Real Time Mask Classification for COVID-19

Muhammet Abdullah Soytürk
Koç University
msoyturk20@ku.edu.tr

Serhan Türkan
Koç University
sturkan20@ku.edu.tr

Bulut Bulgu
Koç University
bbulgu16@ku.edu.tr

Abstract

The world is currently facing a pandemic that was caused by a variant of a coronavirus and the most effective way that is known to prevent this virus from propagating quickly is to wear a mask in public spaces. The mask should cover the mouth, nose, and chin of the user for best protection. In this work, we introduce a real-time deep learning-based system that can detect situations where a mask is worn correctly, incorrectly, or not worn at all.

1. Introduction

According to the report published by the World Health Organization (WHO) on January 11, 2021, coronavirus disease 2019 (COVID-19) has infected 88.8 million people worldwide and caused more than 1.9 million deaths so far [1]. As the number of cases and deaths increased, people started to worry more about their health. Previous studies show that face masks can prevent the spread of the coronavirus [2]. Hence, many governments advise their citizens to use mask in public areas, some countries even made it mandatory to wear masks in public places to prevent the spread of the virus. In countries that made the usage of mask compulsory, citizens who do not comply with the regulation and are detected and penalized. However, it is not feasible to detect people who do not wear masks manually. Therefore, efficient face mask detection platforms are needed to protect global public health, but research on this subject is very new and limited.

While previous works [3,4] focus on binary problem of detecting whether a person wears a mask or not, there is only one study [5] that aims to detect the incorrectly worn masks. Wearing a mask incorrectly does not prevent the virus from spreading. According to the World Health Organization, the mask should cover the mouth, nose, and chin of the user. Hence, it is important to develop a system that can detect all three classes In this work, we aim to develop a system that can classify correctly worn masks, incorrectly worn masks and faces with no mask.

2. Related Work

2.1. Traditional Object Detection:

Traditional algorithms used in studies such as 'Rapid Object Detection using a Boosted Cascade of Simple Features' [6] and 'Histogram of Oriented Gradients' [7] have proven effective for such tasks, but these algorithms are largely based on feature engineering. In the age of deep learning, it is possible to train Neural Networks that perform better than these algorithms.

2.2. Convolutional Neural Networks:

Having superior spatial feature extraction capability and lower computational cost, CNN plays an important role in computer vision related pattern recognition tasks [8]. CNN uses convolution cores to combine with original images or feature maps to extract higher-level features. However, how to design better convolutional neural network architectures is still a matter of study. Residual Network (ResNet) proposed by K. He et al to train much deeper neural networks [9]. ResNet architecture can learn an identity mapping from the previous layer.

2.3. Face Mask Detection:

The study 'System for Medical Mask Detection in the Operating Room Through Facial Attributes' presented by Nieto-Rodríguez et al [3] detects the presence or absence of a medical mask. The main purpose of this study was to trigger an alarm only for medical staff who do not wear a medical mask, by minimizing as many false positive face detections as possible, without missing any medical mask detections.

3. Datasets

Three datasets have been employed to train a classifier: Face Detection Dataset [10], MaskedFaceNet Dataset [11] and Flickr-Faces-HQ (FFHQ) Dataset [12]. After all three datasets were combined, the combined dataset included 3232 correct worn mask, 1123 incorrectly worn mask, and 1717 without mask. The sizes of the train set, validation set, and test set are 60%, 20%, and 20% of the combined dataset, respectively. The three classes are called correct, incorrect and no mask for simplicity.

3.1. Face Detection Dataset

This dataset contains 853 images where some of the images contain more than one person. After the extraction of each person's face in all images, this dataset provides 3232 correct, 123 incorrect and 717 no mask images.

Figure 1: An example image from Face Detection Dataset.

3.2. MaskedFaceNet Dataset

This dataset contains 67193 correct and 66900 incorrect images where the images were automatically created with a face to mask model.

3.3. FFHQ Dataset

This dataset contains 70000 faces with no mask where the images were gathered by crawling Flickr.

4. Preprocessing

After the combined dataset was constructed, all images were resized to 256x256 for consistency. Then, channel-wise means and standard deviations were computed, and normalization applied. In order to be able to detect faces looking right or left better, horizontal flip augmentation was performed. In addition to that, random rotation augmentation was applied to better classify faces that are tilted left or right.

Figure 2: An example image to demonstrate horizontal flip

5. Experiments

	Training Accuracy	Validation Accuracy	Test Accuracy	Time (20 epoch)
Resnet18	98.890	96.410	0.971	2107.964
Resnet34	0.964	0.957	0.964	-
Resnet50	0.965	0.958	0.965	-
Vgg16	0.965	0.945	0.948	2715.669
Vgg19	0.951	0.948	0.953	-
Densenet	0.9714	0.976	0.970	3582.684
Mobilenet_v2	0.971	0.953	0.958	2244.607

Table 1: Accuracies and timings of different models (Reported times are for single GPU execution)

Seven different models have been trained to see which model performs the best fort the given task. The learning rate was dynamically decayed with ReduceLROnPlateau schedular of PyTorch when the validation loss stops decreasing. Batch size and weight decay was optimized with random search. It was found that the optimal batch size was 32 and weight decay was 1e-5. Adam was selected as the optimizer after trying out all optimizers in PyTorch including SGD, SGD + Nesterov, and Adagrad. The results are showing that resnet18 is the ideal choice fort the task at hand because it has the best accuracy (it also has the lowest training time as a side advantage). It has been observed that batch size affects the zig-zag behavior of the validation loss and accuracy significantly.

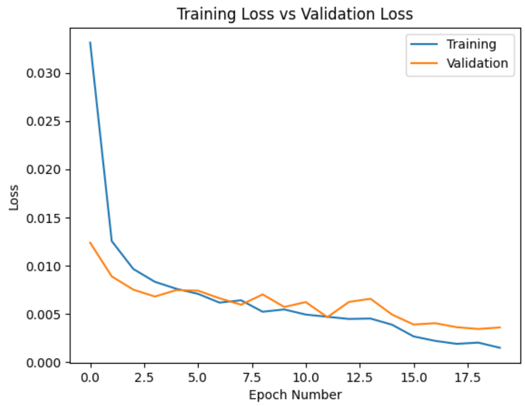

Figure 3:Resnet18 Loss Functions

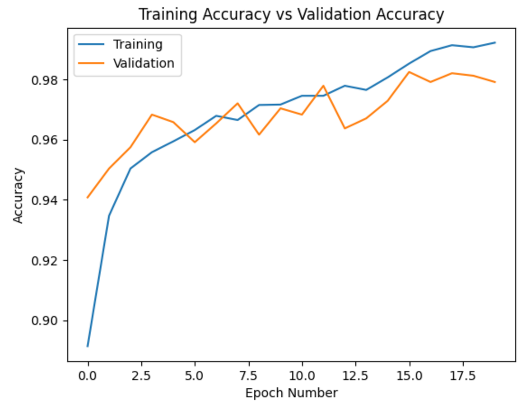

Figure 4:Resnet18 Accuracies

6. Conclusion & Future Work

Past 9 months have shown that measures should be taken to reduce the spread of the COVID-19 pandemic and masks are one of the most effective tools that can be used in public spaces. In this work, we have modeled a real-time face mask classifier using seven different CNN architectures. To train, validate and test the model, we used a combination of three datasets which contain 3232 correct worn masks, 1123 incorrectly worn mask, and 1717 without mask. We have used OpenCV to detect faces and PyTorch to classify these faces into aforementioned three categories. The models were tested with images and real-time video stream from a laptop's webcam.

This real-time face mask classifier can be used in public areas where people are crowded together, such as shopping malls, airports, and workplaces. It can also be expanded with an alarm system to sound and alert when there is a person not wearing their mask properly. By monitoring the public, detecting and preventing those who do not comply with health rules will prevent the spread of the virus.

References

[1] W. H. Organization et al., “Weekly Operational Update on COVID-19” January 11th, 2021.

[2] Y. Liu, A. A. Gayle, A. Wilder-Smith, and J. Rocklöv, “The reproductive number of covid-19 is higher compared to sars coronavirus,” Journal of travel medicine, 2020.

[3] A. Nieto-Rodríguez, Manuel Mucientes, Victor Brea, “System for Medical Mask Detection in the Operating Room Through Facial Attributes” Iberian Conference on Pattern Recognition and Image Analysis, 2015

[4] Mingjie Jiang, Xinqi Fan, Hong Yan, “Retina Face Mask: A Face Mask Detector”, Iberian Conference on Pattern Recognition and Image Analysis, June 2016

[5] Qin, B.; Li, D. Identifying Facemask-Wearing Condition Using Image Super-Resolution with Classification Network to Prevent COVID-19. *Sensors* **2020**

[6] Paul Viola, Michael Jones “Rapid Object Detection using a Boosted Cascade of Simple Features”, IEEE Computer Society Conference on Computer Vision, 2001

[7] Navneet Dalal, Bill Triggs, “Histograms of Oriented Gradients for Human Detection” IEEE Computer Society Conference on Computer Vision, 2005

[8] J. Deng,W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” IEEE conference on computer vision and pattern recognition, 2009

[9] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” IEEE conference on computer vision and pattern recognition, 2016

[10] "Face Mask Detection". kaggle.com. https://www.kaggle.com/andrewmvd/face-mask-detection, 2021

[11] Adnane Cabani,Karim Hammoudi, Halim Benhabiles, and Mahmoud Melkemi. "MaskedFace-Net – A Dataset of Correctly/Incorrectly Masked Face Images in the Context of COVID-19". Smart Health, 2020

[12] Karras, Tero, Samuli Laine, and Timo Aila, “A Style-Based Generator Architecture for Generative Adversarial Networks”. arXiv:1812.04948 [cs, stat], http://arxiv.org/abs/1812.04948, 2019.