ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

A YoloV5-Based Deep Learning System for Detecting Face Masks

[1] Sahiba Sulthana, [2] K. Sravani,

[1]Assistant Professor, Megha Institute of Engineering & Technology for Women, Ghatkesar.
[2] MCA Student, Megha Institute of Engineering & Technology for Women, Ghatkesar.

Abstract—

Wearing a face mask in public during the current Coronavirus Disease 2019 (Covid-19) pandemic might help lower the number of persons affected by limiting the discharge of respiratory droplets from infected individuals. The purpose of this research is to investigate "Yolov5"s" deep learning model, which is a powerful tool for face mask recognition. Different models with 20, 50, 100, 300, and 500 epochs were built for comparison. Among the several deep learning models tested, the one with 300 epochs had the best overall performance, with a 96.5% accuracy rate. Subjects—deep learning, Yolov5, face mask detection

I. INTRODUCTION

A new member of the family of viruses that cause SARS and MERS, the 2019 novel coronavirus (2019-nCoV) is responsible for the 2019 COVID-19 pandemic. Infected and dead patients are piling up at an alarming rate. In December 2020, there were around 82.7 million cases globally, with almost 1.8 million fatalities, as reported by the World Health Organization (WHO) [1]. Direct contact with respiratory droplets from an infected individual is the main mode of viral transmission. Wearing a face mask is strongly recommended, and in some places even mandated, as a means to contain the COVID-19 epidemic. Wearing a face mask may reduce the spread of COVID-19 in the community, as well as the number of hospitalizations and fatalities caused by the virus, according to research. One possible way to lessen the impact of the pandemic, according to Steffen E. Eikenberry et al. [2], is to wear a mask when you're out in public. During the COVID-19 epidemic, researchers Cornelia Betscha et al. [3] set out to examine the behavioral and societal effects of mask legislation. Countries or communities favored mask wearers, and the research found a favorable correlation between mask use and other preventive behaviors. As the COVID-19 epidemic continues to spread, some governments have instituted policies mandating the use of face masks in public spaces, including public transportation. In order to verify if someone is wearing a mask or taking a temperature before entering a facility, security systems often use computer vision. There is evidence that AI-powered face mask detection systems work; for example, Sammy V. Militante and Nanette V. Dionisio [4,5] developed a deep learning-based system to detect when someone is wearing a face mask and to measure physical distance, and Sharma Vinay [6] developed a convolutional neural network (CNN) to do the same. In order to determine whether someone is wearing a face mask correctly, improperly, or not wearing one at all, this study used YoloV5 to conduct a CNN. Section IV provided an analysis and discussion of the experimental findings for face mask identification derived from deep learning models with various epochs (20, 50, 100, 300, and 500), and Section V provided a conclusion.

II. FACE MASK DETECTION USING CNN

A. Face mask dataset

The 853 photos that made up the public face mask dataset included three labels: "With_Mask," "Without_Mask," and "Incorrect_Mask," reflecting the same format as in earlier research [7-8]. Six hundred eighty-two photographs were used for training the model, eighty-five images were used to validate the results, and eighty-six images were used for testing the model. In the PASCAL VOC format, each picture had a label, bounding boxes, and a size that varied. Figure 1 displays several sample pictures from the face mask collection.

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

Fig. 1. Examples of face mask image dataset

B. Face mask detection framework

The YoloV5 model that was constructed for this article was split into two components: the training model and the face mask detection model. The two portions are shown in Figure 2. The face mask dataset was used to train the model using 682 photos. As inputs, the YoloV5 face mask detection model processed a prediction score of three classes: "With_Mask," "Without_Mask," and "Incorrect_Mask." The model then supplied the output photos with their predicted classes and detection scores.

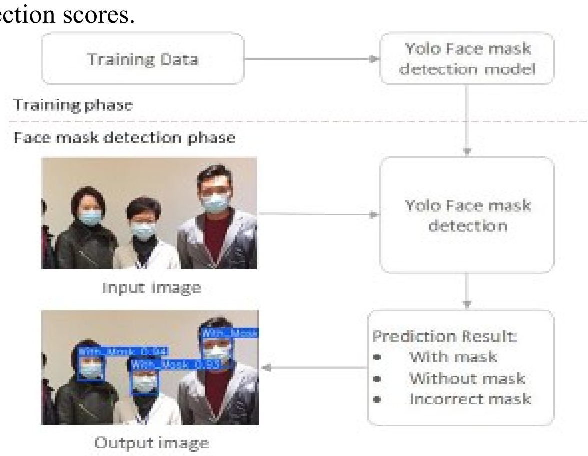


Fig. 2. A diagram of face mask detection framework

C. A YoloV5-based deep learning model for facial masks YoloV5 is an object detection network that uses regions and has one stage. In order to achieve fast processing speeds, the Yolo reframes object detection as a regression issue. Not long ago, YoloV5 found use in a vision system for an apple-harvesting robot and in real-time person search [9]. The backbone, the head, and the detection are the three primary parts of YoloV5. A convolutional neural network (CNN) is the backbone, gathering and shaping picture features at various granularities. When developing picture features, YoloV5 makes use of the CSP Bottleneck. The central processing unit (CPU) is a multi-layered architecture that integrates picture characteristics before passing them on to a prediction engine. To aggregate features, the YoloV5 uses the PA-NET as well. As part of its identification procedure, the algorithm employs both box and class prediction processes, drawing on characteristics from the brain. Figure 3 shows a schematic of the YoloV5 design. The training dataset for the YoloV5 face mask identification model included 685 photos annotated in VOC format. It has been shown that the model's performance is affected by an epoch, which is the number of runs over the full training dataset, in the training model process [11]. Finding the best epoch to use the YoloV5 with the training model that was constructed was the main objective of this work.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

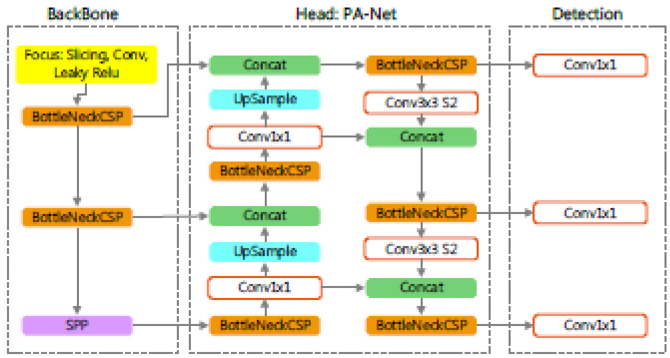
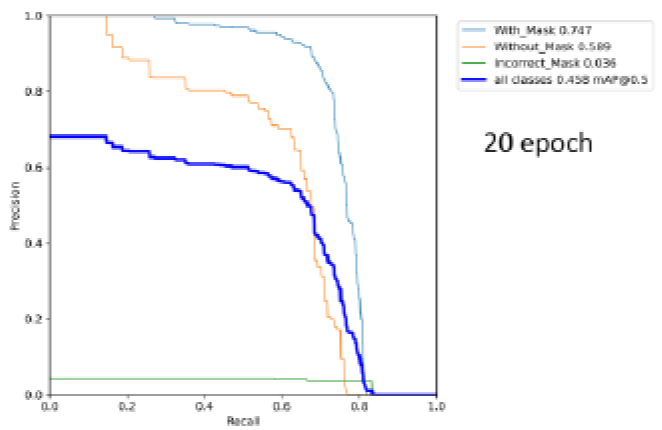


Fig. 3. Overview of YoloV5 architecture

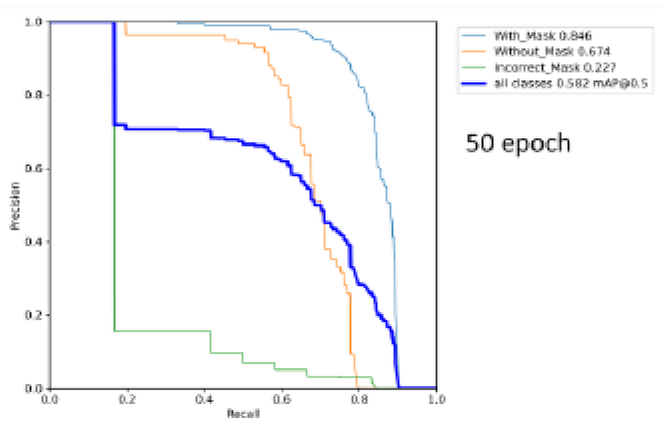
III. EXPERIMENTAL RESULTS

A. Model Training

The training face mask dataset, which consists of 682 photos split into three categories: "With_Mask," "Without_Mask," and "Incorrect_Mask," is subjected to five distinct epochs—20, 50, 100, 300, and 500—in order to choose the best epoch for the YoloV5 training model implementation. Additionally, the recall and accuracy were calculated. A total of 85 photos of face masks were used to verify each model. Each class's recall and accuracy were supplied via the validation procedure. Figure 4 displays the diagrams representing the recall and accuracy of each model. Figure 4 (d) shows that out of all the training models, the one with 300 epoch had the best performance.



(a)



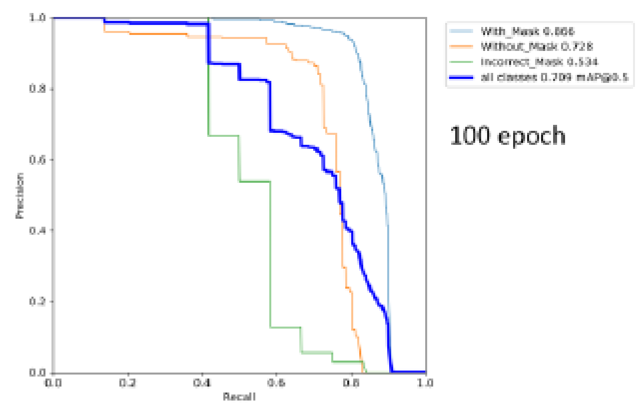
(b)

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

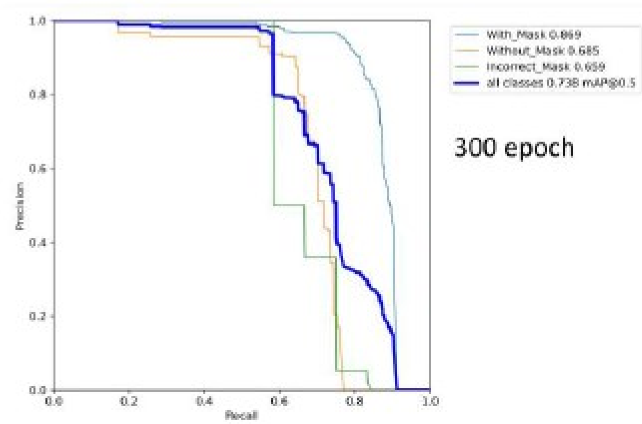
ISSN 2454-9940

www.ijasem.org

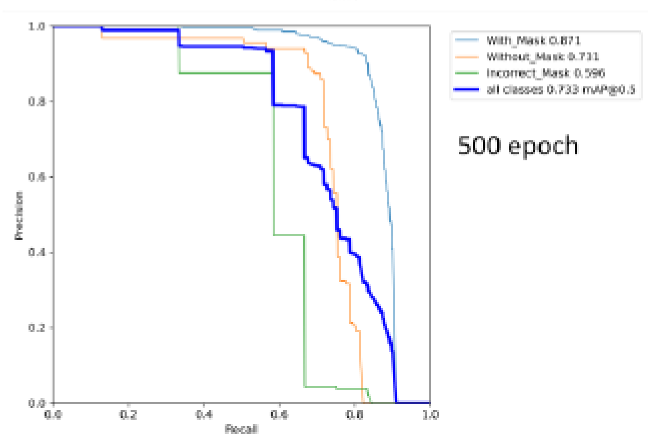
Vol 19, Issue 2, 2025



(c)



(d)



(e)

Fig. 4. Precision and recall of training model at diffierent epoch; (a) 20 epoch, (b) 50 epoch, (c) 100 epoch, (d) 300 epoch, and (e) 500 epoch

Figure 4 shows the experimental data showing that compared to 20 epochs, 50 epochs, and 100 epochs, the training model for each class at 500 epochs produces better outcomes. In addition, when comparing the training models for each class at five different epochs, the one with 300 epochs yields the greatest performance, meaning the highest

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

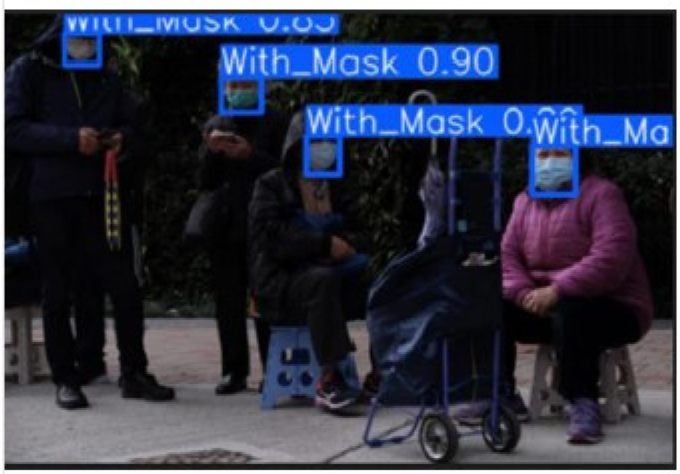
ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

mean average precision (mAP) across all classes. Results for detecting face masks (B) Additionally, the 86-image face mask dataset is used to evaluate the built deep learning models. Figure 5 and Table 1 show the results of the face mask detection.

(a)



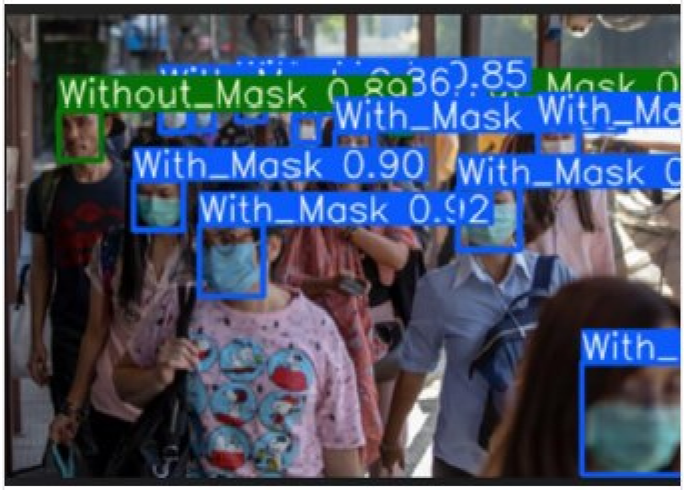
(b)

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025



(c)

Fig. 5. Face mask image detection results with three cases: (a) incorrect mask in one-person image, (b) with mask in people image, and (c) with/without mask in crowd image.

TABLE I. ACCURACY OF FACE MASK DETECTION RESULTS

Cases	*Number of*	
	Correct Detections	*Incorrect Detections*
20 epoch	79	7
50 epoch	81	5
100 epoch	81	5
300 epoch	83	3
500 epoch	82	4

Figure 5 shows the results of testing the built deep learning model with YoloV5 on three image cases: 1) an incorrect mask in a one-person image, 2) a people image with a mask, and 3) a crowd image with and without a mask. The results show that the model performs well in detecting face masks. In addition, the developed model achieved 83 correct detections and 3 incorrect detections out of 300 epochs when tested with 86 images (Table I). This means that the model's accuracy is very high, reaching 96.5%, which is the highest precision and recall achieved from the training model with 300 epochs in each class (Fig. 4 (d)).

C. Result discussion

The performance of mask detection is often enhanced by increasing the number of epochs, since this increases the number of processing steps. Table 1 shows that both models performed similarly after 500 epochs of training, while the model with 300 epochs produced far superior results, with just three false positives.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

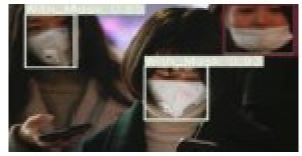
ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

(a)

(b)



(c)

Fig. 6. Incorrect detection resulted from using model trained with 300 epochs (a) "Incorrect_Mask" detected as "Without_Mask", (b) people using white cloth instead of white mask, and (c) "Incorrect_Mask" detected as "Without_Mask" in a partial facial/upper body image

Figure 6 (a) and (b) show examples of erroneous object recognition due to distortion in form; one face mask in Figure 6 (a) doesn't seem to be a face mask, while white fabric was mistakenly identified as a face mask in Figure 6 (b). Figure 6(c) displays a face mask-wearing lady who is mistakenly identified as "Without_Mask" in the upper right corner. A missing upper body or face picture caused the mistaken detection.

IV. CONCLUSION

Research and development of a deep learning model for facial mask recognition is detailed in this article. Using five distinct epoch counts, the YoloV5 trains the model. With a 96.5% accuracy rate tied to the greatest recall and precision, the 300-epoch deep learning model for face mask identification outperformed the other 85 photos that were examined. Using a computer vision system, future work might determine whether a person is wearing a mask correctly or incorrectly in real time.

REFERENCES

[1] WHO, (December 31, 2020). WHO Coronavirus Disease (COVID-19) Dashboard. https://covid19.who.int/

[2] Eikenberry, Steffen E., Marina Mancuso, Enahoro Iboi, Tin Phan, Keenan Eikenberry, Yang Kuang, Eric Kostelich, and Abba B. Gumel. "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* (2020).

[3] Betsch, Cornelia, Lars Korn, Philipp Sprengholz, Lisa Felgendreff, Sarah Eitze, Philipp Schmid, and Robert Böhm. "Social and behavioral consequences of mask policies during the COVID-19 pandemic." *Proceedings of the National Academy of Sciences* 117, no. 36 (2020): 21851-21853.

[4] Militante, Sammy V., and Nanette V. Dionisio. "Real-Time Facemask Recognition with Alarm System using Deep Learning." In *2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC)*, pp. 106-110. IEEE, 2020.

[5] Militante, Sammy V., and Nanette V. Dionisio. "Deep Learning Implementation of Facemask and Physical Distancing Detection with Alarm Systems." In *2020 Third International Conference on Vocational Education and Electrical Engineering (ICVEE)*, pp. 1-5. IEEE, 2020.

[6] Sharma, Vinay. "Face Mask Detection using YOLOv5 for COVID-19." (2020).

[7] MakeML, 2020. Mask Dataset https://makeml.app/datasets/mask

[8] Kaggle Dataset. 2020. Face Mask Detection. https://www.kaggle.com/andrewmvd/face-mask-detection

[9] Li, Ye, Kangning Yin, Jie Liang, Chunyu Wang, and Guangqiang Yin. "A Multi-task Joint Framework for Real-time Person Search." *arXiv preprint arXiv:2012.06418* (2020).

[10] Kuznetsova, Anna, Tatiana Maleva, and Vladimir Soloviev. "Detecting Apples in Orchards Using YOLOv3 and YOLOv5 in General and Close-Up Images." In *International Symposium on Neural Networks*, pp. 233-243. Springer, Cham, 2020.

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

[11] Glenn Jocher, Alex Stoken, Jirka Borovec, NanoCode012, ChristopherSTAN, Liu Changyu, Prashant Rai. (2020, October 29). ultralytics/yolov5: v3.1 - Bug Fixes and Performance Improvements (Version v3.1). Zenodo. http://doi.org/10.5281/zenodo.4154370

[12] Kaggle. 2020. YOLOv5 Facemask. https://www.kaggle.com/rkuo2000/yolov5-facemask