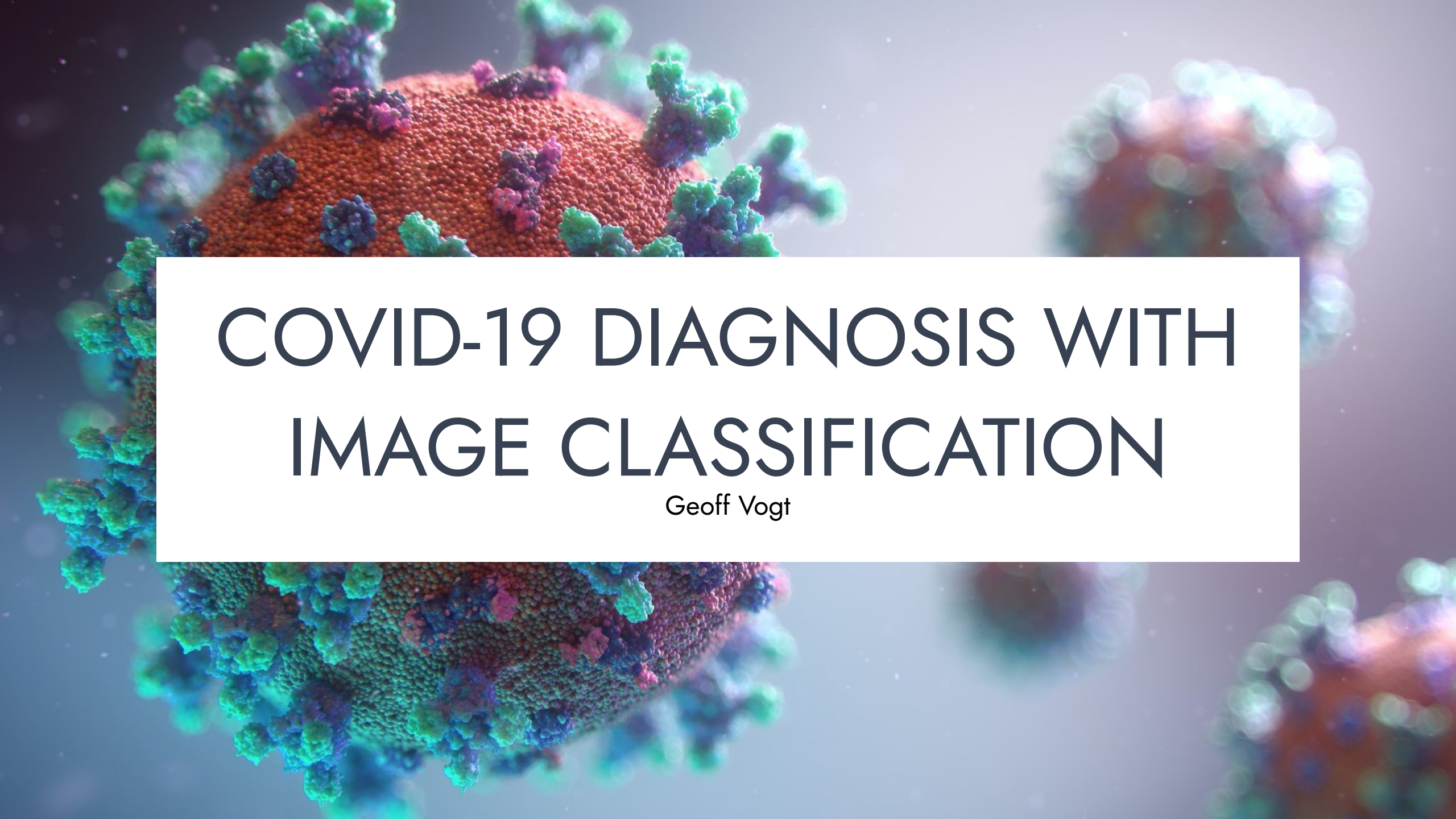


COVID-19 DIAGNOSIS WITH IMAGE CLASSIFICATION

Geoff Vogt

BUSINESS CASE

Develop an algorithm to analyze X-ray/CT images

Implement the algorithm into a user-friendly tool for medical professionals to easily upload X-ray images and receive a classification of COVID-19 or not

Work with medical institutions, hospitals, and clinics to market and distribute the tool to areas with limited access to PCR testing or where PCR tests are not immediately available

Stakeholders:

Medical professionals (e.g. doctors, nurses, radiologists)

FIGHTING THE PANDEMIC: X-RAY IMAGING FOR FAST COVID-19 DIAGNOSIS

COVID-19 has infected over 400 million people and caused more than 5 million deaths worldwide as of March 2023.

Source: World Health Organization

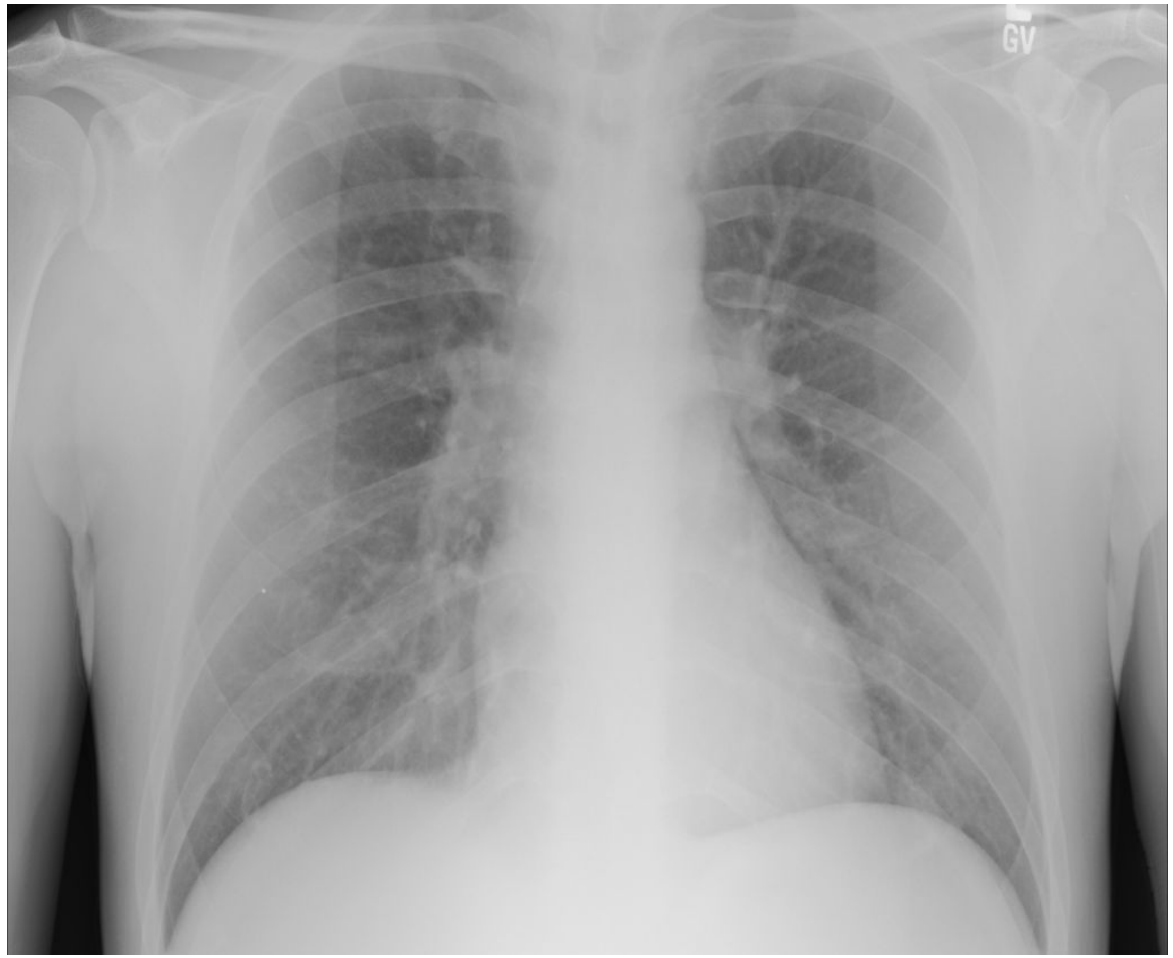
PCR testing, which is the most reliable diagnostic method for COVID-19, may not be readily available in some areas, leading to delayed diagnosis and treatment.

DATA UNDERSTANDING

- The dataset contains X-ray and CT images of both Non-COVID and COVID cases.
- The dataset has been augmented with different techniques to generate approximately 17,099 images.There are two main folders in the dataset, one for X-ray images and the other for CT images.
- The X-ray folder contains two sub-folders of 5,500 Non-COVID images and 4,044 COVID images.The CT folder also contains two sub-folders of 2,628 Non-COVID images and 5,427 COVID images.
- The dataset was published on June 12, 2020, and is currently in Version 3. The contributors to the dataset are Walid El-Shafai and Fathi E. Abd El-Samie. https://data.mendeley.com/datasets/8h65ywd2jr

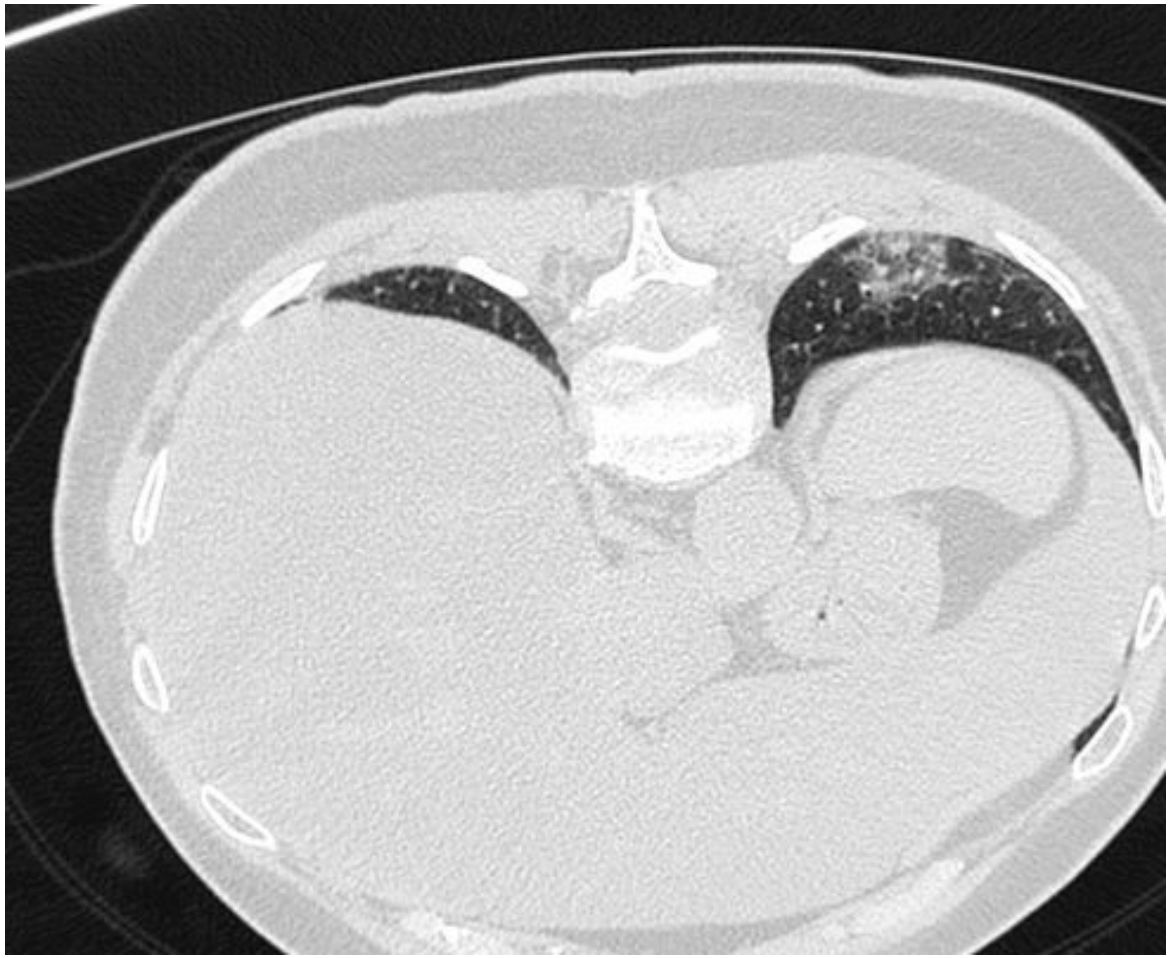
COVID VS. NON-COVID XRAY

COVID

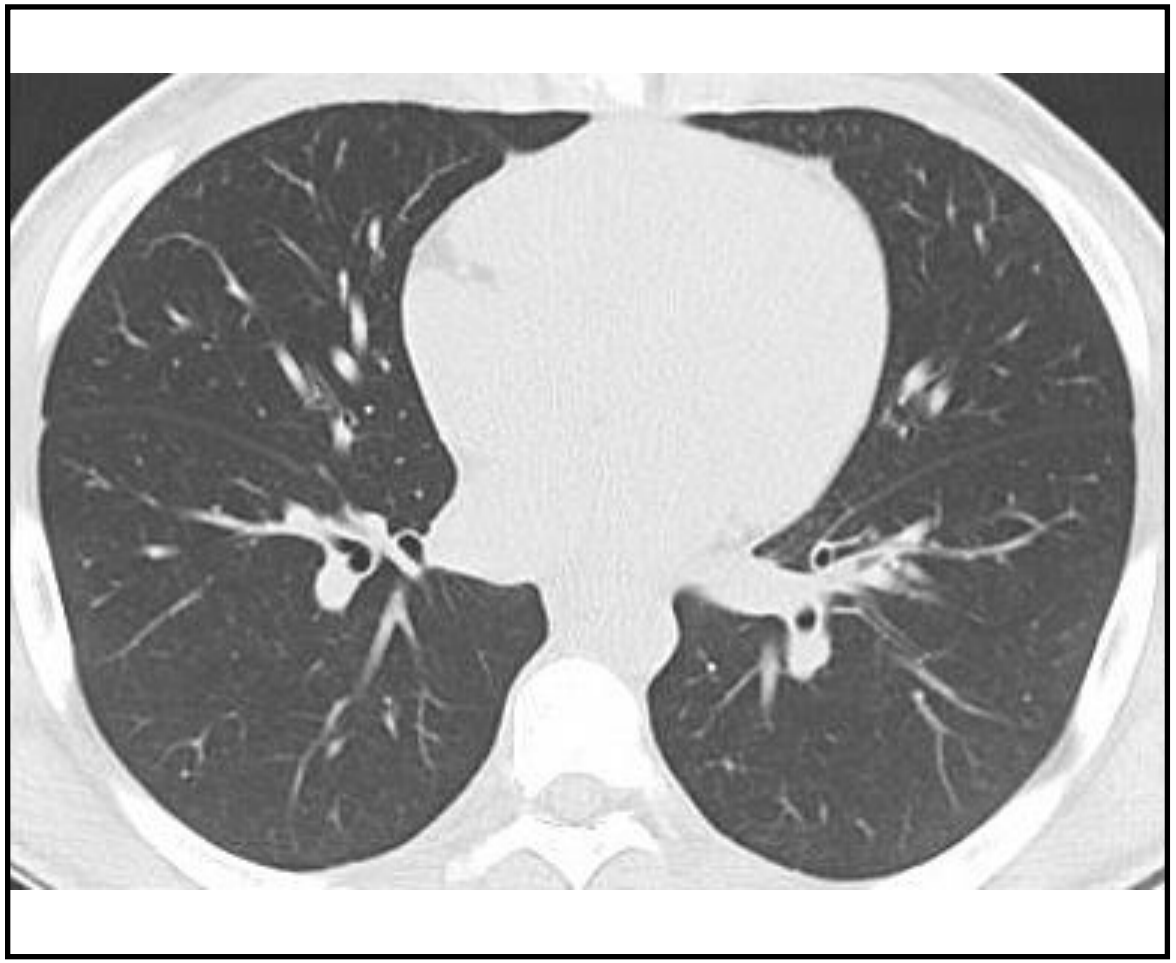


NON-COVID

COVID VS. NON-COVID CT

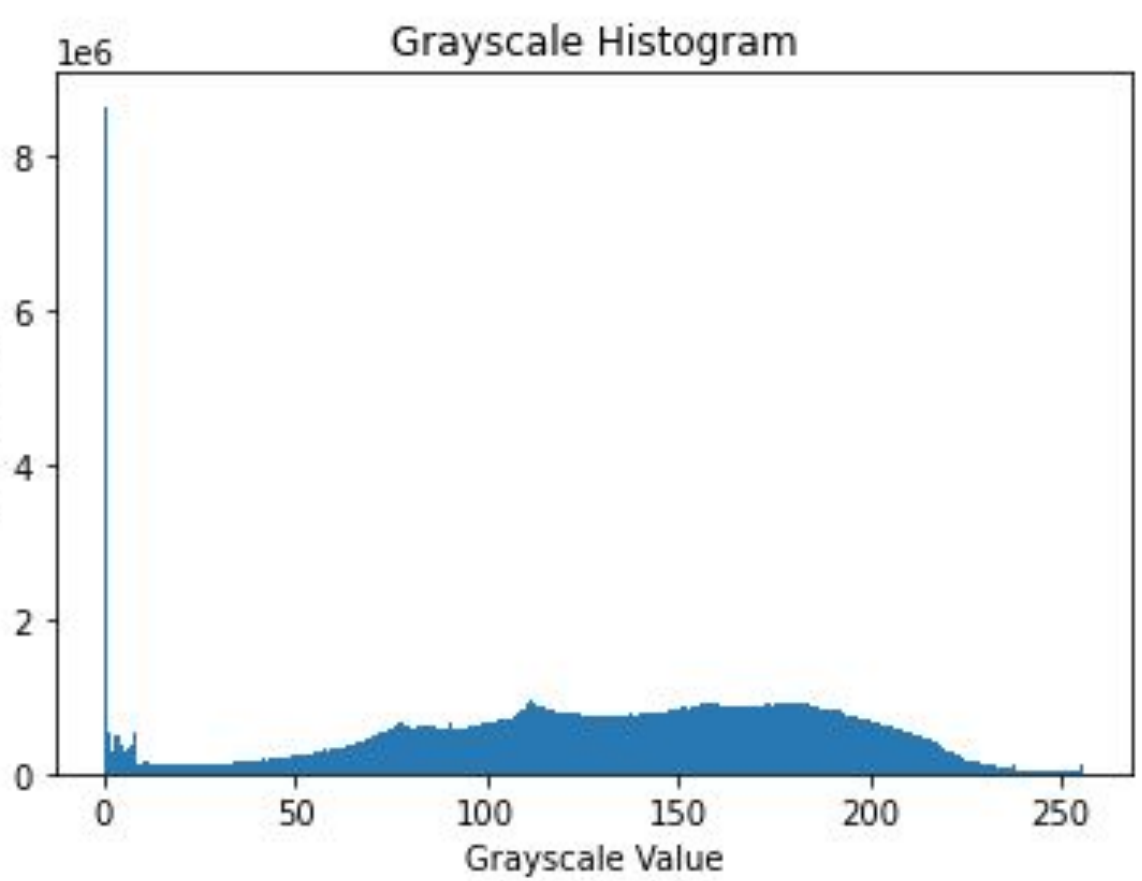


COVID

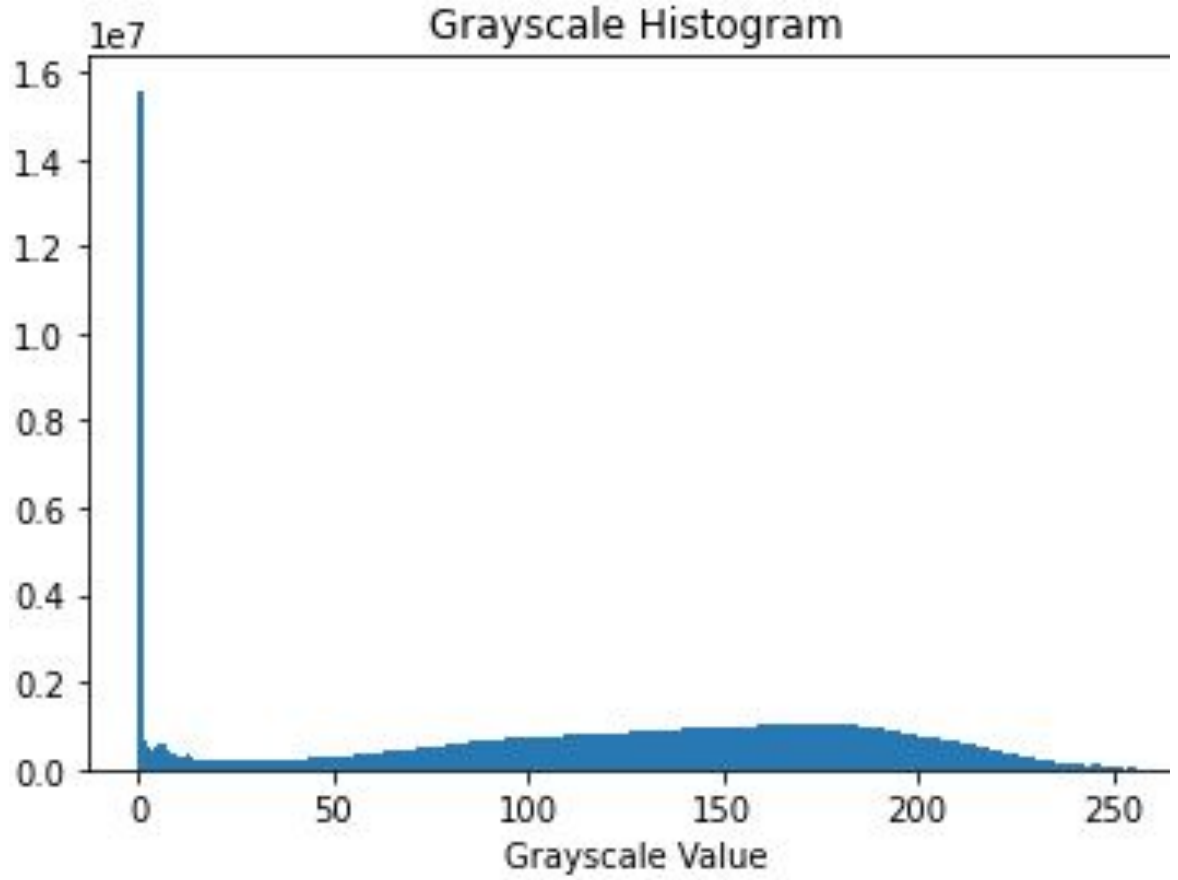


Non-COVID

X-RAY GRAYSCALE DISTRIBUTION

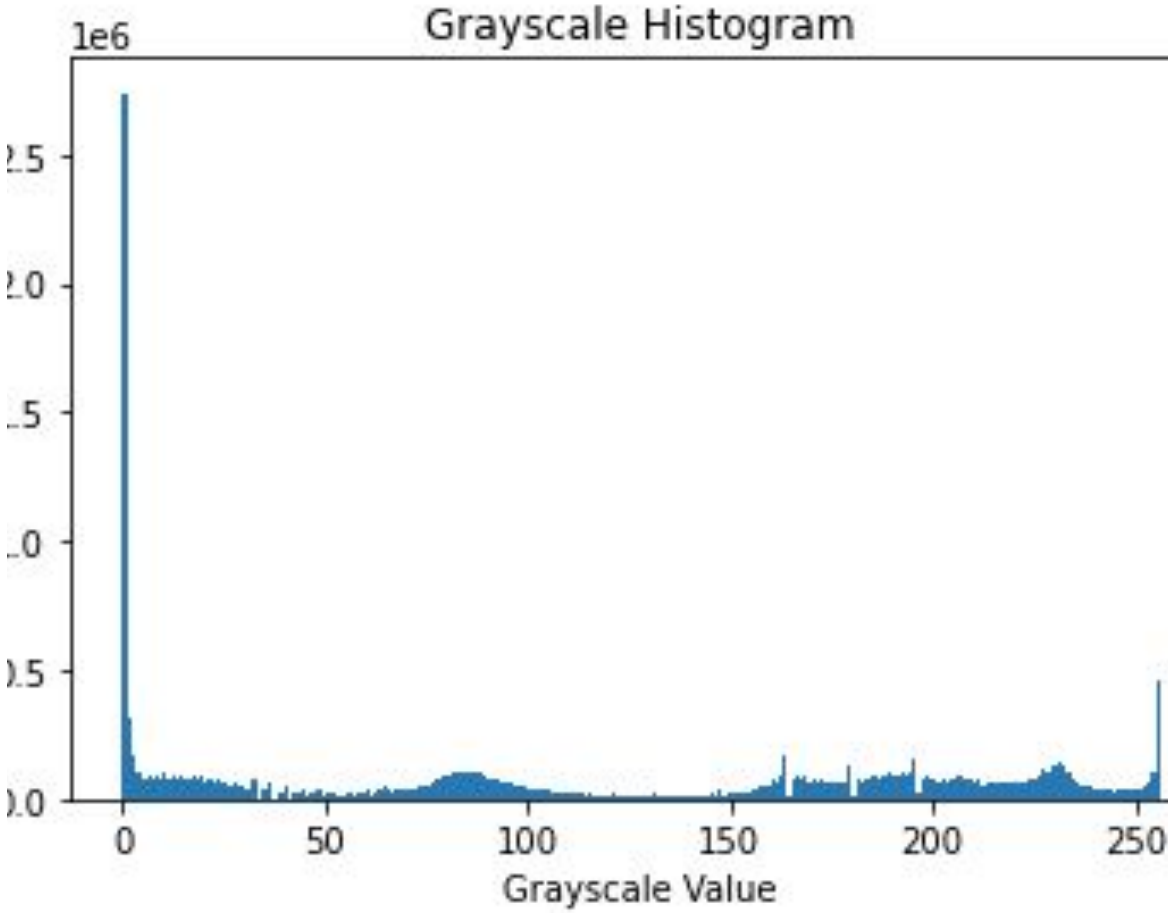


XRAY COVID

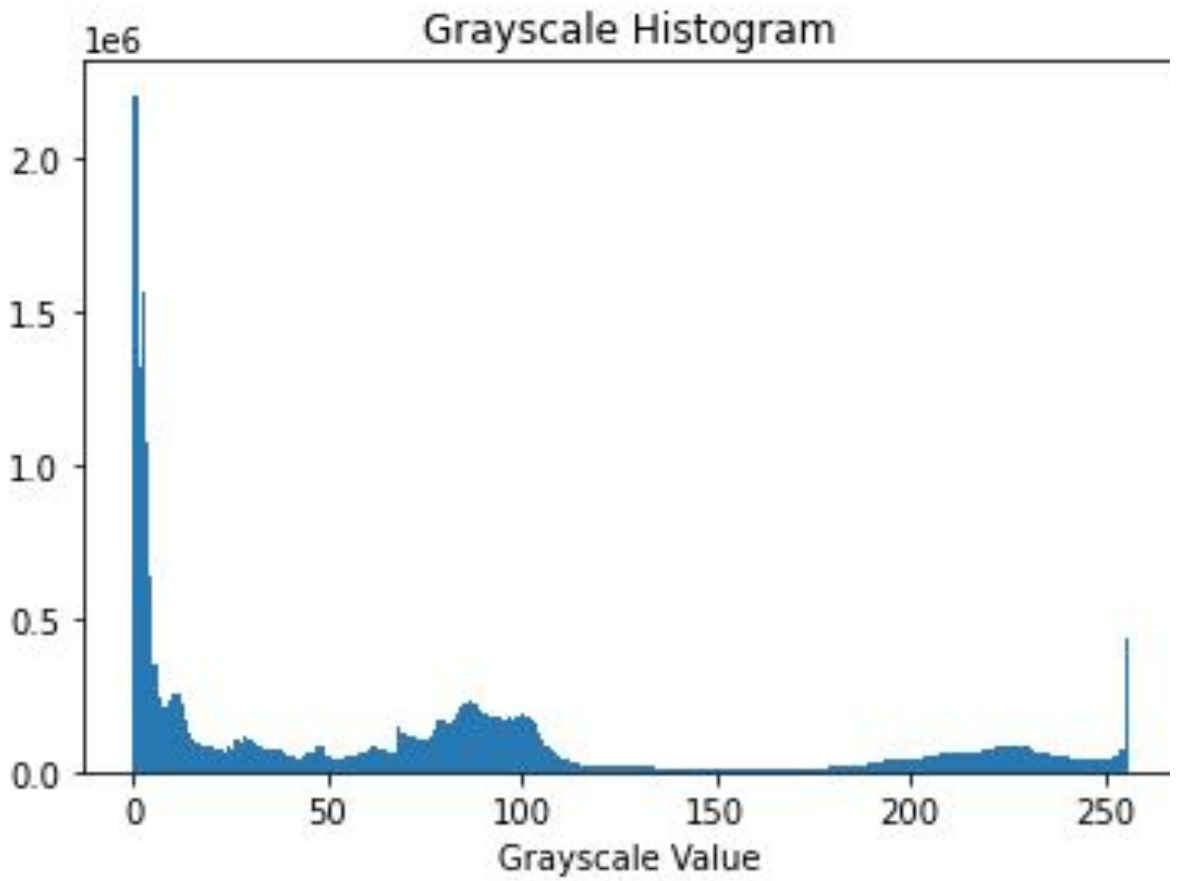


XRAY NON COVID

COVID-19 AND NON-COVID GRAYSCALE DISTRIBUTION



CT COVID



CT NON COVID

DATA MODELING

1. Data processing

Two sets of convolutional neural network models were run for X-Ray and CT images

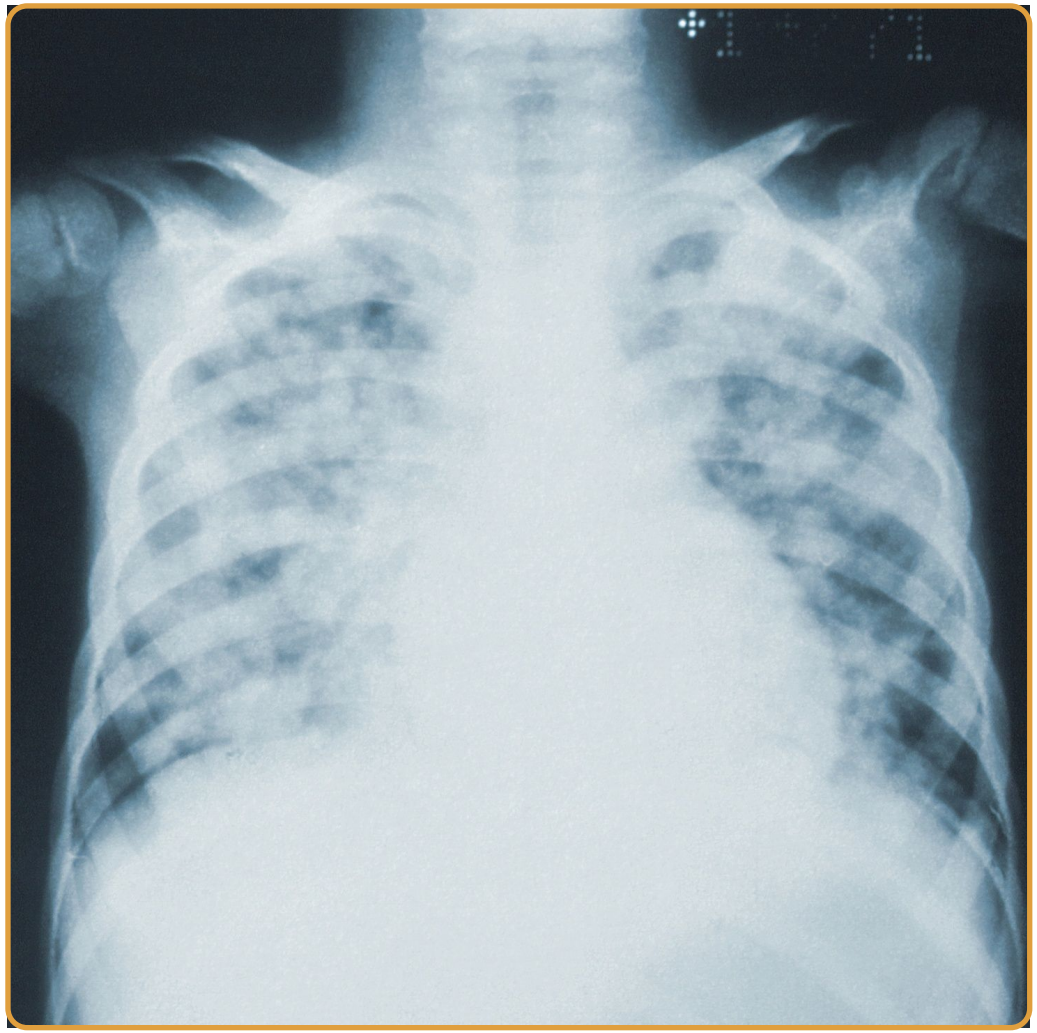
2. Iterative modeling

5 different models per X-Ray and CT iterating model complexity using additional convolutional layers, max-pooling, adding dense layers, regularization and other hyper-parameter tuning

3. Final model

The strongest model for each set was chosen as the final model.

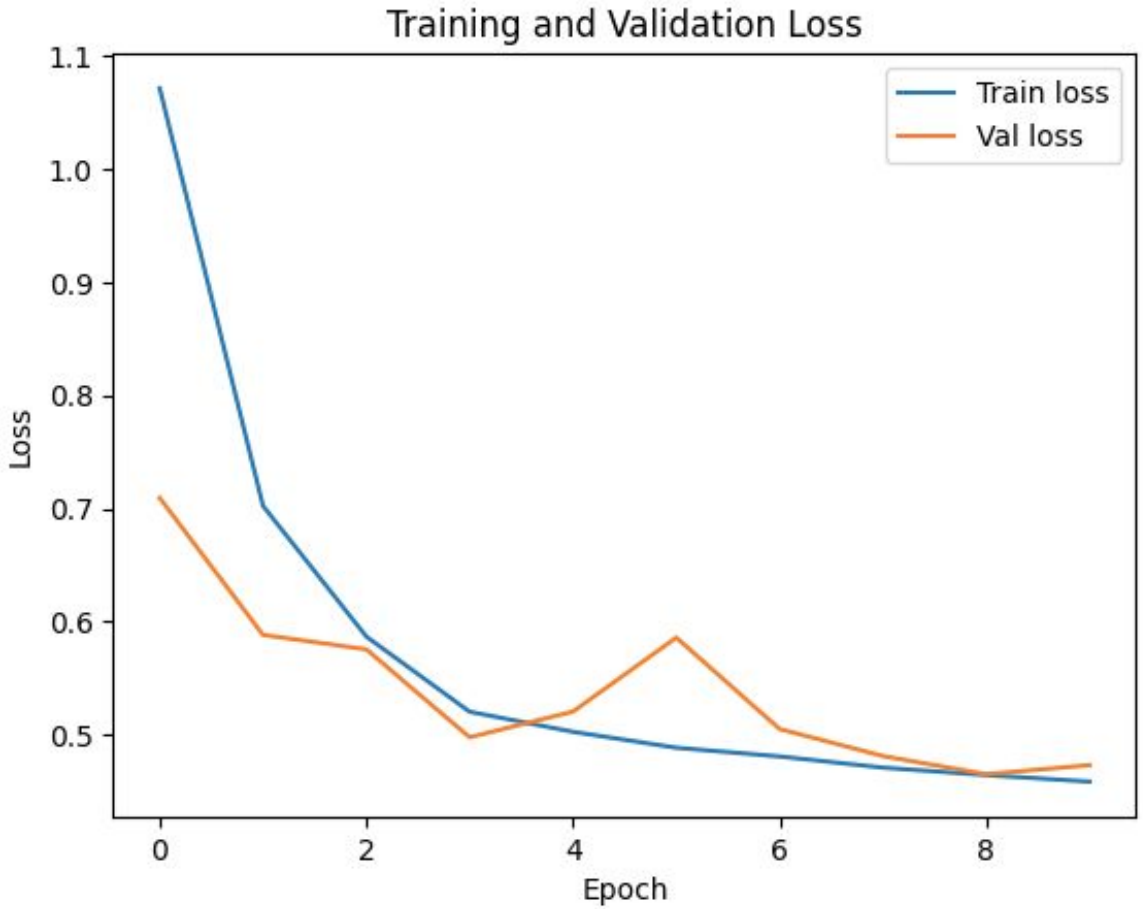
X-RAY FINAL RESULTS



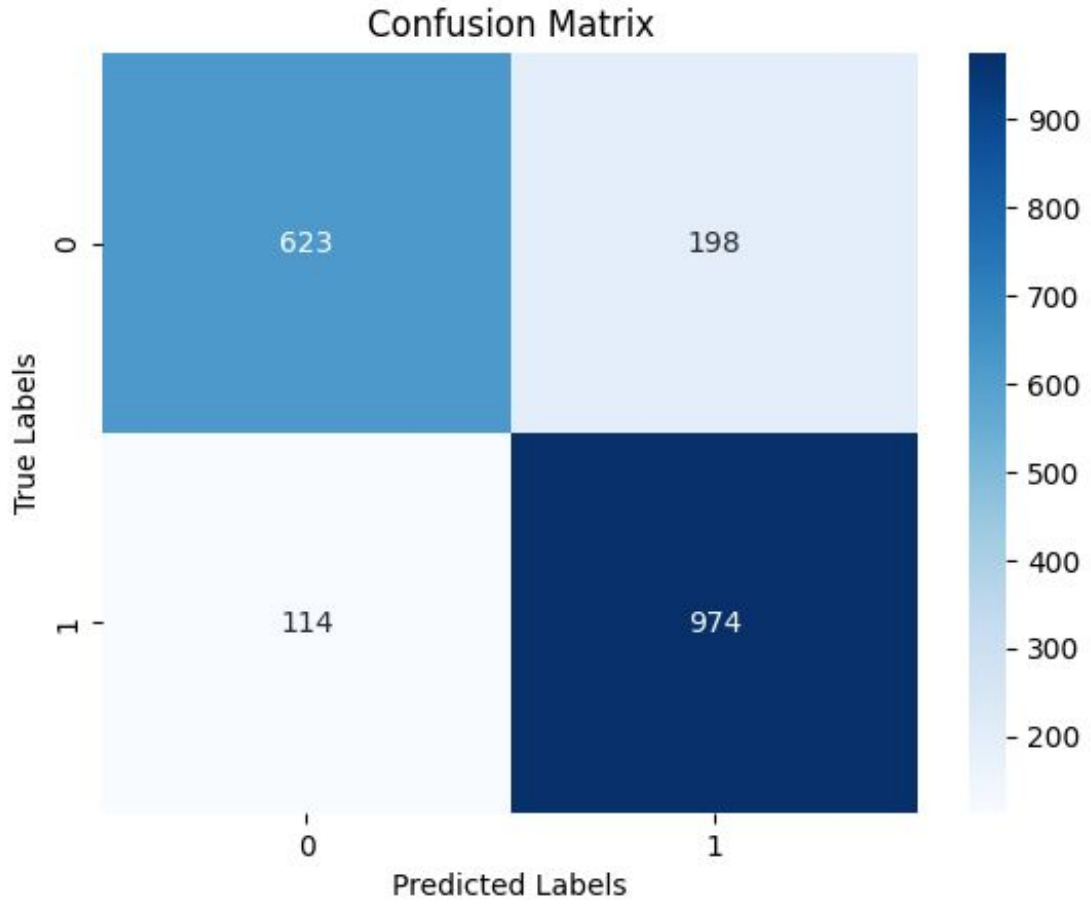
Convolutional Neural Network

- Test Accuracy: .84
- Recall: .83
- Loss: ..43

TRAINING CURVE



CONFUSION MATRIX



CT FINAL RESULTS

Convolutional Neural Network

- Accuracy: 88%
- Recall: .86
- Loss: .42

RECOMMENDATIONS + NEXT STEPS

1. CT scan is better at predicting COVID than X-Ray
2. Model can still be useful for a generalization of risk, but probably not ideal for diagnosis
3. Look at differences between pneumonia and COVID
4. Further improvements can be implemented with additional layers in the neural net, data augmentation or additional hyper-parameter tuning

QUESTIONS?

Email: geoffrey.m.vogt.com

LinkedIn: https://www.linkedin.com/in/geoffvogt/

Github: https://github.com/gvogt2023