INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Face Mask Detection and Social Distance Monitoring with Deep Learning

A Kala[1]*, P Sharon Femi[1], V Rajalakshmi[1], S Kalavathi[2], K Ashwini[3]

1 Associate Professor, Sri Venkateswara College of Engineering, Sriperumbudur
2 Assistant Professor, Sri Venkateswara College of Engineering, Sriperumbudur
3 Assistant Professor, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Chennai

OPEN ACCESS

Received: 10-01-2023
Accepted: 07-06-2023
Published: 27-06-2023

Citation: Kala A, Femi PS, Rajalakshmi V, Kalavathi S, Ashwini K (2023) Face Mask Detection and Social Distance Monitoring with Deep Learning. Indian Journal of Science and Technology 16(25): 1888-1897. https://doi.org/10.17485/IJST/v16i25.70

* **Corresponding author**.

akala@svce.ac.in

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: This work proposes a real-time classification model that can accurately detect whether an individual is wearing a face mask and maintaining social distance with the goal of developing a lightweight and easily deployable system for surveillance purposes. **Methods:** The proposed method easily identifies the human by bounding boxes and wearing of face mask by real-time Face Detection Recognition System. This is a robust model that involves detection, tracking and validation as its features. Pre trained deep learning models like Inceptionv3, DenseNet are used and compared with the proposed Improved MobileNetv2. A tested deep learning model is developed to check social distancing, which uses the YOLO object detector and computes the Euclidean distance between people to confirm the safety of the system. Finally, the proposed method is evaluated in terms of precision, recall, F1-score, support, accuracy, sensitivity and specificity. **Findings:** The analysis of the results reveals that the improved MobileNetv2 model achieves precision of 98%, recall of 98% and accuracy of 98%. Hence, this deep learning system contributes to the management of COVID-19 outbreak in an efficient way and can be installed for operation in public places like shopping malls, stadiums etc. **Novelty:** The proposed system can be effortlessly integrated into embedded devices that have limited computational capabilities to the detection of face masks in photographs and real-time videos.

Keywords: Human Detection; Improved Mobilenetv2; Face Mask Detection; YOLO

1 Introduction

Over the recent past, deep learning methods are employed for object detection, weather prediction and other monitoring system working around the globe. Literature review shows that the usage of deep learning techniques in face mask detection has a remarkable improvement in performance. This provides an opportunity to apply deep learning model for face mask detection with improved accuracy. The main objective of this research work is developing a model that achieves a balance between recognition accuracy and resource limitations by combining Improved MobileNetV2 and YOLOV3 (You Only Look Once) with transfer learning techniques. This model is designed to be used in real-time video surveillance for monitoring public places, detecting whether individuals are wearing face masks and to ensure the maintenance of social distancing.

Ghadekar et al [1] suggested a Keras model that outperforms the standard architectures of VGG-16 and MobileNet-V2. The model incorporates the idea of image localization using the Multi-Task Cascaded Convolutional Neural Network (MTCNN) architecture. The proposed model is evaluated based on the metrics accuracy, support values, precision, recall, and F1-score. The real-time surveillance system suggested by Walia et al [2] have aimed to detect whether people appearing in a video feed are wearing masks and whether they are adhering to social distancing guidelines. The methodology utilized YOLOv5 to identify humans in the CCTV feed and subsequently utilized Stacked ResNet-50 to classify whether they wear masks or not. Additionally, DBSCAN was used to identify proximities among the detected individuals. Kaviya et al [3] put forth a deep learning-based predictive model and live risk analysis application designed to identify high-risk areas based on social distancing compliance and face mask usage among individuals. Joolfoo and Hosany [4] examined the different machine learning approaches that can be utilized to combat COVID-19. The models suggested by them are coupled with computer vision algorithms and able to deploy in various locations to monitor and enforce government-recommended safety protocols. Javad et al [5] aimed to accelerate the development of automated systems for face mask detection and social distance measurement in public areas. In addition, their prediction provided an end-to-end pipeline for performing real-time face mask detection and social distance measurement in an outdoor environment. Ahamed et al [6] focused on human detection using MobileNet Single Shot Multibox Detector (SSD) model. The distance between humans is computed based on pixel value. The distance among central point and overlapping boundary of humans in divided region is calculated. Alerts are raised when social distancing is violated. The developed system triggers warning in the presence of people in restricted areas.

Kaur et al developed a simple technique for face masks detection using ML tools [7]. Initially, their method aims to identify the faces from the video or image and then checks for the presence of mask in it. It recognized the face with mask in any angle. This method used specialized CNN to determine the presence of mask. Teboulbi et al. focused on the implementation of detecting face mask and social distancing as embedded vision system using the models MobileNet, ResNet and VGG16 [8]. The authors have planned to incorporate sensors for collecting the multimedia data and connect to Edge Cloud to capture the detected objects from varying angles. Meivel et al designed a drone made of deep learning for monitoring social distancing and identification of face masks [9]. The persons without masks are detected from video frames with faster R- CNN model and their particulars are sent to the corresponding officials for necessary action. Vu et al. proposed a method that combines a model with local binary pattern (LBP) features for the recognition the wearing of face masks using RetinaFace [10]. RetinaFace is a face detector as well as an encoder that uses supervised learning for detecting different types of faces. LBP features are extracted from the eye, eyebrows and forehead of the masked face. These features are then combined with the features derived from RetinaFace to form a novel framework for reading the masked faces. Goyal et al [11] designed a model for detecting face masks in both static and real-time videos, categorizing the images as either "with mask" or "without mask". The model was trained and assessed using the Kaggle dataset and achieved an impressive accuracy rate of 98%. A deep learning model was introduced by Sethi et al [12] to identify masks on faces in public areas. Their model effectively dealt with different types of obstructions in crowded environments by employing a combination of single and two stage detectors. Through the ensemble technique, they attained remarkable accuracy and significantly enhanced the speed of detection.

The summary of various models with its benefits, findings, performance accuracy are presented in Table 1.

Table 1. Comparative analysis of existing works

Model	Findings	Benefits	Accuracy
Ghadekar et al [1]	Face mask detection - Multi-Task Cascaded CNN for facial localization	Mass screening at public places	Confidence score- 0.9914, F1-score-0.98, Precision- 0.99
Walia et al [2]	Human detection -YOLOv5 Face mask- Stacked ResNet-50 social distancing - DBSCAN	Identify humans from the CCTV feed	Accuracy-87%
Kaviya et al [3]	Human detection-YOLOv3 Face mask- ResNet50v2 Social distancing-Top View Transformation Model	cost effective and radical web application	Accuracy-97.66% precision-97.84%, F1-Score- 97.92%.
Javad et al [5]	Customized YOLOv3 is employed	Performs face mask and social distancing in outdoors	Improvement of 5.3% in accuracy
Ahamed et al [6]	Human Detection- MobileNet Single Shot Multibox Detector	detects the presence of people in restricted areas	Outdoor accuracy- 56.5% to 68% Indoor accuracy-100%

Continued on next page

Table 1 continued

Kaur et al [7]	Face mask – CNN	Helps to separate the people from the crowd who are not wearing mask.	Improved accuracy
Teboulbi et al [8]	Pretrained models - MobileNet, ResNetClassifier, and VGG	Tracks the people and generates an alarm when there is violation	F1-score-99%, sensitivity-99%, specificity-99%, accuracy-100%

As suggested by the literature, the popular deep neural network model such as Convolutional Neural Networks (CNN) [13–15] processes 3D data like images and extracts the features from it for detection. The relevant features are learned while the network trains on a collection of images ensuring high accuracy. The YOLO (You Only Look Once) [16] framework handles object detection differently, it puts the entire image in one instance and predicts bounding box coordinates and class probabilities for these frames. The main contributions of this paper include

- To devise a new approach for object detection that can accurately identify objects in real-time from video streams using transfer learning techniques.
- To assess the effectiveness of the proposed face mask detection method by comparing its performance with various state-of-the-art methods.

2 Methodology

The proposed method analyzes the human actions from live stream videos to determine the existence of face mask and ensure social distancing between humans as a safety measure against COVID-19. The proposed method includes face mask detection module and social distance detection module. The video streaming is continuously monitored for face mask detection and Euclidean distance is measured between the humans for checking social distance. The face mask detection module counts the number of people and label them with either ”with mask” or ”without mask” by bounding boxes. The social distancing detection module checks for social distancing between humans and displays ”safe” or ”not safe” based on the calculated distance in the social distancing module. The architecture of detecting face mask and social distancing is illustrated in the Figure 1.

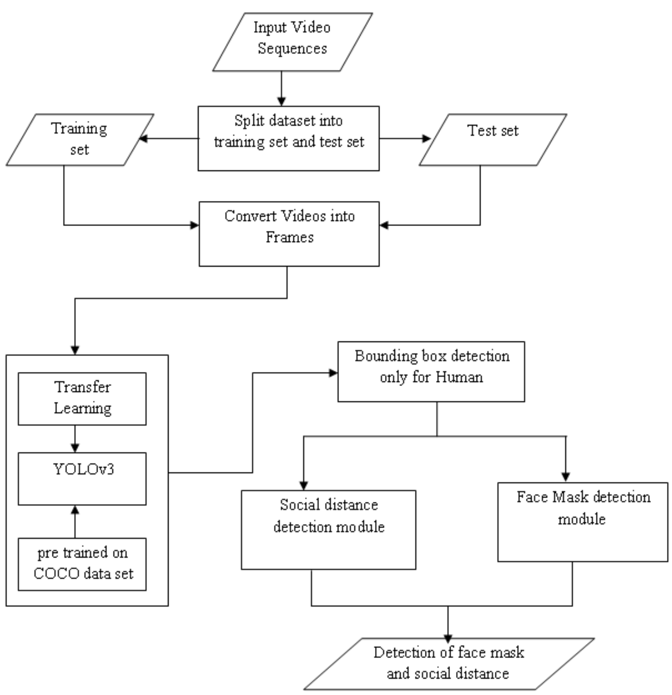

Fig 1. Social distancing and face mask detection architecture

2.1 Dataset

An experimental evaluation was conducted using a dataset from Kaggle, consisting of 3900 images of people's faces captured under varying light conditions and temperatures. The images were divided into two categories based on whether the person was wearing a mask or not. The model was trained on 70% of the images, while 20% were used for testing and another 10% for validation.

2.2 Data Augmentation

Data augmentation expands the amount of data by adding lightly modified copies of existing data or newly generated synthetic data. This helps to improve the performance of deep learning models by generating new and different samples in training dataset. The data is then supplemented so as to extend the size of the dataset. Various operations such as flipping, rotation, shearing, and HSV shift are carried out on the training data to provide the object detector with more characteristics for training.

2.3 Human Detection Model

The real time live video is captured and converted to image frames for detecting face and ensuring social distancing between the humans. These frames are inputted to the human detection model for detecting the humans by enclosing in rectangular boundary boxes.

YOLO framework is employed for human detection [17]. YOLO utilizes Convolutional Neural Networks (CNN) in which a single forward propagation through neural network is used for object detection. YOLO uses single bounding box regression to predict object height, width, center, and class. The boundary points are generated by calculating the boundary regression to draw rectangular boundary boxes around human object in the image frame. Each grid cell is accountable for determining bounding boxes around human object and their confidence values. The value of Intersection over Union (IOU) is 1when the predicted bounding box is same as the real box. IOU makes sure that the predicted bounding boxes are equivalent to the precise human objects. This fact reduces a redundant bounding box that does not match with the human objects. Hence, the bounding box coordinate are predicted with class probabilities for the boxes. The resultant frame with bounded box around humans along with pre-trained weights is fed to the face mask detection model.

2.4 Face Mask Detection

The presence of Face masks in humans can be detected with the assistance of Convolutional Neural Networks (CNNs) [18]. CNNs surmount alternative algorithms by their superior performance with image inputs. The model get the input as the image from video and perform classification based on the classes mask and no mask.

The pre-trained CNN models [19] can be used for a variety of computer vision tasks, including image classification, object detection, and segmentation. In this work, Inceptionv3, DenseNet and MobileNetv2 models are used for detecting face mask. Inception models uses a combination of convolutional layers with different filter sizes and max pooling to capture features at multiple scales. DenseNet uses skip connections to allow information to flow through the network more efficiently. MobileNet is a lightweight CNN model designed for mobile devices that uses depth wise separable convolutions to reduce the number of parameters.

2.4.1 Improved MobileNetV2 Model

MobileNetV2 architecture [20] is improved by adding the Squeeze and Excitation (SE) blocks. These blocks are designed to capture interdependencies between channels and improve the representational power of the network. SE blocks use a global pooling operation to compute channel-wise feature maps, which are then passed through a squeeze operation to reduce their dimensionality. The resulting feature maps are then passed through an excitation operation that learns a set of weights to re-weight the channels based on their importance. The improved MobileNetV2 architecture as shown in Figure 2 is more efficient and accurate, making it a popular choice for various mobile and embedded applications.

To improve the image learning capability of traditional MobileNet, the network structure was modified. This involved abandoning the original classification layer and adding an SE block behind the pre-trained MobileNet. A convolutional layer was then introduced for high-dimensional feature extraction. To enhance the ability to identify face masks, the fully connected layer was replaced with a global average pooling layer, and a new fully connected layer with 1024 neurons and a ReLU activation function was added. Finally, a top layer with a fully connected Softmax layer was used for classification. The parameters of Improved MobileNetV2 model is given in the Table 2.

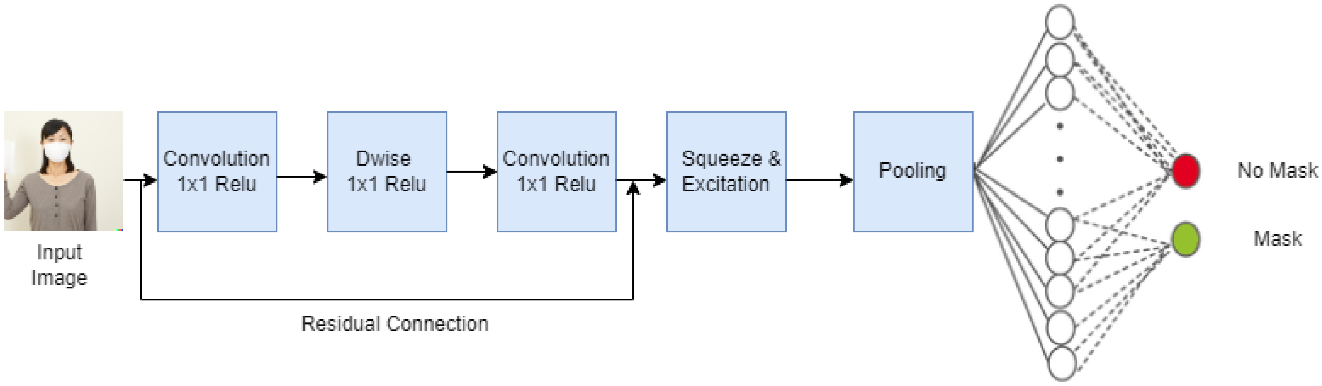

Fig 2. Improved MobileNetV2

Table 2. Parameters of Improved MobileNetV2 model

No. of Layers	Module (Type)	Stride	Filter Shape	Input Size
1	Convolution (Conv)	s2	3x3x3x32	224x224x3
2	Convolution depthwise (dw)	s1	3x3x32 dw	112x112x32
3	Conv	s1	1x1x32x64	112x112x32
4	Conv dw	s2	3x3x64 dw	112x112x64
5	Conv	s1	1x1x64x128	56x56x128
6	Conv dw	s1	3x3x128 dw	56x56x128
7	Conv	s1	1x1x128x256	28x28x256
8	Conv dw	s2	3x3x256 dw	28x28x256
9	Conv	s1	1x1x256x512	14x14x512
10	Conv dw	s2	3x3x512 dw	14x14x512
11	SE Block	s1	7x7x512	7x7x512
12	Average Pool	s1	Pool 7x7	7x7x512
13	Fully Connected	s1	1024	1x1x1024
14	Softmax	s1	Classifier	1x1x25

The squeeze operation reduces the size of feature maps by applying global average pooling to obtain channel-wise statistics, resulting in a contraction of the spatial dimensions (w × h). F_s refer to the squeeze operation, F_e corresponds to the excitation operation, and F_{sc} represents a scaling operation that assigns weights to the features of each channel. Equation (1) outlines the calculation process for each feature map.

$$F_s = \frac{1}{H\ X\ W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i, j) \tag{1}$$

Where u_c represents the c^{th} feature map of previous convolution operation, W and H denote the width and height of u_c.

The purpose of the excitation operation is to comprehensively capture the interdependencies among channels by utilizing the information gathered during the squeeze operation, as indicated in Equation (2).

$$\mathrm{F_e} = \sigma\left(W_2 \delta\left(W_1\left(F_s\right)\right)\right) \tag{2}$$

Where δ represents the rectified linear unit). The parameters W_1 and W_2 are acquired through two fully connected layers positioned around the non-linearity. Consequently, the rescaling of u_c with the activations allows the output of the SE block to be obtained, as demonstrated in Equation (3).

$$\widetilde{X} = F_{sc}(u_c\ ,\ s_c\ ) \tag{3}$$

The model was trained using transfer learning in two steps. In the first step, only the newly added layers' parameters were learned from scratch while keeping the pre-trained weights of the bottom convolutional layers from ImageNet frozen. In the second step, the model was fine-tuned by loading the pre-trained model obtained in the first step and training all the weights using the target dataset.

2.5 Social Distancing Detection

The YOLO (You Only Look Once) object detector is used for detecting people in the image frames and Euclidean distance is calculated between the centroids of bounded boxes. YOLO object detector is an effective real time detector that is considered as a regression problem that determines the class probabilities for each bounding box. YOLO V3 algorithm is used to detect the social distance. The steps involved in the detection of social distance monitoring is illustrated in the Figure 3.

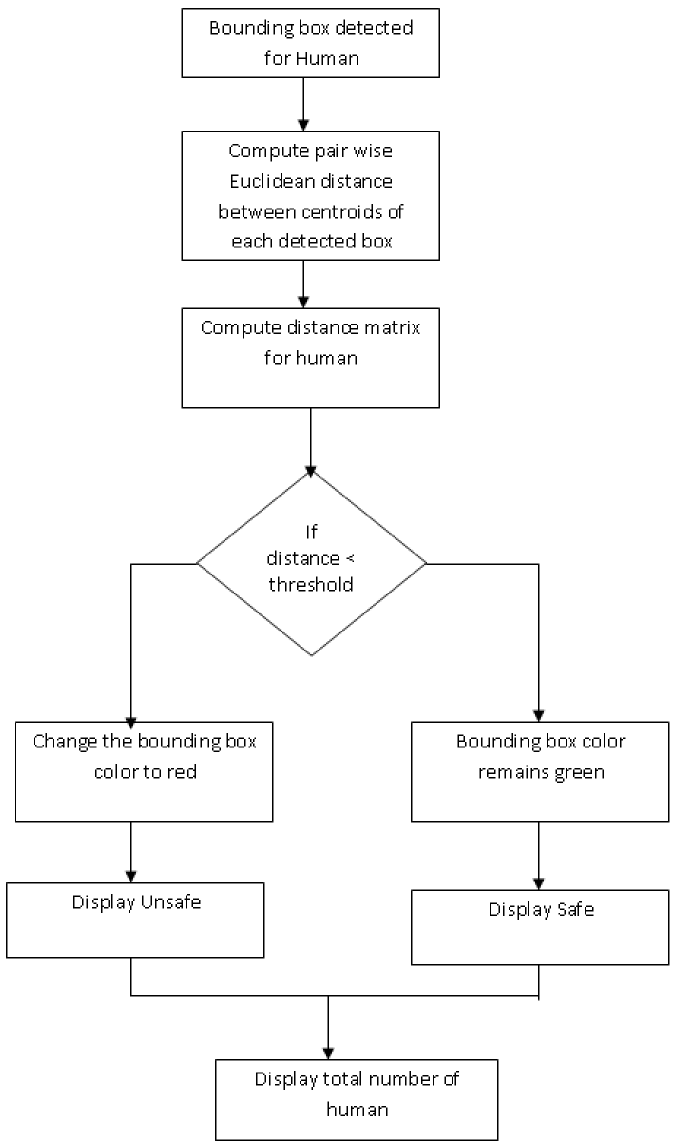

Fig 3. Monitoring of Social distance architecture

The person in a frame is detected by using YOLO Object detection and the Euclidean distance between the centroids of the detected boxes is determined. This determines the probabilities of each class for every bounding box. In a single estimation, the bounding box and class probability for the frame is determined. This method is able to detect the images at an outperforming speed of 45 frames per second (FPS).

In the process of human detection, it checks whether the human count exceeds one. If the human count does not satisfy the requirements, then the model operations are skipped and it displays message "Safe" in the output screen. To evaluate the distance between the more than one human object, the distance differentiate method is used. Calculation of Euclidean distance between the human object and locating the centroids of each detected person helps to find the safest social distance among the people. The centroids of the boundary boxes are extracted and the Euclidean distance is computed between all pairs of centroids.

Euclidean distance is the length between two points of a line segment. If (x1, y1) and (x2, y2) are the point coordinates and d refers to the distance between the points, then Euclidean distance is calculated using equation (4).

$$d = \sqrt{(x2 - x1)^2 + (y2 - y1)^2} \tag{4}$$

If the computed distance is smaller than a threshold value then the person is violating social distancing. The value of the Euclidean distance should be less than 150.

3 Results and Discussion

3.1 Performance Metrics

The performance of the proposed system is evaluated using precision, recall, F1 score and accuracy.

Precision refers to the proportion of correctly identified positive instances

$$\text{Precision} = \frac{\text{tp}}{\text{tp} + \text{fp}} \tag{5}$$

Recall refers to the proportion of identifying the actual positives.

$$\text{Recall} = \frac{\text{tp}}{\text{tp} + \text{fn}} \tag{6}$$

Accuracy refers to the relation of true positives and true negatives and the total number of inputs under consideration.

$$\text{Accuracy} = \frac{\text{tp} + \text{tn}}{\text{tp} + \text{fp} + \text{tn} + \text{fn}} \tag{7}$$

Here, tp, tn, fp, fn represents true positives, true negatives, false positives and false negatives respectively.

F1-score is determined by combining the precision and recall as a single metric by finding the harmonic mean.

$$F1 - Score = 2 * \frac{precision * recall}{(precision + recall)} \tag{8}$$

3.2 Experimental Results

The various pretrained CNN models like Inceptionv3, DenseNet and MobileNetV2 are used and MobileNetv2 outperforms the other two models. The performance metrics of the various models are shown in Table 3.

The softmax prediction is based on the class that has the highest probability. The model is implemented using Adam optimizer with 20 epochs and a learning rate is 0.001. The model detects with an accuracy of 98% on the validation set. The training loss and accuracy of the model is shown in Figure 4.

The performance comparison of Inceptionv3, DensNet and MobileNetv2 based on precision, recall and f1-score by taking with mask, without mask, macro avg and weighted avg is shown in Figure 5.

Once the humans are detected using bounded boxes, the presence or absence of face mask is determined using the face mask detection module. It displays a bounding box in green color with the message "with mask" for people wearing mask and displays a bounding box in red color with the message "No mask" for people without mask. The outputs are shown in Figure 6.

After identifying the presence of mask, the Euclidean distance between the centroids of the bounding boxes are determined to ensure whether social distancing is preserved or not. It displays a boundary box in green or red around humans denoting the message "safe" or "not safe", based on the distance calculated in the social distancing detection module as given in Figure 7.

Thus the proposed system aims to determine the face mask and social distancing, thereby helping to follow the COVID-19 restrictions to avoid the widespread of Coronavirus. Performance of the proposed system is manually evaluated using real time videos. The average accuracy of the system is 98% and average F1 score is 98%.

Table 3. Performance metrics using pretrained CNN models

		Precision	Recall	F1-score	Support	Accuracy
InceptionV3	With mask	0.83	0.85	0.84	520	0.88
	Without mask	0.80	0.77	0.78	385	
	Macro avg	0.81	0.81	0.81	910	
	Weighted Avg	0.82	0.82	0.82	910	
DenseNet	With mask	0.92	0.91	0.92	520	0.91
	Without mask	0.88	0.90	0.89	385	
	Macro avg	0.90	0.91	0.90	910	
	Weighted Avg	0.91	0.91	0.91	910	
Improved MobileNetV2	**With mask**	**0.98**	**0.98**	**0.98**	**520**	**0.98**
	Without mask	**0.98**	**0.98**	**0.98**	**385**	
	Macro avg	**0.98**	**0.98**	**0.98**	**910**	
	Weighted Avg	**0.98**	**0.98**	**0.98**	**910**	

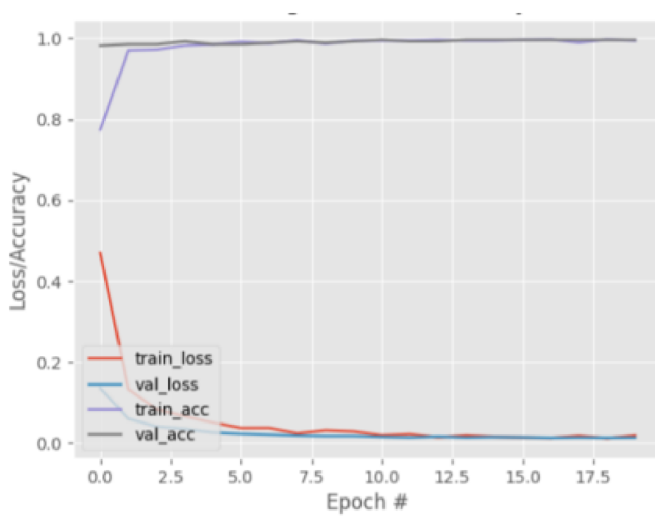

Fig 4. Improved MobileNetV2- Training Loss and accuracy

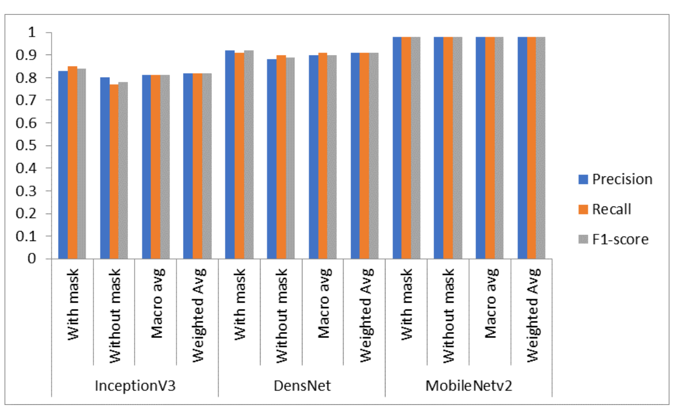

Fig 5. Performance comparison of MobileNetv2 with other models

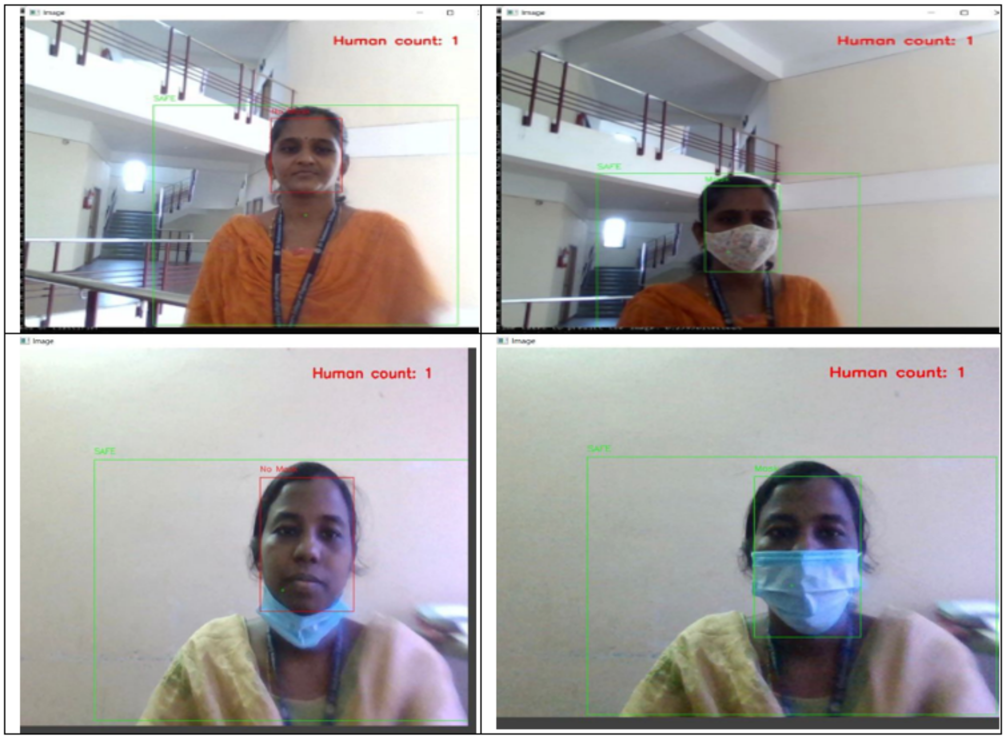

Fig 6. Face Mask Detection

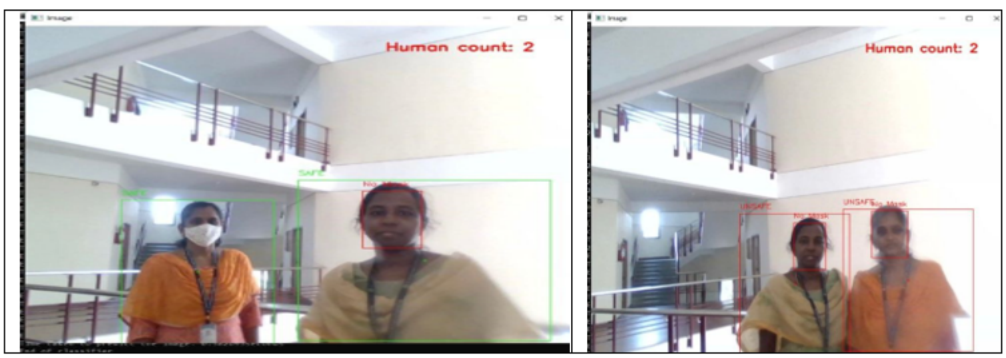

Fig 7. Social Distancing detection module

4 Conclusion

This study presents an improved model on capturing the video of humans and check whether the human wears a face mask in the public places. The proposed model detects the social distancing between humans accurately and the improved MobileNetv2 outperforms DenseNet and Inceptionv3 and produces a precision, recall and f1-score of 98%. The use of Improved MobileNetV2 for face mask detection and Euclidean distance for distance computation also makes it suitable for deployment on embedded systems. However, in order to improve the performance of the model, it may be necessary to optimize the frames per second and inference time involved in making predictions.

References

1) Ghadekar P, Singh G, Datta J, Gupta AK, Anand DS, Khare S, et al. COVID-19 Face Mask Detection using Deep Convolutional Neural Networks & Computer Vision. *Indian Journal of Science and Technology*. 2021;14(38):2899–2915. Available from: https://doi.org/10.17485/IJST/v14i38.996.

2) Walia IS, Kumar D, Sharma K, Hemanth JD, Popescu DE. An Integrated Approach for Monitoring Social Distancing and Face Mask Detection Using Stacked ResNet-50 and YOLOv5. *Electronics*. 2021;10(23):2996. Available from: https://doi.org/10.3390/electronics10232996.
3) Kaviya P, Chitra P, Selvakumar B. A Unified Framework for Monitoring Social Distancing and Face Mask Wearing Using Deep Learning: An Approach to Reduce COVID-19 Risk. *Procedia Computer Science*. 2023;218:1561–1570. Available from: https://doi.org/10.1016/j.procs.2023.01.134.
4) Joolfoo MBA, Hosany MA. Machine Learning Solutions in Combating COVID-19: State of the Art and Challenges. *Athens Journal of Technology & Engineering*;2023:1–18. Available from: https://www.athensjournals.gr/technology/2022-4777-AJTE-Joolfoo-05.pdf.
5) Javed I, Butt MA, Khalid S, Shehryar T, Amin R, Syed AM, et al. Face mask detection and social distance monitoring system for COVID-19 pandemic. *Multimedia Tools and Applications*. 2023;82(9):14135–14152. Available from: https://doi.org/10.1007/s11042-022-13913-w.
6) Ahamad AH, Zaini N, Latip MFA. Person Detection for Social Distancing and Safety Violation Alert based on Segmented ROI. *2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*. 2020;p. 113–118. Available from: https://doi.org/10.1109/ICCSCE50387.2020.9204934.
7) Kaur G, Sinha R, Tiwari PK, Yadav SK, Pandey P, Raj R, et al. Face mask recognition system using CNN model. *Neuroscience Informatics*. 2022;2(3):100035. Available from: https://doi.org/10.1016/j.neuri.2021.100035.
8) Teboulbi S, Messaoud S, Hajjaji MA, Mtibaa A. Real-Time Implementation of AI-Based Face Mask Detection and Social Distancing Measuring System for COVID-19 Prevention. *Scientific Programming*. 2021;2021:1–21. Available from: https://doi.org/10.1155/2021/8340779.
9) Meivel S, Sindhwani N, Anand R, Pandey D, Alnuaim AA, Altheneyan AS, et al. Mask Detection and Social Distance Identification Using Internet of Things and Faster R-CNN Algorithm. *Computational Intelligence and Neuroscience*. 2022;2022:1–13. Available from: https://doi.org/10.1155/2022/2103975.
10) Vu HN, Nguyen MH, Pham C. Masked face recognition with convolutional neural networks and local binary patterns. *Applied Intelligence*. 2022;52(5):5497–5512. Available from: https://doi.org/10.1007/s10489-021-02728-1.
11) Goyal H, Sidana K, Singh C, Jain A, Jindal S. A real time face mask detection system using convolutional neural network. *Multimedia Tools and Applications*. 2022;81(11):14999–15015. Available from: https://doi.org/10.1007/s11042-022-12166-x.
12) Sethi S, Kathuria M, Kaushik T. Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread. *Journal of Biomedical Informatics*. 2021;120:103848. Available from: https://doi.org/10.1016/j.jbi.2021.103848.
13) Kumar A, Kaur A, Kumar M. Face detection techniques: a review. *Artificial Intelligence Review*. 2019;52(2):927–948. Available from: https://doi.org/10.1007/s10462-018-9650-2.
14) Lakshmanan K, John K, Simon RM, Aaditya. Face Mask Detection and Social Distance Monitor. *International Journal of Engineering Research & Technology (IJERT)*. 2022;11. Available from: https://doi.org/10.17577/IJERTV11IS060192.
15) Kalsait S, Pote S, Muskande S, Patil OSS. Face Mask Detection and Social Distance Monitoring using Machine Learning Architecture. *International Journal of Innovative Science and Research Technology*. 2022;7(3):1189–1190. Available from: https://doi.org/10.5281/zenodo.6463272.
16) Swetha S, Vijayalakshmi J, Gomathi S. Social Distancing and Face Mask Monitoring System Using Deep Learning Based on COVID-19 Directive Measures. *2021 4th International Conference on Computing and Communications Technologies (ICCCT)*. 2021;p. 520–526. Available from: https://doi.org/10.1109/ICCCT53315.2021.9711880.
17) Bodini M. A Review of Facial Landmark Extraction in 2D Images and Videos Using Deep Learning. *Big Data and Cognitive Computing*. 2019;3(1):14. Available from: https://doi.org/10.3390/bdcc3010014.
18) Singh R, Ahmed T, Singh R, Udmale SS, Singh SK. Identifying tiny faces in thermal images using transfer learning. *Journal of Ambient Intelligence and Humanized Computing*. 2020;11(5):1957–1966. Available from: https://doi.org/10.1007/s12652-019-01470-4.
19) Yang D, Yurtsever E, Renganathan V, Redmill KA, Özgüner Ü. A Vision-Based Social Distancing and Critical Density Detection System for COVID-19. *Sensors*. 2021;21(13):4608. Available from: https://doi.org/10.3390/s21134608.
20) Nagrath P, Jain R, Madan A, Arora R, Kataria P, Hemanth J. SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. *Sustainable Cities and Society*. 2021;66:102692. Available from: https://doi.org/10.1016/j.scs.2020.102692.