COVID-19 and Pneumonia Classification with X-ray Images Using Combined ResNet50 Model

Daekyung Kim[1], Youngho Kim[2], Jey Kang[3], Byungkon Kang[4], and Deepak Pradhan[5]

I. INTRODUCTION

As of December 5^{th}, 2020, 66,418,437 COVID-19 cases are confirmed as the total number in the world. 1,530,642 people have died, and 46,065,730 have recovered after being infected by the virus. The number of active coronavirus patients is still 19,001,886, and 106,102 have progressed to be at a severe state[12].

The coronavirus disease (COVID-19) originated from Wuhan, China in December 2019. Previous cases of coronavirus outbreaks include Middle East respiratory syndrome (MERS-CoV) and severe acute respiratory syndrome (SARS-CoV). Past examples of automated chest radiography (X-ray) diagnoses are not as easily accessible as current cases due to the comparatively little attention the outbreaks received, and the subsequent lack of interest in the radiography data [2].

Recently, many hospitals have begun to realize the necessity of automated COVID-19 detection due to the rapid growth in the number of COVID-19 patients. Radiography physicians and doctors were able to differentiate between pneumonia-infected and healthy lungs through patients' chest X-ray images. Now, however, doctors have to recognize the differences between COVID-19 and regular pneumonia from X-ray images and CT scans. The problem here is that the symptoms shown in the X-ray and CT scans of COVID-19 and pneumonia are similar enough that a regular doctor cannot easily differentiate the two. Specialized radiologists can diagnose COVID-19 more easily, but there are far fewer specialists than regular doctors [1].

There are two solutions to this problem. The first solution is providing a machine learning classification model to do the differentiation. With less work and involvement from the doctors, the process of accurately diagnosing the patient will be faster. In the cases where the machine learning model fails to provide high confidence in diagnosing, a possible secondary solution would be to use CT scans instead of X-ray images. Using CT scans may provide higher accuracy in classifying COVID-19 compared to X-ray images due to the higher resolution of the produced scans [3]. However, CT scans are more expensive data than X-ray as it takes long for a patient to take a CT scan, whereas X-rays images can be produced quickly. There are balanced trade-offs in these two solutions. In this paper, we will mainly focus on X-ray images. It is also important to re-emphasize the need for automated classification using machine learning as it is already a challenge for the doctors to classify the X-ray images.

II. BACKGROUND

A. ResNet50 Model

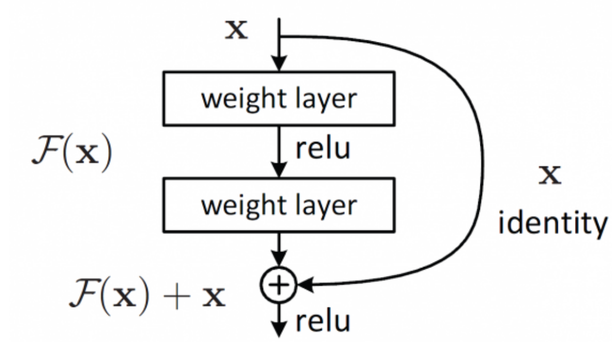

Fig. 1. Residual Network (ResNet)

As development on neural networks progressed, simply adding more and more layers to deeper neural networks seemed to be the solution to improving classification accuracy. However, it gradually became obvious that this introduced additional difficulties in training and lead to scenarios with substantially degraded output. It was around this period that the Residual Network architecture (ResNet) was introduced. The main draw of the ResNet architecture is that it utilizes residuals (the differences between the actual values and the predicted values) rather than features for learning. Additionally, it uses a technique known as shortcut connections—connections bridging the input of a layer to a different layer. There are multiple variations of the base ResNet implementation, of which ResNet50 was selected for use in this paper due to its high relevance to existing X-ray classification problems [10].

III. EXISTING MODEL

A. COVID-Net

COVID-Net is an existing model with the same objectives as our own. The standout features of COVID-Net are its diversity in layer usage and its usage of inter-layer bridges, much like the ResNet architecture. It makes heavy use of the projection-expansion-projection-extension (PEPX) design pattern, which is purported to enhance representational capacity while maintaining computational efficiency [8].

There are several flaws with COVID-Net's proposal. Firstly, the paper introducing COVID-Net does not list the F1 scores of the experiments. This is a critical oversight given that F1 scores better represent a model's capability to adapt to imbalanced dataset like the COVIDx dataset used for

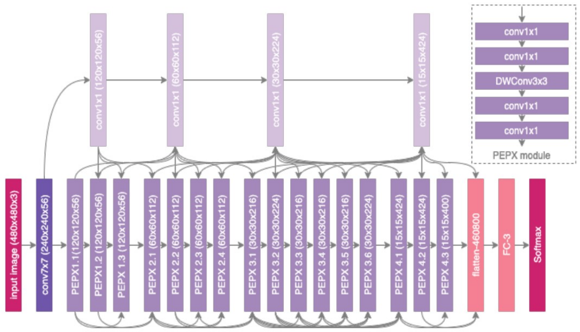

Fig. 2. Architecture of COVID-Net

COVID-Net, which has thousands of non-COVID (normal and pneumonia) X-ray scans but only a few hundred COVID-19 X-ray scans. Additionally, the model itself has many issues regarding both its performance and various errors encountered during evaluation, as can be attested to by the numerous GitHub users who have voiced their complaints on the official GitHub page for the authors' implementation of COVID-Net.

Positive Predictive Value (%)			
Architecture	**Normal**	**Non-COVID19**	**COVID-19**
VGG-19	83.1	75.0	98.4
ResNet-50	88.2	86.8	98.8
COVID-Net	**90.5**	**91.3**	**98.9**

Fig. 3. Positive Prediction

Sensitivity (%)			
Architecture	**Normal**	**Non-COVID19**	**COVID-19**
VGG-19	**98.0**	90.0	58.7
ResNet-50	97.0	92.0	83.0
COVID-Net	95.0	**94.0**	**91.0**

Fig. 4. Sensitivity

IV. DATASET

A. "X-ray Images for COVID-19 Classification"

The COVIDx dataset is composed of open-source images of X-ray scans gathered from five major open access dataset. It has three classes—COVID19, Pneumonia, and Normal. It is continuously updated as new images become available, as X-ray images accessible by the general public remain limited.

Dataset URL: COVID-X Dataset (Link)

B. "X-ray Images of COVID-19, Pneumonia, Normal Conditions"

In order to train our model, we had to filter out poor-quality images from the COVIDx datase, after which we reorganized and processed the data to fit our own format. The images were sorted into 3 directories, accessible via the link below.

Dataset URL: X-ray Images of COVID-19, Pneumonia, Normal Conditions (Link)

V. MODEL CHOICE AND EXPERIEMENTS

We chose Resnet50 as our model of choice as it produced a high performance at varying dataset. We have conducted a training with approximately 3,700 images, and still obtained a high accuracy as shown in Figure 5.

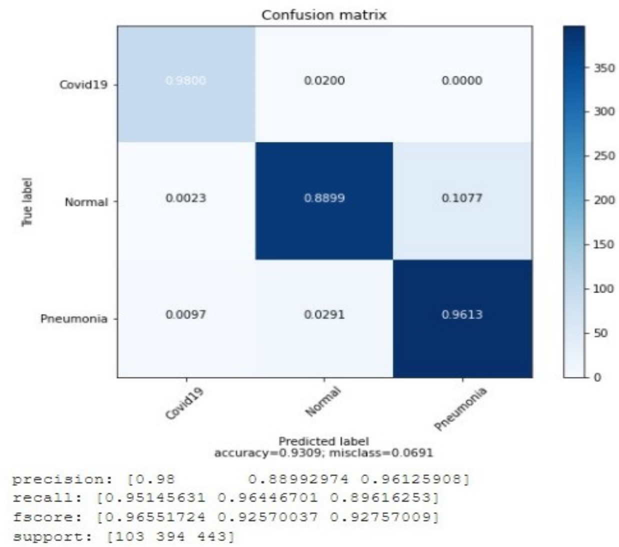

Fig. 5. ResNet50 with 3,736 X-ray Images

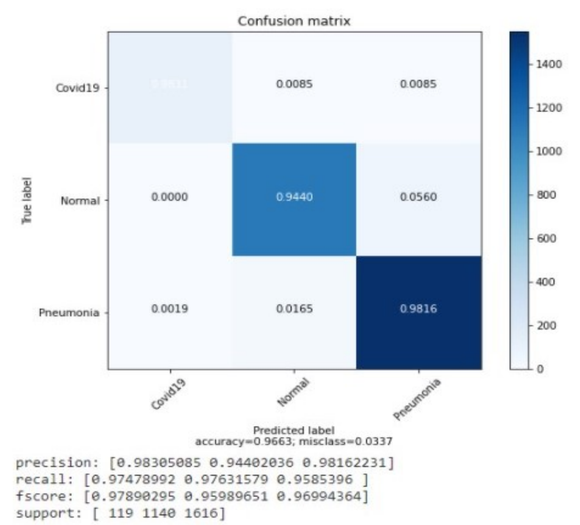

Fig. 6. ResNet50 with 11,506 X-ray Images

In Figure 5's experiment with small dataset of 415 COVID-19, 1,775 Pneumonia, and 1,577 Normal X-ray images, the results were stable across the F-1 scores, ranging from .93 to .96. In addition, we experimented the same model with more data of 478 COVID-19, 6,465 Pneumonia, and 4,563 Normal X-ray images. The results were improved with .96 to .98 F1 scores across the classes as show in Figure 6. We expect to see a better result with more data input.

VI. One versus All ResNet Model

A. Overview

In addition to the simple ResNet model's result, we created our own model utilizing three ResNet instances. Figure 7 displays the planned architecture for the model as made before implementation, and Figure 8 displays our implementation of the model as visualized through Tensorflow.

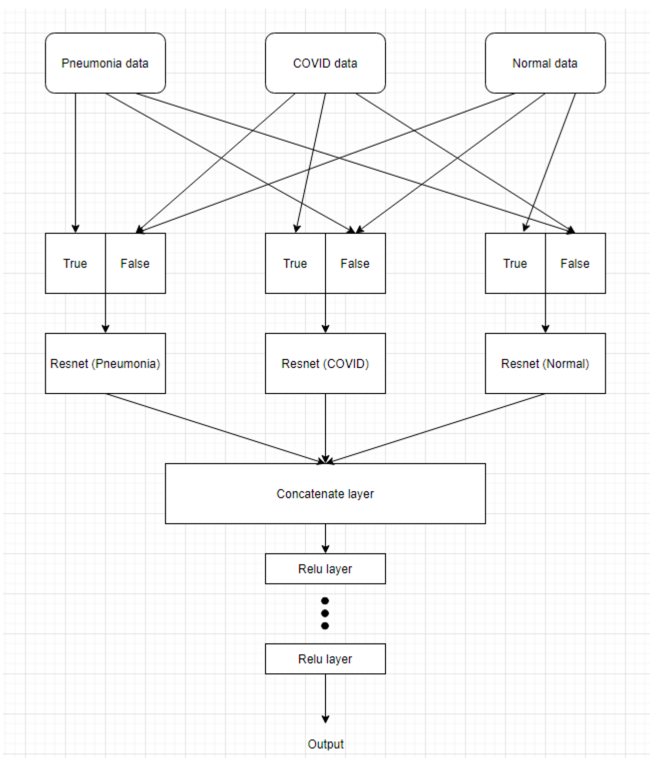

Fig. 7. One Versus All ResNet Model

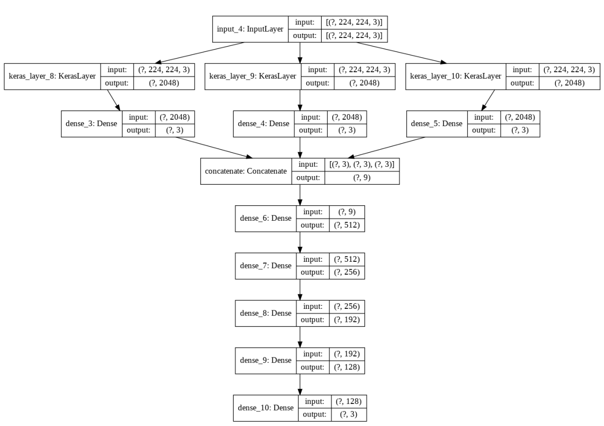

Fig. 8. One Versus All ResNet Model in Tensorflow

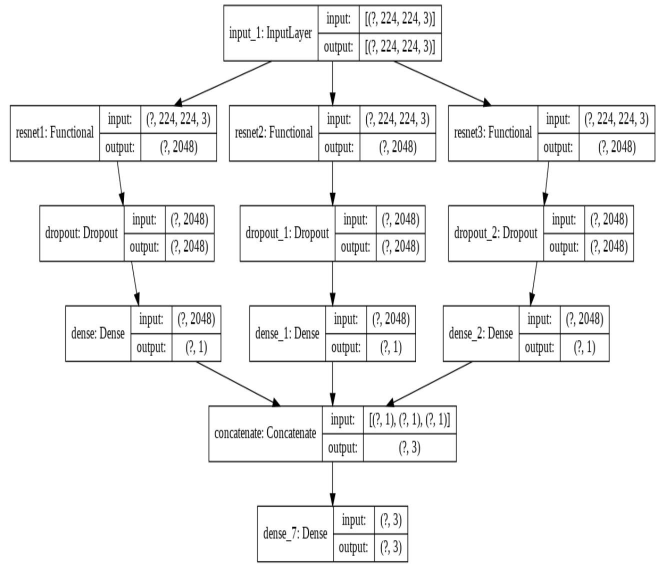

Fig. 9. Triple Resnet Model, One Versus One Variation

Theoretically speaking, by dividing the classification task into sets of two and assigning each binary classification problem to an individual neural network, it is possible to obtain better results than simply using a single model for multi-class classification. There are two main methods of grouping the classes—one versus all and one versus one.

One versus all refers to taking a single class and assigning a label to it, and then assigning the opposite label to all of the other classes. This creates a situation where the binary classifier is trained to pick out images of that single class from the entire dataset. By doing this for every class, it is possible to obtain separate dataset for training N different binary classifiers. This is the method used for the original iteration of the model (Figure 8). First, the dataset was prepared for one versus all training, after which 3 Resnet50 models were trained on the prepared dataset. Finally, the trailing dense layer chain was trained using the concatenated output from the 3 Resnet50 models.

One versus one refers to dividing the entire dataset into groups of two classes and training binary classifiers to differentiate between the two. In theory, this is superior to one versus all classification especially in cases like this where the dataset is imbalanced due to the number of classifiers ($\frac{N(N-1)}{2}$ for one versus one as opposed to N for one versus all). Figure 9 illustrates the current iteration of the model, where the main difference as compared to the previous version is the lack of the trailing dense layer chain in order for faster training during testing.

B. Problems

Unfortunately, our implementation of the model is not currently functioning correctly due to two main problems. The first problem seems to be an internal issue with Tensorflow 2 where frozen layers (layers set as not trainable) remain listed as trainable, causing issues during training. The second problem is that the final dense layer excludes classes during output. To be more specific, the earlier version of the model was incapable of outputting any predictions for class 0 (COVID-19), whereas the current version is incapable of predicting for classes 0 and 1 (COVID-19 and NORMAL). As of now, the exact source of this problem is unknown, but we suspect the structure of the model past the concatenation layer is one of the causes.

VII. CONCLUSION

X-ray images are cheap and easily accessible for analyzing various conditions. In this paper, we used X-ray images to analyze and detect three different classes: COVID-19, Pneumonia, and Normal conditions. Researches show that ResNet50 model has the highest correlation with X-ray images, and our simple experiments with the model has resulted in high performance. Our additional model of one versus all did not have produced a result yet, but the architecture we built is promising and we will continue to develop our ideas further on.

REFERENCES

[1] Al-Karawi, Dhurgham, et al. "Machine Learning Analysis of Chest CT Scan Images as a Complementary Digital Test of Coronavirus (COVID-19) Patients." medRxiv (2020).

[2] Wu, Xiangjun et al. "Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: A multicentre study." European journal of radiology vol. 128 (2020): 109041. doi:10.1016/j.ejrad.2020.109041.

[3] Che Azemin, Mohd Zulfaezal, et al. "COVID-19 Deep Learning Prediction Model Using Publicly Available Radiologist-Adjudicated Chest X-Ray Images as Training Data: Preliminary Findings." International Journal of Biomedical Imaging, 18 Aug. 2020, doi:https://doi.org/10.1155/2020/8828855.

[4] Wang H, Wang L, Lee EH, Zheng J, Zhang W, Halabi S, Liu C, Deng K, Song J, Yeom KW. Decoding COVID-19 pneumonia: comparison of deep learning and radiomics CT image signatures. Eur J Nucl Med Mol Imaging. 2020 Oct 23:1–9. doi: 10.1007/s00259-020-05075-4. Epub ahead of print. PMID: 33094432; PMCID: PMC7581467.

[5] Ioannis D. Apostolopoulos1, Tzani A. Mpesiana, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks", Physical and Engineering Sciences in Medicine, 3 April 2020

[6] Ali Narin, Ceren Kaya, Ziynet Pamuk, "Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks", Version 3, arXiv ,20 Mar 2020

[7] Horry, Michael J., et al. "X-Ray Image based COVID-19 Detection using Pre-trained Deep Learning Models." (2020).

[8] Wang, Linda, Zhong Qiu Lin, and Alexander Wong. "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images." Scientific Reports 10.1 (2020): 1-12.

[9] O'Shea, Keiron, and Ryan Nash. "An introduction to convolutional neural networks." arXiv preprint arXiv:1511.08458 (2015).

[10] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[11] Shoeibi, Afshin, et al. "Automated detection and forecasting of covid-19 using deep learning techniques: A review." arXiv preprint arXiv:2007.10785 (2020).

[12] "Coronavirus Cases:" Worldometer, www.worldometers.info/coronavirus/.