# Augmented COVID-19 X-ray Images Dataset (Mendely) Analysis using Convolutional Neural Network and Transfer Learning

**Preprint** · April 2020

DOI: 10.13140/RG.2.2.20474.24003

CITATIONS
0

READS
264

**2 authors:**

A K M Bahalul Haque
North South University
**36** PUBLICATIONS **7** CITATIONS
SEE PROFILE

Mahbubur Rahman
North South University
**3** PUBLICATIONS **0** CITATIONS
SEE PROFILE

# Augmented COVID-19 X-ray Images Dataset (Mendely) Analysis using Convolutional Neural Network and Transfer Learning

AKM Bahalul Haque, Mahbubur Rahman

bahalul.haque@northsouth.edu, mahbubur.rahman1@northsouth.edu

## Introduction

Coronaviruses are a group of viruses that are spread among mammals and birds. This group causes common cold mild to severe in nature. SARS and MERS were some of its variants while the recent outbreak is of type COVID-19. This strain has its source in Wuhan, China and now has spread throughout the world. WHO has already declared COVID-19 as a pandemic due to its severity of spread and deaths. More than 200 countries have been affected by this. The death toll is increasing and the infection is being spread at an exponential rate. [1]

Some common symptoms of Novel Corona Virus COVID-19 are the common cold, high fever, dry cough, tiredness. The symptoms can be mild to severe. Till now there is no cure or vaccine against this virus. For this reason, the best way to tackle this virus is to prevent it. Lots of prevention mechanisms have been suggested by the World Health Organization(WHO) like frequently washing hands and not touching face, eyes, and nose with a hand. A minimum of 1-meter distance has to be maintained from the infected patients. COVID-19 infects the respiratory parts of the human body especially the lungs. [1]

## Dataset Description & Detection Methodology Used

Recently Mendeley has released Augmented COVID-19 X-ray Images Dataset. This Dataset Contains Augmented X-ray Images for COVID-19 for COVID-19 Disease Detection Using Chest X-Ray images. The dataset is collected from two online available datasets:

- https://github.com/ieee8023/covid-chestxray-dataset
- https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

The dataset contains two folders one for COVID-19 Augmented images while Non-COVID-19 is not augmented and the other folder contains augmented images for both COVID-19 and Non-COVID-19. [2]

We have uses Convolutional Neural Network [3], Inception V3 (Transfer Learning) [4] [5] to find the accuracy of detecting the COVID-19. The dataset is being updated as more and more cases added to the repository. Though this study is not from the medical perspective rather a small effort to build a COVID-19 disease detection system using the dataset available. We understand that the dataset is for experimental and educational

purposes as written in the GitHub repository, but a little effort in detection is all about we tried.

## Result Analysis

The dataset contains 69 images of lung x-ray of infected people with covid19 and 79 lung x-ray images of a healthy person. This is very little data to train a machine learning model. Three different models were used for covid-19 detection using images. The models are listed below.

- Convolutional Neural Network without Augmentation
- Convolutional Neural Network with Augmentation
- Inception V3 using Transfer Learning

### Convolutional Neural Network without Augmentation

This model consists of 2 convolution layers and 2 dense layer.

| Layer Name | Number of Filter | Window Size | Activation |
|---|---|---|---|
| Convolution-64 | 64 | 3X3 | Relu |
| MaxPooling-1 | N/A | 2X2 | N/A |
| Convolution-32 | 32 | 3X3 | Relu |
| MaxPooling-2 | N/A | 2X2 | N/A |

| Layer Name | Neurons | Dropout | Activation |
|---|---|---|---|
| Dense-64 | 64 | 0 | Relu |
| Dense-1 | 1 | N/A | Sigmoid |

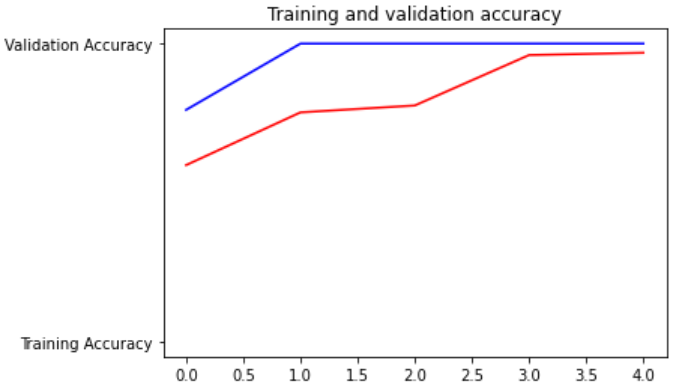


Fig 1. Accuracy of Training vs Validation

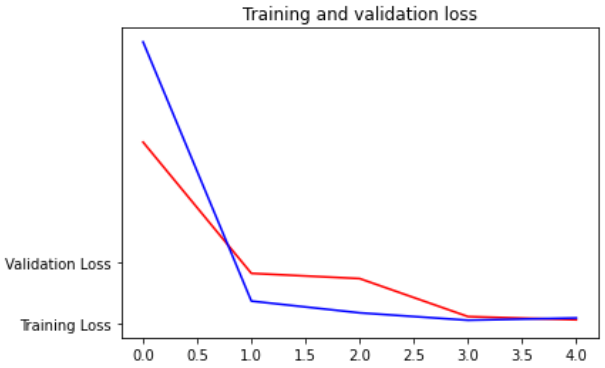


Fig 2. Loss of Training vs Validation

**Convolutional Neural Network with Augmentation**

This model is similar to previous model. But it contains one more convolution layer.

| Layer Name | Number of Filter | Window Size | Activation |
|---|---|---|---|
| Convolution-64 | 64 | 3X3 | Relu |
| MaxPooling-1 | N/A | 2X2 | N/A |
| Convolution-32 | 32 | 3X3 | Relu |
| MaxPooling-2 | N/A | 2X2 | N/A |
| Convolution-32 | 32 | 3X3 | Relu |
| MaxPooling-3 | N/A | 2X2 | N/A |

| Layer Name | Neurons | Dropout | Activation |
|---|---|---|---|
| Dense-512 | 512 | 0 | Relu |
| Dense-1 | 1 | N/A | Sigmoid |

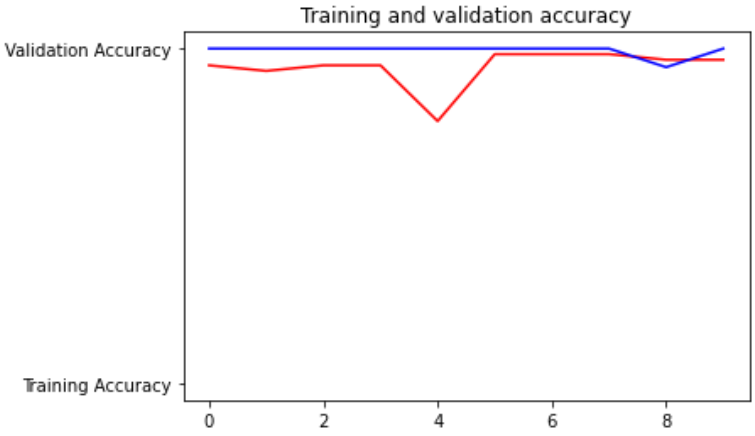


Fig 3. Accuracy of Training vs Validation

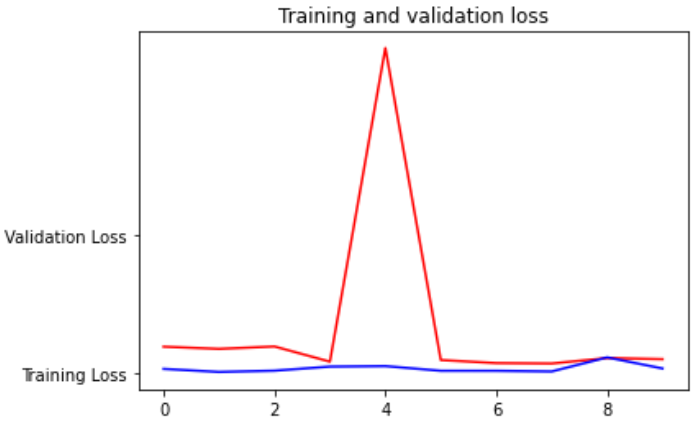


Fig 4. Loss of Training vs Validation

**Inception V3 using Transfer Learning**

The inception model was used up to the mixed7 layer. Its output was used as input for training. Rest of the inception model was not used. The mixed7 is a convolution layer which output is shaped (none, 7, 7, 768) for input shape of 150X150. This was passed to a dense layer.

| Layer Name | Neurons | Dropout | Activation |
|---|---|---|---|
| Dense-1024 | 1024 | 0 | Relu |
| Dense-1 | 1 | N/A | Sigmoid |

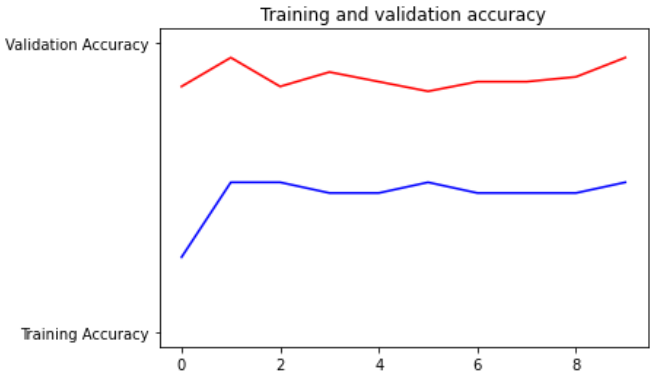


Fig 5. Accuracy of Training vs Validation


Fig 6. Loss of Training vs Validation

## Summary

From a quick glance, we can see that Inception model did worst in terms of predicting covid-19 vs healthy images. Both the CNN model with and without augmentation did have similar score around 97% of training accuracy and 99% of testing accuracy. Even though both of these models were similar we will take the augmented version as it had more training data to work with. However, this we cannot surely say this model that with larger dataset this model will have similar results because of the fact it only had about 130 images to work with. This amount of data is not significant enough for real world use cases.

| Model Name | Epochs | Training Accuracy | Training Loss | Validation Accuracy | Validation Loss |
|---|---|---|---|---|---|
| Convolutional Neural Network without Augmentation | 4 | 0.9752 | 0.06 | 0.8889 | 0.1071 |

| | | | | | |
|---|---|---|---|---|---|
| Convolutional Neural Network without Augmentation | 5 | 1.00 | 0.0108 | 0.8889 | 0.3235 |
| Convolutional Neural Network with Augmentation | 9 | 0.9667 | 0.1079 | 0.9444 | 0.1129 |
| Convolutional Neural Network with Augmentation | 10 | 0.9667 | 0.100 | 1.000 | 0.0323 |
| Transfer Learning Using Inception V3 | 9 | 0.8833 | 0.7511 | 0.4815 | 0.6939 |
| Transfer Learning Using Inception V3 | 10 | 0.9500 | 0.1517 | 0.6938 | 0.5185 |

Table 1. Model Comparison of covid-19 vs chest x-rays

# Future Work

Further improvement of the dataset used can lead to an effective result. Moreover, the worldwide pandemic has reduced the scope of work and research due to inaccessibility of resources. As more data to work on was available in with the augmented version, it showed better accuracy. As said before, more data will lead to more accuracy. Hopefully this effort will be helpful in shedding some lights for the researchers.

# References

[1] https://www.who.int/news-room/q-a-detail/q-a-coronaviruses

[2] Alqudah, Ali Mohammad; Qazan, Shoroq (2020), “Augmented COVID-19 X-ray Images Dataset”, Mendeley Data, v4.http://dx.doi.org/10.17632/2fxz4px6d8.4

[3] Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., & Khan, M. K. (2018). Medical image analysis using convolutional neural networks: a review. *Journal of medical systems*, *42*(11), 226.

[4] Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, *22*(10), 1345-1359.

[5] Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, *22*(10), 1345-1359.