

Face Mask Detection Using Machine Learning

Dhanasekaran S, Pradeep Reddy K, Reddappa Reddy K and Bhargav Reddy K*

Department of CSE, Kalasalingam Academy of Research and Education Madurai, Tamilnadu, India

Received Date: May 06, 2022 Accepted Date: June 07, 2022 Published Date: June 09, 2022

*Corresponding author: Bhargav Reddy K, Department of CSE, Kalasalingam Academy of Research and Education Madurai, Tamilnadu, India, 9381060408, Email: 9918004044@klu.ac.in

Citation: Dhanasekaran S, Pradeep Reddy K, Reddappa Reddy K and Bhargav Reddy K (2022) Face Mask Detection Using Machine Learning. J Comput Sci Software Dev 2: 1-7

Abstract

Object discovery, which is intended to automatically mark links to objects of interest in photos or videos, is an extension of the image classification. In recent years, it has been widely used in smart traffic management, intelligent surveillance systems, military acquisition, and placing surgical instruments in surgical navigation, etc. COVID-19, the outbreak of a new coronavirus at the end of 2019, poses a serious threat to public health. Many countries require everyone to wear a mask in front of people to prevent the spread of coronavirus. To effectively prevent the spread of the corona virus, we are introducing a detection-based device (SSD), which focuses on getting a face mask in real time in the store. We are making contributions to the following three areas: 1) introducing a lightweight spinal network to deliver feature, SSD-based and flexible spatial flexibility, aimed at improving access speed and meeting real-time acquisition requirements; 2) proposing a Feature Development Module (FEM) to reinforce the in-depth features learned in CNN models, which are intended to improve the presentation of small element features; 3) create a COVID-19-Mask, a large database of whether consumers wear a mask, by collecting photos from two supermarkets. Test results show the high accuracy of the real-time performance of the proposed algorithm.

Keywords: COVID19, SSD, Face Mask Detection, FEM (Feature Enhancement Module), CNN models.

Introduction

Rapid advancements in the fields of Science and Technology have led us to a stage where we are capable of achieving feats that seemed improbable a few decades ago. Technologies in fields like Machine Learning and Artificial Intelligence have made our lives easier and provide solutions to several complex problems in various areas. Modern Computer Vision algorithms are approaching human-level performance in visual perception tasks. From image classification to video analytics, Computer Vision has proven to be a revolutionary aspect of modern technology. In a world battling against the Novel Coronavirus Disease (COVID-19) pandemic, technology has been a lifesaver. With the aid of technology, 'work from home' has substituted our normal work routines and has become a part of our daily lives. However, for some sectors, it is impossible to adapt to this new norm. As the pandemic slowly settles and such sectors become eager to resume in-person work, individuals are still skeptical of getting back to the office. 65% of employees are now anxious about returning to the office (Woods, 2020). Multiple studies have shown that the use of face masks reduces the risk of viral transmission as well as provides a sense of protection (Howard et al., 2020; Verma et al., 2020).

Machine-learning technology powers many aspects of modern society: from web searches to content filtering on social networks to recommendations on e-commerce websites, and it is increasingly present in consumer products such as cameras and smart phones. Machine-learning systems are used to identify objects in images, transcribe speech into text, match news items, posts or products with users' interests, and select relevant results of search. Increasingly, these applications make use of a class of techniques called deep learning. Conventional machine-learning techniques were limited in their ability to process natural data in their raw form. For decades, constructing a pattern-recognition or machine-learning system required careful engineering and considerable domain expertise to design a feature extractor that transformed the raw data (such as the pixel values of an image) into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input

However, it is infeasible to manually enforce such a policy on large premises and track any violations. Computer Vision provides a better alternative to this. Using a combination of image classification, object detection, object tracking, and

video analysis, we developed a robust system that can detect the presence and absence of face masks in images as well as videos. In this paper, we propose a two-stage CNN architecture, where the first stage detects human faces, while the second stage uses a lightweight image classifier to classify the faces detected in the first stage as either 'Mask' or 'No Mask' faces and draws bounding boxes around them along with the detected class name.

This algorithm was further extended to videos as well. The detected faces are then tracked between frames using an object tracking algorithm, which makes the detections robust to the noise due to motion blur. This system can then be integrated with an image or video capturing device like a CCTV camera, to track safety violations, promote the use of face masks, and ensure a safe working environment. Declared by the WHO that a potential speech by maintaining distance and wearing a mask is necessary. Wearing a mask captured by the image detection where the machine can cover and translate only the mouth portion of the face part. Computer vision is a following section of Deep learning particularly an area of convolution neural network (CNN). Added with one main thing is CNN supports very high configuration Graphic Processing Units (GPU) thus as real time image or video extraction of visualisation is a bitter task. As we require people mask having or not which call a surveillance system there is a need for powerful validation such as video stream analysis that is fulfilled by advanced CNN.

Literature Review

Traditional Discovery

The problem of getting too many covered and uncovered faces in photos can be solved with a standard object detection model. The acquisition process mainly involves locating objects in images and separating them (in the case of multiple objects). Traditional algorithms such as Haar Cascade (Viola and Jones, 2001) and HOG (Dalal and Triggs, 2005) have proved to be effective in such tasks, but these algorithms are largely based on Feature Engineering. During In-Depth Learning, it is possible to train Neural Networks that perform much better than these algorithms, and do not require any additional Feature Engineering.

Convolutional Neural Networks

Convolutional Neural Networks (CNNs) (LeCun et al., 1998) is an important component of modern Computer Vision functions such as pattern detection, image detection, pattern recognition functions, etc. CNN uses a bunch of convolution to match the images or feature first. maps to extract high-quality features, thus resulting in a very powerful tool for Computer Vision functions.

Modern Acquisition Algorithms

CNN-based object acquisition algorithms can be divided into 2 categories: Multi-StageDetector and Single Stage Detectors.

Multi-Stage Detectors

In multi-stage detector, the detection process is divided into several steps. A two-stage detector such as the RCNN (Girshick et al., 2014) begins to measure and propose a set of interested regions using selected searches. CNN feature veggies are then released in each region independently. Many algorithms based on Regional Proposal Network such as Fast RCNN (Girshick, 2015) and Faster RCNN (Ren et al., 2015) have achieved higher accuracy and better results than most single-stage detectors.

Single Stage Examiners

A single-stage detector makes a single step detection, directly over the dense samples of potential locations. These algorithms override the regional proposal category used in multi-category acquisition receivers and thus are generally considered faster, at the expense of some loss of accuracy. One of the most popular single-stage algorithms, You Only Look Once (YOLO) (Redmon et al., 2016), was launched in 2015 and achieved close

to real-time performance. Single Shot Detector (SSD) (Liu et al., 2016) is another popular algorithm used for object detection, which provides excellent results. RetinaNet (Lin et al., 2017b), one of the leading finders, is based on Feature Pyramid Networks (Lin et al., 2017a), and uses targeted losses.

Face Mask Discovery

As the world began to use security measures against Coronavirus, several plans emerged for the discovery of the Mask Mask. (Ejaz et al., 2019) performed facial expressions with a masked face and uncovered face using Principal Component Analysis (PCA). However, the accuracy of detection decreases to less than 70% when the known surface is covered. (Qin and Li, 2020) introduced a method for identifying conditions for wearing a face mask. Divide the conditions of wearing a mask into three categories: wearing the right face mask, wearing the wrong face mask, and not wearing a face mask. Their system takes a picture, identifies it and implants a face, and then uses SRCNet (Dong et al., 2016) to create a high-resolution image that is edited and edited. The work of (Nieto-Rodríguez et al., 2015) introduced a method that detects the presence or absence of a medical mask. The main purpose of this approach was to raise awareness among only medical professionals who would not wear a surgical mask, by minimizing false facial expressions as much as possible, without losing any medical mask discovery.

(Loey et al., 2021) has proposed a two-component model. The first part makes use of ResNet50 (He et al., 2016) to extract the feature. The next section is a face mask separator, based on a combination of the old Machine Learning algorithms. The authors reviewed their program and estimated that In-Depth Learning methods would achieve better results as the structure, comparison, and selection of the best model among a set of classic machine learning models was a time-consuming process.

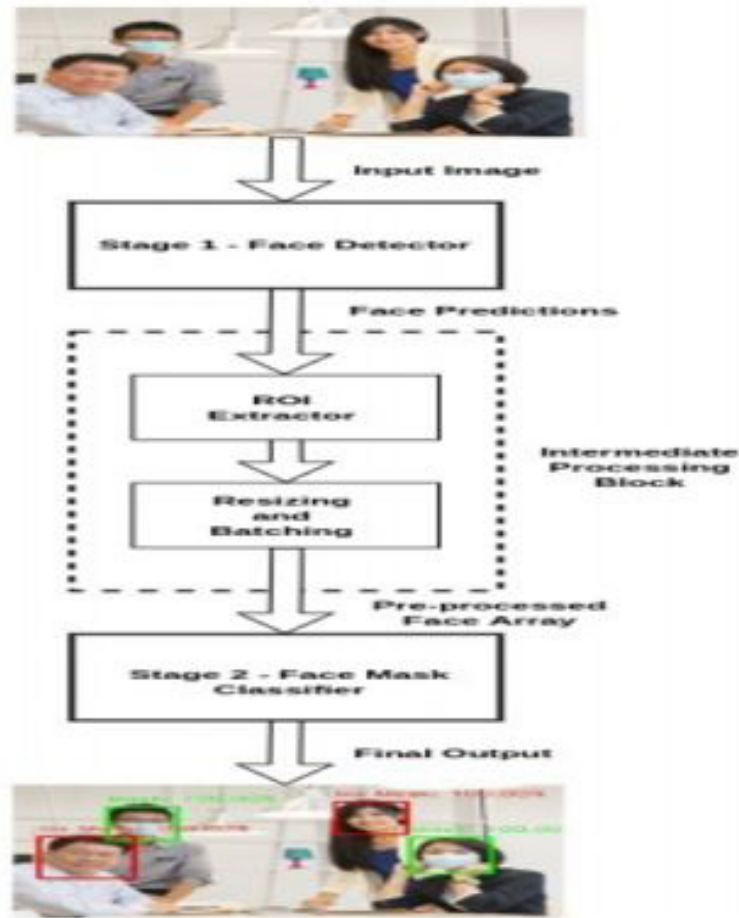


Fig 1: Flow diagram of proposed system.

In the current system, a significant number of managers around the world are also interested in a face recognition system to ensure open spaces, for example, leaves, air terminals, transit stations, and train stations, and so on. Facial expressions are one of the most common health issues that have been investigated around the world. Amazing progress has been made against the invention of facial recognition in all recent years. The small amount of face highlighting on the covered face poses more challenges than other standard facial expressions.

To alleviate the lack of large databases, we have upgraded our Hidden Face Website (MFD). Another level of obstacles and directions are accessible to this knowledge base. We've added more confused facial images to our MFD Database. In this way, our Hidden Face Data Set has 45 topics that cover a different concealment, containing clear and complex foundations.

Methodology

Figure 1 represents the structure of our proposed system (image captured from database by (Larxel, 2020)). It consists of two main sections. The first phase of our architecture includes the Face Detector, which incorporates multiple faces in images of various sizes and detects faces even in isolated conditions. The detected faces (regions of interest) removed from this category are then merged and transferred to the second phase of our architecture, namely the CNN based Face Mask Classifier. The results from the second stage are coded and the last one is a properly acquired image and classified as a covered or unveiled face.

Stage 1 - Face Finder

The face detector serves as the first phase of our system. Raw RGB image is transmitted as input to this section. The face detector removes and removes all the faces found in the image with their connecting box links. The process of getting a face accurately is very important to our makeup. Training a highly accurate face detector requires a lot of labeled data, time, and calculation tools. For these reasons, we have selected a pre-trained database on a large database to make it easier and more stable to access. Three different pre-trained models have been tested in this category:

Dlib (Sharma et al., 2016) - The Dlib Deep Learning face detector offers much better performance than its predecessor, the Dlib HOG-based face detector.

MTCNN (Zhang, K. et al, 2016) - Uses CNN's three-phase cascade architecture to locate and locate face and face key points.

RetinaFace (Deng et al., 2020) - It is a single-stage pixel-wise design that uses a multi-tasking learning strategy to simultaneously predict face box, face points, and face key points. The model detection process is used in this section, as it requires seeing a person's face that may be covered with masks.

Experimental Results and Discussion

Training Dataset

Three face mask separation models are trained in our database. Images of a set of blank and uncovered facial data were collected from a database of images available on a public domain, as well as specific data published online. Hidden images are found in the Real Face Recognition database.

Face Detector

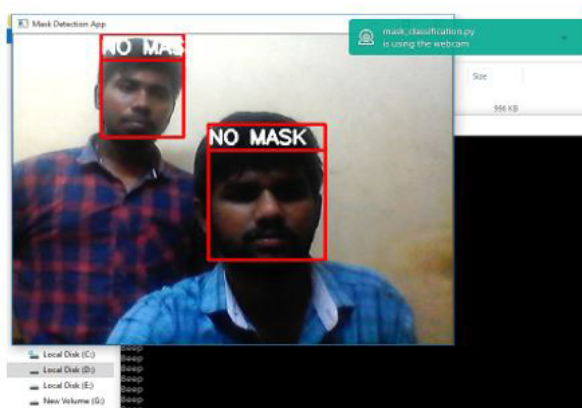
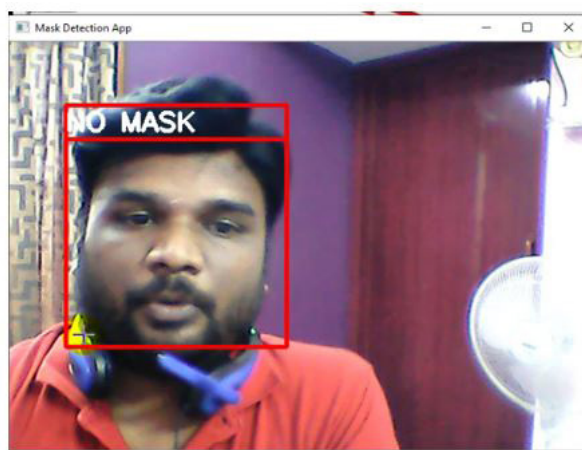
The face detector serves as the first phase of our system. Raw RGB image is transmitted as input to this section. The face detector removes and removes all the faces found in the image with their connecting box links. The process of getting a face accurately is very important to our makeup. Training a highly accurate face detector requires a lot of labeled data, time, and calculation tools. For these reasons, we have selected a pre-trained database on a large database to make it easier and more stable to access.

Face Mask Classifier

This section takes the ROI processed in the Intermediate Processing Block and classifies it as a Mask or No Mask. The CNN-based section of this section was trained, based on three different image classification models. CNN separators are trained to classify images as hidden or open. These models have simple structures that provide high performance with low latency, suitable for video analysis. The output of this category is an image (or video frame) with a local face, which is classified as covered or unveiled.

Face Mask Detector

We tested three pre-trained models to get face Stage: Dlib DNN, Retina Face. The approximate time frame for each model is calculated, based on a set of hidden and uncovered images. It was noted that all three models show positive effects on images taken at a very short distance, with no more than two people in the image. MTCNN and Retina Face are better than Dlib and can see many faces in photos. Both are able to see covered or covered faces. MTCNN has a very high accuracy when it detects a face with a preview. Retina Face can detect side faces for good viewing as well.



Conclusion

This project is used to identify situations for wearing a face mask in a public place. Identifies and provides alerts that correct facial mask wear, improper face mask wear, and do not wear face mask. Using a combination of image classification, object detection, object tracking, and video analysis, we have developed a robust system that can detect the presence and absence of masks in photos and videos. Most people do not

wear masks in public places such as colleges, train stations, bus stops, etc. It encourages the use of this concept to identify and identify people who do not wear masks in public places. The FaceNet model optimized for hidden and uncovered images provides better precision of acknowledging the covered face. It is important for us to improve and extend our work to deal with the unusual coverage of the face mask. To be able to look after yourself, be safe, accurate and competent this helps more creativity and less waste.

References

1. Hui DS, I Azhar EI, Madani TA, et al. (2020) Ongoing 2019-nCoV epidemic threat of a new global coronavirus - Coronavirus outbreak of the latest 2019 novel in Wuhan, China. *International Journal of Infectious Diseases* 91: 264-266.
2. Liu Y, Sun P, Highsmith MR, et al. (2018) Comparison of the application of in-depth study techniques for bird awareness in aerial photography // 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC). IEEE 317-324.
3. Chen X, Kundu K, Zhu Y, et al. (2015) 3d Object Recognition Objects // *Advances in Neural Information Processing Systems* 424-432.
4. Liu W, Anguelov D, Erhan D, et al. (2016) Ssd: Single shot multibox detector // *European conference with computer perspective*. Springer, Cham 21-37.
5. Yang F, Choi W, Lin Y (2016) Use All Layers: A Fast and Accurate CNN Object Finder with Scale-Based Integration and Cascaded Rejection Classifiers // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society.
6. Bell S, Lawrence Zitnick C, Bala K, et al. (2016) Internal-external network: Experimental objects in contexts by cross-linking and duplicate neural networks // *IEEE conference processes for computer visualization and pattern recognition*. 2016: 2874-2883.
7. Li H, Lin Z, Shen X, et al. (2015) Convolutional neural network cascade of facial detection // *IEEE conference procedures for computer vision and pattern recognition*. 2015: 5325-5334.
8. Cao G, Xie X, Yang W, et al. (2018) Featured SSD: Quick Access to Small Items // *Ninth International Conference on Graphics and Photography (ICGIP 2017)*. International Society for Optics and Photonics 10615: 106151E.
9. Li Y, Sun X, Wang H, et al. (2012) Automatic target acquisition of target sensitive images with high resolution using a location-based location model. *IEEE Geoscience and Remote Sensitivity* 9: 886-890.
10. Takacs G, Chandrasekhar V, Tsai S, et al. (2010) Integrated real-time tracking and monitoring with instantaneous switching features // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE 934-941.
11. Zou Z, Shi Z. (2016) Discovery of a spacecraft into space photography by SVD networks. *IEEE Transactions in Geoscience and Remote Sensing* 54: 5832-5845.
12. Cheng G, Zhou P, Han J (2016) Studying convolutional neural networks unchanged to detect object in VHR optical sensing image remote. *IEEE Transactions on Geoscience and Remote Sensing* 54: 7405-7415.

Submit your manuscript to a JScholar journal and benefit from:

- ✦ Convenient online submission
- ✦ Rigorous peer review
- ✦ Immediate publication on acceptance
- ✦ Open access: articles freely available online
- ✦ High visibility within the field
- ✦ Better discount for your subsequent articles

Submit your manuscript at
<http://www.jscholaronline.org/submit-manuscript.php>