Journal of Engineering Sciences
Vol 12, Issue 7, July/ 2021
ISSN NO:0377-9254

Social Deprivation with Protective Mask Detector

DR.Ch.Rajendra Babu.[1], K.JyothirVijayaLakshmi.[2*], K.MeghanaDurga Saisri.[2], Sk.Rafath Anjum.[2]
[1]Professor & HoD, Department of Computer Science And Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada, Andhra Pradesh, India.
[2]Andhra Loyola Institute of Engineering and Technology, Vijayawada, Andhra Pradesh, India.
***Corresponding Author's email:** kjyothirvijayalakshmi@gmail.com

Abstract. The Covid-19 had a huge impact on different sectors in many countries, so to control the wide spreading of any virus, we proposed a social distance with face mask detection system that controls people in public places like colleges, schools, malls, theatres. By installing this system at public places, the manual checking is replaced with automated checking of peoples's social distance and face mask. From the captured video, initially social distance takes place then face mask detection takes place. In social distance,for person detection we are using pretrained darknet modelwhich is a DNN object detector in OpenCV and for mask detection we are using darknet model trained with custom dataset. To develop a model that classifies a person following social distance or not and wears a face mask or not, we are using Yolov4 architecture. If the detector result shows that more than 10 violations then automatically alarm gets activated to alert people that they are in danger and also an email will be sent to management to alert then that there place is in danger to indicate them to be careful.We have developed a web application in which user can input an image/ a video/ connect to camera(live capture). Alleviation in any pandemic situation can be controlled by social distancing as well as wearing face mask which would save lots of lives.

Keywords: OpenCV, YOLO,machine learning, darknet, deep neural network(DNN).

1 Introduction

In thissociety, the pandemic of COVID-19 is causing a worldwide emergency and had a huge impact on the lives of humans. This pandemic caused the loss of billions of human lives. Coronovirus is caused by the SARS-CoV-2 virus, which spreads between people in several different ways. The virus can spread from an infected person's mouth or nose in small liquid particles when they cough, sneeze, speak, sing or breathe. These particles range from larger respiratory droplets to smaller aerosols. Hence, World Health Organization (WHO) stated that by wearing a mask and also by maintainingphysical distance the chance of virus spreading can be reduced.

A mask is NOT a substitute for social distancing. Masks should still be worn in addition to staying at least 6 feet apart, especially when indoors around people who don't live in your household. When you wear a mask, you protect others as well as yourself. Masks work best when everyone wears one.

The probability of virus spread is very high in public places like colleges, schools, malls, etc. In these places to make sure that people following physical distance and wearing a mask, we are developing a alert based system that allowspeopleget intimated the danger situation.

The central objective of the system is to make ensure that people are maintain physical distance and wearing a face mask in public places to prevent the spread of the virus from person to person. The other objective of the system is automating the manual checking of whether rules followed by people or not. To develop this system, we used the concept namely Machine Learning.

Computer Vision is a scientific field that deals with how computers can gain high-level understanding with the data present in the form of images and videos. This trains computers to extract the features present in the images and videos. It also identifies and classifies the objects present in the images. This concept is used in our system to identify person then followed by identifying the face mask present in the images that are taken as input.

2 Literature Survey

Sahana Srinivasan et al.,[1] proposed a comprehensive and effective solution to perform person detection, social distancing violation detection, face detection and face mask classification using YOLOv3, Density-based spatial clustering of applications with noise (DBSCAN), Dual Shot Face Detector (DSFD) and MobileNetV2 based binary classifier have been employed on surveillance video datasets. The system performs with an accuracy of 91.2% and F1 score of 90.79% on the labelled video dataset and has an average prediction time of 7.12 seconds for 78 frames of a video.

Savyasachi Gupta et al.,[2] Elucidated a novel framework called SDMeasure which aims to determine whether a set of people are following 'Social Distancing' guidelines of maintaining a minimum distance of 6 feet (or 1.8 metres) when observed from video footage of a public area.

Afiq Harith Ahamad et al.,[3] Proposed a social distance detector that focuses on detecting people in areas of interest using the MobileNet Single Shot Multibox Detector (SSD) object tracking model and OpenCV library for image processing. The distance will be computed between the persons detected in the captured footage and then compared to a fixed pixel values.

Yew Cheong Hou et al.,[4] Proposed a social distance detector usingthe open-source object detection pre-trained model based on the YOLOv3 algorithm was employed for pedestrian detection. Later, the video frame was transformed into top-down view for distance measurement from the 2D plane.

Joseph Redmon et al.,[5] YOLOv2 performs about 40 Frames Per Second (FPS) and accuracy better than YOLOv1 and other models such as SSD and R-CNN.

Joseph Redmon et al., [6] YOLOv3 is faster and accurate as Single Shot Multibox(SSD). It can recognize 80 different objects from real-time video feed or images.

Alexey Bochkovskiy et al.,[7] YOLOv4 is faster speeds at 65 Frames Per Second.It is incremental model of YOLOv3 and it has YOLOv4 Tiny also.

Chandrappa D.N et al.,[8] Described a face detection system that uses Boosted Cascade Features. Here AdaBoost algorithm, a cascade of classifiers, and Haar classifiers are used to detect a face. Since the classifiers used are of different types, the efficiency of the system varies between different datasets. This system also works efficiently for frontal face detection.

S. Meivel et al.,[9] This paper describes mask detection using Matlab specified the Faster R-CNN algorithm and Social Distancing using Yolov2.

Shashi Yadav[10] Deep Learning based Safe Social Distancing and Face Mask Detection in Public Areas for COVID-19 Safety Guidelines Adherence using Convolutional Neural Networks (CNNs), Single Shot Detector.

3 Proposed System

3.1 System Design

The proposed system automates the manual checking of a person's social distance and face mask. The automation takes place because of linking the OpenCV with email and alarm. The block diagram of the system is shown in Figure 1.

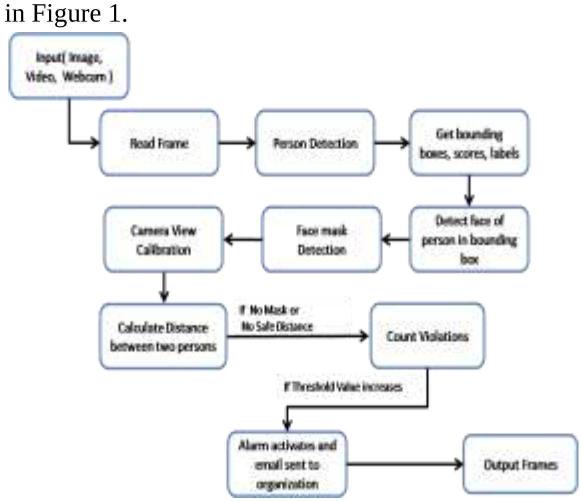

Figure 1- Block diagram of the system.

This system is installed along with a camera which captures video at the public areas. From the captured video, initially person detection and caluculate pairwise diatance then face mask detection takes place. To detect the persons, we are using a darknet model which is a Deep Neural Network Object detector in OpenCV. Only when the person is detected then the mask detection takes place. To develop the model that classifies the person, we are using pretrained Yolov4 and To develop the model that classifies the person wears a face mask or not, we are using Yolov4 architecture. The developed model takes the frame obtained from the video stream and classifies the output into 2 classes namely mask and no_mask. Based on the detetcor result the working of the alarm and email sending depends. If the detector result shows that more than 10 rule violations exits then alarm gets activated and email is sent to organization to alert both people and management else if the result is violations >0 or <10 then the gate warning message else it indicates safe. The workflow of the system is shown in Figure 2.

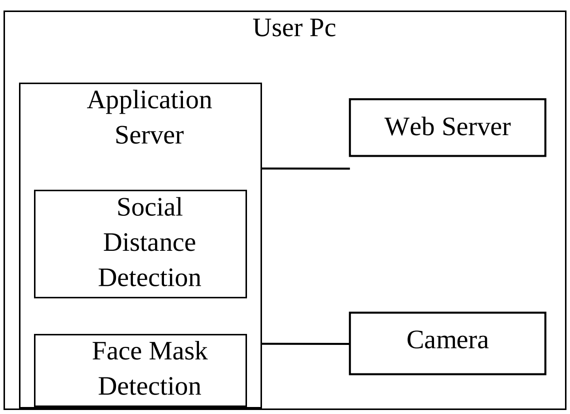

Figure 2 - Workflow of Detector.

3.2 System Requirements

The hardware requirements of the system are:

Camera: It is used for capturing a video.

Intel CORE i3 or above.

The software requirements of the system are:

OpenCV: It is a python library that consists of many in-built functions to perform image processing and machine learning tasks. It mainly focuses on capturing the video, extracting the features. It is useful for object detection and face detection. For loading darknet models.

Playsound: Playsound is a pure Python, cross platform, single function module with no dependencies for playing sounds.

Smtplib:
Python provides smtplib module, which defines an SMTP client session object that can be used to send mail to any Internet machine with an SMTP.

3.3YOLOv4 Architecture

YOLOv4 is a one stage detector. The working of one stage detectors is shown in Figure 3.

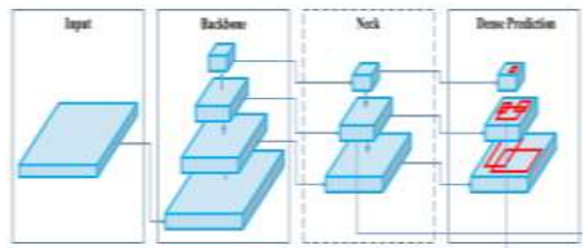

Figure 3 - One Stage Detector Architecture

Backbone: It's a deep neural network composed mainly of convolution layers. The main objective of the backbone is to extract the essential features, the selection of the backbone is a key step it will improve the performance of object detection. Often pre-trained neural networks are used to train the backbone.

The YOLOv4 backbone architecture is composed of three parts:

- Bag of freebies
- Bag of specials
- CSPDarknet53
-

Neck(detector): The role of the neck in the case of one stage detector is to collect feature maps from different stages. This contains additional blocks SPP(Spatial Pyramid Pooling) and path aggregation blocks PAN(Modified Path Aggregation Network).

Head(detector): The role of the head in the case of a one stage detector is to perform dense

Vol 12, Issue 7, July/ 2021
ISSN NO:0377-9254

prediction(YOLO). The dense prediction is the final prediction which is composed of a vector containing the coordinates of the predicted bounding box (center, height, width), the confidence score of the prediction and the label.

The working of YOLOv4 is shown in Figure 4.

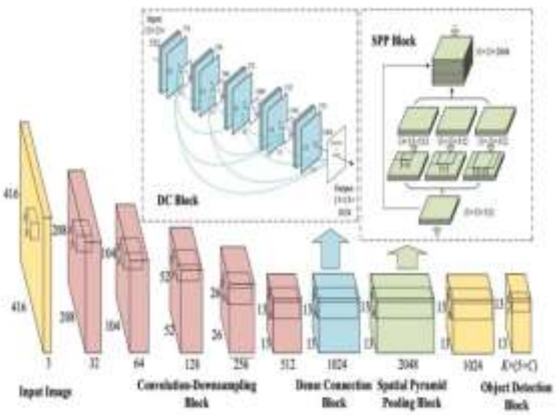

Figure 4 -YOLOv4 Architecture

3.4 Dataset

The Dataset used for model training was collected from Kaggle Repository, Google. It has two class labels namely mask and no_mask. The dataset considered contains nearly 1000 images. Here the concept of image augmentation is used to expand the size of the dataset so that the ability to classify and performance of the model can be improved and labeled using roboflow. The sample images of dataset are shown in Figure 5.

Figure 5 - Sample dataset images

3.5 Implementation:

The proposed system can be developed by following the below steps:

Step1 - Training and saving weights of Face Detection model: Before going to face mask detection, it is necessary to develop a model to detect a face masks. A face mask detection model can be developed using many algorithms like MobileNetV2. But here we are using a DNN detector YOLOv4 using OpenCV, as it can detect a face masks in a single pass. It is a Darknet model that uses a CSPDarknet53 Architecture.

Step2- Developing, training, and testing a Face Mask Detection model. The steps to train and test the model are given below:

1.Images were annoted using Roboflow.

2. Install Darknet for Mac or Windows first.

3. Create configuration file (.cfg), data file (.data) and install pretrained weights of yolo.

4. Train and test the performance using darknet commands.

Step3- Testing with dynamic input:

1. Load the pretrained YOLOv4 object detection model and face mask detection model that was developed in Step 2.

2. Start the web application then user selects input type.

3. If input type is image goto step 6 Else If input type is video goto step 4 Else input type is live goto step 5.

4. Convert video into frames goto step 6.

5. Start the video streaming and web camera starts receiving the video and converts it into frames (i.e., image)goto step 6.

6. Resize the image into 416*416 dimensions as network size and preprocess the image.

7. With the help of a social distance detector module detect persons, calculate distance after camera view calibration andusing face detection model detect the face masks in each image.

8. Count the violations of social distance and face masks.

9. If the violations are greater than 10 alarm gets activated and email sent to organization indictaing danger situation.

4 Results

As we discussed in the previous section, a face mask detection model has been developed and trained. We have trained our dataset using YOLOv2, YOLOv3, YOLOv4.

1.Performance metrics of model training are shown in Figure 6, Figure 7, Figure 8.

The performance metrics that are considered while developing the model are Precision, Recall, F1-score, and accuracy.

Accuracy is the ratio of correct predictions made to the total number of predictions performed. The accuracy obtained for this system is nearly 94%. The formula used to calculate accuracy is:

Accuracy = No. of predictions / Total No. of predicitions

Precision is the ratio between correct positive predictions to the total positive predictions. The precision obtained is 0.89 which is near to 1.0 is a perfect precision value. The formula used to calculate precision is:

Precision = No.of correct positive predictions/Total No. of positive predictions

Recall is the ratio between correct positive predictions to total positive examples. The recall value obtained is also 0.95 which is a high and good value. The formula used to calculate recall is:

Recall = No. of correct positive predictions / Total No. of values in the positive class

F1-score is the average of precision and recall values. The f1-score obtained is 0.92. The formula used to calculate f1-score is:

F1-score = precision + recall

```
detections_count = 7495, unique_truth_count = 659
class_id = 0, name = no_mask, ap = 53.92%        (TP = 97, FP = 83)
class_id = 1, name = mask, ap = 81.33%           (TP = 412, FP = 136)

for conf_thresh = 0.25, precision = 0.70, recall = 0.77, F1-score = 0.73
for conf_thresh = 0.25, TP = 509, FP = 219, FN = 150, average IoU = 50.49 %

IoU threshold = 50 %, used Area-Under-Curve for each unique Recall
mean average precision (mAP@0.50) = 0.676260, or 67.63 %
Total Detection Time: 4 Seconds
```

Figure 6 - Performance metrics of YOLOv2

```
detections_count = 999, unique_truth_count = 659
class_id = 0, name = no_mask, ap = 67.00%        (TP = 107, FP = 27)
class_id = 1, name = mask, ap = 92.66%           (TP = 451, FP = 34)

for conf_thresh = 0.25, precision = 0.90, recall = 0.85, F1-score = 0.87
for conf_thresh = 0.25, TP = 558, FP = 61, FN = 101, average IoU = 69.70 %

IoU threshold = 50 %, used Area-Under-Curve for each unique Recall
mean average precision (mAP@0.50) = 0.798308, or 79.83 %
Total Detection Time: 4 Seconds
```

Figure 7 - Performance metrics of YOLOv3

```
detections_count = 994, unique_truth_count = 659
class_id = 0, name = no_mask, ap = 91.19%        (TP = 140, FP = 24)
class_id = 1, name = mask, ap = 97.11%           (TP = 485, FP = 51)

for conf_thresh = 0.25, precision = 0.89, recall = 0.95, F1-score = 0.92
for conf_thresh = 0.25, TP = 625, FP = 75, FN = 34, average IoU = 71.77 %

IoU threshold = 50 %, used Area-Under-Curve for each unique Recall
mean average precision (mAP@0.50) = 0.941507, or 94.15 %
Total Detection Time: 6 Seconds
```

Figure 8 - Performance metrics of YOLOv4

2. After training the model, The working of the system was tested with the dynamic input. The results obtained during the system working are shown in Figure 9, Figure 10 and Figure 11.

Figure 9 - Output with status warning

Figure 10 - Output with status Danger

Figure 11 - Sending email when danger situation.

3. User Interface of our project looks as Figure 12, Figure 13 and Figure 14 and Camera View Calibration as in Figure 15.

Figure 12 – Main Page.

Figure 13 – Uploading Image Page.

Figure 14 – Uploading Video Page.

Figure 15 – Camera View Calibration Coordinates.

5 Conclusion

In this paper, an automated social diatance and face mask detection system has been developed. From the video collected by a camera or images and videos given by user, social diatance and face mask detection takes place by using the darknet modelYOLOV4, we identified whether people following social distance and wears a mask or not and the working of the alarm and email sending depends on the output predicted by the detection model. The results have shown that the accuracy of the system is nearly 95% which is accurate and efficient. In the future, we can also extend the project by checking improper mask and whether the mask is N95 or not. One Limitation with our model is mannequin's are also detecting as human we want to work on differentiation of dummy object and real object.

References

1. Sahana Srinivasan, R Rujula Singh, RuchitaR Biradar, SA Revathi (2021),"COVID-19 Monitoring System using Social Distancing and Face Mask Detection on Surveillance video datasets", IEEE.
2. Savyasachi Gupta, Rudraksh Kapil, Goutham Kanahasabai, Shreyas Srinivas Joshi, Aniruddha Srinivas Joshi (2021), "SD-Measure: A Social Distancing Detector", IEEE.
3. Afiq Harith Ahamad, Norliza Zaini, MohdFuad Abdul Latip (2020), "Person Detection for Social Distancing and Safety Violation Alert based on Segmented ROI", IEEE.
4. Yew Cheong Hou, Mohd Zafri Baharuddin, Salman Yussof, SumayyahDzulkifly (2020), "Social Distancing Detection with Deep Learning Model", IEEE.
5. Joseph Redmon, Santosh Divvala, Ross Girshick , Ali Farhadi (2016), "You Only Look Once: Unified, Real-Time Object Detection", ArXiv.
6. Joseph Redmon, Ali Farhadi (2018), "YOLOV3: An Incremental Model", ArXiv.
7. Alexey Bochkovskiy (2020), "YOLOv4: Optimal Speed and Accuracy of Object Detection", ArXiv.
8. Chandrappa D.N, Akshay G, RavishankarM (2012), "Face Detection Using a Boosted Cascade of Features Using OpenCV", Springer.
9. S. Meivel, K. Indira Devi, S. Uma Maheswari,and J. VijayaMenaka (2020), "Real time data analysis of face mask detection and social distance measurement using Matlab",PMC.
10. Shashi Yadav (2020),"Deep Learning based Safe Social Distancing and Face Mask Detection in Public Areas for COVID-19 Safety Guidelines Adherence", IJRASET.
11. Adina Rahim, Ayesha Maqbool, TauseefRana (2021), "Monitoring social distancing under various low light conditions with deep learning and a single motionless time of flight camera", PLoS ONE 16(2): e0247440.
12. Saponara, S., Elhanashi, A. &Gagliardi, A. Implementing a real-time, AI-based, people detection and social distancing measuring system for Covid-19. J Real-Time Image Proc (2021). https://doi.org/10.1007/s11554-021-01070-6.
13. Meenpal, Toshan & Balakrishnan, Ashutosh&Verma, Amit. (2019). Facial Mask Detection using Semantic Segmentation. 1-5. 10.1109/CCCS.2019.8888092.
14. SK Kotamraju, PG Arepalli, SS Kanumalli (2021), Implementation patterns of secured internet of things environment using advanced blockchain technologies, Materials Today: Proceedings, 2021.
15. Gopi A.P., Patibandla R.S.M. (2021) An Efficient Methodology for Avoiding Threats in Smart Homes with Low Power Consumption in IoT Environment Using Blockchain Technology. In: Choudhury T., Khanna A., Toe T.T., Khurana M., Gia Nhu N. (eds) Blockchain Applications in IoT Ecosystem. EAI/Springer Innovations in Communication and Computing. Springer, Cham. https://doi.org/10.1007/978-3-030-65691-1_16.
16. RSML Patibandla, SN Mohanty (2021), Need of Improving the Emotional Intelligence of Employees in an Organization for Better Outcomes,Decision Making And Problem Solving: A Practical guide, 2021.
17. Arjya Das, Mohammad Wasif Ansari, RohiniBasak (2020), "Covid-19 Face Mask

Detection Using TensorFlow, Keras and OpenCV", IEEE.
18. S. Susanto, F. A. Putra, R. Analia and I. K. L. N. Suciningtyas, "The Face Mask Detection For Preventing the Spread of COVID-19 at PoliteknikNegeriBatam," 2020 3rd International Conference on Applied Engineering (ICAE), 2020, pp. 1-5, doi: 10.1109/ICAE50557.2020.9350556.