The Face Mask Detection For Preventing the Spread of COVID-19 at Politeknik Negeri Batam

Susanto
Electrical Engineering
Politeknik Negeri Batam
Batam, Indonesia
susanto@polibatam.ac.id

Febri Alwan Putra
Electrical Engineering
Politeknik Negeri Batam
Batam, Indonesia
febry@polibatam.ac.id

Riska Analia
Electrical Engineering
Politeknik Negeri Batam
Batam, Indonesia
riskaanalia@polibatam.ac.id

Ika Karlina Laila Nur Suciningtyas
Electrical Engineering
Politeknik Negeri Batam
Batam, Indonesia
ikakarlina@polibatam.ac.id

2020 3rd International Conference on Applied Engineering (ICAE) | 978-1-7281-9917-7/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICAE50557.2020.9350556

***Abstract*— After the new Coronavirus disease (COVID-19) case spread rapidly in Wuhan-China in December 2019, World Health Organization (WHO) confirmed that this is a dangerous virus which can be spreading from humans to humans through droplets and airborne. As for the prevention, wearing a face mask is essentials while going outside or meeting to others. However, some irresponsible people refuse to wear face mask with so many excuses. Moreover, developing the face mask detector is very crucial in this case. This paper aims to develop the face mask detector which is able to detect any kinds of face mask. In order to detect the face mask, a YOLO V4 deep learning has been chosen as the mask detection algorithm. The experimental results have been done in real-time application and the device has been installed at Politeknik Negeri Batam. From the experimental results, this device is able to detect the people who wear or do not wear the face mask accurately even if they are moving to various position.**

Keywords—COVID-19, mask detector, YOLO, deep learning.

I. INTRODUCTION

The spread of COVID-19 is increasingly worrying for everyone in the world. This virus can be affected from human to human through the droplets and airborne. According to the instruction from WHO, to reduce the spread of COVID-19, every people need to wear face mask, do social distancing, evade the crowd area and also always maintain the immune system. Therefore, to protect each other, every person should wear the face mask properly when they are in outdoor. However, most of selfish people won't wear the face mask properly with so many reasons.

To overcome this situation, a robust face mask detection needs to be developed. In order to detect a face mask, the object detection algorithm can be implemented. The state of art of object detection algorithm which has a robust performance is the You Only Look Once (YOLO). As presented in [1], Susanto, et al., used the YOLO deep learning method to distinguish the white ball and goal which is integrated to humanoid robot soccer. This algorithm has been carried out by using the NVIDIA JETSON TX1 controller board. The other work implemented the YOLO was introduced by Liu, et al. [2]. In this work they implemented the traditional image processing in order to shooting of the noise, blurring and rotating filter in real-world. Then they used the YOLO algorithm to train a robust model to improve the traffic signs detection. On the other hands, Yang, et al., [3] used the YOLO algorithm to detect the face in real-time application with accuracy and fast detection time. In contrast with [3], in [4] they improved the YOLO algorithm for detecting the face in a video sequence and compared the accuracy of detecting to the traditional approach. They also used the FDDB dataset for training and testing out the model. The improvement of YOLO model also has been done by Zhao, et al [5]. They improved YOLO model to detect the pedestrian which address two issues such as leverage real-time saliency through surveillance camera and extract the detail of distinguished feature.

A few years later, YOLO method was improved to YOLO V2 which was able to detect over 9000 object categories. In this version, a novel, multi-scale training method was developed [6]. Thereafter, Kim, et al., implemented the YOLO-V2 for image recognition and other testbenches for a CNN accelerator. In this work, the YOLO V2 method has been applied through the simulation and FPGA experiment [7]. Harisankar, et al. [8] on the other hand, modified the YOLO V2 to detect and localize the pedestrian by adding the Model H architecture. In order to detect the pedestrian precisely, they used ZED camera and created the depth map.

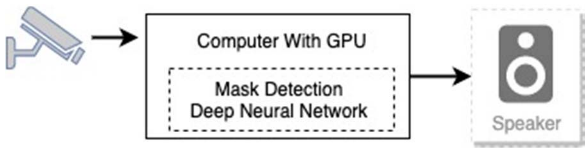


Fig. 1. The block diagram of the face mask detection hardware.

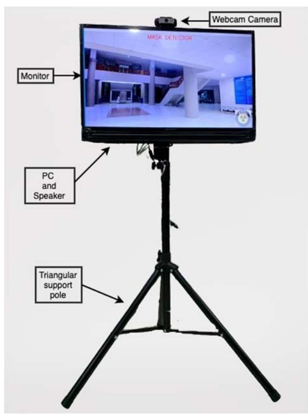


Fig. 2. The prototype of face mask detector.

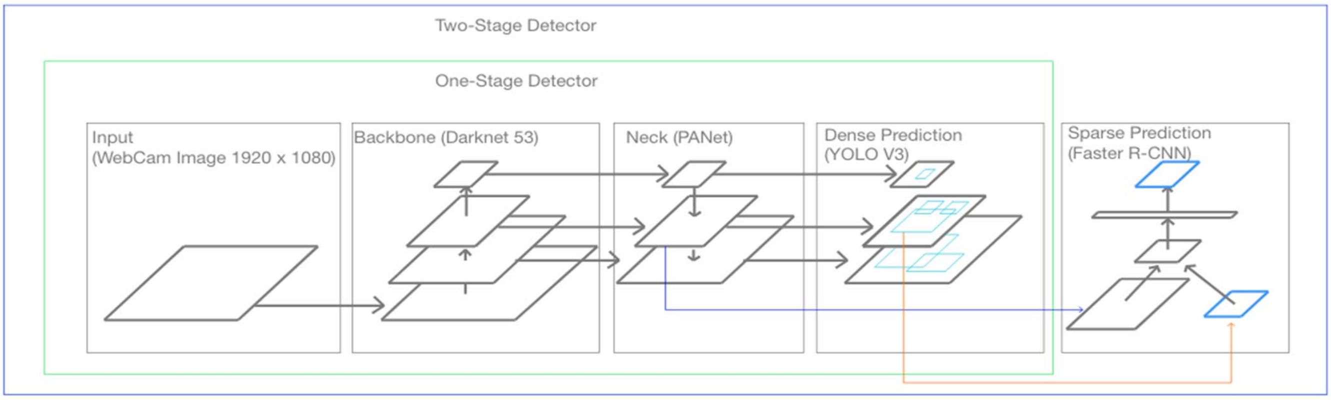


Fig. 3. The face mask detection system.

Within two years Redmon, et al., introduced the YOLO V3 although the architecture is a little bigger than the last time, this version is accurately as SSD algorithm but three times faster [9]. Therefore, Hu, et al. [10] implemented this version to detect the workers with or without helmets in videos. The first, YOLO V3 model was to identify and intercept the worker video then made the sample model of wearing and not wearing helmet. The theorical analysis and experimental results verify that the proposed algorithm is able to detect the helmet with high detection accuracy. Because of the YOLO method is open source, then it allows everyone to improve the algorithm of the YOLO method. As presented in [11], Alexey, et al. introduced the YOLOv4 with optimal speed and accuracy. In this work, they assumed that such universal features such as Weighted-Residual-Connections (WRC), Cross-Stage-Partial connections (CSP), Cross mini-Batch Normalization (CmBN), Self-adversarial-training (SAT) and Mish activation are able to give high detection accuracy. According to the state of arts and the results which have already been tested by some researcher, in this work we developed a face mask detection for COVID-19 prevention by using the YOLO V4 algorithm. In contrast with [11], the YOLO v4 algorithm will be used to detect the face mask as the object by adding several feature which will be explained in section III. The face mask detection will be applied in real-time application to detect all types of commercial face mask. By adding the YOLO V4, it is hoped that this device able to detect whether the users are wearing a face mask or not.

The rest of this paper will be organized as follows: section two illustrates the hardware architecture. Section three describes the face mask detection algorithm and Section four presents the experimental results in real-world condition. This work will be closed by the conclusion and future work which are delivered in section five.

II. The Hardware Elements of Face Maks Detection

This section will explain about the hardware architecture of our system. The hardware diagram system can be seen on Fig. 1. The whole hardware system on Fig. 1 consists of a digital webcam camera, PC, and a speaker. In order to do the face mask detection, a deep neural network has been chosen as a detection method. All the computation of the face detection is done in a computer which is mounted to the GPU to enhance the graphical calculation of the image. This system is running in real-time application when the camera detected the user who wear or un-wear the face mask. When the system detected the user who un-wear or wear face mask from the camera, it will order the speaker to alert him/her to wear the face mask and will be going on until the user put their face mask properly.

The prototype of the face mask detector is presented on Fig. 2. This figure shows the position of each parts of the face mask detector when it is operated in real-time application. The webcam camera, as seen on Fig. 2, was mounted on the top of the monitor, which allowed the detection of the user at around 5 meters. The PC and speaker on the other hand, was put on the back of the monitor for the safety reason and to make it more compact so that it can be moved everywhere freely. The prototype, illustrated on Fig. 2, is equipped by MiniPC embedded with GTX 1060 Nvidia GPU for the object detection method calculation.

III. The Face Mask Detection

As for the face mask detector method, this work implemented the deep neural network known as YOLO V4. According to [11], the YOLO V4 is able to run twice faster than the other deep neural network method which is used to detect the object. The performance of this version is able to improve the YOLOv3' AP by 10% and FPS by around 12%. By reflecting these results, it is suitable to implement the method into the real-time face mask detector where high detection accuracy is needed.

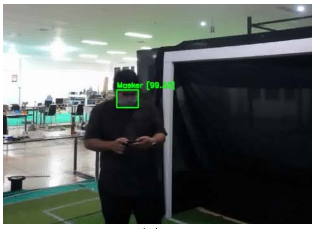
In this work, the object which needs to be detected is the face mask wearer. The YOLO V4 which has been implemented in this work consists of two-stages detector. As seen on Fig. 3, the first-stage detector consists of input, backbone, neck, and dense prediction. Moreover, the second-stage of detector has sparse prediction to predict the object by understanding the bounding boxes and the classes on the object. The input image of this work needs to be handled in the resolution of around 1920x1080 pixel in real-time application to ensure the detection of the moving object who wear the face-mask properly. In this part, the convolutional layer 3x3 is built and the larger number of parameters will be selected as the backbone input layer. While at the backbone parts, the Darknet53 has been chosen as a detector method. In this backbone, the input network resolution will be shrunk into 512x512 with receptive field size of 725x725 and contains 29 convolutional layers by 3x3 and then each layer will be sent to the neck detector. The PANet is applied as the neck detector method for understanding the parameter aggregation from different backbone level detector. All the aggregation layer from the neck detector will be sent to the sparse prediction parts as the input layer in this stage is represented as line blue on Fig. 3. Meanwhile, the last detector process on stage one is done at the dense prediction. The YOLO V3 model is used in this stage to generate the prediction which result will be used as the input prediction at the second-stage predictor. The second stage predictor has sparse prediction which applies the faster R-CNN as the prediction method. In this stage, it got the input layer from the neck which is 3x3 layer and the input prediction from the dense prediction. The result from this stage are two classes such as face-mask wearer and un-wearer of face-mask.

IV. EXPERIMENTAL RESULTS

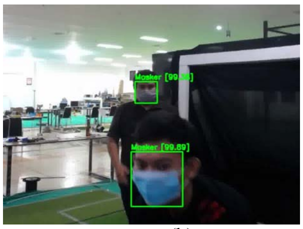
This section will present the experiment results of the face mask detection in real-time application and that has already been installed at Politeknik Negeri Batam. The first experiment which is depicted on Fig. 4 has been done as the first trial before it is implemented for the moving person. Fig. 4(a) illustrates the face detector that detected the single user wearing a face mask accurately even it has some disturbance in the area. As for Fig. 4(b), the user was added slowly from below the camera and the detector was able to detect the mask properly. When the users are standing close to each other as seen on Fig. 4 (c)-(f), this system was also able to detect the face mask even if the user was surrounded by many objects with similar color.

After did the trial with no error, we are ready to verify this device with more user. As seen on Fig. 5(a)-(c), we added the user into three people with different types of face mask such as surgical and fabric face mask. Each person was standing in different position to verify the performance of face mask detection. From the picture it is verify that the face detected remains steady in detecting face mask of the users even the lighting was in different brightness. To make different in brightness, we turned off and on the lamp at our lab as to test the feature of this system.

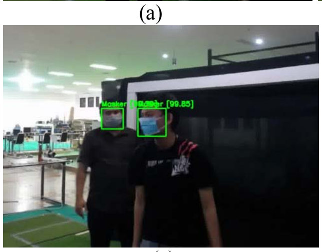
On the other hand, the experiment of detecting a non-wearing mask is presented on Fig. 6(a)-(c). On Fig. 6(a) the first user who wear a white T-shirt attempted to pull off his mask, and the mask detector was able to distinguish the non-wearing mask and mask-wearing user. Fig. 6(b) also presented the mask detection precisely. Moreover, on Fig. 6(c) the first and third users tried to take off their mask, and the mask detector detected the face mask condition steadily.



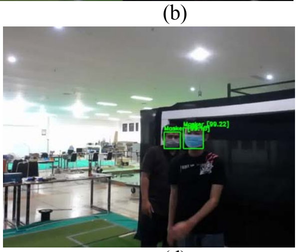
(a)



(b)

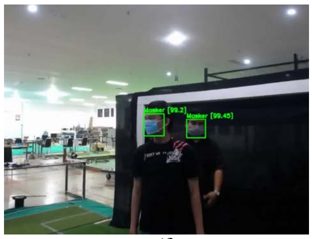


(c)



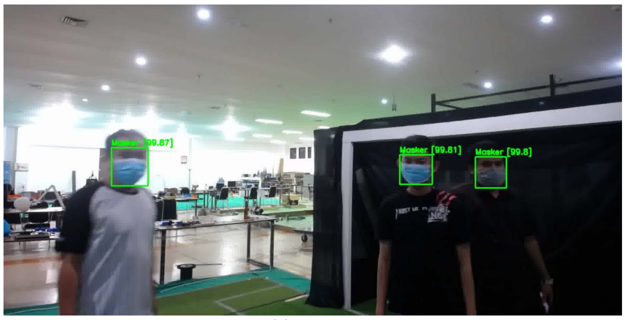
(d)

(e)

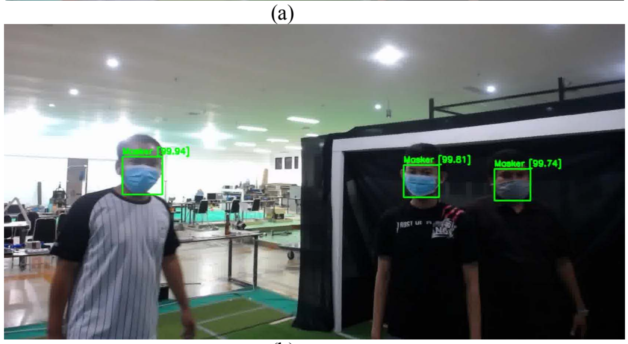


(f)

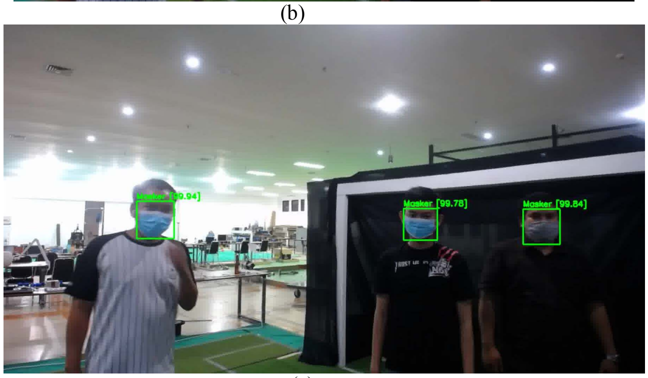
Fig. 4. The face mask detector detected the user who wear the face mask (a) alone in the frame, (b) detected the new user, (c)-(f) detected the face mask where the user was close to each other.



(a)

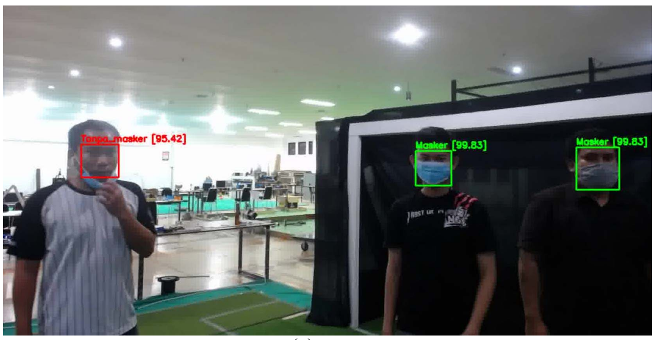


(b)

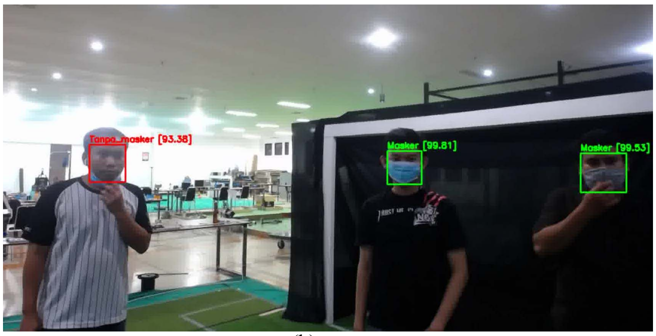


(c)

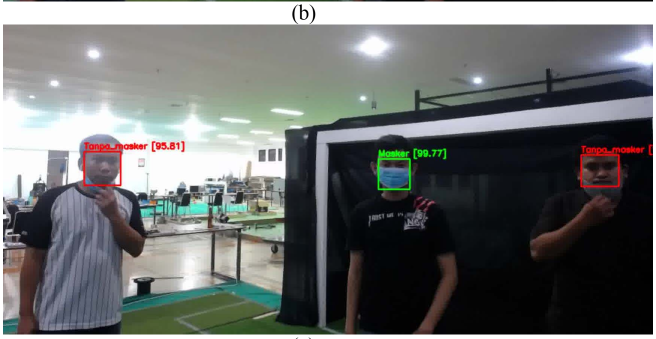
Fig. 5. The detector detected multiple people who are wearing a face mask with different position from each other.



(a)

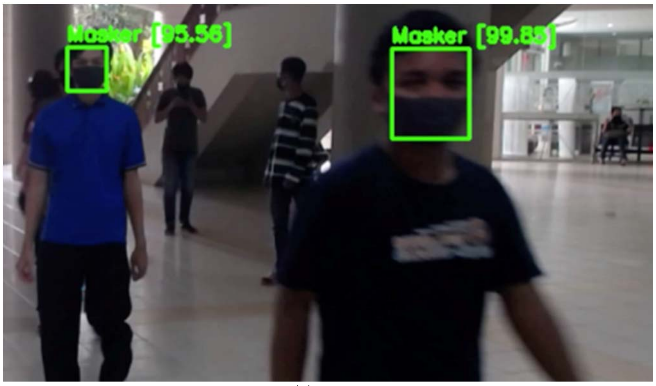


(b)

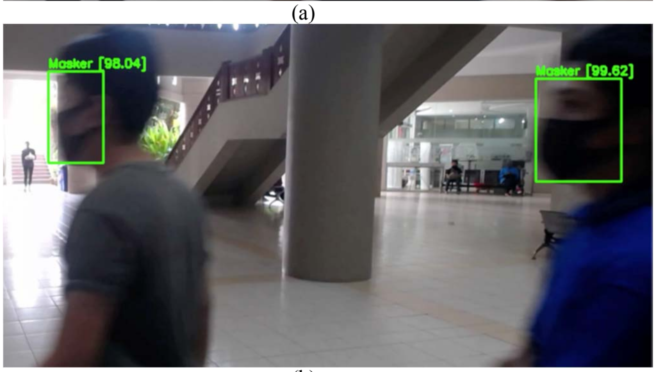


(c)

Fig. 6. The face mask detection detected the non-wearing user with different angle of pose.



(a)

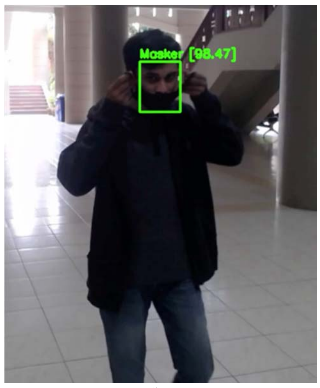


(b)

Fig. 7. The face detection detected the fabric face mask while (a) the user moving towards to the device detector, (b) the user moved away from the device.

When the user attempted not to wear the face mask properly, represented on Fig. 8, this device will announce that the user needs to wear the mask by the speaker plugged into the device. Fig. 8 (a) illustrates the non-wearing user of face mask and Fig. 8 (b) for wearers of face mask. From all these experiments, the average FPS generated by the face detection is about 11,1 FPS which is presented on Fig. 9. From all these experiments, the face detector which is built by YOLO v4 algorithm is able to detect and distinguish a non-wearing and a wearing-mask user properly in every different situation such as lighting, mess up area, and clean area.

(a)



(b)

Fig. 8. The face-mask detected the (a) a non-wearing face mask user, (b) a wearer of face mask.

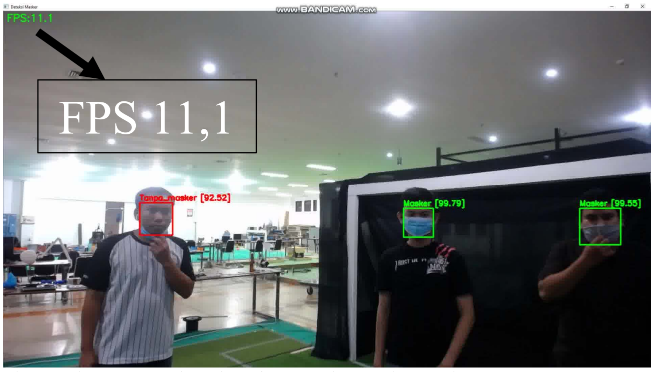


Fig. 9. The FPS results when face mask was detected.

V. Conclusion and Future Work

This work developed the face mask detection by using YOLO V4 algorithm. The YOLO V4 algorithm consists of deep learning method which is able to detect the object properly. This device has already been installed at Politeknik Negeri Batam in real-time application to avoid the spread of COVID-19 in campus area. From the experiment results, the algorithm is able to detect and distinguish a non-wearing and a wearing-mask precisely with any condition of surrounding environment. In the future, we will add the thermal detection on this device to help the guard's work easier. Furthermore, this device is hopped to be installed in other crowd area which need face mask detector.

References

[1] Susanto, E. Rudiawan, R. Analia, D. S. Pamungkas, and H. Soebakti, "The deep learning development for real-time ball and goal detection of barelang-FC," in 2017 International Electronics Symposium on Engineering Technology and Applications (IES-ETA), Surabaya, 2017, pp. 146-151, doi: 10.1109/ELECSYM.2017.8240393.

[2] C. Liu, Y. Tao, J. Liang, K. Li and Y. Chen, "Object detection based on YOLO network," in 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 2018, pp. 799-803, doi: 10.1109/ITOEC.2018.8740604.

[3] W. Yang and Z. Jiachun, "Real-time face detection based on YOLO," in 2018 1st IEEE International Conference on Knowledge Innovation and Invention (ICKII), Jeju, 2018, pp. 221-224, doi: 10.1109/ICKII.2018.8569109.

[4] D. Garg, P. Goel, S. Pandya, A. Ganatra and K. Kotecha, "A deep learning approach for face detection using YOLO," in 2018 IEEE Punecon, Pune, India, 2018, pp. 1-4, doi: 10.1109/PUNECON.2018.8745376.

[5] C. Zhao and B. Chen, "Real-time pedestrian detection based on improved YOLO model," in 2019 11th International Conference on

Intelligent Human-Machine Systems and Cybernetics (IHMSC), Hangzhou, China, 2019, pp. 25-28, doi: 10.1109/IHMSC.2019.10101.

[6] J. Redmon, & A. Farhadi, YOLO9000: Better, Faster, Stronger, 2016.

[7] C. Kim et al., “Implementation of Yolo-v2 image recognition and other testbenches for a CNN accelerator,” in 2019 IEEE 9th International Conference on Consumer Electronics (ICCE-Berlin), Berlin, Germany, 2019, pp. 242-247, doi: 10.1109/ICCE-Berlin47944.2019.8966213.

[8] H. V and K. R, “Real time pedestrian detection using modified YOLO V2,” in 2020 5th International Conference on Communication and Electronics Systems (ICCES), COIMBATORE, India, 2020, pp. 855-859, doi: 10.1109/ICCES48766.2020.9138103.

[9] J. Redmon and A. Farhadi, YOLOv3: An Incremental Improvement, 2018.

[10] J. Hu, X. Gao, H. Wu and S. Gao, “Detection of workers without the helments in videos based on YOLO V3,” in 2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Suzhou, China, 2019, pp. 1-4, doi: 10.1109/CISP-BMEI48845.2019.8966045.

[11] A. Bochkovskiy, C-Y. Wang, and H-Y. M. Liao, YOLOv4: Optimal Speed and Accuracy of Object Detection, 2020.