Int. J. Nonlinear Anal. Appl. 14 (2023) 1, 679–697
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2022.28178.3824

Face mask detection based on algorithm YOLOv5s

Mohammed S.H. Al-Tamimi*, Firas Amer Mohammed Ali

Computer Science Department, Collage of Science, University of Baghdad, Baghdad, Iraq

(Communicated by Mohammad Bagher Ghaemi)

Abstract

Determining the face of wearing a mask from not wearing a mask from visual data such as video and still, images have been a fascinating research topic in recent decades due to the spread of the Corona pandemic, which has changed the features of the entire world and forced people to wear a mask as a way to prevent the pandemic that has calmed the entire world, and it has played an important role. Intelligent development based on artificial intelligence and computers has a very important role in the issue of safety from the pandemic, as the Topic of face recognition and identifying people who wear the mask or not in the introduction and deep education was the most prominent in this topic. Using deep learning techniques and the YOLO ("You only look once") neural network algorithm, which is an efficient real-time object identification algorithm, an intelligent system was developed in this thesis to distinguish which faces are wearing a mask and who is not wearing a wrong mask. The proposed system was developed based on data preparation, preprocessing, and adding a multi-layer neural network, followed by extracting the detection algorithm to improve the accuracy of the system. Two global data sets were used to train and test the proposed system and worked on it in three models, where the first contains the AIZOO data set, the second contains the MoLa RGB CovSurv data set, and the third model contains a combined data set for the two in order to provide cases that are difficult to identify and the accuracy results that were obtained. obtained from the merging datasets showed that the face mask (0.953) and the face recognition system were the most accurate in detecting them (0.916).

Keywords: Mask, Corona Pandemic, Deep Learning, YOLO ("You Only Look Once"), Data Sets
2020 MSC: 68T07

1 Introduction

In December of 2019, a pandemic crisis known as Corona Virus Infection Disease 2019 (COVID-19) was declared in China. It is a sickness similar to the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus, and on 11 March 2020, the World Health Organization (WHO) proclaimed it a global pandemic. Humanity has witnessed numerous pandemics throughout history, including the African obesity pandemic and the pandemic flu pandemic [16]. Humans face a critical moment in their battle against an invisible foe, Coronavirus. And it has infected millions of people worldwide, and many have died. Several people in Wuhan city reported having pneumonia of unknown origin around the beginning of January 2020. The rapid expansion of the COVID-19 outbreak presents significant obstacles to virus containment [8]. With the awful ascent of Corona Virus Infection Disease, the whole world is looking for a method for halting the spread of the infection. New limitations have arisen in day-to-day existence, for example, quarantine, social distancing, wearing masks, continuous hand washing, and Determining the number of people in enclosed spaces.

*Corresponding author
Email addresses: m_altamimi75@yahoo.com (Mohammed S.H. Al-Tamimi), firasamerm@gmail.com (Firas Amer Mohammed Ali)

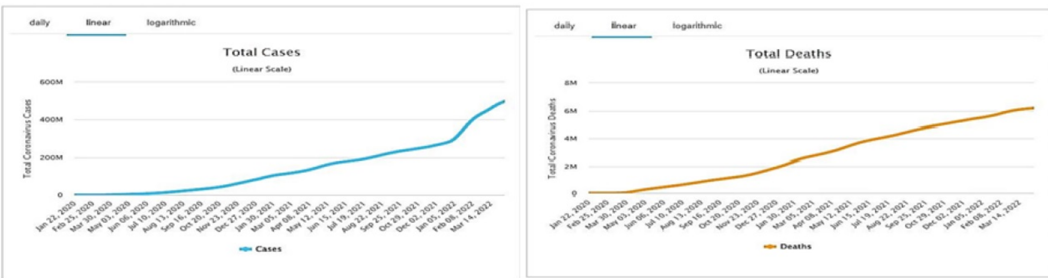
Received: June 2022 *Accepted:* August 2022

Government elements have set up numerous biosafety guidelines to lessen contamination [14]. Among them is the obligatory utilization of wearing masks openly, as they have been demonstrated to be compelling in safeguarding people and everyone around them [19]. From this point of view, the mask has become an essential element that does not leave people because of its important role in protecting people from infection with the Coronavirus. Are people eligible to enter or not by wearing a mask [7]. This paper provides an overview of the effect of the Corona pandemic on mask detection; the change caused, the importance of the catcher used to prevent the spread of COVID-19, and the challenges of masked face detection for prosed system.

2 Corona Pandemic

Covid-19 is an acronym for Corona Virus Infection Disease 2019, an infectious disease caused by (SARS-CoV-2) infection and has a place in the COVID family [10]. The latest update by the organization about the Corona pandemic was Global; the number of new weekly cases has continued to decline since a peak in January 2022. During the week of 6 until 12 June 2022, over 3.2 million cases were reported, similar to the number reported during the previous week. After five weeks of decline, the number of new weekly deaths has risen again, with over 8700 fatalities reported, a 4% increase compared to the previous week [12]. As of 12 June 2022, over 533 million confirmed cases and over 6.3 million deaths had been reported globally. This edition updates the geographic distribution of circulating SARS-CoV-2 variants of concern (VOCs). Also provide a particular focus on mass gathering events during disease outbreaks and WHO's recommended risk-based approach to decision-making for mass gathering events [20].

Figure 1 shows the total number of cases caused by the pandemic, with recoveries among patients and more than 6 million deaths.



(a): Number of total cases **(b):Number of total deaths**

Figure 1: Number of total cases and deaths [20]

General side effects can be loss of smell and taste, cough, fever, weariness, and breathing. In extreme cases, a few people foster intense respiratory pain, Acquired Immunodeficiency Syndrome (ARDS). In these specific cases, longer-term harm to organs has been noticed, and a few patients who have recuperated keep on encountering a scope of impacts including muscle shortcoming, cognitive decline and different side effects for quite a long time a while later [6, 25]. The infection communicates for the most part employing respiratory plot through little drops from coughing, talking, and gasping with a higher transmission factor in shut and inadequately ventilated places. Individuals may become tainted by contacting polluted items and their faces afterwards [2].

3 Famous Problem Background

The face mask detection process challenges come from the great diversity in the real world [24]. If they are found, these greetings or difficulties may cause a weakness in the efficiency of the masked face detection systems [3]. The most critical challenges showed in Figure 2 and summarized as follows points:

1. **Different face orientations:** The face may be in different orientations and reflections. This difference in the direction of the face comes from the difference in the camera's location relative to the face during shooting. The face will have different angles, and this difference affects the recognition of the main features of the face, such as the nose, eyes, mouth, and mask, and this, in turn, leads to a weakness in the efficiency of the masked face detector in general [1, 5].
2. **Non-mask Occlusion:** There may be a blockage of the face with certain things that make the process of discovering the face a difficult task. Where there is a face distortion due to the interference of the face or other

Figure 2: Challenges in face mask detection [5]

items such as clothes, glasses, placing the hand on the face, etc. [22], this makes the process of extracting the properties of those images is complicated [4, 23].

3. **Small or blurred faces:** Small faces, blurred faces, different face sizes, number of faces, Light source and difference in ambient light, contrast dimensions during image formation.
   These factors are caused by different shooting conditions and lead to the loss of details that the detector can deal with [1, 9, 17].
4. **Various Mask Types:** The face masks come in different shapes, colours and patterns [24]. Training the detector on certain types of masks will lead to its inability to detect other types of masks, so the data set used in training must include most of the types of masks that can be used [1, 11].

4 Related work

There is a wide range of studies conducted on the subject of revealing face masks, especially after the Corona pandemic; the following are the most prominent studies that were close to the proposed system:

In [22] the proposed system allows individuals who enter the mall to capture images with a camera and send them to an interface for face mask recognition. The mall's entrance will open if a face with a mask is picked up within two seconds. It is open and available for passage; if not, it will be completed. He repeated identifying the face mask until he was successful. And the proposed system used a method based on YOLOV5 for application to recognize faces, whether wearing masks, people into the store; they just need to stand in front of the camera, people can be identified, if recognition success and interface display can enter the gate open, this method is no longer required to use human crowd control, which saves a substantial amount of time and resources. In [15] the proposed system using SSDv2 is a multi-scale detection algorithm designed to detect obscured and small mask-wearing targets. Multi-scale features help networks detect targets of different sizes, so they can be used to detect face masks in distant and close shots. Experiments show the network accurately detects masks in complex scenes. Real-time face mask detection is another good use for the network. In [18] the proposed System for Face detection with alignment and face mask classification modules. This pipeline lets to use efficient face detectors. To predict mask existence on identified faces and train MobileNetV11.0 as a backbone. Also tried label smoothing, aligning with critical spots, and omitting the upper part of the face when training the face mask classification model. Base training produces the most refined model. This shows that the proposed pipeline is a suitable method based on the current methods. In [13] the proposed system, the First, creates algorithms to detect masks in public locations and temperature monitoring to detect fever. A data set and labels were created to train the algorithms in this component. Given the data collection and sample size, a program has been designed to apply artificial masks to RGB images using predefined models that recognize faces and facial locations. Based on this information, a mask is applied to the face's existing types and textures. This data set's labels came from previously used models. Multiple YOLOv5 algorithms were examined using this data collection. All the models produced good results after training and evaluation. Still, the tiny model was chosen because the proposed system used YOLOv5 with a data set created and released for researchers to work on with Mendeley. In [21] the proposed System is A lightweight, high-performance face mask detector based on YOLOv5 with several technologies where Shuffle CANet combines ShuffleNetV2 and Coordinate Attention for rich feature extraction and relevant information. Next, utilize BiFPN to incorporate features. Combine high- and low-level semantics. To get high-quality fixation sites, we replaced the original localization loss function with CIoU loss during model training.

5 Proposed System for Face Mask Detection

The effects of COVID-19 on the global economy can be seen with the naked eye, as the confinement of people in their homes brings less production and slows down commercial dynamism. Today, after using the personal face mask as an obligatory precaution, there is an urgent need for applications that control the entry of people in certain places, and these applications must adapt to safety conditions to maintain the physical security of the user and allow them to continue daily activities as naturally as possible. This proposed a system for face mask detection based on the latest deep learning techniques to increase the public's awareness and remind people to wear a mask. To restrict entry to certain places and to discover whether people wear face masks or not, the proposed system can work offline (from images) and online (real-time). It contains features that help the system run on devices with limited computing power, a lower financial cost, and is easy to implement using surveillance cameras compared to existing models. The proposed system is applicable in the real world.

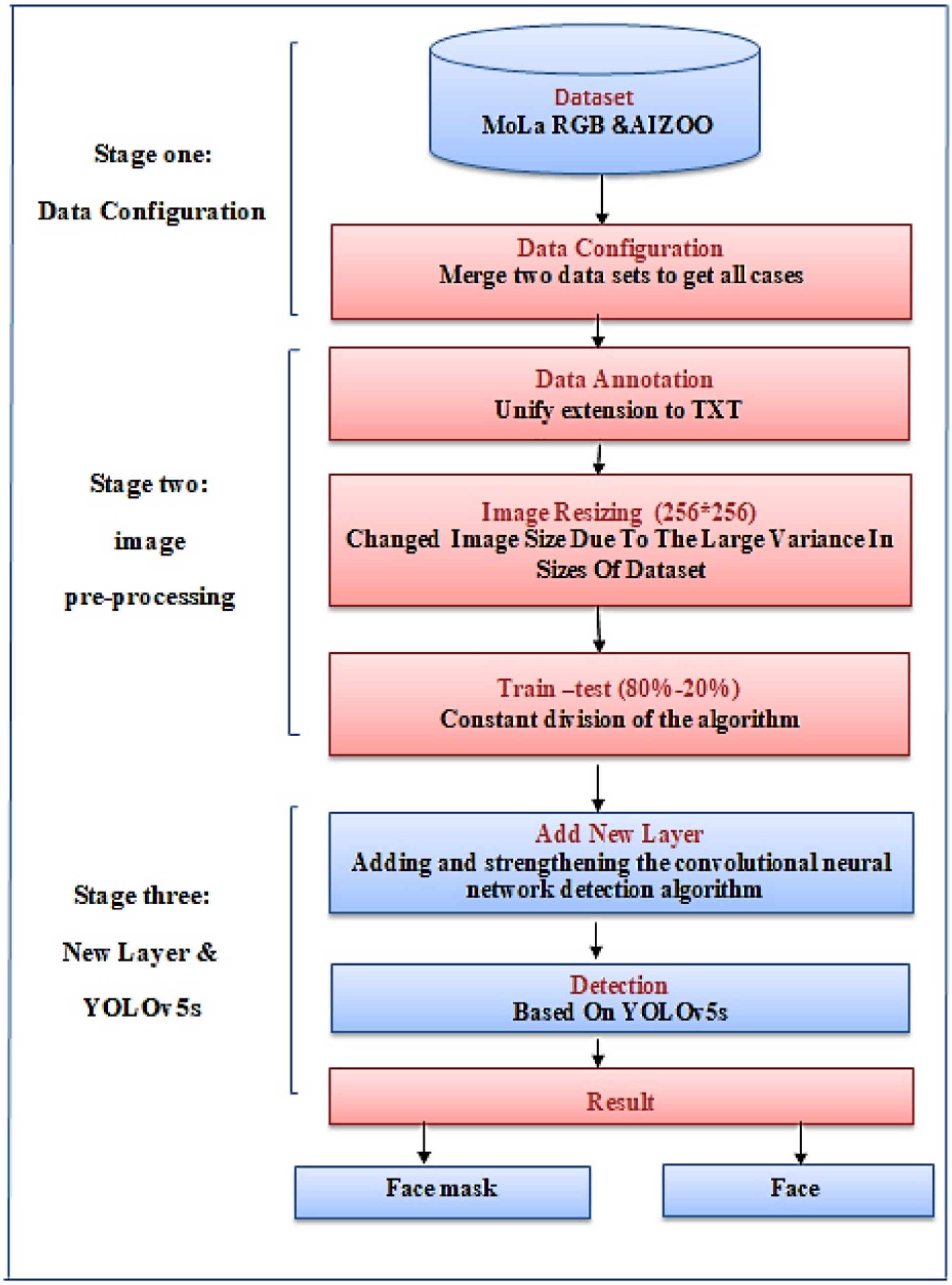


Figure 3: Block Diagram of Proposed System

5.1 Dataset in the Proposed System

The dataset used in the proposed system is explained in detail in the third chapter; two datasets, AIZOO and MoLa RGB CovSurv, are combined into one dataset (Merged dataset) to provide all possible cases for training the system in the following Figure 4 is an analytical chart to show the number of faces and faces mask used in training the proposed system.

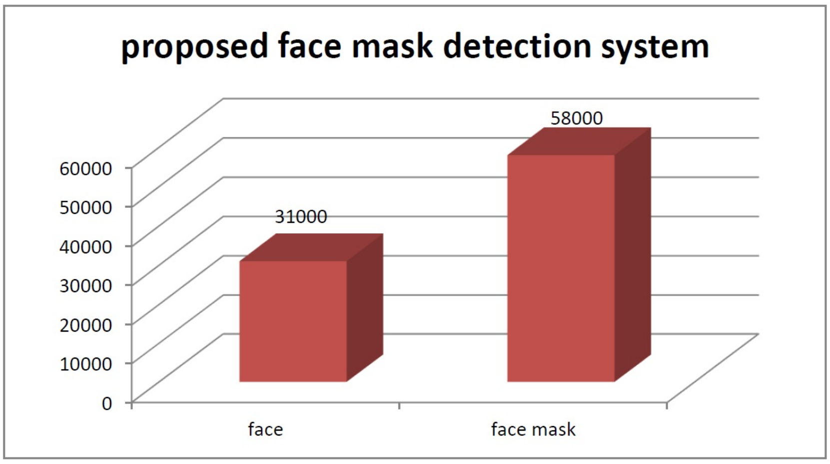


Figure 4: Chart face and face mask in the proposed system dataset

6 Pre-Processing Proposed System

Pre-Processing operations are critical in detection issues, as they assure satisfactory results and high accuracy. Will take a set of images to review the system's work in processing and detection. The following is a discussion of the pre-processing operations' results in the proposed system, which are as follows:

6.1 Resizing the image

In this step, the image size is standardized to 256 × 256 due to the considerable variation in image sizes ranging from 100 × 100. The smallest size and the most significant size are 5482 × 4096. The primary purpose of this step is to standardize the image sizes and reduce the overfitting that occurs due to the contrast of the large data size. The process of changing the image size is by changing the dimensions of the image (width and height) and not by cutting out parts of the image because the purpose of the proposed system is to reveal all the faces in the image, regardless of their number, and not to reveal one face, so cutting any Part of the image may lead to the loss of some essential details of the image. The selection of the size 256 × 256 was based on experience and, the resizing image operation replaced filters because the images in the dataset used were high resolution and did not need any other operations. Table 1 shows the difference between the original images from the dataset in different cases and sizes due to resizing the images.

Table 1: Resizing image from different sizes to size 256 × 256

Original image	**Size**	**Resizing the image to** (256 × 256)
	1024 × 684	

	305 × 204	
	433 × 568	
	1024 × 712	
	600 × 863	
	1200 × 800	
	461 × 531	
	1024 × 802	

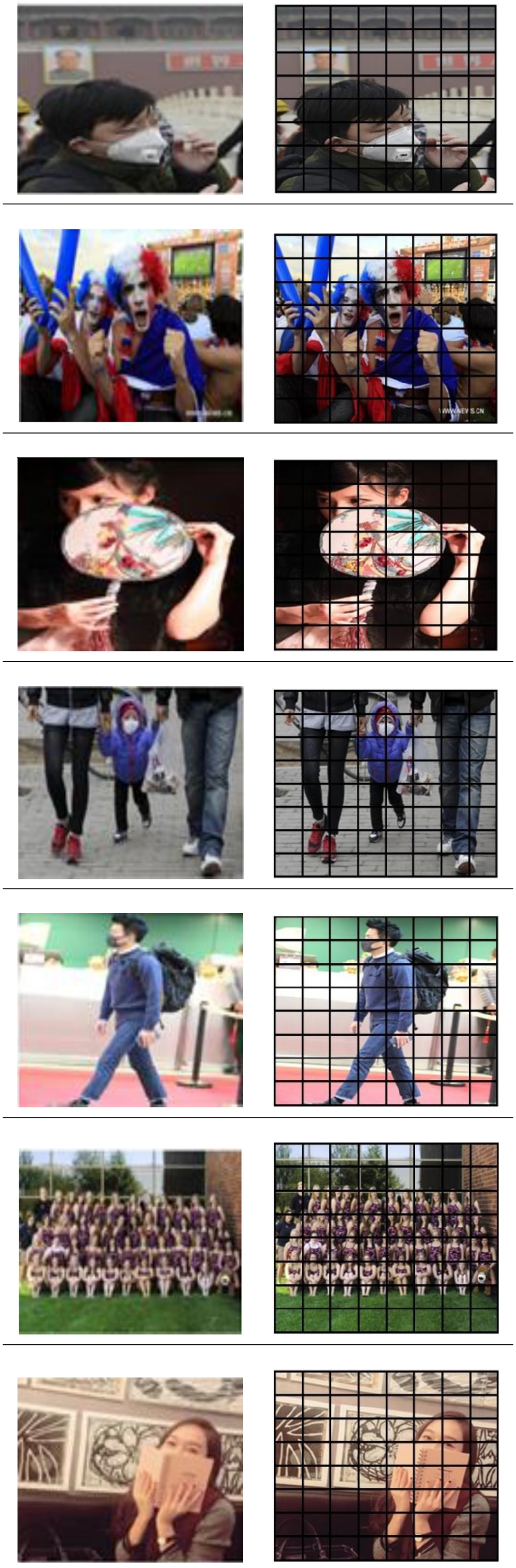
	550×500	
	1024×1415	
	267×161	
	250×377	

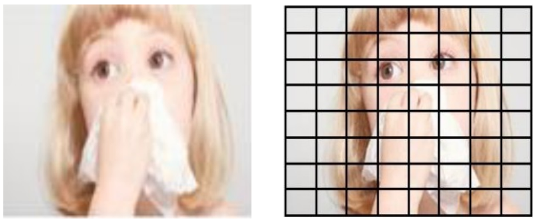
6.2 Grid Concept Applied

Grid applied for creating the bounding boxes around the objects in the images. Bounding boxes are one of the most popular and recognizable image methods used in deep learning in algorithm YOLOv5s that are used in the system; the following Table 2 shows the Annotation in resizing images.

Table 2: Applied The Annotation in Resizing Image

Resizing the image to (256×256)	**Data Annotation in size** (8×8)
	
	



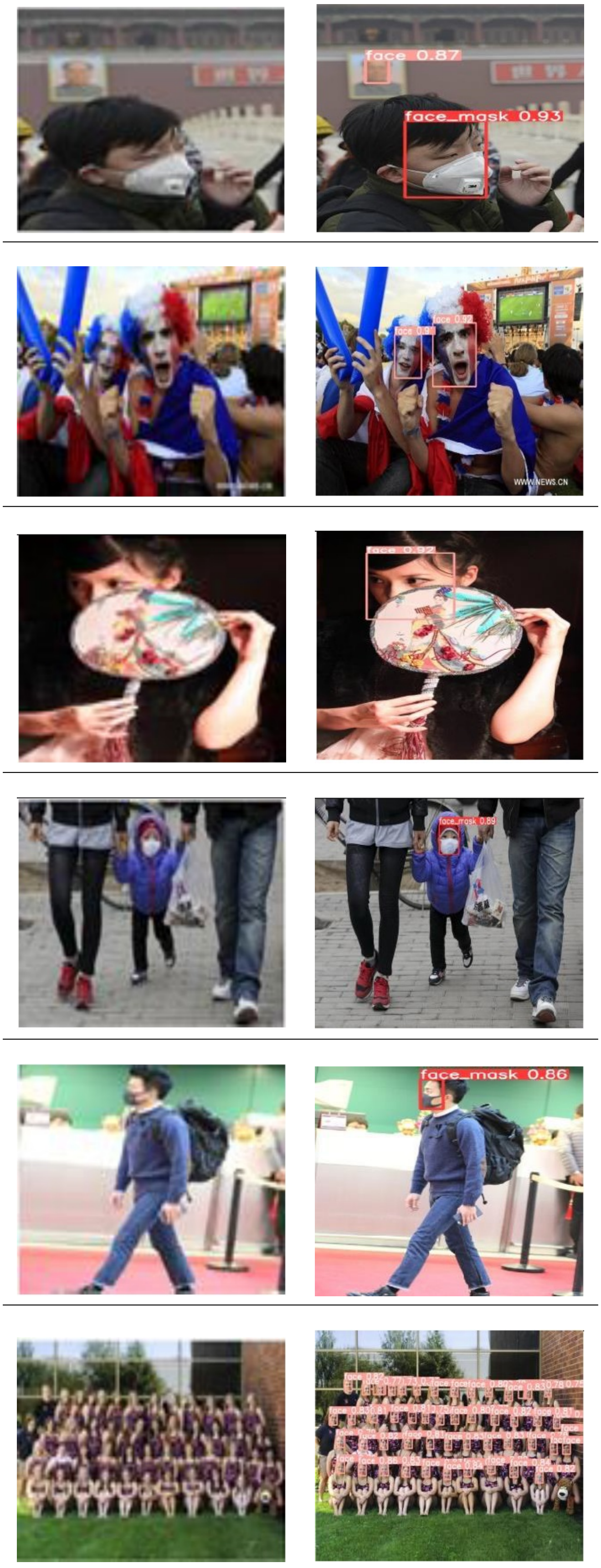


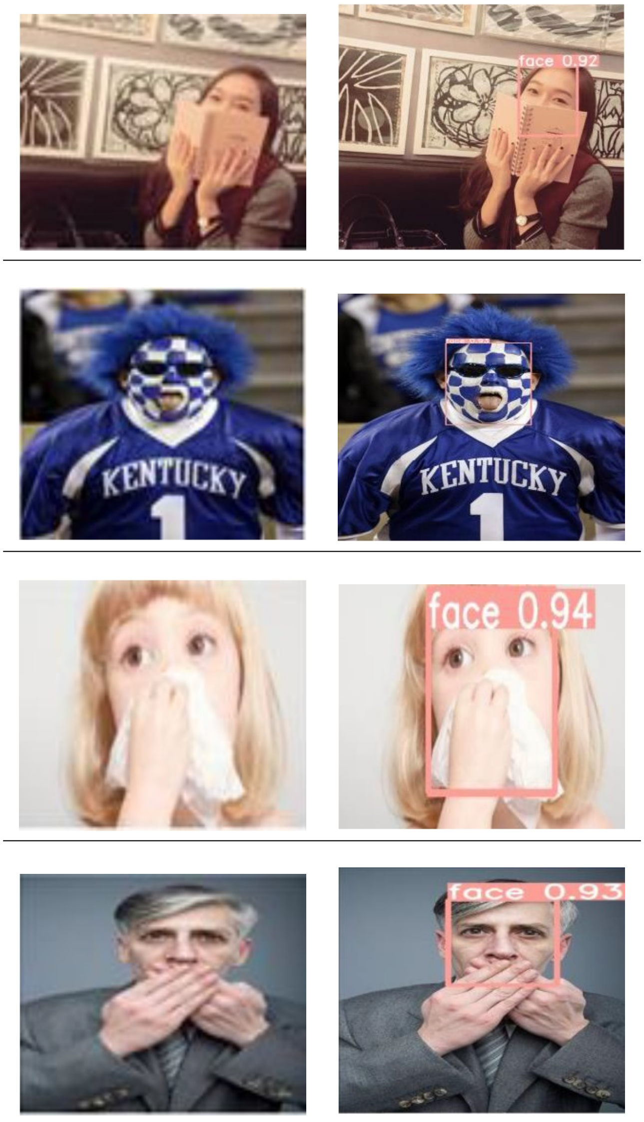
7 Detection Object in Proposed System

Object detection is one of the most popular computer vision models due to its versatility. Mask and non-mask object detection models seek to identify the presence of related objects in images and classify them into two categories. They also train an object detection model to identify all types of face recognition, if it is wearing or not of these objects and classify them correctly. The following Table 3 shows how to discover masks in the previously shown images as models.

Table 3: Overview of Discover Masks and No Mask in Images of Data Set

Image	Detection
	
	


face 0.87
face_mask 0.93
face_mask 0.86



The model can detect the face and the face mask in many options during Real-time, Image, Video, Directory of images, and Stream. The proposed model has overcome many face detection challenges, such as Different face orientations, Non-mask Occlusion, blurred faces, Various Mask Types, small face and face masks, etc. Figure 5 shows the system's ability to detect most of the images that contain challenges correctly.

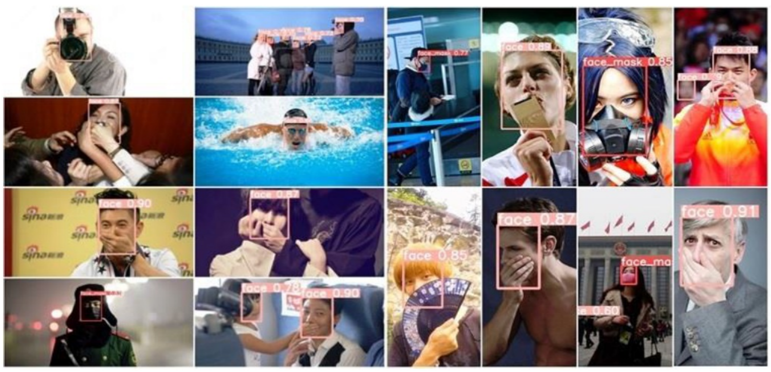


Figure 5: The proposed system has overcome many challenges

Figure 6 shows the system's ability to correctly detect most images containing a small face and face mask.

Figure 6: Detection of the small face and face mask

8 Detection Object In Real-Time Proposed System

In the proposed system, it can detect faces that are wearing a mask in real-time. The goal of the proposed system is to have the ability to work online in real-time and offline by evaluating the results of the data set and the reason for working with this. The method relies on a global data set to assess the system's quality and compare it with the previous work, and the other side works in real-time to use the proposed system in all areas that need to be known. The following images show the system's ability to detect the images in real-time as experiments correctly.

- Experiments one

Figure 7: Experiment One in real-time

- Experiments Two in Video

Figure 8: Experiment Two in video real-time

9 Model Training Results

The results of the training differ according to the number of epochs conducted for each dataset, where four training periods were conducted for each dataset, each period has a different number of epochs, and the results were as follows:

- AIZOO dataset

Table 4: Best results of training the proposed system in different epochs

Epoch Number	**train/ box_loss**	**train/ obj_loss**	**train/ cls_loss**	**metrics/ mAP_0.5**	**metrics/ recall**	**metrics/ precision**
10	0.048966	0.015848	0.0074704	0.87332	0.80215	0.90033
50	0.037971	0.012472	0.0034581	0.92765	0.88934	0.94423
75	0.033379	0.011386	0.002293	0.97212	0.94914	0.97226
100	0.033284	0.011197	0.002248	0.97744	0.95495	0.97349

There are three types of loss shown in Table 4, objectless loss and classification loss. The box loss represents how well the algorithm can locate the centre of an object and how well the predicted bounding box covers an object. Objective loss measures the probability that an object exists in a proposed region of interest. The image window will likely contain an object if the objectivity is high. And Classification loss gives an idea of how well the algorithm can predict the correct class of a given object. The decrease in the values of The box, objectless and classification losses of the training and validation means less loss, which means an improvement in the results and thus an increase in the efficiency of the model. On the other hand, the increase in the values of precision, recall and mean average precision means an increase in the strength of the prediction and, therefore, an increase in the efficiency of the model. he increase in the number of training epochs improves the performance of the proposed system; as shown in Table 4, At the 100 epochs, the system achieved good results, and with these results, the system does not need to increase the number of training epochs any more.

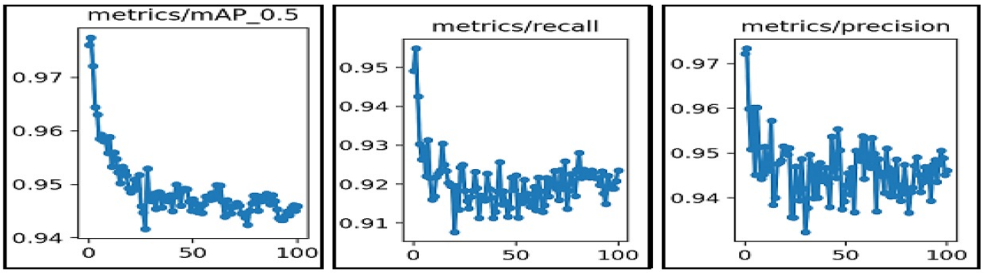


Figure 9: Best results of training the proposed system in the AIZOO dataset

- MoLa RGB CovSurv dataset

Table 5: Best results of training the proposed system in different epochs

Epoch Number	**train/ box_loss**	**train/ obj_loss**	**train/ cls_loss**	**metrics/ mAP_0.5**	**metrics/ recall**	**metrics/ precision**
10	0.023588	0.0098103	0.0020976	0.84375	0.78789	0.92388
50	0.020417	0.0087691	0.0015398	0.84011	0.79014	0.91621
200	0.025171	0.008765	0.0007539	0.94449	0.92705	0.95353
230	0.031123	0.010627	0.0017957	0.94896	0.92646	0.95404

As shown in Table 5, When trying to train the model on 800 epochs, the YOLOv5s algorithm stopped training at 230 epochs because the algorithm noticed that the model results did not improve more by increasing the number of epochs. In figure 10, An epoch consists of one complete cycle through the training data, usually with many steps. A training epoch uses all training data for calculation and optimization (train the model). At the end of each cycle, the epoch is checked in the GPU memory to calculate the boxes, objects, classes, and labels in image size 256×256. And the calculation of the image numbers, label numbers, precision(P), recall(R), and mAP for all classes in the training data to the current epoch. The calculation of the results is repeated for all epochs. In the end, the time the system takes to train all the training data over 231 epochs is calculated, which is 2.192 hours.

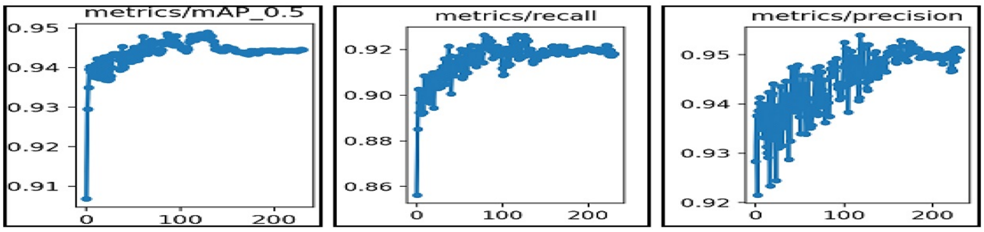


Figure 10: Best results of training the proposed system in the MoLa RGB CovSurv dataset

- Merged dataset

Table 6: Best results of training the proposed system in different epochs

Epoch Number	**train/ box_loss**	**train/ obj_loss**	**train/ cls_loss**	**metrics/ mAP_0.5**	**metrics/ recall**	**metrics/ precision**
10	0.032288	0.010821	0.002108	0.94122	0.90985	0.93628
50	0.031088	0.01056	0.0015883	0.94235	0.91478	0.94508
100	0.029613	0.010171	0.0013873	0.94437	0.91884	0.95334
200	0.025171	0.008765	0.0007539	0.94449	0.92705	0.95353

As shown in Table 6, It has been observed that the training results are constantly improving when the number of epochs is increased. It was also noticed that the results in 100 epochs are very close to the results of 200 epochs, which means that increasing the number to more than 200 will not lead to an increase in the accuracy of the model, so it was stopped at 100 epochs.

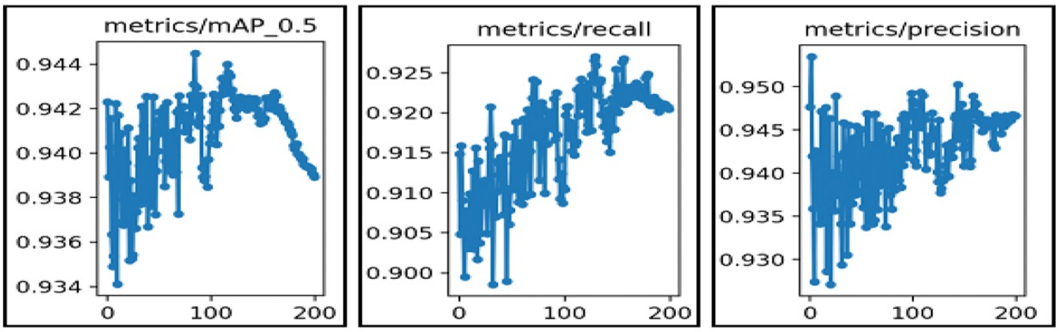


Figure 11: Best results of training the proposed system in the MoLa Merged dataset

10 Evaluation of the proposed detection system

Evaluate the proposed system by using the confusion matrix. That means it is a summary of classification problem prediction outcomes. The number of rights and unsuccessful predictions is totalled and broken down by class using count values. The recommended approach for evaluation will be used, which is the key to the confusion matrix. A detection model's correctly or incorrectly predicted number of occurrences is summarized in a confusion matrix.

Table 7 indicates the values of (TP, FP, and FN) that were extracted after testing the proposed system on the datasets.

Table 7: The result of (TP, FP, and FN) for datasets

Dataset	Class	TP	FP	FN
AIZOO	All	2228	31	23
	Face mask	907	6	7
	Face	1319	28	18
MoLa RGB CovSurv	All	13180	1497	1688
	Face mask	9409	241	1148
	Face	3813	825	508
Merged	All	15455	1417	1700
	Face mask	10316	253	1165
	Face	5121	854	548

10.1 Evaluation the result

And noted When ground truth is present in the image, and the model fails to detect the object, classify it as False Negative(FN).

The following terminology is widely used to refer to counts calculated in a confusion matrix for (datasets) based on the above value of the dataset:

- **Precision:** It is measured by the mean of a data collection and the known value of the thing being quantified, whereas bias is measured by the difference between the mean of a collection of data and the known value of the item being quantified. Precision is the percentage of records in a group the classifier has declared as a positive class, as defined by the precision equation below.

$$Precision(p) = \frac{TP}{TP + FP} \tag{10.1}$$

When the equation (10.1) is applied to all images, the following is the results after testing proposed system on test set:

- Precision (AIZOO dataset)= 0.986
- Precision (MoLa RGB CovSurv dataset)= 0.898
- Precision (Merged dataset)= 0.916

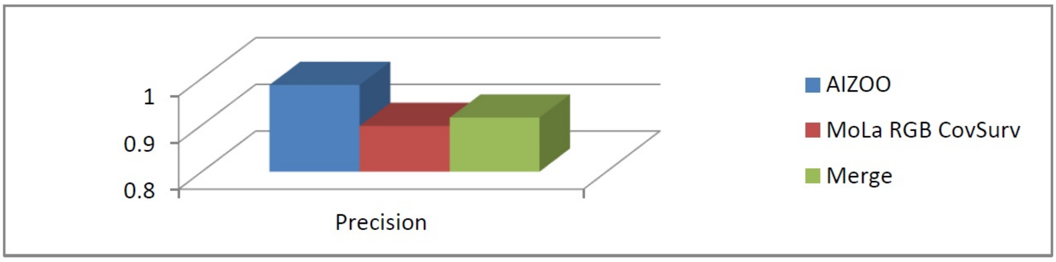


Figure 12: Precision values after testing the proposed system

- **Recall:** The fraction of positive cases correctly predicted by the classifier is measured by the recall, and its value is the same as the actual positive rate.

$$Recall(r) = \frac{TP}{TP + FN} \tag{10.2}$$

When the equation (10.2) is applied to all images, the following are the results after testing the proposed system on the test set:

- Recall (AIZOO dataset)= 0.99
- Recall (MoLa RGB CovSurv dataset)= 0.872
- Recall (Merged dataset)= 0.89

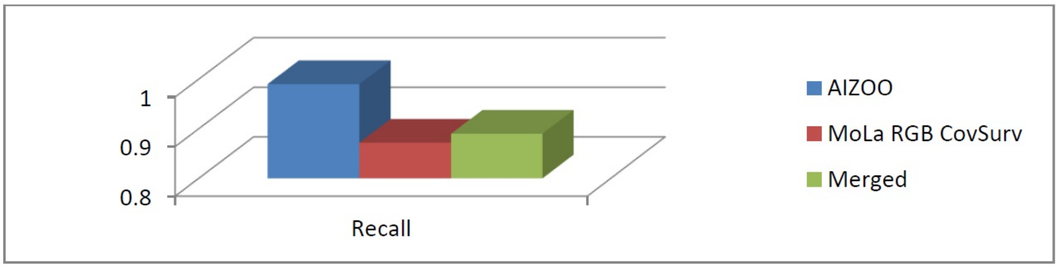


Figure 13: Recall values after testing the proposed system

- **Mean Average Precision (mAP)**

After Precision and Recall are computed, Average Precision (AP), which is the popular matrix measurement for accuracy in object detection, AP can indicate how the accuracy of different algorithms performs with the same dataset. Based on this, it is possible to observe whether the algorithm performs better than other algorithms. It is computed separately for each object category, where mean AP (mAP) is the average of AP (AP is calculated for each class and averaged to get the mAP), which gives the mean Accuracy value.

The following is the results after testing proposed system on test set:

- mAP (AIZOO dataset)= 0.994
- mAP (MoLa RGB CovSurv dataset)= 0.922
- mAP (Merged dataset)= 0.935

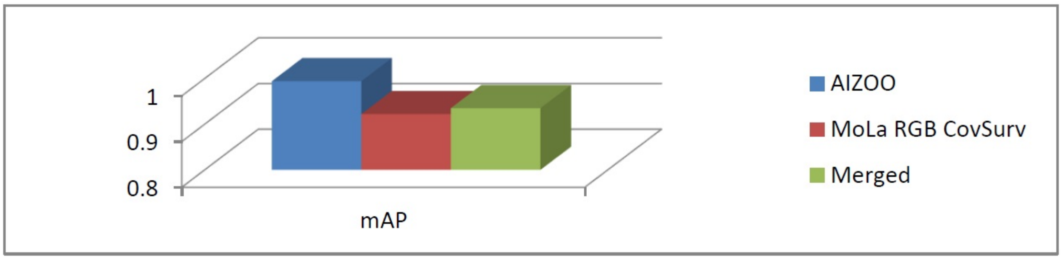


Figure 14: mAP values after testing the proposed system

- **F1:** Score is another evaluation matrix used to measure models' performance. It is also known as F-measure or F-score. It considers both the precision and the recall to be computed, where the F1 Score reaches its best value at 1 (perfect precision and recall) and worst at 0. F1, the harmonic mean of accuracy and recall, as well as the flowing equation, denote F1, and the equation is

$$F1 = \frac{2rp}{r+p} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{10.3}$$

  When the equation (10.3) is applied to all images, the following is the results after testing proposed system on test set:

  - F1 (AIZOO dataset)= 0.99
  - F1 (MoLa RGB CovSurv dataset)= 0.88
  - F1 (Merged dataset)= 0.90

  Figure 15 shows the F1 value for three-class (face mask and face and all class) and the confidence value that optimizes the precision and recall, corresponding to the maximum F1 value for all classes. The value of the F1 increases with the increase in the confidence value.

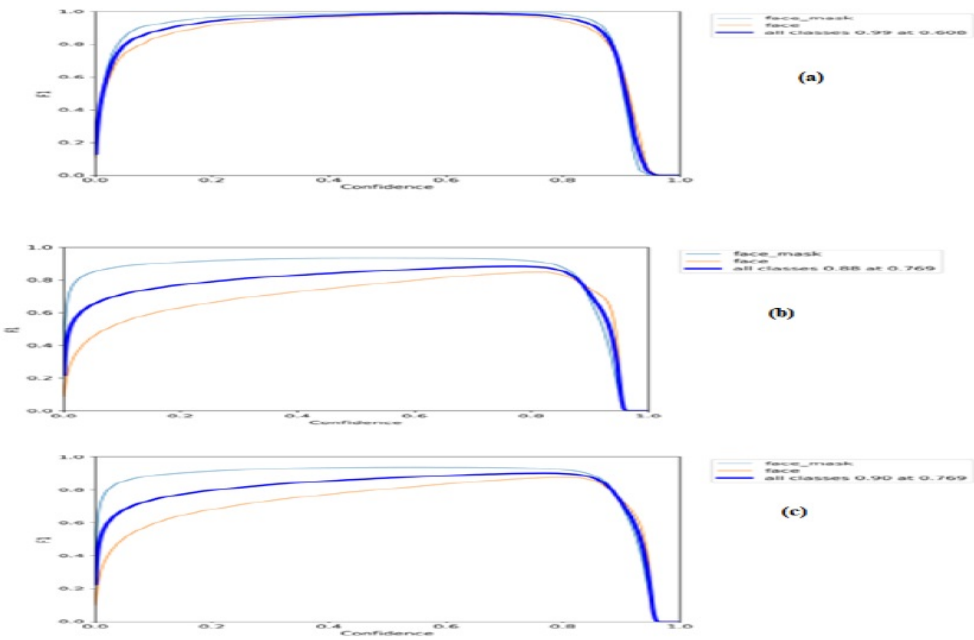


Figure 15: F1 Score after testing the proposed system for: (A) AIZOO dataset, (B) MoLa RGB CovSurv dataset, and (C) Merged dataset

In figure 15 (a, b, c), there are three curves, the blue represents the face mask, the orange represents the face only, and the dark blue represents all cases.

Noted the result was good because in Figure 15(a) at point 0.608, The results higher Accuracy and in Figure 15(b) at point (0.769) and Figure 15(c) in point 0.769 thus satisfactory because the F-measure is the weighted harmonic mean of precision (P) and recall (R) of a classifier, taking $\alpha = 1$ (F1 score). It means that both metrics have the same importance. In the figures above, the confidence value that optimizes the precision and recall.

The following Table 8 is the summary results after testing the proposed system on the test set for each class:

Table 8: Summary result of the proposed system for each class

Dataset	No. of image in test set	Class	No. of labels in class	Precision	Recall	mAP	F1
AIZOO	**1595**	Face mask	914	0.992	0.993	0.995	0.99
		Face	1337	0.979	0.987	0.993	0.98
MoLa RGB CovSurv	**9035**	Face mask	10717	0.975	0.878	0.949	0.93
		Face	4398	0.822	0.867	0.896	0.85
Merged	**10630**	Face mask	11631	0.976	0.887	0.953	0.93
		Face	5735	0.857	0.893	0.916	0.87

11 Conclusion

This study on identifying face and face masks using deep learning techniques and constructing a model that can monitor an area using real-time cameras and a global dataset for evaluation without additional hardware. The conclusion is given below:

1. YOLOv5s algorithm is an algorithm in which it is difficult when any modification is made to it, as many experiments were made to add a neural network to it, which resulted in the algorithm stopping several times before reaching the right place to add the neural network within the algorithm in a way that makes the helpful modification and not leads to stop the algorithm from working to work to increase the accuracy of the results more.
2. Increasing the number of epochs during training is essential in increasing the accuracy of detection in the model. Still, this increase will be to a certain extent, which will not help increase the accuracy of the model but instead stress the system. The results should constantly be monitored to determine the number of periods needed to train the model. Sometimes the algorithm automatically stops training that the model does not learn more as the number of epochs increases, so the automatic stop technique was used when the best epochs were obtained.
3. The proposed system can detect face and face masks with high efficiency and can be used in daily life scenarios; supermarkets, schools, universities, hospitals, and airports are the first candidates in which the proposed model can be severe; it helps to defeat the spread of the Covid-19 virus.

Also, as a future work for the proposed system, it should be strengthened with interfaces to facilitate work on it by all people, even if they are unfamiliar with the calculator.

References

[1] F.A. Alamri, *Contextual information for object detection*, PhD diss. University of Exeter, 2020.

[2] D.A.M. Cota, *Monitoring COVID-19 prevention measures on CCTV cameras using deep learning*, PhD diss. Politecnico di Torino, 2020.

[3] E.R. Davies, *Computer vision principles, algorithms, applications*, Learn. **5** (2018), 2–9.

[4] B. Dhivakar, C. Sridevi, S. Selvakumar and P. Guhan, *Face detection and recognition using skin color*, 2015 3rd Int. Conf. Signal Process. Commun. Network. (ICSCN), 2015, pp. 1–7.

[5] X. Fan, M. Jiang and H. Yan, *A deep learning based light-weight face mask detector with residual context attention and Gaussian heatmap to fight against COVID-19*, IEEE Access **9** (2021), 96964–96974.

[6] Y. Fang, Y. Nie and M. Penny, *Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis*, J. Med. Virol. **92** (2020), no. 6, 645–659.

[7] J. Ieamsaard, S.N. Charoensook and S. Yammen, *Deep learning-based face mask detection using YoloV5*, Proc. 9th Int. Electr. Eng. Congr. iEECON, 2021, pp. 428–431.

[8] K. Karthik, R.P.A. Babu, K. Dhama, M.A. Chitra, G. Kalaiselvi, T.M.A. Senthilkumar and G.D. Raj, *Biosafety concerns during the collection, transportation, and processing of COVID-19 samples for diagnosis*, Arch. Med. Res. **5** (2020), no. 7, 623–630.

[9] P. Li, L. Prieto, D. Mery and P. Flynn, *Face recognition in low quality images: A survey*, arXiv preprint arXiv:1805.11519, 1 (2018), no. 1.

[10] Y. Liu, A.A. Gayle, A. Wilder-Smith and J. Rocklöv, *The reproductive number of COVID-19 is higher compared to SARS coronavirus*, J. Travel Med. **27** (2020), no. 2, 1–4.

[11] A. Liu, C. Zhao, Z. Yu, A. Su, X. Liu, Z. Kong, J. Wan, S. Escalera, H.J. Escalante, Z. Lei and G. Guo, *3D high-fidelity mask face presentation attack detection challenge*, Proc. IEEE Int. Conf. Comput. Vis. **2021** (2021), 814–823.

[12] Q.X. Ma, H. Shan, H.L. Zhang, G.M. Li, R.M. Yang and J.M. Chen, *Potential utilities of mask-wearing and instant hand hygiene for fighting SARS-CoV-2*, J. Med. Virol. **92** (2020), no. 9, 1567–1571.

[13] C. Melo, S. Dixe, J.C. Fonseca, A.H.J. Moreira and J. Borges, *AI based monitoring of different risk levels in covid19 context*, Sensors **22** (2022), no. 1, 1–18.

[14] M.R. Ortiz, M.J. Grijalva, M.J. Turell, W.F. Waters, A.C. Montalvo, D. Mathias, V. Sharma, C.F. Renoy, P. Suits, S.J. Thomas and R. Leon, *Biosafety at home: How to translate biomedical laboratory safety precautions for everyday use in the context of COVID-19*, Am. J. Trop. Med. Hyg. **103** (2020), no. 2, 838–840.

[15] Z. Shanshan, W. Kaisheng, L. Yuxiang and L. Jin, *Face mask detection based on SSDv2 network*, Assoc. Adv. Artif. Intell. **3** (2021), no. 422.

[16] C. Shorten, T.M. Khoshgoftaar and B. Furht, *Deep learning applications for COVID-19*, J. Big Data **8** (2021), no. 1.

[17] S. Singh, U. Ahuja, M. Kumar, K. Kumar and M. Sachdeva, *Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment*, Multimed. Tools Appl. **80** (2021), no. 13, 19753–19768.

[18] B. Sommana, U. Watchareeruetai, A. Ganguly, S.W. Earp, T. Kitiyakara, S. Boonmanunt and R. Thammasudjarit, *Development of a face mask detection pipeline for mask-wearing monitoring in the era of the COVID-19 pandemic: A modular approach*, 19th Int. Joint Conf. Comput. Sci. Software Engin. (JCSSE), IEEE, 2022, pp. 1–6.

[19] J.S. Talahua, J. Buele, P. Calvopina and J. Varela-Aldas, *Facial recognition system for people with and without face mask in times of the covid-19 pandemic*, Sustain. **13** (2021), no. 12, 1–19.

[20] Worldometers, *COVID live*, https://www.worldometers.info/coronavirus/, 2022.

[21] S. Xu, *An improved lightweight YOLOv5 model based on attention mechanism for face mask detection*, arXiv preprint arXiv:2203.16506, (2022), 1–12.

[22] G. Yang, W. Feng, J. Jin, Q. Lei, X. Li, G. Gui and W. Wang, *Face mask recognition system with YOLOV5 based on image recognition*, IEEE 6th Int. Conf. Comput. Commun. ICCC, 2020, pp. 1398-1404.

[23] S. Yang, P. Luo, C.C. Loy and X. Tang, *Wider face: A face detection benchmark*, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn. **2016** (2016), 5525–5533.

[24] C.-W. Yang, T.H. Phung, H.-H. Shuai and W.-H. Cheng, *Mask or non-mask? Robust face mask detector via triplet-consistency representation learning*, ACM Trans. Multimed. Comput. Commun. Appl. **18** (2022), no. 1s, 1–20.

[25] J. Yu and W. Zhang, *Face mask wearing detection algorithm based on improved YOLO-v4*, Sensors **21** (2021), no. 9, 3263.