A Comparative Study of YoloV5.6 Model on Face Mask Detection

[1]G. Prasanth Kumar, [2]Dr. P. Sivakumar, [3]K. Rakesh, [4]PCS Nagendra Setti

Sasi Institute of Technology & Engineringg, Andhra Pradesh, India

[1]*prasanth@sasi.ac.in*, [2]*sivakumarperumal@sasi.ac.in*, [3]*rakeshkancharla@sasi.ac.in*, , [4] *pcns@sasi.ac.in*

Abstract

Worldwide pandemic COVID-19 conditions arose in a plague of risky illnesses everywhere. Wearing a facial covering will assist with forestalling the spread of sickness and hold the individual from getting any aerial irresistible microbes. Numerous strategies were proposed to perceive regardless of whether the individual wearing a cover using various deep learning algorithms. Even though, the existing system did not predict accurately. Hence to solve the challenges a new version of the Yolo algorithm is proposed for face mask detection. The proposed method predicts whether the person is fully masked or improperly or not masked. This paper is to look into a viable technique for facial covering utilizing a deep learning model Yolov5.6. After conducting experimental results using the Kaggle dataset with an alternative number of epochs 20,50, and 100, the model at 100 has the best exhibition precision of 96.8%.

Keywords: *Face Mask Detection, Deep Learning, Yolo algorithm, Yolov5.6, COVID-19, Facial Covering*

1. Introduction

There were a lot of pneumonia sicknesses in the Chinese city of Wuhan in December 2019. A part of the early cases had nitty gritty visiting and working in the fish and live animal market in Wuhan. The disorder was subsequently named COVID-19 when investigations revealed that it was brought about by a newfound Covid. COVID-19 expanded throughout China and the remainder of the world. The WHO proclaimed the outbreak a public health emergency worldwide on January 30th, 2020.

A Covid is an immense infection animal variety. They are comprised of an innate material place included by a lipid envelope with protein spikes. It looks like a crown. The crown is known as a crown in Latin, and this infection is known as the Covid. A few Covids cause disorders in people and creatures. Covid might cause respiratory diseases in people going from the normal cold to additional genuine problems. Among these is the Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV), which was tracked down in China in 2003. The Middle East Respiratory Syndrome Covid (MERS-CoV) was tracked down in Saudi Arabia in 2012, and SARS-CoV-2, the contamination obligated for COVID-19, was found in December 2019. Coronavirus is found to circulate in a variety of species. These viruses can sometimes jump straight from animals to humans; this is known as overflow, and it tends to be brought by a variety of conditions such as a viral mutation or greater interaction between humans and animals.

The drops are delivered through the respiratory framework on the off chance that a contaminated individual is hacking or wheezing. If a human is inside 4 feet, there is a high likelihood, that he can breathe in these disease-causing drops. These drops might stick on the surfaces permitting the infection to get by for quite a while. This way the contaminated individual can turn into an integral justification behind the spread of the infection. To keep the infection from spreading, clinical covers are the most effective way. As per the review, the clinical covers mean careful as well as procedure masks and might be looking like a cup or collapsed. These masks can be joined to the head with strings. They are entirely analyzed to control the filtration, simple breathing, and some time for water resistivity. The exploration looks at the assortment of recordings as well as photos to distinguish individuals who are wearing those clinical masks that are as indicated by the govt. rules. Along these lines, it can significantly help the govt. in making a move against those individuals who don't wear legitimate masks. [1]

Starting from the beginning of the pandemic, wearing a mask in the open has been ordinary in China and other Asian nations. The United States is at present encountering significant pandemic breakouts, with cases and passings growing continuously. People should utilize defensive hardware like masks, as indicated by the CDC (Centers for Disease Control and Prevention). Many individuals, particularly youngsters, who have the disease yet no side effects can spread it to numerous others unintentionally, according to explore. The equivalent is valid for the people who at long last secure side effects anyway dissipated the ailment before being broken down sure. The Centers for Disease Control and Prevention (CDC) has encouraged individuals to wear covers openly in spots where social removing is challenging to forestall the spread of the infection. The CDC's admonition is upheld by a few examinations, incorporating one distributed in the New England Journal of Medicine.

Using a mask when stepping outside during a pandemic has shown to be quite effective in limiting the transmission of coronavirus. It represents being a responsible citizen of a country. Due to the practice of wearing masks daily, some nations, like China and Korea, were able to deal with the Covid-19 in a brief period. It was suggested that you use masks, regardless of the sort of mask you have. Use it and be safeguarded. The mask works as an actual boundary, preventing the infection from entering. Many people unintentionally infected others with the Covid-19 virus. The usage of a mask is required for two reasons. It prevents the virus from entering your mouth or nose through an infected person's sneeze or cough. Due to the spread of the virus, the recklessness of a few people has resulted in the death of countless more. Second, on the off chance, you contact a virus-infected surface then the virus will be stopped by the mask. Many countries mandated the use of masks, and citizens were pushed and monitored to comply. Section II discusses the related work of face mask detection. Section III and IV discuss the methodology and results of the proposed model.

2. Related Works

Face detection research began in 2001, with the creation of handcrafted attributes and the application of standard machine learning algorithms to make an effective discovery and identification-proof classifiers. This approach has several flaws, including a high level of feature complexity and low identification precision. Deep convolutional neural networks (CNN)-based face recognition algorithms have been popular of late as a method for working on the performance of detection.

A. Facial covering Detection utilizing Deep Learning Models

1. CNN

Arijit Malakar et al. (2021) proposed a Convolutional Neural Network (CNN) model and OpenCV for the discovery of facial coverings consistently. This proposed model purposes convolutional layers to prepare a dataset of 12000 pictures of countenances with and without a veil, and the Haar overflow classifier of OpenCV to foresee constant video transfers. As an outcome, using CNN design, the model achieved a precision of 98.8 percent on the preparation dataset and 99.37 percent on the approval set [2].

Ravi Kishore Kodali et al. (2021) built a basic Convolutional Neural Network (CNN) utilizing TensorFlow, Keras, Scikit-learn, and OpenCV to make the algorithm more exact. JavaScript API assists in getting to the webcam for constant facial covering discovery. The suggested work is partitioned into three phases i.e., pre-processing, training of CNN, and identifying faces with and without veils. The primary stage is the pre-processing section, in which the images in the training subject avoid false predictions and upgrade the quality of the images. The subsequent stage is training the CNN model then the CNN group faces with and without masks. Thus, the proposed model accomplished an accuracy of 96% [3].

2. R-CNN and Faster R-CNN

Bingshu Wang et al. (2021) proposed a two-stage technique for overseeing wearing veils using creamer AI computations. The fundamental stage is to use a prepared Faster_RCNN model recognizing facial covers and in the resulting stage, a classifier arranged by BLS is applied to take out foundations. The creator proposes a dataset for Wearing Mask Detection (WMD) that contains 7804 genuine pictures. They foster a profound exchange learning model for the pre-disclosure of wearing veils. They gather a phony dataset containing 17654 train veils, 1936 covers, and 6813 test veils. The proposed model achieved a precision of 93.54% and an exactness of 94.84% [4].

3. LeNet Algorithm

Muhammad Haziq Rusli et al. (2021) utilized Multi-Task Cascaded Neural Network to find the face locale in the dataset. Prior to proceeding to the readiness stage, the AI computation LeNet is utilized to remove the components from the face pictures. LeNet work is used to recognize the face locale and train the specific picture for the best presentation. In this way, the face area was somewhat upset by the facemask, it will isolate the cover includes independently. The proposed model has achieved a precision of 98.61% in the preparation stage and testing stage with the LeNet estimation [5]

4. ResNet 50

Arkaprabha Basul et al. (2021) proposed a computer vision and deep learning framework to perceive facial coverings from images and recordings. It empowers the automatic detection of the face mask. To train the model the most important part is the dataset. The author uses three types of datasets that are publicly available. The Grad-CAM methodology utilizes just limit focuses to take out those from the face and cut the face with the focuses so we can avoid all of the blocks from the image. Later it has been finished on the pre-handled dataset by the ResNet50 plan. Thus, the proposed model has achieved an accuracy of 99.4% compared to other object models [6].

5. MobileNetv2

Samuel Ady Sanjaya et al. (2020) have encouraged a model using AI estimations through the image portrayal technique, MobileNetv2. The model has been made by social occasion information, pre-managing information, isolating the information, testing, and completing the model. The model can be executed on the cameras to hinder the transmission of the Covid by recognizing people who are wearing facial covers. The model has been arranged and gone after for the best show on the first datasets, i.e., the Kaggle dataset and the Real-World Masked Face Dataset (RMFD). The model has achieved a precision of 96.85%. After the execution of the model in 25 metropolitan organizations, the degree of people wearing a veil has a serious solid area of 0.62 [7].

Wuttichai Vijitkunsawat et al. (2020) studies the performance of the three algorithms utilized for masked face detection; KNN, SVM, and Mobile Net calculations. The authors utilized the Prajna Bhandary dataset comprising 1376 images by separating 690 images for people with masks and the rest 686 images for people without masks. After conducting experiments on the three algorithms, the performance of the Mobile Net algorithm is higher than the other algorithms with an accuracy of 88.7% in both image and video datasets. The Mobile Net algorithm is the best and most reasonable for checking regardless of whether an individual wearing a mask [8].

Jesica. J et al. (2021) proposed a new convolutional neural network architecture called Mobile Net Convolutional Neural Network, an effective component for object detection. The first set of data came from Kaggles' Medical Mask Dataset, which includes photographs of people wearing masks as well as descriptions in XML files. The artificially created mask dataset by Prajna Bhandary is accessible at PyImageSearch in the second set of the dataset. A mask is applied to a non-masked image of a person in this duplicate dataset. There are 1376 photos in the dataset, isolated into two classes: 690 images with masks and 686 images without masks. There is a lot of noise and duplicates in the dataset. As a result, the dataset should be pre-processed. The photos are resized, and the MobileNetV2 model transforms the pixel representation of the images into a list. OpenCV is a freely available library for performing computer vision tasks. It delivers machine learning frameworks with libraries, tools, and support and uses several varieties of techniques, including facial recognition and detection video and object movement monitoring, and so on. CNN architectures and other computer vision architectures mostly use it. OpenCV uses the TensorFlow framework to train the neural network. The datasets are partitioned into preparing and testing batches to a 4:1 extent. The training datasets comprise 80% of the aggregate, with the remaining 20% used for testing. As a result, the CNN architecture employing MobileNetV2 for face mask identification achieved 99% accuracy and precision [8].

6. Yolo algorithm

Loey M et al. (2020) proposed an original model for clinical veiled face recognition, zeroing in on clinical cover objects to forestall COVID-19. To create elite execution results for picture acknowledgment, a YOLOv2-based Res Net-50 model was created. To prepare the model, they join the Medical Masks Dataset (MMD) with the Face Mask Dataset to create a new dataset (FMD). In the wake of wiping out bad quality photographs and overt repetitiveness, the dataset now contains 1415 pictures. The proposed identifier incorporates YOLOv2 with ResNet-50 for highlight extraction and distinguishing proof in the preparation, approval, and testing stages. The YOLOv2 acknowledgment network is a convolutional brain network with various convolutional layers, a change layer, and a result layer. As a result, the prescribed identifier was viewed as 81% precise [9].

Sunil Singh et al. (2021) proposed a successful nonstop profound learning-based approach for modernizing the ID of concealed faces, wherein each masked face is perceived persistently through bouncing boxes. They come up with two alternative algorithms for detecting face masks: R-CNN and YOLOV3. Object detection is achieved using both techniques. Both models were developed using a small dataset of 7500 photos. Following thorough model testing, it is clear that the YOLOV3 is quicker than the R-CNN. YOLO algorithms are preferred over R-CNN or other detection algorithms because they perform single-shot detection and have a lot of higher frame rates. In comparison to the R-CNN model, the speed/accuracy is likewise high. In 0.045 seconds, the YOLOV3 model achieved a 55% accuracy, whereas the R-CNN model achieved a 62% accuracy in 0.15 seconds. Training on huge datasets can enhance accuracy even further [10].

Jirarat Ieamsaard et al. (2021) proposed a strong procedure for facial covering area using a profound learning model made by YOLOV5. The YOLOV5 is used to lead the CNN to recognize in case an individual is wearing a veil precisely, improperly, or not by any stretch. There are 853 pictures in the dataset, with three imprints: with a veil, without a veil, and wrong cover. The dataset is divided into three phases: 682 for model planning, 85 for the check, and 86 for model testing. The preparation and facial covering recognition parts of the YOLOV5 system are segregated into two segments. YOLOV5 is used to prepare the model at five particular epochs [20, 50, 100, 300, 500]. Facial covering detection at 300 epochs yields the best results, with a 96.5% precision [11]

B. Others

Xinqi Fan et al. (2021) proposed an original single-shot light-weight facial covering identifier (SL-FMDet), which can identify facial coverings precisely. They proposed another leftover setting consideration module (RCAM) to separate rich setting elements and spotlight on genuine facial coverings, as well as a remarkable helper errand to direct Synthesized Gaussian heatmap relapse for additional segregating highlights for faces with and without veils (SHGR). The AIZOO and Moxa3K datasets were utilized in the examination. The WIDER FACE and Masked Faces (MAFA) datasets were joined to make the AIZOO dataset, which includes more than 8000 pictures. half of the appearances in the WIDER FACE dataset are ordinary, while half of the countenances in the MAFA dataset are cover wearing. For testing, we inspected a subset of 1839 pictures. The Moxa3K dataset contains roughly 3000 pictures, with 2800 being utilized for preparing and the excess 200 being utilized for testing. Investigates Retina Face and Retina Facemask-M were additionally led, and these models performed 1-2% more regrettable than SL-FMDet. At the point when the model was contrasted with other existing models on the AZIOO facial covering dataset, it had a 93.6% exactness, and when it was contrasted with other existing models on the Moxa3K dataset, it had a 55.56% precision [12].

Arjya Das et al. (2020) have proposed a method using AI groups like TensorFlow, Keras, OpenCV, and SciKit-Learn. The proposed methodology separates the face unequivocally and a while later perceives in case it has a cover or not. The proposed procedure involves an outpouring classifier and a pre-arranged CNN which contains two 2D convolutional layers related with layers of thick neurons. The model is prepared, approved, and tried upon two datasets. Relating to Dataset 1, (which comprises 1376 pictures in which 690 pictures are veiled appearances and 686 are exposed countenances) the method accomplished a precision of 95.77%. Dataset 2 has various faces and besides having different shaded concealed faces, (comprising 853 pictures) the procedure accomplished an exactness of 94.58% [13].

After conducting the literature survey, it is observed that the existing machine learning and deep learning algorithms are experimented on artificial face masks as well as in real-time scenarios. Apart from detection, accuracy is the most important thing for model validation. Each model has given different accuracies. To improve the accuracy in detecting a face mask a model is built using the new version on Yolo i.e., Yolov5.6. The proposed version of Yolo gives better exactness contrasted with the current models.

3. Proposed Work

In this paper, the flowchart of the proposed model is introduced beneath. The system follows a series of steps for detecting face mask-wearing from the given dataset.

The dataset has been given as input for data pre-processing and data partitioning. Data pre-handling is a procedure for cleaning and sorting out the crude information to make it reasonable for building and preparing the models. Information apportioning is utilized to separate the information into more modest ones. By information apportioning, the information is parted into three sets i.e., training set, validation set, and testing set.

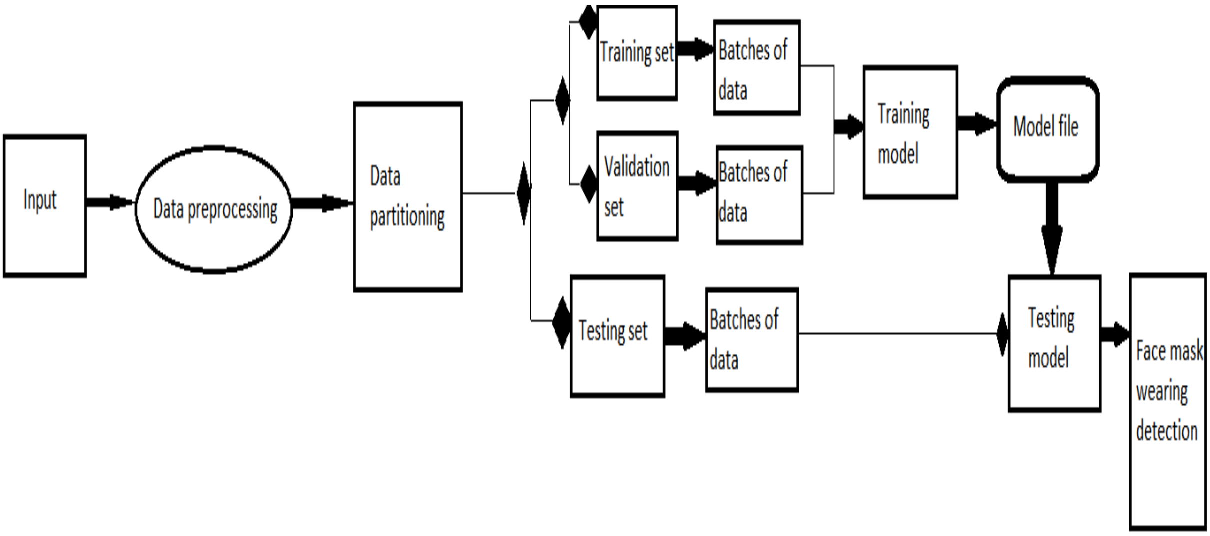

Fig 1. Flowchart of the Proposed System

The training set and testing set are two important parts of the model. While training, it is necessary to teach the proposed algorithm, and testing itself name suggests helps to validate the progress of the proposed algorithm and optimize it for improved results.

A. *Face mask detection using yolov5.6*

For training and testing the model, the possibility of yolov5.6 is utilized to work on the precision of the proposed framework. Yolov5.6 is an updated version of yolov5.

4. Proposed Work

YOLOV5:

Consequences are damned is a best-in-class single-stage object distinguishing proof module that is little and light with the eventual result of being used in edge gadgets. The most recent adaption of this series, YOLOv5, is used here because of its precision, speed, and ability to perceive objects in the essential run. Multi-object tracking, like article identification in an image-based visual examination, is an important part of video analysis. The Pytorch version of YOLOv5 is part of a series of object identification architectures and models that are pre-trained on the COCO dataset and afterward used to track objects using a deep-sort technique. Anything that the YOLOv5 model was trained to recognize might be followed this [14].

Input photos are divided into NxN grids by the YOLO network. To identify the bounding boxes of items in the picture, each grid is handled as a regression problem. In the PASCAL and COCO [15] datasets, this model greatly increased detection performance. The spine, the head, and the indicator are the three central pieces of YoloV5. A CNN fills in as the spine, assembling and shaping picture characteristics at different granularities. To make picture characteristics, the YoloV5 uses the center and scale prediction (CSP) Bottleneck. The head is involved layers that total picture qualities prior to sending them to an estimate computation. The PA-NET is additionally utilized by the YoloV5 for highlight accumulation. The discovery relies upon attributes from the head, as well as box and class expectation steps.

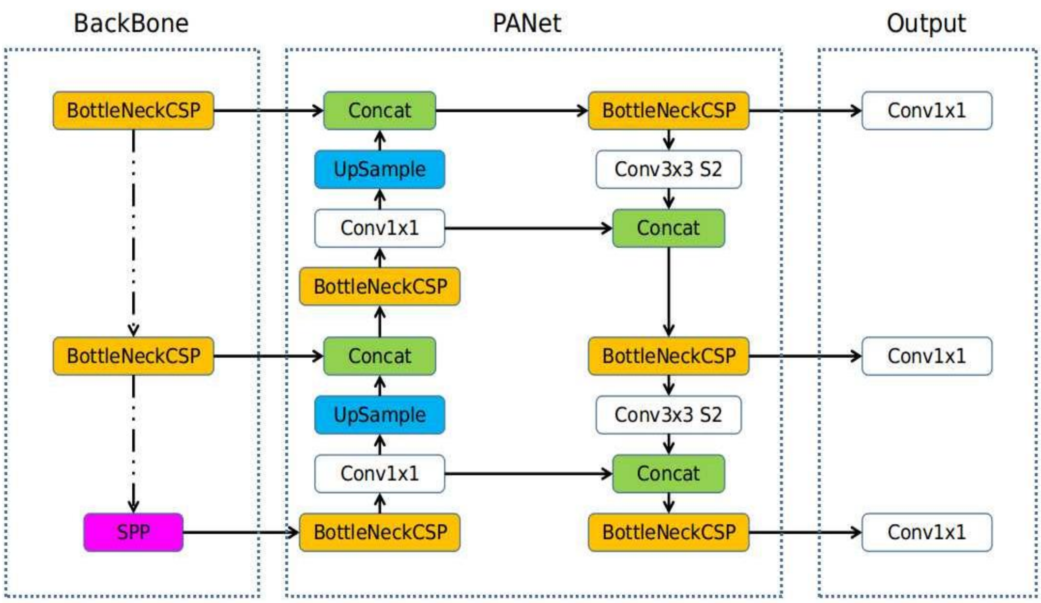

Fig 2. Overview of Yolov5

YOLOv5n, YOLOv5s, YOLOv5m, and YOLOv5x are all versions of the YOLOv5 model, which was introduced in 2020. From YOLOv1 to YOLOv4, the YOLOv5 object detection module experienced several modifications to speed up and increase model performance. The baseline architecture is open-source and can be found from the following git URL: [16]. The yolov5 model's performance is displayed below.

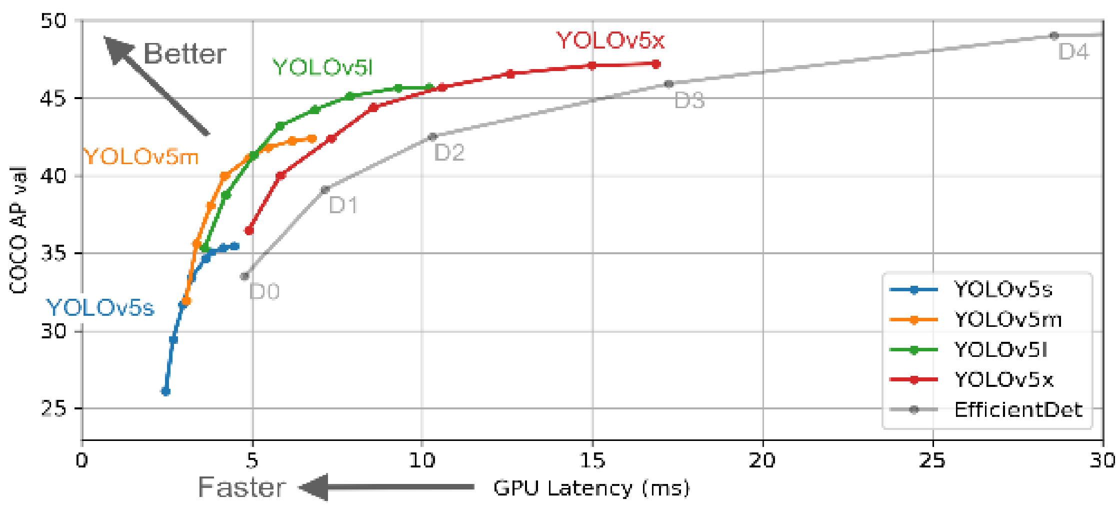

Fig 3. Performance of Yolov5

The Yolov5 model has been modified to increase its speed and accuracy. The performance of the Yolov5.6 has been shown below in a graphical format.

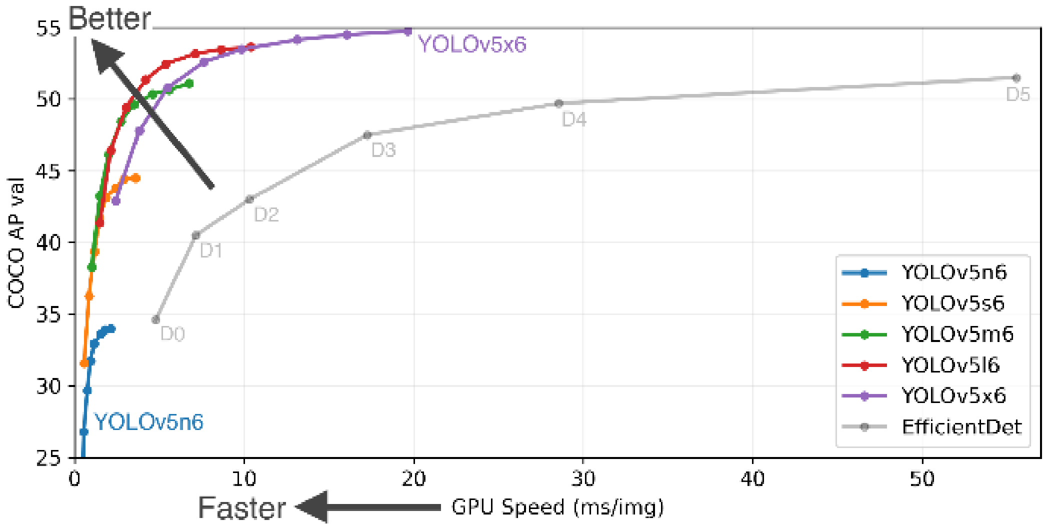

Fig 4. Performance of Yolo5.6

Performance Metrics

To assess object identifier execution across various datasets, average precision (AP) and map are the most by and large utilized measurements. map registers the contrast between the ground truth and expectations of the association. To handle accuracy and review, the accompanying ideas are utilized.

- True positive (TP): right expectation matching ground truth organizes
- False positive (FP): wrong or lost identification of an article
- False Negative (FN): undetected ground truth organizes
- True Negative (TN): expectation when no ground truth exists

Evaluation Metrics:

The performance of the Yolov5.6 model was compared using the precision, recall, loss, and Mean Average Precision (mAP) as evaluation measurements

Precision is a measure of how many of the predicted positives are positive. The mathematical definition is given as

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

A recall is a measure of how many of the true positives are correctly classified. The mathematical definition is given as

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

Average Precision (AP), perhaps the most used metric to measure the accuracy of object detectors, is utilized to find the locale under the accuracy review bend. The mean of the Average Precision (map) is the normal not entirely set in stone for all classes.

5. RESULTS AND DISCUSSIONS

A. *Model Training*:

To develop the facial covering recognition model utilizing yolov5.6, the Kaggle dataset has been utilized which comprises 853 images [41] which include three classes with marks to be specific “With_Mask”, “Without_Mask”, “Incorrect_Mask “.

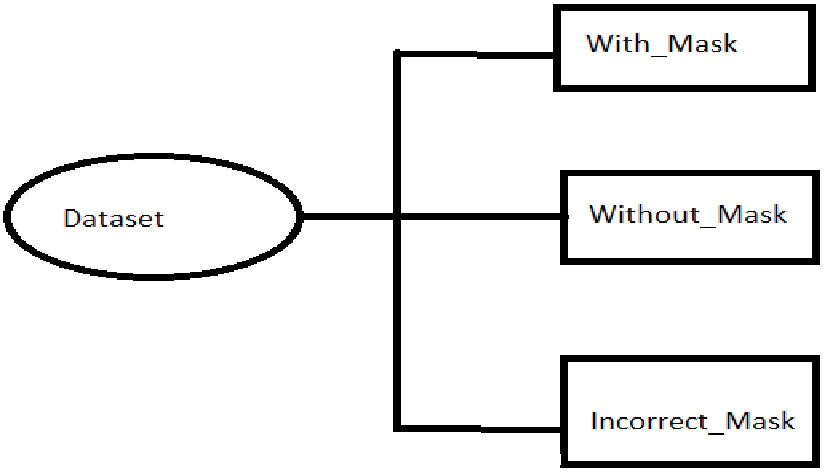

Fig 5. Dataset

Among 853 pictures, 682 pictures are utilized for preparing the model. The model has been prepared at various ages i.e., 20,50, and 100. The accuracy and review were furthermore handled. Each model is approved with 85 pictures. The approval interaction gave the accuracy and review of each class. The presentation of the model at 20, 50, and 100 epochs are displayed beneath.

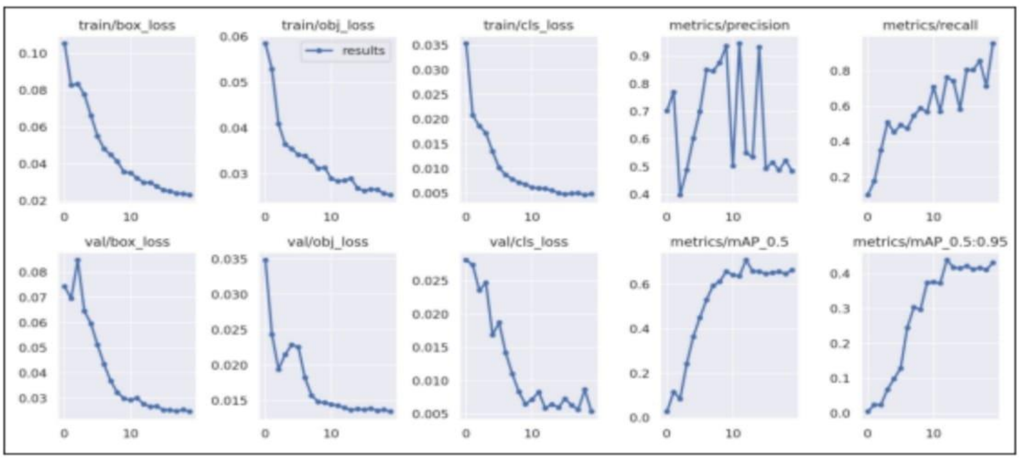

Fig 6. Performance of Yolov5.6 at 20 epochs

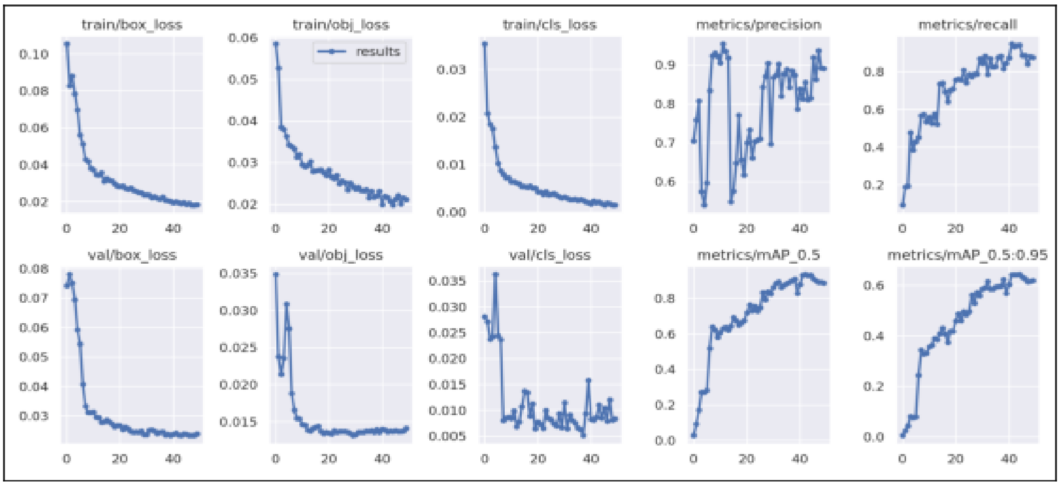

Fig 7. Performance of Yolov5.6 at 50 epochs

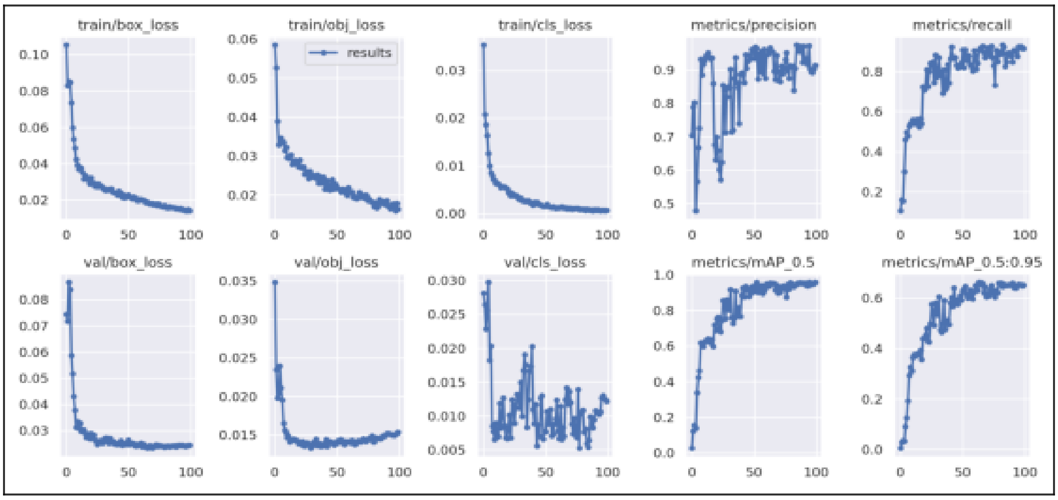

Fig 8. Performance of Yolov5.6 at 100 epochs

The preparation model for each class at 100 epochs gives further developed results than the case at 20 and 50 epochs, as shown above. Also, the preparation model with 100 epochs conveys the most outrageous precision for all classes when appears differently from the preparation model with different epochs. In like manner, the model with 100 epochs gives the best results. A more noticeable number of epochs extends the amount of dealing with steps, which will overall further foster the exhibition of veil identification.

B. *Face Mask Detection Results*

The developed model is approved with the facial covering dataset of 86 pictures. The facial covering identification results are displayed beneath.

Fig 9 (a)

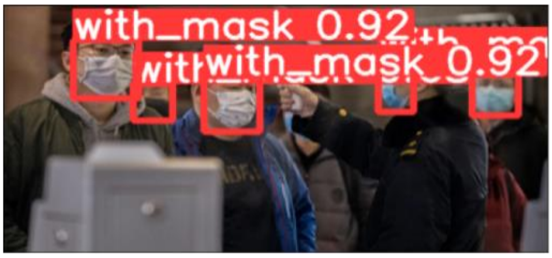

Fig 9 (b)

Fig 9. Face Mask Detection results with three cases shown in (a) and (b)

Comparison of Face Mask Detection Results:

TABLE1: ACCURACY RESULTS COMPARED WITH RELATED WORK

Author	Model	Training dataset	Accuracy
Wuttichai Vijitkunsawat e al.	Mobile Net algorithm	Prajna Bhandary dataset consist of 1376 images.	88.7%
Xinqi Fan et al.	SL-FMDet model	1. AIZOO and Moxa3K dataset were used.	AZIOO = 93.6% Moxa3K = 55.56%
Preeti Nagrath et al.	SSDMNV2 model	Kaggle's Medical Mask include 5521 images	92.64%
Bingshu Wang et al.	Faster_RCNN	Wearing Mask Detection (WMI includes	93.54%
Mohamed Loey et al.	Yolov2 with ResNet-50	Medical Masks Dataset (MMD) Face Mask Dataset FMD)	81%
MINED30	Yolov3 Darknet-53	Mask Dataset	88%
Murat Ucar et al.	Yolov4	Kaggle Dataset of face masks	82.2%
Jirrarat Ieamsaard et al.	Yolov5	Kaggle Dataset of face masks	96.5%
This Study	Yolov5.6	Kaggle Dataset of face masks	96.8%

6. Conclusion and Future Enhancement

In this paper, the deep learning model for facial covering identification utilizing yolov5.6 has been created. The model is prepared at different epochs i.e., 20,50, and 100. When contrasted with the 86 pictures of the dataset, the model at 100 epochs has the best exhibition with an exactness of 96.8%. Aside from detection precision assumes a significant part. This study assists with diminishing the spread of Covid by recognizing the people who are wearing a cover or not. This venture can be done in packed regions like shopping centers, theaters, schools, universities, etc. to lessen the spread of COVID-19. Later on, the model could be recognized dynamically by utilizing a PC vision framework.

References

[1] Sharma, Vinay. "Face Mask Detection using YOLOv5 for COVID-19." Ph.D. diss., California State University San `Marcos, 2020.

[2] A. Malakar, A. Kumar and S. Majumdar, "Detection of Face Mask in Real-time using Convolutional Neural Networks and Open-CV," 2021 2nd International Conference for Emerging Technology (INCET), 2021, pp. 1-5

[3] R. K. Kodali and R. Dhanekula, "Face Mask Detection Using Deep Learning," 2021 International Conference on Computer Communication and Informatics (ICCCI), 2021, pp. 1-5,

[4] B. Wang, Y. Zhao and C. L. P. Chen, "Hybrid Transfer Learning and Broad Learning System for Wearing Mask Detection in the COVID-19 Era," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-12, 2021, Art no. 5009612.

[5] M. H. Rusli, N. N. A. Sjarif, S. S. Yuhaniz, S. Kok and M. S. Kadir, "Evaluating the Masked and Unmasked Face with LeNet Algorithm," 2021 IEEE 17th International Colloquium on Signal Processing & Its Applications (CSPA), 2021, pp. 171-176

[6] Basu, Arkaprabha & Ali, Md. (2021). COVID-19 Face Mask Recognition with Advanced Face Cut Algorithm for Human Safety Measures

[7] S. A. Sanjaya and S. Adi Rakhmawan, "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), 2020, pp. 1-5,

[8] W. Vijitkunsawat and P. Chantngarm, "Study of the Performance of Machine Learning Algorithms for Face Mask Detection," 2020 - 5th International Conference on Information Technology (InCIT), 2020, pp. 39-43

[9] Loey M, Manogaran G, Taha MHN, Khalifa NEM. Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. *Sustain Cities Soc*. 2021;65:102600

[10] Singh, S., Ahuja, U., Kumar, M. *et al.* Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment. *Multimed Tools Appl* **80,** 19753–19768 (2021).

[11] J. Ieamsaard, S. N. Charoensook and S. Yammen, "Deep Learning-based Face Mask Detection Using YoloV5," 2021 9th International Electrical Engineering Congress (iEECON), 2021, pp. 428-431

[12] Yu, J.; Zhang, W. Face Mask Wearing Detection Algorithm Based on Improved YOLO-v4. *Sensors* **2021**, *21*, 3263.

[13] A. Das, M. Wasif Ansari and R. Basak, "Covid-19 Face Mask Detection Using TensorFlow, Keras and OpenCV," 2020 IEEE 17th India Council International Conference (INDICON), 2020, pp. 1-5,

[14] Zhu, L.; Geng, X.; Li, Z.; Liu, C. Improving YOLOv5 with Attention Mechanism for Detecting Boulders from Planetary Images. Remote Sens. 2021, 13, 3776. [CrossRef]

[15] Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft coco: Common objects in context. In European Conference on Computer Vision; Springer: Cham, Switzerland, 2014.

[16] https://www.kaggle.com/datasets/andrewmvd/face-mask-detection