AI-Based Low-Cost Real-Time Face Mask Detection and Health Status Monitoring System for COVID-19 Prevention

CHOON EN YOU, WAI LEONG PANG, KAH YOONG CHAN
Faculty of Engineering,
Multimedia University,
63100 Cyberjaya,
MALAYSIA

Abstract: - The outbreak of COVID-19 had brought a great challenge for the World Health Organization (WHO) in preventing the spreading of SARS-CoV-2. The Ministry of Health (MOH) of Malaysia introduced the MySejahtera mobile application for health monitoring and contact tracing. Wearing a face mask in public areas had been made compulsory by the Government. The overhead cost incurred in hiring the extra manpower to ensure all the visitors wear a face mask, check-in through MySejahtera and the status in MySejahtera is healthy before entering a premise. A low-cost solution is urgently needed to reduce the heavy overhead cost. An AI-Based Low-Cost Real-Time Face Mask Detection and Health Status Monitoring System (AFMHS) is proposed to perform real-time detection for the face mask and MySejahtera Check-In tickets by using artificial intelligence. MobileNetV2 was used for the detection and recognition of face and face masks. YOLOv3 was used for the detection of the region of interest for the MySejahtera Check-In ticket to locate the health and vaccination status of the visitor. Optical character recognition (OCR) is a technique that is used to detect the text captured in an image and encode the recognized text. OCR is implemented to recognize the text extracted from the ticket. Tesseract is used as the OCR engine in AFMHS. Raspberry-Pi-4B (Raspberry Pi Generation 4 Model B) with 4 GB RAM is used as the processing unit of AFMHS. The total cost of the AFMHS is only USD220. Extensive experimental tests were carried out to evaluate the performance of AFMHS. The optimum operation setup conditions are proposed to achieve 100% accuracy. The optimum operating distance for the face mask detector and MySejahtera Check-In ticket detector are 1.5m and 15cm respectively.

Key-Words: - AI, COVID-19, Face mask detection, Machine Learning

Received: June 13, 2021. Revised: September 8, 2022. Accepted: October 8, 2022. Published: November 4, 2022.

1 Introduction

On 11 March 2020, the WHO declared a pandemic that had been caused by the coronavirus disease namely COVID-19. COVID-19 caused the severe acute respiratory syndrome. The pandemic forced the Government of Malaysia to take action to prevent the spreading of the COVID-19 disease. The Government of Malaysia had implemented several standard operation procedures (SOP) to fight the pandemic, i.e. wearing a face mask in public, and introduced the MySejahtera mobile application for contact tracing and health status. MySejahtera assists in suppressing the COVID-19 outbreak by helping users in monitoring their health conditions. Users could report their health status when COVID-19 symptoms were detected. The detailed information of the users that are infected by COVID-19 is recorded and the COVID-19 risk status is updated accordingly in the application.

MySejahtera Check-In function is important for all types of public premises such as restaurants, shops, companies, schools and construction sites. All public premise owners need to display the QR code generated by MySejahtera. Visitors are required to use the MySejahtera QR Code Scanner to scan the QR code displayed before entering any public premises. Upon scanning, a ticket will be generated whereby information such as the name of public premises, name of the individual who entered the public premise, dates and phone number will be listed. The health risk and vaccination status of the visitor would also be shown. Only the visitor with the status of low risk (no symptoms) and fully vaccinated could enter the premises. This is another mandatory rule set by MOH.

The flow of people that would pass in and out of public premises such as schools, shopping malls, shops and restaurants is high. It would increase the difficulty of the premise owner to ensure that the people that enter the premises are healthy, and checked in through MySejahtera with a healthy and vaccinated status. It is a huge challenge to guarantee that the visitors would wear the face mask correctly and have scanned the QR code before entering the premises. The premises' owners hired security guards to ensure that all the visitors are wearing face masks, with normal body temperature and also check the health and vaccination status in the MySejahtera of the visitors at the entrance of the premise. Only visitors with normal body temperature and healthy MjSejahtera status are allowed to enter the premises. It could be a huge financial burden to the premise owner as the extra cost is incurred. From the perspective of an economical view, this effect is greatly obvious if several security guards are being hired to monitor the premise. These have inspired us to propose an AI-Based Low-Cost Real-Time Face Mask Detection and Health Status Monitoring System (AFMHS).

The project objectives are to develop a real-time alert system for the face mask detector and MySejahtera Check-In ticket detector using artificial intelligence. The MobileNetV2 algorithm is used in the face mask detector [3]. Optical character recognition (OCR) is incorporated with the YOLOv3 object detection algorithm to detect the specific targeted character from the MySejahtera Check-In ticket [4]. The AFMHS allows the visitor who wears a face mask and has a record of low risk with fully vaccinated on the MySejahtera Check-In ticket to enter the premises.

2 Literature Review

In the year 2021, the researcher Fushuai Wang et al did a research project on face recognition with MobileNetV2. The processing unit proposed is on Raspberry Pi 4B, [1]. The author mentioned that MobileNetV2 could provide a balance between the accuracy and the network parameter, which allows it to be suitable for usage on mobile devices. The face images were recognized in a total of 5 categories. 100 face images are used in the training set which is 500 images in total. 50 face images in the test set, which means 250 images in total, [1].

In the year 2021, the researcher Ikram ben Abdel Ouahab et al developed a real-time face mask detector with MobilNetV2 and Raspberry Pi, [2]. The face mask model was trained, tested and evaluated through Google Colab with a huge database available online on GitHub. The dataset consists of a total of 3835 images with 1915 images with masks and 1918 images without mask. The selected database consists of people wearing masks with different poses and positions, [2]. The approach of this project work was more towards the ability to perform in real-time whereby speed is a crucial factor. A smooth motion and appearance of the video stream were considered. When the real-time video stream is analysed by the detection model, the frame per second (FPS) value will decrease. The decrease in FPS will be even more obvious in devices with lower processing power. From this research work, the lower processing power device such as Raspberry-Pi-4B could cater for the performance requirement. MobileNetV2 could also be deployed due to the nature of its lightweight model.

In the year 2021, the researcher Samuel Ady Sanjaya et al completed research work on face mask detection by using MobileNetV2, [3]. In this research paper, the author mentioned that face mask recognition was being implemented with the image classification method through MobileNetV2. The proposed face mask detection model was referring to the datasets taken from the Kaggle dataset and the Real-World Masked Face dataset. These datasets consist of 5 thousand masked faces and 90 thousand normal faces photos.

Extra work is carried out in the year 2022 to perform face mask detection. The enhanced Yolo algorithm was proposed to detect the face mask, [9]. Experimental work was carried out to evaluate the performance of the algorithm proposed and this shows the importance of the solution in face mask detection. Convolutional Neural Network was used to monitor the face mask of the workers on the construction site to ensure the workers' safety, [10]. The solution proposed also monitors the physical distance between the workers. The performance of various approaches in face mask detection was reviewed in [11]. However, no related work is reported in monitoring the health status of MySejahtera. An effective solution is urgently needed to monitor the health status in MySejahtera of the people before they enter the premises.

In the year 2021, the researcher R Shashidhar et al produced a project regarding the detection and recognition of vehicle number plates with the method of OCR through YOLOv3, [4]. Since there are different background colours and types of license plate, the YOLOv3 model was trained to localize the vehicle number plate. It is particularly mentioned that YOLO had been used to find the region of interest which is the part of the vehicle number plate only [4]. A dataset that consisted of 6439 images of different alphabet-numerical characters was created. The result showed that an accuracy of 91.5% was obtained for this vehicle number plate detection model ,[4].

In the year 2020, the researcher Chinmaya Kumar Sahu et al did research work on a comparative analysis of the deep learning approach regarding YOLOv3 for real-time recognition of vehicle number plates, [5]. 2500 images of cars that consist of foreign number plates and Indian license plate numbers had been prepared. The images will be labelled with a desktop application namely Labeling. An Extensible Markup Language (XML) file is created for each labelled image. The region of interest will be labelled with a bounding box. Upon labelled, the coordinate and the object class name of the region of interest that contained the target outcome are stored in the xml text file. The coordinate stored in the xml file is compatible with the Darknet Config. Darknet Config data architecture is part of the YOLOv3 model. After the annotation and labelling of the dataset, the YOLOv3 model was trained for the localization of vehicle number plate detection.

In the year 2022, the researcher Srividya Subramanian et al proposed a text extraction model with YOLOV3, [6]. YOLOv3 has been implemented to localise and classify the characters from images. Newspapers and books in the form of digital were collected and used as the dataset. There were a total of 80 distinct characters being used. An annotation process had been carried out to draw the bounding box over the characters. An annotation file that contained the coordinate of the location of the character had been created.

A comparative review had been made on 4 different OCR engines. The researcher Ahmad P. Tafti et al had done a comparative review to evaluate the OCR engine between Google Docs OCR, Tesseract, ABBYY FineReader as well as Transym, [7]. From this research work, experimental evaluation in the form of qualitative and quantitative approaches had been made using these OCR engines and services. In order to perform the experimental comparison, 1227 images from 15 different categories had been used, [7]. The main focus will be on the category of machine-written characters, machine-written digits, noisy images and blurred images. Tesseract is a good choice since the service provided is free of charge.

3 Design of the AFMHS

Fig. 1 shows the two main components of the AFMHS, i.e. the face mask detector and the MySejahtera Check-In ticket detector. The object detection algorithm is modelled with the python programming language. OpenCV is a Python library for image processing to perform face recognition and object detection. OpenCV has been installed in Python to perform face detection and image processing. Imutils is used for image processing and the functions of imutils include rotation and resizing of images. The images will be preprocessed by Imutils before being accessed by the OpenCV. Tensorflow is used for machine learning applications.

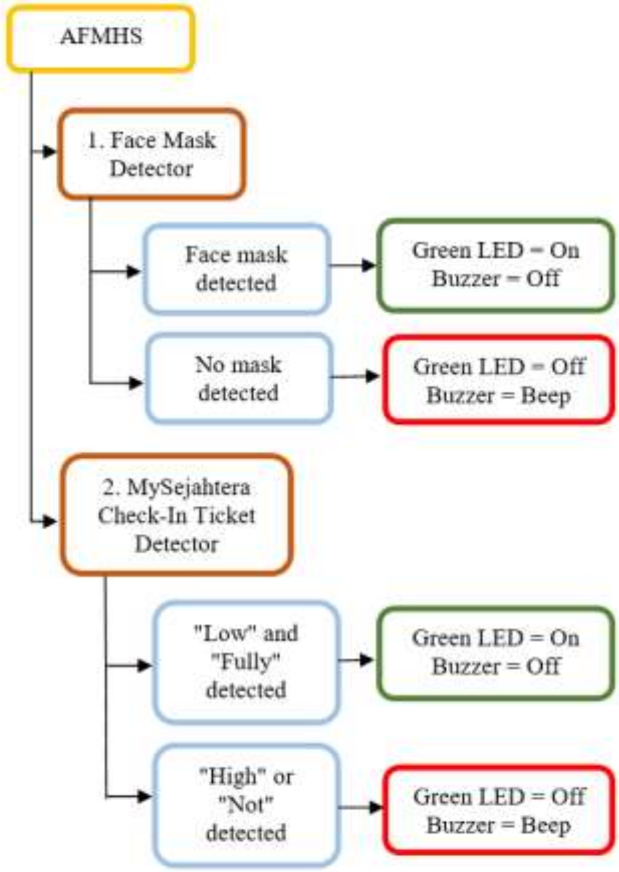


Fig. 1: Two main components of the AFMHS

3.1 Face Mask Detector

The dataset used consists of images of one individual and more than one individual. Different conditions of content are covered in the dataset. Some of the images are slightly blurred and the pixel is not as high. The pictures are taken in various positions and different poses. The images of the users wearing the different types of face masks are collected as well. There are various types and colours of masks collected in the dataset including normal medical face masks and N-95 face masks. A dataset that is rich in a variant of information that is reasonable and related is preferred.

Scikit-learn is used for image classification and regression. The dataset consists of two groups of images. One group of images for faces with masks is saved in a file namely with_mask. Another group of images for faces without masks is saved in a file namely without_mask. The number of images for the file “with_mask” is 504 images and 430 images for "without_mask" which made out a total of 934 images.

The dataset is used by the MobileNetV2 neural network for the training process. All the images had been resized to 224 x 224 pixels. Next, another pre-processing step is the conversion of the images into the NumPy array format. All the images of the dataset had been scaled to the range of -1 to 1. The preprocessing step is to perform one-hot encoding on the labels. 80% of the images from the dataset were split into training and another 20% for testing. The MobileNet model is loaded together with pre-trained ImageNet weights.

The FaceNet model is implemented for face detection. The installed and imported packages are used to grab the dimension of the frame. A blob is created from the frame. Afterwards, the FaceNet model will be deployed for face detection when the created blob is being passed through the network. The confidence in the detection is extracted. Detections with confidence levels lower than 0.5 are filtered out and excluded. Next, the x and y coordinates of the bounding box for the detected object are calculated. The frames from the real-time streaming video are looped over.

Detection of the faces in the frame is carried out. The face in the bounding box is processed and determined whether the face is wearing a face mask or not. The predictions for both labels are compared and displayed on top of the bounding box. The label is either "Mask Detected" on top of the green colour bounding box or "No Mask Detected" on top of the red colour bounding box.

3.2 MySejahtera Check-In Ticket Detector

YOLOv3 will be used as the object detection model with the backbone of Darknet-53. A dataset that consists of the images taken from the MySejahtera Application. A sample of the MySejahtera Check-In Ticket is shown in Fig. 2. A MySejahtera Check-In ticket consists of the information regarding the name of the premise, name, phone number, date, time, risk status and vaccination status. All this information is collected. The main focus is put on the section on risk status and vaccination status. A total of 210 images had been included in the dataset. Images of various conditions regarding the risk and vaccination status of the MySejahtera Check-In ticket had been collected.

Fig. 2: Annotation of the MySejahtera Check-In Ticket image by labelImg

There are unnecessary texts from the ticket, hence it is important to perform annotations to set the region of interest. In order to perform annotation, labelImg (a graphical image annotation tool) is used to draw the bounding box for identifying the area of interest as shown in Fig. 2. The regions of interest are the risk status and vaccination status. However, other information such as name, date, time, location, and phone number are annotated as well for future research. There are a total of 7 bounding boxes drawn for each image. An annotation text file is generated and it consists of the coordinates of the bounding boxes. The format of the position of the bounding box is as follows, <class name> <x coordinate of centre> <y coordinate of centre> <width> <height>.

Google Colab is used for the training process. All the datasets including the images and all the annotation files are uploaded to Google Drive and directed to Google Colab. 80% of data are used for training and another 20% for testing. YOLOv3 with backbone Darkent-53 is used for the training. Each training weight is saved after each epoch. The best weight that has the lowest loss is updated and saved.

Tesseract OCR is used as the character recognition engine. Gray-scaling of the images before OCR detection is implemented. The purpose of this preprocessing step is to make sure that less information is needed for each pixel. It speeds up the recognition of the character. The MySejahtera Check-In ticket detector mainly searches for the word “Fully”, “Low”, “High” and “Not” from the text file.

3.3 The Operation of the AFMHS

A microcontroller is required to deploy the AFMHS and a camera module is needed to capture the video. AFMHS is deployed in a microcontroller, Raspberry-Pi-4B (Raspberry Pi Generation 4 Model B) with 4 GB random access memory. AFMHS used the camera of a Samsung S9 smartphone to capture the video through the DroidCam software, [8]. Droidcam is an application that would provide the service of using a smartphone as a webcam. The Samsung S9 smartphone consists of an 8MP front camera and a 12MP rear camera which can capture video with good resolution for AFMHS. An alert system is proposed in AFMHS. This system consists of a green LED and a buzzer. The green LED acts as an indicator for the user that fulfills the requirement (with face mask, healthy and vaccinated status) to pass through while the buzzer acts as an alarm system to alert those who are not fulfilling the requirement.

Fig. 3 shows the system operation flowchart of the AFMHS proposed. AFMHS consists of two functions, i.e. 1) Face mask detector and 2) MySejahtera Check-In Ticket Detector. The face mask detector captures the video through the camera. The green LED is on once the visitor is wearing the face mask and the system will show “Mask Detected” on the screen. However, AFMHS will show “No Mask Detected”, the green LED is turned off and the buzzer is beeping when the visitor is not wearing a face mask detected.

The second function of AFMHS is MySejahtera Check-In Ticket Detector. AFMHS analyses the risk and vaccination status shown in MySejahtera Check-In Ticket (Fig. 2). Visitors with “Low” risk and “Fully” vaccinated status are granted to enter the premises. The green LED is turned on to indicate that the visitor is healthy, fully vaccinated and safe to enter the premises. The buzzer will beep once “High” risk or “Not” vaccinated are detected. AFMHS alerts the authority that the visitor with either at high risk or not vaccinated is detected and is not allowed to enter the premises.

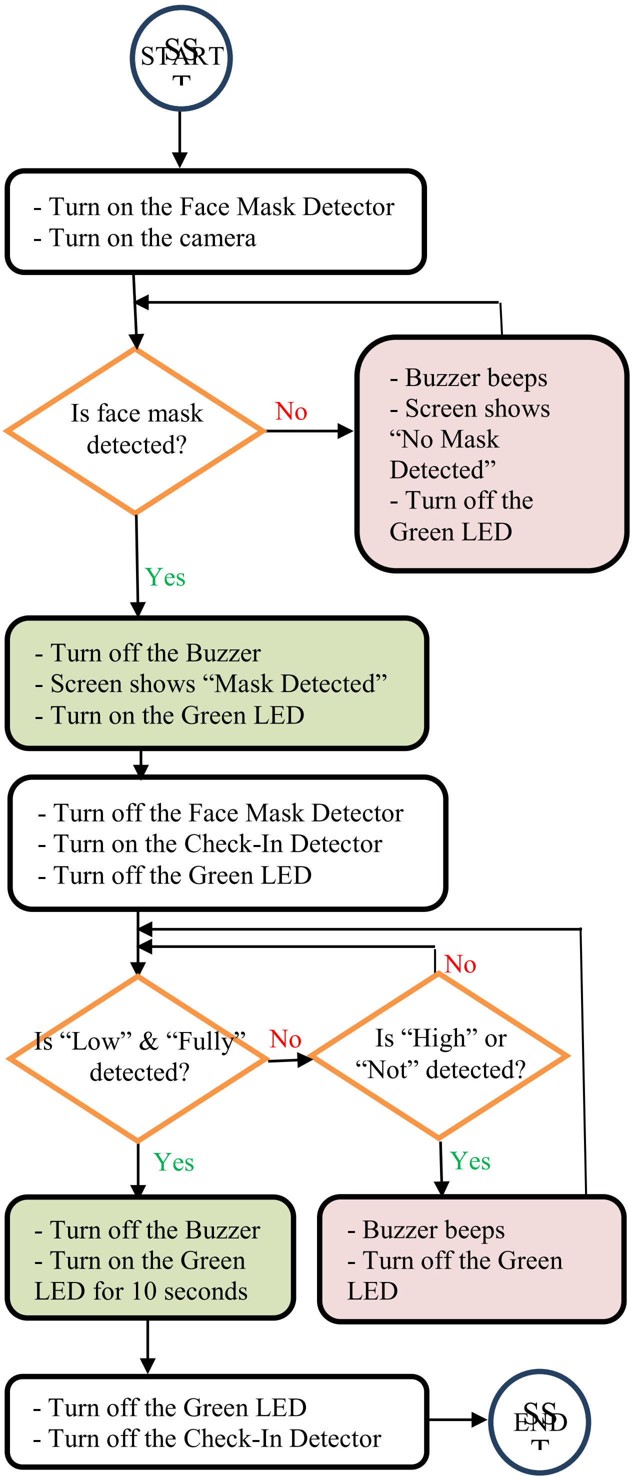


Fig. 3: The operation flowchart of AFMHS

4 Results and Discussion

Extensive experimental tests are carried out to determine the accuracy of the face mask detector with the distance between the individual and camera as the fixed variables and the light intensity of the environment as manipulating variables. Fixed distances of 1m, 1.5m, 2m, 2.5m, and 3m are used. The light intensities of 40~60 Lux, 90~120 Lux, 200~220 Lux, and 300~320 Lux are used. Three different types of face masks are used which are the black mask, the green mask, and the N95 white mask. AFMHS is tested for the accuracy of detecting no mask as well.

Table 1. Accuracy with different light intensities

Distance (m)	Light intensity (Lux)	Accuracy			
		Black	Green	White (N95)	No mask
1.0	300~320	100%	100%	100%	100%
	200~220	100%	100%	100%	100%
	90~120	100%	100%	100%	100%
	40~60	100%	100%	100%	100%
1.5	300~320	100%	100%	100%	100%
	200~220	100%	100%	100%	100%
	90~120	100%	100%	100%	100%
	40~60	100%	100%	100%	100%
2.0	300~320	60%	100%	100%	100%
	200~220	50%	100%	100%	100%
	90~120	50%	100%	100%	100%
	40~60	40%	100%	100%	100%
2.5	300~320	60%	100%	100%	100%
	200~220	40%	80%	100%	100%
	90~120	40%	60%	100%	100%
	40~60	10%	60%	80%	80%
3.0	300~320	20%	20%	40%	40%
	200~220	0%	0%	20%	20%
	90~120	0%	0%	20%	20%
	40~60	0%	0%	20%	20%

Table 2. Accuracy of the MySejahtera Check-In ticket detector with fixed light intensity (60 Lux) and anti-glare hood

Distance (cm)	Accuracy
8	10/10 = 100%
12	10/10 = 100%
15	10/10 = 100%
20	1/10 = 10%
25	0/10 = 0%

In order to test the functionality and limit of the prototype, several parameters and conditions had been evaluated. For each parameter, the experiment is repeated 10 times. The measured output results are recorded. The output is correct when it is the same as the expected output. The output is wrong when it is different from the expected output or no output had been produced. For each parameter, the accuracy rate is based on the correctness of the output result of AFMHS measured.

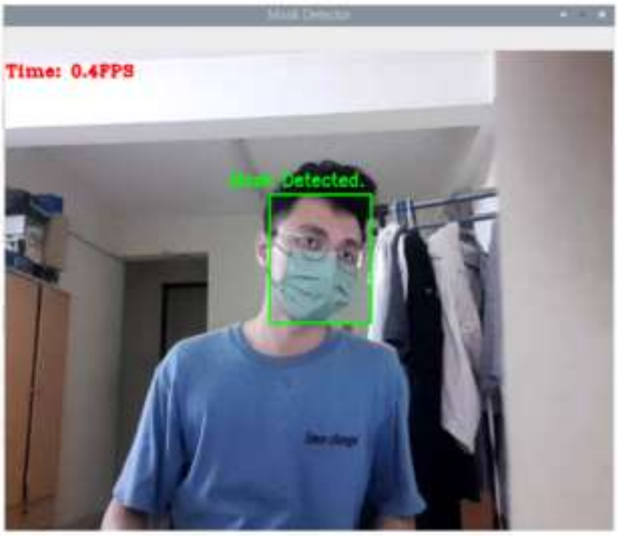
As shown in Table 1, the longer the distance, the lower the accuracy. The distance of 1m and 1.5m is the optimum distance value as the performance in terms of accuracy for all categories is high. The lower the light intensity, the lower the accuracy. The optimum light intensity is 300~320 Lux. Moreover, the face mask detector had been tested for the incorrect way of wearing the mask with a fixed light intensity of 300~320 Lux and the distance is fixed at 1.5m. The accuracies across all categories of masks are between 80% to 90%.

Fig. 4 shows the results of the face mask detector. Fig. 4(a) shows the figure when the face mask is detected. A green colour bounding box with “Mask Detected” on top of the box is shown in the figure when the face mask is detected. Fig. 4(b) shows the figure when no mask is detected. A red colour bounding box with “No Mask Detected” on top of the box is shown in the figure when no mask is detected.

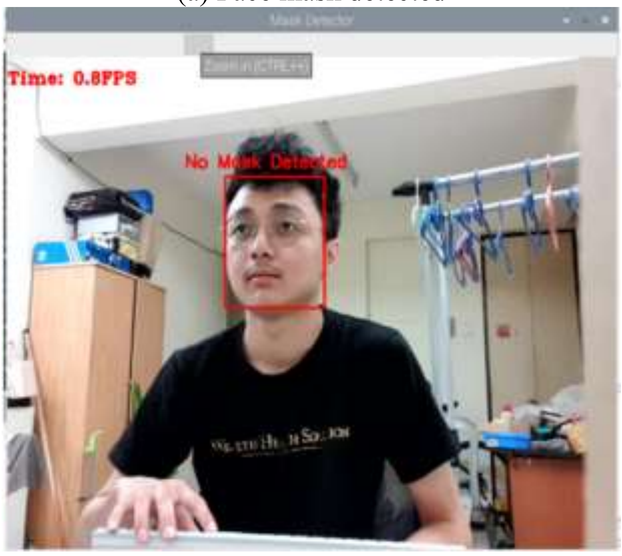
The glare on the handphone screen affected the accuracy level of the detection. The glare on the screen came from the reflection of the light source. The mitigation is to build an anti-glare camera hood to place the camera device in order to enhance the AFMHS performance. The purpose of the anti-glare camera hood is to limit the glare on the handphone screen in order to enhance the image quality of the MySejahtera Check-In Ticket.

Without the anti-glare hood, the light intensity around the screen of the phone is high. With a fixed distance of 0.08m between the camera and the screen of the user’s phone that displays the MySejahtera Check-In ticket, the light intensities of 150 Lux, 200 Lux, 300 Lux, and 500 Lux were used and the accuracies were 80%, 80%, 60%, 50% and 30% respectively. When an anti-glare hood was implemented to avoid the reflection of light and glare on the phone’s screen, the accuracy of the detector increased. With the anti-glare hood, the light intensity around the phone’s screen is reduced to the range of 20 to 60 Lux. Although the light intensity is low, the recorded accuracy of the detector is excellent with an accuracy of 100%.

Furthermore, with an anti-glare hood, the distance of 8cm, 12cm, 15cm, 20cm and 25cm between the camera and the user’s phone and the light intensity of 60 Lux were used to evaluate the performance of the MySejahtera Check-In ticket detector. Fig. 5 shows the results of the MySejahtera Check-In ticket detector. As shown in Table 2, it was observed that the longer the distance, the lower the accuracy. The accuracy starts to drop when the distance is more than 15cm. Hence, the optimum distance is between 8cm to 15cm. It can be deduced that in order to achieve high accuracy for the MySejahtera Check-In ticket detector, the distance should not be more than 15cm and the light intensity of the environment should keep at least around 20~60 Lux with the anti-glare hood.



(a) Face mask detected



(b) No mask detected

Fig. 4: Results of face mask detector

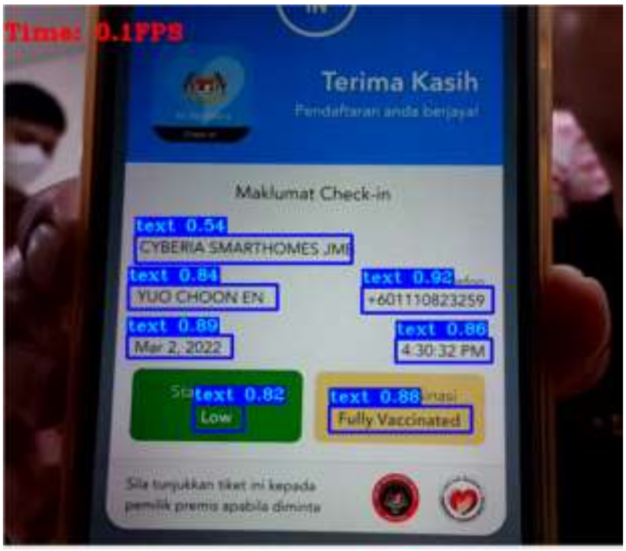


Fig. 5: Results of the MySejahtera Check-In ticket detector

5 Conclusion

The AFMHS proposed is a perfect solution to ensure the users wearing a face mask and the health status in MySejahtera are healthy and vaccinated. No similar work is reported in monitoring the health and vaccination status in MySejahtera. An effective and low-cost AFMHS is urgently needed. AFMHS consists of two major functions, i.e. face mask detector and MySejahtera Check-In ticket detector. Extensive experimental works are carried out to enhance the performance of AFMHS. The anti-glare hood is proposed to improve the accuracy of the AFMHS. The optimum conditions for the face mask detector are fixed light intensity of 300~320 Lux and the distance is fixed at 1.5m. The optimum experimental setup of AFMHS is proposed to enhance its performance. The distance between the camera and the user's phone should not be more than 15cm and the light intensity of the environment should keep at least around 20~60 Lux for the MySejahtera Check-In ticket detector. The accuracy of the AFMHS is 100%. The total cost of the AFMHS is only USD220. It is a promising solution that is ready to be deployed and is more economical compared to the current practice in which extra manpower is used to ensure the visitors are wearing face masks and the MySejahtera status is healthy before entering a premise.

Acknowledgement:
This research was funded by the Internal Research Fund, Multimedia University, Malaysia.

References:

[1] F. Wang, R. Zheng, P. Li, H. Song, D. Du and J. Sun, Face recognition on Raspberry Pi based on MobileNetV2, *2021 International Symposium on Artificial Intelligence and its Application on Media (ISAIAM)*, Xi'an, 2021, pp. 116–120.

[2] O. Abdel, L. Elaachak, M. Bouhorma and Y. A. Alluhaidan, Real-time Facemask Detector using Deep Learning and Raspberry Pi, *2021 International Conference on Digital Age Technological Advances for Sustainable Development (ICDATA)*, Marrakech, 2021, pp. 23–30.

[3] S. A. Sanjaya and S. R. Adi, Face Mask Detection Using MobileNetV2 in The Era of COVID-19 Pandemic, *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*, University of Bahrain, 2020, pp. 1–5.

[4] R. Shashidhar, A. S. Manjunath, R. K. Santhosh, M. Roopa and S. B. Puneeth, Vehicle Number Plate Detection and Recognition using YOLO- V3 and OCR Method, *2021 IEEE International Conference on Mobile Networks and Wireless Communications*, Karnataka, 2021, pp. 1–5.

[5] C. K. Sahu, S. B. Pattnayak, S. Behera and M. R. Mohanty, A Comparative Analysis of Deep Learning Approach for Automatic Number Plate Recognition, *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Palladam, 2020, pp. 932–937.

[6] S. Subramanian, V. Kekatpure, G. Raymond, K. Parab, S. Dugad and A. Shirke, TEYSuR - Text Extraction with YOLO and Super Resolution, *2022 International Conference for Advancement in Technology (ICONAT)*, Maharashtra, 2022, pp. 1–7.

[7] A. P. Tafti, A. Baghaie, M. Assefi, H. R. Arabnia, Z. Yu and P. Peissig, OCR as a Service: An Experimental Evaluation of Google Docs OCR, Tesseract, ABBYY FineReader, and Transym, *Lecture Notes in Computer Science*, Vol. 10072, 2016, pp. 735–746.

[8] “Dev47Apps: DroidCam - Use your phone as a webcam!” https://www.dev47apps.com/ (accessed Oct. 28, 2022).

[9] P. Wu, H. Li, N. Zeng and F. Li, FMD-Yolo: An efficient face mask detection method for COVID-19 prevention and control in public, *Image and Vision Computing*, Vol. 117, 2022, pp. 1-10.

[10] M. Razavi, H. Alikhani, V. Janfaza, B. Sadeghi and E. Alikhani, An Automatic System to Monitor the Physical Distance and Face Mask Wearing of Construction Workers in COVID-19 Pandemic, *SN Computer Science*, Vol. 3, 2022, pp. 1-8.

[11] Vibhuti, N. Jindal, H. Singh and P. Rana, Face mask detection in COVID-19: a strategic review, *Multimedia Tools and Applications*, Vol. 81, 2022, pp. 40013–40042.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

Choon En Yuo carried out the simulation, experimental test and optimization.
Wai Leong Pang has carried out the conceptualization of the project and carried out the formal analysis, validation, supervision, review and edit of the paper.
Kah Yoong Chan has carried out the validation, project management, review and edit of the paper.