FACE MASK CLASSIFICATION

CHIA JUN

17067991

FACULTY OF COMPUTER SCIENCE AND
INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2021

TABLE OF CONTENTS

Table of Contents ii

List of Figures iv

List of Pseudocode v

List of Appendices vi

CHAPTER 1: INTRODUCTION 1

1.1 Objectives 1

1.2 Data Source 2

CHAPTER 2: ANALYSIS AND DESIGN 4

2.1 Literature Review 4

2.1.1 Face Mask Use in the Community for Reducing the Spread of COVID-19: A Systematic Review 4

2.1.2 SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2 5

2.1.3 Building Classification Model 5

2.1.4 The 4 Convolutional Neural Network Models That Can Classify Your Fashion Images 5

2.2 Algorithm Design 6

CHAPTER 3: EXPERIMENT 8

3.1 Own Convolutional Neural Network 8

3.2 MobileNet from Keras 13

CHAPTER 4: DISCUSSION 16

CHAPTER 5: CONCLUSION **17**

References 18

Appendix A 19

LIST OF FIGURES

Figure 1.1: Class Distribution 3

Figure 1.2: Sample Images 3

Figure 3.1: Training and Validating Accuracy 8

Figure 3.2: Training and Validating Loss 8

Figure 3.3: Input Images 9

Figure 3.4: Output Images 10

Figure 3.5: Own Input Images 11

Figure 3.6: Output of own test images 12

Figure 3.7: Training and Validating Accuracy 13

Figure 3.8: Training and Validating Loss 13

Figure 3.9: MobileNet Output Images 14

Figure 3.10: MobileNet Output on my own images 15

LIST OF PSEUDOCODE

Pseudocode 2.1: Data Processing..6

Pseudocode 2.2: Model Training 7

Pseudocode 2.3: Model Evaluation..7

LIST OF APPENDICES

Appendix A …………………………………………………………………... 19

CHAPTER 1: INTRODUCTION

As the Coronavirus disease 2019 (COVID-19) pandemic is getting worse even after 1 year, it comes to question whether technology can help in the effort of reducing the risk of infecting COVID-19. One of the ways to reduce the chance of infecting COVID-19 is to wear mask and maintain social distancing. Mask will reduce the chance of patient to spread the virus and block people to inhaling the virus from the environment.

Here comes of this project, make use of technology and Artificial Intelligence to check whether a person wear a mask and classify if that person wears the mask correctly. As from Coclite at el (2021), the use of face mask in the community helps in reducing the spread of COVID-19. This application will be meaningful in the case like entry control, where the system will detect person with or without mask and prohibit the one without mask to enter.

This project will mainly split into 2 parts, first is to detect faces using face detection model and from the faces detected, perform classification using classification model trained.

1.1 Objectives

1. **To detect faces in images**:

   To detect faces in images by using face detection model whether the faces are with mask or no mask.

2. **To classify faces detected based on the condition of mask worn**:

   From the faces detected from face detection model, classify the faces based on the condition of mask worn (No mask, mask, incorrect, or other covering.) using a classification model.

1.2 Data Source

The dataset used in this project is Face Mask Detection Dataset. This dataset contains 4326 unique images come together with bounding boxes of faces and respective classification annotations. There are 20 classes in the dataset, which are face_with_mask, face_with_mask_incorrect, mask_colorful, face_no_mask, mask_surgical, face_other_covering, scarf_bandana, eyeglasses, helmet, face_shield, sunglasses, hat, hood, goggles, hair_net, hijab_niqab, other, gas_mask, balaclava_ski_mask, and turban. The data distribution in each class is shown in Figure 1 below. The distribution of classes are imbalance and this might affect the accuracy of classification.

The reason why I choose this dataset is because it comes with vice variety of samples which might help to make my classification model more robust.

From the classes listed above, I found out there are classes like scarf_bandana, eyeglasses, hat, hood and goggles which are not helpful in my aim of distinguish people with suitable precautions to reduce the risk of infecting COVID-19. The classes that I will be using in this project are face_with_mask, face_with_mask_incorrect, face_other_covering and face_no_mask. Even though I only use 4 classes in our model, but almost every image in the data will be used given these classes are overlapped with the other classes which can be seen in Figure 2. Data pre-processing step will be taken to select out the classes which are suitable in our case for further training.

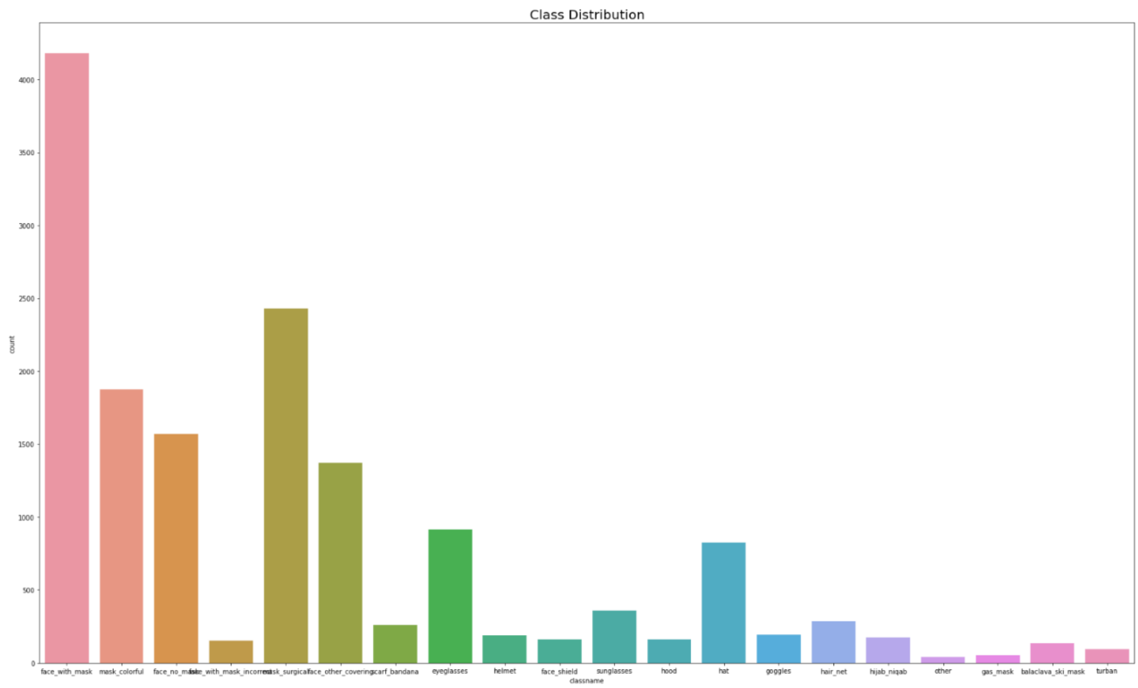

Figure 1.1: Class Distribution

Figure 2 below shows sample of images with respective bounding box annotations from the dataset.

Figure 1.2: Sample Images

CHAPTER 2: ANALYSIS AND DESIGN

2.1 Literature Review

2.1.1 Face Mask Use in the Community for Reducing the Spread of COVID-19: A Systematic Review

From Coclite at el (2021), COVID-19 is a new, rapidly emerging infectious disease caused by a novel coronavirus, SARS-CoV-2 (Severe Acute Respiratory Syndrome CoronaVirus-2), which is primarily transmitted via droplets during close unprotected contact with an infector and fomites. It has higher hospitalization and mortality rate than influenza and is spreading in an immune naïve population. Moreover, there is increasing evidence that people with mild or no symptoms at the pre-symptomatic and early stages of infection can contribute to the spread of COVID-19 too.

With this, Asian countries like China, Japan and Singapore implement the use of face mask whether medical or non-medical, which is adopted since 2003 SARS. Therefore, the team conducted a systematic review of the existing scientific literature, with randomized trials and observational studies, including modelling and experimental studies, on the effectiveness and efficacy of wearing face masks in the community for reducing the spread of COVID-19 in non-healthcare and non-household setting.

The research shows in the best-case scenario, when the mask efficacy is at 95%, the R0 can fall to 0.99 from an initial value of 16.90. Levels of mask filtration efficiency were heterogeneous, depending on the materials used (surgical mask: 45–97%). One laboratory study suggested a viral load reduction of 0.25 (95% CI 0.09–0.67) in favor of mask vs. no mask.

2.1.2 SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2

This paper by Nagrath at el (2021) is the existing work done for the same objective of the project, this paper introduces the architecture of how to combine detection model with classification model while in this case, the authors using Single Shot Detection model for face detection and MobileNetV2 for classification.

Their work gives me an idea on how to achieve the objectives where we will need to perform face detection on images with faces and perform classification on the faces detected from the face detection model.

2.1.3 Building Classification Model

Given the aim of this project is to detect and classify faces based on the present of mask on the faces. Article Building Classification Model (2019) gave me a thorough idea on how to design and implement the project despite the article only mentioning classification model, while this project requires one more model for face detection.

2.1.4 The 4 Convolutional Neural Network Models That Can Classify Your Fashion Images

This article by Le gives me idea on how to create my own convolutional neural network (CNN) from scratch for my classification model.

In this article, Le introduces several CNN-based classification models for image classification like CNN with 1 convolutional layer, CNN with 3 convolutional layers, CNN with 4 convolutional layers and VGG-19 pre-trained model. Based on the model architecture introduced by Le, I implement them in this project and test out which model brings the best result.

2.2 Algorithm Design

As mentioned in Data source session, the image in the dataset comes with bounding box and annotation, so data-processing will be required to crop out the faces and use the processed data for training a classification model. While face detection will be using pretrained face detection model based on the idea from Nagrath at el (2021), which provided solution on DNN-based face mask detection system using Single Shot Multibox dectector and MobileNetV2.

In data processing, we first create label list based on the classes we wanted to use, then the annotated faces belong to the selected labels will be cropped out and stored together with its respective class name. After finished processing, the processed data will be saved in a pickle file, or future use given the processing of data processing took long time.

Pseudocode 1 below shows the general flows of the data processing.

Pseudocode 2.1: Data Processing

```
Dataset Processing
1.  initialize data as empty list
2.  initialize labels as a dict of class name as key and ID as value
3.  initialize assign as dict of ID as key and class name as value
4.  for every unique image in the dataset
5.       j = respective image json file
6.       for every annotation in j:
7.           if class name in annotation is in labels:
8.                get bounding box coordinates
9.                crop out image based on bounding box
10.              append cropped image and class name as a list into
data.
11. randomly shuffle data
12. save data into data.pkl file
13. split data into X and Y numpy arrays
14. one-hot encoding Y
```

After the data is well set, the next step will be model training. Before that, the data will be split into training and test set. Then start training the classification model with experiment of different model like MobileNet[1] or self-create neural network.

The pseudocode below shows the general flows on model training.

Pseudocode 2.2: Model Training

```
Model Training
1.  initialize classification model
2.  setup loss, optimizer and metrics of the model
3.  split X and Y into training(xtrain, ytrain) and
validating(xval,yval)
    set
4.  initiate datagen by loading xtrain into ImageDataGenerator
5.  pass datagen into model fit_generator with batch size for start
    training
6.  plot training and validating accuracy
7.  plot training and validating loss
8.  save the trained model weight
```

After training the classification model, perform testing on the data using the model trained. By using face detection model to predict faces and using the trained model, predict the respective class.

Pseudocode 3 below shows the basic flow of the model evaluation.

Pseudocode 2.3: Model Evaluation

```
Model Evaluation
1.  for image in selected test images
2.       detect faces using selected face detection model
3.       for face detected with confidence higher than threshold:
4.            crop out the detected face from the image
5.            predict the class using classification model.
6.            draw the bounding box and write predicted class name on
the
              image
7.  plot all the images with respective results
```

[1] Keras Team. (2018). *Keras documentation: MobileNet and MobileNetV2*. Keras.io.
https://keras.io/api/applications/mobilenet/

CHAPTER 3: EXPERIMENT

3.1 Own Convolutional Neural Network

First, I carried out an experiment with my own convolutional neural network with 3 convolutional layers from scratch for classification model based on Le (2018). For detection model, I am using OpenCV Caffe model which is based on SSD. Below are the experiment results where Figure 3.1 represent the Training and Validating accuracy, Figure 3.2 represent the Training and Validating loss for my current classification model.

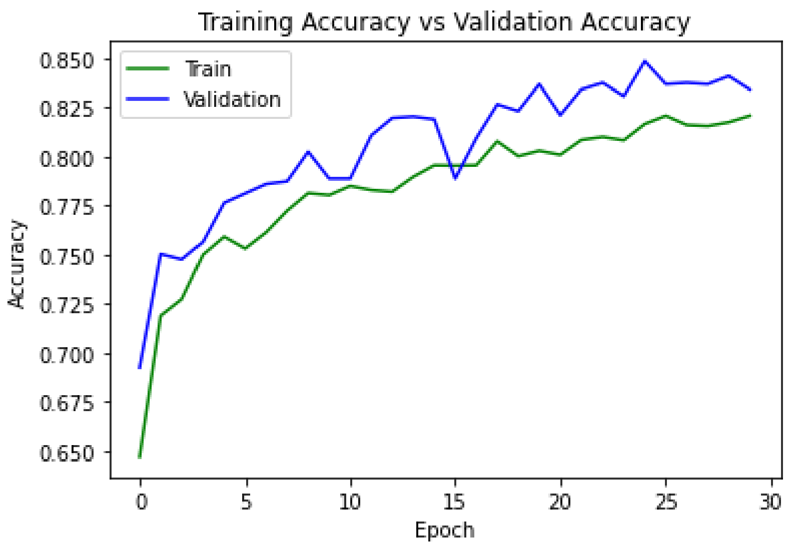

Figure 3.1: Training and Validating Accuracy

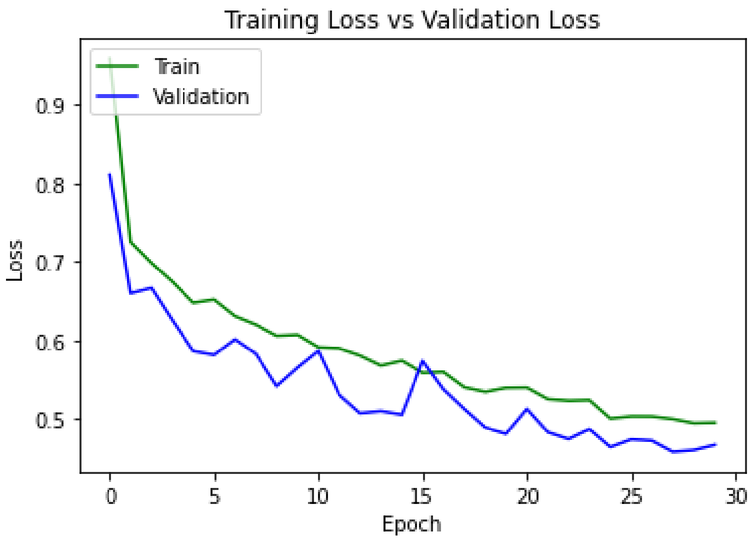

Figure 3.2: Training and Validating Loss

With the trained model, I test out the model with several inputs, Figure 3.3 below shows the input that I used to inference the result, and Figure 3.4 shows the output of my current model.

Figure 3.3: Input Images

Figure 3.4: Output Images

After inferencing the model, I found out the result is acceptable, but there are still some parts need to be pointed out, which is it does not able to detect the child wearing mask not properly in the first image. And I proceed the model to test on my own images as shown in Figure 3.5.

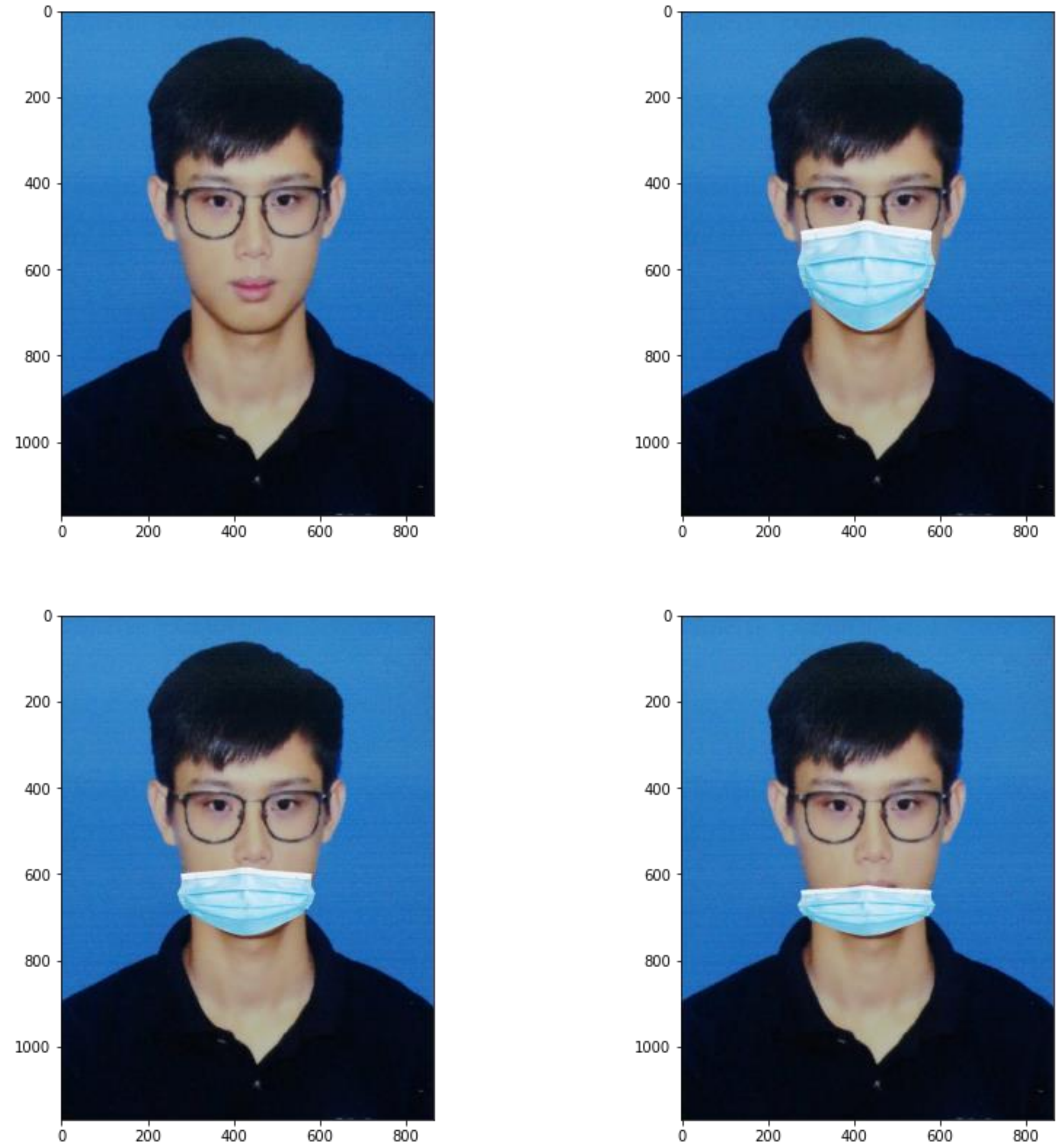

Figure 3.5: Own Input Images

The first image (up left) shows I do not wear mask, second image (up right) shows I am wearing a mask properly. Third image (down left) shows I does not wear mask properly with mask lower than nose. The last image (down right) also shows me do not wear mask correctly with mask at the chin.

The output of my current classification model with own input images are shown in Figure 3.6 in the next page.

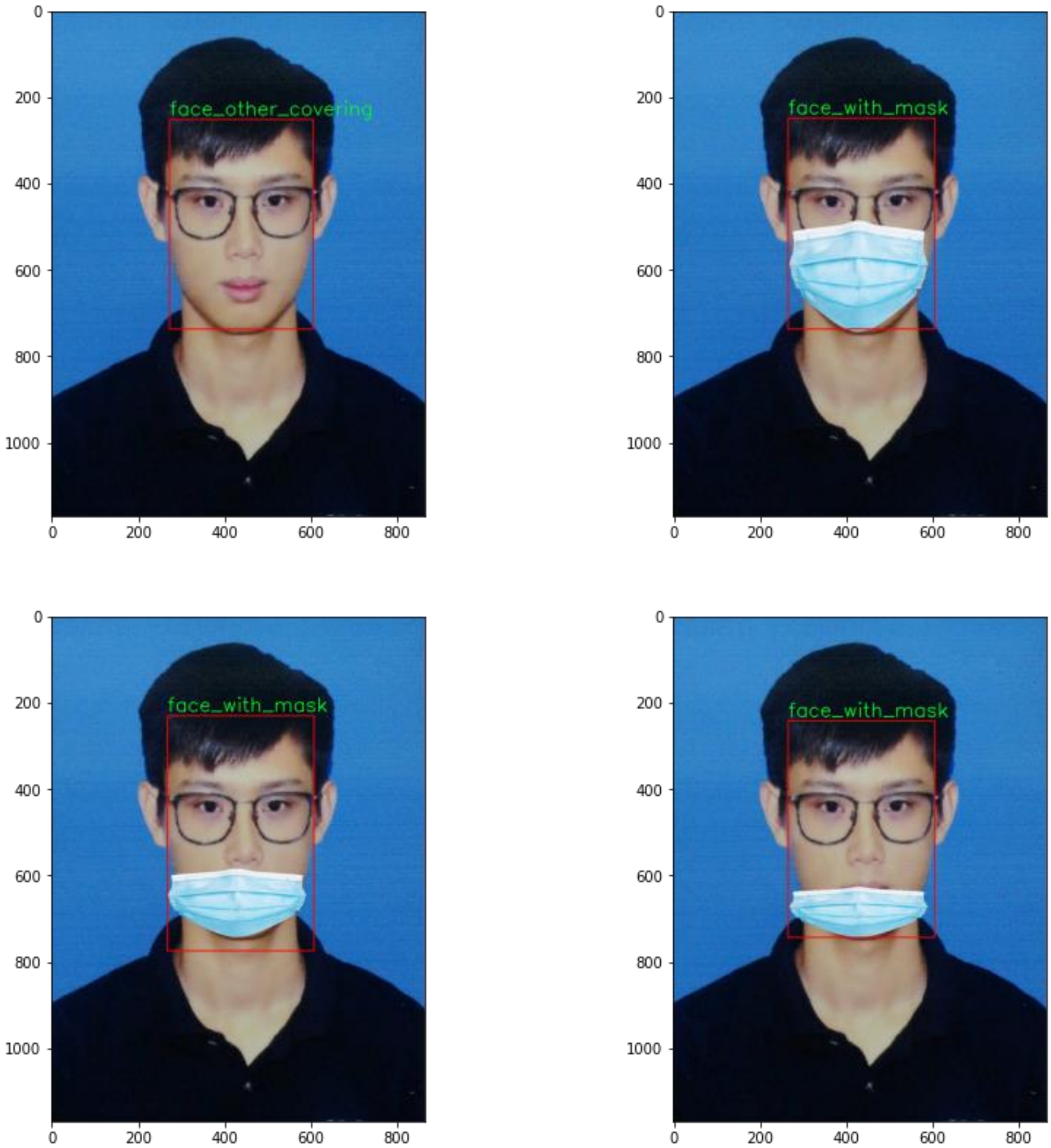

Figure 3.6: Output of own test images

Based on the results of current model, the first output image predicted as "face_other_covering" is suspected because I am wearing spectacles. Second image predicted correctly with "face_with_mask". While the third and last image shows the current model does not classified correctly for face does not wear mask correctly. I will continue experimenting with MobileNet in the next session.

3.2 MobileNet from Keras

In this session I experiment with implementing MobileNet from Keras by using its architecture and retrain with my dataset for classification. Where this idea of transfer learning is gained from Analytics Vidhya (2017). Below are the current experiment results where Figure 3.7 represent the Training and Validating accuracy, Figure 3.8 represent the Training and Validating loss.

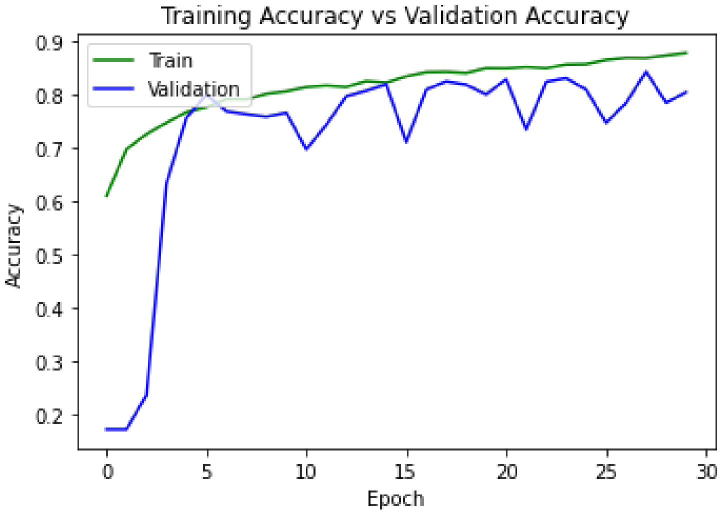

Figure 3.7: Training and Validating Accuracy

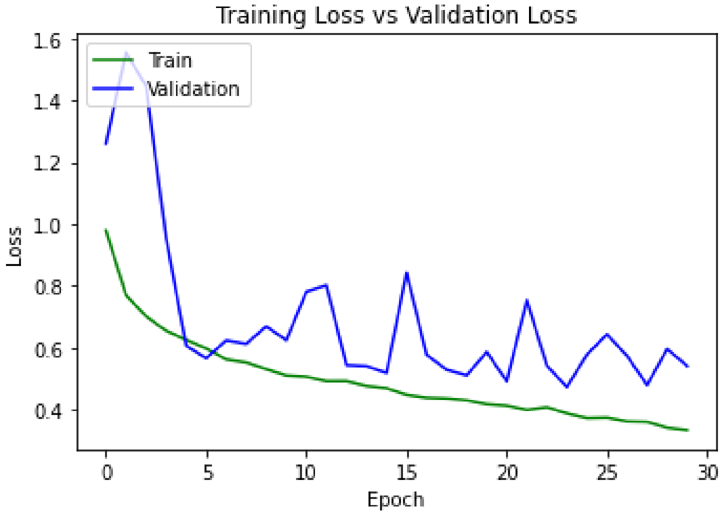

Figure 3.8: Training and Validating Loss

With the trained model, I test out the model with the same inputs as Figure 3.3, and Figure 3.9 shows the output of my current model.

Figure 3.9: MobileNet Output Images

Based on the current result, everything looks the same with the previous experiment, but the only difference is it is able to classify faces which wearing mask incorrectly as shown in first image above.

And I proceed to test on my own images as in Figure 3.5. The output is shown in Figure 3.10 in the next page.

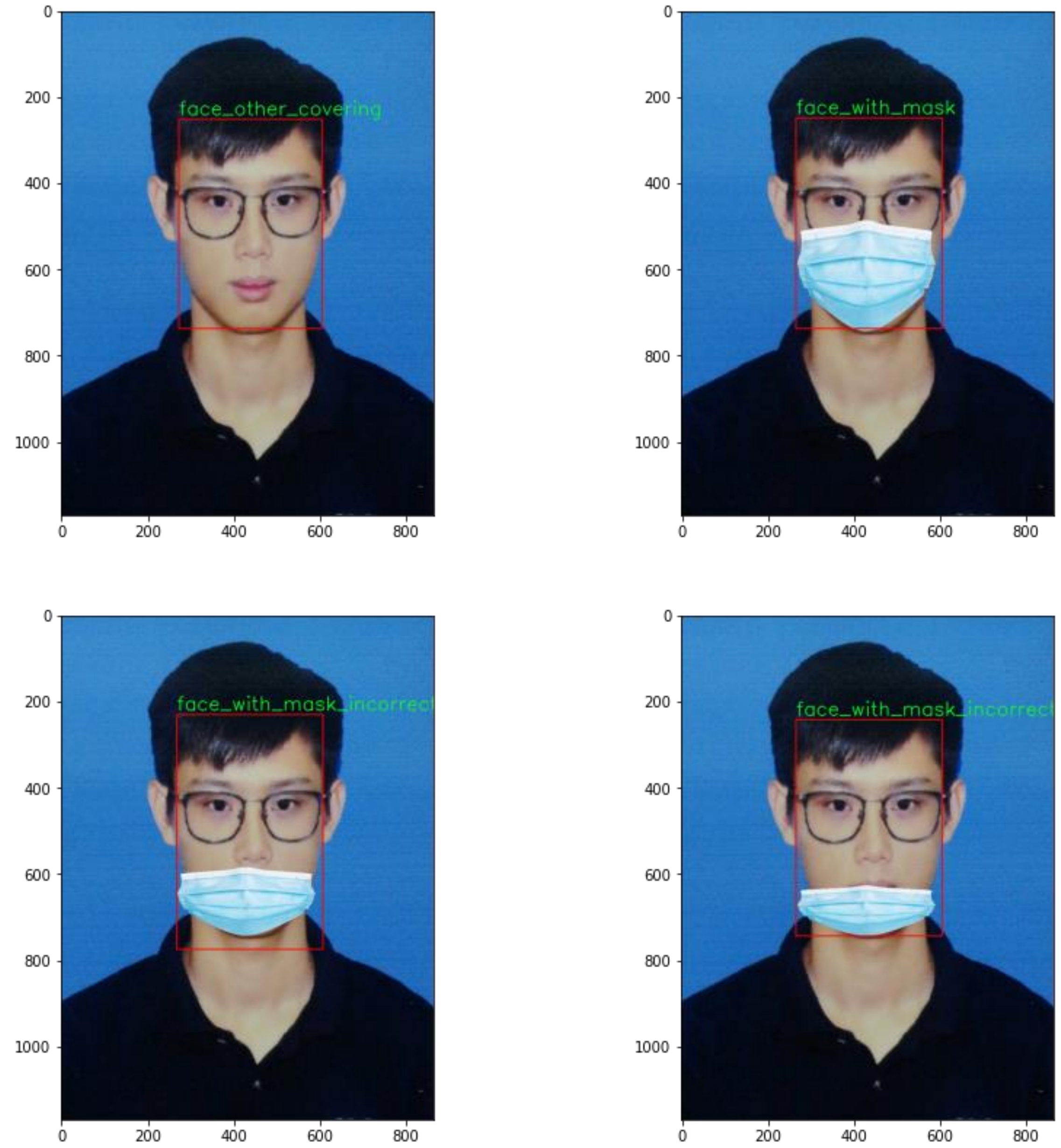

Figure 3.10: MobileNet Output on my own images

Based on both the outputs of MobileNet classifier, it shows that it performs better than my own CNN which build from scratch as it is able to detect my face with wearing mask incorrectly.

CHAPTER 4: DISCUSSION

Throughout the experiment, we can conclude that the classification retrained based on Keras MobileNet architecture works better with higher accuracy and lower loss. And it is able to classify face with wearing mask incorrectly although there are very few training data for this class. Although classification model created from scratch did bring acceptable results, but it could not distinguish face wearing mask incorrectly.

The strength of my current approach is it is simpler and fast to be trained given I am using pretrained model for face detection and face classification model is simpler as it is the same with other image classification model which using convolutional neural network.

The limitations of this project are my current approach required 2 different model handling different job which are face detection and another for classification, while it will be better and more straightforward if I can create an end-to-end model which will detect and classify faces at once. While another limitation is the current approach could not detect face mask worn by the faces if any. Given the pretrained model used is trained on face datasets.

The future improvement can be done is to create and end-to-end model which will be able to detect both human face and mask worn. This will be more practical in reducing the chance of infecting COVI-19 by using such technology in entry control for specific venue which require higher level of precaution as not every type of mask brings the same level of performance in reducing chance of infection.

CHAPTER 5: CONCLUSION

This project demonstrates how to create an image classification model with both create from scratch convolutional neural network and implementing transfer learning by using pre-trained MobileNet model from Tensorflow Keras to classify faces based on condition of mask worn.

This project included model building process and steps in building image classification model, starting from loading and preprocessing data, defining model architecture, training model and estimation of performance by inferencing.

In defining model architecture, I have to define how my model will look based on the number of layers, which activation function should be used, and how many hidden units each layer should have. While the most important criteria here is that the activation function of output layer must be Softmax given this project is to perform classification for more than 2 classes based on Brownlee, J. (2020).

While in training, we have to choose the right loss function used. As mentioned above my model will handle classification for more than 2 classes, I implement categorical cross-entropy loss function during my training.

The process of designing a neural network including the number of layers, activation function and loss function had been strengthen my theoretical knowledge studied in lectures and improved my understanding for future work.

REFERENCES

Analytics Vidhya. (2017, June). *Transfer Learning | Pretrained Models in Deep Learning*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2017/06/transfer-learning-the-art-of-fine-tuning-a-pre-trained-model

Analytics Vidhya. (2019, January 10). *Image Classification | Building Image Classification Model*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2019/01/build-image-classification-model-10-minutes

Brownlee, J. (2020, October 18). Softmax Activation Function with Python. Machine Learning Mastery. https://machinelearningmastery.com/softmax-activation-function-with-python/

Coclite, D., Napoletano, A., Gianola, S., del Monaco, A., D'Angelo, D., Fauci, A., Iacorossi, L., Latina, R., Torre, G. L., Mastroianni, C. M., Renzi, C., Castellini, G., & Iannone, P. (2021). Face Mask Use in the Community for Reducing the Spread of COVID-19: A Systematic Review. *Frontiers in Medicine*, *7*. https://doi.org/10.3389/fmed.2020.594269

Le, J. (2018, October 7). *The 4 Convolutional Neural Network Models That Can Classify Your Fashion Images*. Medium; Towards Data Science. https://towardsdatascience.com/the-4-convolutional-neural-network-models-that-can-classify-your-fashion-images-9fe7f3e5399d

Nagrath, P., Jain, R., Madan, A., Arora, R., Kataria, P., & Hemanth, J. (2021). SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. *Sustainable Cities and Society*, *66*, 102692. https://doi.org/10.1016/j.scs.2020.102692

APPENDIX A

Link to program folder

https://drive.google.com/drive/folders/18_6jg5ZKsIEnrXuuWH6eGcOlO17kCmYf?usp=sharing

Link to program code

https://colab.research.google.com/drive/1b8ptsah4ErG7-3wHVdKEuRauxOWWya_M?usp=sharing