Indian Institute of Information Technology Allahabad

Summer Project Report On

An AI Enabled Social Distancing Detection Tool

Under the Guidance of :
Dr. Vrijendra Singh, HOD, IT
IIIT Allahabad

Submitted by :
Manav Vallecha : IRM2017007
Nayan Agarwal : IWM2017501
Shubham : IWM2017004
Aman Gupta : IWM2017006

Contents

1 ABSTRACT **2**

2 INTRODUCTION **2**

3 MOTIVATION **4**

4 PROBLEM DEFINITION **5**

5 LITERATURE REVIEW **6**

6 PROPOSED METHODOLOGY **6**
6.1 Object Detection 7
6.2 Why we Chose YOLO Algorithm for Object Detection? 7
6.3 Non Max Suppression 8

7 REQUIREMENTS **11**
7.1 Hardware Requirements 11
7.2 Software Requirements 11

8 IMPLEMENTATION PLAN **11**

9 RESULTS **12**

10 Comparative Study **14**

11 FUTURE WORK **14**

12 CONCLUSION **15**

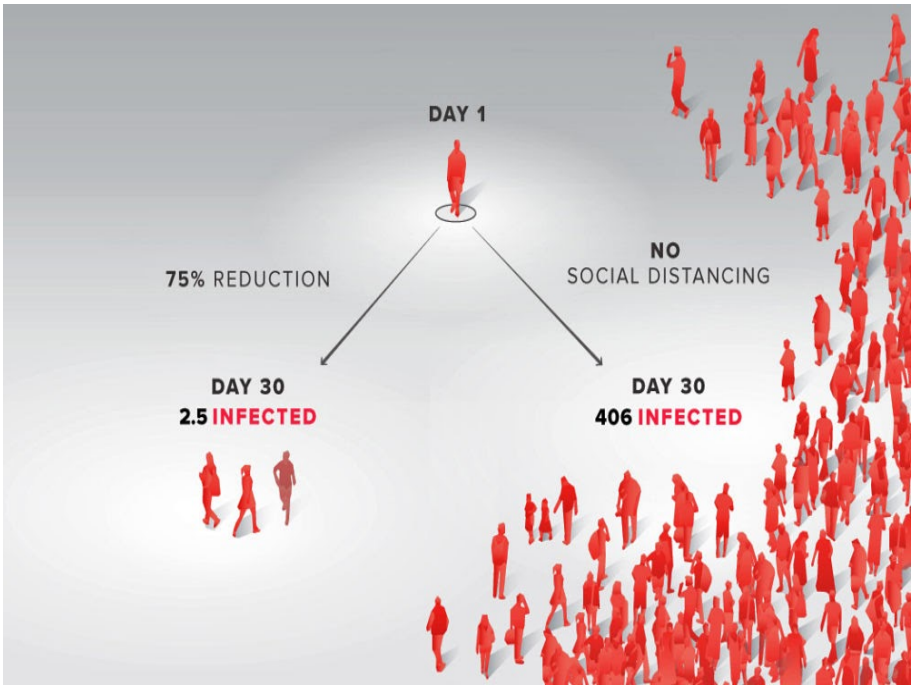
13 REFERENCES **15**

1 ABSTRACT

Considering the enormous change that the world is facing, because of the COVID-19, we present the project on an AI based tool to examine the social distancing for roads, parks, workplaces, and even other crowded and jam-packed spots like shopping centers, train stations, etc as social Distancing has proven to be an effective method to curb the spread of CoronaVirus.To get the results in real time we made use of the Yolo algorithm which is a popular object detection algorithm that gives pretty accurate results.

2 INTRODUCTION

Social distancing is a method used to control the spread of contagious diseases.As the name suggests, social distancing implies that people should physically distance themselves from one another, reducing close contact, and thereby reducing the spread of a contagious disease (such as Coronavirus) . The below picture highlights the importance of social distancing in today's world.



Social Distancing is the key element in our fight against COVID-19 . It may remain an essential part of our lives for months and reshape how we interact with the outside world forever. In the present scenario , we can see many people violating

the social distancing norms and not following the rules. . The current way of social distancing at stores is not a viable long term solution, and we need to come up with a better way to restore the shopping experience while ensuring the safety of everyone. We could have a solution build in order to ensure that social distancing norms are being followed in busy places.With the help of Artificial Intelligence (AI), the same technology that is the backbone of self-driving Tesla and Netflix recommendations, combined with edge computing, the technology that is reshaping the Internet of Things (IoT), we can practice social distancing with minimal disruption to our daily lives . Using modern deep learning ,image processing techniques and using modern technologies we have tried to come up with a simple solution that can monitor if social distancing norms are being followed at a certain place. We use Deep Learning Models , and image processing libraries to achieve this objective. Computer vision along with deep learning, machine learning image processing provide effective solutions to measure social distancing among humans across the moving frames .
The project deals with three main points :-

- **Object Detection** :
  In order to identify if the social distancing norms are being followed or not , first our algorithm needs to detect the persons in an image or a video feed. We use YOLO (You Only Look Once) real time object detection algorithm , which is one of the most effective object detection algorithms that also encompasses many of the most innovative ideas coming out of the computer vision research community. Object detection is a critical capability of autonomous vehicle technology. It's an area of computer vision that's exploding and working so much better than just a few years ago. Other than YOLO there are 2 other main object Detection algorithms namely R-CNN and their variants, including the original R-CNN, Fast R- CNN, and Faster R-CNN , Single Shot Detector (SSDs)

- **Object Tracking**: Only require the object detection phase once (i.e., when the object is initially detected) . Object tracking is extremely fast — much faster than running the actual object detector itself . Be able to handle when the tracked object "disappears" or moves outside the boundaries of the video frame..

- **Finding Distance**: In order to determine when the social distancing norms are violated we need to compute the distance between two detections (i.e. two persons) in the image or video frame . We also need to determine a certain threshold based on whether our algorithm can give proper output if it's a violation of social distancing or not.

To visualize it, we used Red Green color bounding box across the pedestrians in

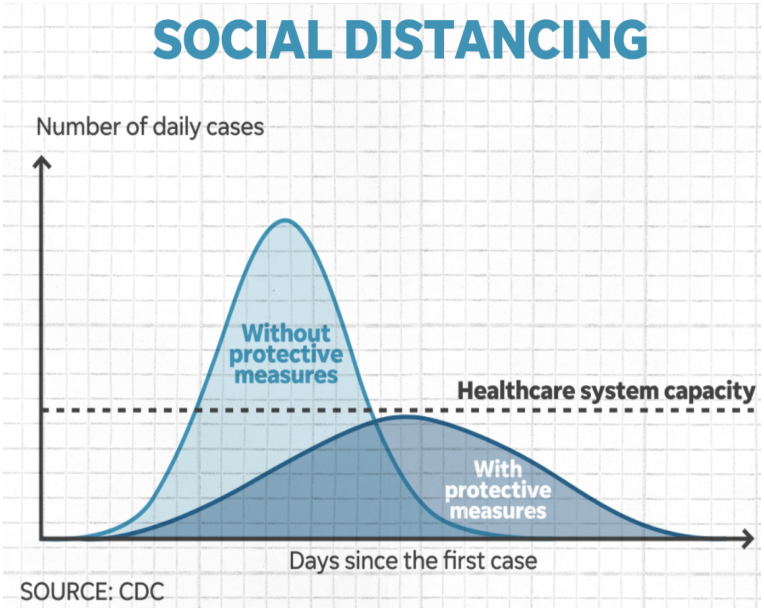
the image .The project aims to cover the following key features :-

- Live video surveillance to fight against covid-19 spread
- Tracking the crowd movement across the day time
- Hot-spot areas can be monitored by security forces from central station
- If AI based solution used by authority then there will be less chance get infected security forces

3 MOTIVATION

In the present scenario, life has stopped due to the fear of a deadly virus. As a result of this some of the most powerful economies of the world have been hit. Governments in various countries have locked the cities in order to prevent the spread of coronavirus. But this is not the permanent solution to the problem at all, sooner or later people are going to start coming out of their houses in order to resume their jobs, lives, as the world cannot be stopped forever. In such a scenario, solutions must be found out in order to resume the daily lives of people and prevent them from getting the coronavirus. There are many things that can be done, but the most common thing that comes to mind is Social Distancing. Everyone has a role to play in slowing the spread and protecting themselves, their family, and their community. In addition to practicing everyday steps to prevent COVID-19, keeping space between you and others is one of the best tools we must avoid being exposed to this virus and slowing its spread in communities.

Now this motivated us to think about social distancing, and how this problem can be solved if we think from the domain of Computer Science. One application that comes to mind is to ensure that proper social distancing standards are being followed at various places (that have a high probability of getting crowded). This can be achieved if we have a way to detect when the norms are being broken.This led us to the application of Computer Vision and Deep Learning techniques on real world scenarios in order to implement a social distancing detection tool. The below curve shows the impact of social distancing :



4 PROBLEM DEFINITION

In the above sections we stated our motivation behind trying to implement an AI based approach for social distancing using Computer Vision , Deep Learning .The task is to build an AI based approach for Detecting Violation of Social Distancing Norms . This is achieved with the help of Object Detection , Object Tracking and Distance Measuring .

For better results we aim to use the state of the Art Object detection algorithm. YOLO is popular because it achieves high accuracy while also being able to run in real-time. The algorithm "only looks once" at the image in the sense that it requires only one forward propagation pass through the neural network to make predictions.

We used Oxford Town Center Dataset for this project.The Oxford Town Centre dataset is a CCTV video of pedestrians in a busy downtown area in Oxford used for research and development of activity and face recognition systems.The CCTV video was obtained from a surveillance camera at the corner of Cornmarket and Market St. in Oxford, England and includes approximately 2,200 people.

Since its publication in 2009 the Oxford Town Centre dataset has been used in over 80 verified research projects including commercial research by Amazon, Disney, OSRAM, and Huawei; and academic research in China, Israel, Russia, Singapore, the US, and Germany among dozens more.

5 LITERATURE REVIEW

Due to this pandemic everybody tries to come up with the solution of social distancing with different techniques. The Paper by S Yadav Deep learning based safe social distancing and face mask detection in public areas for covid-19 safety guidelines adherence . [1] In their method they check for people wearing masks at public places and also whether they are following social distancing or not . Their system gives an image of the violator to nearby control systems and also gives alert to the violator . Their system uses raspberry pi4 with a camera to automatically track the people in public places . Their model is trained with a custom data set installed in the system.then their system is fed with real time video of public spaces . To which if any violator is found they make alerts. The accuracy of their system was 90%.

In this Paper of N Singh Punn, S K Sonbhadra, S Agarwal Monitoring covid-19 social distancing with person detection and tracking via fine-tuned yolo v3 and deepsort techniques [2] . They calculated the violation index term as the number of peoples to the number of groups formed. They used different object detection techniques like RCNN,YOLOV3 and SSD with balanced performance of mAP and FPS score.

The paper by F Sener, N Ikizler-Cinbis, Two-person interaction recognition via spatial multiple instance embedding [3] . In their system they achieve the accuracy of 93.3% . Their experimental findings show the efficient way to infer the form of interaction using the two benchmark dataset. They used embedding of spatial multiple instances , which implicitly integrates the distance between people into the learning process of multiple instances. As the relative spatial position of communicating people likely to complement the visual descriptor. They extracted the motion of communicating people from each region of the detected individual and the visual descriptor for two people was created.

In this paper of L Ramadass, S Arunachalam, Z Sagayasree Applying deep learning algorithms to maintain social distance in public places through drone technology [4] . In this paper they used a drone and yolov3 to monitor social distancing . They can check if a person is wearing a mask or not and social space is maintained or not. The yolov3 algorithm runs with drones with custom data collection. After reading and understanding this papers we get the inspiration of our project .

6 PROPOSED METHODOLOGY

As we described above in this project, we implemented Object Detection on Persons, then we Computed their distance and based on certain parameters we checked in the video stream whether there is Social Distancing Violation or not.

6.1 Object Detection

For the object detection part we made use of the You Only Look Once (YOLO) algorithm. With YOLO, a single CNN simultaneously predicts multiple bounding boxes and class probabilities for those boxes. YOLO trains on full images and directly optimizes detection performance.

6.2 Why we Chose YOLO Algorithm for Object Detection?

Yolo Algorithm has certain advantages over other object detection algorithms.

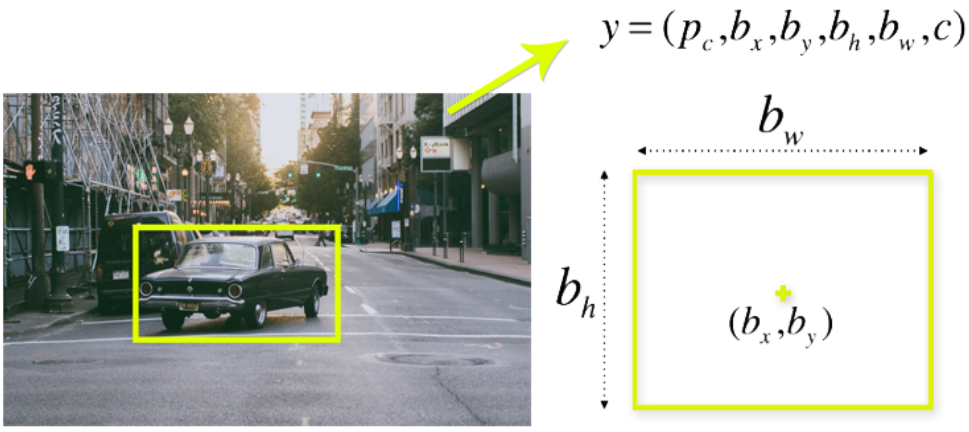
- YOLO is extremely fast.
- YOLO sees the entire image during training and test time so it implicitly encodes contextual information about classes as well as their appearance.
- YOLO learns generalizable representations of objects so that when trained on natural images and tested on artwork, the algorithm outperforms other top detection methods.

The neural network uses features from the entire image or video frame to predict each bounding box. The bounding box coordinates are predicted for each class along with their probability.

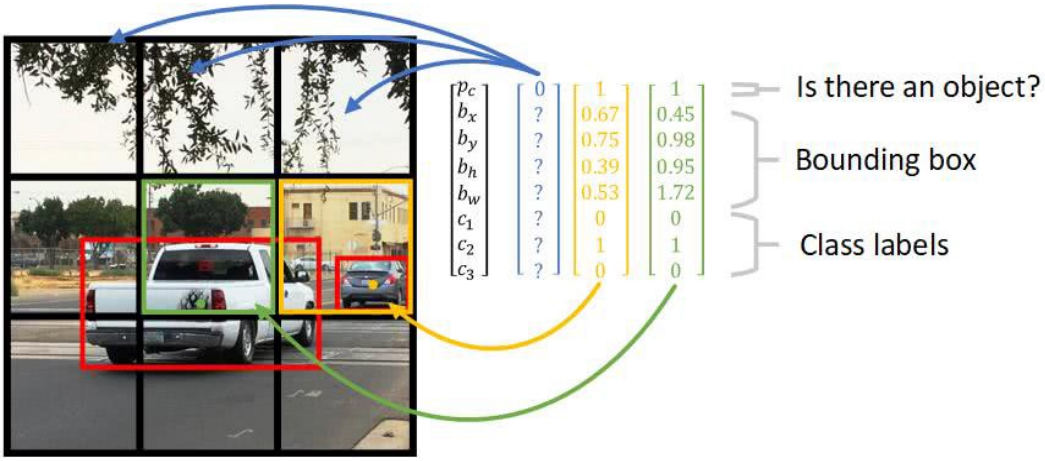
In order to understand the working of Yolo algorithm we need to understand what it predicts in an image frame. Using Yolo our main aim is to predict the class (i.e. what is the object) and where is the object (Object Localisation). Each bounding box is described by centroid of that bounding box and the height and width of that bounding box.

- bx – centroid x coordinate
- by – centroid y coordinate
- bw - width
- bh - height
- c – for the class id (i.e. a person, cat, car etc.)

Also, for each bounding box a confidence score is predicted, that tells us how confident is the algorithm that the object denoted by class c is actually there in that particular bounding box.



The Yolo algorithm is not searching for the desired region in an image, instead it splits the image into grid cells . Each grid cell is responsible for predicting 5 bounding boxes, and confidence scores for those boxes. These confidence scores reflect how confident the model is that the box contains an object and also how accurate it thinks the box is that it predicts. The below figure shows an image and grid cells and how the yolo predicts the label for corresponding grid cell.



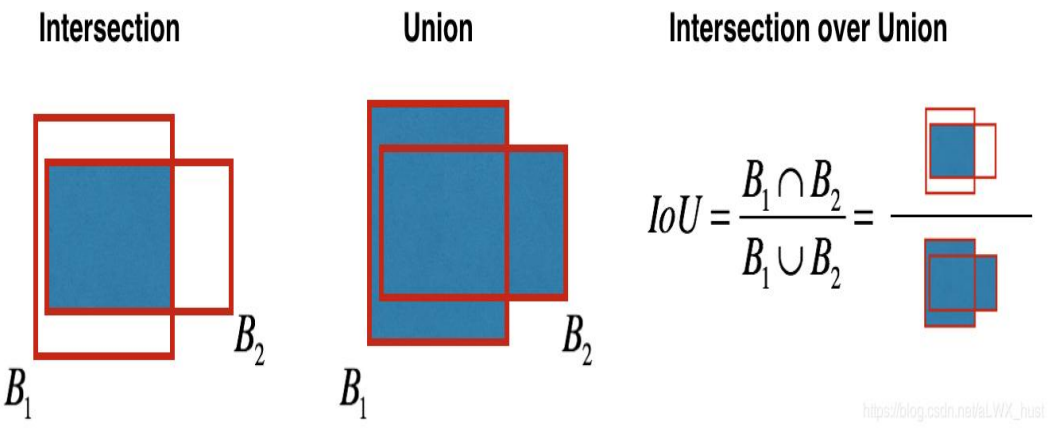
YOLO is a clever convolutional neural network (CNN) for doing object detection in real-time.

6.3 Non Max Suppression

Since YOLO outputs multiple bounding boxes , and all of them are not useful , we need to filter bounding boxes in order to only have the best detections with us. As our algorithm will output multiple detections for the same object category.

For this we use Non – Max Suppression . The candidate regions for our object of interest are called proposals. Non max suppression makes sure we only detect an object one time.

Basically what NMS (Non Max Suppression) does is that for each object category detected , it finds the bounding box with highest confidence score and then all the bounding boxes having a high Intersection over Union (IOU) with the highest confidence box get suppressed. This is done with a help of a threshold value for IOU, that we set initially.
IOU (Intersection Over Union) is defined as the ratio of Area of Intersection of bounding boxes to the Area of their Union.

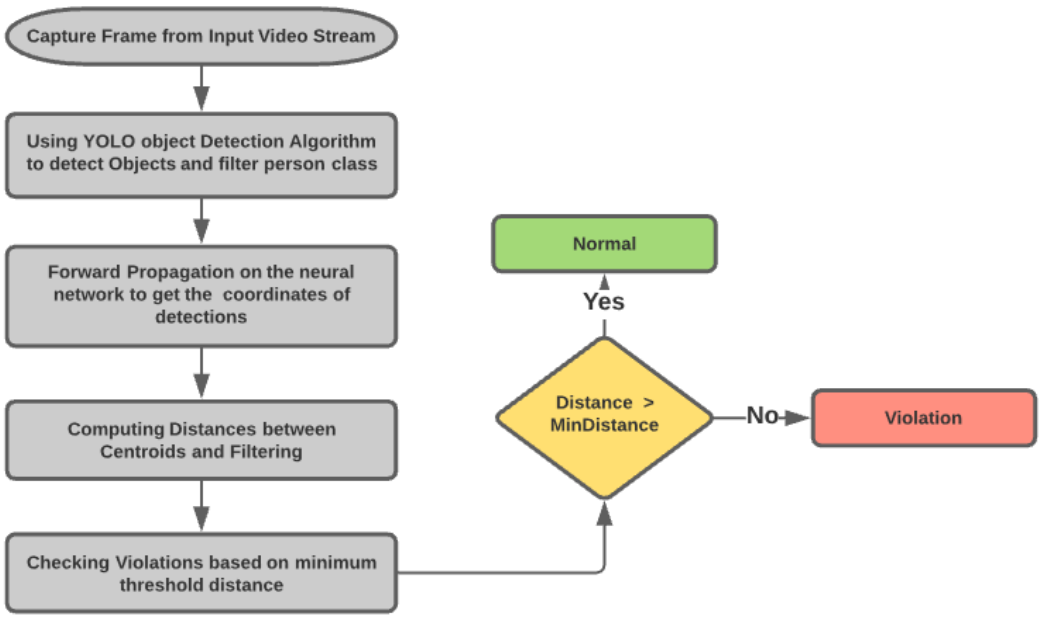


In this project we apply the principles of object detection and object localisation

to build our Social Distance Detector. First we created a function that detects the people in our image frame and returns us the detection results. The detection results consist of confidence scores, coordinates of the bounding box (top left corner coordinates and bottom right coordinates) and the coordinates of their centroids. A minimum confidence criterion is set up that discards all the bounding boxes not meeting that threshold. Since we are using Yolo algorithm , it predicts the boxes for 80 classes but for the application of our project we are only interested in the person class , so we filter the bounding boxes based on the object detected. This can be said as the first line of filtering. Second filtering is done in order to remove all the boxes that predict the same object using **Non-max Suppression.** For a video stream one by one we grab the video frame and pass this frame to the function that we built to get the detection results. After we have these detection results. Since we also return the centroid coordinates, we compute Euclidean distance between each pair of bounding box centroids and compute a distance matrix D. Suppose the centroid at ith index is c1 and centroid at jth index is c2, then

D[i][j] = distance between the bounding boxes of c1 and c2.

Based on this we set a certain threshold and find out all the boxes that violate that threshold criteria, we add the bounding box to the social distancing violation category. The flow of our methodology can be summarised with the help of the below shown flow chart:



7 REQUIREMENTS

7.1 Hardware Requirements

- Machine Supporting GPU , for running YOLO object detection model
- Core i5 CPU
- Minimum 8GB ram

7.2 Software Requirements

- Python 3.7
- OS supporting Google Chrome Web Browser
- Numpy
- Time
- OpenCV
- OpenCV_Contrib
- Math
- Jupyter Notebook / Google Collab

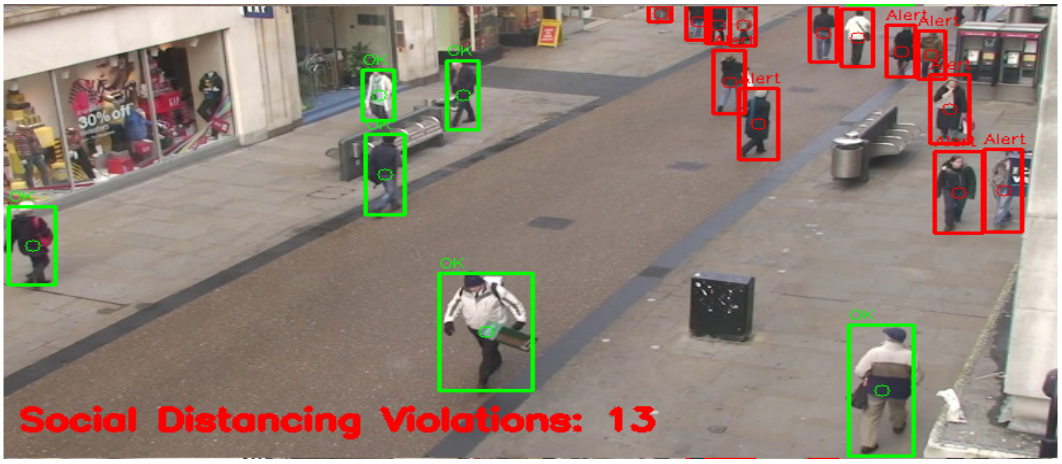
8 IMPLEMENTATION PLAN

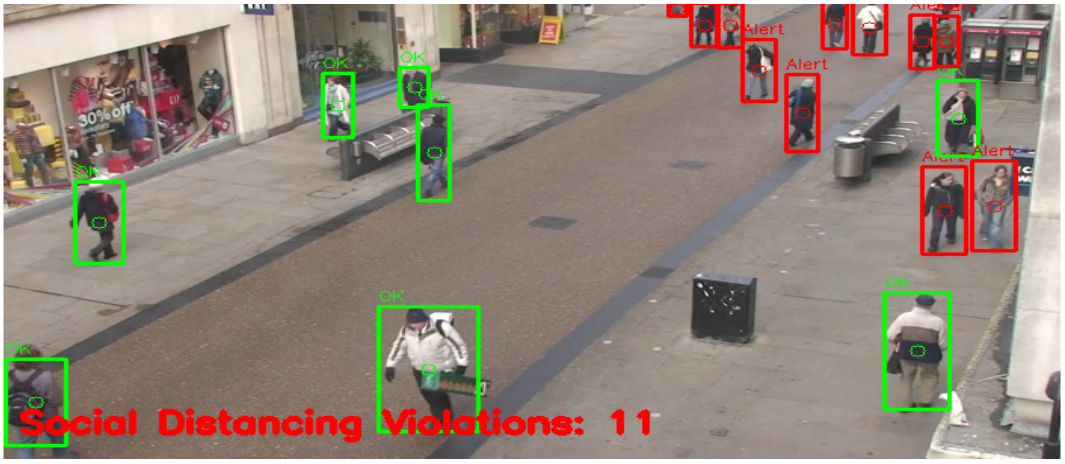
Task	Start Date	End Date
Literature review	20-06-2020	25-08-2020
Finalizing Problem Statement and Goals	22-06-2020	24-06-2020
Studying different Object detection techniques	24-06-2020	25-07-2020
Working on code	25-07-2020	26-08-2020
Preparing report	26-08-2020	02-09-2020

Table 1: Implementation Timeline

9 RESULTS

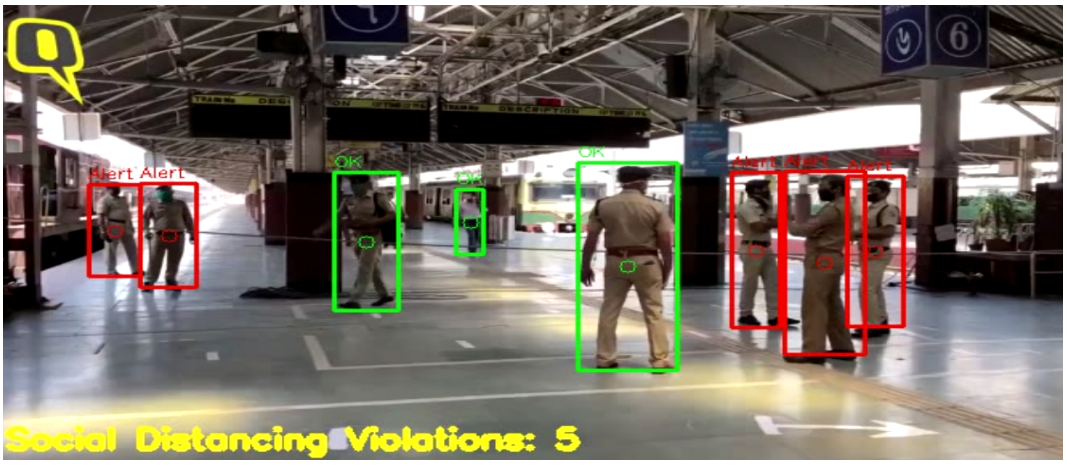
In this project the function that we built for performing object detection on persons and compute the distance can take input images and perform computation on them. Also with the help of OpenCV library we can pass the video stream as the input and in turn generate video output in real time that shows the social distancing violations happening in each frame. We also mark the person's bounding box as red in order to depict that social distancing norms are violated. And mark those bounding boxes with Alert. Rest all the bounding boxes are marked green. We ran our algorithm on the CCTV footage of pedestrians in a busy downtown area in Oxford that was taken from the Oxford Data Center. As a result we got the output video depicting the result of the social distancing detector that we developed. Some of the frames from that result are shown below.





Social Distancing Violations: 9

INDIAN RAILWAYS
Social Distancing Violations: 0


Social Distancing Violations: 5

10 Comparative Study

We have developed our code to work on webcam stream ,video input and images. Since we have used the YOLO algorithm we are getting results on object detection. We can compute the accuracy of pure object detection in this case. But since our problem is based on Social Distancing,which is a random problem we cannot compute exact evaluation metrics for Social Distancing, Though we have the accuracy for YOLO's object detection. In social distancing we have no way to know the ground truth beforehand i.e. whether or not the person violated the social distancing norms or not, Since it is meant to work in real time , we have no way of determining the distances before-hand. But since we can do accuracy comparison for object detection we have presented a comparative study of other object detection models besides YOLO.

Model	**Accuracy%**	**Time**	**fps**
R-CNN	73.2	0.11	9.1
Faster-RCNN	83.8	2.24	0.4
SSD300	74.3	0.01	46
SSD 512	76.8	0.05	19
YOLO	84.4	0.03	38
YOLOv2	88.6	0.03	40
YOLOv3	89.3	0.03	45

Table 2: Comparative Study

11 FUTURE WORK

In this project we combined the knowledge of Computer Vision , Deep Learning in order to build an AI based Social Distancing Solution. Though we have implemented a simple implementation that tells whether the social distancing norms are being violated or not based on object tracking and object detection , it certainly can be extended and upscaled in order to add more features and become more robust.

Some of the things that can be added are :-

- In order to improve our social distancing detection we need to utilize a proper camera calibration . Doing so will yield better results and enable us to compute actual measurable units (rather than pixels).
- We can also have an alert system that gets triggered in a place whenever the number of social distancing violations exceed a certain threshold. If this type of solution is deployed at places then less police workforce have to monitor the

situation personally , which in turn will also guarantee the safety of security personnel.

- This project can be extended to further give alerts on a mobile via a mobile app , The registered users of the mobile app can be given an alert whenever there are situations of social distancing violations in a place.

12 CONCLUSION

Through the entire project phase we came across various studies and research papers regarding this topic and other topics specific to the uncertain time we are living in. We briefly discussed those in the Literature Review section. We discussed how we applied the knowledge of computer vision and deep learning to solve a real world problem . Later we discussed the object detection algorithm that we used and how we built our project. The results that were obtained from our algorithm are presented. We also discussed how we can improve our social distancing detector to become more robust and more applicable in the real world.

13 REFERENCES

1. S Yadav "*Deep learning based safe social distancing and face mask detection in public areas for covid-19 safety guidelines adherence*" `http://ijraset.com/fileserve.php?FID=30560`
2. N Singh Punn, S K Sonbhadra, S Agarwal "*Monitoring covid-19 social distancing with person detection and tracking via fine-tuned yolo v3 and deepsort techniques*" `https://arxiv.org/abs/2005.01385`
3. F Sener, Ikizler-Cinbis "*Two-person interaction recognition via spatial multiple instance embedding Journal of Visual Communication and Image Representation, volume 32*" `https://linkinghub.elsevier.com/retrieve/pii/S104732031500142X`
4. L Ramadass, S Arunachalam, Z Sagayasree "*Applying deep learning algorithm to maintain social distance in public place through drone technology*" International Journal of Pervasive Computing and Communications `https://www.emerald.com/insight/content/doi/10.1108/IJPCC-05-2020-0046/full/html`
5. Online Article Social Distancing `https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html`
6. Oxford Town Center Dataset `https://megapixels.cc/oxford_town_centre/`

7. Oxford Town Center Dataset https://megapixels.cc/oxford_town_centre/
8. Open-Cv Deep Neural Network Module Available: https://docs.opencv.org/3.4/d6/d0f/group__dnn.html
9. open-cv net class reference https://docs.opencv.org/3.4/db/d30/classcv_1_1dnn_1_1Net.html#ae8be9806024a0d1d41aba687cce99e6b