Covid-19 Mask Protocol Violation Detection Using Deep Learning and Computer Vision

Mitali Deshmukh[1], Gauri Deshmukh[2], Pradnya Pawar[3], Priyanka Deore[4]

[1-4] UG Students, Dept. of Computer Engineering K K Wagh Institute of Engineering Education and Research, Maharashtra, India

---***---

Abstract - *The Covid-19 pandemic has spread far and wide, infecting millions of people all over the world. Wearing a mask is one of the most basic safety measures for protection against infection. Several government authorities have made mask usage compulsory in public places. For efficient implementation of mask usage policy there is a need of a system capable of detecting mask protocol violations in real-time. To solve this problem, we present a system that uses Haar Cascades for real-time face detection from a live video stream and a light weight MobileNetV2 architecture for identifying mask protocol violations. The system was able to detect mask wearing, non-mask wearing and faces wearing masks improperly with 99 percent accuracy. This system can help in tracking safety violations, promoting the use of facemasks, and ensuring a safe working environment.*

Key Words**: Covid-19, Mask Detection, Deep Learning, Computer Vision, MobileNetV2, Haar Cascades**

1.INTRODUCTION

The end of 2019 witnessed the outbreak of Coronavirus Disease, which has continued to be the reason for hardship for several of lives and businesses even in 2021. Studies have proved that wearing a mask significantly reduces the danger of viral transmission [1]. The primary role of face masks is to prevent the transmission of virus via respiratory droplets, which may simply enter through the mouth and nose to infect new hosts. Therefore, maintaining social distance and wearing masks is necessary in order to gain control of the situation in the pandemic.

Many governments all over the world have made use of mask in public places mandatory. Currently it is the responsibility of a guard posted at the entrance of buildings or police in public places to check whether people are wearing a face mask or not. Apart from being labor intensive this approach puts the guards at risk of viral exposure from individuals not wearing face masks. Technology holds the key here.

This paper introduces a face mask detection system capable of differentiating between users wearing a mask, not wearing a mask and wearing mask improperly in real-time video stream. This system can be used by authorities to track adherence to mask usage guidelines and take necessary actions. This system can be ported to surveillance cameras to identify individuals not wearing masks. Wearing a face mask is also mandatory in some workplaces such as hospital operating rooms, pharmaceutical industry and food industry. Proposed system can also be deployed in aforementioned workplaces to ensure maintenance of hygiene standards.

2. RELATED WORKS

There are several different approaches adopted by researchers in the field of computer vision for mask detection. Sammy V. Militante Et al. [2] proposed the use of a convolutional neural network to design the face mask classifier. A VGG-16 CNN model was trained and then ported to Raspberry Pi to test it's performance in an embedded system. Loey Et al. [3] introduced a masked face detection model that is based on deep transfer learning and classical machine learning classifiers. Their study presented an integration between deep transfer learning and classical machine learning algorithms. They combined ResNet-50 with Support Vector Machine to build a hybrid deep transfer learning architecture for masked face classification.

A. Nieto-Rodríguez Et al. [4] proposed a system to identify breach protocols in operating rooms. The system executes 4 steps - First MoG (Mixture of Gaussians) technique was used for background subtraction. The second step involved classification and tacking. Face classification is triggered in the foreground and in the regions estimated by the tracking block. The third step involves association of new detections, either mask or maskless faces, with the estimated position of the trackers. In step 4 either new trackers are created when new face detections do not match existing trackers, or alternatively, the position of the trackers are updated to those of the existing faces. Viola Jones classifier was used for face and mask detection. The system sorted faces out in the gray-scale space before the second stage, where color filters were used. Finally, the system made decision about the final class based on the highest number of
detections across the scale space.

Rafiuzzaman Bhuiyan Et al. [5] proposed a deep learning based assistive system to classify COVID-19 Face Mask for Human safety with YOLOv3. Image was passed as input to the YOLOv3 model. The object detector went through the image and divided it into a grid to analyze it's features and locate the target object. Amit Chavda Et al. [6] proposed a multi-stage CNN architecture for face mask detection. Their system is composed of two stages - the first stage detects user's faces while second stage uses a lightweight image classifier to classify the faces detected as 'Mask' or 'No Mask' and draws bounding boxes around them with class names. The system was extended to videos as well by using an object tracking algorithm. RetinaFace model was used for face detection while NASNet was used for classification of face into mask wearing or non-mask wearing category.

3. METHODOLOGY

3.1 Data Gathering and Pre-Processing

RMFD dataset [7] and MaskedFaceNet [8] were used for gathering images depicting different mask usage conditions. 'with mask' data included around 2000 images of people correctly wearing different types of masks. 'without mask' and 'incorrectly worn mask' data included around 2000 images each of people not wearing mask and wearing mask incorrectly. To ensure a fixed shape input all images in the dataset were resized to 224*224 pixels. As MobileNetV2 requires inputs from interval (-1,1), data was normalized accordingly.

3.2 Model Building and Training

MobilenetV2 is a light weight architecture that can be applied to embedded devices with limited computational capacity such as smart surveillance cameras [9]. MobileNetV2 model was loaded with pre-trained ImageNet weights excluding the head of the network. New head was constructed and appended in place of the old head to customise the model for the application. Around 70 percent of the dataset was used for training the model. During training of the model data augmentation was used to introduce diversity in the training data by zooming, rotating and pivoting image parameters. Model was trained for 20 epochs with a batch size of 32.

3.3 System Overview

Fig-1 depicts working of the system. The system is divided into two phases - face detection and classification. Input to the system is a live video feed. First, video frame undergoes grey scale transform and Haar Cascade is used to detect presence of a face in a video frame. If a frame is found to contain a face Haar Cascade returns co-ordinates corresponding to the face region. This face region acts as the input to our trained MobileNetV2 model which in turn predicts the appropriate class label. Bounding box with appropriate annotation is displayed around the face.

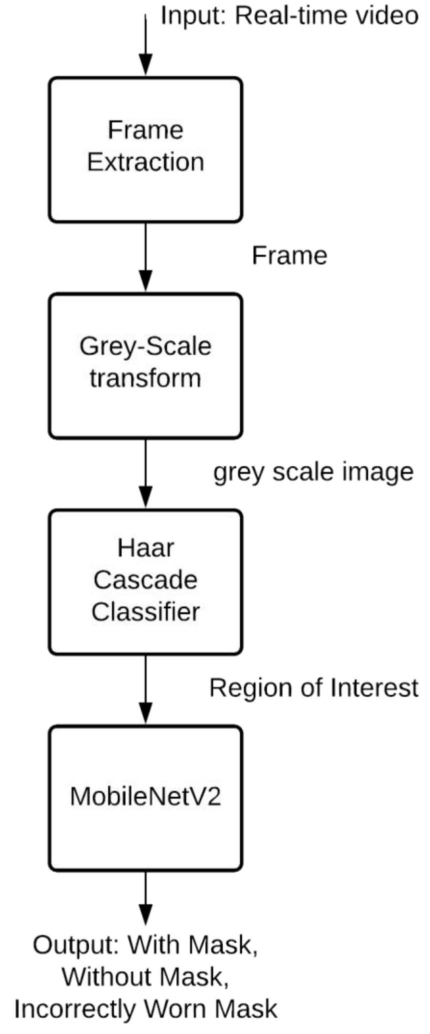

Fig-1: System Overview

2. RESULTS

Acceptable validation accuracy was achieved during the testing of the model. Several experiments were conducted with batch size fixed to 32 and 20 iterations for epochs for observing the performance. Test results by visualization of accuracy and loss metrics are illustrated in Chart-1. Accuracy of the system was measured in terms of precision and recall. Table-1 displays the evaluation metrics of the model. System was tested for detecting various mask wearing conditions. Fig-2, Fig-3, Fig-4 display different test results on the performance of the model in detecting mask wearing conditions. Instances of person not wearing mask, wearing mask and wearing mask improperly are all successfully detected.

Table -1: Accuracy Report

	Precision	Recall	F-1 Score
With Mask	1.00	0.99	0.99
Without Mask	0.99	0.99	0.99
Incorrect Mask	0.99	1.00	1.00
Accuracy			0.99
Macro Average	0.99	0.99	0.99
Weighted Average	0.99	0.99	0.99

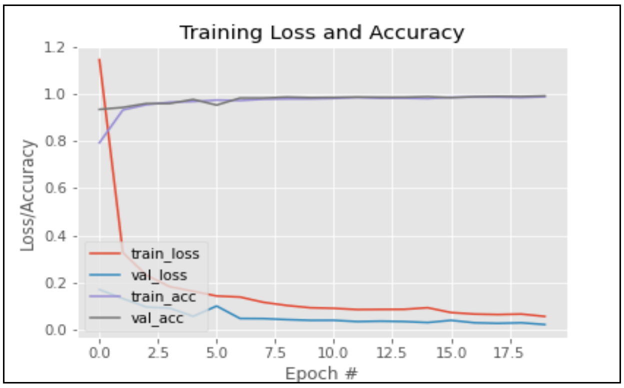

Chart -1: Training Loss Vs Accuracy

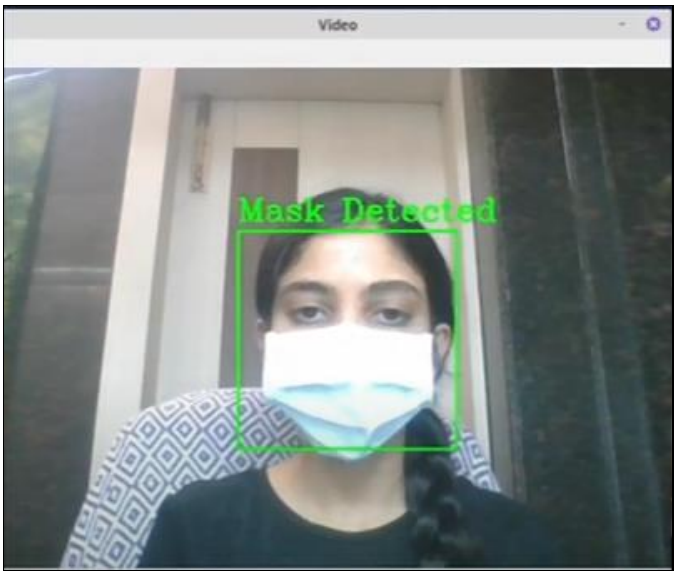

Fig -2: Test result for "With Mask"

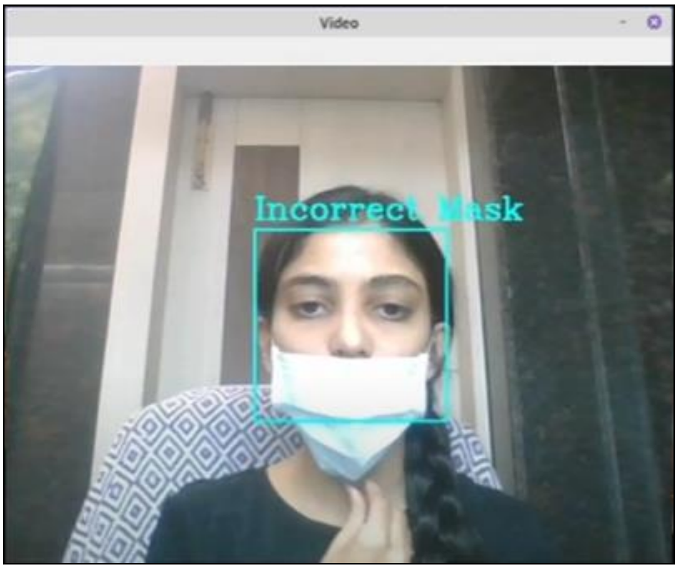

Fig -3: Test result for "Incorrect Mask"

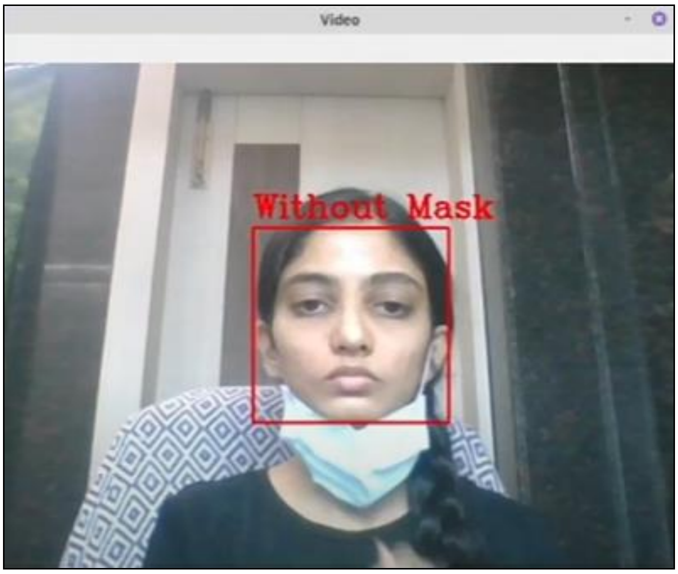

Fig -4: Test result for "Without Mask"

3. CONCLUSIONS

Manually monitoring mask usage requires human effort and poses threat of infection from unmasked people. The system presented makes use of deep learning based to distinguish between users following mask protocol and users violating mask protocol. The system employs Haar Cascade for face detection and a light weight MobileNetV2 architecture for distinguishing between proper and improper mask usage. It is capable of recognizing mask worn under nose and mask worn on chin as a violation of mask protocol. The model was able to detect mask protocol violation correctly. As the system uses a MobileNetV2 architecture it is also suitable for deployment on embedded devices with limited computational capacity. Using such a system for monitoring mask usage can help to track safety violations, promote the use of facemasks, and ensure a safe working environment.

REFERENCES

[1] Wang, Yu, et al. "Reduction of secondary transmission of SARS-CoV-2 in households by face mask use, disinfection and social distancing: a cohort study in Beijing, China." BMJ global health 5.5 (2020): e002794.M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[2] Militante, Sammy V., and Nanette V. Dionisio. "Real-Time Facemask Recognition with Alarm System using Deep Learning." 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC). IEEE, 2020.

[3] Loey, Mohamed, et al. "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic." Measurement 167 (2021): 108288.

[4] Nieto-Rodríguez, Adrian, Manuel Mucientes, and Victor M. Brea. "Mask and maskless face classification system to detect breach protocols in the operating room." Proceedings of the 9th International Conference on Distributed Smart Cameras. 2015.

[5] Bhuiyan, Md Rafiuzzaman, Sharun Akter Khushbu, and Md Sanzidul Islam. "A Deep Learning Based Assistive System to Classify COVID-19 Face Mask for Human Safety with YOLOv3." 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2020.

[6] Chavda, Amit, et al. "Multi-stage cnn architecture for face mask detection." 2021 6th International Conference for Convergence in Technology (I2CT). IEEE, 2021.

[7] Wang, Zhongyuan, et al. "Masked face recognition dataset and application." arXiv preprint arXiv:2003.09093 (2020).

[8] Cabani, Adnane, et al. "MaskedFace-Net–A dataset of correctly/incorrectly masked face images in the context of COVID-19." Smart Health 19 (2021): 100144.

[9] Sandler, Mark, et al. "Mobilenetv2: Inverted residuals and linear bottlenecks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.