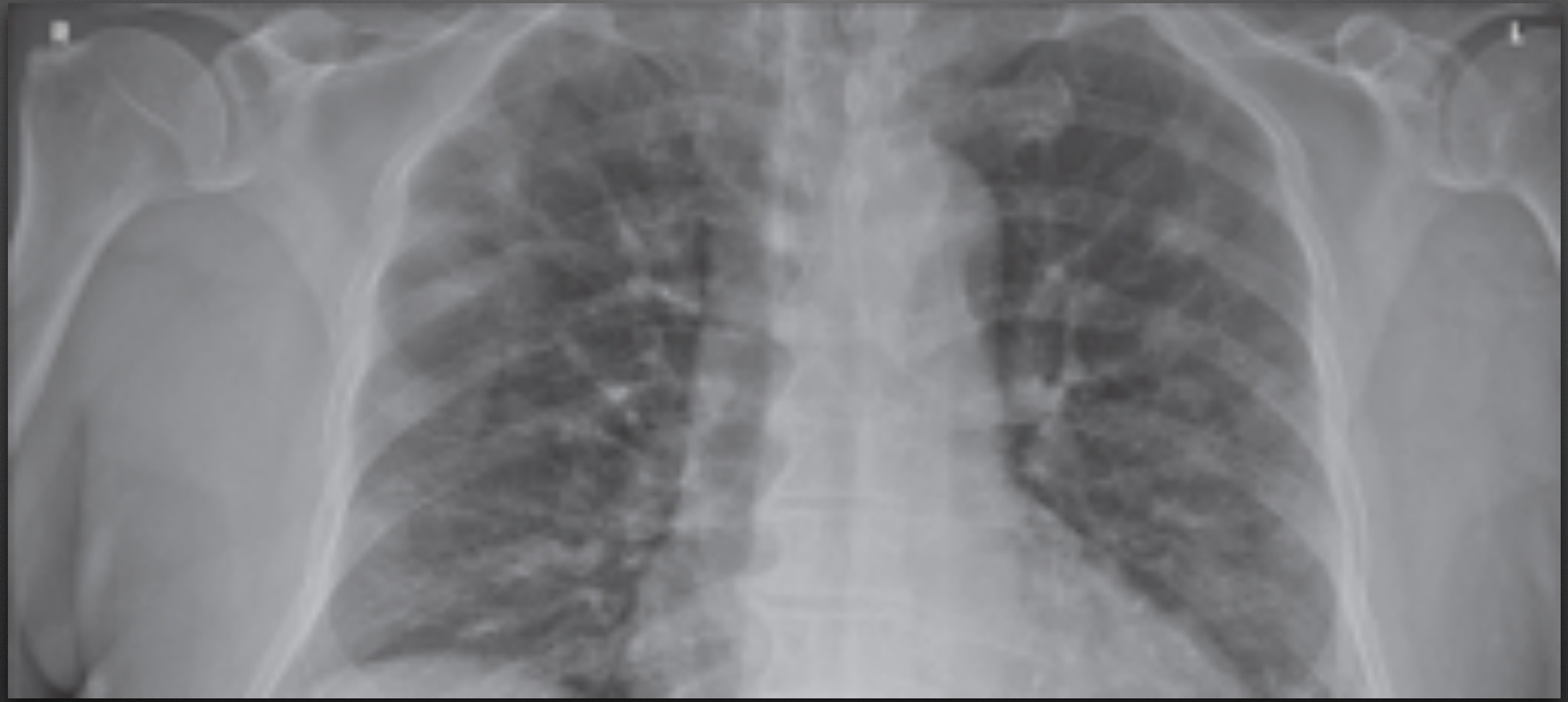
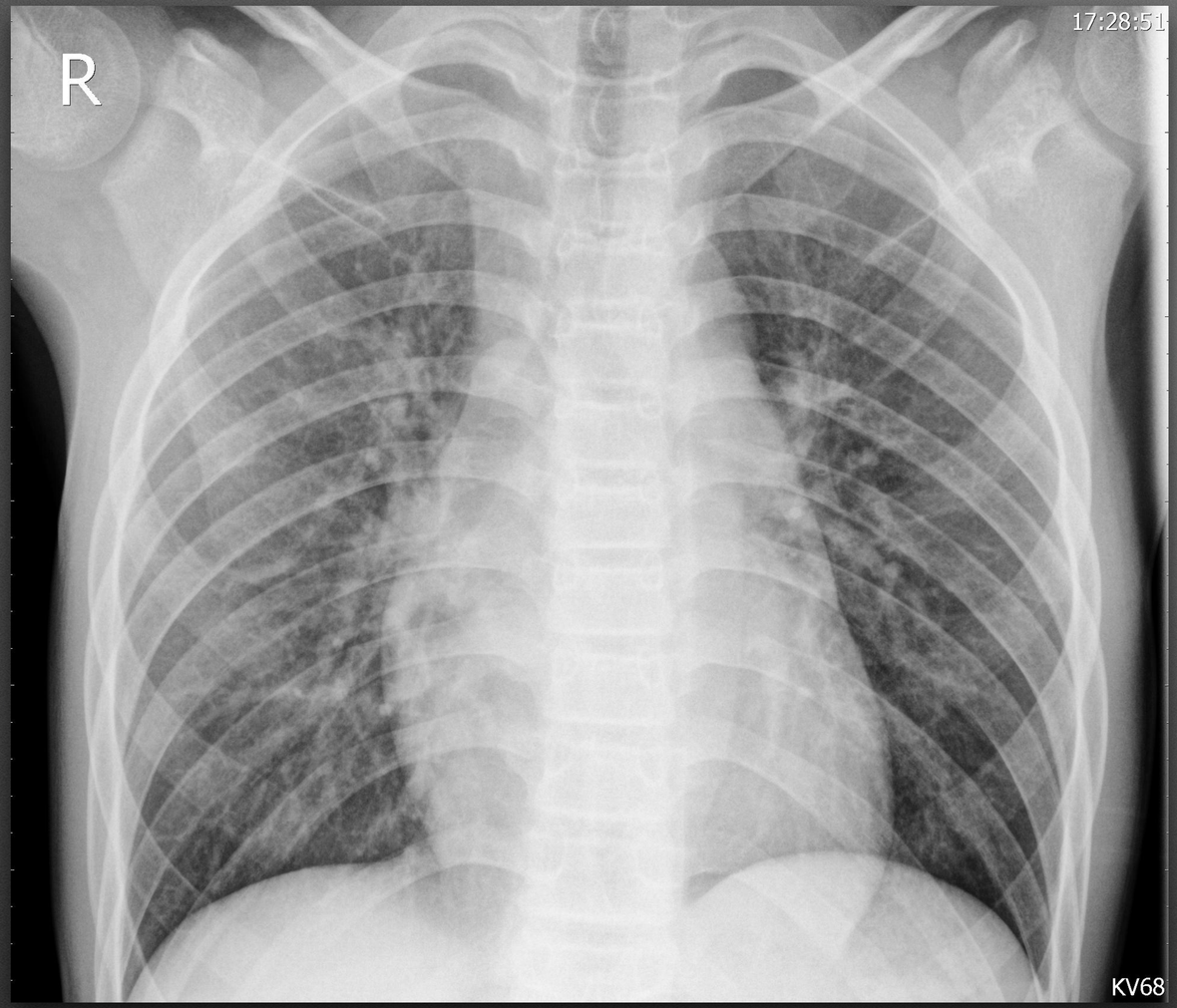
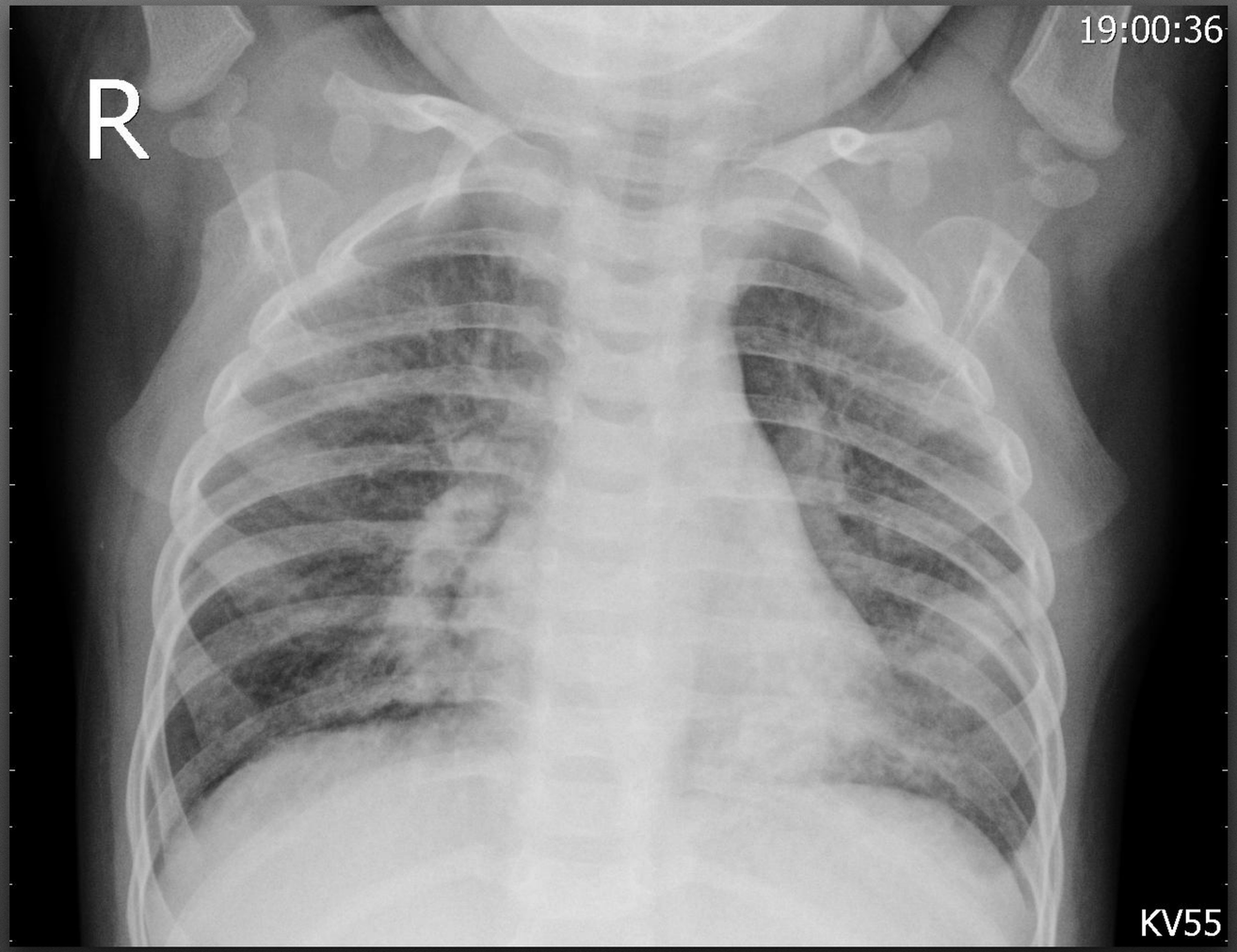
Chest X-Ray Classification

Brent Thayer



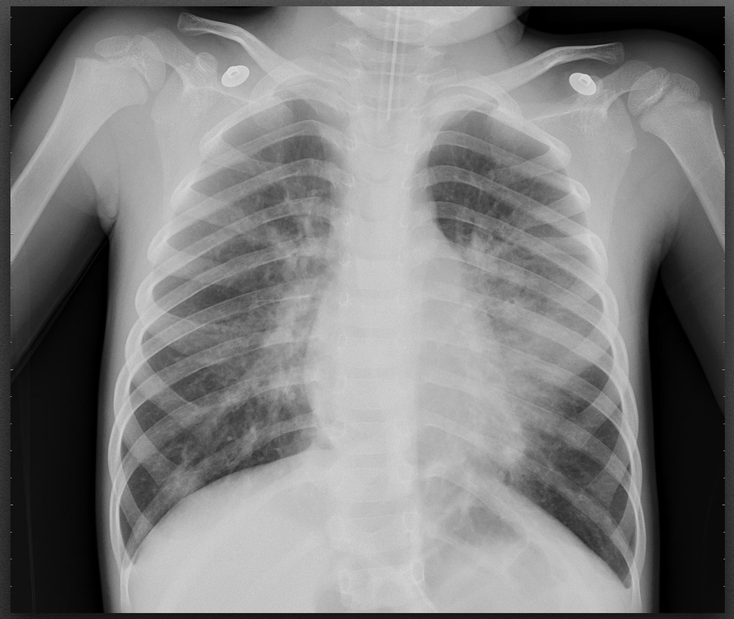


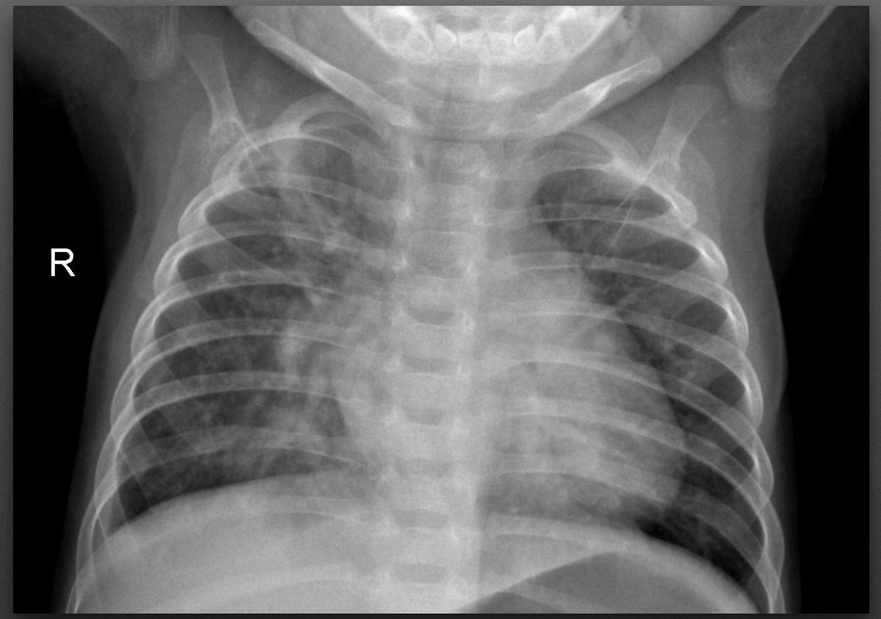


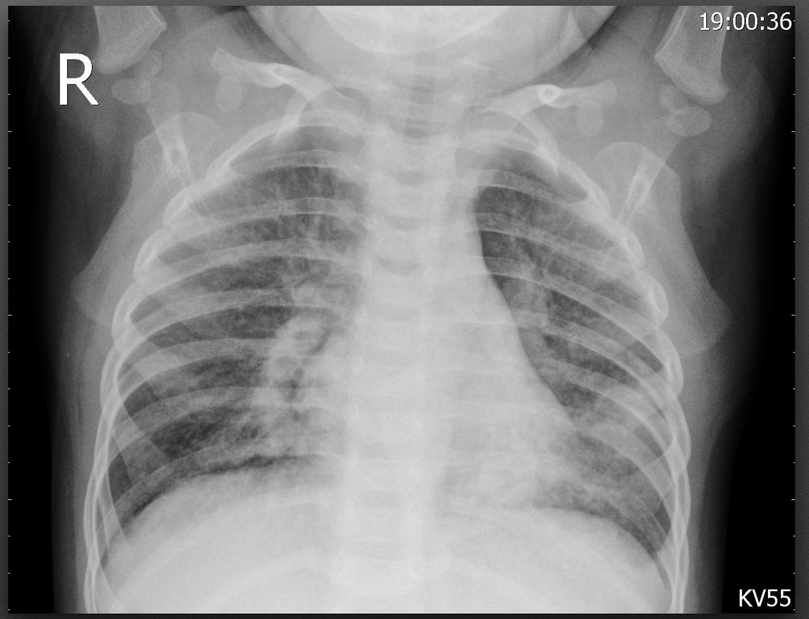
Is it possible to train a model to classify pneumonia from a chest X-ray?

An X-ray allows a radiologist to see the lungs, heart and blood vessels. When interpreting the X-ray, the radiologist will look for white spots in the lungs (called infiltrates) that identify an infection.

Pneumonia

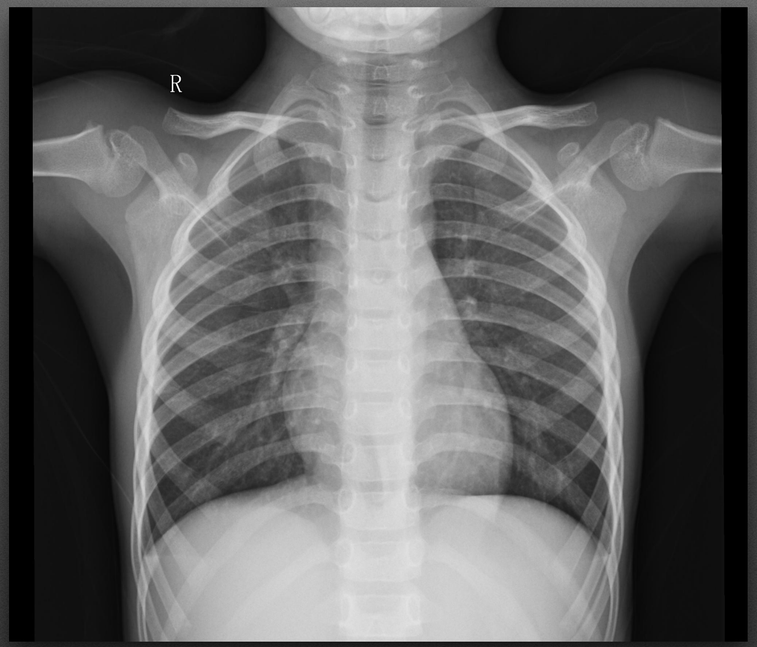


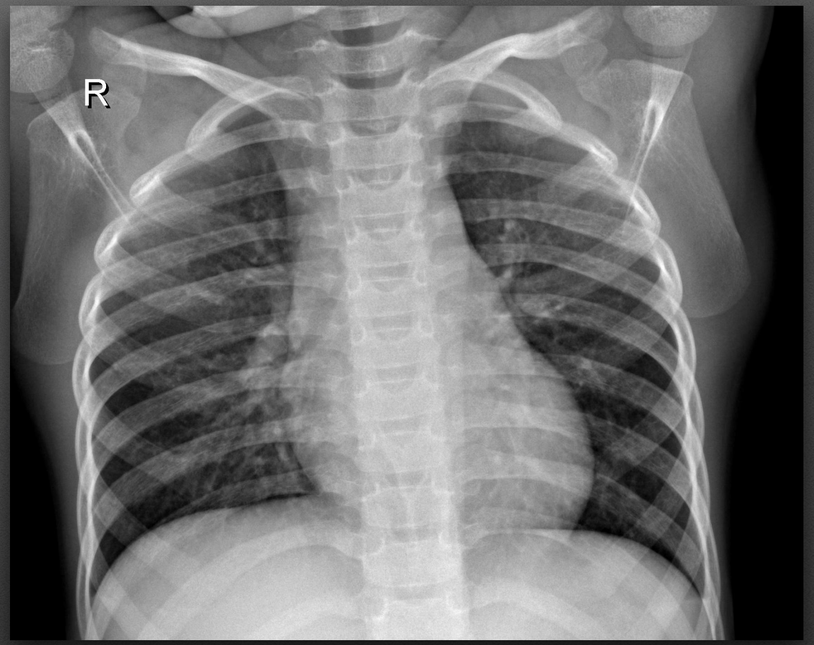




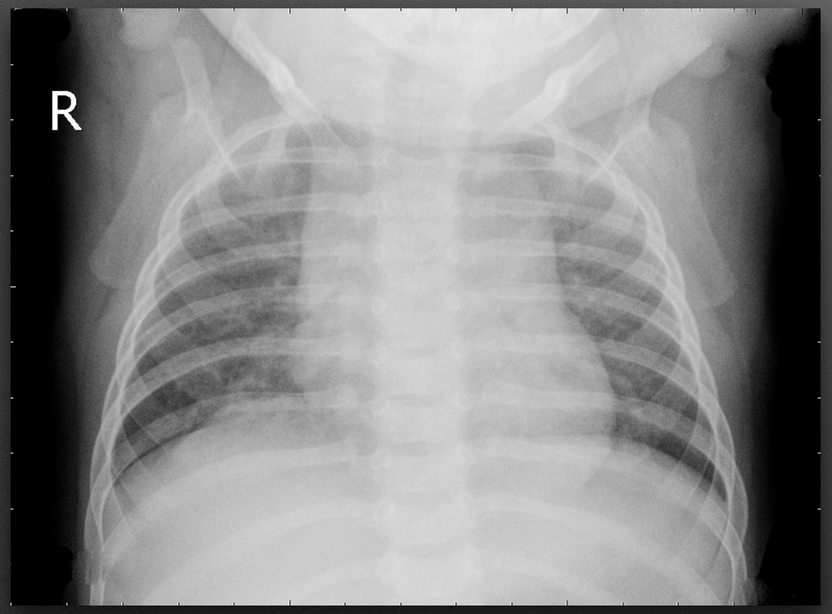
Pneumonia is an infection that inflames the air sacs in one or both lungs. The air sacs may fill with fluid or pus (purulent material), causing cough with phlegm or pus, fever, chills, and difficulty breathing. A variety of organisms, including bacteria, viruses and fungi, can cause pneumonia.

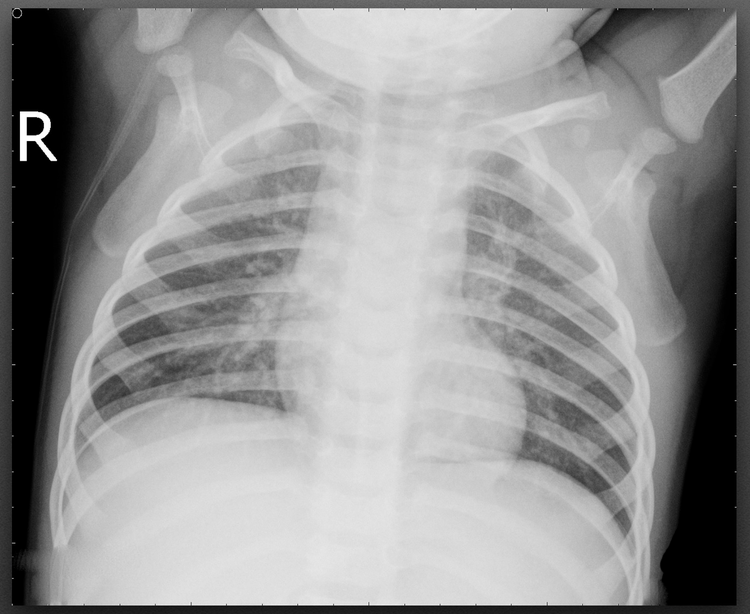
Normal



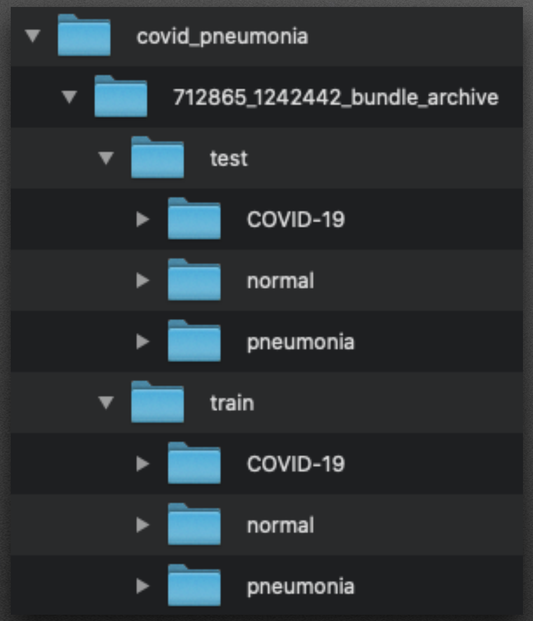


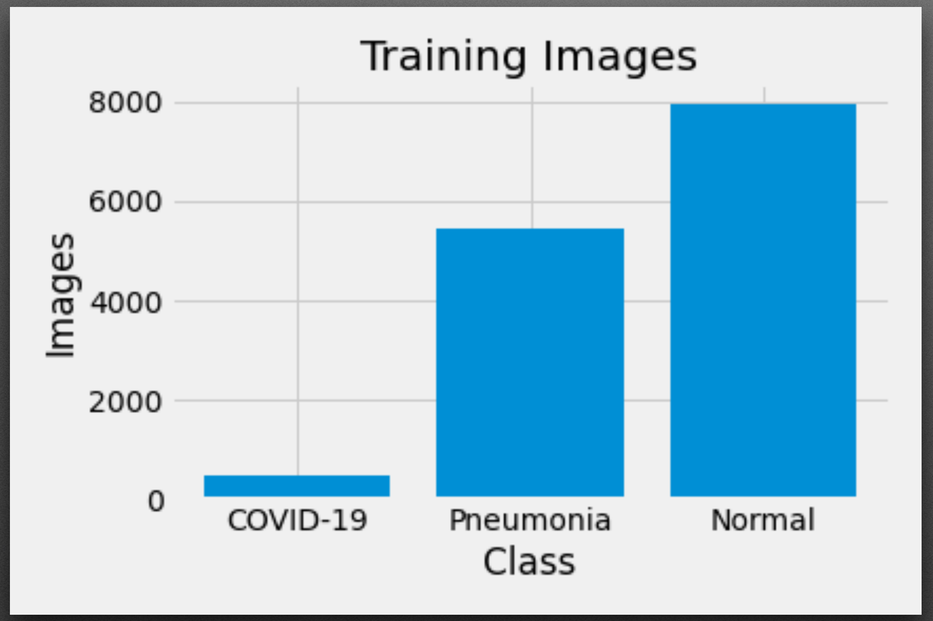
Anterior X-rays

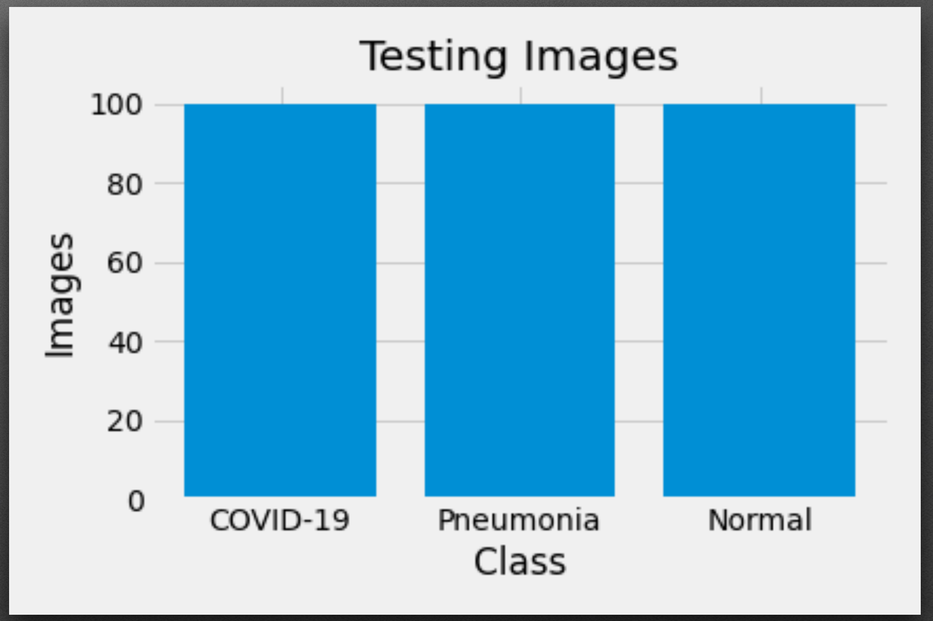




First Dataset





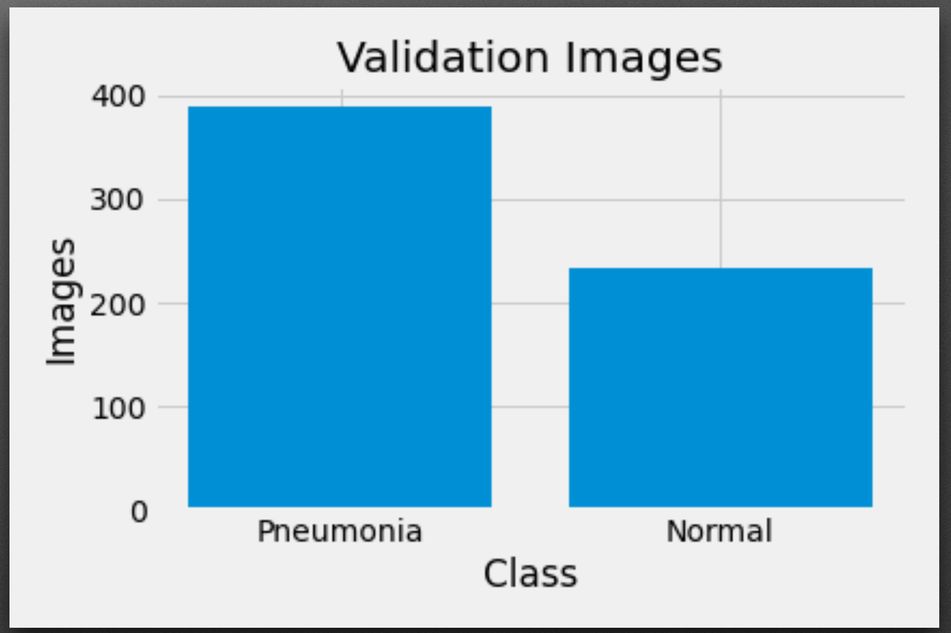


COVID-19 was removed to create a binary classification problem.

The model was learning to always predict normal.

New Dataset

	Images
test	
NORMAL	8
PNEUMONIA	8
train	
NORMAL	1341
PNEUMONIA	3875
val	
NORMAL	234
PNEUMONIA	390

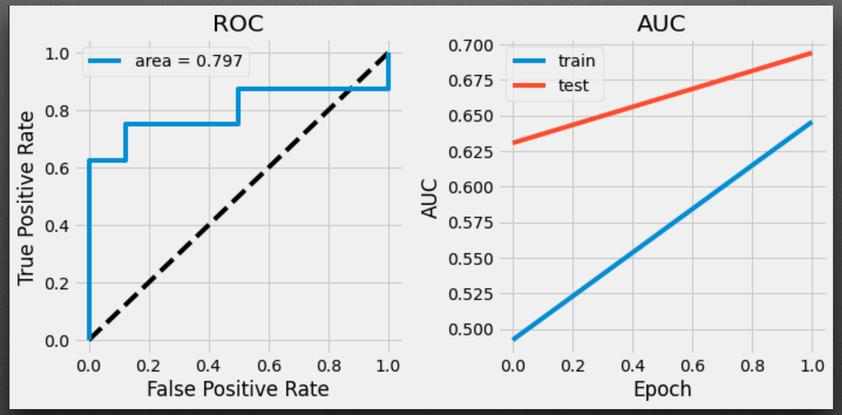


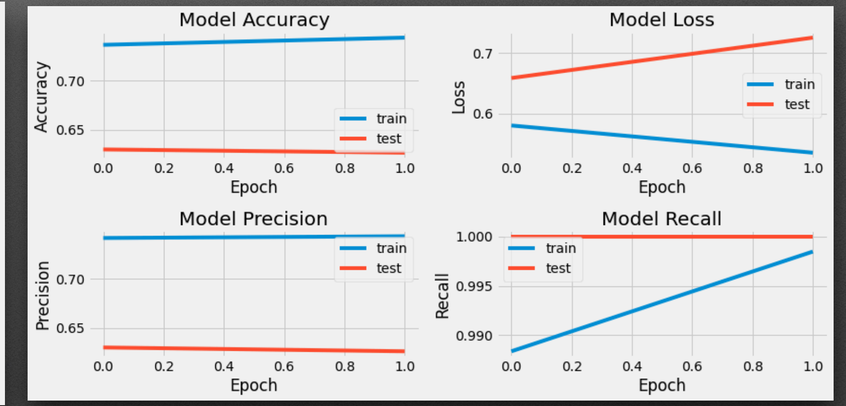
Interestingly enough, the model was now learning to always predict pneumonia after only a couple batches had processed.

Model Progress

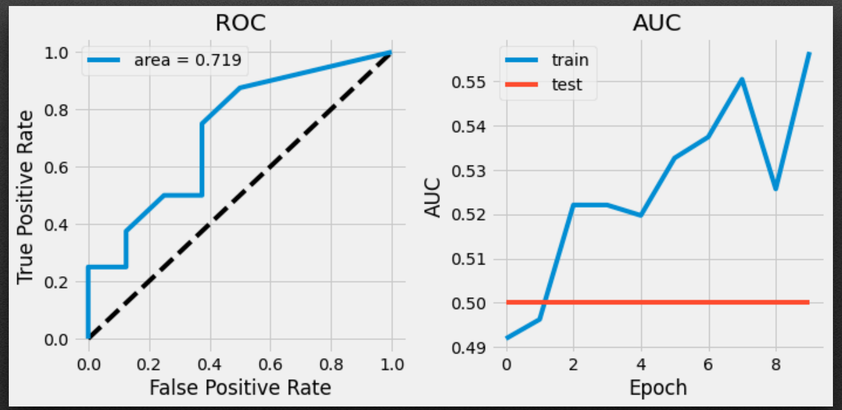
After Switching Datasets

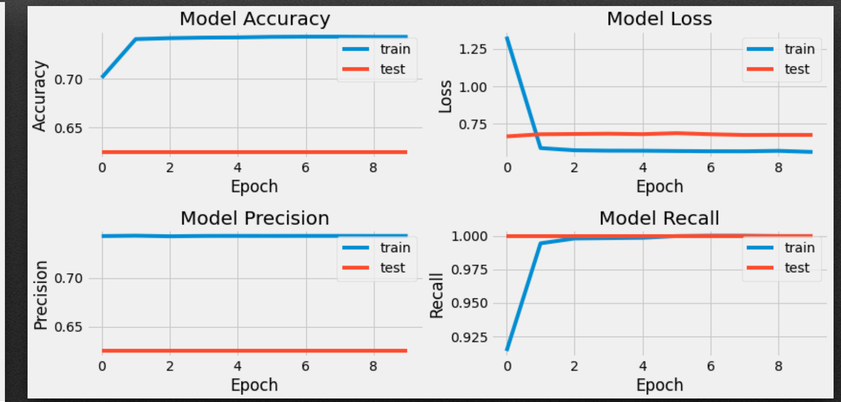
Initial



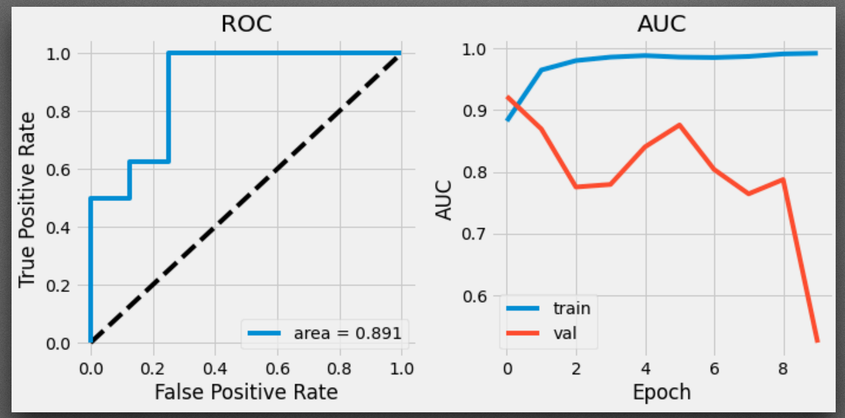


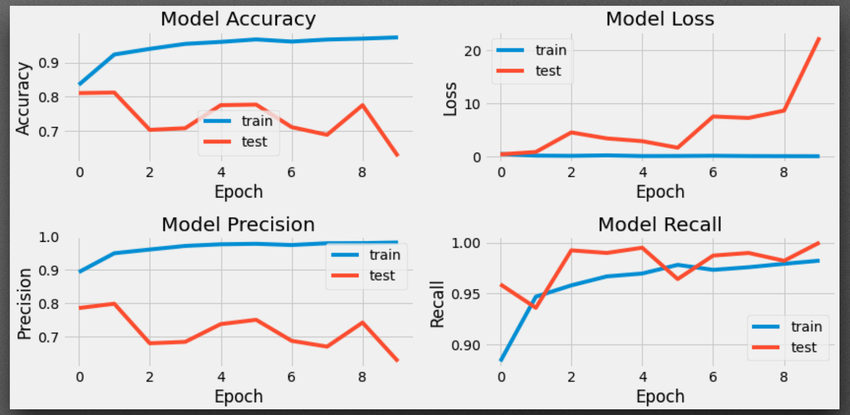
Removed Grayscale and Augmentation



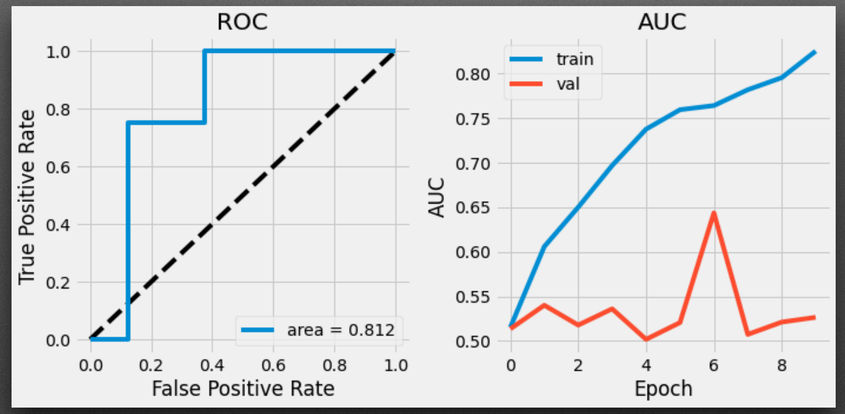


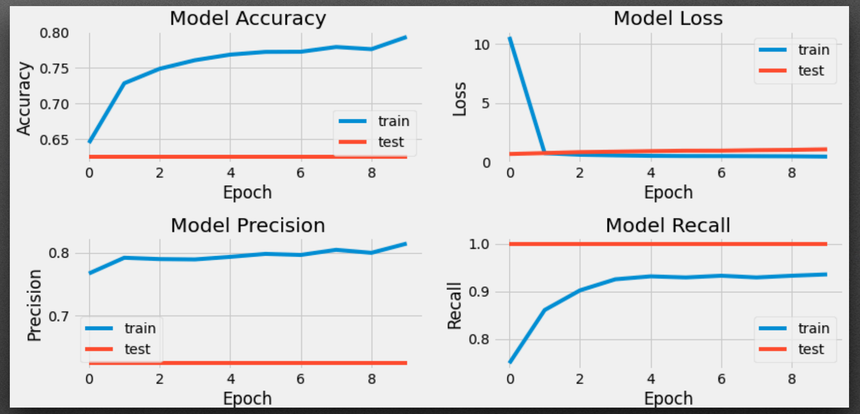
Unscaled Images /
Removed Dropout



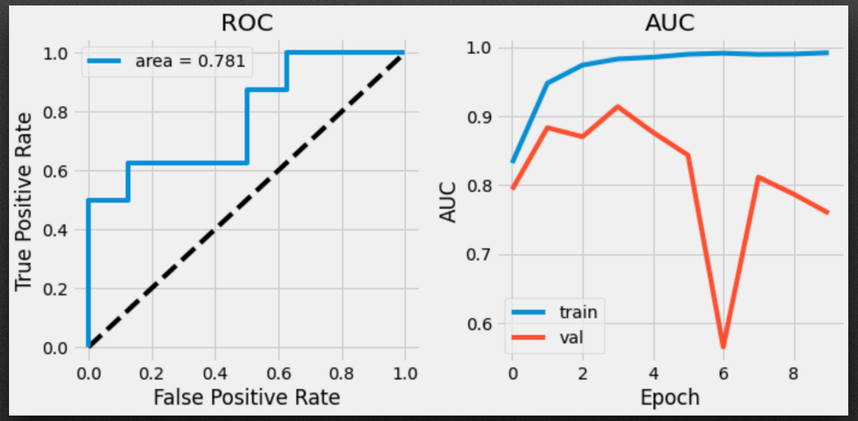


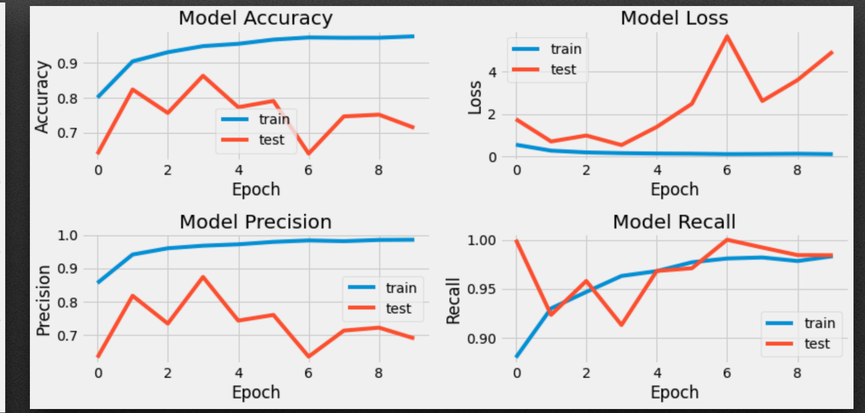
ReLU Activations /
Added Dropout



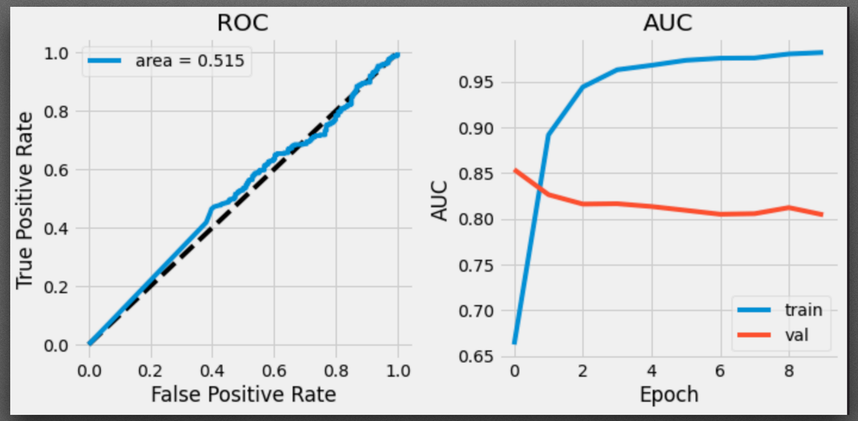


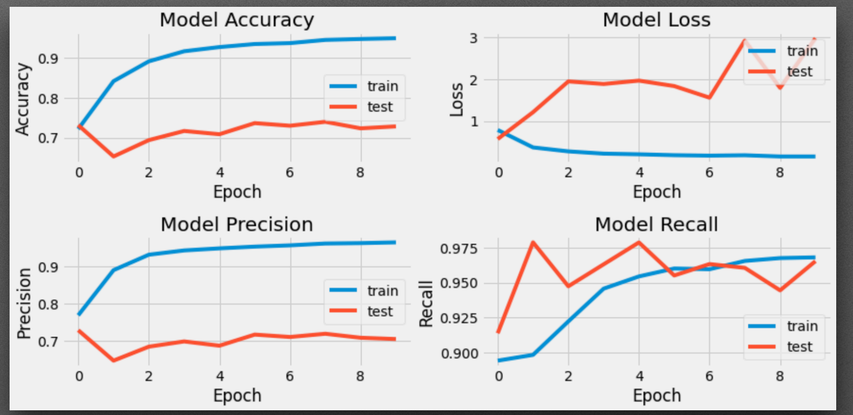
Weighted
Normal to 2



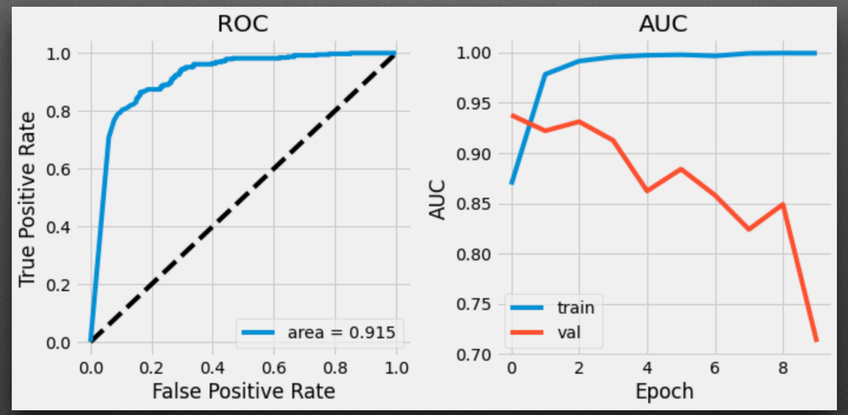


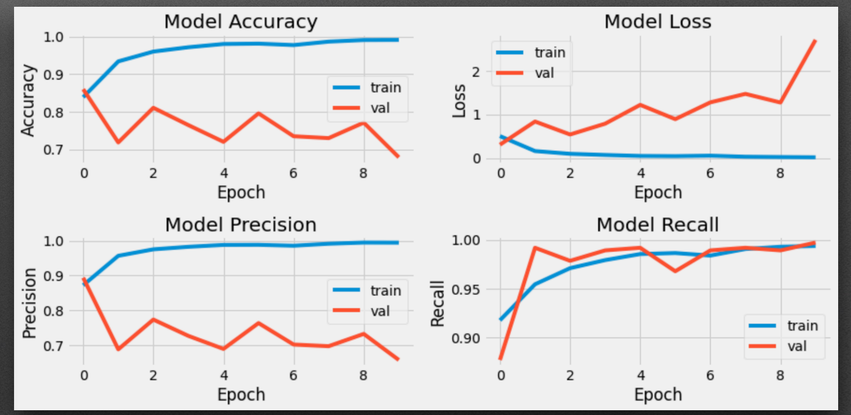
Changed Validation Images



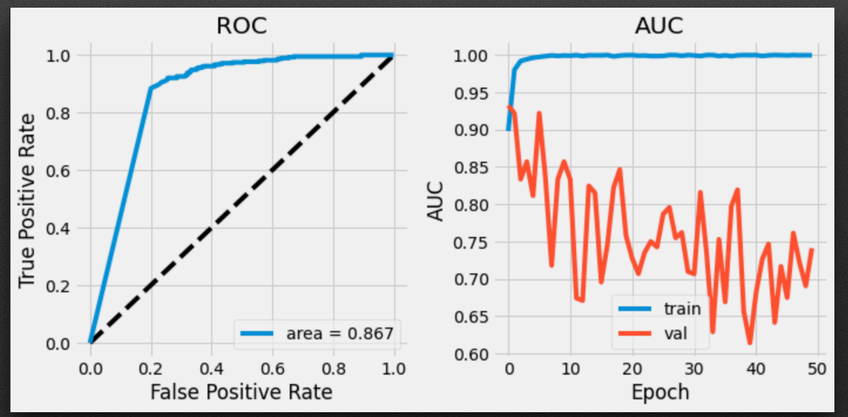


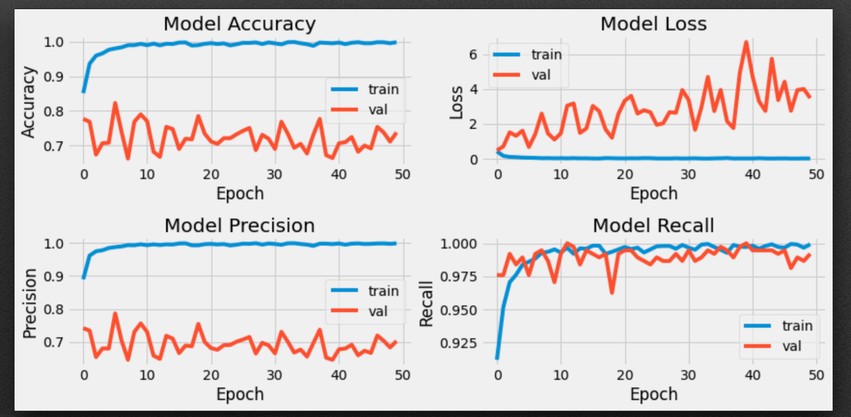
Removed Validation Shuffle



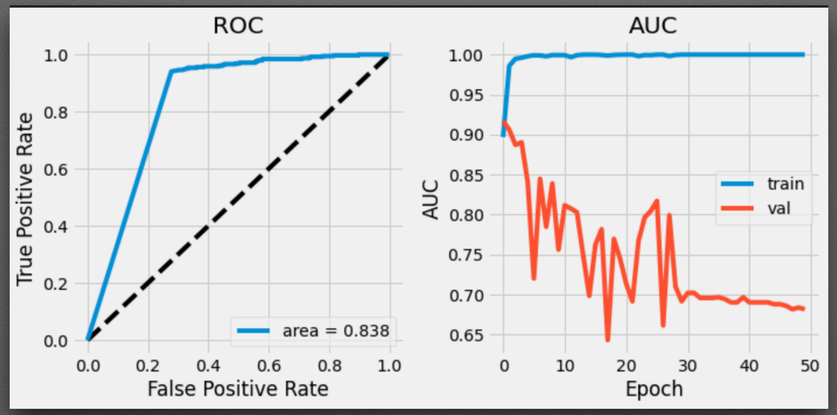


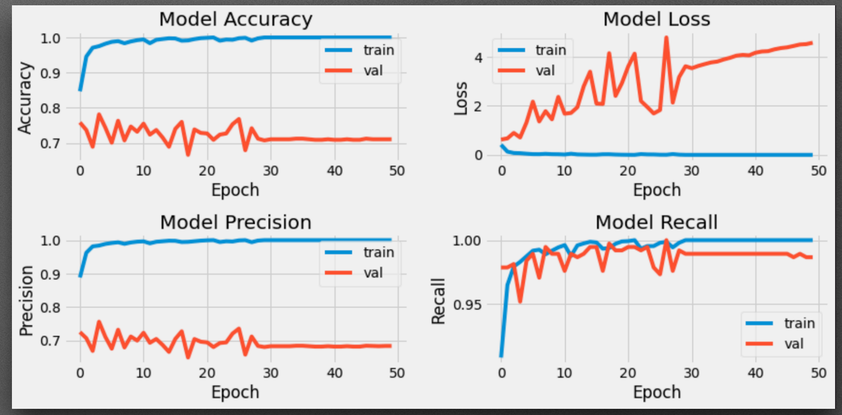
50 Epochs / 1 Dense Layer / Weighted Normal to 1.2



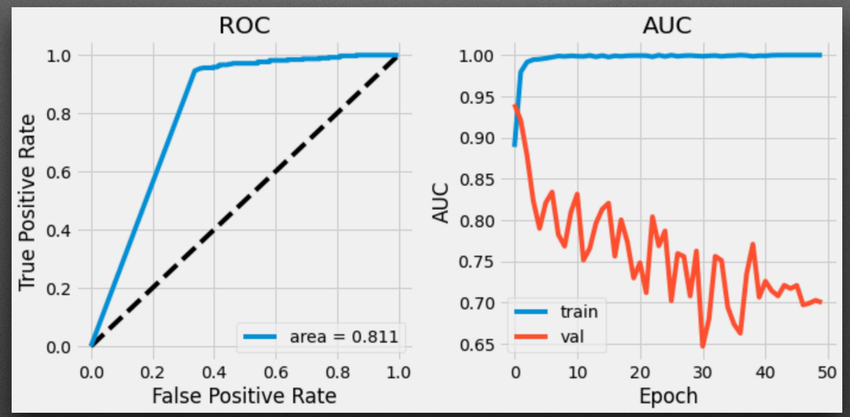


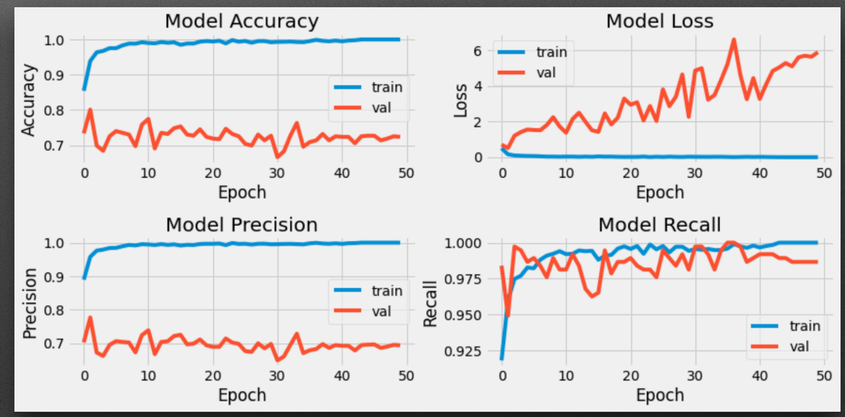
50 Epochs /
All Activations Swish



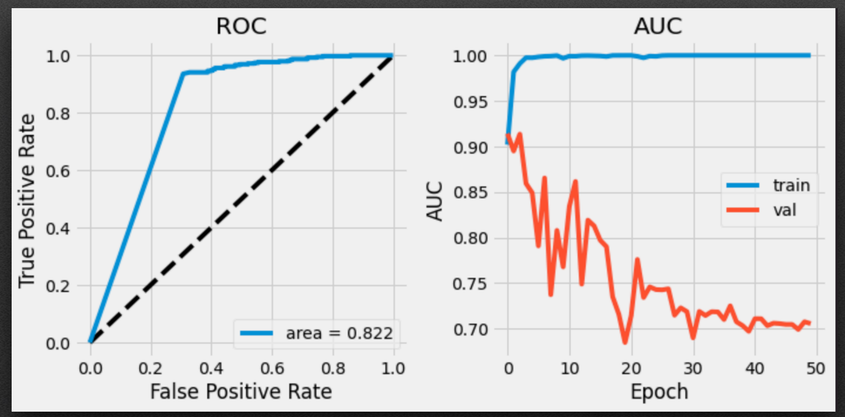


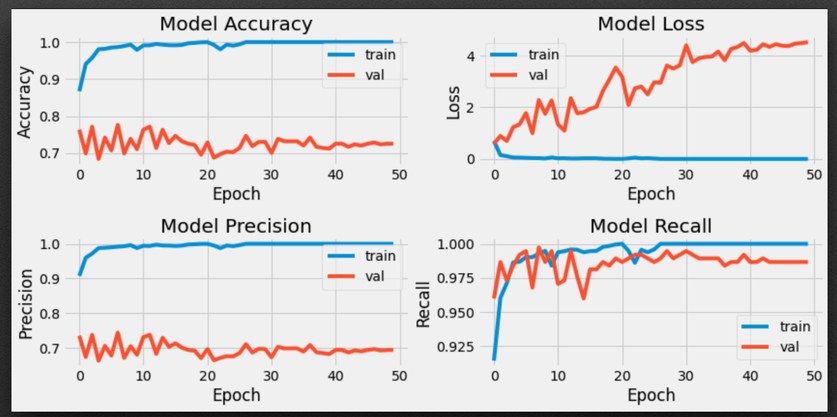
50 Epochs /
All Activations ReLU



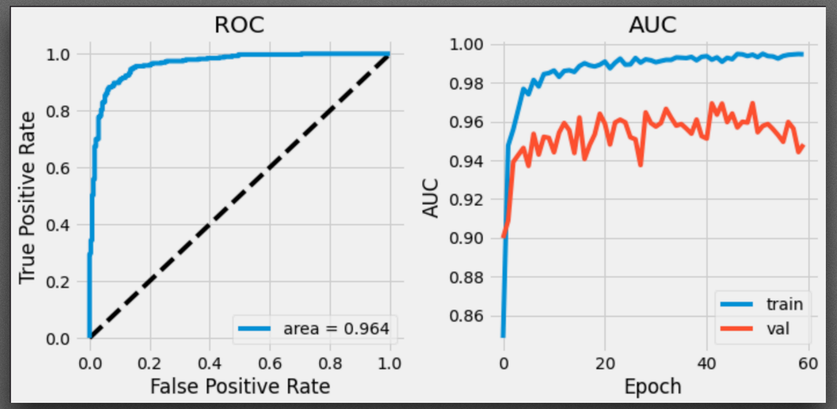


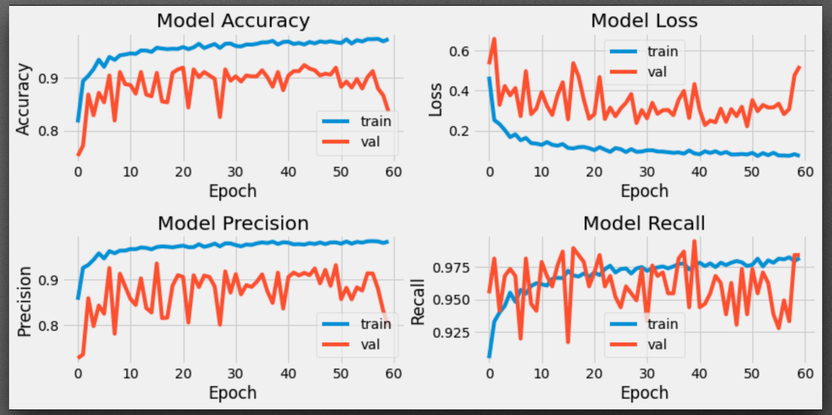
Removed last conv
layer. Dense units
512 to 256



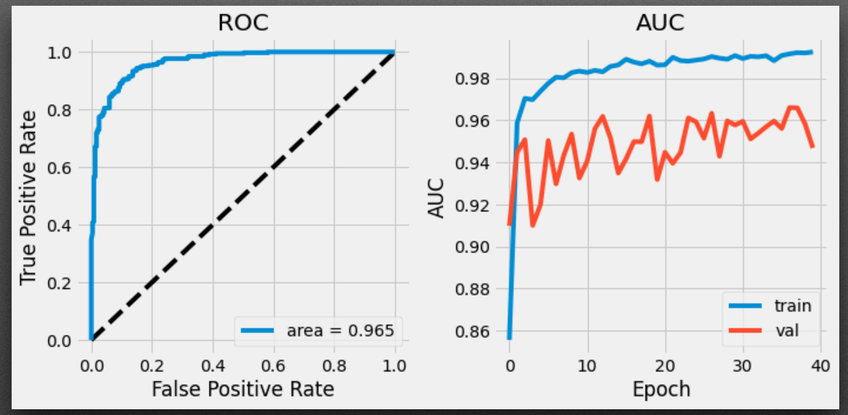


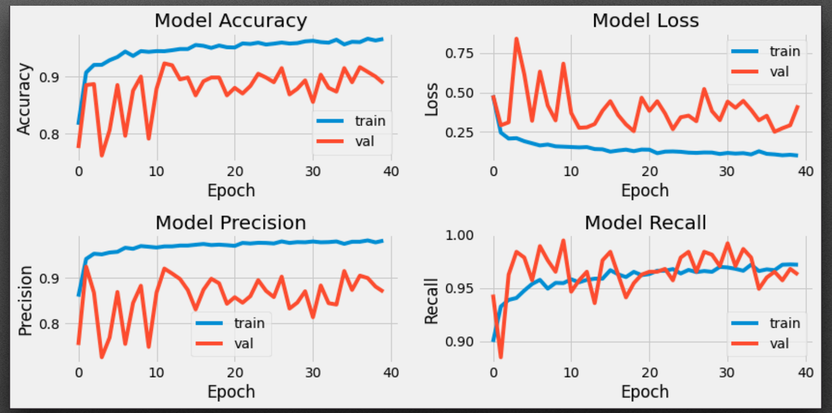
Added Augmentation / 60 Epochs / Swish Activation





Added Second Dense Layer





Added Second Dropout

Final Model

Final Model

```
Model: "sequential_18"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_90 (Conv2D)           (None, 298, 298, 16)      448
_________________________________________________________________
max_pooling2d_90 (MaxPooling (None, 149, 149, 16)      0
_________________________________________________________________
conv2d_91 (Conv2D)           (None, 147, 147, 32)      4640
_________________________________________________________________
max_pooling2d_91 (MaxPooling (None, 73, 73, 32)        0
_________________________________________________________________
conv2d_92 (Conv2D)           (None, 36, 36, 64)        18496
_________________________________________________________________
max_pooling2d_92 (MaxPooling (None, 18, 18, 64)        0
_________________________________________________________________
conv2d_93 (Conv2D)           (None, 8, 8, 128)         73856
_________________________________________________________________
max_pooling2d_93 (MaxPooling (None, 4, 4, 128)         0
_________________________________________________________________
conv2d_94 (Conv2D)           (None, 2, 2, 128)         147584
_________________________________________________________________
max_pooling2d_94 (MaxPooling (None, 1, 1, 128)         0
_________________________________________________________________
flatten_18 (Flatten)         (None, 128)               0
_________________________________________________________________
dense_42 (Dense)             (None, 512)               66048
_________________________________________________________________
dropout_34 (Dropout)         (None, 512)               0
_________________________________________________________________
dense_43 (Dense)             (None, 1)                 513
=================================================================
Total params: 311,585
Trainable params: 311,585
Non-trainable params: 0
```

	Train	Validation
Loss	0.1440	0.1440
Accuracy	0.9511	0.9511
Precision	0.9741	0.9741
Recall	0.9497	0.9497
AUC	0.9849	0.9849
True Pos	3742	3742
True Neg	1237	1237
False Pos	104	104
False Neg	133	133

Predictions

Key Takeaways

Class Imbalance

True Positives | True Negatives

ROC Curves | Area Under the Curve

Model Complexity

Overfitting / Image Augmentation

Cloud Computing

Next Steps

Apply this model to X-rays from the local hospital

Go back and try COVID again