Social Distancing Detection for COVID-19 using Deep Learning

Divyanshi Bhojak[1], Tarushi Jat[2]

Abstract— **In addressing the worldwide ongoing rampant the novel Covid-19 pandemic situation, the process of flattening the curve for coronavirus infected cases will be difficult if the citizens do not take action to prevent the spread of the virus. One of the most crucial practices is to ensure a safe distance between people around us in public. This project proposed a methodology to mitigate the spread of the coronavirus using social distance detection using deep learning techniques. The detection tool was developed to alert people to maintain a safe distance by evaluating a pre-recorded video feed of pedestrians walking in the road. We used an open-source pre-trained model Yolov3 and Yolov4 for object detection in the video frame used as input. The euclidean distance between any non compliant pair of pedestrians and red bounding boxes in the frame indicate the violation of precautionary distance. Furthermore, we have also implemented an alert mechanism in which mail will be sent if the number of violations exceeds the defined limit.**

Keywords: **COVID-19, Social distancing, pedestrian detection and tracking, distance estimation, deep learning.**

I. INTRODUCTION

With the generation of the novel Coronavirus disease, COVID-19 first reported in the last month of 2019, and within a few months the rampant virus became a global epidemic affecting millions of people. In May 2020 The World Health Organisation (WHO) declared the situation as pandemic. The statistics by WHO on 8th October 2020 confirm 36 million infected people and a scary number of 1,056,000 deathsin 200 countries.

In India, The Ministry of Health and Family Welfare has recommended to the citizens to maintain a safe distance of at least 6 feet among themselves in public and crowded places. Social distancing, as shown in Fig. 1 enforced in public places in India to prevent the proliferation of the disease, by minimising the proximity of human physical contacts in covered or crowded public places to stop the widespread accumulation of the infection risk. In the month of March 2021, India witnessed a second wave of coronavirus took devastating toll by cases count crossed 1 lakh setting off the most explosive phase of the pandemic during which the daily count of new infections reached as high as 4.14 lakh on May 6 2021. Recent research has confirmed that infected people with mild or no symptoms may also be carriers of the novel coronavirus infection. Social distancing is essential, particularly for those people who are at higher risk of serious illness such people with age more than 80 years and infants from COVID-19. By decreasing the risk of virus transmission from an infected person to a healthy one, the virus spread and disease severity can be significantly reduced. Therefore, it is important all individuals maintain controlling behaviours and observe social distancing. Many research works have proved social-distancing as an effective non-pharmacological approach and an important inhibitor for limiting the transmission of contagious diseases such as H1N1, SARS, and COVID-19.

Fig. 1: Social Distancing Scenario in India.

The organization of this report is as follows. Section II describes some existing notable related work in this field. Section III describes the Proposed Methodology of the project. Section IV describes the proposed experimental results. Section V summarizes the proposed work and results.

II. LITERATURE SURVEY

Since the end of the year 2019, when the first case of covid19 was reported in Wuhan, China, many research has been done in various aspects to take precautions against the virus. Punn NS, Sonbhadra SK and Agarwal S in [1] proposed a deep learning based framework for automating the task of monitoring social distancing using surveillance video. Their framework utilizes the YOLO v3 object detection model to segregate humans from the background and Deepsort approach to track the identified people with the help of bounding boxes and assigned IDs. Mahdi Rezaei and Mohsen Azarmi in [2] proposed DNNmodel in combination with an adapted inverse perspective mapping (IPM) technique and SORT tracking algorithm leads to a robust people detection and social distancing monitoring. Their model has been trained against two most comprehensive datasets by the time of the research—the Microsoft Common Objects in Context (MS COCO) and Google Open Image datasets. Imran Ahmed, Misbah Ahmad, Joel J P C Rodrigues, Gwanggil Jeon and Sadia Din in [3] proposed the framework that uses the YOLOv3 object recognition but this model does not perform well with input videos having challenging outdoor environments. Afiq Harith Ahamad, Norliza Zaini and Mohd Fuad Abdul Latip in [4] proposed the model that helps detecting people in areas of interest using "MobileNet Single Shot Multibox Detector (SSD)" object tracking model and "OpenCV" library for image processing. Their results show that the object detection model used for detecting persons is having the difficulty in detecting people correctly in the outdoor environment and difficult scenes with distant scenes. Inspired from the ideas, we present a computer vision technique for social distancing detection using YOLO version 3 and version 4 models. We have utilized the concept of euclidean distance to calculate how far the two people are from each other and the application will highlight if the people are unsafe.

III. Methodology

The proposed methodology is a three stage model involving object detection that is detecting pedestrians in input video frame, then monitoring distance and estimation that is monitoring social distance between the pedestrians then distance estimation and lastly alert mechanisms through email when distance is violated. In this project, the social distance detection and estimation are computed using pre-trained models and proposed methodology Fig. 2 is discussed in the subsequent sections.

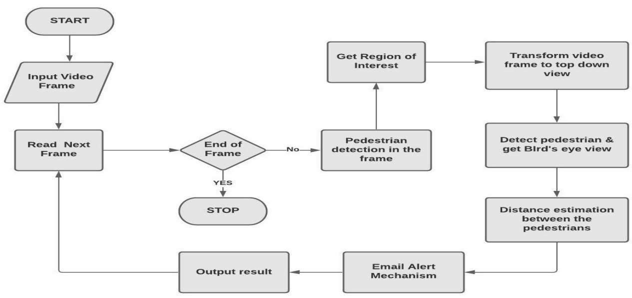

Fig. 2: Workflow Architecture of Social Distancing Detection

A. Object Detection in Video

Object detection is one of the challenging tasks involving a method that is used to recognize and detect different objects present in an image or video and label them to classify these. In the early phase object detection was one of the main technical challenges. But, after 2014, with the increase in technical advancements using deep learning tasks were accomplished and solved the problem. Object detection is a process of finding all the possible instances of real-world objects, such as human faces, flowers, cars, etc. in images or videos, in real-time with utmost accuracy. In our project, from the input pre-recorded video frame will detect pedestrians in the frame with their unique localisation bounding boxes.

B. Pretrained Models

In this project to perform experiments used pre-trained CNN-based models. We used the pre-trained model YOLO that is You Only Look Once, a family of CNN models that are a series of end-to-end deep learning models designed for fast object detection, developed by Joseph Redmon, et al. and first described in 2015. The approach involves a single deep convolutional neural network that splits the input into a grid of cells and each cell directly predicts a bounding box and object classification. In this project we used a comparative analysis on YOLOv3 and YOLOv4. The illustration of the YOLO model for object detector pipeline is shown in Fig. 3.

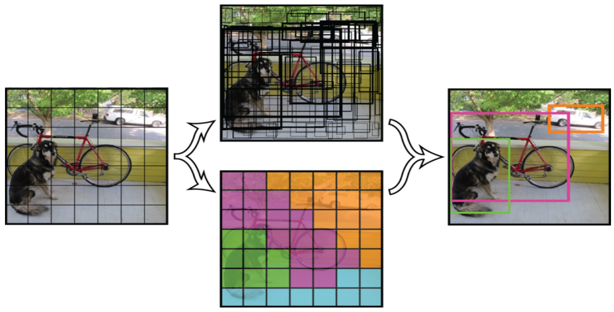

Fig. 3: The illustration of YOLO model for object detector pipeline

1. YOLOv3:
YOLO version 3 most salient feature is that it makes detection at three different scales which are precisely given by downsampling the dimensions of the input image by 32, 16 and 8 respectively.. YOLO is a fully convolutional network and its eventual output is generated by applying a 1 x 1 kernel on a feature map. In YOLO v3, the detection is done by applying 1 x 1 detection kernels on feature maps of three different sizes at three different places in the network. YOLOv3 has a Darknet53 backbone and has class prediction through logistic regression.

2. YOLOv4:
YOLO version 4 composed of CSPDarknet53 as a backbone, spatial pyramid pooling additional module, PANet path-aggregation neck and YOLOv3 head. YOLO version 4 is an important improvement over version3 implementation of a new architecture in the Backbone. YOLOv4 adds a SPP block after CSPDarknet53 to increase the receptive field and separate out the most important features from the backbone.

C. People Class Detection

The next phase is people class detection and unique ID assignment for each individual. From the inputted dataset and pre-recorded video our aim was to detect only pedestrians in the frame other object classes are ignored in this application. Consequently, after unique pedestrian detection in the frame bounding boxes will be created around each pedestrian in Fig 4. Hence, the bounding box best fits for each detected pedestrian can be drawn in the image, and this data of detected pedestrians will later used for distance estimation.

D. Camera Perspective Transformation

For camera setup, the camera is captured at fixed angle as the video frame, and the video frame was treated as a perspective view in Fig 5. and transformed into a two-dimensional top-down view for more accurate estimation of

Fig. 4: Detection of only Humans in Video

distance measurement. In the implementation, we assumed that all the pedestrians are walking on the same flat plane. The region of interest (ROI) of an image focused on the pedestrian walking street was transformed into a top-down 2D view in Fig. 6 that contains 480×480 pixels. Then we selected the four coordinate points from the frame to transform into the top-down view mapping them to the corners of a rectangle in the 2D image view. This perspective transformation is important to calculate the real world distances in the frame. Then getting a bird's eye view by localizing pedestrian video from a top-down perspective in the region of interest. Based on this bird's eye view in Fig 7. distance will be estimated. This bird's eye view then has the property of points being equidistant no matter where they are. All it needs is a multiplier that maps the distance between two points in pixels to distance.

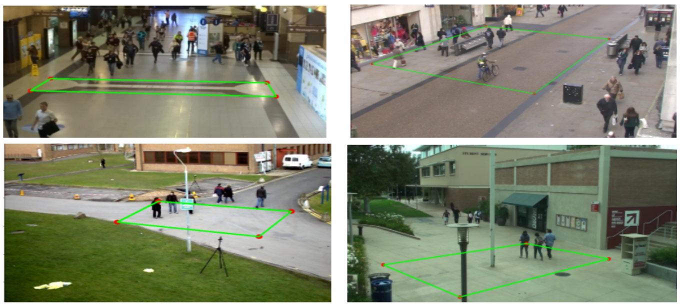

Fig. 5: Camera Perspective Transformation of sample videos

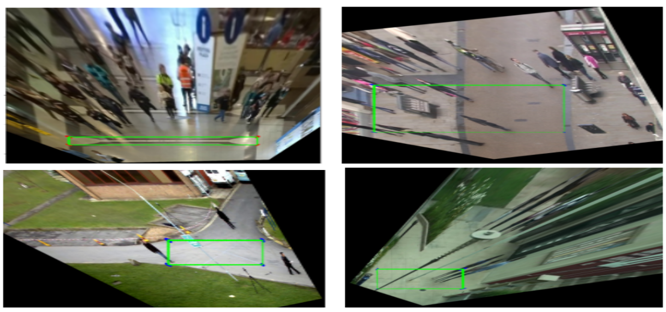

Fig. 6: Top-down transformation of sample videos

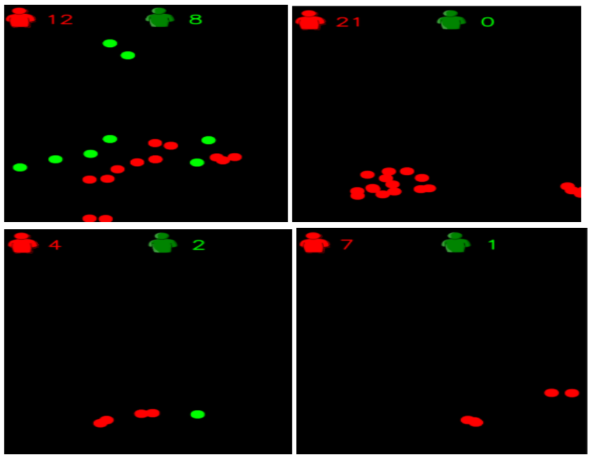

Fig. 7: Bird's eye view transformation of sample videos

E. Distance Estimation between Pedestrians

After the perspective transformation, the location for each pedestrian can be estimated to correspond to the number of pixels in earlier generated bird's eye view. In this phase, the bounding box position (x,y,w,h) where x and y are two-dimension coordinates, w and h are the width and height of the frame view. For each pedestrian, the position in the top-down view is estimated based on the bottom-center point of the bounding box. Then distance is estimated between two non-complaint pairs of pedestrians using euclidean distance. Given the position of two pedestrians in an image as (x1, y1) and (x2, y2) respectively, the distance between the two pedestrians can be computed as given in Fig. 8.

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Fig. 8: Euclidean Distance Formula

F. Bounding Box generation

After the distance estimation between the pedestrians. Then red and green color bounding boxes will be generated based on the distance estimation. Depending on the preset minimum precautionary distance if the distance between any two non complaint pairs of pedestrians is less than acceptable distance will be indicated with a red bounding box for those pedestrians that serve as precautionary warnings. And when the maintain the safe and acceptable distance then bounding box will be depicted by green color.

G. Email Alert Mechanism

After all the above phases, to create more awareness and to immediately prevent the violations made between the pedestrians. So, if the distance between the pedestrians is below the acceptable distance and the bounding box color is red then we added a new mechanism to send an alert message through Email. After detecting violation of social distancing in video frame alert message through Email will be sent stating the text message "Social Distancing Violations

Exceeded! ” that is, the number of violations exceeds the maximum violations limit which was earlier set . in Fig 9.

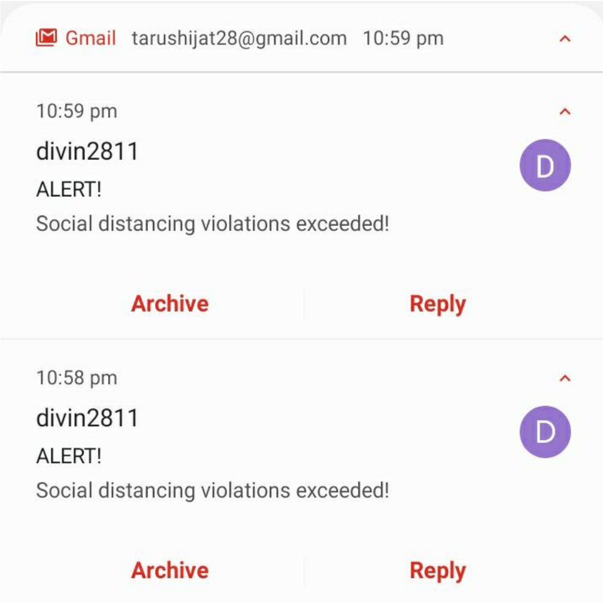

Fig. 9: Email Alert Mechanism

IV. Experimental Analysis

For video object detection, we have used the two versions of YOLO pretrained models, YOLO-V3 and YOLOV4. We have used these models which are pretrained on the COCO Image dataset, which is a large scale object detection, segmentation and captioning dataset. This dataset contains 80 object categories and 91 stuff categories with around 1.5 million object instances and 330k images. We have experimented on four input videos which show the pedestrians walking on streets, malls, outside a university and in a town. To calculate distance between pedestrians in each of the videos, the perspective view of the video frame is transformed into a top down view for more accurate calculation of distance between each detected pedestrian. Each of the detected pedestrians is then represented by a dot. Thereafter we calculate the distance between each of the detected pedestrians which are now represented by dots. To identify which of the detected pedestrians are violating the distance criteria, we have defined a threshold value. If any of two pedestrians are having a distance less than the defined threshold then they will be identified with the help of red boxes, otherwise pedestrians will be identified with green boxes. To test the performance of the social distancing detection model, implemented with yolov3 and yolov4, on the input video with unknown ground truth values, confidence score of each of the detected objects in the video frames is being calculated. To get the overall confidence score of each of the four input videos, we took the average of all confidence scores. Table I shows the confidence score of each of the four input videos in case of yolov3 and yolov4.

In each of the input videos, along with the detection of pedestrians violating the social distancing criteria with the help of red and green boxes, the total number of people who are safe and the total number of people who are unsafe is also calculated. Fig. 10 shows the output video snapshot of the “MOT20” input video implemented with yolov4 and yolov3 models respectively. Fig. 11 shows the output video snapshot of the “TownCentre” input video implemented with yolov4 and yolov3 models respectively. Fig. 12 shows the output video snapshot of the “PedestrainWalking” input video implemented with yolov4 and yolov3 models respectively. Fig. 13 shows the output video snapshot of the “StudentVideo” input video implemented with yolov4 and yolov3 models respectively.

TABLE I: Confidence Scores

Video Name	YoloV4	YoloV3
MOT20	0.8440056110991888	0.7503122930372382
TownCentre	0.8816814148867955	0.7883997548728089
PedestrianWalking	0.9709215107691926	0.8729871214650065
StudentVideo	0.8703143704892438	0.8085678247857485

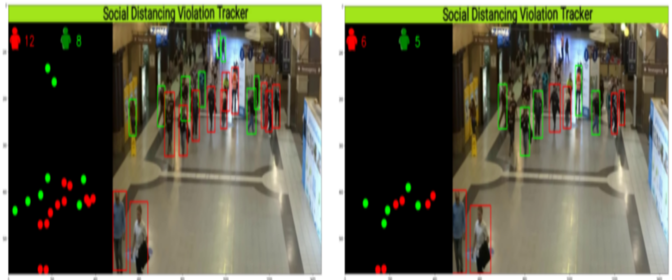

Fig. 10: MOT20 output snapshot with YoloV4 and YoloV3

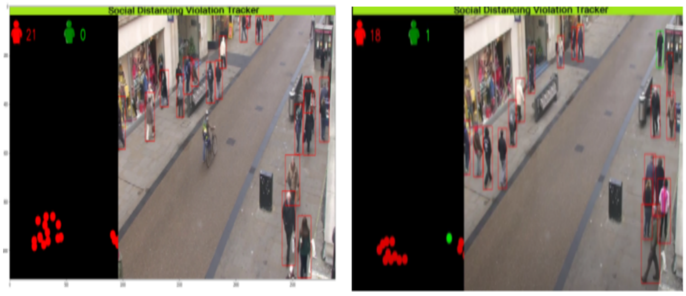

Fig. 11: TownCentre output snapshot with YoloV4 and YoloV3

Conclusion

In this project work, a methodology of social distancing detection for covid19 using deep learning model is proposed. With the help of computer vision technology, the distance between people can be estimated and any pair of pedestrians violating the distance criteria will be covered with the help of red boxes. The proposed methodology is validated using four different input videos showing waking pedestrians. The visualization results showed that the proposed method is capable of determining the social distance measure between people

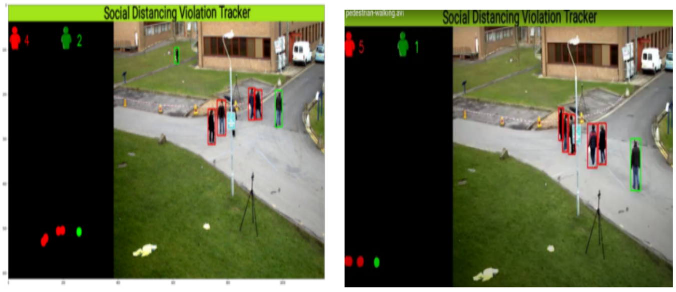

Fig. 12: PedestrianWalking output snapshot with YoloV4 and YoloV3

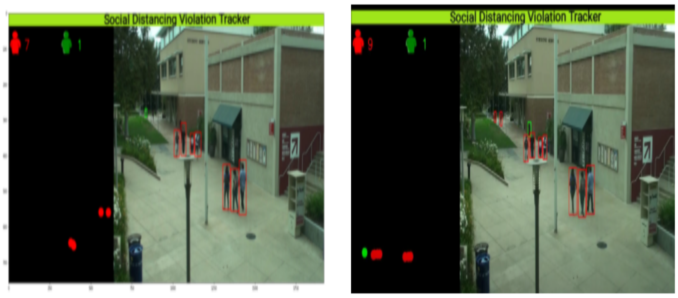

Fig. 13: StudentVideo output snapshot with YoloV4 and YoloV3

and therefore this model can be implemented in real life environments such as universities, public transport stations and many other places as a future work. Furthermore, this work can be improved by implementing many other criterias that are important, such as human body temperature detection and mask detection, to fight against covid19 pandemic.

REFERENCES

[1] Punn NS, Sonbhadra SK and Agarwal S, "Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques", arXiv preprint arXiv:2005.01385, April 2021.
[2] Mahdi Rezaei and Mohsen Azarmi, "DeepSOCIAL: Social Distancing Monitoring and Infection Risk Assessment in COVID-19 Pandemic", arXiv preprint arXiv:2008.11672, 2020.
[3] Imran Ahmed, Misbah Ahmad, Joel J P C Rodrigues, Gwanggil Jeon and Sadia Din, "A deep learning-based social distance monitoring framework for COVID-19", https://doi.org/10.1016/j.scs.2020.102571, 2021.
[4] Afiq Harith Ahamad, Norliza Zaini and Mohd Fuad Abdul Latip, "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.
[5] Lakshmisri Surya, Ravi Teja Yarlagadda, "AI Economical Smart Device to Identify COVID-19 Pandemic, and Alert on Social Distancing Who Measures", International Journal of Creative Research Thoughts, Feb 2021.
[6] Sizhen Bian, Et al, "A wearable magnetic field based proximity sensing system for monitoring COVID-19 social distancing", Proceedings of the 2020 International Symposium on Wearable Computers. SEP 2021.
[7] Dongfang Yang, Ekim Yurtsever, Et al, "A Vision-based Social Distancing and Critical Density Detection System forCOVID-19", published in JUL 2020.
[8] AbelBrodeur, NikolaiCook, Et al, "On the effects of COVID-19 safer-at-home policies on social distancing, car crashes and pollution", Journal of Environmental Economics and Management, March 2021.
[9] Adarsh Jagan Sathyamoorthy, Utsav Patel, Et al, "COVID-Robot: Monitoring Social Distancing Constraints in Crowded Scenarios", published in 21 Aug 2020.