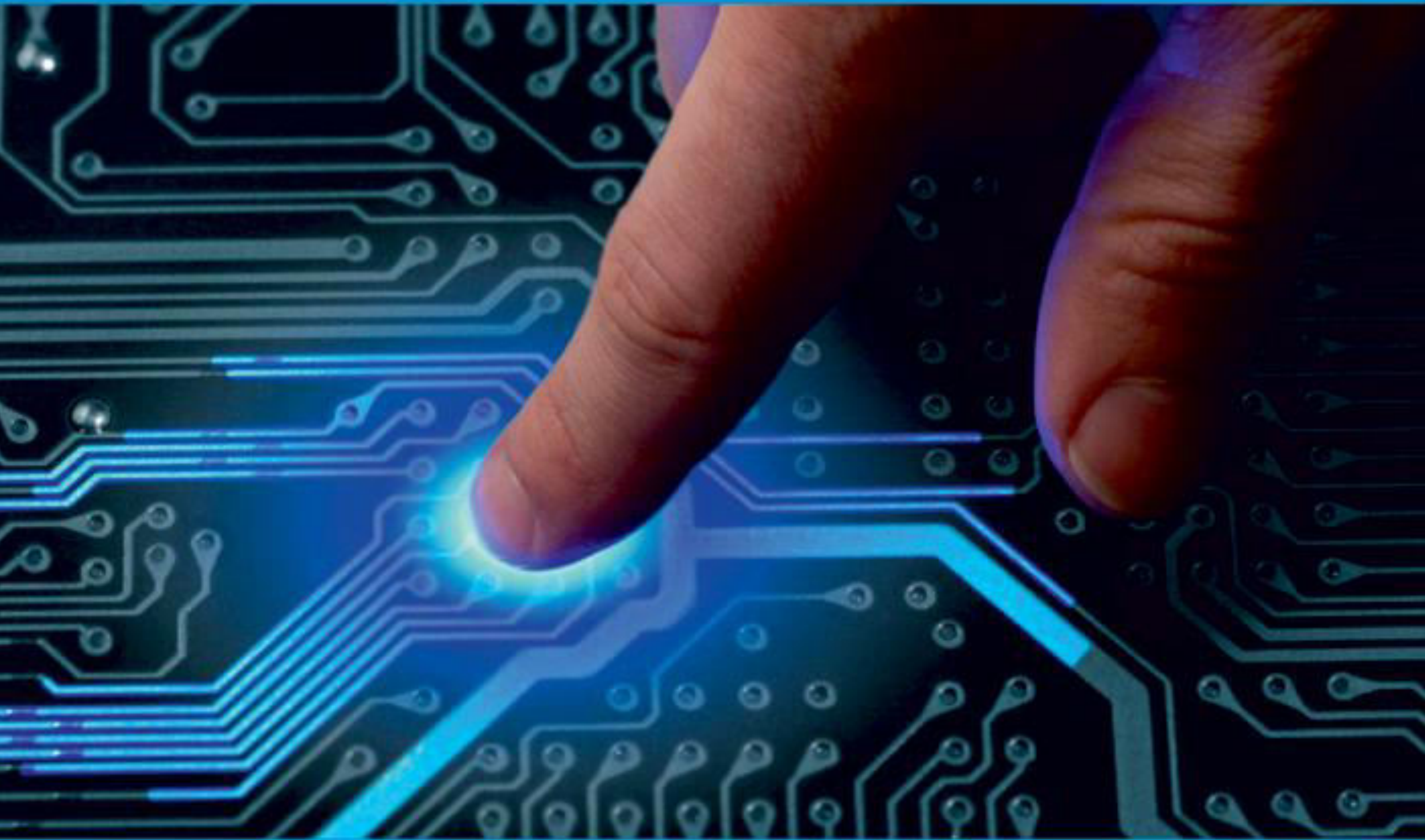
e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 6, June 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

YoloV5 Based Social Distancing and Face Mask Detection for Covid-19

Potluri Sairaj

PG Student, Department of Computer Science and Engineering, Faculty of Engineering and Technology, Jain(Deemed to be University),Bangalore, India

ABSTRACT: COVID-19 virus remains a source of anxiety and danger in the recent timeline. With a huge population in a mobile state and inadequate human power and resources to monitor them, monitoring the social distance norms manually is identified to be impractical. A lightweight, durable, and 24-hour video monitoring system that optimizes this process is needed. This article presents an absolute solution to how the COVID 19 disease spread can be reduced. A YOLOv5-based Facemask and Social Distance detection Framework for Covid-19 is proposed to reduce the spread of the disease. The suggested method processes both images and videos using the OpenCV library. The method employs a pre-trained SSD (Single Shot Detector) to identify faces in the frame, and MobilnetV2 serves as a classifier to identify faces hidden behind masks. To estimate the net space in between the enclosing features, the Euclidean Space calculating algorithm has been utilized. YoloV5 is then used to recognize persons in the frame and create bounding boxes around them. The system performance is evaluated through some particular levels of performance metrics. These are referred to as accuracy along with precision and recall. The accuracy of this model is 99.74%, higher than another previous model.

KEYWORDS: Covid'19, face mask, social distancing, Disease, CNN, YOLOV5, SSD

I. INTRODUCTION

A health crisis in the world is COVID-19 Pandemic Disease that has been able to create a great impact on humanity and also on everyday lives. This is also majorly about the aspect of several respiratory syndromes that are also being initiated by coronavirus 2 (SARS-CoV-2), a global pandemic that emerged in Wuhan, a city in China in December 2019. This has also infected about 7,711 people and 170 reported deaths which were named COVID-19 by the World Health Organization. World Health Organization WHO has recommended wearing face masks and also maintaining a particular social distance. This social distance should be about 2m. This is also to be maintained and to be identified among several individuals for avoiding the risk of virus transmission. Customers should be wearing masks and following the aspects of safe social distancing. They should also demonstrate to use of the services of many public service providers [1]. It's very important to wear a mask and follow social distancing for prevention purposes due to COVID-19. Since the virus is spread through physical or close contact, people should limit their interactions with each other [2]. Covid-19 illness can be reduced by following social distancing and wearing a face mask. The world's governments suggested a variety of solutions. These solutions range from consignment during lockdowns to being capable of preventing the rapid growth of the infection of the pandemic. Nations shut down businesses and public places to avoid people's interactions. The huge number of COVID-19 cases along with a huge range of deaths was reduced in all infected countries by applying the lock-down for their communities.[3]. Washing hands frequently should be followed [4]. To slow down the growth rate of new cases we have to maintain social distance. To live a safer and healthier life, social distancing is a necessity as recent findings show that droplets from a sneeze or a deep breath will fly more than six meters [5,6]. Therefore, the computer vision task that is terms to be important within the global society is the understanding of face mask detection along with the monitoring of the social distancing through safe distances.

Masks should be worn by people who are identified to be at a particular level of greater risk based on the severe illnesses that they suffer from. Mostly Covid 19 is considered to have spread by people who are unaware that they are infected and do not have any symptoms and through contact that is being placed among one another for about six feet [7]. Hence Face mask rules are enforced by many organizations. Classifying people who are particularly using or wearing masks and people who are not in a particular organization is identified as cumbersome and possibly conflicting if it is done manually. We need to save lives until everyone is vaccinated. Hence, a deep learning system is

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

being determined to monitor social distancing and classify people without facial masks [9]. This paper describes the basic identification as well as monitoring of not only the real-time but also a person who is capable of ensuring safe social distancing and also the identification of wearing the face masks. And it also checks whether every person in a public is wearing a face mask while the deep learning method is being used [9]. Efficient and autonomous results can be obtained using deep learning and machine learning models. Computer vision systems generate real-time data [11]. To detect the face with or without a mask deep learning method can be used [8]. The objective is to surveillance social distance and mask detection consecutively by the use of deep learning so that the transmission of this disease is minimized [10]. Tasks that are related to computer image and vision are also demanding through the object direction and classification that is all related to the instance segmentation have shown excellent performances in deep learning [12]. It is economical, easy, and effective to use a computer vision-based automatic system to check the spread of Covid -19 [13]. This system can also be used within the CCTV cameras so as there is a close eye is kept on people. This will detect images and videos and identify who is not following the regulation by drawing a bounding box around them. The bounding box around the person who is violating the regulation will be marked as red. Hence the officials can be alert and take the necessary actions. [14,15].

The main contributions of this paper are described below:

- The proposed method consists of the following steps such as pre-processing, feature selection, bounding box insertion, and checking whether the Distance or mask violations are maintained or not
- The live stream image is recorded and the data is pre-processed using the OpenCV library in python. The pre-processed data is given into the feature selection model.
- The features are extracted from the face using a pre-trained Single Shot detector.
- The features are selected and the selected features are given into the detection model and the cropped images are sent into it and the bounding boxes are latterly added using YOLOV5.
- Then the image is classified with MobileNet V2 for face mask detection and using YOLO V5 the social distance detection is done. The data are checked, whether the condition is yes or no. The results are displayed on is monitor
- The proposed method is evaluated by certain metrics which are accuracy, Precision, Recall, F1-score, Loss, and Specificity. The presented model has an accuracy of 99.74%, greater than any other previous method.

Moreover, the details of this work are discussed in further sections, section 2 deals with the findings, and an analysis of related existing work of the proposed model is detailed below. And the details of the data set in described in section 3. Moreover, in Section 4, the work that is being introduced is experimented with and analyzed. Section 5 deals with the result analysis and the comparative discussion of the work. Section 6 shows the conclusion of the work.

II. RELATED WORK

"Some of the recent research works related to social distancing and face mask detection using deep learning were reviewed in this section"

*Sethi, et.al.[*16] have proposed a strategy for decreasing the risk of Coronavirus spread using deep learning face mask detection. With the speak to the identification of a very precise as well as a system that is real-time so as the non-mask faces are being identified within the public, it is just about the proposed technique that is to be interviewed. This also means that position as well as real-time on the faces of the public along with a particular proposed technique. This is just about the ensemble of the one-stage and about the doctors that are capable of everything. Several detectors are also being introduced especially relating to the low inference and also the high accuracy which are capable of promoting 0.44% recall and 11.07%. higher precision is, therefore, to be achieved.

Dondo, et.al.[17] have also identified a particular strategy for deep learning usage which is face mask detection. This is also considered to be an important implementation along with the performance analysis to help in determining everything in real-time. It Is also with the masks and several image sequences that are all being introduced or long with a detective phase and also with the faces. These do not have any mask so has been presented as a detection of an object detection network that is also generic and holes together with an accuracy rate of 90%.

Loey, et.al. [18] show that there is a presence of medical face masks which helps in the novel deep learning for the model-based aspects, important content for the deep transfer. Several other important components are being introduced rating to the ResNet-50. All these are capable of understanding the transfer of the deep learning model and also capable of generating certain futuristic extractions. Moreover, a second component is being determined to be YOLO V2. It is about the medical reasons for masks and also about the anchoring of boxes so that a particular estimation is being

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

derived. This is also being represented as a part of the object recognition and its process for identification of the detector and also knowing the fact about the optimizer so as the percentage level is found about 81%.

Singh, et.al. [19] have proposed that face mask detection and the use of YOLOV3 and even the models like R-CNN can be used in this make environment. This is also because of the drawing of the bounding boxes for the identification of the fact about them wearing masks or not wearing masks. This is to keep proper track of all the daily ratios of several people who are wearing the masks.

Nagrath, et.al. [20] also suggested the fact regarding the real-time aspect of the DNN-based detection system. This detection system is not only capable of introducing a multi-box detected with a single shot but it is also about the MobileNetV2 that is relatively resourced. The general denomination is also about the utilization of some of the extremely safety considerations. The employment of the single-shot multi-box detector is being done by the method of SSDMNV2 which also relates to the MobileNetV2 architecture. This is for deriving certain important aspects within the framework embedded along with certain utilization of devices and also conducting a detection for the real-time mask.

Fan, et.al. [21] identify facts regarding the deep learning aspect which also relates to several important connotations about the face mask detector and even with the attention that is provided as a residual context. It is also majorly related to the Gaussian heatmap there is capable of implementing the attention module as a part of the residual context. This is also about the suggested model that is also being introduced and terms to be strong potential with the contribution of the public health care and its contribution that relates to the coronavirus illness battles. There are separate other discriminating characteristics that are also being learned from this article. This is by representing the major characteristics and for introducing the faces and differentiating them from the faces that are very masks.

Jun zhang, et.al. [22] have presented a method of COVID-19 Control Using a particular determination that relates to the novel detection framework. This framework is also being determined all along with the considerations that are being made for wearing a particular mass can also for identifying context attention. This also relates to the face mask-wearing conditions with the help of the R-CNN. Several other components are also being introduced by wearing a face mask and even during with the basic frameworks that are being determined for these situations. Hence talking about the infraclass distance to the shortest intro class distance is also about the aspect of extracting certain distinguishing features.

Mohan, et.al. [23] have suggested a procedure for Medical detection for face masks using a Tiny CNN Architecture for Resource-Constrained Endpoints. Impacts also relate to the recognition of the resource constraint type of endpoints and also about the architecture that is being dealt with in the medical face mask. This equally means that it is not just about the memory footprints but it is also about the qualified construct. There are other important components just like the memory footprints and also the quantity sizes of the model that is being introduced as a part of the tensor flow light framework for minimizing the aspects of the size. Certain other models are also being introduced just like the 138 KB for a particular 30 FPS inference rate.

Meivel, et.al. [24] have proposed we aspect that with the detection of the face mask along with the identification of social distance in it. It is not just about the real-time data analysis that is to be made but it is also about solving along with multitasking several problems that are related to picture detection. Several other denominations are also being introduced starting from the methodological aspect that has been used with the context of a data set for the denomination of the phase database and also to help in the detection of the UN mask person.

Mahdi Rezaei and Mohsen Azarmi [25] have presented a CCTV security camera and computer vision that is hybrid along with the deep neural networks or Yolov4 for depending upon the automatic context of people detection. This not only works outside but also inside or indoors as well letting to the inverse perspective that is also being introduced here as a part of the IPM technique and also when dealing with the tasking algorithms which have become some of the important concepts of social distance in monitoring. The inverse perspective that is, therefore, being tackled here as a major part of the algorithm is also much more related to people detection in a reliable way and also the identification of the monitoring systems through social distance.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

III. PROPOSED ALGORITHM

In recent times much of the study has already been done for the detection of face masks along with the tracking of social distancing through violation techniques. There is a smart approach that is being represented here in this paper to introduce the detection of face masks on humans and monitoring certain violations of social distancing. This is also about the two important factors of recognition of its mask and also the identification of social distancing during the period of covid-19 which might reduce its spread. This paper discusses a method to appropriately detect persons, face masks, and social distancing in a live stream image. Figure 3 explains the presented model of the face mask detection system, live stream video is given as data and image, and video is pre-processed using OpenCV. Then these processes images collected from the live video stream are sent to the mobilnetv2 architecture for recognizing if an individual in the live stream is wearing a facemask or not. Here mobilenetV2 is identified as a particular type of classifier for the face masks. The face detection is done through a pre-trained SSD model which is trained to detect people's faces. The dataset used in the project to detect masks and people are based coco dataset. As for person detection, YOLO V5 is used, to produce bounding boxes on the persons. Figure 4 illustrates the YOLOV5 model for the social distancing detection process. For measuring social distance, calculate the center point or centroid of a particular bounding box of the image along with its measurement of the distance it is also about bonding to different boxes with the help of the euclideandistance.If the distance is less than 200 meters, change the color of the bounding box to red and alert to keep distance.

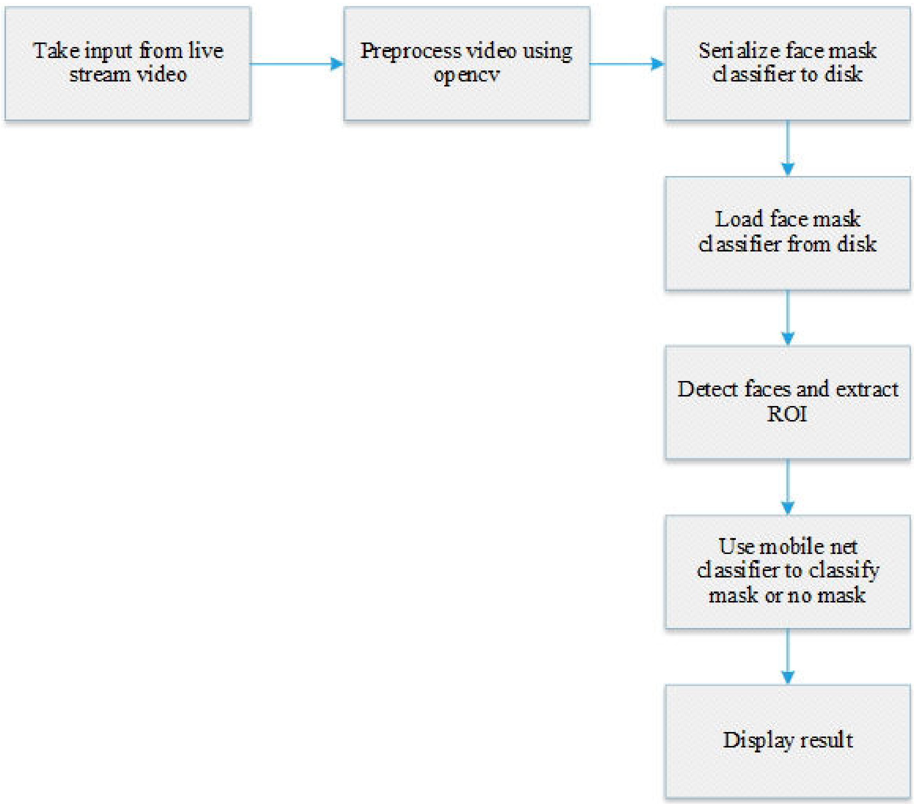


Figure 3: *a proposed model for detection* of face mask

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

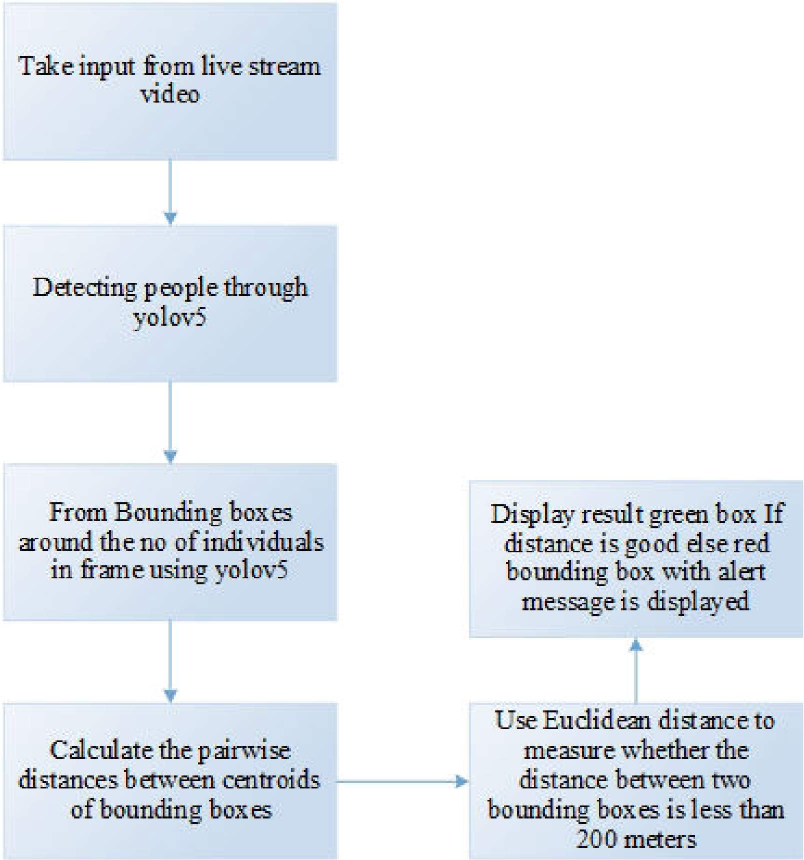


Figure 4: *a proposed model for social distancing detection*

A. Person Detection

The fifth iteration of YOLO is utilized for the person detection architecture. The image is divided into a grid using YOLO, which concurrently produces boxing box confidence and class probability for each grid cell. This method approaches detection as a regression problem. The final bounding box output and classification are then generated by the model after it aggregates these data. This architecture is inherently a strong fit for this work because of its performance and efficiency on real-time video data. The original YOLO architectures have seen numerous revisions and advancements. This takes us to YOLO v5, the fifth iteration.This most recent version uses Path Aggregation Network (PANet) as the neck for feature aggregation and Cross Stage Partial Network (CSPNet) as the model backbone. Better feature extraction and a large increase in the mean averaged precision score are the outcomes of these enhancements. YOLOv5 is the next version of YOLO v4 with has the same head. It performs a region-based object detection network and a stage detector. The Yolo recasts object detection as a regression problem that leads to high-speed processing. The model design of YOLOv5 is given in figure 3. It composes of three vital components.

- ❖ Backbone
- ❖ Head
- ❖ Neck

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

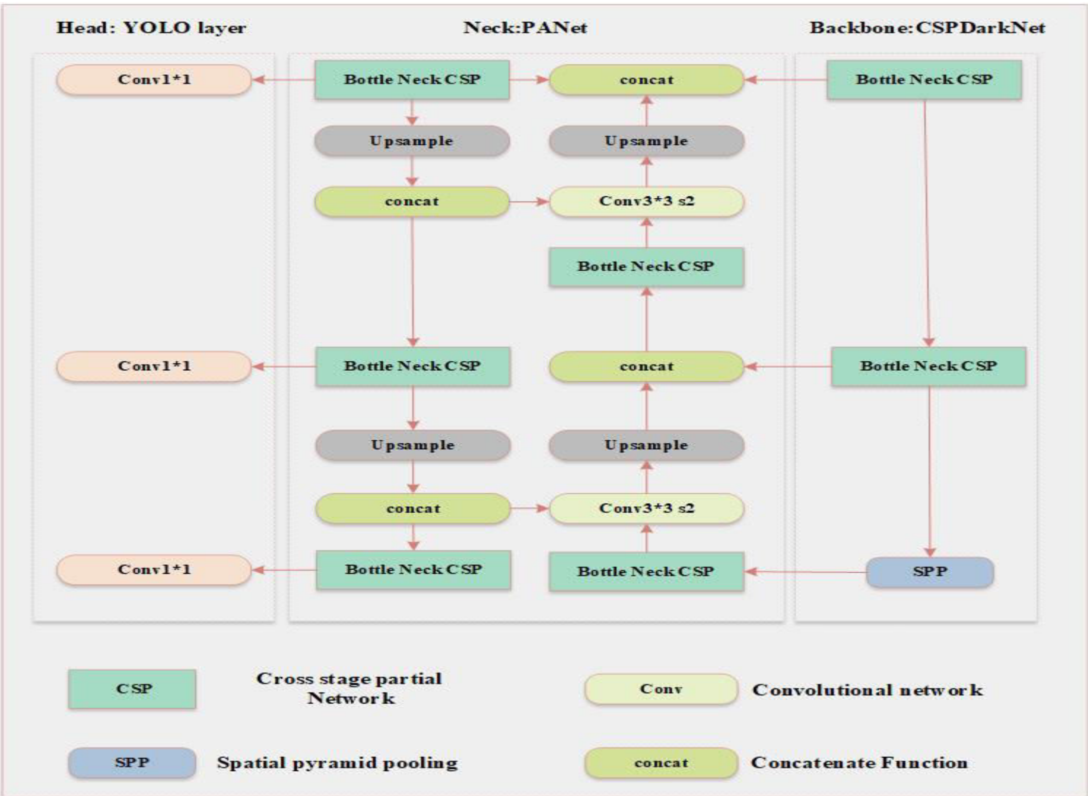


Figure 5: *Model design of YOLO v5*

Backbone: Backbone aids in the extraction of important features from an image input. YOLO v5 uses CSP (Cross Stage Partial Networks) as a backbone for the extraction of important characteristics from an image input.

- ***Neck***: Creation of feature pyramids is done by the Neck. The scaling of objects is well handled in models by the feature pyramids. This particularly helps in certain identifications that are being made about the same objects but presented through a variety of scales and sizes and also work well on previously unknown data. Models like BiFPN, PANet, and FPN use Feature pyramids. YOLO v5 also uses PANet as a neck to attain feature pyramids.
- ***Head:*** The final detection step of the model is the Head in which anchor boxes are used to create final output vectors with bounding boxes, abjectness scores, and class probabilities feature pyramids are generated by the Neck in YOLO v5 and the object scaling is handled by feature pyramids. The methodology utilized CSPDarknet53, EfficientNet-B3, and CSPResNext50, as the YOLOv5 object detectors backbone. The CSPDarknet53 and CSPResNext50 used DenseNet. YOLOv5 used a data augmentation procedure augmenting 4 images altogether to trace small objects accurately. The detection of the persons from the video is performed by the algorithm by placing the four persons in 4 different coordinate locations. Later, the data loader is capable of performing three different methods. These include the color space changes along with various aspects of data argumentation and scaling of the argumentation. YOLOv5 relates to facts about the three different predictions not only in various scales but also in different spatial points. Several other relations are also being introduced with the overall loss that is being defined for the function. This remains much more localized and cross-entropy along with the aspect of confidence loss for the determination of classification score as identified (8).
-

$$\beta_{doose} \sum_{n=0}^{q^2} \sum_{n=0}^{B} 1_{n,m}^{0al} \left(\left(R_y - R_y\right)^2 + (R_x - R_x)^2 + (R_v - R_v)^2 + (R_l - R_l)^2 \right) + \sum_{n=0}^{q^2} \sum_{n=0}^{B} 1_{n,m}^{0al} \left(- \log \log \left(\alpha(R_\theta)\right)\right) + \sum_{L=1}^{D} BDE\left(x_L, \alpha(\mu_L)\right) + \beta_{doose} \sum_{n=0}^{q^2} \sum_{n=0}^{B} 1_{n,m}^{0al}\left(- \log \log \left(\alpha(R_o - 1)\right)\right) \quad (8)$$

where β_{doose}indicates the basic determination of the co-ordinate error while on the other hand, p2 is capable of indicating various grids that are present within the images. A helps in the generation of the bounding boxes as represented in every grade relating to the objects that are being described with the fact of nth range of bonding box present within the grid m. Without this, the equation will result in 0.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

B. Social distance detection

The object detector YOLO v5 can detect people in each frame. In figure 4, the YOLO Architecture is illustrated. It counts the people in the frame and calculates the separation between the basic aspect of the centroid. This is present within the bounding boxes that surround the detected people. Through Euclidean formulation, the distance is computed. The efficient aspect of the euclidean distance calculation relates to the identification of distance that is present between two humans. This is also initiated as a model that is capable of generating a boundary box that is present around *Ec*→person. On the other hand with the aspect of successful detection, this is much more related to the pair represented for EC, and the distance that is calculated in between also relates to the euclidean distance which is D within the plane.

$$D\left(E_c^i E_c^j\right) = \sqrt{(x_j - x_i)^2 (y_i - y_i)^2} \qquad (9)$$

$$\{risk\ ;\ D < T\ \ No\, risk; \quad D > T \qquad (10)$$

where *Xi*, *Yi,* and *Xj*, *Yj*are identified to coordinate which are represented for the centroid coordinates that are represented within the bounding box. This is determined for *Eic*and *Ejc* respectively. As identified the year decides to be the identification for the threshold value along with the minimum social distance that is also being maintained and enhanced. The difference also relates to a particular situation which is all about the calculation of the Euclidean distance that is D and finding out that it is greater than a particular specified value which is T. Initialize the color of the bounding boxes, under the threshold limit it appears to be in green color and beyond the limit display the box in red color.

C. Face Detection

Face detection, which involves locating and validating any human faces present in the video frame or the real-time data inputs, is the first stage of the face recognition pipeline. To extract features from the selected faces, such as the features of the nose, eyes, and lips, the algorithm must first analyze the images and identify all human faces that are there but are subject to various constraints, such as positions, lighting, and occluded faces. The features are finally matched with the person's known name. The human brain instantaneously performs this with any time-lapse. When a face is discovered, bounding rectangles or boxes are drawn around it using the BGR image as an input. For the next stage of Face recognition, the bounding boxes with the highest certainty are chosen. This article has implemented the Single Shot Multibox Detector (SSD) framework, which was designed based on the model named ResNet-10 neural network model, to recognize faces in a given video, image, or real-time input. A well-known library called Open-Source Computer Vision (openCV) and its aspect of deep learning module was used to work on real-time computer vision techniques linked to deep learning. Tensorflow, Caffe, PyTorch, and other deep learning frameworks are handled by this OpenCV module. With the aid of the pre-trained neural network models that were provided alongside OpenCV, face detection can be achieved with remarkable accuracy.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

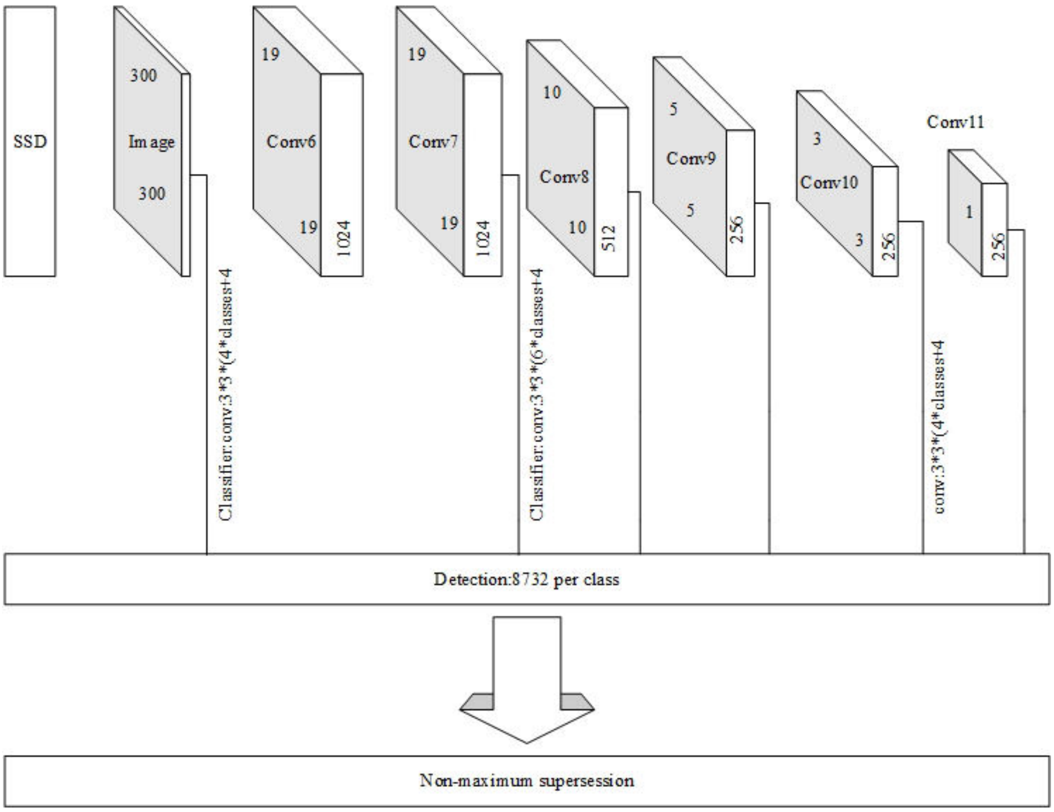


Figure 6: Structure of SSD model

D. Face mask classification using Mobile Net V2

The Mobile Net V2 model is considered to be a convolutional neural network method that is being built up by the Google corporation consisting of some enhanced efficiencies. The pre-processing phase is comprised of filtering of images, input pre-processing with MobileNetV2, along with hot label encoding. The determination of the image is found to be an integral part of the critical pre-processing step. However at the same time with the effectiveness determine for the training models this becomes also an important part of the computer vision. The model runs better in small images hence the Image is resized in 224x224 pixel size. The processing of images into an array from the data set is the next step. To call the image in a loop function, conversion of the image into a certain array is done. After that, there will be a use of the image for the preprocessing input but utilization of the mobile net V2. In the final step hot encoding on labels is done because data labels cannot be used by machine language. Every variable of the input and output along with the labeled data will also be particularly transformed to a certain numerical label to introduce the understanding of the algorithm along with its data processes. MobileNetV2 is found to be a bottleneck depth that has separable convolution construction of fundamental blocks with residuals architecture. MobileNetV2 is a bottleneck depth separable convolution construction of fundamental blocks with residuals architecture. 1x1 convolutions with "ReLU6" is the first. The determination of the second layer includes depth-wise "convolution," whereas the third layer includes a 1x1 "convolution" with no non-linearity. A one-stride residual block is an initial stride. The second layer is a residual block with stride 2 that is utilized in shrinking. Object detection is accomplished by classification to determine the input class and regression to update the bounding box. Most backbone networks for detection are networks for classification tasks, except for the last entirely connected layers. The images are taken as output and feature images are produced by the backbone network which performs the feature extraction of the tasks related to Object detection. Feature maps with high-quality classification issues are frequently extracted using established training approaches. This is the base model which is MobileNetV2 n ImageNet. The comparison of the “ground truth boxes” with the evaluated” bounding boxes” for the changing of the trainable parameters, when necessary, in the back propagation process. Every feature space uses a kernel for the outcome production which shows the pixel score to find the existence of an item and the dimensions of the bounding box. The Mobile Net is designed with two parts: a base model and a classifier. The reuse of the base model and trimming of the head along with the two fully connected layers is done.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

IV. RESULTS AND DISCUSSION

This section shows the result and discussion of the proposed methodology in two parts, classification of face masks and monitoring of social distancing. The performance which is determined for this introduced algorithm is being compared and analyzed with the help of certain existing algorithms. There are other performance metrics like Accuracy, Precision with Recall, F1-score, Loss, and Specificity values are compared using different algorithms. The existing algorithm CNN, MobileNetV3, Inception V3, and Resnet-50 for the aspect of face mask detection along with Stacked ResNet50 are analyzed and compared with the proposed algorithm in face mask detection system and Faster R CNN and even Yolo V3, Yolo V4 tiny and Yolo V2 are the existing algorithms compared with the proposed model in social distance detection model. The performance metrics are analyzed below.

A. Performance Metrics

The proposed model has certain performance measures which are all being explained here. Some of them have been validated while others have been tested and there is an analysis that is being derived from everything. This means the performance measures that are also being included here determine certain existing analysis and also helps in the planning of a mathematically expressed context.

❖ **Accuracy:**

Accuracy is determined to be a matric that is capable of describing the performance of a model throughout the classes. This is vividly useful with the identification of classes along with equal importance and it is also about its calculation through a particular ratio as expressed between correct predictions and the number of predictions.

$$Accuracy = Tp + TN * TP + TN + FP + FN \quad (9)$$

❖ **Precision**

Calculation of the precision is being done by explaining a ratio and finding out the difference that exists between the total positive samples that are classified correctly and also the total samples that are classified as positive. The accuracy of the model is therefore being measured with the help of this classification.

$$precision = \frac{TP}{TP+FP} \quad (10)$$

❖ **Recall**

This is being calculated by the identification of the positive samples with a ratio to the positive samples that are correctly classified. This means that it is also about the ability of the model to detect certain positive samples. Moreover, with higher records, the detection of the positive samples is also higher in number. $Recall = \frac{Tp}{TP+FN}$ (11)

❖ **F1-Score**

F1-score is considered to be the harmonic mean as identified for the recall and precision.

$$F1 - score = \frac{2*(precision*recall)}{(precision+recall)} \quad (12)$$

❖ **Sensitivity**

It is called a true positive rate as the various proportion of the positive and the measures that are being initiated are correctly identified.

$$Sensitivity = \frac{TP}{TP+FP} \quad (13)$$

❖ **Specificity**

This is identified as the true negative rate as the proportion of negative as the measures are correctly being identified.

$$Specificity = \frac{Fp}{Fp+TN} \quad (14)$$

❖ **Loss**

This is being determined as a penalty that is derived for times of bad prediction. It means that this is not just about a number that is capable of indicating the bad prediction of the model but it is about finding out if there is a perfect prediction being made. With a perfect prediction, the loss would be 0 and otherwise, the loss is found to be greater.

❖ **Intersection of Union**

The Intersection of Union (IoU), is identified as an evaluation metric that is commonly used for the estimation of the regression quality. This is also very much related to the already predicted bounding box along with the assignment that

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

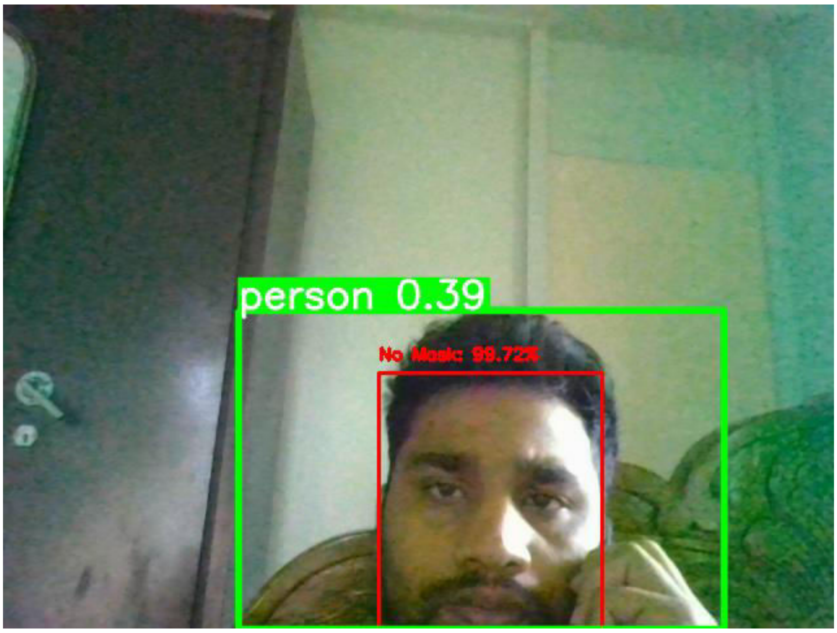
corresponds it with the ground truth box. This is also being determined for the overlapping of the two boundaries that are being measured detailed through the boundaries of the real object. This boundary however is considered to be the ground truth that is to be considered as the predicted object boundary. $IOU = \frac{Area\ of\ overlap}{Area\ of\ union}$ (15)

❖ **Mean Average Precision:**

Mean Average Precision (mAP) is identified for its performance determination that is detailed with models for object detection. This means that starting from classification to localization of a particular image everything is being determined specifically.

$$mAp = \frac{1}{N}\sum_{i=1}^{N} AP_1 \quad (16)$$

(a)



(b)

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

(c)

Figure 6: *Face mask and social distancing detection results in three case*

(a) Social distancing high risk in live stream video

(b) Result in a single person in the frame without the mask

(c) Results with two people in the frame

Figure 6 illustrates the three cases of the result after the face mask and social distance detection system were applied. In case (a), the live stream video is detected and the result shows red box on the person's images and demonstrates the moving persons are getting closer beyond the threshold limit triggers the red box. In case (b), shows that a live image is detected and the person in that image is not wearing the mask with social distance within the threshold limit. Case (c) demonstrates the image detected by the proposed model and the result shows that out of two people in the image both are violating social distancing and a person has no mask.

B. Performance analysis

This section describes the detailed performance analysis along with the comparison of the existing and proposed methods.

C. Performance metrics analysis in Face mask detection model

The performance metrics include accuracy along with precision and also recall, F1-score, the aspect of Loss, and Specificity. All these are analyzed and compared for the existing algorithms CNN, Mobile Net V3, Inception V3, Resnet – 50, and Stacked ResNet – 50. The confusion matrix is often used for the determination of accuracy along with the recall of the existing algorithms. Table 1 presents the performance and comparison of current and proposed methods in the face mask detection model.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

Table 1:*Performance comparison of the previous and proposed model*

Algorithm	Performance metrics					
	Accuracy	Precision	Recall	F1- Score	Loss	Specificity
CNN	87.45	4.7	66.12	61.12	1.5	53.23
Mobile Net V3	89.56	5.9	93.12	72.23	1.3	82.23
Inception V3	94.45	6.5	94.45	77.11	1.0	92.45
Resnet- 50	95.34	7.4	91.23	76.45	1.8	78.33
Stacked ResNet - 50	96.12	8.3	92.34	84.23	1.26	83.11
Proposed method	99.72	9.1	96.23	90.23	0.8	95.32

a) Accuracy, precision, and Recall

The Accuracy, precision, and Recall have a greater comparison that is present between the proposed as well as the already existing methods as shown in figure 6. As identified the accuracy along with precision as well as a recall of certain methods that are current as well as that are being suggested just like the convolution neural network the mobile network V3 and along with its inception V3 and resnet 50 are all used for the face mask detection. Stacked ResNet50 is shown in Figure 7. Accuracy and also precision along with Recall are often dealt with as a suggested approach which is dealt to be superior in several current strategies being introduced.

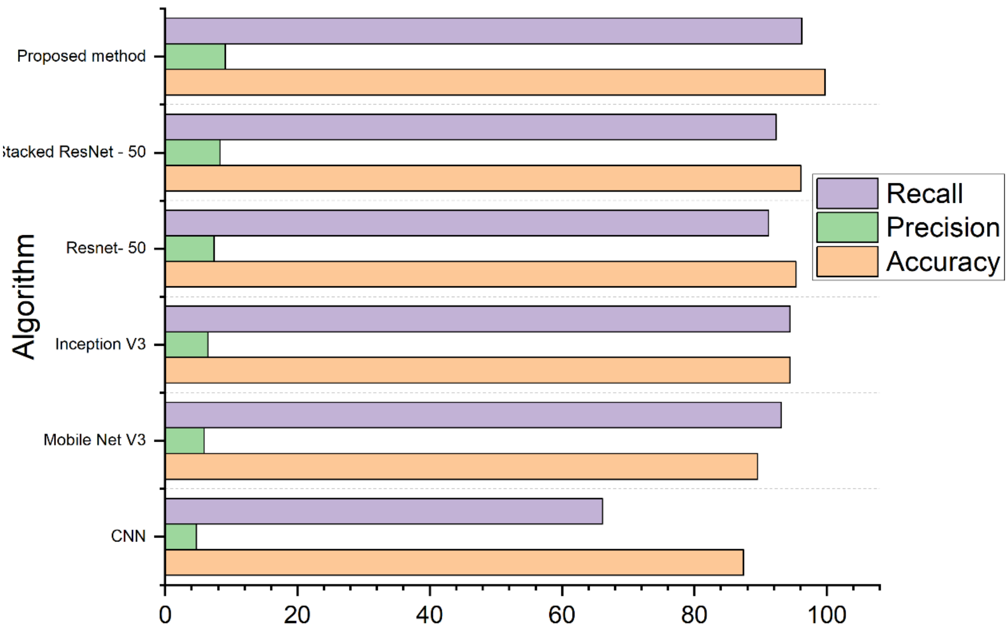


Figure 7:*comparison of the proposed and existing methods of Accuracy, precision, and Recall.*

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

From the figure, the sensitivity determined for the particular proposed method is high when it is being compared to the already existing method, precision is 9.1 and Recall is 96.23%, and performance is better than the existing method. Accuracy is 99.72%, higher than other existing methods.

b) F1- Score, Loss, and Specificity

The F1- Score, Loss, and Specificity comparison that is being made between the proposed as well as the existing methods are shown here in this figure . The F1- Score, Loss, and Specificity are identified for suggested as well as current methods just like CNN, MobileNetV3, Inception V3, Resnet-50 used for the face mask detection. The Stacked ResNet50 is also represented in Figure 8. The F1- Score, Loss, and Specificity determined for this particular approach that is being suggested is considered to be superior to all the various other current strategies.

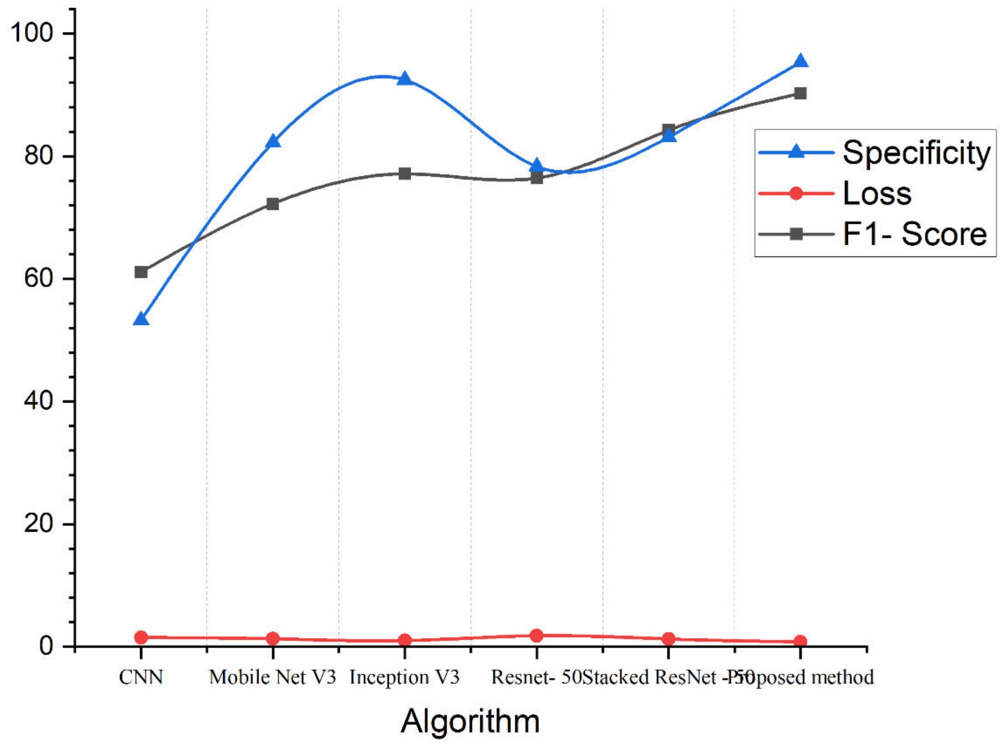


Figure 8:*comparison of the proposed and existing methods of F1- Score, Loss, and Specificity*

From the figure, the sensitivity of the proposed method is high when compared to the existing method, F1-score is 90.23%and specificity is 95.32, and performance is better than the existing method. Accuracy is 0.8, lower than other existing methods.

D. Performance metrics analysis in social distancing detection model

The performance metrics as referred to Accuracy, Precision, F1-score, Loss, Recall, and Specificity are analyzed and compared for the existing algorithms Faster R CNN, Yolo V3, Yolo V4 tiny, and Yolo V2. The confusion matrix is therefore being used to determine the accuracy as well as recall of the existing algorithms. Table 2 presents the performance and comparison of current and proposed methods in the social distancing detection model.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

Table 2: *Performance and comparison of the previous and proposed methods*

Algorithm	Performance metrics					
	Accuracy	Precision	Recall	mAP	loss	FPS
Faster R CNN	88.52	6.9	88.12	69.34	2.3	2.44
Yolo V3	92.45	5.5	90.45	55.45	1.3	34.11
Yolo V4 tiny	94.34	6.4	87.23	64.44	2.8	22.34
Yolo V2	95.12	5.3	91.34	53.23	1.26	45.45
Proposed method	98.72	9.5	95.23	95.34	0.7	55.67

a) Accuracy, precision, and Recall

The Accuracy, precision, and Recall comparison of certain methods that are already existing and the method that is being proposed is represented in figure 8. The Accuracy, precision, and recall of both the suggested as well as the current methods just like Faster R CNN, Yolo V3, Yolo V4 tiny, and Yolo V2 are shown in Figure 8. The Accuracy, Precision, and Recall of the approach that is being suggested are also considered to be superior to all the various other current strategies.

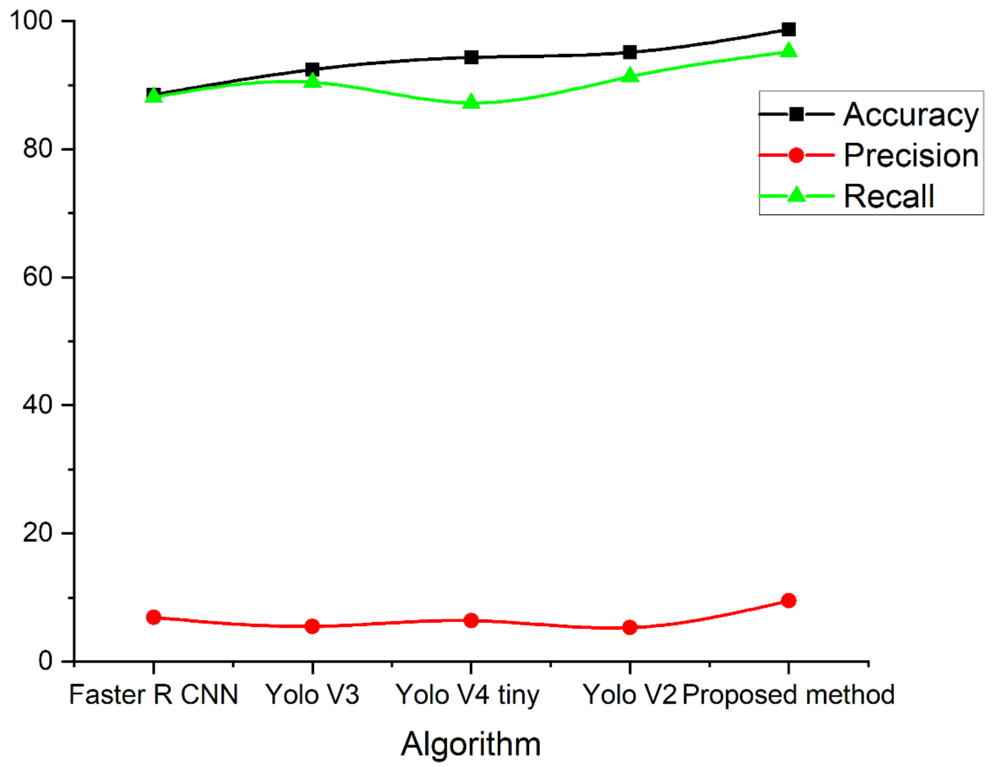


Figure 8:*comparison of the proposed and existing methods of Accuracy, precision, and Recall.*

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |
|| Volume 10, Issue 6, June 2022 ||
| DOI: 10.15680/IJIRCCE.2022.1006178|

From the figure, the sensitivity determined by the proposed method is high when compared to the already existing method, precision is 9.5 and Recall is 95.23%, and performance is better than the existing method. Accuracy is 98.72%, higher than other existing methods.

b) mAP, Loss, and FPS

The mAP, Loss, and FPS comparison made between the already existing methods and the method that is being proposed is represented in figure 9. The mAP, Loss, and FPS of all the methods that are suggested along with the current methods just like CNN, MobileNetV3, Inception V3, Resnet-50 for the face mask detection, and Stacked ResNet50 are represented in Figure 9. The F1- Score, Loss, and Specificity for the approach that is suggested is considered to be superior to what the other current strategies show.

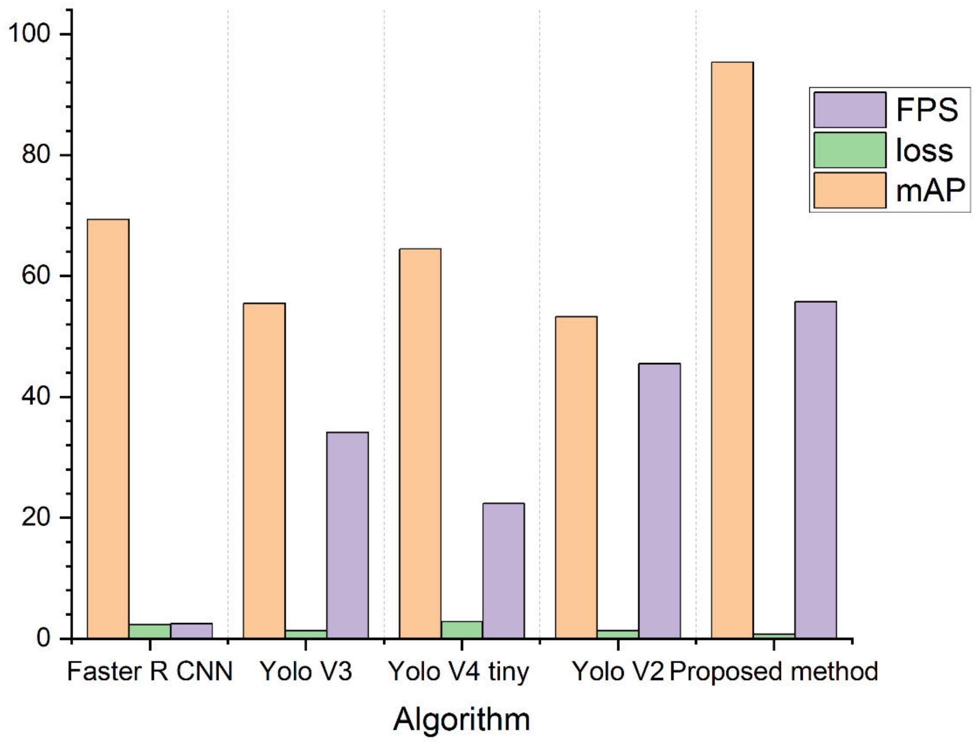


Figure 9:*comparison of the proposed and existing methods of Accuracy, precision, and Recall.*

From the figure, the sensitivity of the method that is proposed is high when compared to the already existing method, mAP is 95.34 and FPS is 55.67%, so performance is better than the existing method. loss is 0.7, low loss with high performance than other existing methods.

Table: 3 comparisons of performance matric with the existing papers.

Author	Performance matric
	Accuracy (%)
Sethi, et.al.[16]	**98.2**
Dondo, et.al.[17]	**90.0**
Loey, et.al. [18]	**70.01**
Nagrath, et.al. [20]	**97.74**
Proposed method	**99.72**

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

The above table 3 shows the comparison of the proposed approach with existing papers with the performance matric parameter like accuracy, analyzed by comparing with existing approaches.

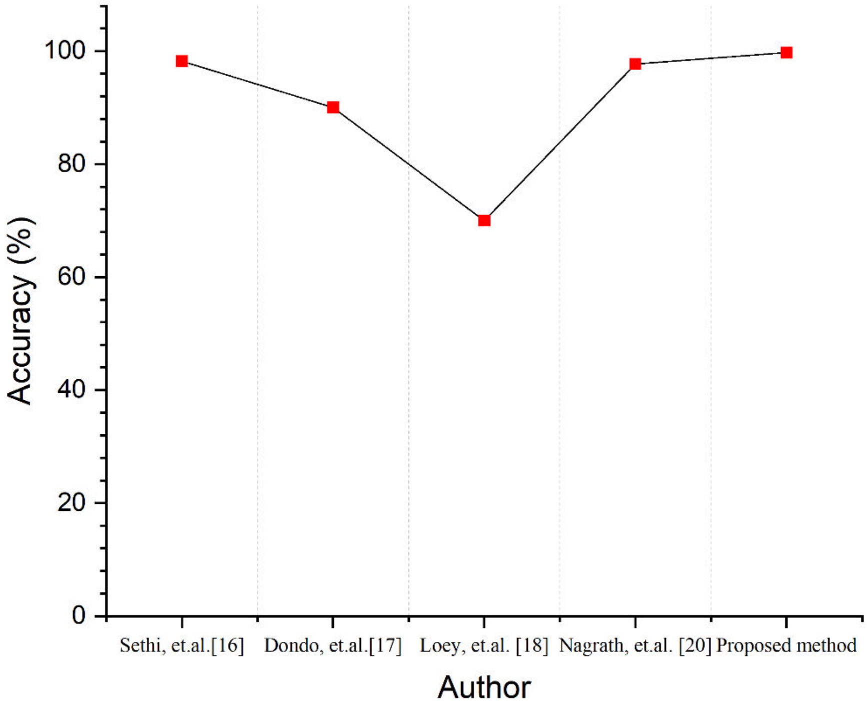


Figure 10:*comparison of existing papers with the current method*

From figure 10, the comparison of existing with the proposed method is plotted graph based on performance matric, and accuracy. The proposed is way higher than that of the mentioned existing methods.

IV. CONCLUSION AND DUTURE SCOPE

This paper provides a solution to detect the face mask and also social distancing automatically and also from the live stream video. The video is recorded and the data is given as input for pre-processing and features are selected. The selected features are classified and checked and the output is displayed. The presented model is evaluated through performance metrics such asAccuracy, Precision, Recall, F1-score, Loss, and Specificity. The accuracy given by the presented model is better than the previous model. The aggregation method assists in achieving high accuracy, and also greatly enhances detection speed. Various potent models have helped in determining a particular transfer learning usage which is not only a part of the extensive experimentation but it is also about the unbiased data set that is being resulted. It is also majorly dependent on minimal cost as well as high reliability. With the future direction being identified the techniques that is suggested or never limited to certain identification for video surveillance device or musk detection but it is much more than that. Hence starting from facial landmarks to the face mask or for any other biometric reasons this model can also be used.

REFERENCES

1.Yadav, Shashi. "Deep learning based safe social distancing and face mask detection in public areas for covid-19 safety guidelines adherence." International Journal for Research in Applied Science and Engineering Technology 8, no. 7 (2020): 1368-1375.
2.Pagare, Reena, PradnyaKedari, Pradum Kumar Dubey, and Sahil Khanolkar. "Face Mask Detection and Social Distancing Monitoring." Int. J. Res. Appl. Sci. Eng. Technol 9, no. 1 (2021): 374-379.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

3.Jayashri, S. "Video analytics on social distancing and detecting mask." Turkish Journal of Computer and Mathematics Education (TURCOMAT) 12, no. 9 (2021): 2916-2921.
4.Saponara, Sergio, AbdussalamElhanashi, and Alessio Gagliardi. "Implementing a real-time, AI-based, people detection and social distancing measuring system for Covid-19." Journal of Real-Time Image Processing 18, no. 6 (2021): 1937-1947.
5.Meivel, S., Nidhi Sindhwani, Rohit Anand, Digvijay Pandey, Abeer Ali Alnuaim, Alaa S. Altheneyan, Mohamed Yaseen Jabarulla, and MesfinEsayasLelisho. "Mask Detection and 6.Social Distance Identification Using Internet of Things and Faster R-CNN Algorithm." Computational Intelligence and Neuroscience 2022 (2022).
7.Keniya, Rinkal, and NinadMehendale. "Real-time social distancing detector using socialdistancingnet-19 deep learning network." Available at SSRN 3669311 (2020).
8.Nagrath, Preeti, Rachna Jain, Agam Madan, Rohan Arora, Piyush Kataria, and Jude Hemanth. "SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2." Sustainable cities and society 66 (2021): 102692.
9.Teboulbi, Safa, SeifeddineMessaoud, Mohamed Ali Hajjaji, and AbdellatifMtibaa. "Real-Time Implementation of AI-Based Face Mask Detection and Social Distancing Measuring System for COVID-19 Prevention." Scientific Programming 2021 (2021).
10.Krishna, K. P., and S. Harshita. "Social Distancing and Face Mask Detection Using Deep Learning." (2021).
11.Walia, Inderpreet Singh, Deepika Kumar, Kaushal Sharma, Jude D. Hemanth, and Daniela Elena Popescu. "An Integrated Approach for Monitoring Social Distancing and Face Mask Detection Using Stacked ResNet-50 and YOLOv5." Electronics 10, no. 23 (2021): 2996.
12.Ottakath, Najmath, Omar Elharrouss, Noor Almaadeed, Somaya Al-Maadeed, Amr Mohamed, Tamer Khattab, and Khalid Abualsaud. "ViDMASK dataset for face mask detection with social distance measurement." Displays (2022): 102235.
13.Saponara, Sergio, AbdussalamElhanashi, and Qinghe Zheng. "Developing a real-time social distancing detection system based on YOLOv4-tiny and bird-eye view for COVID-19." Journal of Real-Time Image Processing 19, no. 3 (2022): 551-563.
14.Prasad, Janvi, Arushi Jain, David Velho, and Sendhil Kumar. "COVID Vision: An integrated face mask detector and social distancing tracker." International Journal of Cognitive Computing in Engineering (2022).
15.Zope, Vidya, Nikhil Joshi, Srivatsan Iyengar, and Krish Mahadevan. "COVID-19 care: checking whether people are following social distancing and wearing face masks or not using deep learning." (2020).
16.Hebbale, Swaroop, and V. Vani. "Real time COVID-19 facemask detection using deep learning." learning 6, no. S4 (2022): 1446-1462.
17. Sethi, S., Kathuria, M. and Kaushik, T., 2021. Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread. Journal of biomedical informatics, 120, p.103848.
18.Dondo, D.G., Redolfi, J.A., Araguás, R.G. and Garcia, D., 2021. Application of deep-learning methods to real time face mask detection. IEEE Latin America Transactions, 19(6), pp.994-1001.
19.Loey, M., Manogaran, G., Taha, M.H.N. and Khalifa, N.E.M., 2021. Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. Sustainable cities and society, 65, p.102600.
20.Singh, S., Ahuja, U., Kumar, M., Kumar, K. and Sachdeva, M., 2021. Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment. Multimedia Tools and Applications, 80(13), pp.19753-19768.
21.Nagrath, P., Jain, R., Madan, A., Arora, R., Kataria, P. and Hemanth, J., 2021. SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. Sustainable cities and society, 66, p.102692.
22.Fan, X., Jiang, M. and Yan, H., 2021. A deep learning-based light-weight face mask detector with residual context attention and Gaussian heatmap to fight against COVID-19. Ieee Access, 9, pp.96964-96974.
23.Zhang, J., Han, F., Chun, Y. and Chen, W., 2021. A novel detection framework about conditions of wearing face mask for helping control the spread of covid-19. *Ieee Access*, *9*, pp.42975-42984.
24.Mohan, P., Paul, A.J. and Chirania, A., 2021. A tiny CNN architecture for medical face mask detection for resource-constrained endpoints. In Innovations in Electrical and Electronic Engineering (pp. 657-670). Springer, Singapore.
25.Meivel, S., Devi, K.I., Maheswari, S.U. and Menaka, J.V., 2021. Real time data analysis of face mask detection and social distance measurement using Matlab. Materials Today: Proceedings.
26.Rezaei, M. and Azarmi, M., 2020. Deepsocial: Social distancing monitoring and infection risk assessment in covid-19 pandemic. Applied Sciences, 10(21), p.7514.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006178|

27.Sharma, P., Singh, A., Raheja, S., & Singh, K. K. (2019). Automatic vehicle detection using spatial time frame and object based classification. *Journal of Intelligent & Fuzzy Systems*, *37*(6), 8147-8157.

28..Barshandeh, Saeid, Mohammad Masdari, Gaurav Dhiman, Vahid Hosseini, and Krishna K. Singh. "A range-free localization algorithm for IoT networks." *International Journal of Intelligent Systems* (2021).

29.Singh, Narendra, Pushpa Singh, Krishna Kant Singh, and Akansha Singh. "Machine learning based classification and segmentation techniques for CRM: a customer analytics." *International Journal of Business Forecasting and Marketing Intelligence* 6, no. 2 (2020): 99-117.

30.Jangra, Manisha, Sanjeev Kumar Dhull, and Krishna Kant Singh. "Recent trends in arrhythmia beat detection: A review." In *Communication and Computing System. Proceedings of the International Conference on Communication and Computing Systems, ICCCS*, vol. 2016, pp. 177-184. 2017.

31.Sahoo, Arun Kumar, Tapas Kumar Panigrahi, Gaurav Dhiman, Krishna Kant Singh, and Akansha Singh. "Enhanced emperor penguin optimization algorithm for dynamic economic dispatch with renewable energy sources and microgrid." *Journal of Intelligent & Fuzzy Systems* 40, no. 5 (2021): 9041-9058.

32.Malik, Anju, Mayank Aggarwal, Bharti Sharma, Akansha Singh, and Krishna Kant Singh. "Optimal Elliptic Curve Cryptography-Based Effective Approach for Secure Data Storage in Clouds." *International Journal of Knowledge and Systems Science (IJKSS)* 11, no. 4 (2020): 65-81.

33.Izonin, Ivan, Roman Tkachenko, Nataliya Shakhovska, Bohdan Ilchyshyn, and Krishna Kant Singh. "A Two-Step Data Normalization Approach for Improving Classification Accuracy in the Medical Diagnosis Domain." *Mathematics* 10, no. 11 (2022): 1942.

34..Shandilya, Smita, Ivan Izonin, Shishir Kumar Shandilya, and Krishna Kant Singh. "Mathematical modelling of bio-inspired frog leap optimization algorithm for transmission

S

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.165

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details