INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Real Time Face Mask Detectionusing Yolov5 and OpenCv

Rakshit Umesh Shetty, Prof. Ghanashyam Phadke, Shubham Niraj Kumar, Ashish Kantilal Tiwari
[1]*Department of Electronics EngineeringKJSIEIT, 75 B.Tech ETRX*
[2]*Project GuideKJSIEIT*
[3]*Department of Electronics EngineeringKJSIEIT, 43 B.Tech ETRX*
[4]*Department of Electronics EngineeringKJSIEIT, 81 B.Tech ETRX*

Abstract: COVID-19 cases have been a big threat to human mankind with different variants found worldwide infecting many people alike. The whole world struggled through different waves of coronavirus and is struggling to reduce the spread effectively. The use of mitigation and suppression was practised throughout and it was concluded use of suppression couldn't come without economic downfall for long term. With results from China and South Korea it was clear that use of masks helped reduce spread within social gatherings ,to ensure everyone wears a mask is one thing and to wear a mask properly covering their nose and mouth fully leaving no part exposed was other problem to tackle. This paper attempts to develop effective yet simple approach to monitor real time data of people wearing masks. The main aim of this project is to help identify and monitor gatheringin public places where wearing mask is essential as it takes a week to show symptoms and anyone can get infected, results show if two individuals are wearing a mask the risk of virus spreading is at its lowest. This model successfully recognizes ifany individual is wearing face mask or not and also determinesif the face mask covers nose and mouth only then it recognizesit as an masked individual.
Keywords: Computer Vision, Real Time Detection, Deep Learning, Yolo Approach, Video Detection.

I. INTRODUCTION

The coronavirus pandemic has produced an environment of uncertainity and terror as this virus can transmit through the respiratory system or through phyiscal contact.COVID-19 has killed more than six million people around the globe, in India alone it has killed over 500000 people and infected over 43 million people. According to the research conducted by medical health experts people's daily behaviour played a very important role in multiplication of the virus from one person to another. Chinaas a country being the first to experience the outbreak of COVID-19 has now set a benchmark in prevention from the disease. Wearing masks and maintaining social distance were the two main important steps to decrease the spread of virus. The main problem was people were unable to follow the guidelines properly so wearing masks in public areas needed supervision. So in this process to replace manual inspection method with a machine learning method and use YOLOv5 which is the most powerful object detection at present which could be a very accurate method especially in the supervision of wearing masks in public places. Computer vision technology uses a variety of imaging systems instead of visual organs as input means using computers to replace the brain to complete the processing and interpretation of visual information. WHO report points out that there are two ways of coronavirus spread, one being the respiratory droplets and other any type of physical contact. Droplets are produced when a infected person coughs or sneezes these droplets are then passed to the surrounding environment. If any other individualis present less then 4 feet away from the infected person then, there are high chances he can inhale these infection-causing droplets. Virus passed through these droplets can also live onto metal or plastic surfaces for days even without any human interaction. To prevent spread of virus, medical masks arethe best bet.Masks are most effective when both individual wear masks the chances of tranfer of virus through respiratory channels decrease significantly.

A. Objective of Developed Model

There are no applications which detect face mask in real time with accuracy most of the efficient face mask detection applications detect either a person wearing mask or not, but ourmodel classifies and detects individual wearing mask correctlyor not and even if someone is not wearing mask altogether,an efficient system which analyzes video feed in real time to detect if people are wearing mask or not and if mask doesn't cover the whole mouth and nose its detected as improper way of wearing mask this helps in awareness and can alsocan be helpful in residential districts or densely populatedareas or public transport spots. This project uses deep learning classification using Computer vision through use of OpenCv and Yolov5 to detect face masks on people in three different categories.

II.   RELATED WORKS

Face detection is a computer technology that determines the locations and sizes of human faces in arbitrary (digital) images. It detects facial features and ignores anything else, such as buildings, trees and bodies. Human face perception is currently an active research area in the computer vision community[2]. A structural face construction and detection system is presented in [12]. The proposed system detects the different lightning, rotated facial image, skin color etc . Human face localization and detection is often the first step in applications such as video surveillance, human computer interface, face recognition and image database management. Locating and tracking human faces is a prerequisite for face recognition and/or facial expressions analysis, although it is often assumed that a normalized face image is available[2]. COVID-19 epidemic has affected many people across the world creating a health crisis around the globe. World Health Organization (WHO) recommends that covering face with mask is one of the best protection to minimize COVID infection. COVID-19 pandemic has forced governments across the planet to impose strict lockdown to stop the spread of virus. Many health experts says that wearing face masks while at public places clearly reduces the spread of the virus. A Machine Learning (ML) based face mask detection with face recognition and alert has been proposed in [4].Putting on a face mask can restrict the spread of the virus. In manycases, coronavirus can be asymptomatic too. As it is rightly said, prevention is better than cure, one should wear a face mask while coming in contact with people. By doing this, an individual ensures his safety, another person's safety, and in this way helps to curb the spread of the disease. The World Health Organization (WHO) as well as the Centers for Disease Control and Prevention (CDC) has suggested the use of face masks for decreasing the spread of the virus[7].Taking maximum advantage of our webcam, the writer used OpenCV (Open Source Computer Vision Library) to perform face detec-tion in real-time from a live stream. It is common knowledge that videos are comprised of frames that are still images. From a video, face detection was carried out in every frame. There is no significant difference between face detection in still images and video streams of real-time[11].there is a gap between current face detection performance and the real world requirements. To facilitate future face detection research, we introduce the WIDER FACE dataset, which is 10 times larger than existing datasets. The dataset contains rich annotations, including occlusions, poses, event categories, and face bounding boxes.[9]

III.   METHODOLOGY

A.   Data Collection

Dataset was created and labelled for three classes using ImgLib python library which converts and creates xml data format to be used around the object to be detected using bounding boxes but due to the constraint of not many sample size present we decided to use Kaggle Repository which consisted of many images with multiple people, different angles and different lightning angles as well.

B.   Dataset

Dataset used from kaggle repository labeled face mask detection contained 853 images belonging to 3 different classes as well as bounding boxes in the PASCAL VOC format.The 3 classes of images are :
- With mask;
- Without mask;
- Mask worn incorrectly.

C.   Dependencies and Libraries

- *Python:* Python is used as a Programming language
- *Numpy:* Numpy is a library in python which is used for comprehensive mathematical functions, random number generators, linear algebra routines, Fourier transforms and much more this helps make sense of the mathematical functions to operate on arrays
- *Xml:* This library is used to extract information from pictures
- Torch PyTorch or torch is an open source machine learning framework based on the python Torch library which is mostly used for complex operations such as NLP and Computer Vision.
- wandb
- *Yolov5:* Yolov5 is the lastest version of Yolo which consists of a object detection architecture and pretained models on coco128 which consists of 128 images which are useful to work with before developing a model which has a large dataset. Yolo is short for You Only Look Once which is meant to mean how fast it is with respect to other algorithms it divides any image into a grid system. Different cells in the grid are responsible for detecting objects within itself. This approach uses a

single neural networkto process the entire picture, then separates it into parts and predicts bounding boxes and probabilities for each component.

- *OpenCv:* OpenCV is a library of programming functions mainly aimed at real-time computer vision. OpenCv makes it easyfor a machine to understand and interact with real world environment and make sense of it. Opencv has different functions from getting video and image feed throughcamera and also saving and extracting information from images

D. Training Model

For Training we use Yolov5s model on coco128 by spec- ifying Dataset, Batchsize, image size and epochs. Either pre- trained weights or randomly initialized weights can be usedbut here we use pretrained weights as these are recommended.

Training model needs labeling ,annotation, data cleaning anddata acquisition before training can be started

E. Object Detection

Object detection is one of the most prominent computer vision tasks. In a nutshell, given an image, an object detector will find:
- the objects in image which is face mask in our system
- types or classes
- bounding box containing coordinates of face mask in theimage

F. Real Time Detection

Real-time object detection is detecting object in real-time with fast inference while also maintaining high level of accuracy with respect to time. Object detectors like YOLOv5s are pretrained algorithms to detect objects. Training consists ofusing a bunch of images with respective annotations or labels containing information to adjust the model and teach it howto detect objects. We use this to our advantage and extract datafrom video as video is a series of images frame by framewe gather information in real time and detect face masks as objects
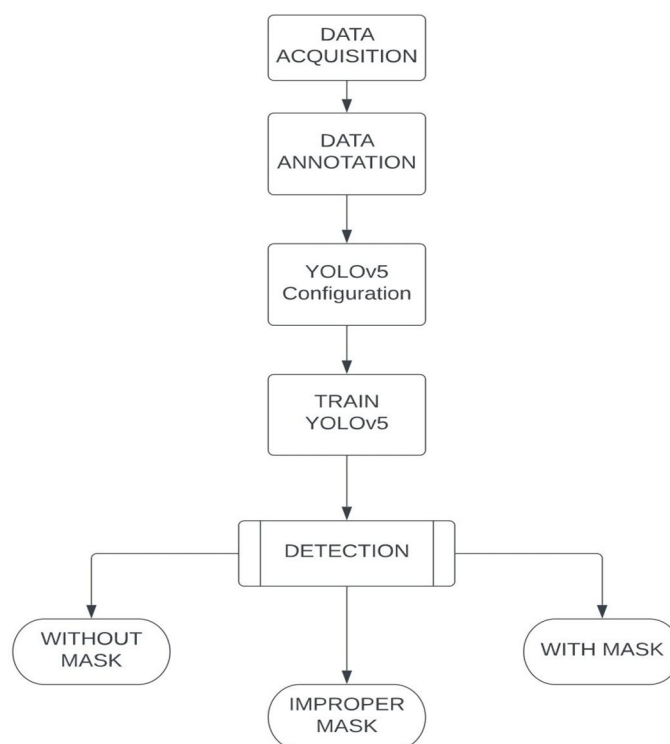

Fig. 1. Project Implementation

Real time Detection for our trained model using OpenCv gathers information with video capture function inbuilt within cv2 package. We resize the frame according to the width and height we need and use different functions to render information from image we then display video feed frame by frame with bounding boxes detecting face mask on the samesystem

IV. RESULTS

In this section we will go through our analysis and findings of our developed model. This model works on test data and classified accurately people in three different classes which were people who wore mask to those who were not wearing any mask and also those who wore the mask but incorrectly not fully covering their nose or mouth. Below are the training and validation results
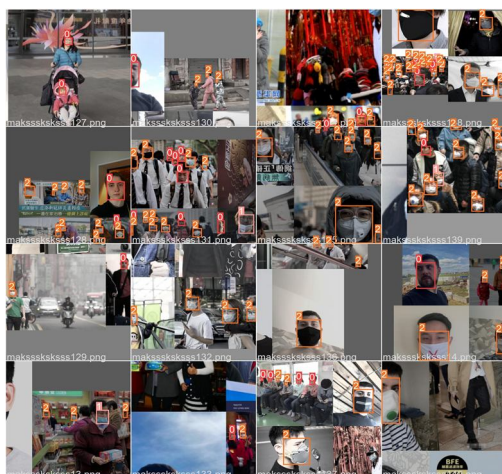

Fig. 2. Validation batch 2 labelling results

Fig. 3. Validation batch 2 labelling results

A. Graphs for visual representation of Precision, recall, mean Average Precision

Looking at the statistics of the results, Considering the precision of the developed model. Precision is a ratio between the number of positive results and the number of positiveresults predicted by the classifier. From the precision graph shown it can be said that the model is performing well asthe precision value for all three types of images from dataset increase overtime in the training process. Confidence level being 1 which is the highest at 0.915 precision. PrecisionGives the fraction of correctly identified as positive out ofall predicited positives. Precision $=TP/TP+FP$
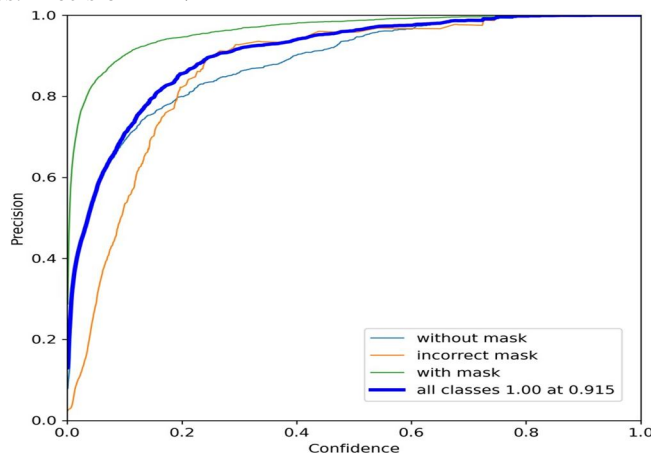

Fig. 4. P Curve Graph

Recall graph of the developed model shows the recall value of the developed model over time .It is the ratio between the correct positive results found from and all samples found tobe positive.
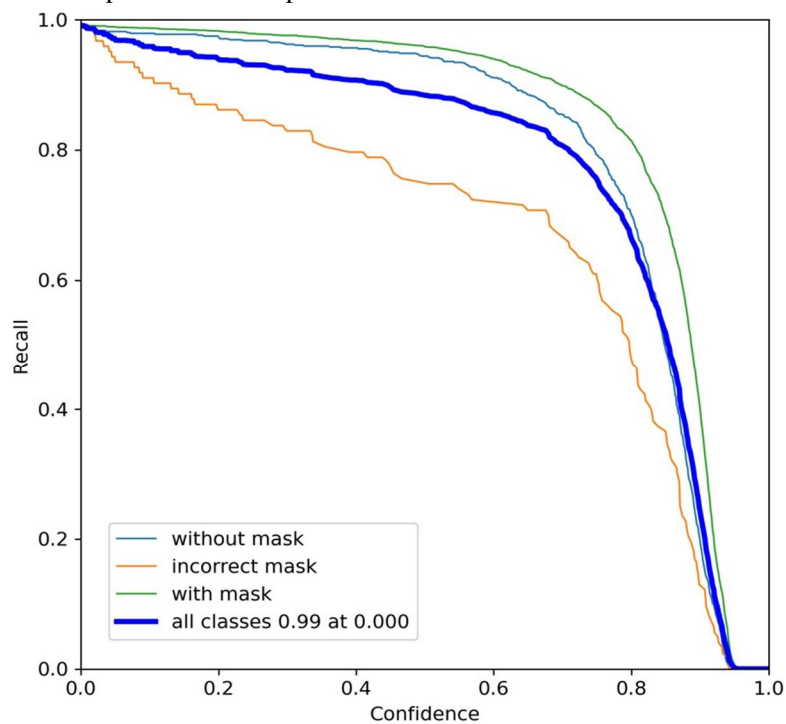

Fig. 5. R Curve Graph

We use confusion matrix because as the classes increases which in our case is 3 We can not rely on a single valueof accuracy in classification when the classes are imbal- anced.Confusion matrix consists of :

True Positive: You predicted positive and it's true.

True Negative: You predicted negative and it's true.

False Positive (Error): You predicted positive and it's false. False Negative(Error): You predicted negative and it's false.
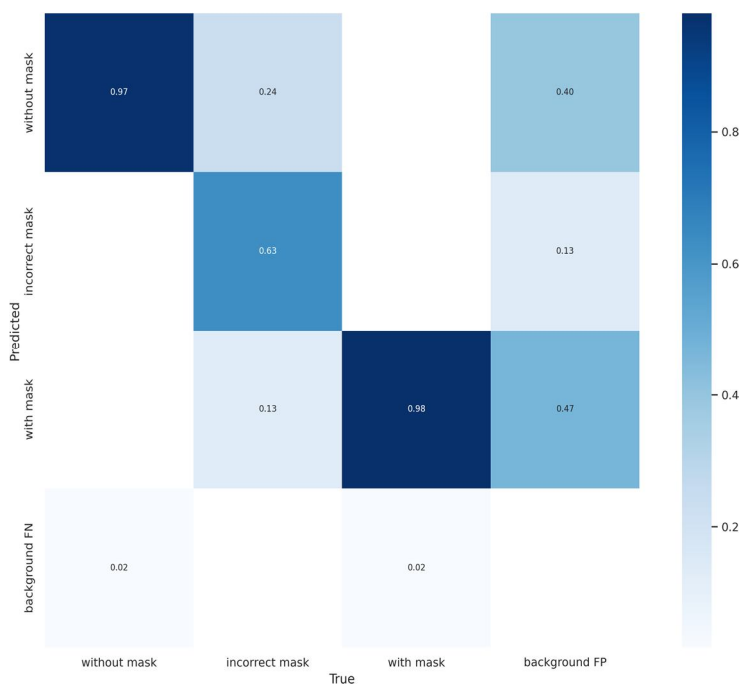

Fig. 6. Confusion matrix

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

B. Weights and Biases

1) Wandb is an Experiment Tracking Tool: WandB is a central dashboard to keep track of your hyperparameters, system metrics, and predictions so you can compare models live

TABLE I

SPECIFICATIONS OF BEST MODEL SAVED

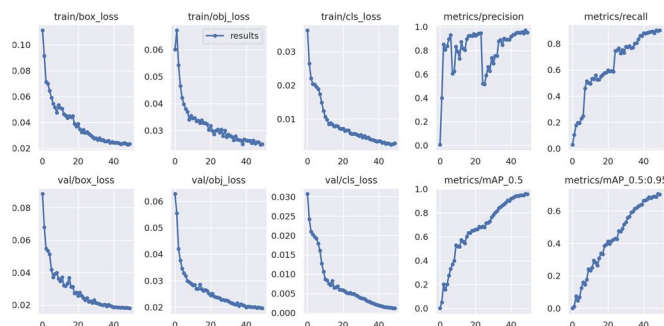| Image Annotation	METRICS		
	Precision	*Recall*	*mAP @ 0.5*
ALL IMAGES	95.2%	90.2%	95.7%
WITHOUT MASK	91.2%	95.1%	97.4%
INCORRECT MASK	96%	78.9%	90.5%
WITH MASK	98.3%	96.6%	99.1%

Fig. 7. Analysis of test batches

Fig.7 graphs show all the results acquired while training the developed model through the use of wand for visualization. This model has been tested using YOLOv5.Training process was completed by running the model with the use of Google Colab which is a Virtual machine developed by google that also provides Gpu memory for faster execution as our project was fully software oriented and needed high computational power which was possible due to use of GoogleColab's high RAM and GPU memory to execute and saveour model for then using the best saved model for Real time Detection using OpenCv for computer vision. Below Fig 8shows results obtained in real time through the system.

Fig. 8. Final Results

System doesn't produce accurate results when light condi- tions are not good and cannot detect face mask or gives results which are not accurate as seen in Fig. 9

CONDITION	EXPECTED RESULT	OBSERVATION	RESULT OBTAINED
Low Light Detection	To observe if detection is possible with low light		UNSUCCESSFUL

Fig. 9. Low Light Detection Results

V. CONCLUSION

With proper supervision and training model checking all the parameters with different weights and analysing them according to different graphs and matrix we conclude that our developed system works as intended with high base level of accuracy. We have developed a system which can continuously and accurately monitor an unsupervised area where there may be large gatherings happening a crowded part of a city in general and with use of real time detection our system can continuously analyse the situation and gather information about people wearing mask or not and also analyse if people are not wearing mask properly as recommended by official authorities like covering mouth and nose properly. This system can be installed any where such as supermarkets or public transport or public places as it requires no extra hardware equipment only a real time camera.

This system can be used to check or verify all customers wearing face masks and can be used as a mitigation tool. As we know from past couple year suppression and lock- downs slow the economic growth and also forced a lot of people to quarantine like they were detained against their will. Suppression can be a useful technique but not for long and with huge groups of people this is where our system comes in use as to help check people wearing masks properly or not this can be used at entrances. There is no need to process the video to check the footage afterwards detecting in real time saves time and with high accuracy our developed model uses real time feed from any camera devices which can provide real time video from anywhere which decent lightning conditions to detect and analyse a person wearing mask or not.

This system has some limitations also. For example, some- times it detects accurately if a person has worn the mask or not only when the person is directly forward facing towards the camera . For example, it is quite useful in supermarkets, movie theatres and airports where people have to go through a single direction to enter or exit. One open problem is to improve this system to detect the faces of the people who are not directly facing the camera or if the place is not well lit this can be because of the quality of the camera used where it is hard to render information out of video where quality of video is not clear.

REFERENCES

[1] B. S. Manjunath, R. Chellappa and C. von der Malsburg, "A feature based approach to face recognition," Proceedings 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1992

[2] Mamata s. Kalas, REAL TIME FACE DETECTION AND TRACK- ING USING OPENCV , International Journal of Soft Computing and Artificial Intelligence, ISSN: 2321-404X, Volume-2, Issue-1, May-2014

[3] Walid Hariri, Efficient Masked Face Recognition Method during the COVID-19 Pandemic, pp.1-7, July 2020

[4] B. A, N. K. M, A. Kumar Sivaraman, R. Vincent and R. M, "Mask Detection in Crowded Environment using Machine Learning," 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), 2021, pp. 1202-1206

[5] F. M. J. Mehedi Shamrat, S. Chakraborty, M. M. Billah, M. A. Jubair, M. S. Islam and R. Ranjan, "Face Mask Detection using Convolutional Neural Network (CNN) to reduce the spread of Covid-19," 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), 2021, pp. 1231-1237

[6] V. KumarB.V.P, N. S. Murthy Sharma and K. Lal Kishore, "A Technique to Reduce Glitch Power during Physical Design Stage for Low Power and Less IR Drop", International Journal of Computer Applications, vol.39, no. 18, pp. 62-67, 2012. Available: 10.5120/5086-7450 [Accessed 28 August 2020].

[7] M. Inamdar and N. Mehendale, "Real-Time Face Mask Identification Using Facemasknet Deep Learning Network", SSRN Electronic Journal, 2020. Available: 10.2139/ssrn.3663305.

[8] world2020announces,WHO announces COVID-19 outbreak a pandemic.2020,World Health Organization and others,2020

[9] S. Yang, P. Luo, C. C. Loy, and X. Tang, "Wider face: a face detection benchmark," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 5525–5533, Las Vegas, NV, United States, 2016.

[10] A. Mikolajczyk and M. Grochowski, "Data augmentation for im- proving deep learning in image classification problem", 2018 Inter-national Interdisciplinary PhD Workshop (IIPhDW), 2018. Available: 10.1109/iiphdw.2018.8388338 [Accessed 28 August 2020]

[11] "Real-time Object Detection and Recognition Using Deep Learning with YOLO Algorithm for Visually Impaired People", Journal of Xidian University, vol. 14, no. 4, 2020. Available: 10.37896/jxu14.4/261

[12] V. S.V, M. Katti, A. Khatawkar and P. Kulkarni, "Face Detection and Tracking using OpenCV", The SIJ Transactions on Computer Networks Communication Engineering, vol. 04, no. 03, pp. 01-06, 2016. Available:10.9756/sijcnce/v4i3/0103540102

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ⊙ (24*7 Support on Whatsapp)