Portable Face Mask Detector Using Jetson Nano and USB Webcam QHD Based on OpenCV

Hadi Supriyanto, Ruminto Subekti, and Adri Almas Afwansyah
Department of Automation and Mechatronics Engineering,
Bandung Polytechnic of Manufacturing, Bandung, Indonesia

Keywords: Face Mask Detector, Yolov4-tiny, Computer Vision, Object Detection.

Abstract: ARI (Acute Respiratory Infection) is still a major cause of infectious disease morbidity and mortality in this world. The mortality rate for ARI reaches 4.25 million every year in the world. Similar to ARI, Covid-19 is also a contagious disease. By implementing health protocols, especially the use of masks, can reduce the risk of contagious disease transmission so that all activities can run well. The portable face mask detector is needed to prevent ARI and Covid-19. It can monitor and ensure that mask-wearing rule is followed by each individual. The face mask detector can work with the YOLO algorithm. The research was conducted using a USB webcam to capture images, a USB speaker, and a Jetson Nano as an image processing device with the Yolov4-tiny algorithm to detect masked and non-masked objects. The developed system successfully detects the use of masks when the distance between the object and the camera varies from 2.68 m to 5.38 m and the position of the object's face is 110˚ (right to left) and 83.33˚ (up and down). The system alerts each time a subject is not wearing a mask. The system can detect masked and non-masked objects with an error value of 0,67%.

1 INTRODUCTION

ARI (Acute Respiratory Infection) is still a major cause of infectious disease morbidity and mortality in this world. The mortality rate for ARI reaches 4.25 million every year in the world. Based on data from the World Health Organization (WHO) in 2019 lower respiratory tract reduces life expectancy by 2.09 years in sufferers. (World Health Organization, 2019) The group most at risk is toddlers. About 20-40% are hospitalized among children due to ARI with about 1.6 million deaths due to pneumonia itself in children under five per year. In adults, the mortality rate (25-59 years old) reached 1.65 million. (Susilowati et al., 2021)

The incidence of ARI is influenced by intrinsic and extrinsic factors. Intrinsic Factor includes age, breastfeeding, nutritional status, low birth weight, nutritional status immunization. While the extrinsic factors include knowledge, educational factors, occupancy density, physical condition of the house, house ventilation, cigarette smoke, socio-economic and profession (Susilowati et al., 2021). Environmental influences such as forest fire smoke, dust, and other air pollution can make the situation worse. Similar to ARI, Covid-19 is also a contagious disease.

Coronavirus (SARS-CoV-2) is a virus that can attack the respiratory system (Wu et al., 2020). This virus has evolved to produce new virus variants. These variants include delta and omicron. From the Weekly Epidemiological Update on Covid-19 published by the World Health Organization (WHO) on 28 December 2021, it appears that the risk of the new variant of concern, Omicron, remains high. Evidence found suggests that this variant has a growth advantage over the delta variant with a doubling time of 2-3 days, leading to an increase in cases of virus infection in several countries (World Health Organization, 2021).

Early prevention for ARI and Covid-19 can be done by maintaining physical distance, mask-wearing, washing hands regularly with soap and water or hand sanitizer, not touching the facial area before washing hands, and increasing stamina through a healthy lifestyle (Fatia et al., 2020).

A person can contract contagious disease by accidentally inhaling droplets that come out when a sufferer coughs or sneezes, by touching the facial area without washing hands first after touching an object

Supriyanto, H., Subekti, R. and Afwansyah, A.
Portable Face Mask Detector Using Jetson Nano and USB Webcam QHD Based on OpenCV.
DOI: 10.5220/0011811300003575
In *Proceedings of the 5th International Conference on Applied Science and Technology on Engineering Science (iCAST-ES 2022)*, pages 409-416
ISBN: 978-989-758-619-4; ISSN: 2975-8246

that has been exposed to a splash of saliva from a sufferer, and through contact with a sufferer (Jayaweera et al., 2020).

To implement a portable face mask detector, an object recognition algorithm is required. Various methods such as Convolutional Neural Network, You Only Look Once (YOLO), Cascade Classifier and others can be used as object recognition methods. Based on some of these methods, it can be determined which method is suitable and effective for object recognition by considering its speed and accuracy.

Based on this concept, the face mask detector can use the YOLO algorithm. This algorithm is a real-time object detection system. It was first introduced by Redmon and friends in 2015. Their article entitled You Only Look Once: Unified, Real-Time Object Detection describes an object detection system capable of detecting objects in real-time up to 45 FPS on hardware with a graphics processing unit (GPU) (Redmon et al., n.d.).

Based on previous research in a journal written by H. Supriyanto, N.W. Nugraha, and H. Febianto, good results were obtained in facial mask recognition. The developed system successfully detected the use of masks at the Polman Bandung Informatics Laboratory with a different number of people. The results showed that the recognition system worked with 100% accuracy at a light intensity of 80 lux."(Supriyanto et al., 2021).

Finally, after recognizing the above problems, facial mask recognition research was conducted. The research was conducted using a USB webcam to capture images in the lab and Jetson Nano as an image processor with the Yolov4-tiny algorithm to distinguish between masked and unmasked object.

2 RELATED WORKS

2.1 Face Mask Detection Using Jetson Nano

In a study entitled "Face Mask Detection Using Jetson Nano" conducted by Talagadadeevi Pooja, Badithaboina Chandra Mouli, Doodi Sri Lakshmana Harika, Bonu Yoganand, dan Nikkula Hemanjali in 2021(Pooja et al., 2021). The researchers managed to create a mask-wearing detection system using NVIDIA Jetson Nano, HD Logitech Camera, ResNet, and VNC worker.

2.2 YOLO v3-Tiny: Object Detection and Recognition Using One Stage Improved Model

In a study entitled " YOLO v3-Tiny: Object Detection and Recognition using one stage improved model" conducted by Pranav Adarsh, Pratibha Rathi, dan Manoj Kumar in 2020 (Pranav Adarsh et al., 2020). This research shows a comparison of several algorithms for detecting and identifying objects and found that YOLO-v3-Tiny can increase the speed of the object detection process while still striving for the accuracy of the detected objects (Pranav Adarsh et al., 2020).

2.3 Face Mask Detection Based on Transfer Learning and PP-YOLO

In a study entitled "Face mask detection based on Transfer learning and PP-YOLO" conducted by Wang Jian, dan Lin Lang in 2021 (Jian & Lang, 2021). This research shows that the PP-YOLO-Mask model gets the mAP value of 86.69%, and succeeds in obtaining good mask wearing detection in public places."(Jian & Lang, 2021)

3 METHODOLOGY

3.1 System Overview

The system we have developed is used to observe students who are studying in the informatics laboratory, if there are students who do not wear masks for a certain period of time, the system will raise an alarm, and students who violate it will be captured.The system we have developed looks as shown in Figure 1.

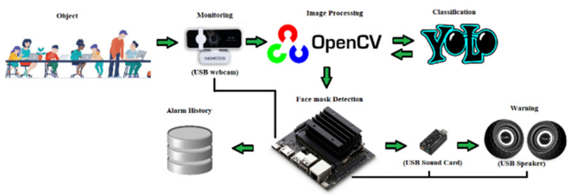


Figure 1: System Overview.

The USB webcam, USB sound card, and USB speaker are connected serially to the jetson nano. The USB webcam is used to monitor objects. Objects with and without masks are recognized by the system through OpenCV data processing and YOLO classification data. The USB sound card is used as a connection between the Jetson nano and the USB speaker. USB speaker cannot be connected directly to the Jetson nano because this microcontroller does not have the audio connector. Then the alarm history is used to store violation data when the object is not wearing a mask. Violation data can be accessed by moving the database file to a PC or laptop and then opening it with the DB Browser software.

Table 1: System Demand.

No	System Demand
1	Classification of human objects wearing masks or not
2	Speaker as a warning for violations of not wearing masks
3	Images in .jpg format for every time there is a violation
4	Video in .mp4 format for every time there is a violation
5	History of violations that stored in a database

Figure 2: Portable Face Mask Detector concept design.

This portable system consists of several modules including, a microcontroller module, speaker module, webcam module, and power module. All of these modules can be stored in a portable box to make it easier for users when moving the tool from one place to another. In the microcontroller module, there is a microSD as a data storage memory and Operating System from Jetson Nano. Then the ethernet cable is used for screen projection or controlling the Jetson Nano by other devices if needed. Screen projection is done using VNC Server and VNC Viewer software.

3.2 System Flow Chart

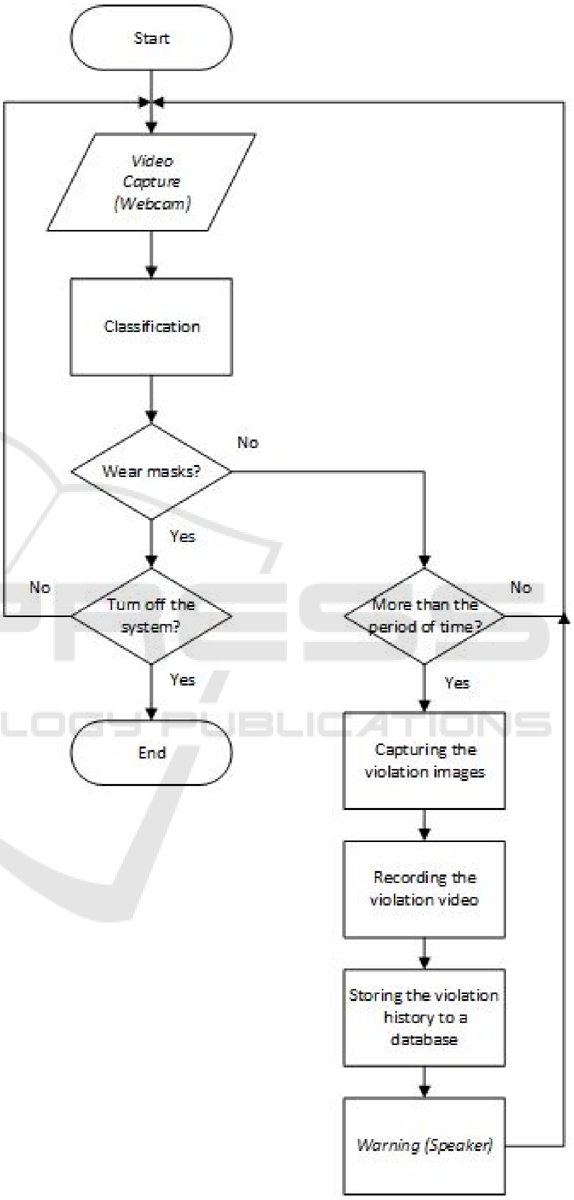


Figure 3: Face Mask Detector Flowchart.

In Figure 3, when the system starts operating, the USB Webcam will start taking video. Then the system will detect objects caught by the USB Webcam. When there is an object that does not wear a mask for a predetermined time, the system will detect the event as a violation of health protocols.

Jetson nano will process the data by taking pictures and recording videos of the health protocol violation. The captured data will be saved into the database as alarm history which is needed to know how many violations have occurred while the system is activated. After the data is stored, the alarm will be activated by issuing a voice warning and an appeal to wear a mask via USB Speaker.

3.3 Yolo Flow Chart

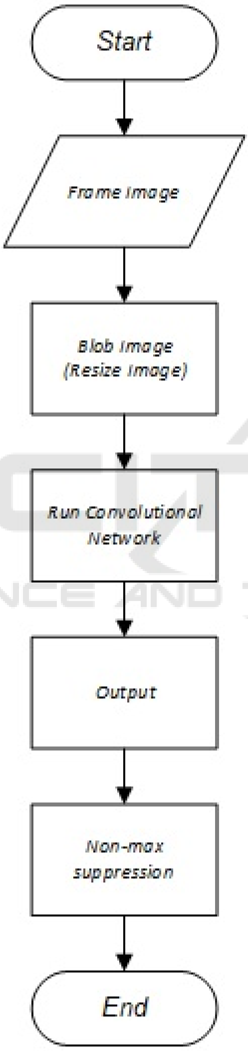


Figure 4: Yolo Flowchart.

Figure 4 shows the flowchart of the YOLO algorithm used to classify human objects wearing masks and not wearing masks. First, the image frame captured from the camera will be resized and made into a blob image, the input image will run a convolutional network to classify the prediction class. The output of the convolutional process is an array with a vector length of 7. where 4 bounding boxes, 1 box confidence, and 2 class confidence, non-max suppression is used to eliminate bounding boxes that have low confidence scores, and the last process is to create a box for each detected object that has a high score with the coordinates and width generated from the convolutional process.

3.4 Yolov4-tiny Model Making

This section describes the process of creating a yolov4-tiny model using google colab VM (Virtual Machine). Making the yolov4-tiny model can be done online by utilizing the colab VM facility.

Figure 5: Dataset example.

The first step begins with labeling the dataset with 2 classes, mask, and no-mask using labeling software. Labeling uses the YOLO format. Then clone darknet into the colab VM, this is done because for training models darknet is needed as a modeling framework. Then create a yolov4-tiny folder and a training folder in it to store the model results when the training process is complete. Then upload the dataset file, cfg file, obj.data, obj.names, and process.py which will be needed in the training process into the yolov4-tiny folder. Then build darknet to activate the framework function. Then run process.py to create train.txt and test.txt files which contain training image directories and validation image directories. then configure the obj.names file which contains 2 classes, the obj.data file which contains the train.txt, test.txt directories, the model directory generated in the training process and the number of classes and configure the yolov4-tiny-custom.cfg file to determine the size of the image to be processed, the number of classes, the number of iterations and others. Then load the pre-trained model and perform the training process until it reaches the specified max_batch value.

3.5 Dataset

The dataset used in this study was several images containing objects wearing masks and not wearing masks. In one image, it can contain objects wearing a mask, not wearing a mask, or both of them. The specification of the image for the dataset is in .jpg format. After the labeling process is completed, the .txt file will be generated. Each image has a .txt file that can describe the object classes based on YOLO format.

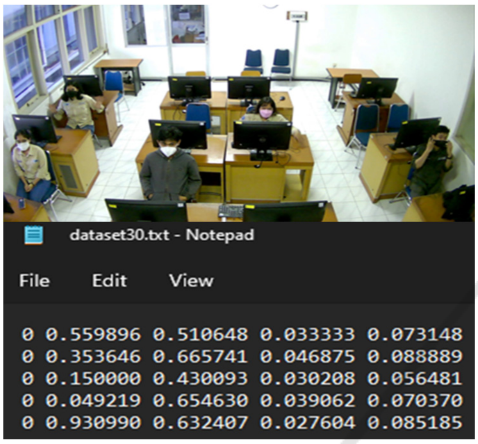


Figure 6: .jpg and .txt file.

The first digit in the .txt file indicates the class of the object, if the value is 0 then the object is wearing a mask and if it is 1 then the object is not wearing a mask. The next digits are the coordinates of the bounding box location according to the YOLO format.

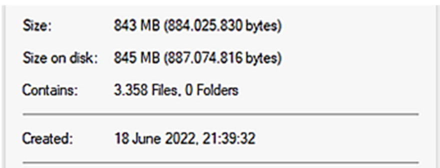


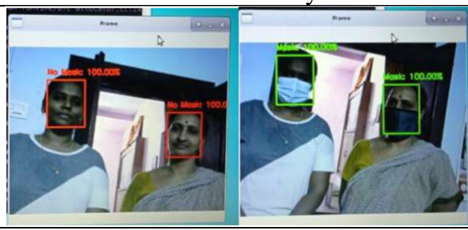
Figure 7: Dataset folder.

The number of datasets used in this study is 1679 .jpg images. Then 1679 other files in the form of .txt. These .jpg and .txt files were used for training and testing. After the training and testing process completed, then it will generate the yolov4-tiny model that used in this study.

3.6 Difference with Previous Study

In the previous study entitled "Face Mask Detection Using Jetson Nano" conducted by Talagadadeevi Pooja, Badithaboina Chandra Mouli, Doodi Sri Lakshmana Harika, Bonu Yoganand, dan Nikkula Hemanjali in 2021 (Pooja et al., 2021). The researchers succeeded in making a mask detection system for short distances between objects and the camera. while in this study, the system is implemented to detect objects with a longer distance and a greater number of objects. The difference lies in the dataset used and the scope of its application.

Table 2: The result in previous study and this study.

Previous Study

This Study


In this study, the system developed to have 5 functions. There are classification function, warning alarm (when there was an object not wearing mask), capturing photo, capturing video, and violation history function.

3.7 YOLO and Mobilenetv2 Algorithm

Based on a paper entitled "Comparison of Object Detection Algorithm for Street-Level Objects" in 2022 (Naftali et al., 2022), yolo algorithm is better than mobilenetv2 in terms of quantitative testing result. But mobilenev2 has a faster inference time for detecting object.

Table 3: Quantitative Testing Result.

Measure	Mobilenetv2	Yolov3	Yolov4	Yolov5l	Yolov5s
Precision	0.354	0.767	0.569	0.780	0.670
Recall	0.228	0.535	0.589	0.545	0.501
mAP@0.5	0.315	0.583	0.582	0.593	0.530
Inference Time (ms)	6.30	27.60	27.90	25.40	8.50

The important things in this study is the precision and accuracy of the detection result. So, yolo algorithm is more suitable to be applied in this study.

4 RESULT

4.1 System Function Testing

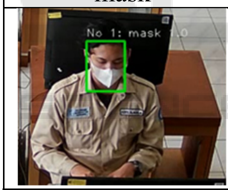
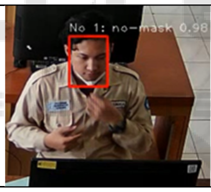
This test is carried out to see the ability of the system that has been made against the demands of the system that has been determined.

Table 4: System function testing.

No	System Demand	Status
1	Classification of human objects wearing masks or not	Success
2	Speaker as a warning for violations of not wearing masks	Success
3	Images in .jpg format for every time there is a violation	Success
4	Video in .mp4 format for every time there is a violation	Success
5	History of violations that stored in a database	Success

- Classification of human objects wearing masks or not

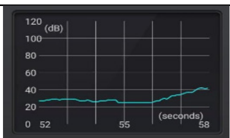
Table 5: Classification.

mask	no-mask	Status
		OK

This table shows that the system can classify objects that wear masks and objects that do not wear masks. A green box will appear when the object is wearing a mask, and a red box will appear when the object is not wearing a mask. The performance of the system has met the demands of system number 1

- Speaker as a warning for violations of not wearing masks

Table 6: Classification.

Condition	Db Meter
No-Violation	
Violation	

In the table 6 shows that the warning speaker will actively make a sound when there is a violation. The performance of the system has met the system demands number 2.

- Images in .jpg format for every time there is a violation

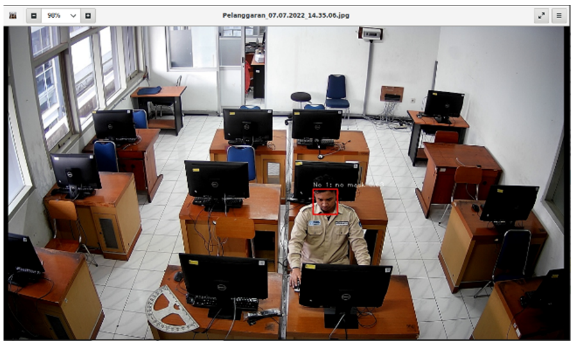


Figure 8: Violation image.

The system has successfully captured a .jpg image when detecting a violation, naming the image file according to the time of the violation. The performance of the tool has met the demands of system number 3.

- Video in .mp4 format for every time there is a violation

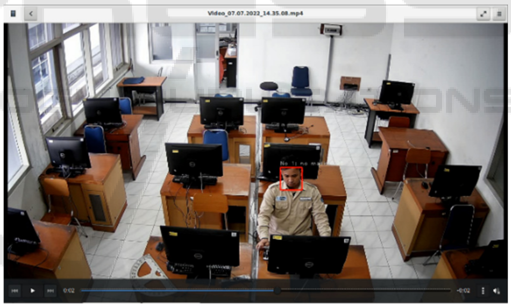


Figure 9: Violation video.

The system has successfully recorded the .mp4 video when it detects a violation, the name of the video file generated according to the time of the violation. The performance of the tool has met the demands of system number 4.

- History of violations that stored in a database

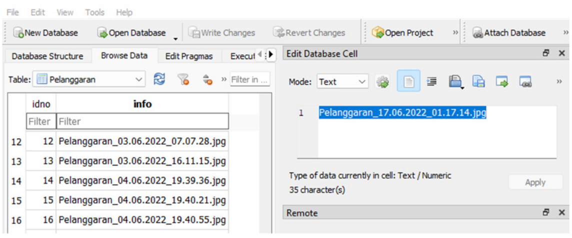


Figure 10: Violation history.

The system has successfully saved the data of violation history into a database. The performance of the tool has met the demands of system number 5.

4.2 Main Scheme Testing

This test is carried out to see the performance of the system that has been made in detecting students who are studying in the informatics laboratory. This test aims to see the system's ability to detect the use of masks simultaneously and the accuracy of its predictions.

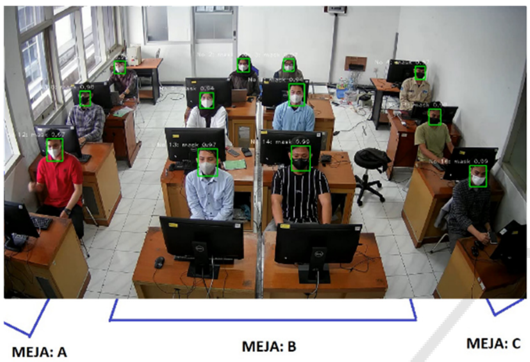


Figure 11: System testing environment.

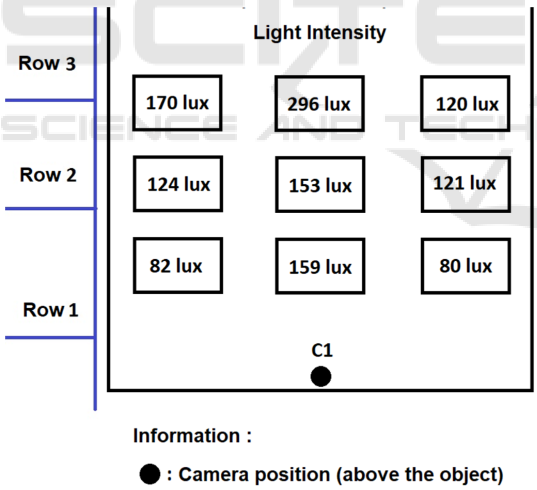


Figure 12: Light intensity in each desk and camera position.

Table 7: Main Scheme Result Table.

Main Scheme					
No	Total Objects		Prediction		Result
	mask	no-mask	mask	no-mask	
1	12	0	12	0	Correct
2	12	0	12	0	Correct
3	10	2	10	2	Correct
4	10	2	10	2	Correct
5	10	2	10	2	Correct
6	10	2	10	2	Correct
7	10	2	10	2	Correct
8	10	2	10	2	Correct
9	11	1	11	1	Correct
10	10	2	10	2	Correct
11	10	2	10	2	Correct
12	10	2	10	2	Correct
13	10	2	10	2	Correct
14	11	1	11	1	Correct
15	11	1	11	1	Correct
16	10	2	10	2	Correct
17	10	2	10	2	Correct
18	9	3	9	3	Correct
19	9	3	9	3	Correct
20	9	3	9	3	Correct
21	9	3	9	3	Correct
22	8	4	8	4	Correct
23	8	4	8	4	Correct
24	8	4	8	4	Correct
25	0	12	0	10	Wrong

Performance metric:

		Actual Class	
		Mask	No-Mask
Predicted class	Mask	TP 237	FP 0
	No-Mask	FN 2	TN 61

$$\text{Accuracy} = \frac{237 + 61}{237 + 61 + 0 + 2} \times 100\% = 99{,}33\%$$

$$\text{Precision} = \frac{237}{237 + 0} \times 100\% = 100\%$$

$$\text{Recall} = \frac{237}{237 + 2} \times 100\% = 99{,}16\%$$

$$\text{F} - \text{Measure} = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 2 = 99{,}58\%$$

$$\text{Error} = \frac{0 + 2}{237 + 61 + 0 + 2} \times 100\% = 0{,}67\%$$

From the results of system testing using the confusion matrix, the accuracy value is 99.33%, and the precision value of mask-wearing detection is 100%. While the average recall value is 99.16%. From the precision data and recall data, the F-Measure value is 99.58% and the error is 0.67%.

5 CONCLUSIONS

Based on the system test results on the portable face mask detector system during this research process, there are the following conclusions:

1. Portable Face Mask Detector using jetson nano has a warning feature every time a violation occurs via USB Speaker, recording with video output in .mp4 format and .jpg images, and storing violation history in the database. The system successfully detects the use of masks when the distance between the object and the camera varies from 2.68 m to 5.38 m and the position of the object's face is 110˚ (right to left) and 83.33˚ (up and down).
2. The system successfully detects objects in the main scheme with an error value of 0,67%.

REFERENCES

Fatia Zulfa, & Henni Kusuma. (2020). Upaya Program Balai Edukasi Corona Berbasis Media Komunikasi Dalam Pencegahan Penyebaran Covid-19. *Jurnal Abdimas Kesehatan Perintis*, *2*(1), 17–24. https://www.jurnal.stikesperintis.ac.id/index.php/JAKP/article/view/445/251

Jayaweera, M., Perera, H., Gunawardana, B., & Manatunge, J. (2020). Transmission of COVID-19 virus by droplets and aerosols: A critical review on the unresolved dichotomy. In *Environmental Research* (Vol. 188). Academic Press Inc. https://doi.org/10.1016/j.envres.2020.109819

Jian, W., & Lang, L. (2021). Face mask detection based on Transfer learning and PP-YOLO. *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering, ICBAIE 2021*, 106–109. https://doi.org/10.1109/ICBAIE52039.2021.9389953

Naftali, M. G., Sulistyawan, J. S., & Julian, K. (2022). Comparison of Object Detection Algorithms for Street-level Objects. *ArXiv.Org*. http://arxiv.org/abs/2208.11315

Pooja, T., Chandra Mouli, B., Sri, D., Harika, L., Yoganand, B., & Hemanjali, N. (2021). Face Mask Detection Using Jetson Nano. *International Research Journal of Modernization in Engineering Technology and Science*. www.irjmets.com

Pranav Adarsh, Pratibha Rathi, & Manoj Kumar. (2020). *YOLO v3-Tiny: Object Detection and Recognition using one stage improved model*.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (n.d.). *You Only Look Once: Unified, Real-Time Object Detection*. http://pjreddie.com/yolo/

Supriyanto, H., Nugraha, W., & Febianto, H. (2021). *Face Mask Detector Using USB Webcam WC300 BASED ON OPENCV*. 1–6.

Susilowati, E., Meiranny, A., & Salsabilla, D. (2021). ISPA DAN FAKTOR PENYEBABNYA. *Proceedings National Seminar*, 161–177. https://doi.org/103.154.143.189

World Health Organization. (2021). *Covid-19 Weekly Epidemiological Update*.

Wu, Y. C., Chen, C. S., & Chan, Y. J. (2020). The outbreak of COVID-19: An overview. In *Journal of the Chinese Medical Association* (Vol. 83, Issue 3, pp. 217–220). Wolters Kluwer Health. https://doi.org/10.1097/JCMA.0000000000000270

World Health Organization. (2019). *The WHO special initiative for mental health (2019-2023): universal health coverage for mental health* (No. WHO/MSD/19.1). World Health Organization.