Using Convolution Neural Networks to Trace Mask Wearing as a CDC Compliance/Health Metric

Henry Vasquez

Table of Contents

1. Abstract

The spread of SARS-CoV-2 has cause caused the shutdown of countries across the world and despite stay-at-home orders and facemask mandates there are still people that refuse to follow these health precautions. Studies have shown that facemasks are effective in preventing the spread of the virus from person to person while other analyses have noticed that countries who imposed a facemask mandate early on saw a slower spread rate of the novel virus. The purpose of this project will be to explore deep learning methods to build a model that detects the presence of a facemask. More specifically the paper will focus on architecture, training and evaluation of a Convolutional Neural Network (CNN). The potential use case in building this CNN is to track the percentage of people wearing their facemask in public compared to the number of cases in the corresponding area.

2. Introduction

SARS-CoV-2, or more commonly known as COVID-19 or novel coronavirus, has altered the way public gatherings are treated by requiring sets of rules in order for businesses to stay open. Although the rules are for the safety of those most at risk of having severe symptoms if falling ill to the virus, there are still some people that forget these rules exist or just disregard these guidelines all together. A tracking poll by YouGov/Economist has been following the percentage of people who say they wear a face mask in public, among other pandemic related questions. In the U.S., the tracker reports 29% of people saying that they wear a face mask in public as the beginning of April while by late July and onward the number skyrockets to the mid-high 70's (Smith, 2020).

Since the outbreak, the Centers for Disease Control and Prevention (CDC) has recommended that use of masks when out in public as a simple barrier to "help prevent droplets from traveling into the air and onto other people when the person wearing the mask coughs, sneezes, talks or raises their voice" (Centers for Disease Control and Prevention [CDC], 2020). The recommendation, according to the CDC is based on knowledge and evidence from clinical and laboratory studies that show the reduction of sprays (Centers for Disease Control and Prevention [CDC], 2020). Studies have been conducted to confirm the CDC's claims and have resulted in similar results. A study conducted by Chris Beyrer, a Professor at John Hopkins Bloomberg School of Public Health, concluded that facemasks protect the person wearing it and the surrounding people (Cimons, 2020).

Based on the studies that have been done, wearing masks in public and following the guidelines set by the CDC would slow the spread of the virus. Many big-name retailers such as Walmart, Amazon, Kroger, Costco and more have adopted mask policies and enforce them in order for customers to enter their stores (Markowitz, 2020). However, there are people out there that do not follow the guidelines for not wanting to or by mistake/forgetting the requirements. Some of these instances can be simply addressed by having employees on the floor watching out for customers in violation, but other times it can be overwhelming to keep up and catch those that refuse to follow the guidelines in place.

Being able to identify when people are following CDC guidelines would allow for a better understanding of how the public is reacting to the situation and how it varies across the country. The study conducted by Chris Beyrer was inspired by researchers that noted differences in the spread of the virus and the deaths between countries that adopted mask-wearing policies from the start and those that did not (Cimons, 2020). The outcome of this project would allow for a way measure what percentage of the population is following CDC guidelines and compare against the spread of the virus and the rates of sever disease and deaths. There are a number of datasets made up of faces and a few datasets that include faces wearing facemasks. This project will focus on utilizing those datasets along with deep learning methods to be able to identify whether a given individual is wearing a facemask.

The purpose of this project is to leverage deep learning techniques by creating a Convolutional Neural Network model that is capable of identifying the presence of a facemask given an image. By combining the model with tools such as OpenCV, there is the capability of identifying in real-time whether individuals are following procedures put in place by the CDC and businesses in order to provide a measurement against cases, severe disease and deaths as a result of the virus.

3. Literature Review

3.1. COVID-19 Studies

Many people in the United States oppose the idea of wearing a mask believing that they are ineffective in preventing the spread of the virus. The effectiveness of wearing a mask has been questioned despite the CDC's guidelines issued and other studies conducted showing the blocking of particles by wearing a mask. A study conducted using high-speed video showed the effectiveness of wearing a mask and the blockage of particles ranging from 20 to 500 micrometers. The study tested the effectiveness of masks using different methods and concluded that masks and respirators can be used in the prevention of the spread of COVID-19 (Forouzandeh et al. 2020).

The type of mask worn is also a debated issue as many people believe they need to wear N95 masks to be safe from the contracting the virus when in fact that is not the case. Most scientists and health care professionals recommend that the masks just cover the mouth and nose and those masks like the N95 may be overkill. In an article by the University of California San Francisco, it states that the most important consideration for masks for the general public may be comfort (Bai, 2020). The N95 respirators are only required in medical situations and surgical masks tend to be more protective and lighter/comfortable that cloth masks. The most important thing is that the mask covers the nose and the mouth to prevent the spread of the virus.

A study conducted by *Health Affairs* showed similar results to the study on face mask mandating and virus spread rate that influenced Chris Bryer's own study. The study from *Health Affairs* compared the growth rate of COVID-19 before and after face mask mandates across 15 states and found that mandates in these states resulted in a decline of the daily growth rate by 0.9 percentage points within the first 5 days and estimates that over 200,000 cases were averted by late May (Lyu and Wehby 2020).

3.2. Deep Learning and Neural Networks

There are several types of techniques in deep learning, such as Artificial Neural Networks (ANNs), Deep Neural Networks (DNNs) and Convolutional Neural Networks (CNNs). Each of the different techniques have their own advantages in different fields. CNNs are most commonly used in analyzing visual imagery (Saha, 2018). CNNs work by passing an image through several layers which then process and perceive data in order to classify the image (Bansari, 2019).

CNNs models can be built by utilizing Keras, an open-source neural network library. OpenCV is also helpful for real time computer vision. It is also open-source and has Python support allowing for real-time images to be processed using the CNN model built using Keras. There are also predefined algorithms for face detection, such as Haar Cascade. Haar Cascade classifiers work by analyzing an image for a set of features. If an image matches enough of the features, the algorithm will determine whether the image includes a face(s), or eye(s) depending on the classifier used.

These resources and tools are useful in building new models that can detect objects of interest or for the purpose of this project, a facemask. By combining machine learning and deep learning techniques, this helps businesses, governments and other entities in identifying those individuals that may not be following public health guidelines that have been put in place to prevent the spread of COVID-19. The benefit of being able to identify these individuals following public health guidelines or not is being able to calculate the percentage of people following preventative measures. In doing so, a comparison can be drawn between COVID-19 cases and deaths vs. the percentage of people taking precautions when out in public. For now, all that has been available are statistics corresponding to countries enforcing masks and the number of cases before and after, not the number of people taking the necessary precautions vs. the number of cases.

4. Data

4.1. Kaggle and Student Code-In

The two datasets being used for this project are from Kaggle and Student Code-In (SCI). Both datasets were created for the purpose of creating face mask detection applications in response to health guidelines recommended to reduce the spread of COVID-19. The datasets differ slightly in that the Kaggle dataset includes a third class "mask worn incorrectly" whereas the SCI dataset only includes the classes "with_mask" and "without_mask." For simplicity, images that have a true value for the mask worn incorrectly variable in the Kaggle dataset will be removed to stay consistent with the SCI dataset.

The SCI dataset consistent of 3,835 images of which 1,917 are classified as "with_mask" and 1,920 are classified as "without_mask" (GitHub). After removing images from the Kaggle dataset with the mask worn incorrectly class there are a total 10,000 broken out by 5,000 "with_mask" and 5,000 "without_mask" (Kaggle). The combine dataset leaves 13,837 total images broken out by 6,917 classified with a mask and 6,920 without a mask. The breakout for the training, validation and test sets are as follows:

- 11,066 training images
- 2,213 test images
- 554 validation images

It is important to note that some of the images were not processed correctly due to some error during the image preprocessing. Any images that were not converted to grayscale or resized were not added to the final dataset.

4.2. Image Preprocessing

The dataset of images from both sources split images into two folders containing the respective classes, with mask and without mask. All images from both sources were combined into two folders for simplicity based on the category they belonged to. Once the images were in their respective folders, they were renamed to have a common naming format. Since the nature of the project is identifying the presence of a face mask, color is not important when extracting features from the image. Part of the image preprocessing included first transforming the images to gray scale and then performing a resize on all images so that they were the same size.

After performing the color changes and resizing, images were saved to a list and a separate list was created that contained the label for each image, i.e. with mask or without mask. The images were then normalized so that the min and max values are 0 and 1 respectively (values of white and black).

5. Methods

5.1. Convolutional Neural Network

CNNs work well in feature extraction and are more often used in image recognition and classification. Images in their simplest forms are matrices consisting of 1's and 0's, for that reason the images were converted to grayscale and normalized. The image is then processed through a set of layers, each layer extracting features and transforming the image until a classification is made.

The layers used in this project consist of convolution, activation, max pooling, flatten, dropout and dense layers. The convolutional layer is specified using a kernel size or filter and scans the image/matrix. An activation function is added in each layer which introduces non-linearity to the network allowing it to learn better. A pooling layer, depending on the one used, analyzes the filter created from the convolution layer and returns a value within the specified pool size. For example, in max pooling the layer would return the max value of the specified pool size over the filter size specified in the convolutional layer.

After adding the layers, or several layers the image classification begins by first flattening the image. The flatten layer turns the image/matrix into an array. Then, the dense layer and dropout layers are applied and a final dense layer to produce the image classification based on the number of available classifications.

5.2. The Model Architecture

For building the CNN the Keras library was utilized in Python. Building the CNN is rather simple whereas most of the time spent was on running the models. Keras includes the sequential API allowing to create models layer by layer. Both models followed similar builds in their layers with model 2 having slightly different parameters and an extra layer. Each layer included a convolutional 2D layer, followed by the activation function and then the max pooling layer. Model 1 included two sets of these layers and a set of fully connected layers and model 2 included three sets of these layers and the fully connected layers. See below for a breakout of the parameters used in each model:

Model 1	**Model 2**	**Model 3**
Input Image: 100x100x3	Input Image: 100x100x3	Input Image: 100x100x3
Conv2D: Filters = 32, Kernel Size = 3x3, Activation = Relu	Conv2D: Filters = 32, Kernel Size = 3x3, Activation = Relu	Conv2D: Filters = 32, Kernel Size = 3x3, Activation = Relu
MaxPooling2D: Pool Size = 2x2	MaxPooling2D: Pool Size = 2x2	MaxPooling2D: Pool Size = 2x2
Conv2D: Filters = 64, Kernel Size = 3x3, Activation = Relu	Conv2D: Filters = 64 Kernel Size = 3x3, Activation = Relu	Conv2D: Filters = 64 Kernel Size = 3x3, Activation = Relu
MaxPooling2D: Pool Size = 2x2	MaxPooling2D: Pool Size = 2x2	MaxPooling2D: Pool Size = 2x2
Flatten	Conv2D: Filters = 64 Kernel Size = 3x3, Activation = Relu	Conv2D: Filters = 128 Kernel Size = 3x3, Activation = Relu
Dropout: Rate = 0.5	MaxPooling2D: Pool Size = 2x2	MaxPooling2D: Pool Size = 2x2
Dense: Units = 32, Activation = Relu		
Dense: Units = 2, Activation = Softmax	Flatten	Conv2D: Filters = 128 Kernel Size = 3x3, Activation = Relu
		MaxPooling2D: Pool Size = 2x2
	Dropout: Rate = 0.5	
	Dense: Units = 32, Activation = Relu	Flatten
	Dense: Units = 2, Activation = Softmax	
		Dropout: Rate = 0.5
		Dense: Units = 32, Activation = Relu
		Dense: Units = 2, Activation = Softmax

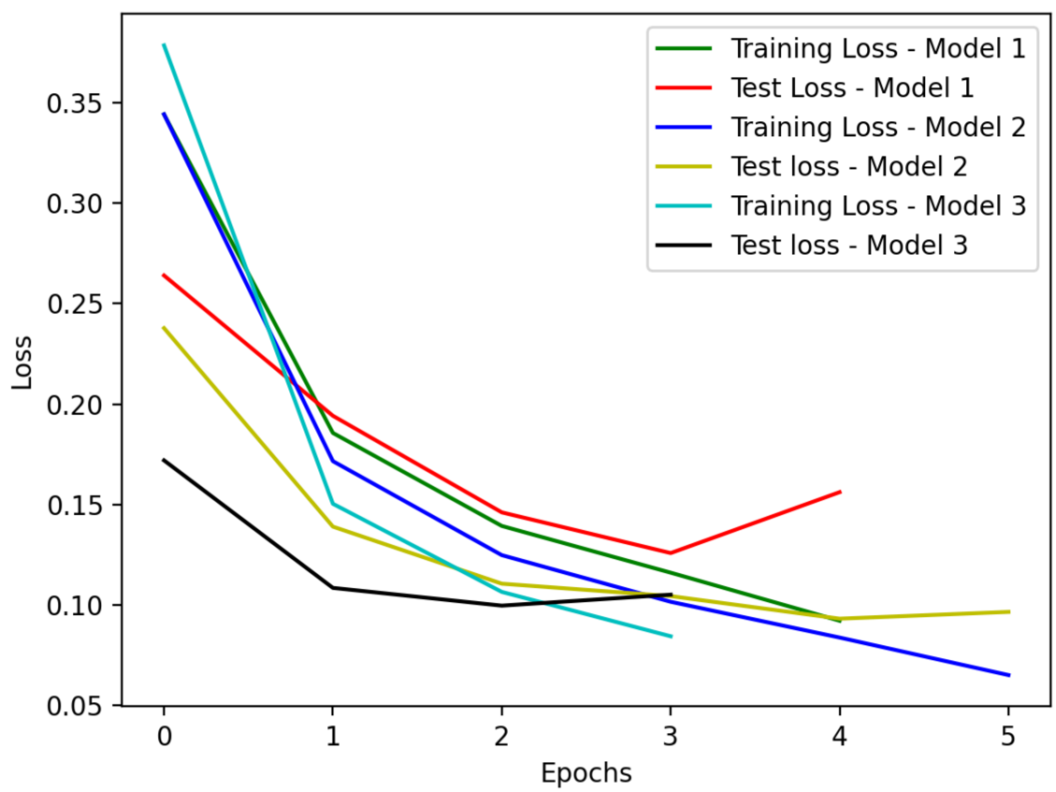
5.3. Model Optimization

When compiling the model, the optimization function used was Adam. The Adam function monitors the loss metric and in the case of both models the loss metric specified was Categorical Cross-Entropy. For the training of each model, 100 epochs were specified. An epoch is one cycle through the training set, however specifying too many epochs can lead to overfitting (Sarkar, 2020). To avoid over fitting the model, early stopping actions can be declared using Keras' callback library. This allows for the monitoring of a specified performance measure and will stop the training if the model does not improve on the specified metric. For the early stopping action, the validation loss was specified meaning if the model stopped improving on validation loss then it would stop it from continuing on to the next epoch and save the best fit model.

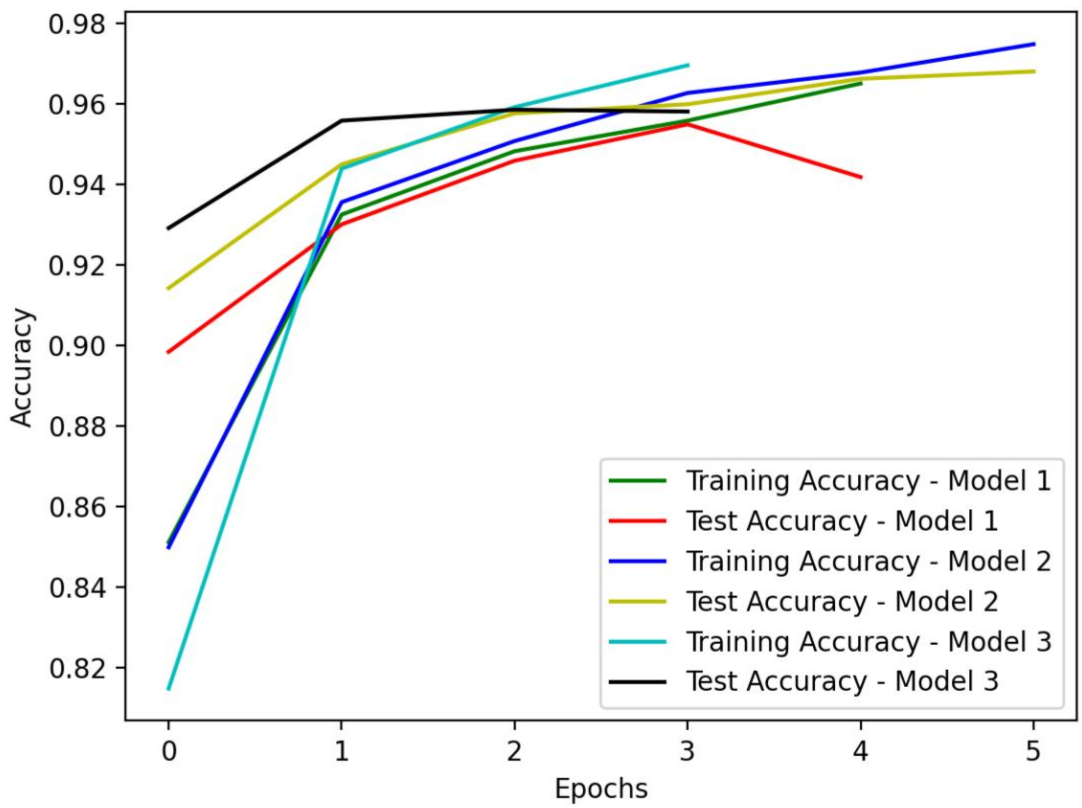
6. Findings

6.1. Model Evaluation

With the specified early stopping actions in place, model 1 was stopped after four epochs, model 2 after five epochs and model 3 after three epochs due to not improving on validation loss. Overall, models 2 and 3 had better results compared to model 1 with lower loss for the training and test data. See below for the comparison of loss between the three models:



Model 2 reported a better accuracy on the test data against the other models. The graph below shows the accuracies of each model on training and test data. The accuracy for model 1 appears to drop off after three epochs while for models 2 and 3 it continues to rise after three epochs.



The models were also evaluated against a validation set consisting of data that was unseen previously. The validation dataset consisted of 554 images, of which model 1 reported 94.40% accuracy with a 13.27% loss. Model 2 recorded a slightly higher accuracy at 95.85% and a much lower loss of 9.62%. Model 3 reported 97.11% accuracy and a loss of 7.31%. Based on accuracy and loss in both the model fit and on the validation set, model 3 was determined to be the better fit.

7. Conclusion

7.1. Comparing the Models

Models 2 and 3 differ from model 1 in that it they have extra layers, which adds to the complexity of the model. In terms of accuracy, both models 2 and 3 performed better than model 1 on the training, test and validation sets. The accuracy results from both models as well as the implementation of the early stop function during the model fitting suggest the models are not overfitting. Overall, model 2 outperforms model 1 in terms of accuracy by nearly 2% and model 3 outperformed model 2 by over 1%.

The improved accuracy of models 2 and 3 compared to model 1 can be attributed to the additional hidden layers allowing the model to extract more features and resulting in improved accuracy. In addition, the added layers in model 3 have a filter size that is double that of model 2's last layer whereas model 2 has only the extra layer with a filter of 64. By adding these extra layers along with increased filter sizes, the models are given more ways of extracting features, which then allows the learning to improve. Model 3 is built on 4 sets of layers, the last two containing filters of size 128. Model 2 is built using 3 layers, the last 2 having filter of size 64. Model 1 is built with 2 layers, the last layer having a filter size of 64. Model 2 performs better than the other models because of the added layers and also the increased filter value in the additional layers. Similarly, model 2 performs better than model 1 for the same reason.

7.2. Further Research

Both models perform relatively well, however they have their limitations. Although the models may detect the presence of a face mask, they do not detect whether the face mask is being worm correctly. A model to predict such instances might require additional hidden layers, making the model more complex and allowing it to extract more features. Additionally, using a different activation function, such as a Parametric ReLu or Sigmoid/Logistic, might be an interesting approach to improve on the current models or expand the current model to detect incorrectly worn facemasks.

7.3. Applications of Research

The models can be applied to produce a facemask tracking report to show the number of people actually wearing facemasks when out in public. The benefits of producing such a report would allow for the comparison of COVID-19 cases, deaths and severe cases vs the percentage of people wearing their masks in public. Countries that introduced a facemask policy resulted in the slowing of cases and deaths (Cimons, 2020). By tracking and reporting actual numbers, cities and governments can identify potential problem areas based on how the public is following facemask policies. However, this again depends on finding a relationship between facemasks and COVID-19 cases.

8. Sources

Bai, Nina. "Still Confused About Masks? Here's the Science Behind How Face Masks Prevent Coronavirus." *Still Confused About Masks? Here's the Science Behind How Face Masks Prevent Coronavirus | UC San Francisco*, 15 Oct. 2020, www.ucsf.edu/news/2020/06/417906/still-confused-about-masks-heres-science-behind-how-face-masks-prevent.

Bansari, Simran. "Introduction to How CNNs Work." Medium, Data Driven Investor, 12 Feb. 2019, medium.com/datadriveninvestor/introduction-to-how-cnns-work-77e0e4cde99b.

Centers for Disease Control and Prevention. Considerations for Wearing Masks. https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/cloth-face-cover-guidance.html#:~:text=(source%20control).-Evidence%20for%20Effectiveness%20of%20Masks,talks%2C%20or%20raises%20their%20voice. Accessed September 1, 2020.

Cimons, Marlene. "Wearing a Mask Protects You and Those around You." The Hub, John Hopkins University, 24 Aug. 2020, hub.jhu.edu/2020/08/24/masks-protect-wearer-and-others/. Markowitz, Andy. "Face Masks Required at 10 Biggest Chains, Including Walmart." AARP, 11 Aug. 2020, www.aarp.org/health/healthy-living/info-2020/retailers-require-face-masks-coronavirus.html.

Parnia Forouzandeh, Kris O'Dowd, Suresh C. Pillai. (2020) Face masks and respirators in the fight against the COVID-19 pandemic: An overview of the standards and testing methods. *Safety Science* 133, 104995.

Saha, Sumit. "A Comprehensive Guide to Convolutional Neural Networks - the ELI5 Way." Medium, Towards Data Science, 15 Dec. 2018, towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53.

Sarkar, Prantik. "How to Choose Number of Epochs to Train a Neural Network in Keras." *CodeSpeedy*, 20 June 2020, www.codespeedy.com/how-to-choose-number-of-epochs-to-train-a-neural-network-in-keras/.

Smith, M. (2020, June 08). International COVID-19 tracker update: 8 June. Retrieved November 22, 2020, from https://yougov.co.uk/topics/international/articlesreports/2020/06/08/international-covid-19-tracker-update-8-june

Wei Lyu and George L. Wehby, et al. "Community Use Of Face Masks And COVID-19: Evidence From A Natural Experiment Of State Mandates In The US." *Health Affairs*, 16 June 2020, www.healthaffairs.org/doi/10.1377/hlthaff.2020.00818.

9. Datasets

GitHub: https://github.com/chandrikadeb7/Face-Mask-Detection

Kaggle: https://www.kaggle.com/andrewmvd/face-mask-detection

10. GitHub Repository

https://github.com/hvasquez81/Data-698-Project