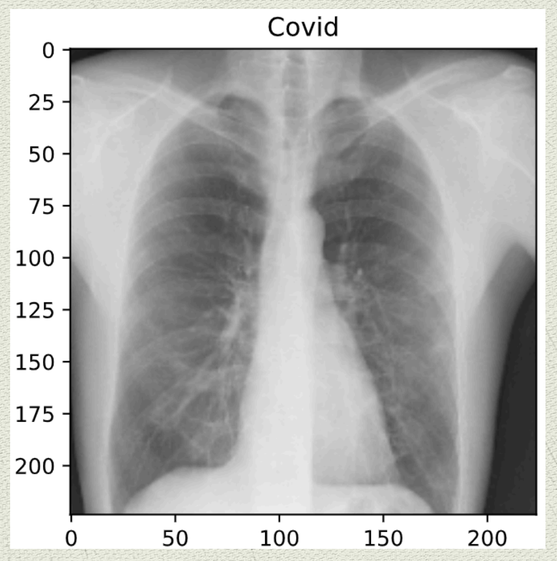
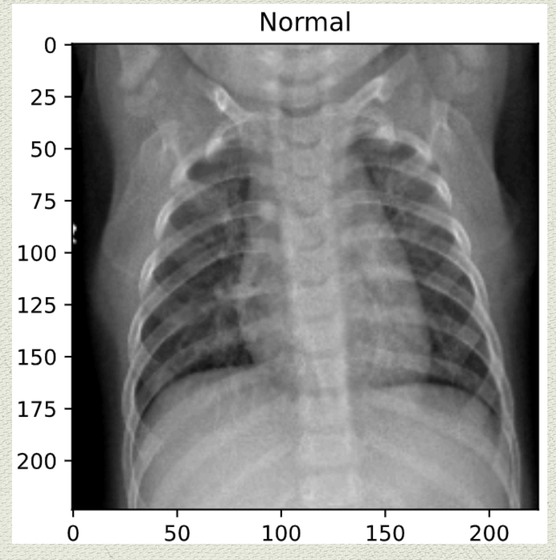
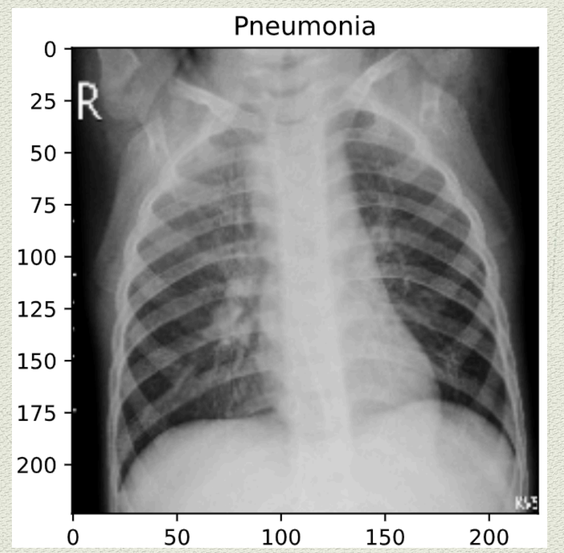
Medical Computer Vision

Ronen Reouveni







The Goals

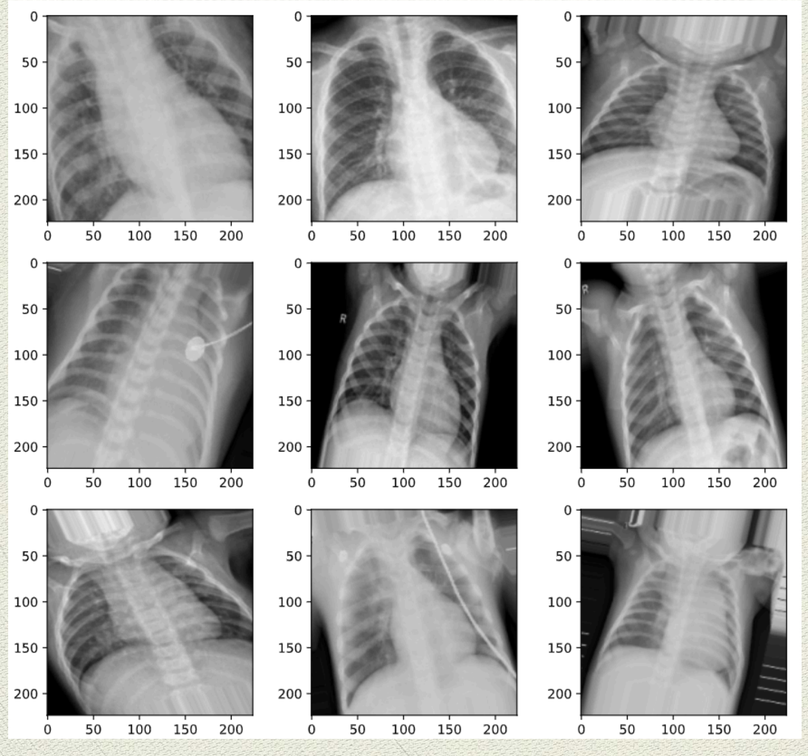
- Can a CNN actually tell the difference between Covid and Pneumonia?
- A secondary but critical goal is to visually depict which features each model is learning.

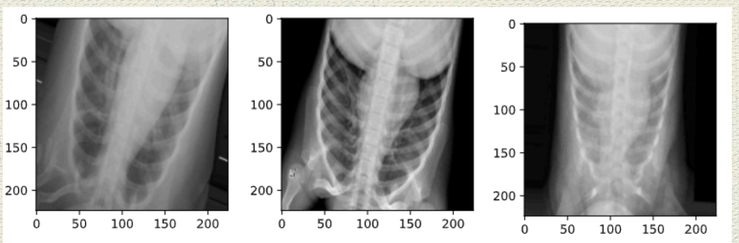
The Data

- The main problem is that there are only about 6,000 X-Rays, there are two ways to mitigate this.
    - Data Augmentation
    - Transfer Learning

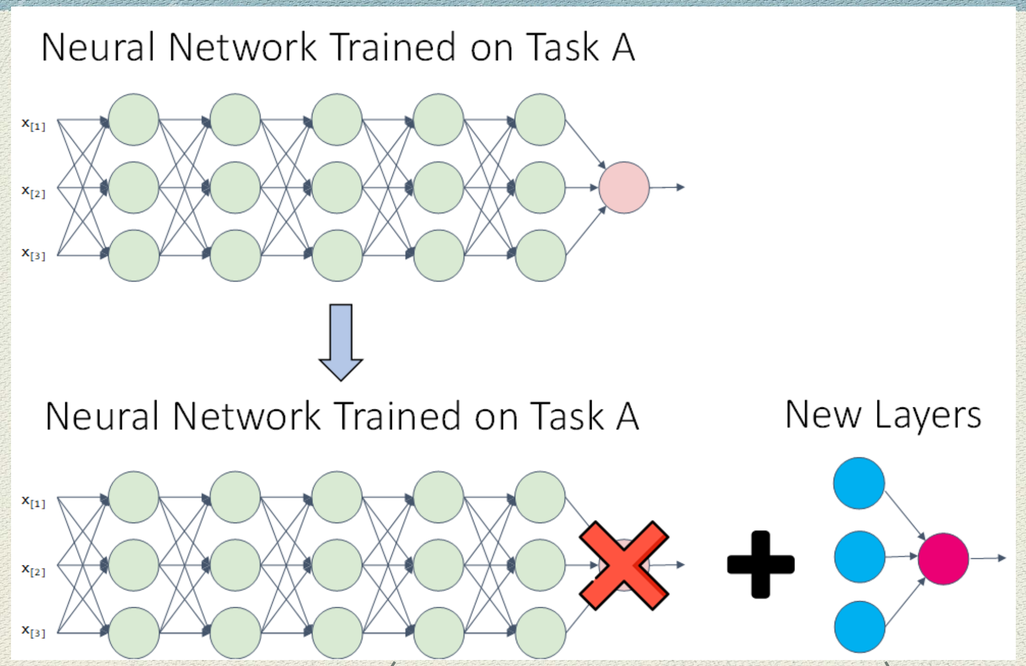
What is Data Augmentation

```
train_datagen = ImageDataGenerator(rescale = 1./255,
                                   zoom_range = 0.3, #.2
                                   rotation_range=20, #15
                                   vertical_flip = False,
                                   horizontal_flip = False)

test_datagen = ImageDataGenerator(rescale = 1./255)

inputShape = (224, 224)
training_set = train_datagen.flow_from_directory(data_path + '/train',
                                                 target_size = inputShape,
                                                 batch_size = 16,
                                                 class_mode = 'categorical',
                                                 shuffle=True)

test_set = test_datagen.flow_from_directory(data_path + '/test',
                                            target_size = inputShape,
                                            batch_size = 16,
                                            class_mode = 'categorical',
                                            shuffle = False)
```





What is Transfer Learning



- We can leverage neural networks trained by large institutions using super computers and massive data sets.

This section will only do feature transformation

This section will learn how the transformed features predict our specific problem.

Which models will be leveraged?

- VGG16
- ResNet152V2
- mobilenet_v3
- A custom DNN as a baseline

New Head

```
max_pooling2d_9 (MaxPooling2D)
dropout_16 (Dropout)
conv2d_6 (Conv2D)
batch_normalization_6 (BatchNor
conv2d_7 (Conv2D)
batch_normalization_7 (BatchNor
dropout_17 (Dropout)
flatten_5 (Flatten)
dense_17 (Dense)
dropout_18 (Dropout)
dense_18 (Dense)
dropout_19 (Dropout)
dense_19 (Dense)
```

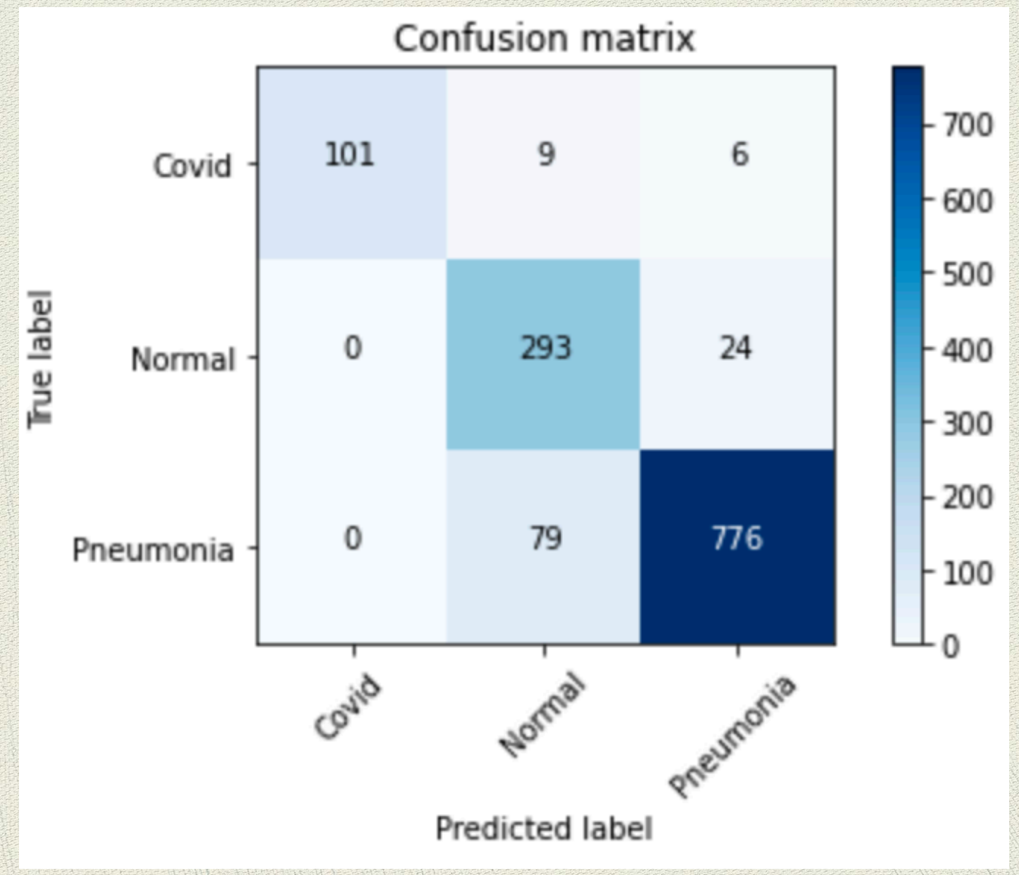
How did Each Model Perform?

	Accuracy	Params
VGG16	89%	36M
ResNet152V2	94%	130M
mobilenet_v3	91%	49M
DNN	74%	38M

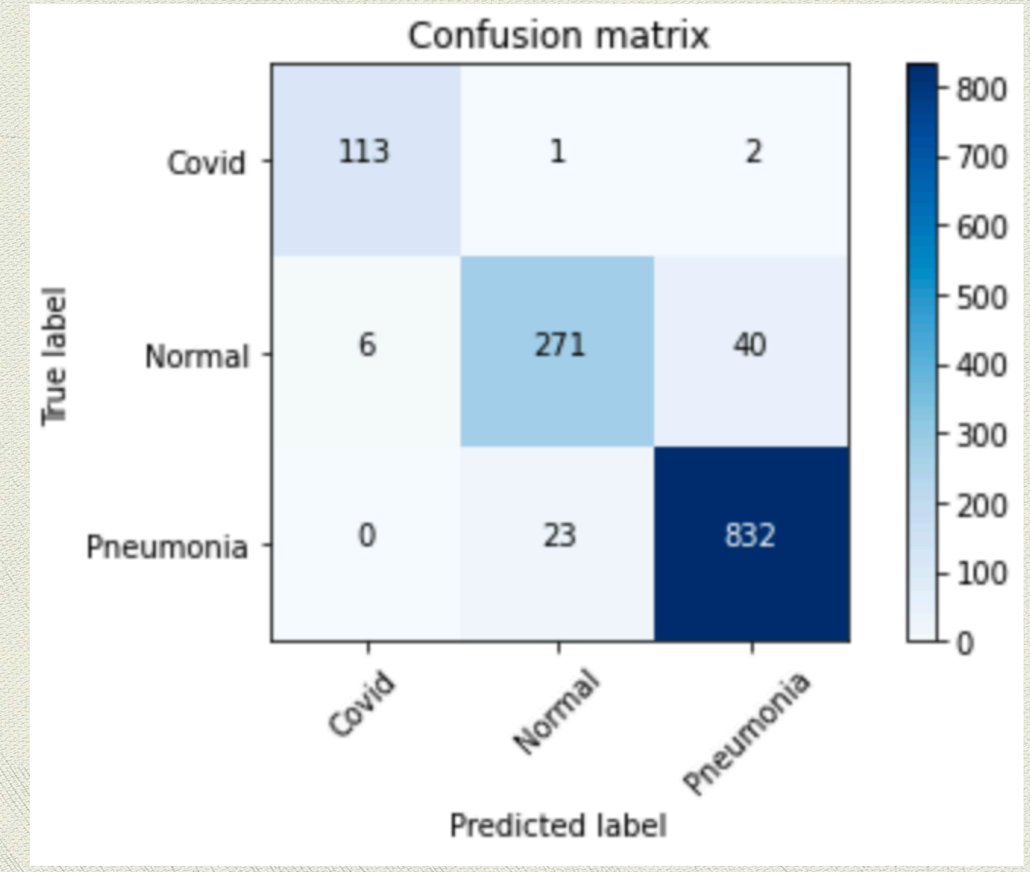
All of these took under 1 hour with no GPU

Confusion Matrices

Mobilnet



Resnet



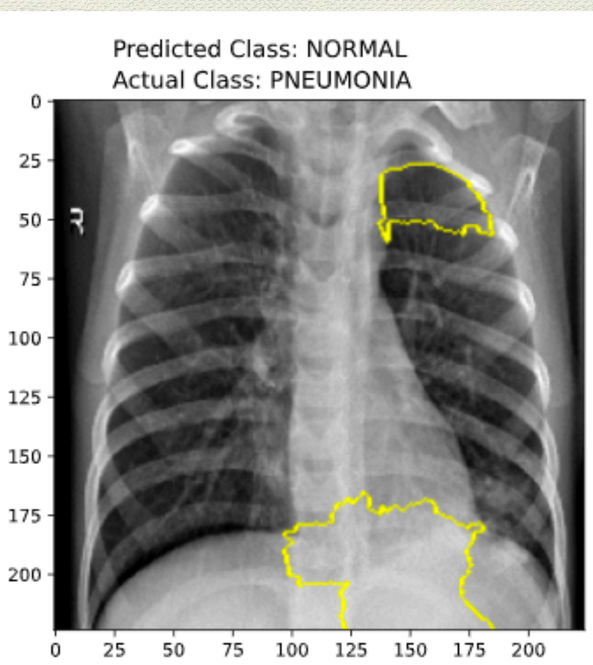
What features are being learned?

- “LIME is a python library that tries to solve for model interpretability by producing locally faithful explanations.”
- An observation is selected and passed into a model a number of times. Each time it is changed, by tracking the changes it can explain which features are important.

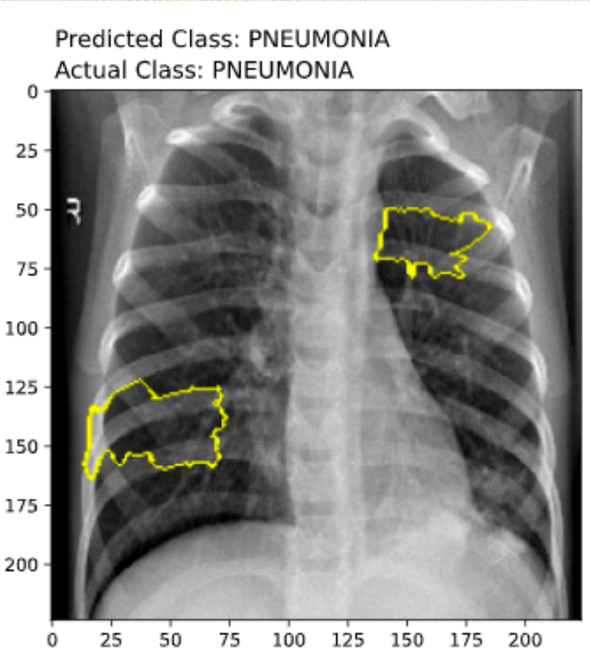
LIME in Action

Only VGG seems to notice the correct features

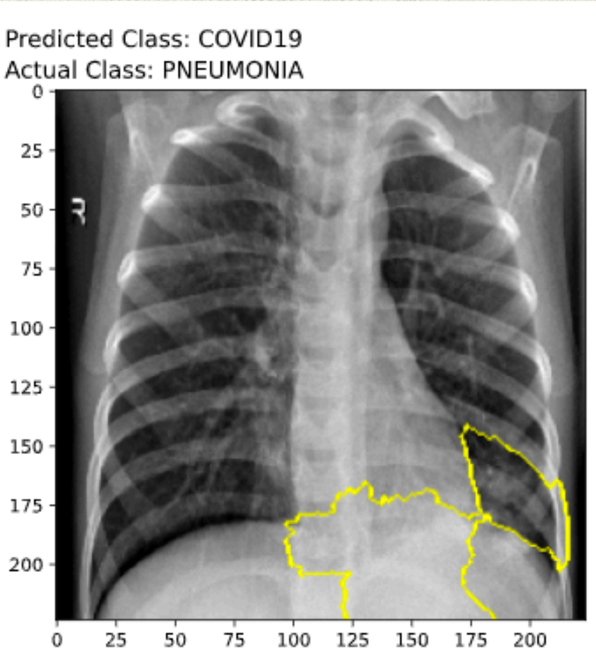
DNN



VGG16



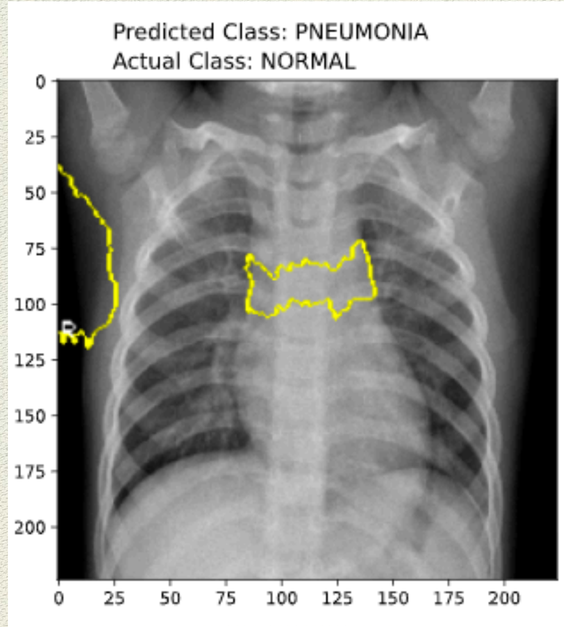
Mobilnet



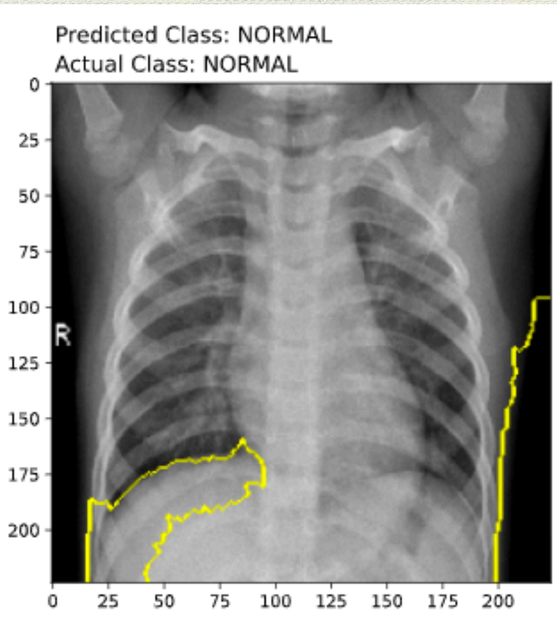
LIME Exposes Problems

This shows that some features it learns are not the disease itself. These X-Rays were classified based on features outside the lungs themselves.

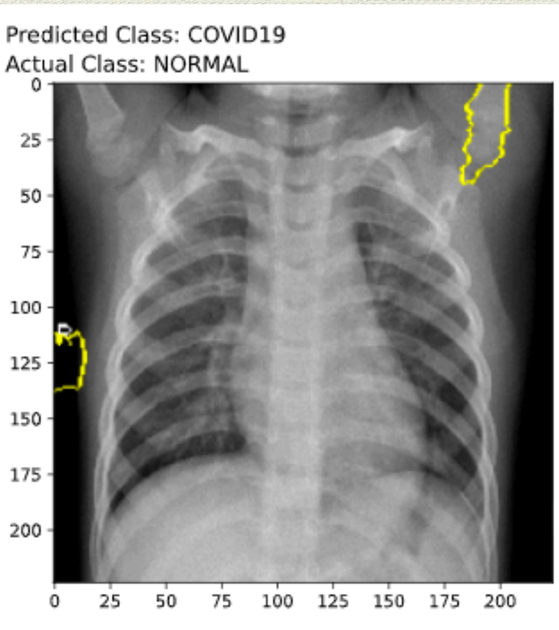
DNN



VGG



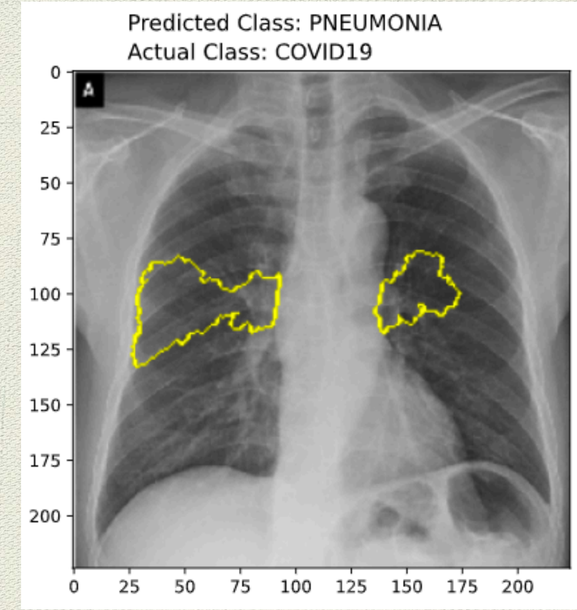
Mobilnet



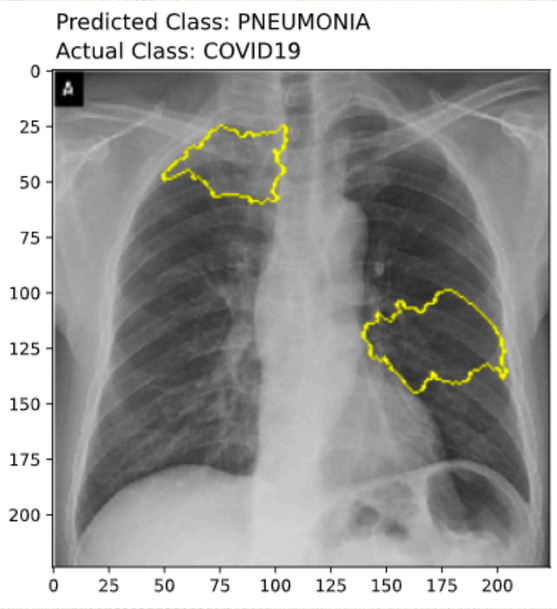
LIME Again

What happened to Resnet?

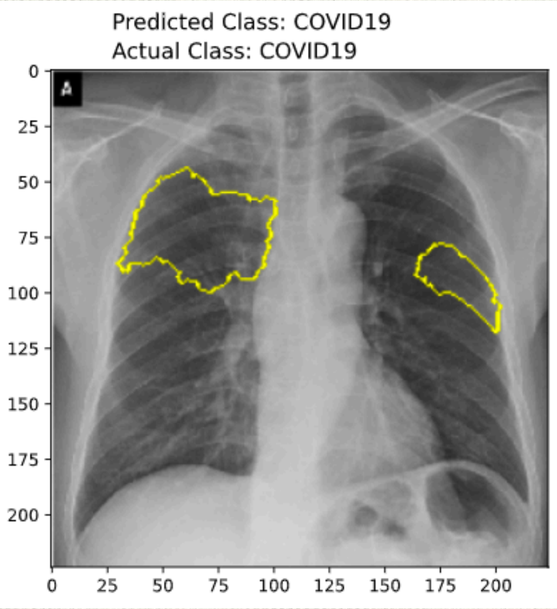
DNN



VGG16



Mobilnet



LIME did not actually work on any Resnet implementation. My hypothesis is that is due to the number of params and style architecture

Recommendation

- I was able to obtain a nice classification rate, but LIME showed it may not be trusted.
- It does however, prove that the concept and idea is feasible. This has implications in medical computer vision that are broad.
    - This has the ability to revolutionize healthcare

Improvements and Questions

- The main way this can be improved is more data.
- The second would be cropping out everything in the X-ray but the lungs themselves
- Questions?