Face Mask Classifier

Utsav Negi (UG)
Gahan Sudhir (UG)
Vidyuth Mallappa (UG)
Group 10

Fall 2022
ECE/CS/ME 539 Artificial Neural Networks
University of Wisconsin-Madison

Abstract

Face masks play a crucial role in the post-pandemic situation. Keeping in mind the importance of the face mask, the project aims to use the concepts of neural networks and its algorithms to design a model which can designate whether a person is wearing a mask or not. The project trains on a dataset consisting of face-cropped images and predicts whether a face in an image is wearing a facemask or not. To check the accuracy of our predictive model, we have a testing dataset that reveals the percentage of correct classifications our model predicts. The trained model is combined with a pre-trained face detection model to detect multiple faces in a frame and classify them in real-time.

Table of Contents

Introduction

During the year 2020 and 2021, the world faced one of the worst pandemics in human history. The pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-COV 2) resulted in more than 1 million deaths in the United States (CDC, 2022). Even so, the case mortality rate of the virus is around 10% (Munster et. al, 2020), which is substantially less than its severe variants, the virus is highly contagious. This virus usually spreads through droplets and small particles breathed out by an infectious person, which can possibly be inhaled by a healthy person, or land on eyes, noses, mouth, and other surfaces (CDC, 2022). Such virus transmissions can be slowed by wearing a face mask.

Face masks prevent the contagious droplets from entering the nose and mouth, and thus preventing the transmission of the virus. According to studies conducted by Alihsan et. al, the probability of getting infected while wearing a face mask was around 7%, while the probability of getting infected without a face mask was more than 50% (2022). This suggests that the face mask can prevent the spread of virus in the community. However, there are several instances where people avoid wearing a face mask properly in public places.

The major reason behind such negligence is the discomfort of wearing a face mask for a prolonged period of time. Furthermore, wearing a mask also interferes with verbal and non-verbal communication among the people, and formation of mists across the glass lens when mask and glasses are worn together. In fact, vaccines against the virus are available, which resulted in less consumption of face masks. However, one must note that this virus can mutate itself and the resulting variants can make vaccines less effective. Hence, wearing a face mask is a precautionary measure to prevent transmission of the virus in a heavily crowded environment and avoid another surge of virus infection.

Inspired by the current scenario of the post-pandemic world and keeping in mind about the importance of face mask, the project aims to build an optimized face mask detector model. This model can assist the concerned authority to detect people not wearing their face mask properly.

The project uses two different datasets. These datasets are used to train the classification model separately and their respective performances are compared. The performance of the model will be evaluated on the basis of confusion matrix and the area under the RUC curve. Furthermore, the classification model will be tested along with the face detection model to detect face mask in real time.

Data Processing

The project uses two different datasets for the face classification model. The aim behind this approach is to compare the performance of the face classification model under different datasets.

Face Mask Dataset A

The Dataset A is a dataset known as Face Mask Detection ~12K Images Dataset (Jhangra, 2021). Sourced from Kaggle, this dataset consists of 11792 cropped images of faces. Out of the 12000 images, 10000 images are used for training the classification model, 800 images are used for validation step and 992 images are used for testing the classification model. Note that, the number of images include the number of masked face images and unmasked face images. In fact, the dataset consists of augmented images and hence the project doesn't require any data augmentation steps.

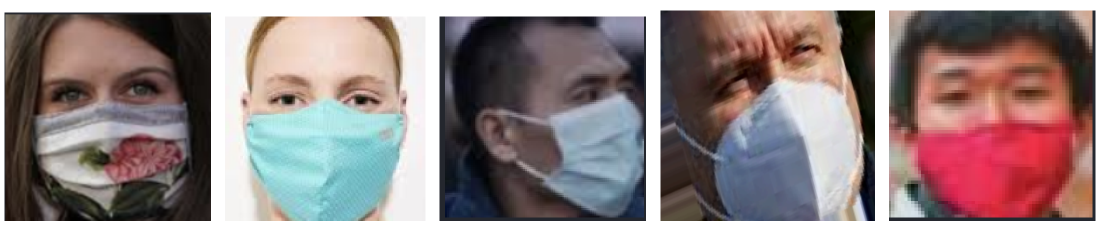

Figure 1: *Examples of masked face images from Face Mask Dataset A.*

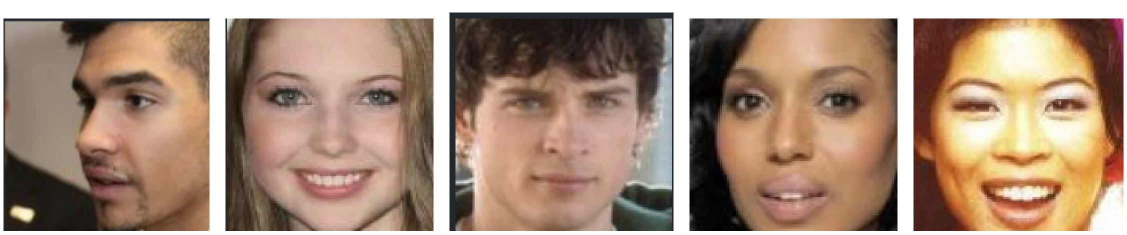

Figure 2: *Examples of Non-Masked Face Images from Face Mask Dataset A. Note that the dataset includes faces consisting moustache and beard as part of non-masked faces*

Face Mask Dataset B

The Dataset B is a dataset known as Face Mask Lite Dataset (Kottarathil, 2020). Sourced from Kaggle, this dataset consists of 20000 images where each image is having a resolution of 1024 by 1024. Considering the size of dataset, the project uses 10800 random images from the dataset for training and validating the model. For testing purposes, the project uses the test dataset from Face Mask Dataset A for comparing the respective performance of both datasets. Similar to Face Dataset A, the dataset consists of augmented images and hence the project doesn't require any data augmentation steps.

Figure 3: *Examples of masked face images from Face Mask Dataset B.*

Figure 4: *Examples of Non-Masked Face Images from Face Mask Dataset B. Note that the dataset includes faces consisting moustache and beard as part of non-masked faces*

The data preprocessing steps involves resizing the image to 100 by 100 pixels as the image is highly detailed and on resizing, the resultant image is able to hold the essential features, which results in faster training process. In addition to resizing the images, each image in the dataset is normalized to prevent any outliers affecting the training process of the classifier. In addition to image processing, each image is labelled by a 1 by 2 tensor such that if the label is [1, 0], then the face is wearing a mask and if the label is [0, 1] then the face is not wearing a mask.

Applied Methodology

Face Detection Model

Face Detection Model is used to detect the face from a given image. This model is a pre-trained model created using dlib face recognition library. The face recognition library detects the important features of the face by comparing the features of the image against the 68 facial landmarks. If the features match with the landmarks, the face recognition model returns the bounding box coordinates of the detected face. These facial landmarks are beneficial for this project as it can detect masked faces more efficiently.

Figure 5: *Visualization of 68 facial landmarks produced by dlib facial recognition (Rosebrock, 2017)*

The main purpose of the face detection model is to detect faces during the real time application and resize the cropped face image to 100 by 100 pixels. Subsequently, these cropped images serve as the input to the classifier model. Furthermore, the face detection model is able to generate multiple cropped images if it detects multiple faces in the same image.

Face Mask Classification Model

The face mask classification model is a convolutional neural network specified in a research paper by Abbasi et. al (2021). Considering 100 by 100 by 3 as the input layer, the model consists of three Convolution-Relu-Pooling blocks and three dense layers. The output layer is 1 by 1 SoftMax layer which is the prediction determined by the model. The classification model is trained using Face Mask Dataset A and Face Mask Dataset B

Constructed using TensorFlow framework, the model is trained for 150 epochs with batch size of 8 and learning rate of 0.001. The model uses Adam Optimizer for optimizing the weights and CategoricalCrossEntropy as the loss function.

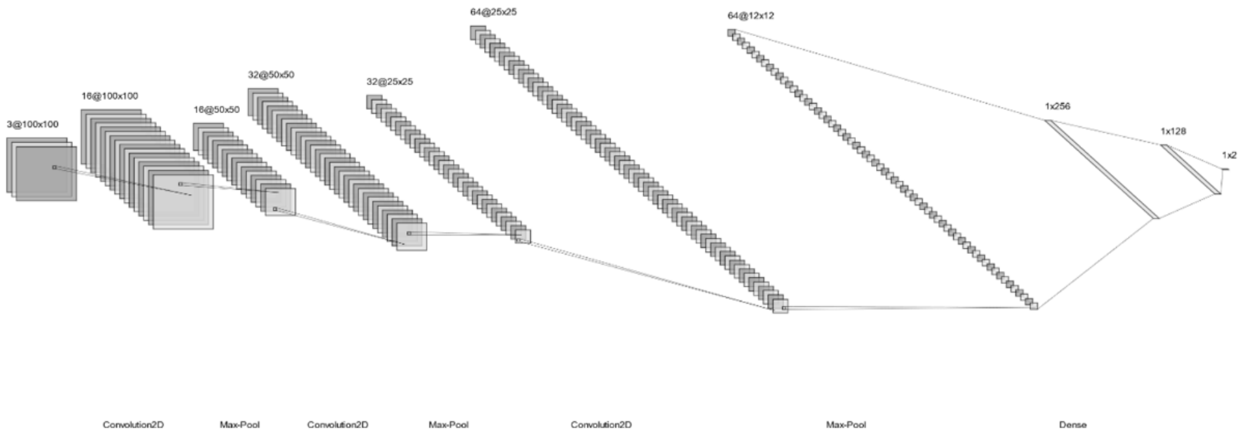

Figure 6: *Neural Network Architecture of The Classifier Model (Abbasi et. al, 2021).*

The above model can be summarized in the following table

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 100, 100, 16)	448
activation (Activation)	(None, 100, 100, 16)	0
max_pooling2d (MaxPooling2D)	(None, 50, 50, 16)	0
conv2d_1 (Conv2D)	(None, 50, 50, 32)	4640
activation_1 (Activation)	(None, 50, 50, 32)	0
max_pooling2d_1 (MaxPooling2D)	(None, 25, 25, 32)	0
conv2d_2 (Conv2D)	(None, 25, 25, 64)	18496
activation_2 (Activation)	(None, 25, 25, 64)	0
max_pooling2d_2 (MaxPooling2D)	(None, 12, 12, 64)	0
flatten (Flatten)	(None, 9216)	0
dense (Dense)	(None, 256)	2359552
dense_1 (Dense)	(None, 128)	32896
dense_2 (Dense)	(None, 2)	258

Total params: 2,416,290
Trainable params: 2,416,290
Non-trainable params: 0

Real-Time Application

The face detection model and the face mask classification are used together to detect whether multiple faces in a given frame are wearing a mask or not wearing a mask. The face detection model returns a list of values representing the bounding boxes coordinates around the detected faces. Using these coordinates, the detected face is cropped and resized into 100 by 100. The resultant image is used as an input to the face mask classification model. Based on the prediction made by the face mask classification model, the color of bounding box around the detected face changes. If the face mask is detected, then the color of bounding box is green, else the color of bounding box is red.

Results

Face Classification Model Results Using Face Mask Dataset A

The classifier model yields the training accuracy of 100% with 0% average loss, an average validation accuracy of 99.25 % with an incremental validation loss, and 98.5% testing accuracy with 3.46% average loss.

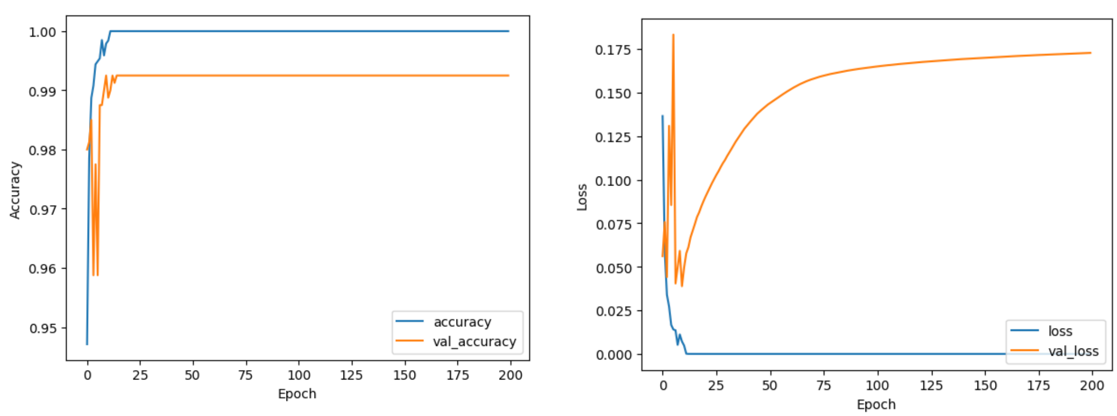

Figure 7: *The graphs on the left shows the trend of training accuracy and validation accuracy with respect to the number of epochs; the graph on the right shows the trend of training loss and validation loss with respect ot the number of epochs.*

The classifier's prediction on the test dataset generated the following confusion matrix

True / Predicted	With Mask	Without Mask
With Mask	501	8
Without Mask	7	476

The Area Under Curve (AUC) for this classifier is around <u>0.985</u>

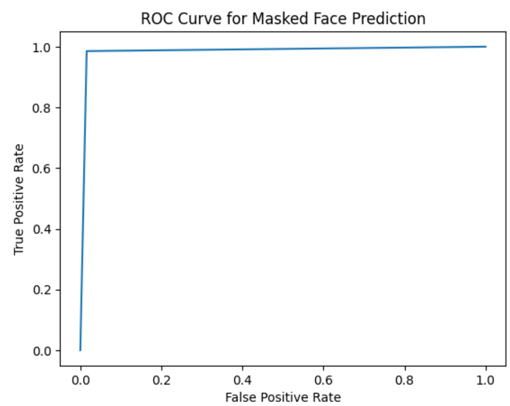

Figure 8: *The ROC Curve of the Classifier Model trained using Face Mask Dataset A.*

Real-Time Results of Model Trained Using Face Mask Dataset A

On using the face detection model and face classification model together, the prediction results are varied. Most of the time, the prediction is classifying the unmasked face as mask. Furthermore, the prediction is inconsistent with real time application and only detects accurately when a fixed frame is used. Furthermore, on applying stock images consisting of multiple faces, the resulted predictions are inconsistent. In fact, there are instances where the face is not getting detected at all, which resulted in the face not getting used as the input to the classifier. The average frames per second of the real time application is around 8 fps on NVIDIA RTX 3060 mobile GPU.

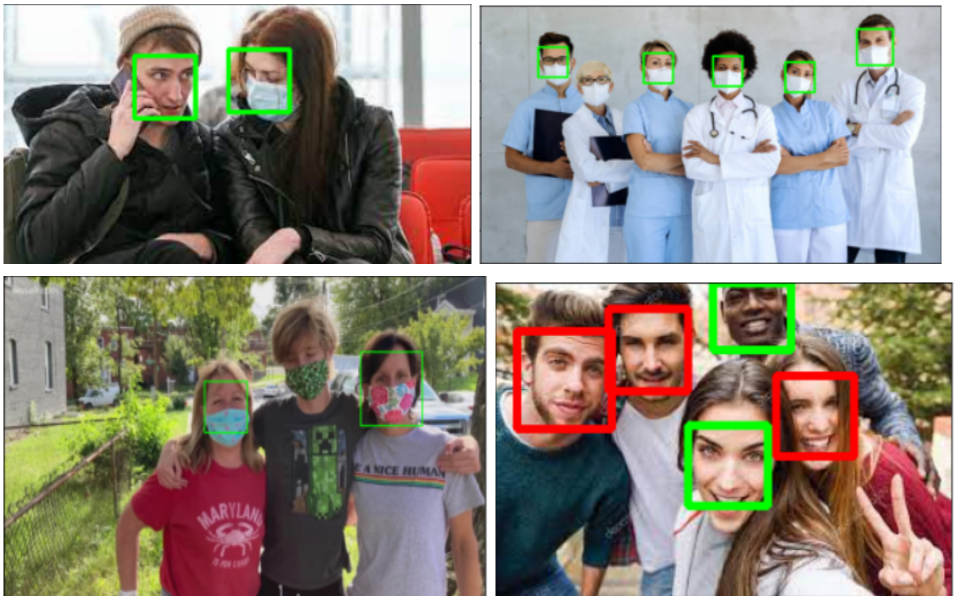

Figure 9: *The real time application of combined models. Here Green Border signifies mask detected, while the red border signifies mask not detected.*

Face Classification Model Results Using Face Mask Dataset B

The classifier model yields the training accuracy of 100% with 0% average loss, an average validation accuracy of 100 % with a decremental validation loss, and 82.6% testing accuracy with 5.81% average loss.

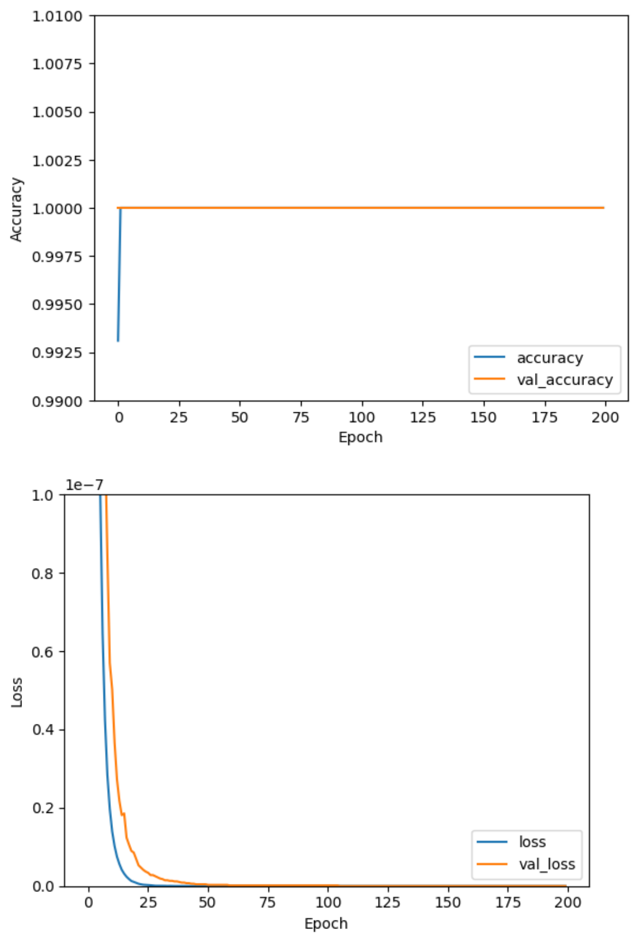

Figure 10: *The graphs on the left shows the trend of training accuracy and validation accuracy with respect to the number of epochs; the graph on the right shows the trend of training loss and validation loss with respect ot the number of epochs.*

The classifier's prediction on the test dataset generated the following confusion matrix

True / Predicted	**With Mask**	**Without Mask**
With Mask	450	0
Without Mask	156	294

The Area Under Curve (AUC) for this classifier is around 0.83

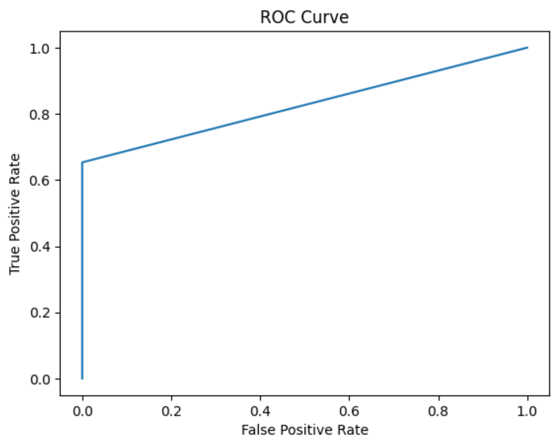

Figure 11: *The ROC Curve of the Classifier Model Trained Using Face Mask Dataset B.*

Real-Time Results of Model Trained Using Face Mask Dataset B

On using the face detection model and face classification model together, the prediction results are less varied compared to the model trained using Face Mask Dataset A. The model is able to classify a face as a masked or unmasked in the real time. But, on applying stock images consisting of multiple faces, the resulted predictions are inconsistent and there are instances where the face is not detected at all, which resulted in the face not getting used as the input to the classifier. In fact, the classifier classifies a face unmasked if the top section of nose is detected while wearing mask. The average frames per second of the real time application is around 9 fps on NVIDIA RTX 3060 mobile GPU.

Figure 12: *The real time application of combined models. Here Green Border signifies mask detected, while the red border signifies mask not detected.*

Discussion

From the above results, it is noted that the real time performance of model trained using Face Mask Dataset B is better than the model trained using Face Mask Dataset A, despite of having contrasting results in stock images consisting of multiple faces. This infers the fact that area under the curve is not sufficient enough to prove the real time performance of a model.

Furthermore, there is still a room of improvement towards the performance of the model. If given enough hardware capabilities, the face classification model could be trained using the complete Face Mask Dataset B as it turns out that the performance of model trained using a subset of Face Mask Dataset B is performing better in real life, compared to the entire dataset of Face Mask Dataset A. In fact, given enough time, there is a possibility of developing a custom face detection model that uses transfer learning from some of the object detection models such as FaceNet and ImageNet.

For further reference, the report has attached the link to the working Google Drive directory, which consists of datasets, trained models, training Jupyter notebooks, Google Collaborate notebook for real time face mask detection application and a real time face mask detection Python code for local systems.

Reference

1. ANN_Project Shared Drive, https://drive.google.com/drive/folders/1QdZ5VEtFtVhEqHbp57EZQ1-mqiVhUiQ7?usp=sharing
2. Alihsan, B., Mohammed, A., Bisen, Y., Lester, J., Nouryan, C. & Joseph, C. (2022), The Efficacy of Facemasks in Prevention of COVID-19: A Systematic Review, https://doi.org/10.1101/2022.07.28.22278153
3. Center of Disease Control and Prevention, CDC (2022), How COVID-19 Spreads
4. Center of Disease Control and Prevention, CDC (2022), CDC COVID Data Tracker
5. Munster, V., Koopmans, M., van Doremalen, N., van Riel, D. & de Wit, E. (Feb, 2020), A Novel Coronavirus Emerging in China — Key Questions for Impact Assessment, https://www.nejm.org/doi/full/10.1056/NEJMp2000929

6. Kottarathil, P. (2020), Face Mask Lite Dataset, Kaggle, https://www.kaggle.com/datasets/prasoonkottarathil/face-mask-lite-dataset?resource=download
7. Jangra, A. (2019), Face Mask Detection ~12K Images Dataset, Kaggle, https://www.kaggle.com/datasets/ashishjangra27/face-mask-12k-images-dataset