Covid-19 Social Distance Monitoring using YOLO Algorithm

Veerkumar Desai 191IT256
Information Technology
National Institute of Technology Karnataka
Surathkal, India 575025

Niranjan Hegde191IT235
Information Technology
National Institute of Technology Karnataka
Surathkal, India 575025

Ekansh Gaikwad 191IT217
Information Technology
National Institute of Technology Karnataka
Surathkal, India

***Abstract*—The main objective of this project is to provide easy to malls , high density roads to identify and take actions accordingly about the social distancing criteria. Through the last 2 years the world has been under immense pressure to counterattack Covid-19 and its effects. Creating a strong partnership between self preventive measures and medical treatment and vaccines a group can try overcoming the spread. Situations can be even critical in night gatherings with no one to monitor. Monitoring is required because you can't trust people to follow the guidelines given by the government even if it's done for their betterment. So using the recently developed YOLO v5 algorithm for the same increases the speed of detecting high risk connections.**

I. INTRODUCTION

Covid 19 could be a around the world misfortune posturing a major risk to mankind. In these times,social removing is the least demanding and exceptionally viable degree against the spreading of Covid 19.As per the WHO, to hone fitting social removing, people ought to keep 3ft or 1m separate between one another. So,we require a framework that can be utilized to monitor social distancing. This framework that we utilize could be a security basic framework since its disappointment or glitch can cause expanded spreading of Covid 19 in this way postures a risk to human life. The security basic framework here's protest location utilizing YOLOv5 to recognize individuals in a video. The show recognizes individuals bunches utilizing recognized bounding box information. The separations between the centroid of the bounding boxes of the people are calculated as the Euclidean distance. By calculating and handling the distance between the individuals they are classified into bunches based on the level of risk.

This framework is exceptionally accommodating in anticipating or abating down the spread of Covid 19 whereas immunization is still going on and modern compelling immunizations are however to be synthesized. This security basic framework can too offer assistance to track the individuals who might have had come in contact with a covid positive quiet and in this way playing an awfully critical part.

II. LITERATURE SURVEY

After the rise of the COVID-19 pandemic since late December 2019, Social distancing is very important to prevent the contagious virus transmission .The number of cases rises exceptionally, with one thousand to two thousand new cases reported per day in the first week of February 2020. Now also there are many cases in the world, So to prevent COVID-19 many AI developers made machine learning projects that tell us about social distance through CCTV and videos. People have a positive response over this social distance detection.

[1]This research paper is based on an intelligent surveillance system for people. Here tracking and social distancing is performed based on thermal images. Here convolutional neural network detector(CNN) is used and in object detection deep learning include other algorithms such as YOLO and YOLOv2 is used in this paper .The code or technique that use here achieved results for people detection and having good accuracy. An algorithm which is implemented on bounding boxes that show the difference in between safe and unsafe conditions, respectively, marking as green box and red box that the bounding box for detected persons. The proposed approach can be used in a distributed video surveillance system.

[2].This research paper proposes an efficient real-time deep learning to automate the process of social distancing with the help of object detection and tracking techniques,that are identified in the real-time with the help of bounding boxes. The bounding boxes that are made around the group of people that are close to each other the property computed with the help of pairwise vectorized approach. The number of violations are there when it's computed the number of groups formed and violation index terms are there when it's computed as the number of people /to the number of groups. The extensive trials were conducted with object detection models: The important file Faster RCNN,YOLO v3,SSD where YOLO v3 architecture is used for the efficient performance with balanced FPS and mAP score.

[3]This paper gives a YOLOv4-based show for social separation. The demonstration is taking a video/picture as input and making notices of SD infringement. In this paper they utilized YOLOv4 in this show that identifies pedestrian’s individuals in open places based on profound learning methods. The show employs a SD limit (SDTH) and an infringement file that take occurrence activity when the infringement happens and trigger a caution mindfulness activity promptly. protest discovery strategies, and SD checking

have moreover been given in this paper. The show is assumed to function persistently in swarmed places to screen individuals, in this way diminishing the effect of COVID-19 spread.

III. METHODOLOGY

- **Deep learning**

There are plenty of deep learning algorithms available for the problem we consider solving and every new method tries to modify and solve something new to the previous one. Conventional object detection algorithms use classifier based procedure, i.e sliding window procedure. In R-CNN , classifiers work on bounding boxes. These algorithms give very good performance, specifically Faster R-CNN has accuracy of around 73% maP, but this method turns down for real time object detection around 7 fps. This is where YOLO (You Only Look Once) comes into play. This is a real time object detection which is based on the Regression problem. YOLO has proven to give better results compared to previous methods. As of now there are 5 versions of YOLO with the 5th one recently released in 2021.
Average precision and FPS got noticeably increased from 10 to 12 percent in YOLO v4 when compared to its previous version. YOLO v4 presented a fast and accurate object detector that can be trained with a smaller batch size on a single GPU. YOLO v5 is merely based on YOLO v3.
Furthermore YOLO v5 turns out to be better than YOLO v4 in some of the areas discussed further. YOLO v5 is a PyTorch implementation from the original Darknet and .yaml files are used here for the same work as .cfg in YOLO v4 to just specify the different layers in the network. Some similarities such as CSP backbone and PA-NET neck remain the same. Results imply that YOLO v5 has higher inference speed compared to the rest of other detectors.

- **YOLO v5 Architecture**

Like any other object detector YOLO v5 has 3 parts -

1. Backbone - basically utilized for extraction of critical highlights from any given picture. Cross Stage Partial network is used in YOLO v5 for the same.

2. Neck - basically utilized to produce and include conglomeration. It makes a difference to demonstrate to generalize well on question scaling. It permits recognition of the same protest with diverse sizes and scales. PANet is utilized as neck in YOLO v5.

3. Head - is used for final detection to generate the final vectors with confidence probability and bounding box ,objectness score.

- **IOU and Non-Max Suppression**

These are techniques used in computer vision methods. Here the main motive is to classify the objects precisely by selecting one bounding box out of several overlapping ones. These functions are important because the classifier or the detector gives a large number of bounding boxes for any particular object and then these algorithms help to get the one with maximum probability and which distinguishes the object present.

IOU in mathematical terms can be expressed as -
IOU = (Target ∩ Prediction) / (Target U Prediction)

NMS Algorithm -

- Steps -
  1. Select prediction S with highest confidence and add it to the list.
  2. Comparing remaining S with the one in the list and substituting if found one greater confidence.
  3. Go To step 1 until all the predictions are checked.

- **Training Dataset**

For object detection and classification as people we have used the original pretrained YOLOv5 weights which are trained on COCO dataset.
MS COCO dataset is the dataset containing labeled images with around 80 classes which is used for training the model.
In the COCO dataset class list, we can see that the COCO dataset is heavily focused on people so we preferred it over other datasets.

- **Pixels and Real world distance, EUCLIDEAN distance**

Ed gives the euclidean distance between T1 and T2 in the form of pixels equivalent to real world unit distance D1D2.

Ed is calculated between centroids of T1 and T2

$$E_d = \sqrt{(x_{T2} - x_{T1})^2 + (y_{T2} - y_{T1})^2}$$

$$T_D = E_d$$

Constant k = D1D2 / Td

- **Processing and Calculation**

1. Cloning original repository of Ultralytics for using YOLO v5 facilities (Ultralytics are the developers of YOLO).

2. Training the model with COCO dataset and Yolo v5 weights.

3. Taking random videos displaying crowded places as input for monitoring social distancing.

4. Firstly we will use the videocapture function for accessing frames from the input video.

5. Then we will use the dnn.darknet function to load the frame and YOLO model.

6. YOLO v5 is further used for human detection and classification of other objects.

7. Non-Max suppression and IOU functions to obtain precise bounding boxes.

8. Detecting bounding boxes and generating vectors with confidence probabilities and their centroids.

9. Centroid plays the most important role while object identification and to determine the most accurate bounding box.

10. Computing pairwise distance between two centroids of two different bounding boxes and checking if they are less than N (threshold distance) apart.

11. Adding the computed distance to the distance matrix and further checking if they are violating the threshold distance and then adding accordingly to the violating and non violating set.

12. Set the color of low risk(non violating set) to green and high risk (violating set) to red.

13. The count of number of people violating the distance rules is also displayed in the output or even in display screen.

14. Once the frame is completed the next frame is grabbed and further processed until there are no more frames left to be processed.

Work flow chart of our model is given below.

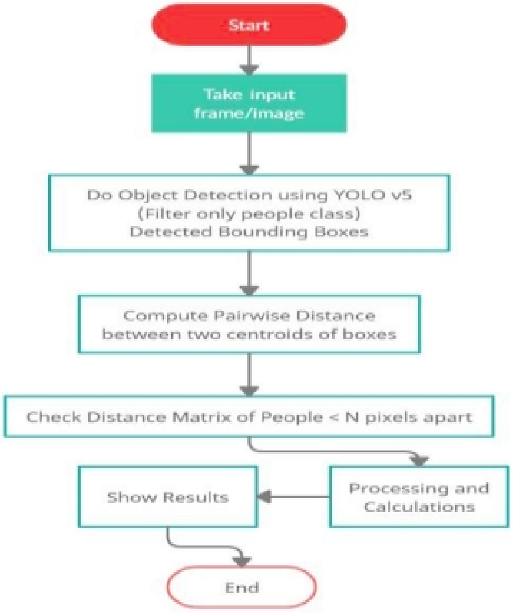
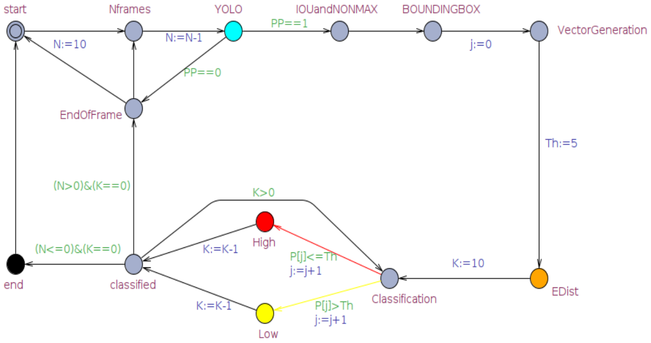

Start
Take input frame/image
Do Object Detection using YOLO v5 (Filter only people class) Detected Bounding Boxes
Compute Pairwise Distance between two centroids of boxes
Check Distance Matrix of People < N pixels apart
Processing and Calculations
Show Results
End

Fig. 1 Overview

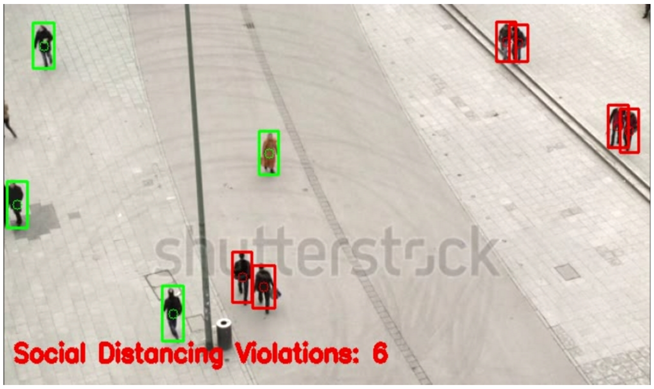
Below is the State wise representation of our model -



IV. EXPERIMENTAL RESULTS

Random snapshots of our output videos are shown below -

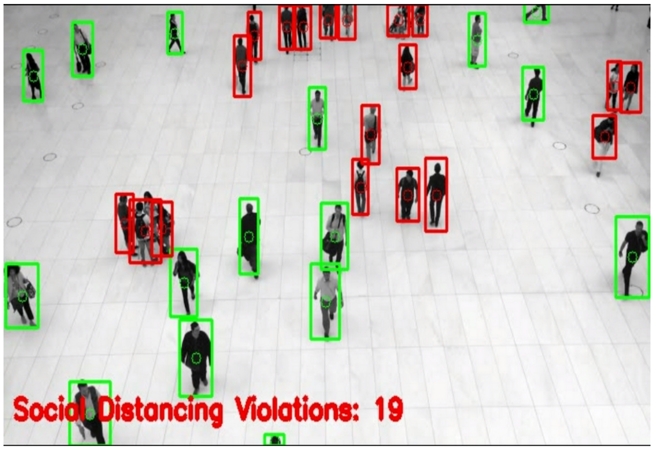
Snapshot 1 -



In above snapshot we can see that 3 pair of people are violating the rules of distancing.

i.e total 3 people are violating.

Snapshot 2



Here we can see that our model is showing number of violations as 19 but it should be 20 as our model did not precisely detect a person standing in group of 4 in the left part of snapshot 3. This arises because the person is covered by other one by around more than 50%.

- **Confusion Matrix**

We had an input video which had -21 people and 9 objects in it.

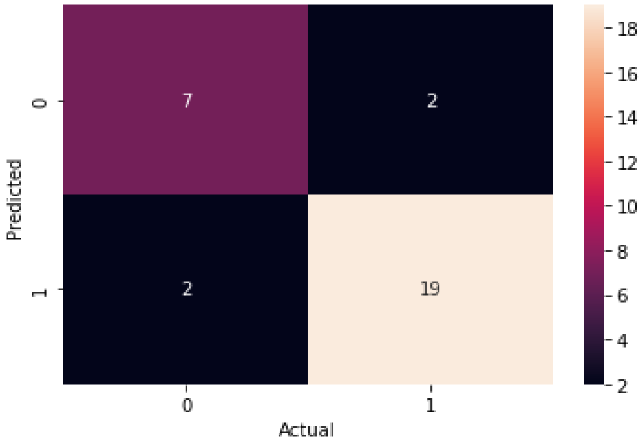
Our model predicted and classified the people with an accuracy of 70.

TP (True Positives) = 19

TN (True Negatives) = 2

FP (False Positives) = 2

FN (False Negatives) = 7



Accuracy = (19+2)/(19+2+2+7)

= 21/30

= 70%

Recall = 19/(19+7)

= 0.73

Precision =19/(19+2)

= 0.90

F1 score = (2.(0.73).(0.70))/(0.73+0.70)

= 0.7146

V. CONCLUSION

This report proposes a fast and efficient deep learning based framework that helps us to monitor the social distance by object detection in real time. We are using YOLOv5 model for the purpose of detecting people. Bounding boxes are generated and distance between the individuals are calculated as per our methodology. Compared to yolov4 ,yolov5 is faster. There are 4 different models in the repository and we are using YOLOv5s.

This project is very helpful in countries like India because of the huge population and lack of personnel to manually monitor social distancing and tracking. Accuracy can be improved for low light detection in future.

ACKNOWLEDGMENT

A Special Thanks to Professor Anand Kumar M for giving valuable opinions and guiding us with this project.

REFERENCES

[1] SOCIAL DISTANCING TRACKER USING YOLO V4
https://turkjphysiotherrehabil.org/pub/pdf/322/32-2-228.pdf

[2] YOLOv4-Based Monitoring Model for COVID-19 Social Distancing Control
https://link.springer.com/chapter/10.1007/978-981-16-2877-1_31

[3] Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques
https://arxiv.org/abs/2005.01385