IJAMEC

INTERNATIONAL JOURNAL OF APPLIED METHODS IN ELECTRONICS AND COMPUTERS

www.ijamec.org

International Open Access

Volume 12
Issue 03

September, 2024

Research Article

YOLO-V4 Based Real-Time Face Mask Detection via Unmanned Aerial Vehicle

Mucahit Akar [*]***, Muhammet Fatih Aslan*** ***, Kadir Sabanci***

Electrical and Electronics Engineering, Karamanoglu Mehmetbey University, Karaman, Türkiye

ARTICLE INFO

Article history:
Received 21 July 2024
Accepted 25 September 2024

Keywords:
YOLO-v4
object detection
mask detection
unmanned aerial vehicle
COVID-19

ABSTRACT

COVID-19, which started in Wuhan city of China's Hubei province and then affected the whole world, continues to spread despite the measures taken. One of the most important of these measures is to use a mask. In some countries, while wearing a mask is mandatory in crowded environments, it is just as difficult to control it. Failure to detect individuals violating the mask causes the virus to spread, resulting in an increase in the number of cases and an increase in the number of deaths. Therefore, detecting the mask and taking action against it is an extremely important issue. In this study, in addition to making mask detection easily and quickly, mask detection is made by using unmanned aerial vehicles (UAVs) using images taken from different angles and different heights. For mask detection from UAV images, training and validation processes were applied using the YOLO-v4 algorithm on a public dataset containing 1510 masked and unmasked human face images. As a result of the training on this dataset, mean-average precision (mAP) was achieved with a success rate of 92.06%. Then, real-time mask detection was performed on the images taken with the DJI Ryze Tello quadcopter using the trained network. The results showed that the UAV is applicable in crowded environments required for autonomous mask detection and gives successful results.

1. Introduction

In December 2019, pneumonia started in Wuhan City of China's Hubei province for an unknown reason, and then it was not only limited to China but began to spread all over the world [1]. Investigations revealed that what caused this was caused by an infectious virus called coronavirus (2019-nCoV or COVID-19) also known as the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). According to the World Health Organization (WHO) report on 26 July 2021, 4.162.304 deaths and 194.080.019 confirmed cases are reported globally [2] and these data continue to increase at varying rates, even if precautions are taken.

In today's world where human interaction is inevitable, even if it is indirect or direct, the virus continues to spread through these interactions. Epidemiological data have revealed that droplets are expelled when people are talking face to face. It has also shown that this phenomenon is more widely spread during talking, coughing, and sneezing [3]. These are the ways that should be especially taken into account so that the virus does not spread and infect more people. Therefore, the use of masks is one of the measures to be taken to prevent the transmission and spread of the coronavirus [4]. Cheng, Wong, Chuang and et al. [5], in their study on the importance of wearing masks for coronavirus in Hong Kong Special Administrative Region (HKSAR); conducted an epidemiological analysis for people who caught COVID-19 in mask-wearing and non-mask-wearing environments. The masked population and the countries that do not wear masks were compared. As a result of the study, they concluded that the amount of COVID-19 caught through saliva and respiratory droplets in communities wearing masks have decreased and that wearing a mask can contribute to the control of COVID-19. Also, in their examinations of mask-wearing and non-mask-wearing samples, Leung, Chu, Shiu et al. [6] detected respiratory droplets and coronavirus in 30% and 40% of aerosols taken

* Corresponding author. E-mail address: *mchtakar13@hotmail.com*
DOI: 10.58190/ijamec.2024.106

from samples not wearing masks, respectively. Likewise, they detected no respiratory droplets and viruses in aerosols from samples wearing masks.

While wearing a mask is so important, it has become possible at a certain level to control whether people are wearing masks in the public. While this inspection can be provided by people such as security guards and health workers, this inspection becomes difficult in crowded environments. This situation increases the spread of the virus due to violations by people who do not wear masks. For this reason, researchers are directed to make mask inspections autonomously with artificial intelligence-based applications. Loey et al. [7] designed a hybrid model using deep and classical machine learning for face mask detection. They used ResNet-50 as a feature extractor and decision tree, Support Vector Machine (SVM), and ensemble algorithm for the classification process. They achieved 99.64%, 99.49%, and 100% testing accuracy in three different datasets, respectively. Furthermore, they [8], performed a model using ResNet-50 deep transfer learning for feature extraction and YOLO-v2 for the detection of medical face masks. They achieved the highest average precision percentage of 81% as a detector. Chowdary, Punn, Sonbhadra, and Agarwal [9] achieved 99.9% accuracy during training and 100% during testing using transfer learning of InceptionV3 model on Simulated Masked Face Dataset (SMFD). In addition, face mask detection applications are applied to achieve high accuracy results by using different state-of-art deep learning models [10], [11]. In their study, Wang, Chen, Wei and Ling [12] proposed a method for manual inspection with deep learning-based YOLO-v4 algorithm using pruning and center loss algorithms for face mask detection. Yu and Zhang [13] proposed a face mask recognition method to improve low-accuracy results, low real-time performance, and poor robustness with improved YOLO-v4. The achieved mean average precision (mAP) of face mask recognition can reach 98.3%.

In this context, it is vital to detect mask violations in universities, city centers, or in short, places where the number of people is high because the risk of spreading the virus is higher in such places. In this case, since a person's mask violation affects many people negatively in cases of COVID-19 and similar situations, it is necessary to conduct a strict inspection. Unlike other studies, this study will be able to detect more people with fewer detection error images from UAV thanks to data taken from different angles and heights. Thanks to the images taken from different angles and heights, the inspection will be more robust. Instead of many cameras used for images from different angles in crowded environments, this process can be done with the help of an unmanned aerial vehicle with less cost. In addition, while mask inspection was carried out manually in many places via UAV, as in the example of the application made by security guards in Aydın, Turkey [14]. Inspired by this application, in this study, the mask inspection process provides autonomous detection via UAV. Therefore, this study will facilitate mask detection with the help of UAVs using the deep learning-based approach, by governments, companies, and businesses that require special supervision, as well as by the employees in charge of mask inspection.

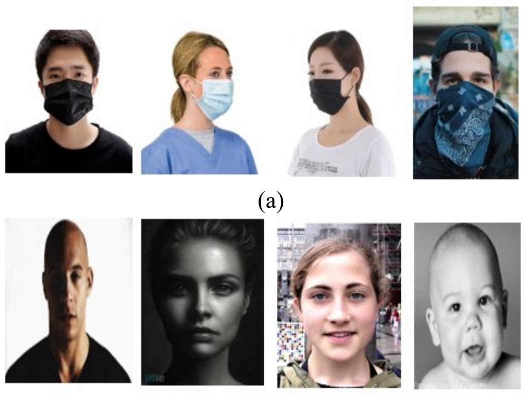
2. Dataset

This section introduces the public and our UAV dataset.

2.1. Labeled Mask Dataset

This research originally bases on images collected from three different datasets. The dataset contains 1510 images along their YOLO labeled text files [15] also known as Labeled Mask Dataset. The first dataset is the Real-World Masked Face dataset (RMFD) [16] which is the most commonly used dataset in face mask detection. RMFD dataset contains 5,000 masked faces of 525 people and 90,000 normal faces. The second dataset is the Mask Wearing Dataset [17], which is an object detection dataset of individuals wearing the various types of masks and those without masks. The images were originally collected by Cheng Hsun Teng from Eden Social Welfare Foundation, Taiwan and relabeled by the Roboflow team. The third dataset is the Simulated Masked Face Dataset (SMFD) [18], which consists of 1570 images, 785 for simulated mask faces, 785 for unmasked faces. Some samples of masked and unmasked faces among the images in the dataset are illustrated in Figure 1.

In this study, a total of 1510 images belonging to two classes (masked and unmasked faces) were used 1359 of them were used for training the model, while the remaining 151 were used for validating the model.



(a)

(b)

Figure 1. Labeled Mask Dataset samples [14]

a) masked faces b) unmasked faces

2.2. DJI Tello Dataset

This dataset, collected in real-time by the UAV, was used to test the trained model. The mask detection process was applied on the video frames collected by the UAV in the university, and people wearing masks were surrounded by

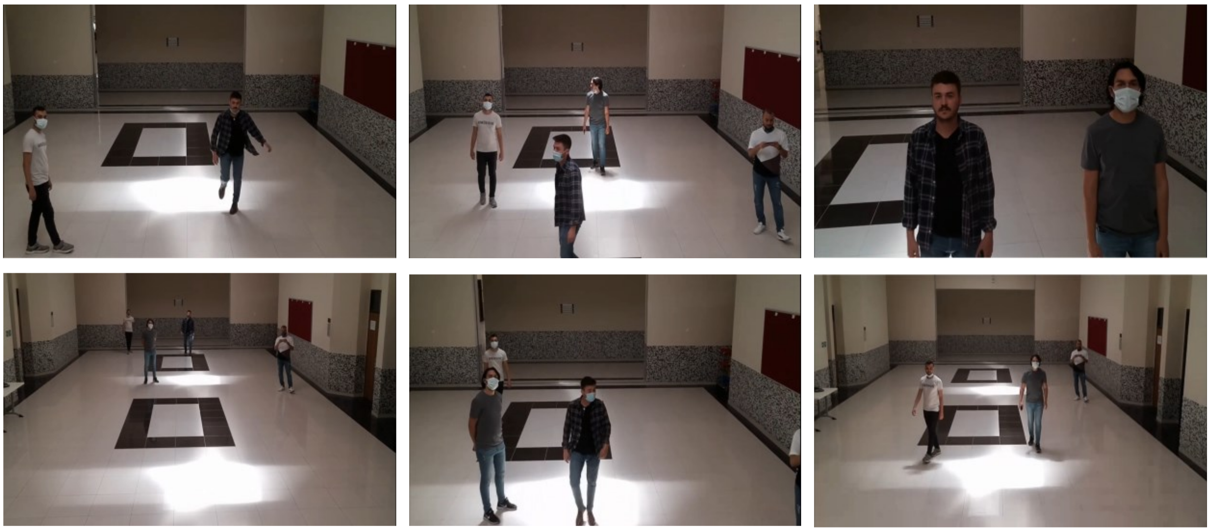


Figure 1. DJI Tello dataset (Some real-time video frames taken from DJI Ryze Tello quadcopter)

green bounding boxes, while those who did not wear masks were framed by red bounding boxes. Some of the video frames used in the application are shown in Figure 2.

3. Methodology

To implement the face mask detection method, this study is based on a deep learning model known as YOLO-v4 [19]. In this study, the algorithm detects objects belonging to two classes – masked faces and unmasked faces. One aspect of the YOLO algorithm that puts it ahead of other algorithms is the fast operation of the algorithm. Therefore, it is frequently used in real-time object detection applications.

In this study, real-time face mask detection is performed over video frames obtained from the UAV. All steps about the YOLO-v4 algorithm are discussed in more detail below.

3.1. YOLO-v4

Object detection is a method that detects the location and classes of an object on an image. YOLO versions are one of the most preferred object detection algorithms in real-time applications. YOLO-v4 is an important improvement of YOLO-v3 [20], the implementation of new architecture in the backbone and the modifications in the neck have improved the mean average precision (mAP) by %10 and the number of FPS (frame per second) by %12. The pipeline consists of three parts, the backbone, the neck, and the head as shown in Figure 3. The backbone is to extract essential features. This is the most important part of improving the performance of object detection. There are three parts of backbone – bag of freebies, bag of specials, CSPDarknet53. On the other side, the main role of the neck is to collect feature maps from different stages. In the case of a single-stage detector, the head is the part that makes the predictions with the dense. The result determined by these predictions is a vector containing the coordinates (height, width, confidence score, and class/label). In this way, the predictions with the highest confidence score are enclosed in a bounding box with its height and width values and displayed with the class label and confidence score.

2.1. Experimental Setup

In this study, DJI Ryze Tello Mini Quadcopter was used as a UAV. The images are converted into images before the network is trained. Afterward, real-time images from UAV were converted dimensions in the same way and given as an input to the trained model. Thus, a fixed model input is provided for the model. The YOLO-v4 model created by this process performs mask detection.

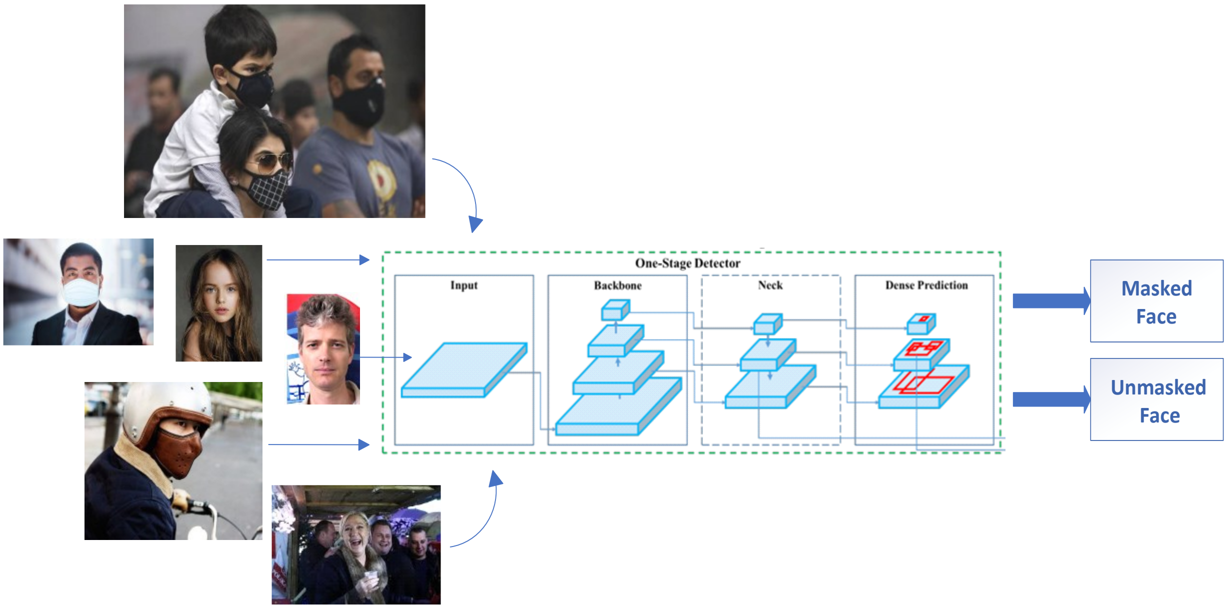


Figure 2. YOLO-v4 algorithm architecture from [18]

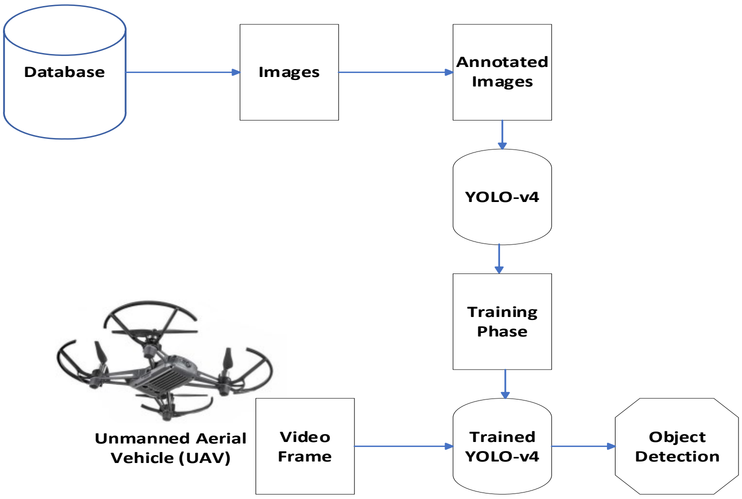


Figure 3. Flow diagram of YOLO-v4 algorithm

Table 1. YOLO-v4 model training parameters.

Batch size	64
Subdivisions	16
Momentum	0.949
Learning rate	0.001
Maximum batches	6000
Steps	4800, 5400

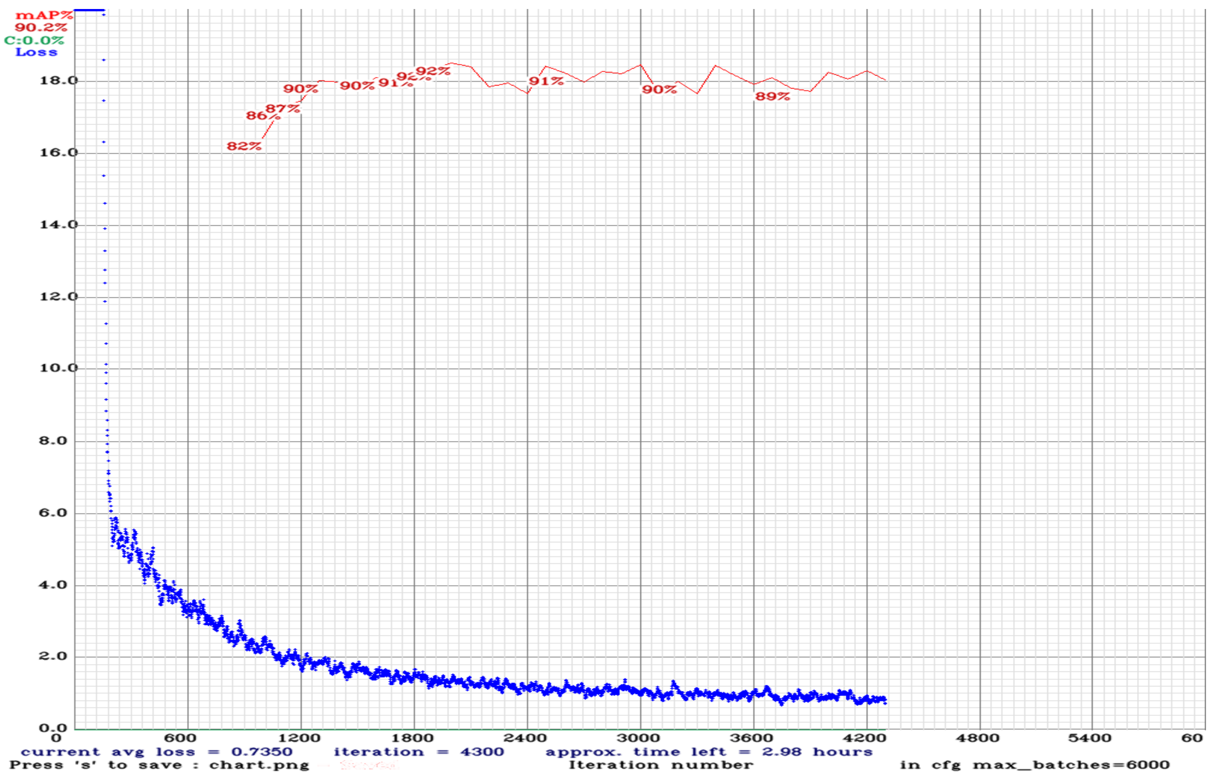


Figure 5. mAP and loss values of the proposed YOLO-v4 model

The flow diagram of the object detection is as in Figure 4. The parameters selected for the model training are shown in Table 1.

All the practices were performed on a laptop with Intel Core i7-7700HQ NVIDIA GeForce GTX 1050 8GB, 16 GB RAM. The results of the model after training and the results during real-time object detection will be explained in detail in the following section.

3. Results

In this section, results are discussed based on two parts. The first part is the training part where the neural networks are trained. In the mask detection made with the YOLO-v4 algorithm, mAP@0.50, Precision, Recall and F1-Score values were obtained as 92.09%, 0.88, 0.88 and 0.88, respectively.

In addition, and loss values that change depending on iteration are shown in the graph in Figure 5.

The second part is the experimental results. In this part, object detection results are collected on real-time video frames from the UAV. Objects belonging to two classes are shown with different colors while performing the object detection process. While people wearing masks are seen in the green bounding box, people not wearing masks are seen in the red bounding box. In Figure 6, some samples of video frames where mask detection is made are shown.

4. Conclusion

In places where the population is high, it is very difficult to control face masks without the help of technology. It would be more logical to do this check with the images taken from the cameras. However, since the cameras standing in a fixed position take images from a single angle, it will reduce the detection of those who violate the mask during mask inspection. In addition, in motion cameras, this situation can be overlooked in crowded environments unless viewed from a certain height. Therefore, performing this inspection at a certain height (changeable) and from different angles (moving) will reduce mask violations. For exactly this reason, in this study, deep learning models were trained using the YOLO-v4 algorithm and mask detection was made with this model, thanks to real-time images from the UAV.

5. Discussion and Future Works

In this study, unlike the real-time applications made with the YOLO-v4 algorithm, real-time object detection obtained from the UAV was performed.

Detections from different angles and heights yielded successful results, but careful control of UAVs in closed areas must be ensured. In addition, the charging capacity of the UAV and the stable flying characteristic of the UAV are aspects of this study that needed to be developed or considered.

In future studies, it is aimed to carry out this control in larger areas by providing the communication of more than one unmanned aerial vehicle. The training success of the model will be increased by increasing the number of data and/or using data augmentation methods. It is also planned to include route planning and navigation applications for unmanned aerial vehicles.

Figure 6. Object detection on real time video frames

Acknowledgment

This research was supported by the project number 18-YL-21, which was accepted by the Karamanoglu Mehmetbey University Scientific Research Projects Commission.

Declarations

Funding: The authors do not have any funds.

Availability of data and material: The authors confirm that the data supporting the findings of this study are available within the manuscript.

Code availability: Code can be requested from the corresponding author via e-mail.

Conflict of Interest: The authors declare that they have no conflict of interest

References

[1] D. P. Scheel., "Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-," *Ann Oncol*, no. January, pp. 19–20, 2020.

[2] WHO Coronavirus(COVID-19) Dashboard, [Online]. Available: https://covid19.who.int

[3] W. J. Wiersinga, A. Rhodes, A. C. Cheng, S. J. Peacock, and H. C. Prescott, "Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review," *JAMA - J. Am. Med. Assoc.*, vol. 324, no. 8, pp. 782–793, 2020, doi: 10.1001/jama.2020.12839.

[4] Coronavirus disease (COVID-19) advice for the public: When and how to use masks, [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/when-and-how-to-use-masks

[5] V. C. C. Cheng *et al.*, "The role of community-wide wearing of face mask for control of coronavirus disease 2019 (COVID-19) epidemic due to SARS-CoV-2," *J. Infect.*, vol. 81, no. 1, pp. 107–114, 2020, doi: 10.1016/j.jinf.2020.04.024.

[6] N. H. L. Leung *et al.*, "Respiratory virus shedding in exhaled breath and efficacy of face masks," *Nat. Med.*, vol. 26, no. 5, pp. 676–680, 2020, doi: 10.1038/s41591-020-0843-2.

[7] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic," *Meas. J. Int. Meas. Confed.*, vol. 167, no. May 2020, p. 108288, 2021, doi: 10.1016/j.measurement.2020.108288.

[8] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection," *Sustain. Cities Soc.*, vol. 65, no. June 2020, p. 102600, 2021, doi: 10.1016/j.scs.2020.102600.

[9] G. Jignesh Chowdary, N. S. Punn, S. K. Sonbhadra, and S. Agarwal, "Face Mask Detection Using Transfer Learning of InceptionV3," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12581 LNCS, pp. 81–90, 2020, doi: 10.1007/978-3-030-66665-1_6.

[10] S. A. Sanjaya and S. A. Rakhmawan, "Face Mask Detection Using MobileNetV2 in the Era of COVID-19 Pandemic," *2020 Int. Conf. Data Anal. Bus. Ind. W. Towar. a Sustain. Econ. ICDABI 2020*, 2020, doi: 10.1109/ICDABI51230.2020.9325631.

[11] M. Jiang, X. Fan, and H. Yan, "RetinaMask: A Face Mask detector," 2020, [Online]. Available: http://arxiv.org/abs/2005.03950.

[12] X. Wang, Z. Chen, B. Wei, and M. Ling, "Application of Pruning Yolo-V4 with Center Loss in Mask Wearing Recognition for Gymnasiums and Sports Grounds of Colleges and Universities," *2020 IEEE 6th Int. Conf. Comput. Commun. ICCC 2020*, pp. 1373–1377, 2020, doi: 10.1109/ICCC51575.2020.9345257.

[13] J. Yu and W. Zhang, "Face mask wearing detection algorithm based on improved yolo-v4," *Sensors*, vol. 21, no. 9, 2021, doi: 10.3390/s21093263.

[14] Anadolu Ajansı."Aydın'da maske takmayanlar drone ile tespit edilip uyarıldı"(22.05.2020), 1. Available: https://www.aa.com.tr/tr/turkiye/aydinda-maske-takmayanlar-drone-ile-tespit-edilip-uyarildi/1850647

[15] Labeled Mask Dataset, [Online]. Available: https://www.kaggle.com/techzizou/labeled-mask-dataset-yolo-darknet

[16] Z. Wang et al., "Masked Face Recognition Dataset and Application," no. March, 2020, [Online]. Available: http://arxiv.org/abs/2003.09093.

[17] Mask Wearing Dataset, [Online]. Available: https://public.roboflow.com/object-detection/mask-wearing

[18] Simulated Face Masked Dataset (SFMD), [Online]. Available: https://github.com/prajnasb/observations

[19] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," 2020, [Online]. Available: http://arxiv.org/abs/2004.10934.

[20] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," 2018, [Online]. Available: http://arxiv.org/abs/1804.02767.