Adam Ten Hoeve

COMP 4449 – Data Science Capstone

Midterm Assignment Writeup

10/13/2021

**Introduction**

Covid-19 is a contagious disease caused by the SARS-CoV-2 virus. The disease primarily spreads by infecting a person’s lungs and then spreading to other host cells. Considering this, chest radiograms play a crucial role in the detection and diagnosis of Covid-19. Within chest x-rays, Covid-19 pneumonia is observed as a white density that obscures the expected lung markings. In typical cases, these white densities are spindly and slightly obscure the underlying markings. In more severe cases, they are much denser and completely obscure the underlying markings. An example of these densities is show in Figure 1. That raises the question, can we train a computational model to detect these densities and indicate whether Covid-19 pneumonia is present in the patient?

A close-up of a fetus

Description automatically generated with medium confidence

Figure 1: Example of Covid-19 Pneumonia

**Dataset Description and Exploration**

The dataset used can be found on Mendeley [[here](https://data.mendeley.com/datasets/fvk7h5dg2p/1)]. The data consists of 603 chest x-ray images, with 234 images of healthy patients, 221 images of patients with Covid-19, and 148 images of patients with bacterial pneumonia. An example of each can be seen in Figure 2.

A picture containing text, x-ray film, different, male

Description automatically generated

Figure 2: Examples of X-ray Images from the Data

The images range in size from 237x328 to 4300x4298 pixels. A full plot of the image dimensions can been seen in Figure 3. During preprocessing, all images were resized to 224x224 pixels. 80% of this dataset was used for training the multiclass classification models, while 20% was reserved as a test set. The training data was then further split into training and validation sets, that were of 80% and 20% of the data, respectively.

Chart, scatter chart

Description automatically generated

Figure 3: Image Dimensions

It should be noted that this dataset is small for a deep learning problem, with only 406 images in the training set. To account for this, I used image augmentation to artificially increase the number of samples in the training set. Augmentation is a process that slightly shifts, rotates, zooms, etc. the training images randomly to produce “new” images that are of the same class as the original. An example is show below in Figure 4.

A picture containing text, cat, looking, x-ray film

Description automatically generated

Figure 4: Example of Zoom Augmentation on the Same Image

Another method that may help with our small dataset problem is Transfer Learning (TL). TL is the process of reusing a model that has already been trained for one task as the starting point for a model of a second task. So long as the two tasks are relatively similar, then the models can leverage the existing network to create an accurate model with only a small dataset. In particular, I used the ResNet50 model, which has already been trained on a large dataset to do generic image classification, as the starting point in the hopes that it will generalize to our pneumonia classification task.

**Data Modeling**

Deep learning models are particularly effective at image classification tasks. I built several Convolutional Neural Network (CNN) with Adam optimizers and categorical cross-entropy loss functions. I tried many different model architectures, and the most successful model architecture is specified in Figure 5.

3x3 Conv, 32

3x3 Maxpool

3x3 Conv, 32

3x3 Conv, 64

2x2 Maxpool

3x3 Conv, 32

3x3 Conv, 64

2x2 Maxpool

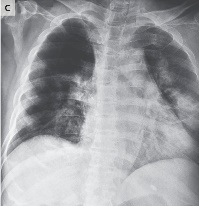
FC 64

FC 3

Normal

Covid-19

Pneumonia



INPUT

Figure 5: CNN Architecture

After 20 epochs, this model had a validation accuracy of 67.8% and a test accuracy of 61.1%. These results are good, especially considering the model is classifying between 3 different classes and the small number of epochs. The full training history can be seen in Figure 6.

Chart, line chart

Description automatically generated

Figure : Training History of CNN Trained From Scratch

But could we do better? I applied transfer learning to the ResNet50 model, including some with feature extraction others with fine-tuned layers. Unfortunately, the transfer learning model did not perform well as I was never able to get the model to perform better than random. The best validation accuracy was 35.6% and the best test accuracy was 39.8% after 40 epochs.

**Conclusion**

The CNN produced a final test accuracy of 61.1%. For a multi-class classification problem with three labels, that is a reasonable result. However, for the application of detecting Covid-19 pneumonia and diagnosing a patient, we would want a much more accurate model. With such a high error rate, this model could cause more harm than good. However, I do see the possibility of this model performing extremely well given more training and tuning.

If I were to continue working on this project, I would continue tuning hyperparameters. In particular, I believe the main reason the Transfer Learning models were unsuccessful was because of a bad learning rate and an inability to run enough epochs. Similarly, I would also move my models from my local laptop, which only has a CPU, to a platform with GPUs such as Google Colab so I could speed up the training times. That amount of time spent hurt my models, as they couldn’t train long enough, and likely resulted in the poor results. I will take steps to remedy these issues in the future.