Using ML for Facial Mask Detection

Harsh Saraiya [a], Prajakta Mhaske [b], Saloni Rane [c], Kiran Deshpande [d], Kaushiki Upadhyaya [e], Nahid Shaikh [f]

[a] A.P Shah Institute of Technology, Thane, India, harshsaraiya@apsit.edu.in

[b] A.P Shah Institute of Technology, Thane, India, prajaktamhaske@apsit.edu.in

[c] A.P Shah Institute of Technology, Thane, India, salonirane@apsit.edu.in

[d] A.P Shah Institute of Technology, Thane, India, kbdeshpande@apsit.edu.in

[e] A.P Shah Institute of Technology, Thane, India, ksupadhyaya@apsit.edu.in

[f] A.P Shah Institute of Technology, Thane, India, nashaikh@apsit.edu.in

Abstract

The COVID-19 pandemic has severely impacted our daily lives and continues to spread over the globe. Wearing a protective face mask to prevent the spread of infection has become the accepted norm. Since people have become a bit careless towards wearing facial masks outdoors, this system detects the human individual through videos not wearing a mask and informs the authorities about it. This paper proposes a simplified approach to accomplishing this goal utilizing TensorFlow, Keras, and OpenCV, as well as some fundamental Machine Learning packages. We have used the Res10 SSD Caffe model for face detection and Image classification is done through MobileNetV2 architecture for the implementation of our system. By attaining high precision, real-time detection, and classification, the suggested methodology established its effectiveness in identifying facial masks.

Keywords

COVID-19; Face Detection; Deep Learning; convolution neural network; facial masks; unmasked; video footage; Image pre-processing

1. Introduction

Due to COVID-19 Pandemic, many individuals have died as a result of this illness, and thousands more are affected every day. The 2019 Coronavirus Infection (COVID-19) has infected more than 20 million people worldwide, resulting in over 0.7 million deaths, according to the official Situation Report of the World Health Organization - 205.

Coronavirus has infected 213 countries, including all industrialized countries such as the United States, the United Kingdom, Russia, China, Japan, Italy, and many others. People's carelessness and lack of consciousness were the main causes of the infected virus. Every day, people go to work or live in other apartments without wearing a mask. Surveillance is extremely difficult and time-consuming all of the time. This study is primarily aimed at resolving this issue and assisting people in self defense. It's critical, especially in COVID-19, to safeguard ourselves from other individuals. To prevent the transmission of Coronavirus, it is strongly advised that the general public wear face masks. Furthermore, with the lifting of the COVID19 quarantine, government and public health agencies are suggesting face masks as vital precautions to take when going out in public. To make the usage of facemasks mandatory, some measures must be devised that require people to put on a mask before entering public spaces. Face mask detection is the process of determining whether or not someone is wearing a mask. In actuality, the issue is reverse engineering of face detection, in which the face is detected using various machine learning algorithms for security, authentication, and surveillance purposes. In the domain of Computer Vision and Pattern Recognition, face detection is crucial. Deep convolutional neural networks (CNN)-based face recognition algorithms have become increasingly popular in recent years as a way to improve detection performance. Furthermore, detecting faces with or without a mask in public is difficult due to the little dataset available for detecting masks on human faces, making the model difficult to train. As a result, for a similar face identification task, the notion of transfer learning is used to transfer learned kernels from networks trained on a big dataset. Detecting mask faces can be done in a variety of ways, including from various perspectives. The goal of this research is to create a system for detecting mask faces using CNN's classifiers. Working with CNN enhances the accuracy of identifying the mask face in a certain region, and it's much more responsive when the face enters the webcam's field of vision, resulting in speedier detection. In some ways, the community is safe while more flu stops are dispersed in the air and create entry barriers into the human body.

2. Literature Review

In this section, we take a look at some of the similar efforts that have been done in this field. Despite the fact that face detection research has been ongoing for decades and has yielded significant results, only a few approaches and algorithms are explicitly developed for face mask recognition. The proposed framework in [1] uses the MTCNN facial recognition model to identify the faces in the video frame and their relevant facial landmarks. These images and cues are then examined by a neoteric classifier that detects masked regions using the MobileNetV2 architecture as an object detector. The purpose of Literature [2] is to give a general review of many methods and algorithms for human recognition using a face mask. The Haar cascade, Adaboost, VGG-16 CNN Model, and other techniques are described. A study of various strategies is conducted to evaluate which technique is feasible. The author of the paper [3] developed a deep learning architecture based on a dataset of images of people wearing and not wearing masks obtained from various sources. For previously unreported test data, the trained architecture distinguished facial mask wearers and non-mask wearers with 98.7% accuracy. Two distinct approaches to detect real-time masking and unmasking of faces are proposed in this study [4]. An object detection model is used in the first technique to discover and distinguish between masked and uncovered faces. A YOLO face detector recognizes faces, whether masked or not, and then categories them into masked and unmasked groups using a dataset in the second method. With a minimal number of training samples, Mask R-CNN, an established object detector, has been trained for face detection as well as instance segmentation and object bounding box detection in this study [5]. The results show that the trained Mask R-CNN outperforms the baseline detector in terms of detection rates. On the basis of the SSD algorithm, literature [7] proposes a method for detecting face masks. To increase the model's ability to exhibit essential features, SSDMask incorporates a channel attention technique. At the same time, the loss function is optimized and the information from different feature levels is fully exploited. The effectiveness of three algorithms: KNN, SVM and MobileNet was compared in the literature [6] to discover the best algorithm for detecting someone is wearing a mask in a real-time context. MobileNet has the maximum accuracy for both input photos and input videos from a camera, according to the results.

Across the full latency spectrum, the MobileNetV2 models are faster for the same accuracy. MobileNetV2 is a powerful feature extractor for detecting and segmenting objects. For example, when paired with the recently released SSDLite, the new model is around 35 percent faster than MobileNetV1 while maintaining the same accuracy. As we've shown, MobileNetV2 is a powerful mobile-oriented model that may be utilized to solve a variety of visual identification problems.

3. Methodology

In this proposed system we have three phases, each with its own respective sub steps. First phase is training the face mask detector model. The second phase is Detection and Recognition. The third phase is Deployment on cloud. We have used the MobileNetV2 classifier model, a highly efficient architecture in order to identify a face mask.

3.1 Training the Face mask model

We have collected 1396 images of faces [9] out of which 1000 are faces with masks and 396 images are without masks. These images are then loaded with a target size of 224x224 and then preprocessed using Keras mobilenetv2 library. We have split the dataset for the training and testing phase. We trained the model on 80% of the data and tested it on 20% of the data. For Data Augmentation, we created a training picture generator. After some trial and error, we found that 1e-4 offered us the best outcomes and the quickest reduction in losses. With a batch size of 32, the model was trained for 20 epochs. A classification report was created during the training procedure and the trained model was preserved.

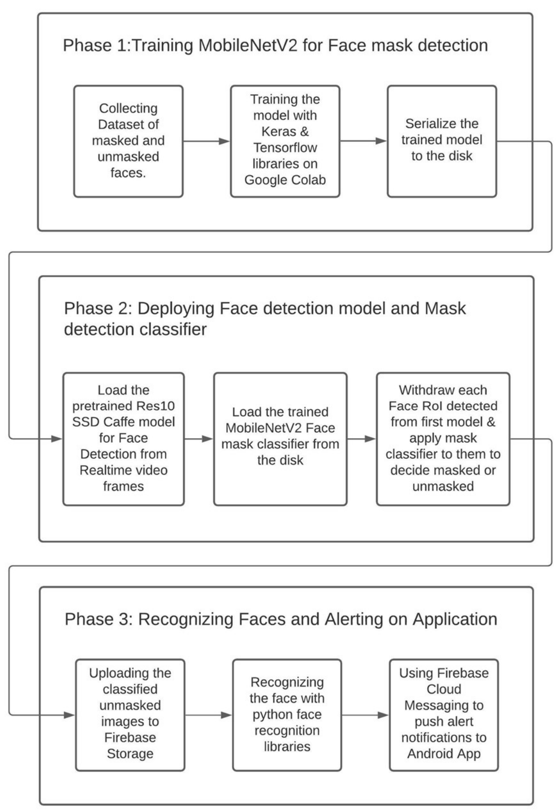

Figure 1. Block Diagram of Face Mask Detection System

3.2 Detection & Recognition

Face detection from the video frames is carried out by using a pre-trained Res10 SSD Caffe model with the OpenCV DNN module. Initially both the models are loaded from the disk and Video Stream from the required input is started using OpenCV and frames are read and resized. The Facial ROI is detected from the blob of the image and these are passed through the Caffe model. The faces and their corresponding location on the frame are then stored into a list. Then using the (startX, startY, endX, endY) coordinates of the Facial ROI, and converting it from BGR to RGB channel, ordering it, resizing them to 224x224 which was desired for the MobileNetV2 model and preprocessing the image using Keras libraries. Passing this list into the MobileNetV2 model will return the predictions made on the Facial images whether masked or unmasked. Then looping over the coordinates and predictions list, a bounding box is appended over the real time frames of the Video Stream. If prediction = "Mask" bounding box color is assigned to Green and if prediction = "Unmask" bounding box color is assigned to Red.

Figure 2. Image of masked and unmasked people from real time video footage

From the above classification, the unmasked faces will be used for facial recognition. We used the Python face recognition 1.3.0 package for face recognition, which is based on dlib's state-of-the-art face recognition developed using deep learning [8]. This model will be trained on a known faces dataset of a company, a hospital staff, college campus and so on. The unmasked faces will be classified by this deep learning library and the authority will get to know the details of the unmasked person viz. Name, Unique Identification number, Department and so on.

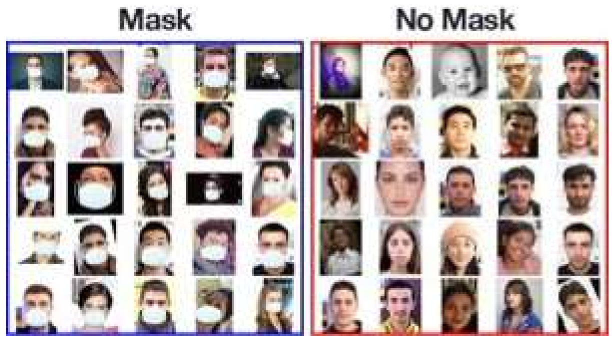

Figure 3. Classification into masked & unmasked category

3.3 Deployment on Cloud For cloud storage

Firebase is a useful platform developed by Google which we will implement for application development. It is the greatest mobile backend as a service since it allows mobile apps to connect to cloud storage and APIs on the backend. In the database, Firebase maintains all data in real time. As a result, data transfer to and from the database is simple and speedy. Google integrated Firebase Cloud Storage with the Firebase SDK, letting users save the files on the cloud in a matter of seconds. We will be using these Firebase SDKs for storing video/images of the people without masks. The identical files on the server will be accessed via Google Cloud Storage. Every file will be saved in a Google Cloud Storage Bucket and will be accessible from both Firebase and Google Cloud. This allows us to use Firebase SDKs to retrieve and upload files from mobile clients, as well as do serverside computing using Google Cloud. The Firebase Real time Database (Firebase RTDB) is a NoSQL database management system maintained by Firebase over the cloud. The images of unmasked people which were collected earlier for the mask detection will be uploaded to the cloud storage. FCM (Firebase Cloud Messaging) connects the server and clients such that receiving and sending alert notifications/messages is simple and dependable. We'll use FCM to alert a relevant authority using the Firebase admin SDK. A notification will be sent to the application using the Firebase admin SDK. A custom notification will be sent to the concerned authority recognizing the details of the person without mask. When a person without a mask is identified, the application will be notified. The application will be developed in Flutter as it provides better user experience, robust performance and enormous time and effort saving.

4. Conclusion

As the technology is developing with emerging trends, therefore the novel face mask detector using deep learning approach for detecting face masked people from real-time videos in public places is proposed using Convolution Neural Network & MobileNetV2 classifier to curtail the community spread of Coronavirus. To select this base model, we evaluated metrics such as accuracy, precision, recovery and selected the MobileNetV2 architecture with the best performance with 96% accuracy and 99% recall. A face mask detection architecture is included in the system, which employs a deep learning algorithm to recognize the mask on the user's face. The model is trained on a real dataset of masked and unmasked facial photos. This face mask detector will act as a valuable tool that can be used in many areas such as shopping malls, airports and other high-traffic locations to keep an eye on the public and prevent the spread of Virus by checking who is and who is not following the basic rules and informing the concerned authority. This collaborative technique not only aids in reaching high accuracy, but it also significantly improves detection speed.

5. Acknowledgment

We would like to express our gratitude to Google Collaboratory for providing access to the computing resources employed in this investigation. We also thank Prof. Kiran Deshpande, Prof. Kaushiki Upadhyaya and Prof. Nahid Shaikh for their expertise and assistance throughout all aspects of our study and research of our project.

6. References

[1] Joshi, A. S., Joshi, S. S., Kanahasabai, G., Kapil, R., & Gupta, S. (2020, September). Deep Learning Framework to Detect Face Masks from Video Footage. *2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)*. https://doi.org/10.1109/cicn49253.2020.9242625

[2] Kalas, M. (2014). Real Time Face Detection And Tracking Using Opencv. *International Journal of Soft Computing And Artificial Intelligence*, *2*(1), 41–44.

[3] Rahman, M. M., Manik, M. M. H., Islam, M. M., Mahmud, S., & Kim, J. H. (2020, September). An Automated System to Limit COVID-19 Using Facial Mask Detection in Smart City Network. *2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*. https://doi.org/10.1109/iemtronics51293.2020.9216386

[4] Abbasi, S., Abdi, H., & Ahmadi, A. (2021, March). A Face-Mask Detection Approach based on YOLO Applied for a New Collected Dataset. *2021 26th International Computer Conference, Computer Society of Iran (CSICC)*. https://doi.org/10.1109/csicc52343.2021.9420599

[5] Cakiroglu, O., Ozer, C., & Gunsel, B. (2019). Design of a Deep Face Detector by Mask R-CNN. *2019 27th Signal Processing and Communications Applications Conference (SIU)*, 1–4. https://doi.org/10.1109/siu.2019.8806447

[6] Vijitkunsawat, W., & Chantngarm, P. (2020, October). Study of the Performance of Machine Learning Algorithms for Face Mask Detection. *2020 - 5th International Conference on Information Technology (InCIT)*. https://doi.org/10.1109/incit50588.2020.9310963

[7] Xu, M., Wang, H., Yang, S., & Li, R. (2020, November). Mask wearing detection method based on SSD-Mask algorithm. *2020 International Conference on Computer Science and Management Technology (ICCSMT)*. https://doi.org/10.1109/iccsmt51754.2020.00034

[8] *face-recognition*. (2020, February 20). PyPI. https://pypi.org/project/face-recognition/

[9] Bhandary, P. (n.d.). *GitHub - prajnasb/observations*. GitHub. https://github.com/prajnasb/observations