Quarterly Journal

Volume 14, Special Issue 3 (2022)

RNI No. : UPENG04179

ISSN : 2229-7111 (Print)
ISSN : 2454-5767 (Online)

SAMRIDDHI

An International Journal

(A Journal of Physical Sciences, Engineering & Technology)

Indexing information:

UGC - Journal List,
Journal No. 64061

Crossref

PKP|INDEX

ULRICHSWEB™ GLOBAL SERIALS DIRECTORY

ProQuest

Summon

J-Gate

OPEN ARCHIVES

SMS LUCKNOW

RESEARCH ARTICLE

SAMRIDDHI Volume 14, Special Issue 3, 2022 Print ISSN : 2229-7111 Online ISSN : 2454-5767

Social Distancing and Face Mask Detection using Yolov4 and CNN

D.M. Ujlambkar[1*], Vedashree Bhat[2], Atharva Jadhav[3], Prajwal Desai[4], Siddhesh More[5]

[1*-5] Department of Computer Engineering, AISSMSCOE, Pune, India ; e-mail*: dmujalambkar@aissmscoe.com

Abstract

Social distancing and wearing face mask is the urge of year 2022 due the possible spread of coronavirus . Social distancing is one of the possible approach with which we can continue fighting this pandemic. Motivated by this notion, the ground plan proposed a deep knowledge based framework for automating the task of covering social distancing using surveillance tape. We proposed COVID-19 Social Distancing and face mask detector system which is a one-stage detector, which consists of a machine learning module for calculating the distance among the different classes ,and a CNN model for detecting whether a person is wearing mask or not. YOLO stands for You Only Look Once, this algorithm is employed for Object Finding also as Object Tracking. After we performed recognizing with YOLOv4, we calculated the Euclidean distance between all the detected boxes and filter out or flag the people that are closest to every other indicating that they're at risk. The high the Euclidean distance score would be the better model was within the spotting of objects. The minimum distance to keep while adhering to social distancing is 6 feet , keeping this as the base for calculating distance, the model was trained and used for object detection as well as object tracking. For the mask detection a CNN model was made using Deep learning and OpenCV. The custom datasets were used to understand facemasks and were trained on those datasets for detection and tracking. Predictive Analysis on the static images as well as in the videos to detect face masks has been done. In addition, the system also used existing IP cameras(web cameras) combined with computer vision to detect people without Mask and violence of social distancing..

Keywords: Machine learning, Object detection, Yolov4, Open CV, Face mask detection, Social distance detection.

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology, (2022); DOI : 10.18090/samriddhi.v14spli03.20

Introduction

From the year 2020, it brought us a lot of challenges, especially in the working sectors. Many establishments were switching to a work from home-based environment [3]. The only way to prevent the spread of covid-19 that was suggested by WHO (World Health Organization) was that, people around the world should wear masks to prevent the risk of getting infected by novel coronavirus which would also help in preventing its transmission and also people should maintain social distance of at least 3m among each other to prevent the spread of virus[6]. Though government announced lockdown and started taking preventing measures then to people were just carelessly gathering together which led to result of 2nd wave in month of April 2021,and 3rd wave in Jan 2022. Talking about the current situation of

Corresponding Author : D.M. Ujlambkar, Department of Computer Engineering, AISSMSCOE, Pune, India ; e-mail: dmujalambkar@aissmscoe.com

How to cite this article : Ujlambkar, D.M., Bhat, V., Jadhav, A., Desai, P., More, S. (2022).Social Distancing and Face Mask Detection using Yolov4 and CNN.

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology, Volume 14, Special Issue (3), 442-450.

Source of support : Nil

Conflict of interest : None

India,though the patients are reducing day by day and recovery rate has also increased due to vaccinations but because of the recent less number of covid patients, the covid norms are been lifted by the government and considering this the covid cases may rise in the future.

So this problem can be easily handled and we can yet stop another wave coming to India if we do something by which people will maintain proper distance between them and will also wear masks for their safety.

Thus the goal of this project was to fight against the corona-virus.As social distancing and wearing masks has proven to be a very effective measure to slow down the spread of the disease, thus this was achieved by providing the monitoring system that keeps track of the peoples by using existing IP cameras and CCTV cameras combined with Computer Vision to detect people without Mask and not following rule of social distancing[1].This research-based on Face Mask Detection & Social Distancing used computer vision to understand various aspects of the images or videos based on frames that was provided as an input to the algorithms.*The algorithm used here was YOLO4 which is an Object detection algorithm in machine learning and it was used to detect people in a frame and check for social distancing by calculating the Euclidean distance between the centroids of the detected boxes*[5].The basic concept behind this was to find the bounding boxes related to the classes, the classes could be anything that would range from a Dog to Car depending on the datasets.*Now for the mask detection a CNN model was made using Deep learning and OpenCV.*

Predictive Analysis on the static images as well as in the videos to detect face masks was done. Thus this social distancing and face mask detector provided an additional easy-to-access tool during the most delicate phase of the fight against the pandemic.

We referred several research papers based on the topics Social distancing,Face mask detection,Object detection,Covid-19 trend analysis,Covid-19 impact on people,etc. From them we observed that, object detector models have been consistently developed with the likes of new algorithm yolov4 as compared to their predecessors like yolov3 and many others. Yolov3 algorithm resulted having lesser mean average precision by as much as 10% and the number of frames per second by 12% as compared to Yolov4 [14]. Yolov4 was improvement on Yolov3. All of the papers mentioned in references have used yolov3 and R-CNN algorithm but object detection has been more accurate with the enhancement in the algorithm with the introduction of yolov4 which helped us overcoming the drawbacks that the earlier papers had. Some authors used thermal rear camera in order to calculate distance, but this method for calculating distance was considered less accurate as compared to Euclidean distance calculation method that we used here.

In consecutive sections of our paper methodology of social distance detector and face mask detector is discussed and research is further concluded with acknowledgement and references.

METHODOLOGY

Yolo Framework

The YOLO framework (You Only Look Once) is used for object detection in a very unique style. It considers the whole image in a single time and calculates the bounding box coordinates and class probabilities for these boxes. YOLO is very fast when compared to others.[9]

We have used yolo-coco in our project which is basically trained for identifying human with a great accuracy.

Figure 1: Object Detection using Yolo

This is one of the best algorithms for object detection and has shown a comparatively similar performance to the R-CNN algorithms. Providing category in the configuration of coco. intersection over union. The evaluation metric used to measure the accuracy of object. [6]

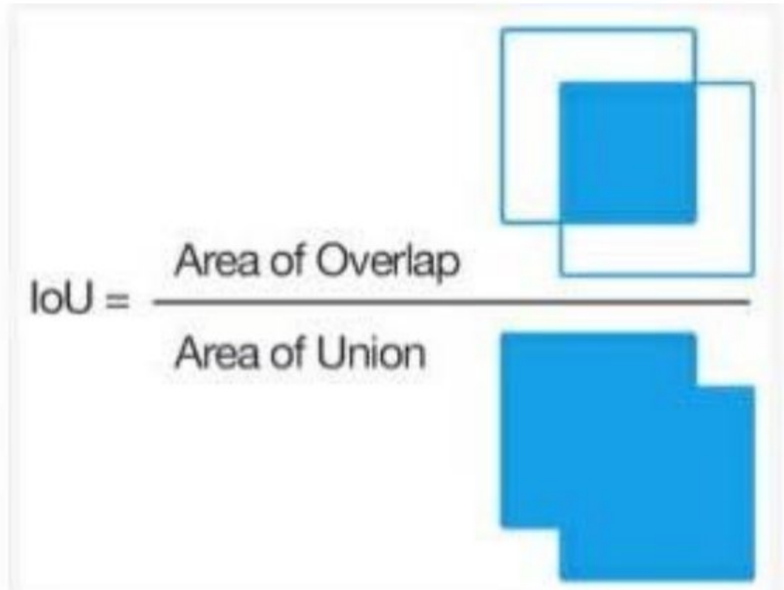


Figure 2: Intersection over union

Convolutional Neural Network

CNN is a deep learning algorithm that considers a input image and allots the learning weights and bias value for various classes in an image and tries to differentiate them from one to another [10]. The convolutional neural network has been made major advancement which can be implemented on an embedded system with a low-resolution input image and low very complexity. We have used the following layers in our project:

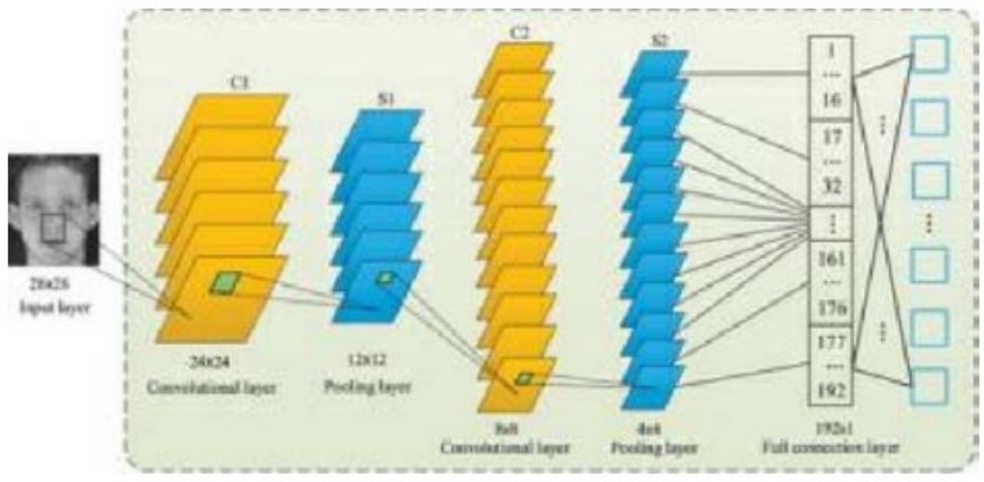


Figure3 : Convolutional neural network architecture

Social Distance Detector

In order to calculate the distance between two people, we need to calculate the Euclidean distance. Using Euclidean distance, the distance between the centroids can be calculated using :-

D = dist.cdist (centroids, centroids, metric = " euclidean").

Where D is a 2-D array as shown in the table below:- (table giving a small example)

	(x1,y1)	(x2,y2)	(x3,y3)
(x1,y1)	1	3	0
(x2,y2)	7	9	4
(x3,y3)	4	5	1

Where each element in the array is calculated according to the formula mentioned below:

$\text{Dist}((\text{xi},\text{yi}),(\text{xj},\text{yj})) = \sqrt{(xi - xj)^2 + (yi - yj)^2}$ where i,j = 1,…N detections

From the above equation, i and j represents rows and columns in 2-D array respectively. If this distance is less than a certain threshold value, the person is violating social distancing norms, or else he is not violating the norms.

Table 1: 2-D array containing the Euclidean distance between each point. Where each element in the array is calculated according to the formula mentioned below: From the above equation, i and j represents rows and columns in the 2-D array, respectively. If this distance is less than a certain threshold value, the person is violating social distancing.

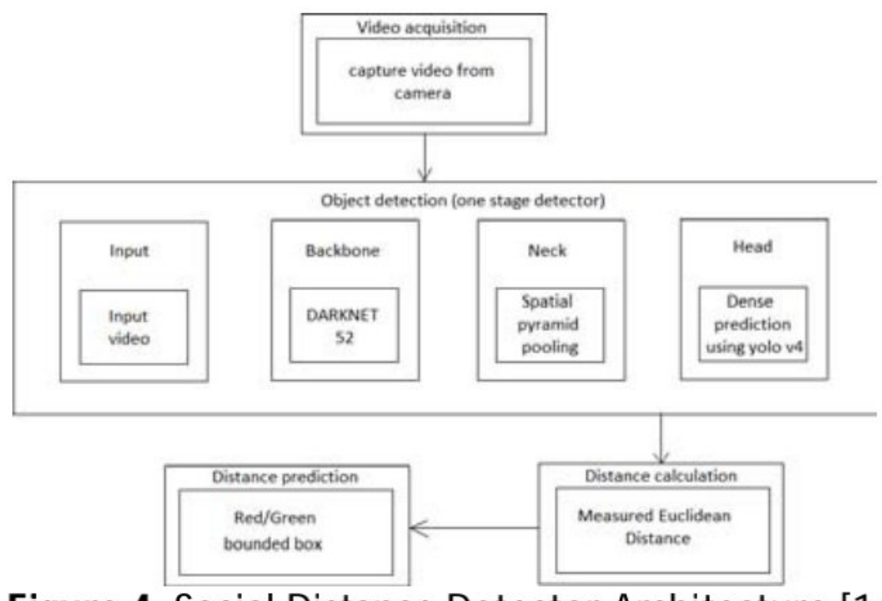


Figure 4: Social Distance Detector Architecture [14]

In social distance detector architecture there are total 4 steps,Video Acquisition,Object detection one stage detector), distance calculation, distance prediction which predicts the result of safety. In first step Camera will be capturing a video and acquisition means firstly capturing a video and then sending it to local storage. After that whole architecture of one stage detector comes into picture. A one stage detector consists of input video, backbone (Darknet 52), neck(spatial pyramid pooling) and head (yolov4). Input consists of input video which we give, backbone refers to the network that takes as an input your image and extract the feature map. The neck (spp) removes the fixed-size constraint of the network i.e CNN does not require a fixed-length output, which are then fed into fully connected layers, and finally the role of head in case of a one stage detector is to perform dense prediction .The dense prediction is the final prediction which is composed of a vector containing the coordinates of predicted bounding box(center, height,width) the confidence score of prediction and the label.[14]

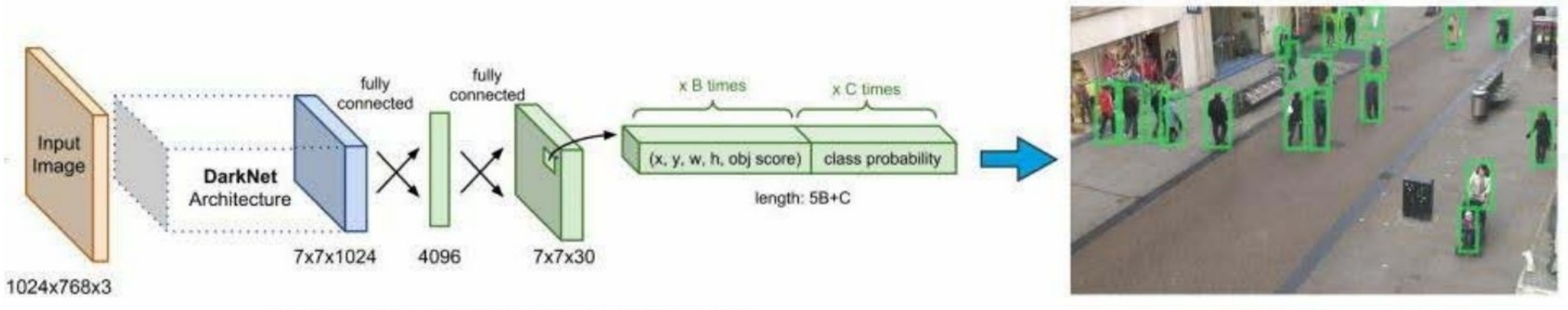


Figure 5: A Yolo based DNN & People detection [14]

The methodology works here is using algorithm used for object detection as well as object tracking. In this figure image layers which are responsible for taking the inputs that would be passed to further layers, input can be any image or video depending upon the cases. Along the input layer comes the darknet architecture, this is an open-source neural network for which framework is created and this framework features Yolo for object detection and object tracking.

Further, the figure consists of the flattered layer which is densely connected with the convolutional layer which is also densely connected to pass the data from each node to other nodes in this architecture. Similarly, this is posted to the output layer which gives 4-port values, there 4 ports describe the predicted value for the bounding box, denoted by x,y,w,h along with the object detection score plus the probability of the predicted class. This Yolo is part of the one-shot object detector family which is accurate and fast.

Face Mask Detector

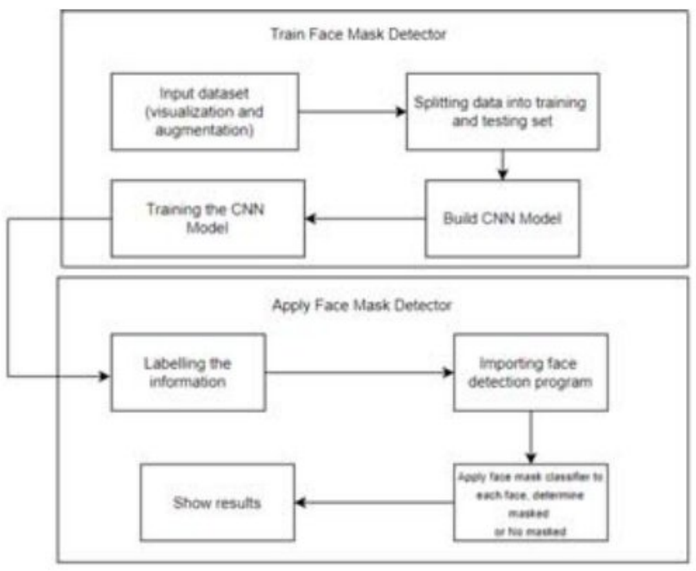


Figure 6: Face Mask Detector Architecture [14]

The face mask detection architecture is divided into two main parts:-

1 When mask detector will be trained

2) When mask detector will be applied

In this firstly, data visualization and data augmentation will take place. Visualization visualizes the total number of images in the dataset in both categories and augmentation will augment the dataset to include more number of images for our training. In next step data will be splitted into training set which will contain the images on which our model will be tested. Then a sequential CNN model is built. The next step is training the built CNN model and it is the main step where we fit our images in training set and test set to our sequential model, which will be built using Keras library. Then labeling of information will be done with the probabilities 0 and 1. After this we intend to use it to detect if we are wearing a face mask using our IP cameras[14].

The methodology works here is to build a CNN model using Tensor Flow to detect if you are wearing a face mask by using the webcam of your PC. The detection of face and telling whether a person is wearing a mask or not is done in 7 main steps, they include :-

1. Data Visualization :- In the first step, the total number of images in our dataset in both categories i.e person with and without masks are visualized.
2. Data Augmentation :- In the next step, we augment our dataset to include more number of images for our training. In this step of data augmentation, we rotate and flip each of the images in our dataset.
3. Building the model :- In the next step, we build our Sequential CNN model with various layers.
4. Training the CNN model :- This step is the main step where we fit our images in the training set and the test set to our Sequential model .
5. Labeling the information :- After building the model, we label two probabilities for our results,'0' as without mask and '1' as with mask.

6. Importing the face mask detection program :- After this, we intend to detect if we are wearing a face mask using our PC's webcam.
7. Detecting the faces with and without masks :- In the last step, we use the Open CV library to run an infinite loop to use our web camera in which we detect the face with and without masks[14].

RESULTS

Social Distance Detector

Below mentioned are three social distance detector features that we have shown in our application :- Image based detector, Video based detector and Web cam based detector.

Image Based Detector

Its main function is to detect the image using one-shot detector (Yolov4 algorithm), after detection put them in proper bounding boxes by filtering others classes and just selecting the person class, then calculating the euclidean distance between these bounding boxes and then finally showing the count of people violating the norm of social distancing.

Below are some figures depicting how actually the image based detector of our social detector works.

Figure 7 : Image based detector (a)

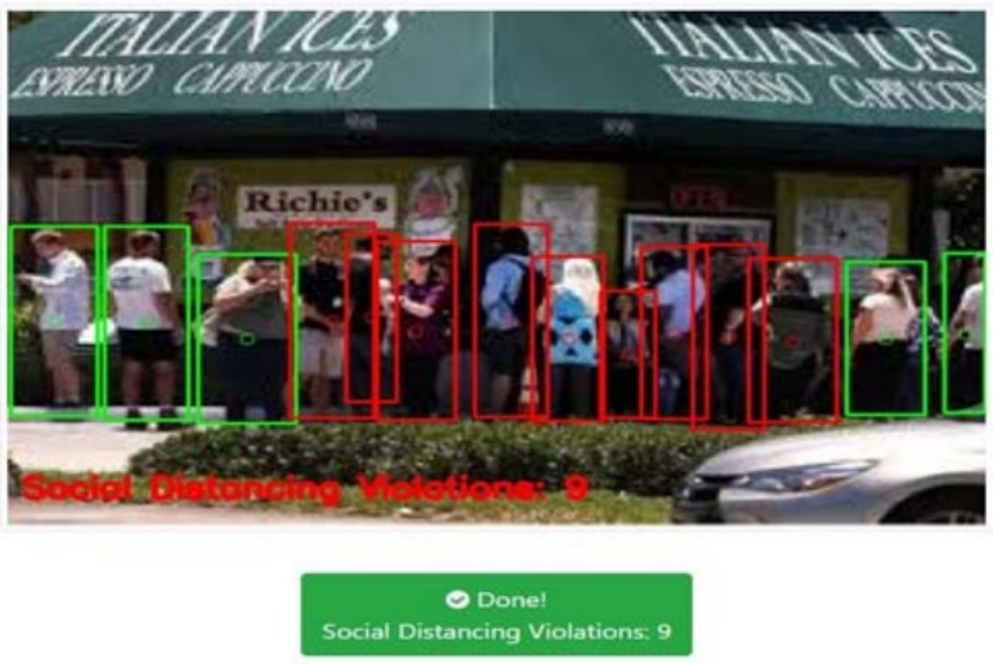


Figure 8 : Image based detector (b)

Video Based Detector

The video detector takes input a video and loads it frame by frame. In each of it's frame Yolov4 algorithm comes into picture and helps in proper detection of objects. With proper detection, it also assigns the bounding boxes over the classes. These bounding boxes are generally of two colors:- red and green. The red color indicates that the people detected in that particular frame are violating the social distancing norm, while the green color indicates that the people detected in that particular frame are following the social distancing norm.

Below are some figures depicting how actually the video based detector of our social detector works.

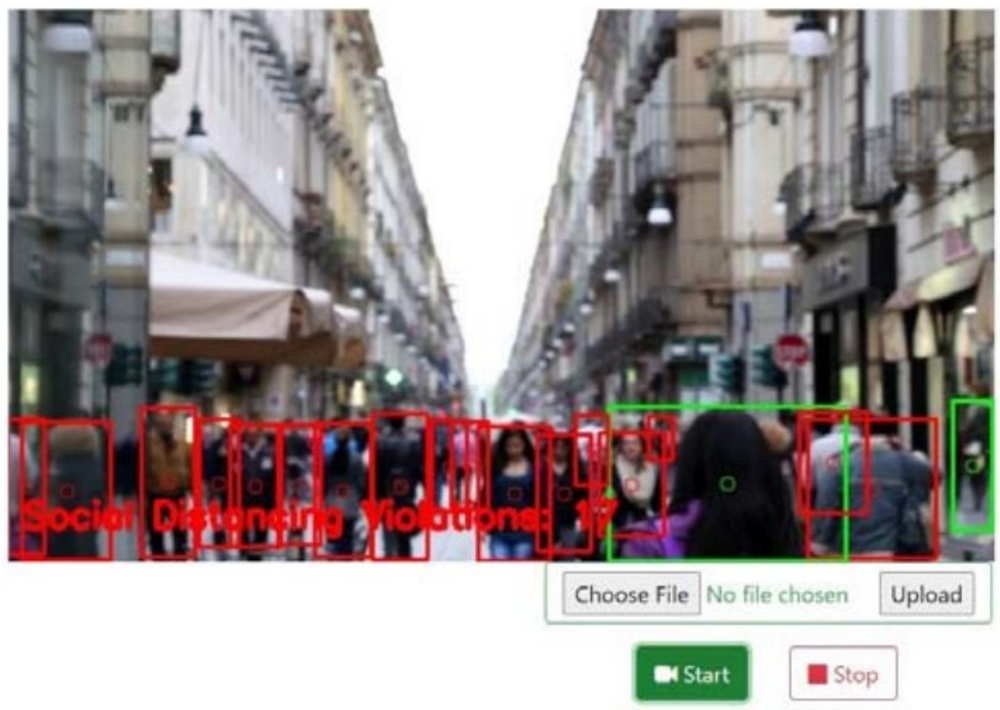


Figure 9 : Video based detector (a)

Figure 10 : Video based detector (b)

Web-cam Based Detector

The web-cam detector takes input through an IP camera of the respective personal computer and loads it frame by frame. In each of it's frame Yolov4 algorithm comes into picture and helps in proper detection of objects. With proper detection, it also

assigns the bounding boxes over the classes. The classes detected here are only the person class and rest of the classes are filtered out. These bounding boxes are generally of two colors :- red and green.

The red color indicates that the people detected in that particular frame is violating the social distancing norm, while the green color indicates that the people detected in that particular frame are following the social distancing norm.

Below are some figures depicting how actually the web-cam based detector of our social detector works.

Figure 11 : Web-cam based detector (a)

Figure 12 : Web-cam based detector (b)

Graph Comparing Different Algorithms

Referring different research papers and by implementing Social distance detector using Yolov4 it was observed that, Fast-RCNN which is known to use selective search method for generating region proposals, Faster-RCNN which is known to use use region proposal network method, Mask-RCNN which is known to add a branch to Faster R-CNN that outputs a binary mask that says whether or not a given pixel is part of an object had low precision, recall and accuracy value as compared to that of Yolov4. Yolov4 achieves the best detection results and thus it proves to be one of the fastest algorithm used in computer vision for object detection.

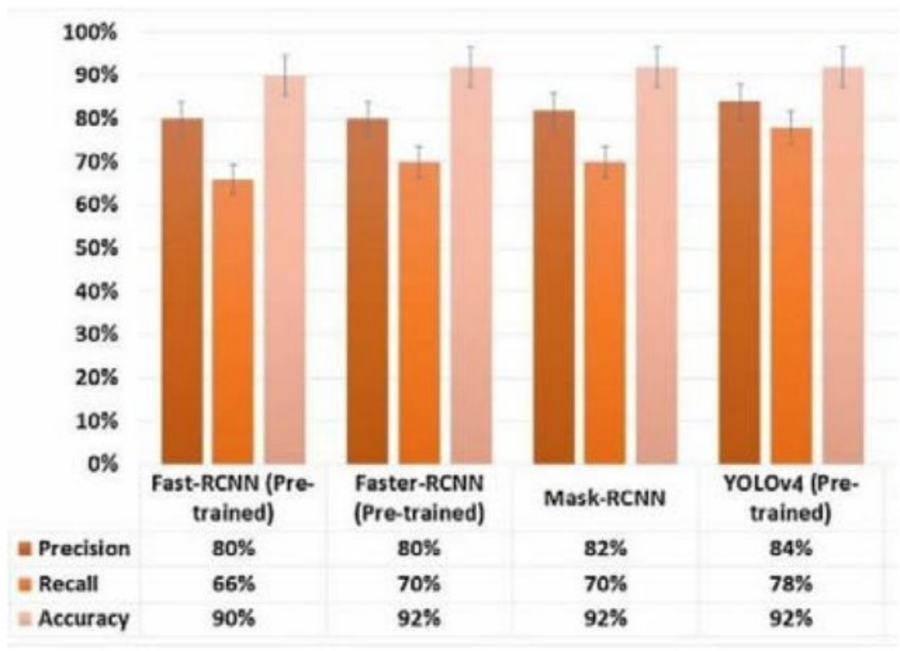


Figure 13 : Results with different algorithms

Face Mask Detector

Below mentioned are three Face mask detector features that we have shown in our application :- Image based detector, Video based detector and Web cam based detector.

Image-based detector

Its main function is to take that image and convert it into gray scale image and then pass onto the face cascade classifier for face recognition. After face recognition we save the face portion from that image. Then we will load our image and make it to the scale as we did in our trained mode. Then we use our trained cnn model for the prediction of the mask.

Below are some figures depicting how actually the image based detector of our face mask detector works.

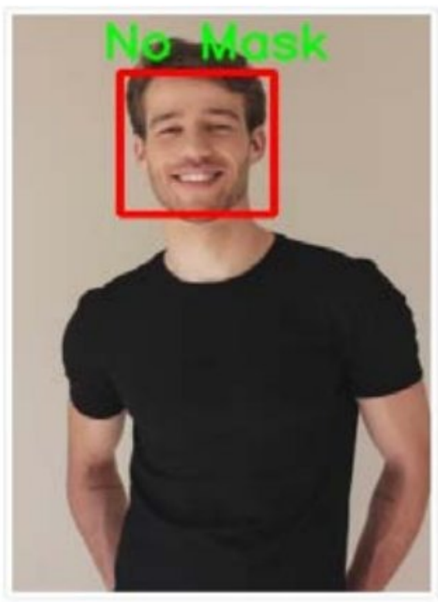


Figure 14 : Image based detector (a)

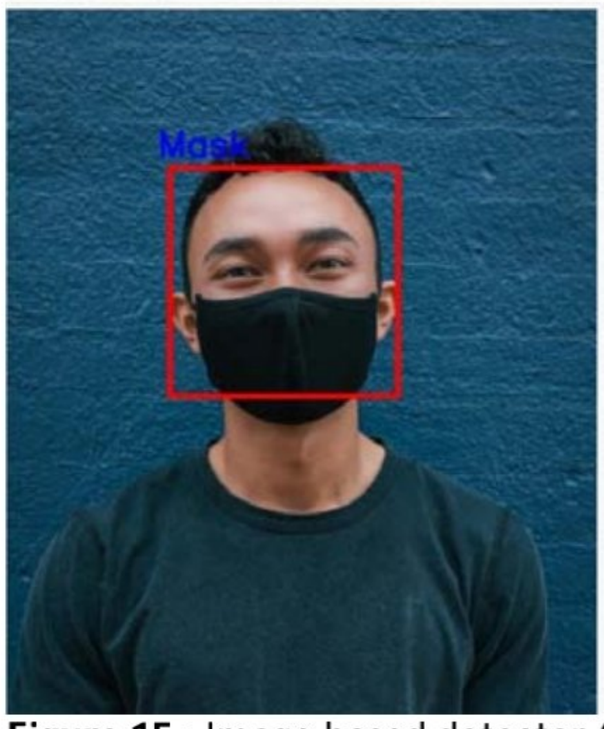


Figure 15 : Image based detector (a)

Video-Based Detector

In video based detector we take each frame from the video and treat it as a separate image. After the image gets loaded in our system we need to convert it into gray scale image for recognizing faces. We use harcascade classifier having different points for checking faces. After face recognition we save the face portion from that image. Then we will load our image and make it to the scale as we did in our trained mode. Then we use our trained cnn model for the prediction of the mask.

Below are some figures depicting how actually the video based detector of our face mask detector works

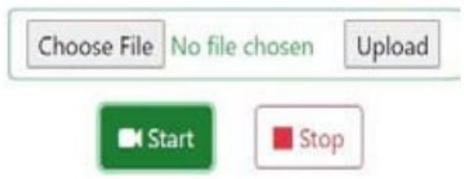


Figure 16 : Video based detector (a)

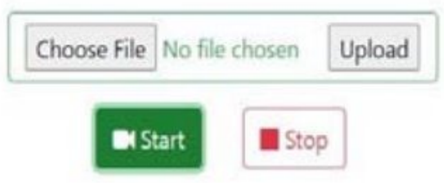


Figure 17 : Video based detector (b)

Web-Cam Based Detector

In the web cam based detector we use the live feed for face mask recognition. This works in the same principle as the video based detector system. Here as we capture each frame from the live feed it is sent for face recognition where that image is converted into gray scale image and then the face is recognized using cascade classifier and then our trained cnn model works on it to predict mask or not. If the model predicts it to be mask then it is shown in blue text otherwise green text is used to show no mask.

Below are some figures depicting how actually the web-cam based detector of our face mask detector works.

Figure 18 : Web-cam Based Detector (a)

Figure 19 : Web-cam Based Detector (b)

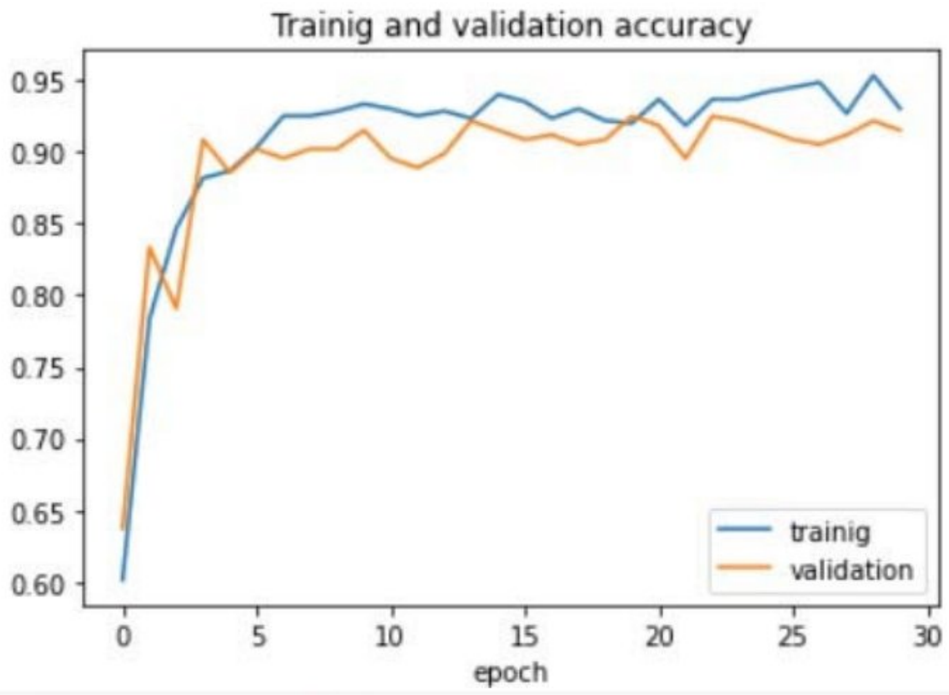
Training and Validation Accuracy and Loss

After training to know how the model performed during that phase we plot two graphs one for the training and validation loss, and the other one for the training and validation accuracy.

First we visualize the keys and then we get 'loss', 'accuracy', 'val_loss', 'val_accuracy'. In the first graph of training and validation loss, we first plot for loss then we plot for validation loss and then we plot for

training and validation named as legend , then we name the title as 'Training and validation loss'. The graph is labeled as 'epoch'.

In the graph, the loss went down which is good for the model. For every epoch, the loss is going down and here the data is well fitted by keeping training and validation loss to a minimum. The 2nd graph will have the same keys as used for loss . Only here we will check accuracy rather than loss. In this second graph the training accuracy is 92 percent. Over number of epochs, the accuracy has increased. The accuracy of validation has reached around 90 percent . As there is no huge gap in the graphs there is nothing wrong about the model. In the 1st graph for the loss, both the loss for training and validation are going in the same direction. If in case the training loss goes down continuously and the validation loss goes to a certain amount and then start to increase , then we can conclude our model is over-fitting but here there is no such case.




```
[ ] test_loss , test_acc = model.evaluate(test_generator)
    print('test loss:{} test acc:{}'.format(test_loss,test_acc))

4/4 [==============================] - 2s 363ms/step - loss: 0.1825 - accuracy: 0.9500
test loss:0.18252113461494446 test acc:0.949999988079071
```

Figure 20 : Training and Validation Accuracy

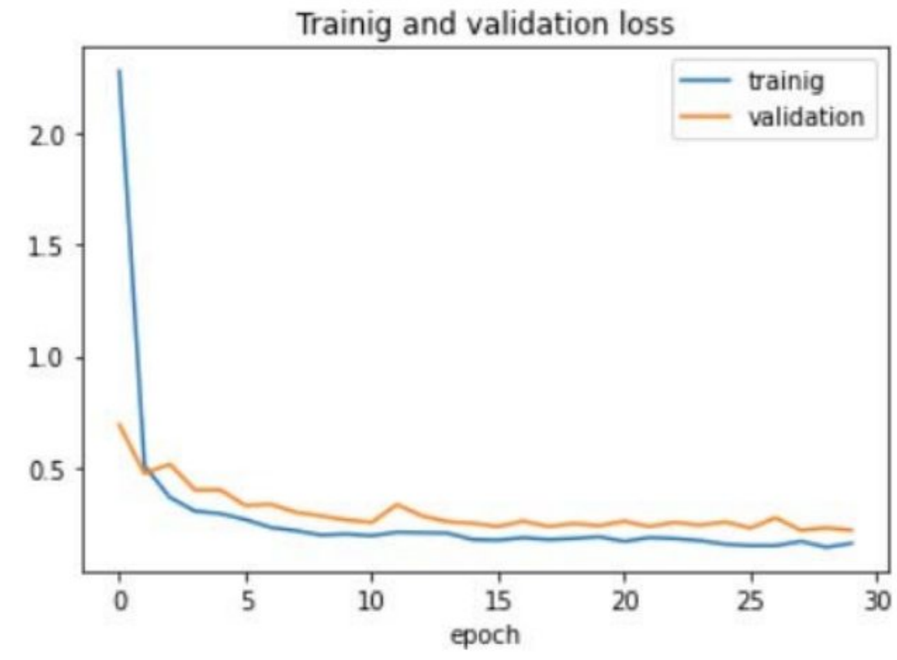


Figure 21 : Training and validation loss

Confusion Matrix

The confusion matrix compares the model's predicted values to the actual target values. Figure 22 shows the confusion matrix for face mask detection.

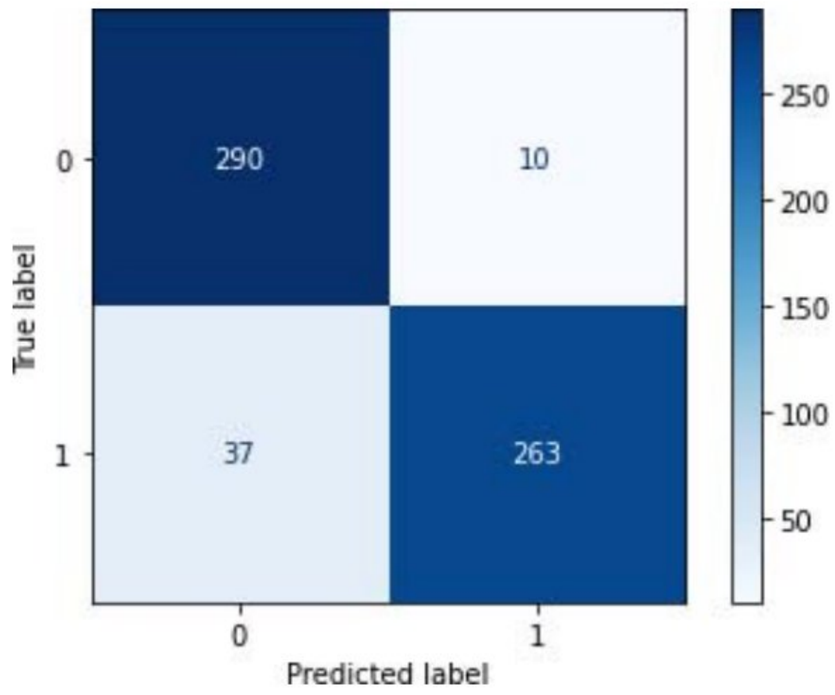


Figure 22: Confusion matrix

Additional Features

People Counter

People counting technology forms the basis for a range of high-tech solutions, including retail analytics, queue management and space utilization applications. To make effective business decisions, you need good intelligence - and that's what footfall counters provide.

For its application, we used Yolov4 object detection algorithm itself for detecting objects, and then just initialized a variable 'total' which consisted of total number of people detected by Yolo algorithm in the bounding boxes.

Conclusion

The study of this research was to understand the social distancing & face mask detection for the events of Covid-19. The social distancing was done using Yolov4 algorithm which is a machine learning module for calculating the distance among the different classes. Here the classes included different people on the streets while the face mask detection was done by creating a CNN model using Open CV which helps to identify whether people are wearing mask or not ; using red box- indicating a person is not wearing a mask and a green box - indicating person is wearing a mask.

The research work is helpful in new normal social and industrial life. The researchers could study prominent ML algorithms and suggested usability of CNN and Yolov4 as the simplest and appropriate

methods to build an application on. Due to strong library support and being apt language for machine learning, Python language can be used for implementation.

A web application was created using the above mentioned methodology that helped the end-user to check the whether the people are following the covid safety norms in order to prevent from infections .A social distance detector and a face mask detector will thus help the cause of reduce in number of covid-19 patients at some extend.

ACKNOWLEDGEMENT

Researchers would like to convey gratefulness to their project mentor, Prof. D.M Ujlambkar for her competent assistance and advice throughout the survey and implementation process of this project.

REFERENCES

[1] BhavikLimbasiya and ChinmayRaut ,2021, COVID-19 Social Distancing Detector using Machine Learning, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume: 08 Issue: 05 (May 2021)

[2] Savita S , Nikhil D K , Vishwesh M S, 2021, Social Distancing and Face Mask Detection From CCTV Camera, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10, Issue 08 (August 2021)

[3] Kumar G. and Shetty S. (2021). Application Development for Mask Detection and Social Distancing Violation Detection using Convolutional Neural Networks. In Proceedings of the 23rd International Conference on Enterprise Information Systems - Volume 1: ICEIS, ISBN 978-989-758-509-8, pages 760-767. DOI: 10.5220/0010483107600767

[4] Nair, Aishwarya&Potgantwar, Amol. (2018). Masked Face Detection using the Viola Jones Algorithm: A Progressive Approach for less Time Consumption. International Journal of Recent Contributions from Engineering, Science & IT (iJES). 6. 4. 10.3991/ijes.v6i4.9317.

[5] S. Li et al., "Multi-angle Head Pose Classification when Wearing the Mask for Face Recognition under the COVID- 19 Coronavirus Epidemic," 2020 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), 2020, pp. 1-5, doi: 10.1109/HPBDIS49115.2020.9130585.K. Elissa, ¯Title of paper if known, unpublished.

[6] G. Deore, R. Bodhula, V. Udpikar and V. More, "Study of masked face detection approach in video analytics," 2016 Conference on Advances in Signal Processing(CASP),2016,pp.196-200,doi: 10.1109/CASP.2016.7746164.

[7] A. Rahman, M. S. Hossain, N. A. Alrajeh and F. Alsolami, "Adversarial Examples—Security Threats to COVID-19 Deep Learning Systems in Medical IoT Devices," in IEEE Internet of Things Journal, vol. 8, no. 12, pp. 9603-9610, 15 June15, 2021, doi: 10.1109/JIOT.2020.3013710.

[8] Rahim, Adina &Maqbool, Ayesha &Rana, Tauseef. (2021). Monitoring social distancing under various low light conditions with deep learning and a single motionless time of flight camera. PLOS ONE. 16. e0247440. 10.1371/journal.pone.0247440.

[9] S. Srinivasan, R. Rujula Singh, R. R. Biradar and S. Revathi, "COVID-19 Monitoring System using Social Distancing and Face Mask Detection on Surveillance video datasets," 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), 2021, pp. 449-455, doi: 10.1109/ESCI50559. 2021. 9396783.

[10] S. S. Muley, AkshaySuryawanshi, ApurvaSonawane, PriteshPatil, Reuben Varghese,Real Time Face Mask Detection using Deep Learning ,GIS Science Journal,ISSN NO:1869-9391,DOI: 20.18001.GSJ. 2021.V8I 5.21.37063

[11] Sahana Srinivasan,Rujula Singh R,Ruchita R Biradkar, Revathi SA, Covid-19 Monitoring System using Social Distancing and face mask detection on surveillance video datasets,Department of computer science and engineering,R.V College of Engineering,Bangalore,Published:Mar 5-7,2021. jkkk

[12] K.Yagna, SAI SURYA, T.GEETHA RANI, B.K. TRIPATH, Social Distance Monitoring and face mask detection using deep learning, School of Information Technology and Engineering,VIT-Vellore- 632014, TN, India,Published:December 2021

[13] Mahdi Rezael and Mohsen Azrami, Deep Social:Social Distance Monitoring and Infection risk assessment in Covid-19 Pandemic,Faculty of Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University,Qazvin,Iran,Published:26 October 2020.

[14] Prof. D.M Ujlambkar, Vedashree Bhat, Atharva Jadhav, Siddhesh More, Prajwal Desai, Social distancing and Face Mask detector using Yolov4 and CNN, International Journal of All Research Education and Scientific Methods (IJARESM), ISSN: 2455-6211 Volume 10, Issue 3, March-2022, Impact Factor: 7.429.

Published by:

SMS LUCKNOW

School of Management Sciences
(SMS Institute of Technology)
CAMPUS: 19th KM Stone, Sultanpur Road,
Lucknow - 227125, U.P. (INDIA)
Ph.: 9792633331, 9506033170
CITY OFFICE: 2nd Floor, Sky-Hi Chambers, 5-Park Road,
Hazratganj, Lucknow - 226001, U.P. (INDIA)
Ph.: 0522-2238116, 2237273
Webiste : smslucknow.com