Data Augmentation for Covid-19 Classification

Allen Chang
Dornsife College of Letters Arts and Sciences
University of Southern California
Los Angeles, California
achang30@usc.edu

Abstract

In December 2019, a series of acute atypical respiratory disease, which would later be known as Covid-19, was observed among a small group of people in China. However, the disease spread rapidly, infecting many people in a short period of time [1]. At the time, there was little known about the virus and no known treatment and vaccines. At present, more research has been conducted, and there is more abundant data for disease analysis. Here, I create a robust deep learning model architecture to classify whether a patient has coronavirus based on chest X-ray scans. I also experiment with scaling and blurring the training data and observe that both augmentation methods improve certain performance metrics. Lastly, I observe that differences in the air content of lungs is the primary marker for Covid-19 diagnosis through chest X-rays.

Introduction

The novel SARS-CoV-2 virus is thought to have originated from a seafood market in Wuhan, China. The disease caused by this virus is referred to as Coronavirus disease 19 (Covid-19). The disease spread rapidly and was later declared a pandemic by the World Health Organization. Covid-19 mainly affects the respiratory system with a wide variety of related symptoms. Symptoms can be severe, such as hypoxia and acute respiratory distress syndrome [1].

According to the WHO, as of July 2021, there have been approximately 190,833,853 confirmed Covid-19 cases and 4,100,087 confirmed Covid-19 related deaths [2]. With Covid-19 being a widespread pandemic, having a quick and accurate method to determine whether a patient is infected is crucial. Typically, deep learning models benefit from increased training data. Thus, I will experiment with augmentation methods including scaling and blurring on chest X-ray images to increase training sample size and examine whether they result in better model performance. Furthermore, I will characterize differences between the X-ray images of healthy and infected individuals.

Materials and Methods

My cohort includes 4500 individuals in total, sampled from the original dataset (here) to accommodate memory and runtime constraints [3], [4]. **Table 1** shows specific sampling metrics. A balanced test set was used for an accurate area under the receiver operator curve (AUROC) assessment.

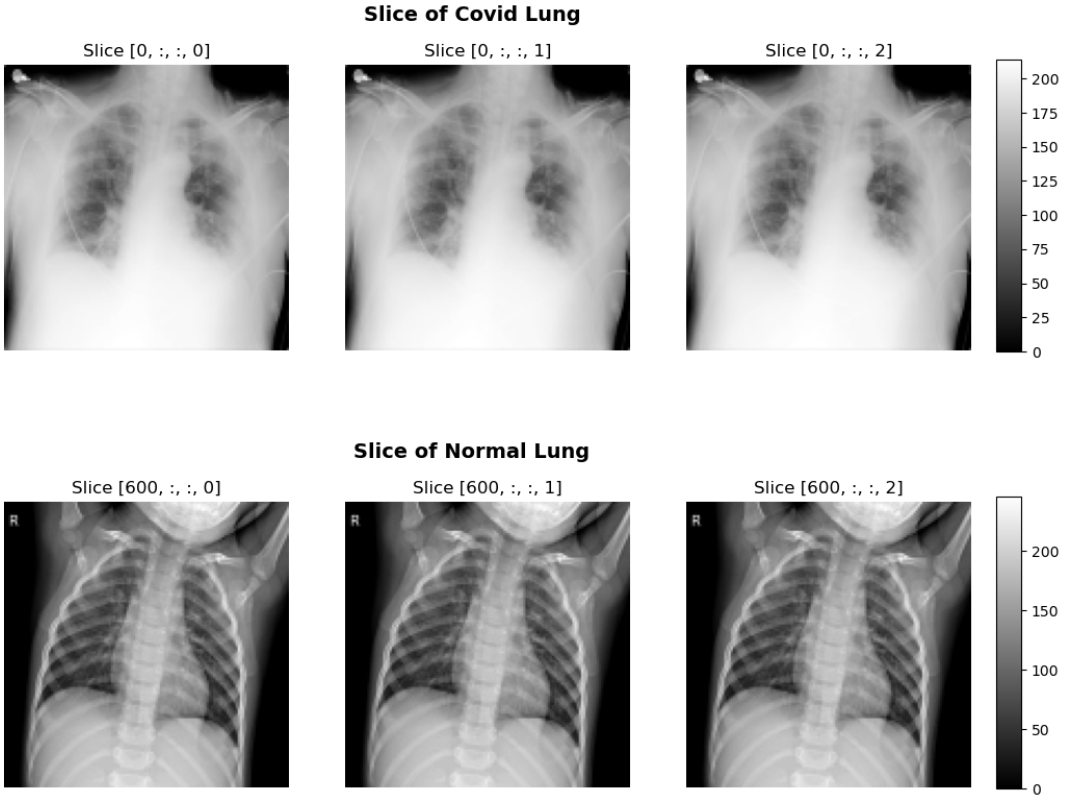
	Training	Validation	Test
Covid-19 Positive	500	500	500
Covid-19 Negative	1500	1000	500

Table 1 *Training, validation, and testing splits are shown.*

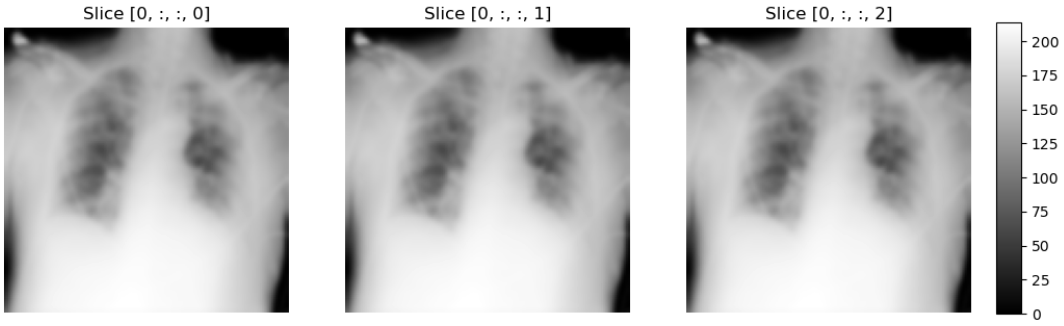
My model architecture comprises of an initial input layer followed by 4 2D convolutional blocks. Each convolutional block consists of a convolutional layer with increasing units (8, 16, 32, and 64) using a 3x3 kernel and rectified linear unit (ReLU) activation. Max pooling (2x2) and batch normalization enhance feature extraction. Dropout layers with rates of 0.4, 0.4, and 0.5, respectively, are used on the last three convolutional blocks. These blocks are followed by a flattening layer followed by another block consisting of a 128-unit dense layer with ReLU activation and a dropout layer with a 0.5 dropout rate. Finally, results are output through a single unit dense layer using sigmoid activation. The Adam optimizer minimizes the binary cross-entropy loss function over 100 epochs with a batch size of 16. The epoch with peak validation accuracy is saved.

My control model was trained on duplicated unaltered training data. My second model was trained on unaltered data concatenated with Gaussian blurred data using standard deviation of 1.05 and a kernel size of 5x5. My third model was trained on unaltered data concatenated with training data scaled by +5%. An example of augmented lungs is shown in **Figure 1**. AUROC, f1-score, precision, recall, and accuracy were reported for each model.

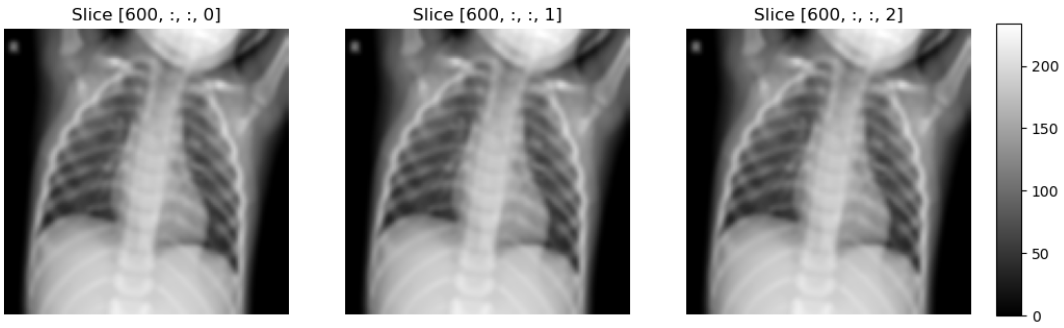
The test set was analyzed to visually identify differences between Covid-19 positive and healthy lungs. Two images (**Figure 5**) were generated for infected and healthy lungs by taking the average voxel intensities over all infected and healthy lungs in the test set.



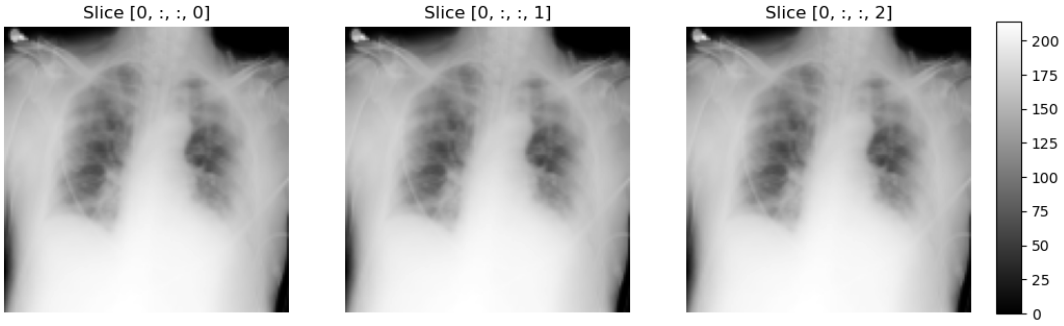
Slice of Blurred Covid Lung



Slice of Blurred Healthy Lung



Slice of Scaled Covid Lung



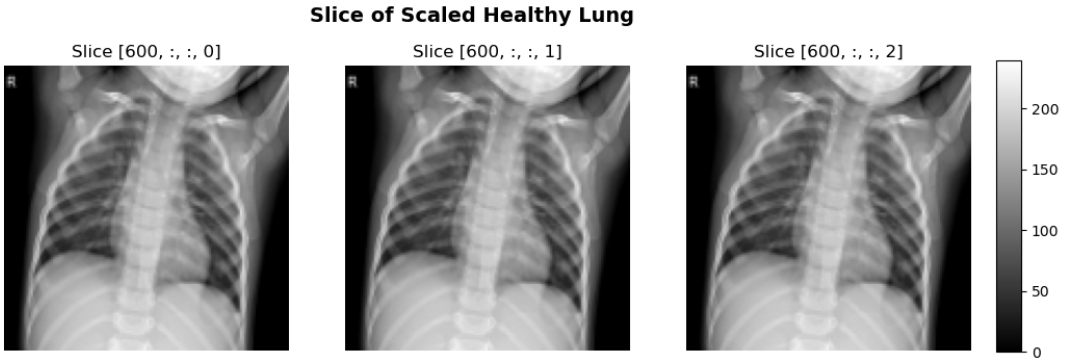
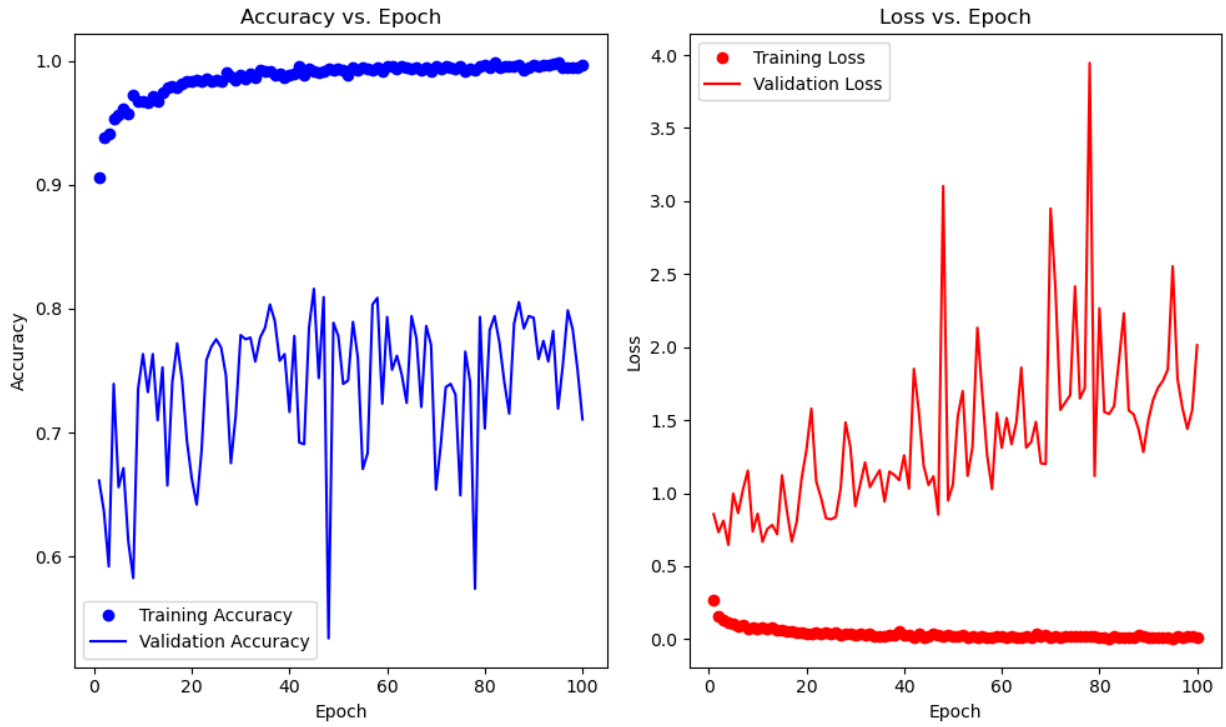


Figure 1 *Unaltered and augmented images of a healthy and Covid positive lung are shown.*

Results and Discission

In **Figure 2**, overfitting can be observed near the 30th epoch for all tested models where validation loss begins to increase, and validation accuracy ceases to increase. Nevertheless, the fluctuations of validation accuracy and testing make it difficult to pinpoint the exact epoch where performance stops increasing.



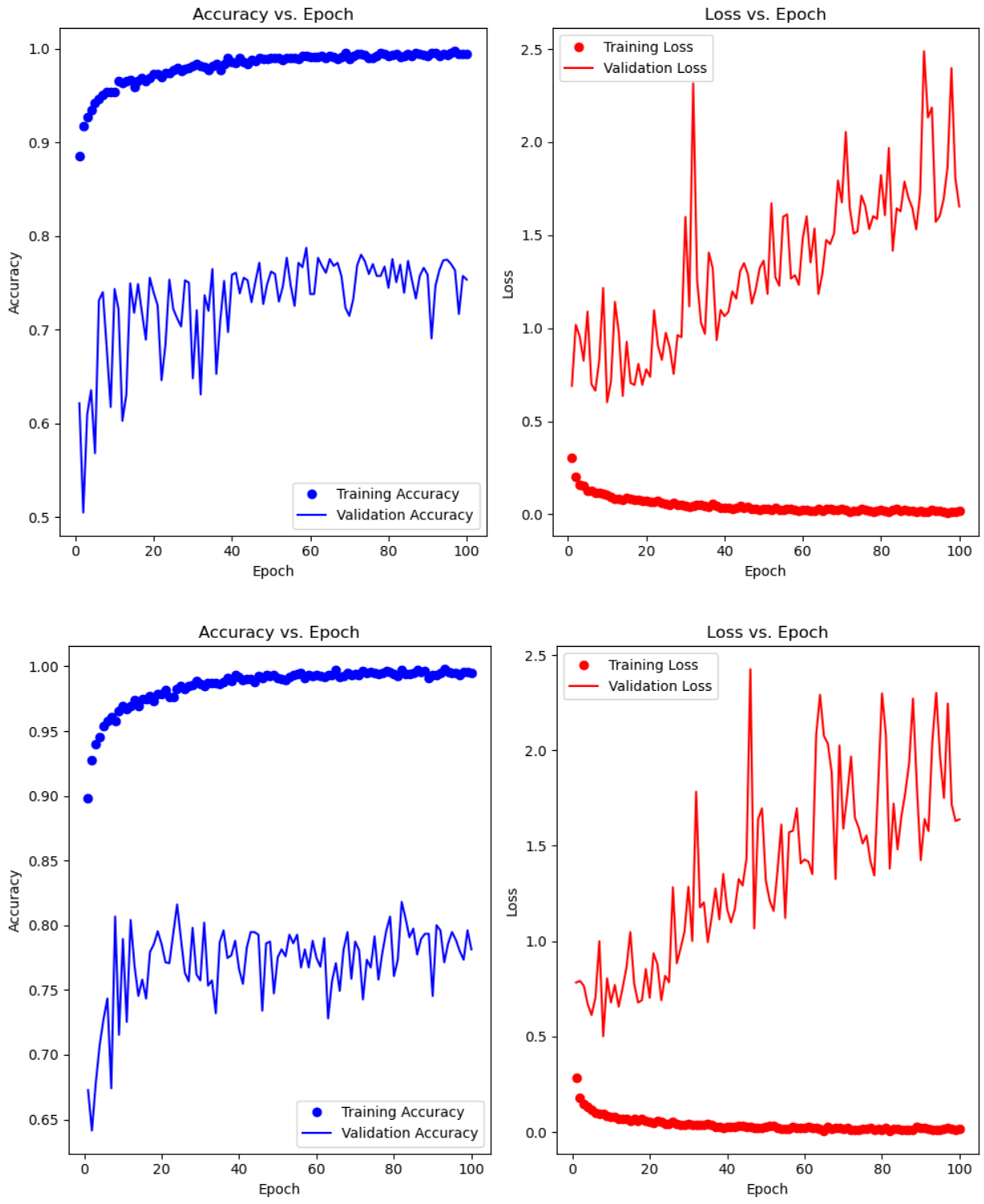


Figure 2 *Accuracy and validation loss for training and validation sets are shown.*

As seen in **Table 2**, the control model has the highest F1-score and accuracy. It has the second highest AUROC (**Figure 3**), precision, and recall. The model trained on blurred data has the lowest AUROC, F1-score, and precision. It has the highest recall and second highest

accuracy. The model trained on scaled data has the highest AUROC and precision. It has the second highest F1-score and lowest recall and accuracy. Interestingly, the model trained on scaled data also has a much lower recall and higher false positive rate as shown in **Table 2** and **Figure 4**.

	Control Model	Model Trained on Blurred Data	Model Trained on Scaled Data
AUROC	0.949	0.933	0.952
F1-score	0.894	0.879	0.881
Precision	0.952	0.896	0.974
Recall	0.842	0.863	0.805
Accuracy	0.887	0.877	0.869

Table 2 *Performance of the models is shown.*

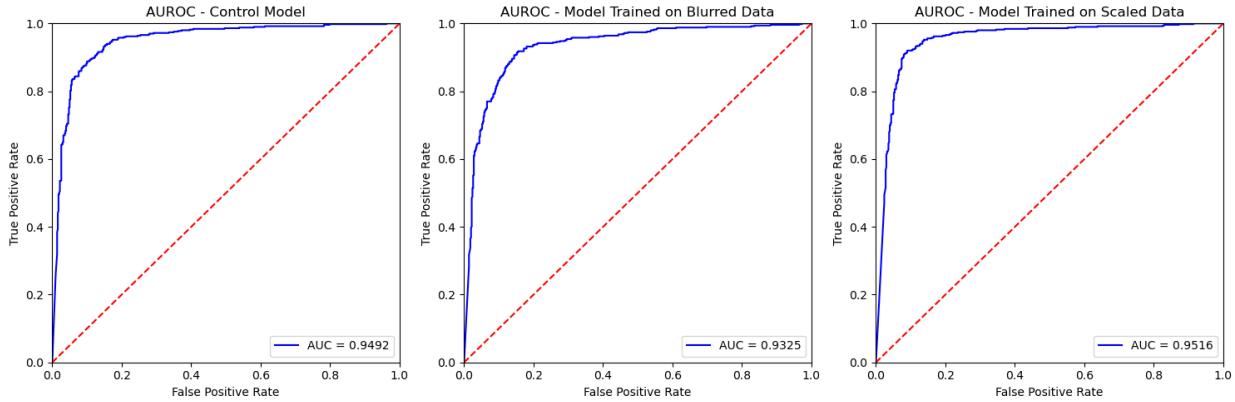


Figure 3 *AUROC curves for the models are shown.*

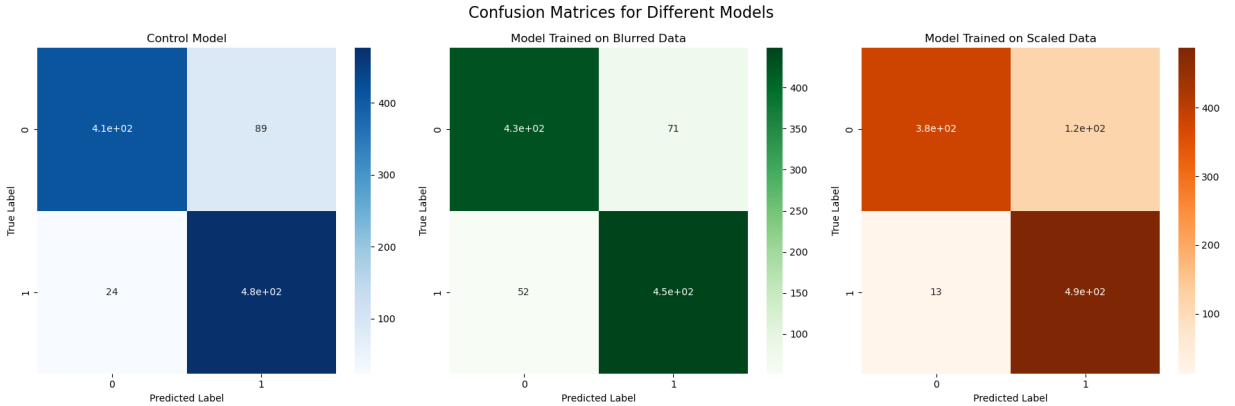


Figure 4 *Confusion matrices for the models are shown.*

I found that choosing an augmentation method depends on the specific uses of the model. For instance, if a false positive is more detrimental than a false negative report, using scaling as an augmentation method will be more valuable, since it has the highest precision. However, if a false negative is more detrimental, then using blurring as an augmentation method will be better,

since it has the highest recall. Nevertheless, all tested models have similar F1-scores and AUROC, indicating similar overall performance.

The most notable difference between the X-rays is where the lungs are located. This region appears to have lower voxel intensities with a noticeable smaller black area in **Figure 5**. Since the lungs are predominantly air, they should appear black on the X-ray. In fact, Covid-19 pneumonia is most noticeable in radiographs when there is a loss of a normal black appearance in the lung [5]. An increased whiteness, typically referred to as a "ground glass appearance," due to increased density is typically seen in Covid-19 infected lungs [5]. Nevertheless, typical markings, such as blood vessels, within the lungs are still noticeable in infected patients, which can be visualized in **Figure 1** [5].

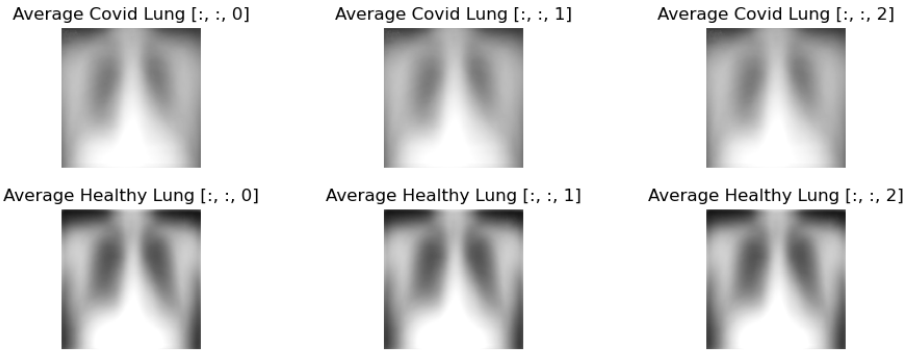


Figure 5 *Averages of healthy and infected lungs are shown.*

Conclusion

My results have shown that Gaussian blurring and scaling are viable augmentation methods, and the utility of each augmentation depends on the desired performance metric to be maximized. Nevertheless, more research needs to be conducted on a wider selection of augmentation parameters and model architectures. Furthermore, a limitation is that it is not feasible to have a test set that represents the true ever evolving distribution of Covid-19 positive individuals in a population.

References

[1] K. Yuki, M. Fujiogi, and S. Koutsogiannaki, "COVID-19 pathophysiology: A review," *Clin. Immunol.*, vol. 215, p. 108427, Jun. 2020, doi: 10.1016/j.clim.2020.108427.

[2] S. U. Rehman, S. U. Rehman, and H. H. Yoo, "COVID-19 challenges and its therapeutics," *Biomed. Pharmacother.*, vol. 142, p. 112015, Oct. 2021, doi: 10.1016/j.biopha.2021.112015.

[3] M. E. H. Chowdhury *et al.*, "Can AI Help in Screening Viral and COVID-19 Pneumonia?," *IEEE Access*, vol. 8, pp. 132665–132676, 2020, doi: 10.1109/ACCESS.2020.3010287.

[4] T. Rahman *et al.*, “Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images,” *Comput. Biol. Med.*, vol. 132, p. 104319, May 2021, doi: 10.1016/j.compbiomed.2021.104319.

[5] J. Cleverley, J. Piper, and M. M. Jones, “The role of chest radiography in confirming covid-19 pneumonia,” *BMJ*, p. m2426, Jul. 2020, doi: 10.1136/bmj.m2426.