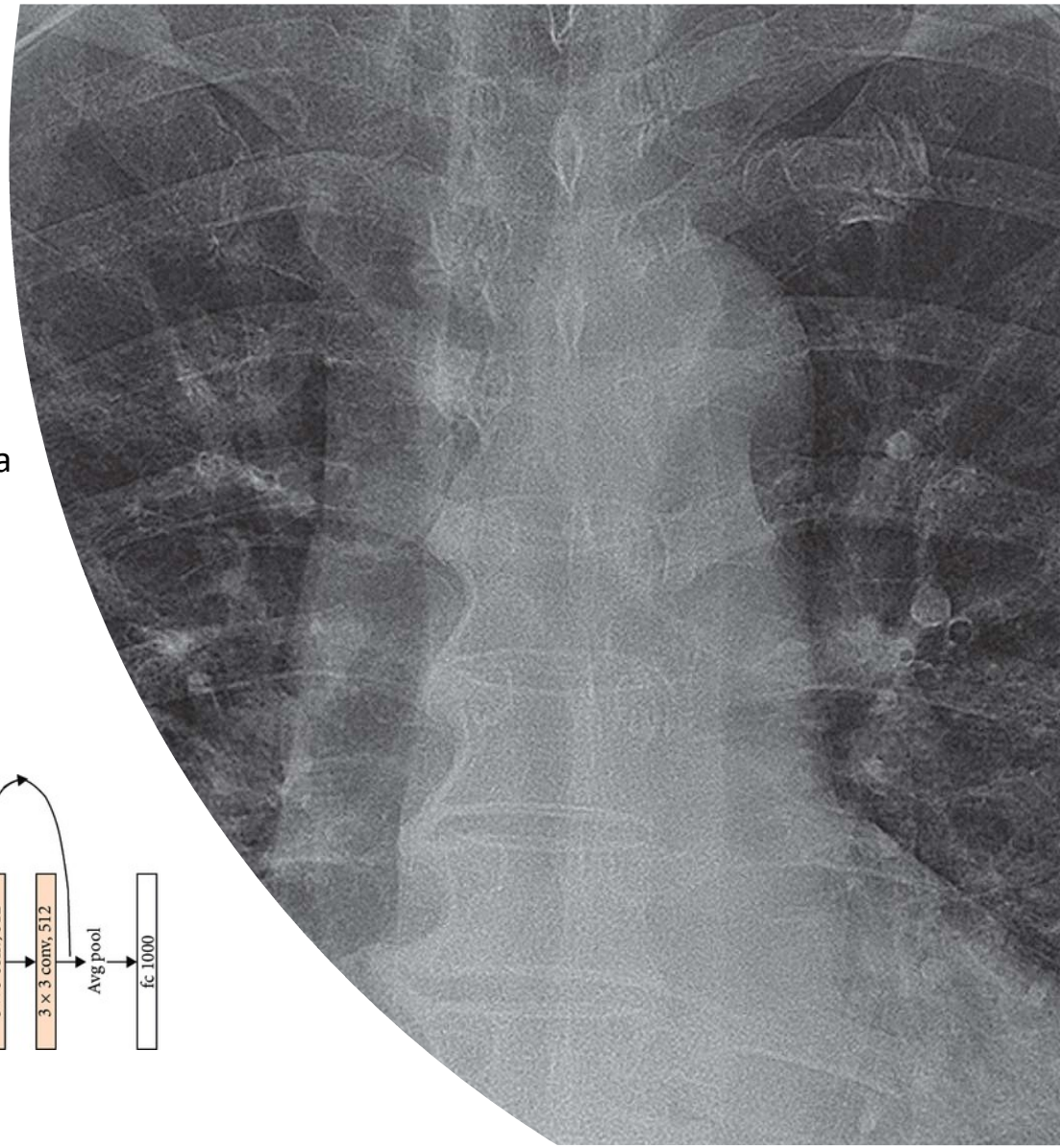
Pneumonia
Water in lungs due to pathogen

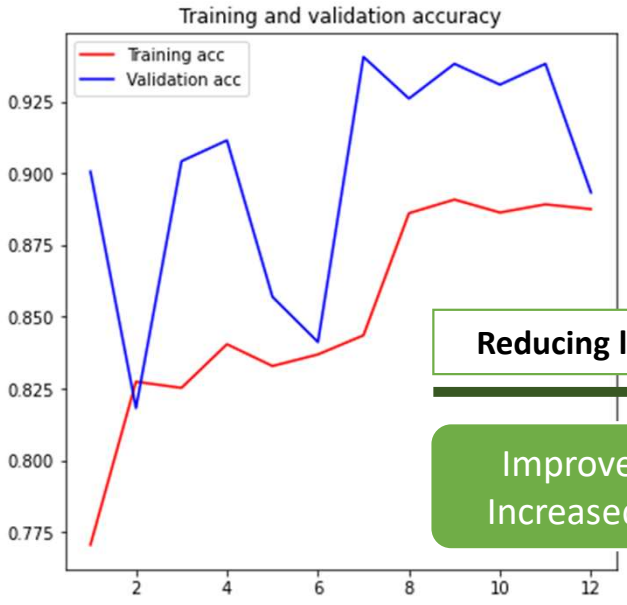
Pneumothorax
total lung collapse

RESNET18 Model

- Dataset: Image net (1million images) → Custom Data
- Input Size: 224 x 224 ← various sizes
- Normalization: normalized ← none
- Classes: 1000 Classes → 3 Classes
- Layers: 18

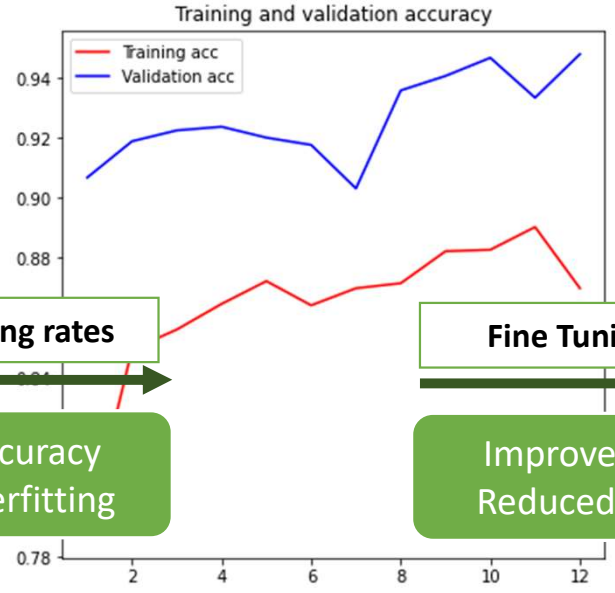






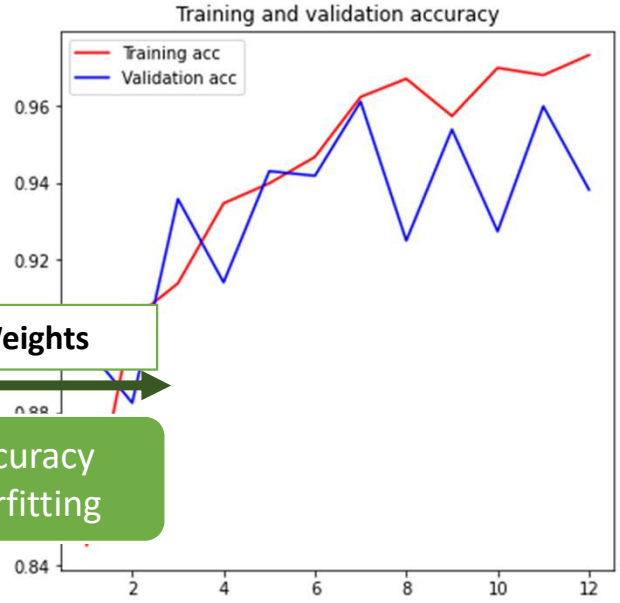
Reducing learning rates

Improved accuracy
Increased overfitting



Fine Tuning Weights

Improved accuracy
Reduced overfitting



- Frozen layers
- Initial LR = 0.01
- Reduce LR on Plateau by a factor of 0.1 with patience 3

Training time: 158.785 minutes

- Frozen layers
- Initial LR = 0.001
- Reduce LR on Plateau by a factor of 0.1 with patience 3

Training time: 183.703 minutes

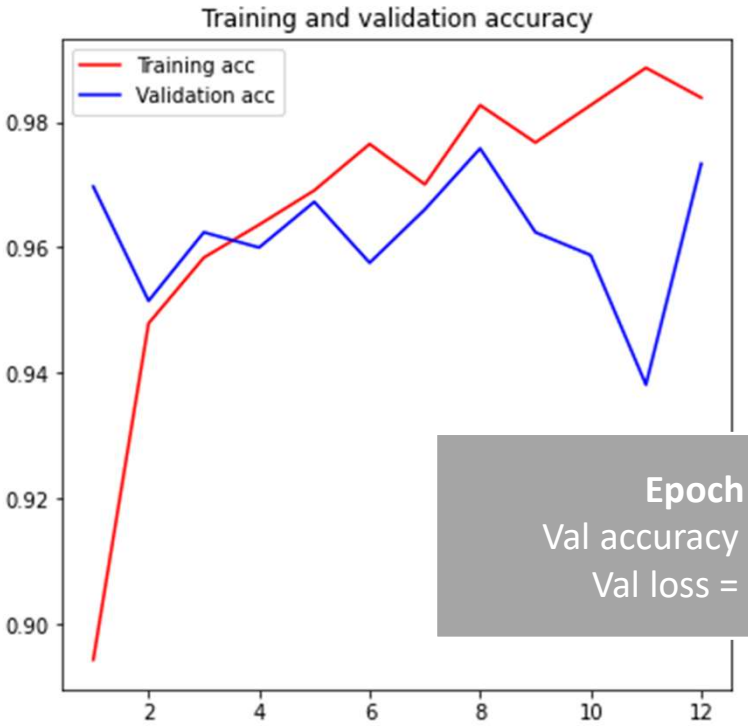
- Un-frozen layers
- Initial LR = 0.001
- Reduce LR on Plateau by a factor of 0.1 with patience 3

Training time: 379.154 minutes

Final Model

- Frozen layers
- Initial LR = 3e-4
- Reduce LR on Plateau by a factor of 0.1 with patience 3

- Training time: 201.908 minutes (Paid Google Collab)



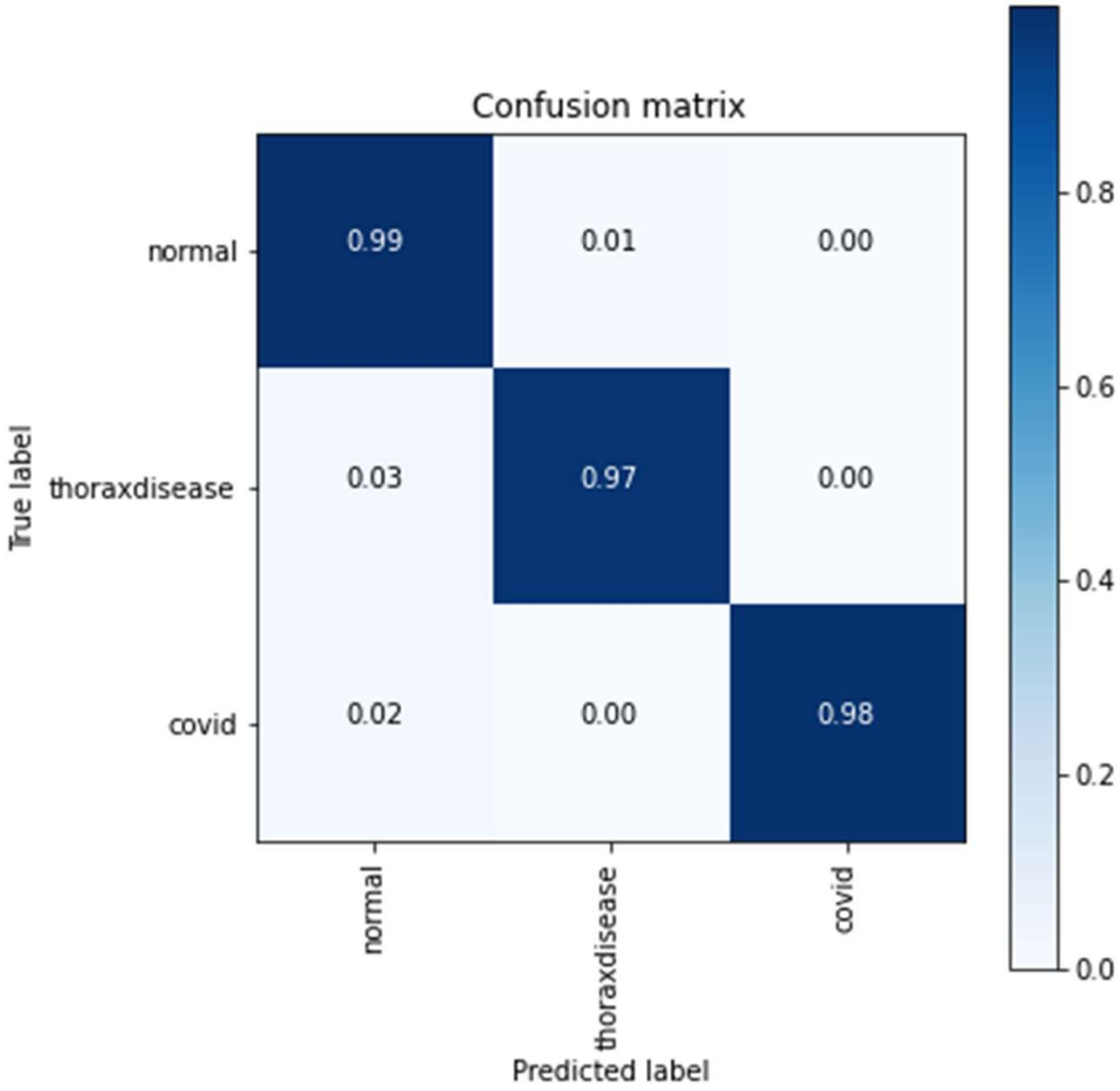
Epoch 8
Val accuracy = 97.5%
Val loss = 7.3%

Evaluation

- **Accuracy**
  Normal: 99%
  Thorax Disease: 97%
  Covid: 98%

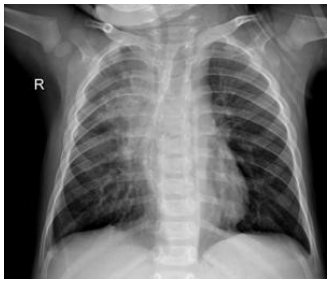
- **COVID-19 Sensitivity: 0.98%**
- **Industry Sensitivity: 84%**

- **Best Kaggle Score:**
- Test F1 : 97%
- Kaggle F1: 90.4%



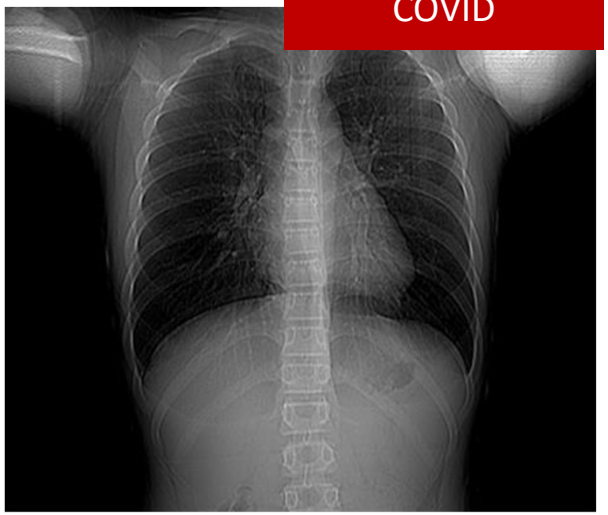
Misclassification Analysis

1. Generalize high contrast images → Normal
2. Poorly taken X-Rays
   bad inspiration, bad exposure → Thorax/Covid
3. Actual misclassifications of image
4. Very subtle findings



Examples of misclassifications

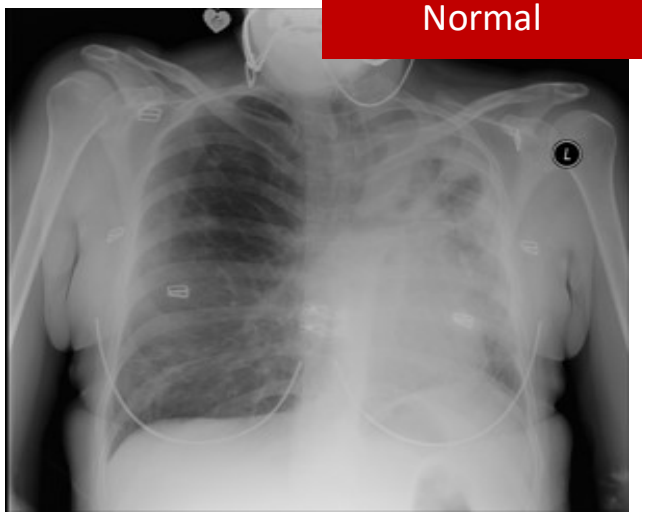
COVID



normal	**51.8%**
thorax	42.7%
covid	5.5%

Doctor review: Normal
Actual misclassification?

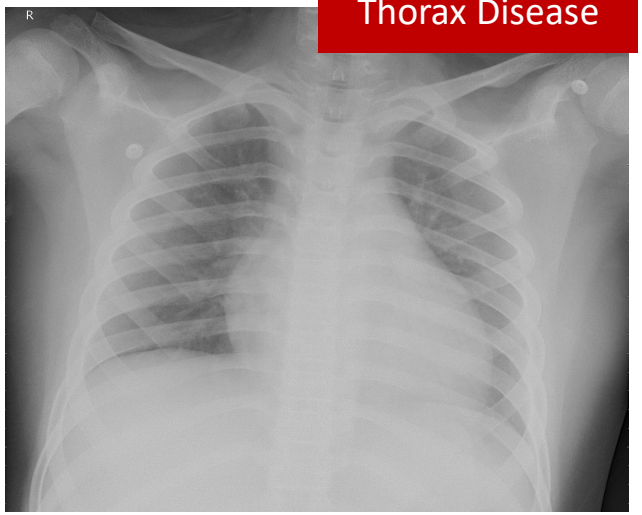
Normal



normal	0.4%
thorax	**97.7%**
covid	0.02%

Doctor review: Abnormal
Actual misclassification?

Thorax Disease



normal	16.3%
thorax	31.5%
covid	**52.2%**

Doctor review: Poor quality, not RIPE

Conclusion

- The model can better detect COVID-19 than the average radiologist
- It cannot tell bad X-Ray from abnormalities like a radiologist can

& Future Works

- Expose model to more classes of lung abnormalities
- Create a model to determine good quality vs bad quality CXR

Thank you