Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
International Conference on Identification, Information and Knowledge in the internet of Things, 2021

AI-based Prevention Embedded System Against COVID-19 in Daily Life

Xuebin Yue, Hengyi Li, Lin Meng*
Department of Electronic and Computer Engineering, Ritsumeikan University,
1-1-1 Nojiigashi, Kusatsu, Shiga, Japan

Abstract

Since the prevalence of COVID-19, the virus has spread all over the world. A large number of people have been infected and died, and countries all over the world have experienced the most severe crisis. Vaccination can effectively resist the virus. However, it does not mean that vaccination can suppress virus spread completely. Hence, wearing a mask correctly and keeping the social distance become emergency methods for reducing the risk of infection. This paper proposes an AI-based prevention embedded system against COVID-19 in daily life by keeping the function of the emergency method. The system consists of two functions, mask-wearing-status detection, and social-distance measurement. Mask-wearing-status detection employs YOLO and realizes the detection and classification of three mask-wearing-status, corrected-wearing, non-corrected-wearing, and without-wearing. Social-distance measurement equips a depth camera for measuring the distance between humans. The system gives an alert when people do not wear a mask correctly or do not keep their social distance. The system has been implemented on Jetson-nano, a compact embedded board, and achieves 6 fps. The experimental results also show that the mask-wearing-status detection accuracy archives at 93.21% and the error of social-distance measurement are within 3 cm, which have proved the effectiveness of the system.

© 2022 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the International Conference on Identification, Information and Knowledge in the Internet of Things, 2021

Keywords: Mask-wearing-status detection; Social-distance measurement; YOLOv4-tiny;

1. Introduction

COVID-19 has swept the world again in 2021. Preventing the spread of the epidemic and vaccination are effective ways to reduce mortality and protect people in safety [1]. However, because of the unbalanced distribution of vaccines, some countries cannot be vaccinated normally. Furthermore, mutant strains of COVID spread diminish the effectiveness of the vaccine. Therefore, preventing the spread of the COVID-19 has become the priority[2].

* Corresponding author.
E-mail address: menglin@fc.ritsumei.ac.jp
Researchers show that wearing a mask and keeping social distance are effective ways to reduce COVID-19 spread and protect people. Some regions and countries stipulate wearing-mask and keeping social distance. However, people often forget these regulations when they enjoy something. With the rapid development of artificial intelligence (AI) technology, the deep learning models have been applied in lots of research fields such as cultural heritage protection [3], environmental protection [4], and so on. It also means the AI technology has become a trend to automatical monitoring mask-wearing status and social-distance [5]. This article proposes an AI-based COVID-19 prevention system, which equips two functions: mask-wearing-status detection, and social-distance measurement. The first function uses YOLOv4-tiny [19] to identify mask-wearing status: corrected-wearing, non-corrected-wearing, or without-wearing. The second function measures the social distances between human bodies. The position information of the human body is obtained by using Intel RealSense Depth Camera D435 [8]. When the social distance between human bodies which are calculated by the system, is less than 1 m, the system gives a prompt warning. Furthermore, for compacting the deep learning models on Jetson-nano, a compact embedded board, the YOLO model is optimized by quantization.

The two major contributions of this paper are as follows:

- Social-distance measurement is added to remind indoor people to keep a certain distance to avoid gathering.
- The trained model is quantified and implemented on Jetson-nano, which can be efficiently deployed in indoor venues such as conference rooms.

The remaining part of this paper is organized as follows. Section 2 introduces the current research on mask-wearing-status detection, as well as the related detection based on YOLO. The proposed AI-based prevention embedded system is introduced in section 3. The experimental conditions and results are shown in section 4. Finally, we conclude the paper in section 5.

2. Related work

In recent years, the target detection algorithm has made great breakthroughs. There are two popular algorithms. One algorithm is based on classification, which obtains the candidate at first and then does classification, such as R-CNN [9], Fast R-CNN [10] [11]. This algorithm has high accuracy, unfortunately, with slow speed; The other algorithm is based on the regression algorithm which does not need to search for candidate regions and directly performs detection and classification, such as YOLO [12], and SSD [13]. This method is fast, only lost a little accuracy. Hence, this paper employs the YOLO for detection.

Since the outbreak of the COVID-19, researchers try to realize mask-wearing-status detection for reducing the COVID-19 spread and protecting people. [14] proposed the RETINAFACE algorithm-based method, which effectively realizes the mask-wearing detection, only judging whether the mask is worn correctly. [1] proposed a novel dataset and two different methods to detect masked or unmasked faces in real-time. These methods are based on the SSD, which introduce a channel attention mechanism to improve the ability of the model to express salient features and strengthen the ability to detect masks [15].

[16] used eight variants of the YOLO algorithm to determine its effectiveness. The results show that original YOLOv4 achieved an mAP value of 71.69% which was the highest among all the original YOLO variants, YOLOv4-tiny achieved an mAP value of 57.71% which was the highest among all tiny variants. It is proved that YOLOv4 and YOLOv4-tiny are efficient in YOLO series algorithms. [17] proposed a new approach to mask-wearing-status detection by replacing Mask-R-CNN with a more efficient model “YOLO” to increase the processing speed of real-time mask-wearing-status detection and not compromise on accuracy.

3. AI-based prevention embedded system

Figure 1 shows the overview of the proposed system which consists of three stages: Primarily, YOLOv4-tiny is trained for generating a mask-wearing-status detection model with three statuses. Furthermore, the Intel RealSense Depth Camera D435 is used to obtain real-time videos and provide in-depth information about indoor personnel. The video information is used for mask-wearing-status detection through a quantified detection model, and depth information is used to monitor the social distance between indoor people. The system is implemented on Jetson-nano.
3.1. Mask-wearing-status detection based on YOLOv4-tiny

YOLO, which regards target detection as a regression problem, is an effective real-time target detection algorithm. In the evaluation process of YOLO, the neural network directly obtains the position, categorizes all of the objects and the corresponding confidence probability from the image with a high speed. YOLOv4 [18], which is based on YOLO target detection architecture and launched recently, adopts the best optimization strategy at present in the field of CNN and finally improves performance and speed.

In practical applications, due to the limited computing resources, many researchers have proposed lightweight target detection algorithms to meet the demand for real-time detection on mobile devices. YOLOv4-tiny [19] is a lightweight model version of YOLOv4. The model takes less computation and memory resources and thus achieves high speed. Since the final model in this study is to be deployed on Jetson-nano, this article takes the lightweight YOLOv4-tiny for model training.

3.2. Social-distance measurement

The Intel RealSense Depth Camera D435 is a stereo solution, which provides the depth value of the object with high precision. The sensor offers quality depth for multiple applications with a range of up to 10 m.

The depth pixel value $Y$, which can be obtained from the depth camera, is the vertical distance from the parallel plane of the sensor to the object and not the absolute range $R$ from the sensor to the object, as Fig. 2 illustrates. The resolution of the depth field of view (FOV) at any distance is $1280(resolution_w) \times 720(resolution_h)$. The degrees of the depth FOV are $\rho_H$ in horizontal and $\rho_V$ in vertical.

Assuming that there are two objects A and B detected as Fig. 3 shows. Let $FOV1(Owh)$ and $FOV2(O'w'h')$ denote the depth FOV where objects A and B are located. The pixel coordinates of the two objects can be expressed as $(w_A, h_A)$ and $(w_B, h_B)$ in $FOV1$ and $FOV2$ respectively, which can be got from the sensor. Let $y_A$ and $y_B$ denote the depth pixel values of objects A and B.

The actual width of $FOV1$ and $FOV2$ can be calculated as follows:

$$FOV1(w) = y_A \times tan\left(\frac{\rho_H}{2}\right) \times 2$$  \hspace{1cm} (1)

$$FOV2(w) = y_B \times tan\left(\frac{\rho_H}{2}\right) \times 2$$  \hspace{1cm} (2)
3.1. Mask-wearing-status detection based on YOLOv4-tiny

YOLO, which regards target detection as a regression problem, is an effective real-time target detection algorithm. In the evaluation process of YOLO, the neural network directly obtains the position, categorizes all of the objects and the corresponding confidence probability from the image with a high speed. YOLOv4 [18], which is based on YOLO target detection architecture and launched recently, adopts the best optimization strategy at present in the field of CNN and finally improves performance and speed.

In practical applications, due to the limited computing resources, many researchers have proposed lightweight target detection algorithms to meet the demand for real-time detection on mobile devices. YOLOv4-tiny [19] is a lightweight model version of YOLOv4. The model takes less computation and memory resources and thus achieves high speed. Since the final model in this study is to be deployed on Jetson-nano, this article takes the lightweight YOLOv4-tiny for model training.

3.2. Social-distance measurement

The Intel RealSense Depth Camera D435 is a stereo solution, which provides the depth value of the object with high precision. The sensor offers quality depth for multiple applications with a range of up to 10 m. The depth pixel value \( Y \), which can be obtained from the depth camera, is the vertical distance from the parallel plane of the sensor to the object and not the absolute range \( R \) from the sensor to the object, as Fig. 2 illustrates. The resolution of the depth field of view (FOV) at any distance is \( 1280 \times 720 \). The degrees of the depth FOV are \( \rho_H \) in horizontal and \( \rho_V \) in vertical.

Assuming that there are two objects A and B detected as Fig. 3 shows. Let FOV_1(\( O_{wh} \)) and FOV_2(\( O'_{w'h'} \)) denote the depth FOV where objects A and B are located. The pixel coordinates of the two objects can be expressed as \((w_A, h_A)\) and \((w_B, h_B)\) respectively, which can be got from the sensor. Let \( y_A \) and \( y_B \) denote the depth pixel values of objects A and B. The actual width of FOV_1 and FOV_2 can be calculated as follows:

\[
x_A = (w_A - \text{resolution}_w) \times \frac{\text{FOV}_1(w)}{\text{resolution}_w}
\]

\[
x_B = (w_B - \text{resolution}_w) \times \frac{\text{FOV}_2(w)}{\text{resolution}_w}
\]

The distance from object A to object B can be converted to the length \( D \) in the XY plane, as Fig. 4 shows. Length \( D \) can be calculated as follows:

\[
D = \sqrt{l_1^2 + l_2^2} = \sqrt{[y_A - y_B]^2 + [x_A - x_B]^2}
\]

3.3. Quantification and deployment

The life cycle of deep learning model development includes five parts: target confirmation, task modeling, data acquisition and annotation, model training, and model deployment. Model deployment is key to the process of model
development. Deploying the trained mask-wearing-status detection model on Jetson is a very critical step. The model trained by YOLOv4-tiny reaches 11 \(fps\) on Jetson-nano. However, in actual operation, higher speed is required to ensure real-time detection. In this paper, Tensor RT introduced by NVIDIA is used to accelerate the reasoning and quantification of the model, which greatly improves the running speed.

Tensor RT is a high-performance neural network inference engine, which is used to deploy deep learning applications for the production environment and can provide maximum inference throughput and efficiency. The engine optimizes the floating-point trained neural networks (usually using 32-bit or 16-bit data) into precision int 8 models with acceptable accuracy loss by quantization, thus achieving the purpose of accelerating the inference process. After the quantization operation, the model reached a speed of 20 \(fps\) on Jetson-nano and increased by 9 \(fps\) compared with the original model.

4. Evaluation

4.1. Dataset

The Face Mask Detection dataset\(^1\) is adopted in this paper. The dataset contains 853 images belonging to 3 classes (corrected-wearing, non-corrected-wearing, and without-wearing). Among them, there are 708 images of corrected-wearing, 59 images of non-corrected-wearing, 86 images of without-wearing. At the same time, there are multiple targets of different categories for some images.

4.2. Experimental result

During the experiment, the distance between two people is measured and compared with the actual distance measured manually. For each distance, two people change with three mask states. The experimental results are compared with the actual status. The results are shown in Table 1 and Table 2.

In Table 1 and Table 2, 0° means horizontal to the sensor direction, and the other angle values indicate the angle of counterclockwise rotation from 0°. Table 1 shows the absolute error between the actual value and the measured value of distance detection.

Table 1. Social-distance detection error(cm).

| Distance | 0.5 | 0.8 | 1  | 1.2 | 1.5 | 1.8 | 2  | 2.5 | 3  | Average |
|----------|-----|-----|----|-----|-----|-----|----|-----|----|---------|
| 0°       | 0.2 | 0.8 | 0  | 2   | 2   | 0.15| 0.2| 1   | 0.13| 0.72    |
| 45°      | 0.8 | 0.5 | 1  | 2.4 | 2   | 1   | 1.3| 0.9 | 1.6 | 1.28    |
| 135°     | 0.4 | 4   | 2.5| 1.2 | 8   | 8.8 | 12 | 10  | 2   | 5.43    |

Table 2. Mask-wearing-status detection accuracy(%).

| Distance | 0.5 | 0.8 | 1  | 1.2 | 1.5 | 1.8 | 2  | 2.5 | 3  | Average |
|----------|-----|-----|----|-----|-----|-----|----|-----|----|---------|
| 0°       | 100 | 100 | 100| 100 | 83.33| 100 | 100| 100 | 100| 98.15   |
| 45°      | 100 | 100 | 100| 100 | 83.33| 83.33| 66.67| 83.33| 66.67| 88.89   |
| 135°     | 100 | 66.67| 83.33| 100 | 83.33| 100 | 100| 100 | 100| 92.59   |

4.3. Discussion and future work

The experimental results show that the detection speed of the model before quantization is 11 \(fps\), and the speed after quantization is increased to 20 \(fps\). With the addition of ranging monitoring, the overall system detection speed is 6 \(fps\), which fits human senses. The system can detect the mask-wearing-status in the crowd and the social distance between people in real-time, with the mask-wearing-status detection accuracy of 93.21% and the social-distance

\(^1\) https://www.kaggle.com/andrewmvd/face-mask-detection
measurement error is no more than 3 cm. In future work, this research plans to construct a new dataset that can detect the distance between people when the human body is facing away from the detection device.

5. Conclusion

At present, COVID-19 is spreading faster and faster. The infected population has broken through historical data day by day and hit a new height. Effective ways to prevent the virus from spreading include vaccines, wearing masks, being less outgoing, and gathering. This paper develops an AI-based prevention embedded system against COVID-19 in daily life. Initially, YOLOv4-tiny is adopted as the mask-wearing-status detection model and trained with three states. Additionally, Intel RealSense Depth Camera D435 is used to get real-time videos and provides in-depth information about indoor people as well. Video information is used to effectively detect mask-wearing status. The depth information is used to monitor the social distance of indoor people. The system makes an alert in cases that the distance between the crowd is less than 1 m, someone does not wear a mask or wears a mask incorrectly. Eventually, the mask-wearing-status detection model is quantified by TensorRT and deployed on Jetson-nano. The system can detect the mask-wearing status in the crowd with high accuracy of 93.21% and the social distance between people in real-time with an error of less than 3 cm, and can be deployed conveniently and quickly in the process of resuming work and production.

References

[1] Sahand Abbasi, Haniyeh Abdi, and Ali Ahmadi. (2021) “A Face-Mask Detection Approach based on YOLO Applied for a New Collected Dataset.” 2021 26th International Conference, Computer Society of Iran (CSICC) : 1-6.
[2] Steffen E. Eikenberry, Marina Mancuso, et al. (2020) “To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic.” Infectious Disease Modelling 5: 293-308.
[3] Lin Meng, Bing Lyu, Zihiu Zhang, C.V. Aravinda, Naoto Kamitoku, Katsuhiro Yamazaki. “Ocracle Bone Inscription Detector Based on SSD” International Conference on IMAGE ANALYSIS AND PROCESSING (ICIAP2019):126-136
[4] Lin Meng, Hirayama Takuma and Oyanagi Shigeru. (2018) “Underwater-Drone With Panoramic Camera for Automatic Fish Recognition Based on Deep Learning” IEEE ACCESS, Vol.6, No.1:17880-17886
[5] Gokul Kumar, and Sujala Shetty. (2021) “Application Development for Mask Detection and Social Distancing Violation Detection using Convolutional Neural Networks.” In Proceedings of the 23rd International Conference on Enterprise Information Systems 1: 760-767.
[6] Minxuan Wen, Kentaro Yokoo, Xuebin Yue and Lin Meng. (2021) ‘An AI-based Mask-Wearing Status Recognition and Person Identification System” 2021 International Symposium on Advanced Technologies and Applications in the Internet of Things (ATIAT).
[7] Ala A Alsanabani, Salah A Saeed, Moeen Al-Mkhafi, and Mohammed Albishari. (2021) “A Low Cost and Real Time Vehicle Detection Using Enhanced YOLOv4-Tiny.” 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA) : 372-377.
[8] Min Sung Ahn, Hosik Chae, Donghun Hoh, Hyunwoo Nam, and Dennis Hong. (2019) “Analysis and Noise Modeling of the Intel RealSense D435 for Mobile Robots.” 2019 16th International Conference on Ubiquitous Robots (UR): 707-711.
[9] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. (2013) “Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation.” Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition
[10] Bin Liu, Wencang Zhao, and Qiaoqiao Sun. (2017) “Study of object detection based on Faster R-CNN.” 2017 Chinese Automation Congress (CAC) : 6233-6236.
[11] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. (2017) “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.” IEEE Transactions on Pattern Analysis and Machine Intelligence 39 : 1137-1149.
[12] Joseph Redmon, S. Divvala, Ross B. Girshick, and Ali Farhadi. (2016) “You Only Look Once: Unified, Real-Time Object Detection.” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) : 779-788.
[13] Wei Liu, Dragomir Anguelov, Dumitru Erhan, et al. (2016) “SSD: Single Shot MultiBox Detector” Computer Vision – ECCV 2016:21–37
[14] Bin Xue, Jianpeng Hu, and Pengming Zhang. (2020) “Intelligent detection and recognition system for mask wearing based on improved RetinaFace algorithm.” 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDI) : 474-479.
[15] Mingyuan Xu, Heng Wang, Shuqun Yang, and Rui Li. (2020) “Mask wearing detection method based on SSD-Mask algorithm.” 2020 International Conference on Computer Science and Management Technology (ICCSMT) : 138-143.
[16] Akhil Kumar, Arvind Kalia, Kinshuk Verma, Akashdeep Sharma, and Manisha Kaushal. (2021) “Scaling up face masks detection with YOLO on a novel dataset.” Optik 239 : 166744.
[17] Ren Liu, and Ziang Ren. (2021). “Application of Yolo on Mask Detection Task.”
[18] Alexey Bochkovskiy, Chienyao Wang, and Hongyuan Liao. (2020) “YOLov4: Optimal Speed and Accuracy of Object Detection.”
[19] Donglin Zhu, Guanghui Xu, Jie Zhou, Enbiao Di, and Mingcan Li. (2021) “Object Detection in Complex Road Scenarios: Improved YOLOv4-Tiny Algorithm.” 2021 2nd Information Communication Technologies Conference (ICTC) : 75-80.