Automatic Mask Detection at Public Places

Aiman Ahmed Mohammed Al-Gamal

Department of Computer Science, St. Joseph First Grade College (University of Mysore) Mysore-570015 India

Abstract

COVID-19's fast spread has resulted in tens of millions of individuals becoming infected around the world. Because there is no specific cure for COVID-19, wearing masks has proven an effective technique of preventing transmission and is mandatory in most public places, resulting in an increase in demand for automatic real-time mask detection systems to replace manual reminders. However, there are just a few investigations on the detection of face masks. The performance of mask detectors must be improved urgently. However, there have been few studies on the detection of face masks. Mask detector performance must be improved immediately. The Properly Wearing Masked Face Detection Dataset (PWMFD) was proposed in this paper, and it includes 9205 photos of mask-wearing samples divided into three groups. Squeeze and Excitation (SE)-YOLOv3, a mask detector with roughly balanced effectiveness and efficiency, was also proposed. By incorporating the SE block into Darknet53, I was able to incorporate the attention mechanism and obtain the relationships between channels, allowing the network to focus more on the relevant feature. To increase the stability of bounding box regression, I used GIoU loss, which can better express the spatial difference between predicted and ground truth boxes. The significant foreground-background class imbalance was solved via focal loss. In addition, I used image augmentation techniques to boost the model's robustness on the challenge. In comparison to YOLOv3, SE-YOLOv3 outperformed YOLOv3 and other state-of-the-art detectors on PWMFD, achieving a higher 8.6 percent mAP while maintaining a comparable detection speed.

Keywords: Deep Learning, Deep Neural Network, Machine Learning, Python, VNC, Logitech Cam, Servo Motor, Raspberry pi3

Article Publication

Published Online: 15-Dec-2021

*Author's Correspondence

Aiman Ahmed Mohammed Al-Gamal

Department of Computer Science, St. Joseph First Grade College (University of Mysore) Mysore-570015 India

aymn3100[at]gmail.com

© 2021 The Authors. Published by RESEARCH REVIEW International Journal of Multidisciplinary. This is an open access article under the CC BY-

NC-ND license

(<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

1. Introduction

1.1 Introduction

"Prevention is better than treatment" is one of the most effective ways to stop COVID-19 from spreading and safeguard humanity. Many scientists and clinicians are working on corona treatments and vaccines. COVID-19 is transferred primarily through droplet infection, which occurs when people cough or when we touch someone who is ill and then touch our face (i.e. wiping eyes or nose). The current pandemic demonstrates that it is far more contagious and spreads quickly. We have two scenarios depending on how the illness spreads: Spread both quickly and slowly. A rapid pandemic will be devastating and result in the loss of countless lives. It arises as a result of a quick pace of infection with no countermeasures in place to slow it down. This is due to the fact that if the number of infected persons grows too great, healthcare systems would be unable to handle it. We will be short on resources such as medical personnel and equipment such as a ventilator. To avert the disaster described above, we must do all possible to make this into a slow epidemic. Only the correct interventions, especially in the early stages, can slow down a pandemic. In this period, everybody who is sick can receive treatment, and there is no need for hospitals to be flooded. We need to engineer our conduct as vaccination in this pandemic. "Not getting infected" and "not infecting others" are the two goals. Washing our hands with soap or a hand sanitizer is the finest thing we can do. Social detachment is the next best thing. It is critical to use a face mask when leaving the house, especially in public locations such as markets or hospitals, to avoid becoming sick or spreading the disease.

1.2 Objectives

The system aims to recognize faces and assess whether or not a person is wearing face masks. Using the above data, we can determine whether the person in question can be allowed to enter public places such as a market or hospital. Hospital, market, shopping malls, bus stops, restaurants, hotels and other public gatherings that require monitoring can benefit from this project. This project consists of a camera that captures the image of people entering public places and uses facial features to determine whether or not they are wearing face masks.

1.3 Existing System

Coronavirus disease 2019 (COVID-19) appeared suddenly in 2019 and had a global impact. COVID-19 has infected more than 125 million individuals worldwide and killed more than 2.7 million people as of March 26, 2021. When an infected individual coughs or sneezes, COVID-19 spreads through droplets of saliva or nasal secretions, which are highly contagious and can be Worse in crowded places. Since COVID-19 does not have a specific treatment, infection must be controlled through preventive measures. Wearing masks has been found to reduce the risk of transmission of the coronavirus and therefore wearing masks is currently one of the most effective ways preventative techniques. The correct way to wear a mask, according to the World Health Organization, is to adjust it so that it covers the mouth, nose and chin. If the masks are not worn correctly, the protection is severely reduced. Security officers are currently stationed in public places to advise individuals to wear masks. However, this method not only exposes the guards to air contaminated with the virus, but also causes congestion at the doors due to its inefficiency, and as a result, a quick and effective solution is needed to solve the situation.

1.4 Proposed System

It introduced SE-YOLOv4 as a new object detector to achieve real-time mask recognition and improve the accuracy of YOLOv4. To start, I added an attention mechanism to the core network, which can help the network allocate more resources to important features. Then I used GyoU and Loss of Focus to speed up the training process and boost performance even more. Finally, in order to achieve robustness, appropriate data augmentation methods were used for face mask detection. In order to evaluate the effectiveness and feasibility of the proposed solution, I present a prototype of an access control gate system with SE-YOLOv3 that can be installed on public gates in this section. The camera, Raspberry Pi, monitor, retractable arm, Arduino, and access gate are shown as part of the access control gate system prototype structure. The Raspberry Pi was used as a terminal to capture video frames, which were then uploaded to a server for model inference. EYOLOv4 was installed on a GPU server that was identical to the Raspberry Pi. The Raspberry Pi then displayed the inference information on the screen and issued hardware control commands to operate the access gateway and camera. To control the devices, the circuit was connected to a motor. The camera will rotate to a location where the detected face is fixed in the middle of the image using the location information. As a result, people did not have to turn their bodies toward the camera, allowing them to pass through the portal faster. Technology will react to people's movements as well as the results of discovery. People who did not wear their masks properly were not allowed to enter through the gate, and the average system processing time for one image frame was 0.13 seconds, which indicated that it could recognize face masks in real time and was useful in practical applications.

1.5 Advantages

- During a pandemic, this is a critical domain that requires extra attention. This device can be used to determine whether a patient is wearing a mask to protect the health of doctors and nurses. A CCTV camera can be installed to detect mask and social distancing breaches, reducing congestion while patients wait in lines for their appointments.
- Many destinations around the world have begun to ease travel restrictions. This mechanism can be performed at airports where people wait in queues at check-in facilities, security clearance gates, passport control and lobbies.

- Because many businesses reopen after a lockdown, this approach can be used to guarantee that customers and workers adhere to all COVID 19 safety standards.

2. Technology Survey

The image is read and one or more elements in the image are classified in object detection. A boundary called a bounding box also determines the position of those objects, and researchers have traditionally used pattern recognition to predict faces based on models of previous faces. The Viola Jones detector, an improved approach to employing Haar's digital image feature used in object recognition, was developed as a tool for detecting faces, however, it failed because it didn't work well on faces in dark places or on faces that weren't. In front of me. Researchers have been eager to create new deep learning algorithms to improve models ever since. Deep learning allows us to learn features in a comprehensive way, eliminating the need to generate feature extracts based on existing information. There are two types of deep learning-based object recognition methods: single-stage and second-stage object detection devices.

Figure 1: complete structure for the project

Two-stage detectors identify objects using two neural networks, such as area-based convolutional neural networks (R-CNN) and faster R-CNN. Coarse-to-fine detection is performed using the first neural network to generate region recommendations and the second neural network to improve these region proposals. This technique achieves good detection performance at the expense of speed. R-CNN pioneered work. R-CNN suggests some potential areas that may contain objects via selective search. The ideas are then loaded into a CNN model, which extracts the features, and a Support Vector Machine (SVM) is used to learn about object classes. By implementing a region of interest (ROI) pooling layer to simultaneously input all motion regions, Fast R-CNN solves this difficulty. Faster RCNN is a descendant of R-CNN and Fast R-CNN, and its training and testing speeds are faster than its predecessors. Whereas R-CNN and Fast R-CNN use selective search methods to limit the speed of discovery, Faster R-CNN uses a region proposal network to identify regions of proposed objects (RPN).

On the other hand, a single-stage detector uses only one neural network for suggestions and area detection; Examples include SSD (Single-shot detection) and YOLO (you only look once). The bounding boxes must be checked in advance to do this. The image is divided into several cells by YOLO, which then matches the boxes surrounding the items in each cell. This, however, is ineffective for the little things. As a result, SSD offers multi-scale detection, which can detect objects of various sizes in the image. Later, in order to increase detection accuracy, Retina Network (RetinaNet) is introduced as a way to improve detection accuracy and eliminate class

imbalance by combining SSD and FPN (Hierarchical Distinctive Network). Single-stage detectors are faster than two-stage detectors, but sacrifice detection performance. However, they are preferred over two-stage detectors.

3. System Analysis and Design

We proposed SE-YOLOv3 as a new object detector to accomplish real-time mask recognition and improve the accuracy of YOLOv3. To begin, we added an attention mechanism to the backbone network, which can assist the network in allocating more resources to critical features. Then we used GIoU and focus loss to speed up the training process and increase performance even more. Finally, in order to achieve robustness, we used appropriate data augmentation approaches for face mask detection.

Figure 2: a and b

a. Real-time Mask detection b. Complete project

4. Implementation

4.1 Experimental Discussion Results

The performance of SE-YOLOv3 on the PWMFD is demonstrated in this section. The TensorFlow deep learning framework was used for the implementation and experiments. On a local PC with an RTX 2070 GPU and 8 GB memory, the network was trained and tested. We used Adam to optimise the model during the training phase. To dynamically change the learning rate, a warm-up and attenuation technique was used. The learning rate was initially to increase after 100 rounds of warm-up. We also employed an exponential decay to avoid overfitting the model. Every 100 epochs, the attenuation was 0.9 of the initial learning rate.

4.2 Wearing Face Mask in Detection Dataset

The suggested dataset's photographs were gathered from a variety of public databases as well as images found on the internet. The Properly Wearing Masked Face Detection Dataset is what we called it (PWMFD). WIDER Face had 2951 images, MAFA had 2581 images, RMFD had 58 images, and the Internet had 3615 photographs. wereannotated all of the data because each dataset's annotation format was different and some data did not fit our requirements. We tagged the samples in such circumstances with the red label "without mask," to prevent the models from being fooled by people who covered their faces with other things. The PWMFD, as well as more information, may be found in. We divided the dataset into two parts to train and test the model. There were 7385 photos in the training set and 1820 images in the testing set. There were 7695 properly masked faces and 10,471 unmasked ones among them.

Figure 3: a and b

a. Face without mask b. Face with mask

5. Test Case

No	Test Case	Description	Result
1	Power up	Power on pi board, supply 5v 2A	Pass
2	Boot	Check OS boot on board	Pass
3	Runing	Capture image from cam	Pass
4	Mask	Detecting face-mask	Pass
5	No-mask	Detecting no face-mask	Pass
6	Cam	Cam connection lost	Fail
7	Door open	Mask detected	Pass
8	Door close	No mask detected	Pass

Table 1: Testing Case

6. Conclusion

The usefulness and importance of real-time mask detection are increasing as the urgency of controlling COVID-19 grows. To solve this problem, we created the PWMFD with 9205 high-quality masked face photos and developed SE-YOLOv4, a fast and accurate mask detector with a channel attention mechanism that improved the backbone network's feature extraction performance. We also applied GIoU and focal loss, as well as the corresponding data augmentation, to increase the model's accuracy and resilience.

7. Future Enhancement

This paper presents a solution based on deep learning to identify masks on faces in public places in order to reduce the spread of coronavirus in the community. At the pretreatment level, the proposed technique efficiently deals with blockages in dense environments through the use of a combination of single- and two-stage detectors. Group policy not only helps to reach high accuracy, but also greatly improves detection speed. Furthermore, the use of transfer learning on pre-trained models combined with large experiments on an unbiased data set has resulted in a highly reliable and low-cost system. Benefits.

Finally, the study guides scientists down exciting new paths. First of all, the proposed technology is not limited to mask detection and can be embedded in any HD video surveillance device. Second, the model can be used with a face mask to detect facial features for biometric reasons

References

- COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Available online: <https://coronavirus.jhu.edu/map.html> (accessed on 1 March 2021).
- Cai, Q.; Yang, M.; Liu, D.; Chen, J.; Shu, D.; Xia, J.; Liao, X.; Gu, Y.; Cai, Q.; Yang, Y.; et al. Experimental treatment with Favipiravir for covid-19: An open-label control study? *Engineering* 2020. [CrossRef]

- Eikenberry, S.; Mancuso, M.; Iboi, E.; Phan, T.; Eikenberry, K.; Kuang, Y.; Kostelich, E.; Gumel, A.B. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infect. Dis. Model.* 2020, 5, 293–308. [CrossRef] [PubMed]
- Leung, N.H.; Chu, D.; Shiu, E.Y.; Chan, K.; McDevitt, J.J.; Hau, B.J.; Yen, H.; Li, Y.; Ip, D.; Peiris, J.S.; et al Respiratory virus shedding in exhaled breath and efficacy of face masks. *Nat. Med.* 2020, 26, 676–680. [CrossRef] [PubMed]
- World Health Organization (WHO) Q&A: Masks and COVID-19. Available online: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/q-a-on-covid-19-and-masks> (accessed on 4 August 2020).
- Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep Learning for Computer Vision: A Brief Review. *Comput. Intell. Neurosci.* 2018. [CrossRef]
- Zhengxia, Z.; Shi, Z.; Yuhong, G.; Jieping, Y. Object Detection in 20 Years: A Survey. *arXiv2019*, arXiv:1905.05055.
8. Jiang, M.; Fan, X. RetinaMask: A Face Mask detector. *arXiv2020*, arXiv:2005.03950.
- Detect Faces and Determine Whether People Are Wearing Mask. Available online: <https://github.com/AIZOOTech/FaceMaskDetection> (accessed on 21 March 2020)