AI against COVID-19: Screening X-ray images for COVID-19

Part II - A lighter model for Covid-19 Detection

Jack (Jianxiang) Xu
University of Waterloo
projectbyjx@gmail.com

Michael (Hongjian) Zhou
University of Waterloo
michael.chou1998@outlook.com

Abstract

The main focus in this project would be further investigating and optimizing our hypothesis on pre-augmentation strategy to produce an improved model. In the prior round, we utilized three channels of the Resnet34 with basic morphological operators, which improved the model performance significantly from 80% to 99%. In this project, we will try to further evaluate the effect of different pre-augmentation on covid-19 detection particularly.

1 Introduction

Since 2016 when ResNet[1] was firstly introduced to ease deep learning for image recognition, it quickly gained dramatic popularity and explicit successful evidence in different real life application[2]. Unsurprisingly, ResNet50 was also introduced to help solve the challenges with COVID-19 detection from chest X-ray images [3].

2 Methodology

In this challenge, we were thinking in a different direction: if we can pre-augment the dataset such that it would be much easier for a selected model to learn the detection. In the first round, we realized that the chest scan is a grayscale image, which does not fully utilize the full capability of the ResNet model, and it takes quite a bit time to train and incapable to achieve extremely high score.

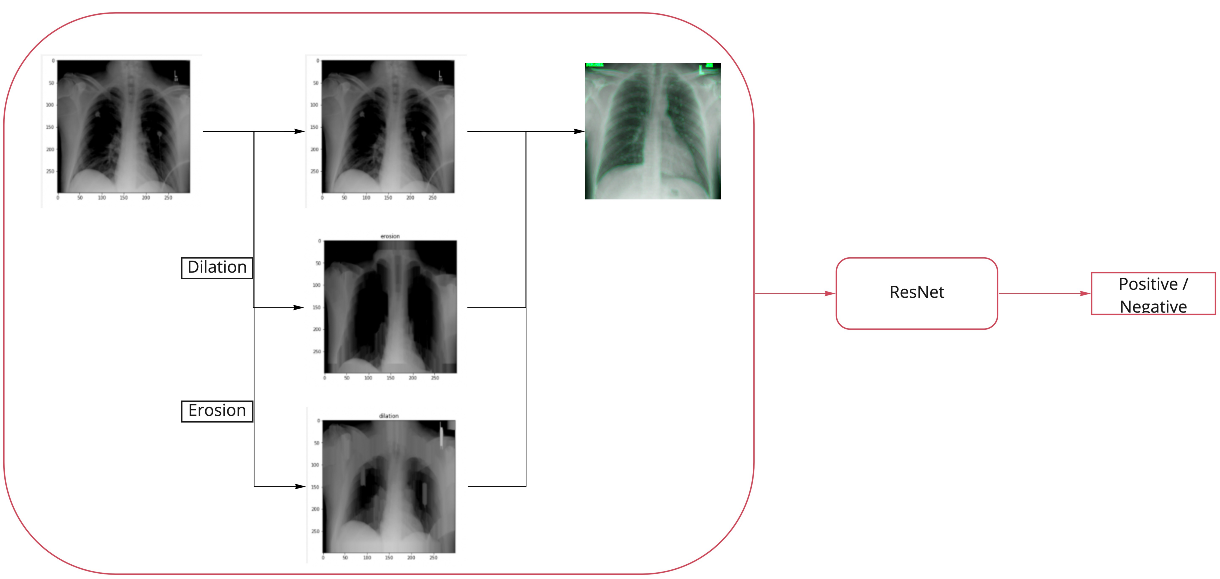


Figure 2-1: Our Proposed Model

Submitted to IEEE SIGHT Montreal 2021- AI against COVID-19 (Part II). Do not distribute.

So, what if we replace two channels with some basic homography operation: one to attenuate noises and enhance the chest contour, and the other to amplify small features. Hence, we proposed a new augmentation method to simply pre-process the given chest images with a dilated and eroded images stabbed in two different channels as shown in Figure 2-1 above. The performance of the resultant model was proved to be promising explicitly.

2.1 Dataset

The model is trained based on the open source dataset from [4]. The dataset is initially pre-processed with channel augmentation to improve the feature difference cross channels as shown in Figure Appendix -4.

3 Results

3.1 ResNet Variance

In order to reduce the complexity of the model, the number of repeated layers were reduced in ResNet model to generate different models. As the complexity reduces, we may find some trending results as discussed in the following sections. As shown in Table 3-1 below, we may find the model tends to learn slower and less stable comparing to a more complex ones, as the early stopping can still capture later improvements in a longer epochs for a much simpler models.

Table 3-1: ResNet Model Reduced Variance

Model	**ResNet 34**	**ResNet 18**	**ResNet 10**	**ResNet 10 (Ablation)**
Layer 1	3	1	1	0
Layer 2	4	2	1	1
Layer 3	6	3	1	1
Layer 4	3	2	1	1
Total Parameters	21.3**M**	12.3**M**	4.9**M**	4.9**M**
Epochs	58	96	193	195
Training Accuracy	99.13%	99.52%	99.32%	99.5%
Validation Accuracy	99%	99.25%	99.5%	99.5%

3.2 Evaluation Over External Dataset

In order to further evaluate the generalization performance of the pre-trained models, we evaluate the model over a completely different dataset from [5], with more categories in terms of non-covid scans as discussed in [6].

Table 3-2: ResNet Model Reduced Variance

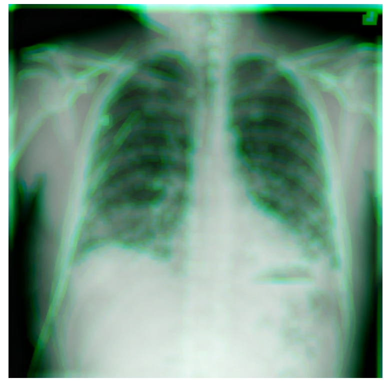
Model	**ResNet 34**	**ResNet 18**	**ResNet 10**	**ResNet 10 (Ablation)**
Epochs	58	96	193	195
Training Accuracy	99.13%	99.52%	99.32%	99.5%
Validation Accuracy	99%	99.25%	99.5%	99.5%
Covid (3616) Accuracy	47.70%	51.41%	53.13%	51.91%
Lung Opacity (6012) Accuracy	51.98%	45.54%	46.84%	48.44%
Normal (10192) Accuracy	54.89%	49.30%	52.01%	50.69%
Viral Pneumonia (1345) Accuracy	55.32%	51.60%	57.17%	47.96%

Interestingly, the performance is not great as shown in Table 3-2 above, since the input images are in 299×299 instead of a high resolution images over 1000×1000. The pre-processing kernels of dilation and erosion may over-exaggerate the features and noises, as shown in Figure 3-2 below. As a result, the performance is not as good. In the later investigation (as later discussed in Section 3.3), we will be exploring the options, such as two-stage resizing to ensure the same rate of dilation and erosion on a similar dpi of the image.

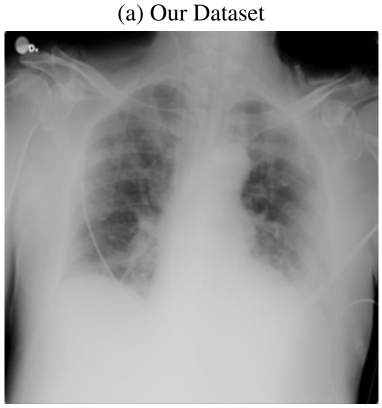
3.3 Reduced ResNet with the latest dataset

The models above were initially trained with a balanced dataset with augmentation on all 2158 positive and 2158 negative samples. The latest dataset contains more positive images scans, so that we are able to train the model with up

(a) Our Dataset



(b) Our Dataset (Augmented)

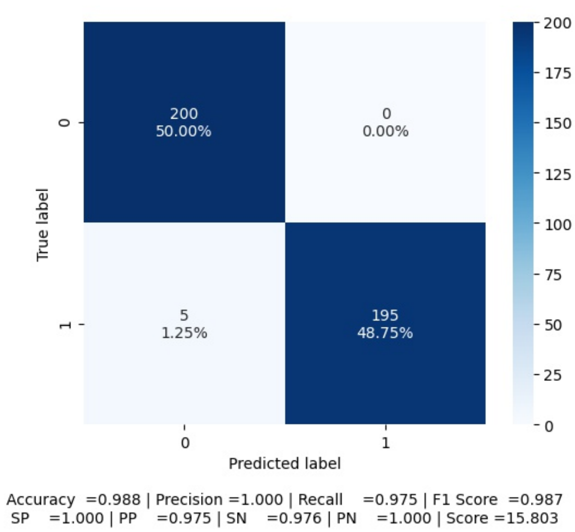


(c) External Dataset

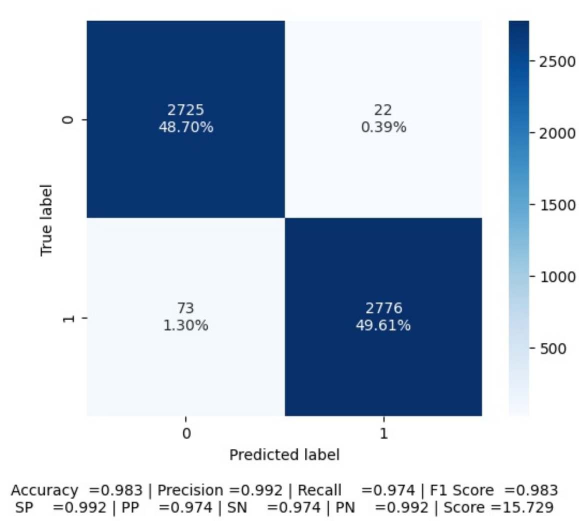
(d) External Dataset (Augmented)

Figure 3-2: Dataset Difference

to 13992 images for each positive and negative samples. The model is evaluate on the same 400 balanced dataset as before for early stopping as shown in Figure 3-3a. To note, in this training, we apply the basic homography operations post the image resizing, reducing the inconsistency due to the original image resolution.



(a) Validation at 78 epochs



(b) Evaluation

Figure 3-3: Confusion Matrices

To further investigate the performance of the model, the training dataset was split into 80% training and 20% evaluation. The result for the selected ResNet 10 was trained and evaluated as shown in Table 3-3 below, with confusion matrix in Figure 3-3b above.

Table 3-3: Reduced ResNet Model Trained On Larger Dataset

Model	**ResNet 10**
Parameters	4.9M
Input Size	$(3 \times 320 \times 320)$
Epochs	78
Training Accuracy	98.63%
80% Validation Accuracy	98.75%
Covid (3616) Accuracy	47.87%
Lung Opacity (6012) Accuracy	51.65%
Normal (10192) Accuracy	54.36%
Viral Pneumonia (1345) Accuracy	49.67%
20% Evaluation Accuracy	98.30%

It appears the model is not performing well on the external data sources even with the post homography operators. However, the model works well with the dataset provided.

4 Discussion

Throughout our investigation, we found our pre-processing techniques work great with even reduced ResNet models, which maintains the accuracy with four times smaller model comparing to the ResNet 34 initially implemented in the first round. However, the pre-trained model did not perform well on the external dataset found from Kaggle, and further investigations are needed. In short, we indeed find a significant performance boost by simple erosion and dilation homography operations prior to the grayscale dataset, to fully utilize the capability of ResNet with cross-channel correlation.

References

[1] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[2] F. He, T. Liu, and D. Tao, “Why resnet works? residuals generalize,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 12, pp. 5349–5362, 2020.

[3] S. Karakanis and G. Leontidis, “Lightweight deep learning models for detecting covid-19 from chest x-ray images,” *Computers in Biology and Medicine*, vol. 130, p. 104181, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482520305126

[4] “Covidx cxr-2 dataset.” [Online]. Available: https://www.kaggle.com/andyczhao/covidx-cxr2

[5] “Covid-19 radiography database.” [Online]. Available: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database

[6] M. E. H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. A. Emadi, M. B. I. Reaz, and M. T. Islam, “Can ai help in screening viral and covid-19 pneumonia?” *IEEE Access*, vol. 8, pp. 132 665–132 676, 2020.

Glossary

Resnet34 Resnet34 is a 34 layer convolutional neural network

Acronyms

CNN Convolutional Neural Network

ResNet Residual Neural Network

MLP Multi-Layer Perceptron

74

Appendix

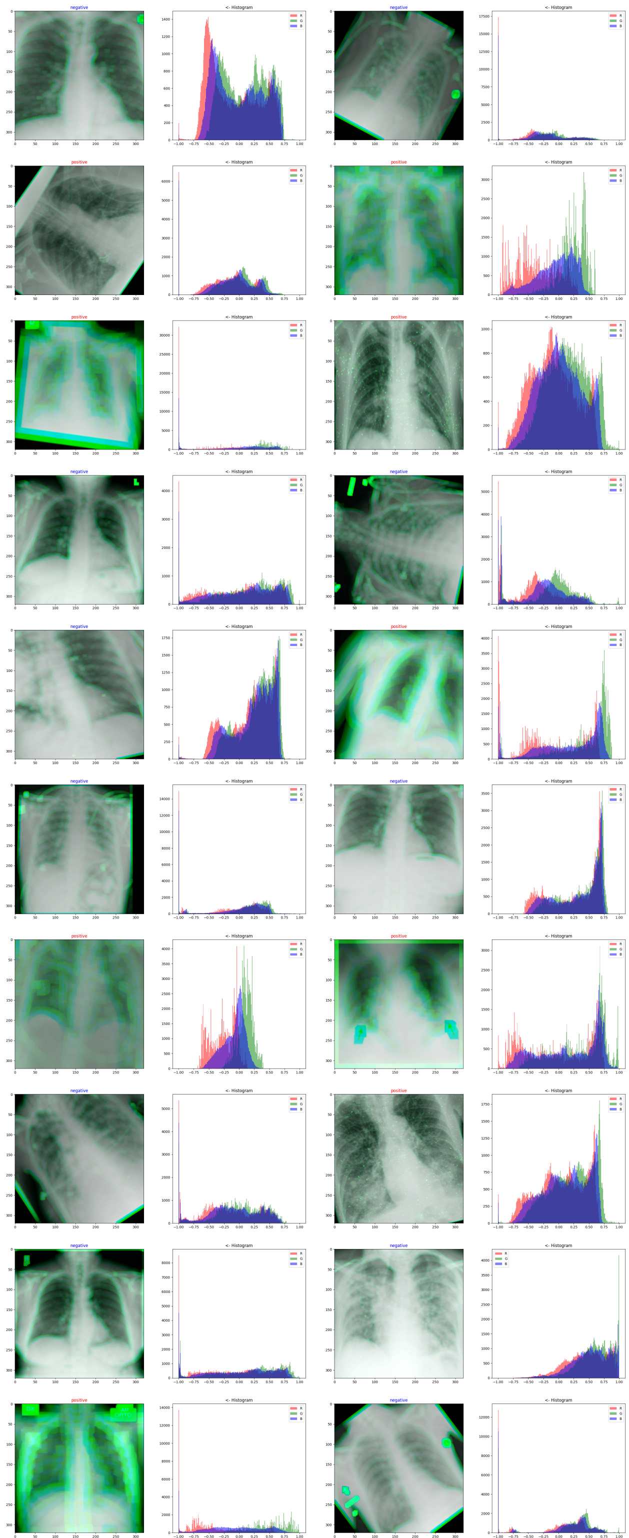


Figure Appendix -4: Pre-processed Training Samples