IJARSCT
Impact Factor: 6.252

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Volume 2, Issue 2, March 2022

Social Distancing Detector Using Deep Learning

Mr D.B Khadse[1], Mayur Shahu[2], Yogesh Karemore[3], Sanidhya Sonkusare[4], Dhananjay Kannake[5]

Assistant Professor, Department of Computer Science & Engineering[1]

UG Students, Department of Computer Science & Engineering[2,3,4,5]

Priyadarshini Bhagwati College of Engineering, Nagpur, Maharashtra, India

Abstract: *The paper presents a methodology for social distancing detection using deep learning to evaluate the distance between people to mitigate the impact of this coronavirus pandemic. The detection tool was developed to alert people to maintain a safe distance with each other by evaluating a video feed. The video frame from the camera was used as input, and the 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. The distance between people can be estimated and any noncompliant pair of people in the display will be indicated with a red frame and red line. The proposed method was validated on a pre-recorded video of pedestrians walking on the street. The result shows that the proposed method is able to determine the social distancing measures between multiple people in the video. The developed technique can be further developed as a detection tool in real time application.*

Keywords: Social distancing, Pedestrian Detection, Deep Learning, Streaming Media, Social Factor

I. INTRODUCTION

When the novel coronavirus (Covid-19) pandemic emerges, the spread of the virus has left public keep anxiety if they do not have any effective cure. The World Health Organization (WHO) has declared Covid-19 as a pandemic due to the increase in the number of cases reported around the world. To contain the pandemic, many countries have implemented a lockdown where the government enforced that the citizens to stay at home during this critical period. The public health bodies such as the Centers for Disease Control and Prevention (CDC) had to make it clear that the most effective way to slow down the spread of Covid-19 is by avoiding close contact with other people. To flatten the curve on the Covid-19 pandemic, the citizens around the world are practicing physical distancing.

To implement social distancing, group activities and congregations such as travel, meetings, gatherings, workshops, praying had been banned during the quarantine period. The people are encouraged to use phone and email to manage and conduct events as much as possible to minimize the person-to-person contact. To further contain the spread of the virus, people are also informed to perform hygiene measures such as frequently washing hands, wearing mask and avoiding close contact with people who are ill. However, there is a difference between knowing what to do to reduce the transmission of the virus and putting them into practice[1].

The world has not yet fully recovered from this pandemic and the vaccine that can effectively treat Covid-19 is yet to be discovered. However, to reduce the impact of the pandemic on the country's economy, several governments have allowed a limited number of economic activities to be resumed once the number of new cases of Covid has dropped below a certain level. As these countries cautiously restarting their economic activities, concerns have emerged regarding workplace safety in the new post-Covid-19 environment. To reduce the possibility of infection, it is advised that people should avoid any person-to-person contact such as shaking hands and they should maintain a distance of at least 1 meter from each other.[4]

Individuals, communities, businesses, and healthcare organizations are all part of a community with their responsibility to mitigate the spread of the Covid-19 disease. In reducing the impact of this coronavirus pandemic, practicing social distancing and self-isolation have been deemed as the most effective ways to break the chain of infections after restarting the economic activities. In fact, it has been observed that there are many people who are ignoring public health measures, especially with respect to social distancing. It is understandable that given the people's excitement to start working again, they sometimes tend to forget or neglect the implementation of social distancing. Hence, this work aims to facilitate the enforcement of social distancing by providing automated detection of social distance violation in workplaces and public areas using a deep learning model. In the area of machine learning and computer vision, there are different methods that can be used for object detection. These methods can also be applied to detect the social distance between people[9].

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 2, March 2022

II. Literature Survey

Human detection can be considered as an object detection in the computer vision task for classification and localization of its shape in video imagery. Deep learning has shown a research trend in multi-class object recognition and detection in artificial intelligence and has achieved outstanding performance on challenging datasets. The survey mainly focuses on human descriptors, machine learning algorithms, occlusion, and real-time detection. For visual recognition, techniques using deep convolutional neural network (CNN) have been shown to achieve superior performance on many image recognition benchmarks.

Deep CNN is a deep learning algorithm with multilayer perceptron neural networks which contain several convolutional layers, sub-sampling layers, and fully connected layers. Later, the weight in the whole layers in the networks are trained for each object classification based on its dataset. For object detection in image, the CNN model was one of the categories in deep learning which are supervised feature learning methods robust in detecting the object in different scenarios. CNN has achieved great success in large-scale image classification tasks due to the recent high performance computing system and large dataset such as ImageNet, COCO. Different CNN models for object detection with its object localization had been proposed in terms of network architecture, algorithms, and new ideas. In recent years, CNN models such as Alex-Net, VGG16, InceptionV3, and ResNet-50 are trained to achieve outstanding results in object recognition. The success of deep learning in object recognition is due to its neural network structure that is capable of self-constructing the object descriptor and learning the high-level features which are not directly provided in the dataset.

The current state-of-the-art object detectors with deep learning had their pros and cons in terms of accuracy and speed. The object might have different spatial locations and aspect ratios within the image. Hence, the real-time algorithms of object detection using the CNN model such as R-CNN and YOLO had further developed to detect multi-classes in a different region in images had been developed. YOLO (You Only Look Once) is the prominent technique for deep CNN-based object detection in terms of both speed and accuracy.

Adapting the idea from the work, we present a computer vision technique for detecting people via a camera installed at the roadside or workspace. The camera field-of-view covers the people walking in a specified space. The number of people in an image and video with bounding boxes can be detected via these existing deep CNN methods where the YOLO method was employed to detect the video stream taken by the camera. By measuring the Euclidean distance between people, the application will highlight whether there is sufficient social distance between people in the video.

III. Methodology

3.1 Yolo Algorithm

YOLOv3 (You Only Look Once, Version 3) is a real-time object detection algorithm that identifies specific objects in videos, live feeds, or images. YOLO uses features learned by a deep convolutional neural network to detect an object detection. Human detection can be considered as an object detection in the computer vision task for classification and localization of its shape in video imagery. Deep learning has shown a research trend in multi-class object recognition and detection in artificial intelligence and has achieved outstanding performance on challenging datasets. Nguyen et al. presented a comprehensive analysis of state -of-the-art on recent development and challenges of human detection. The survey mainly focuses on human descriptors, machine learning algorithms, occlusion, and real-time detection. For visual recognition, techniques using deep convolutional neural network (CNN) have been shown to achieve superior performance on many image recognition benchmarks[5].

YOLO algorithm employs convolutional neural networks (CNN) to detect objects in real-time. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects. This means that prediction in the entire image is done in a single algorithm run. The CNN is used to predict various class probabilities and bounding boxes simultaneously.The illustration for the YOLO model is shown in Fig1 [11].

YOLO algorithm is important because of the following reasons:

- **Speed:** This algorithm improves the speed of detection because it can predict objects in real-time.
- **High accuracy:** YOLO is a predictive technique that provides accurate results with minimal background errors.
- **Learning capabilities:** The algorithm has excellent learning capabilities that enable it to learn the representations of objects and apply them in object detection.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 2, March 2022

3.2 Working of Yolo

YOLO algorithm works using the following three techniques:

- Residual blocks
- Bounding box regression
- Intersection Over Union (IOU)

A. Residual Blocks

First, the image is divided into various grids. Each grid has a dimension of S x S. The following image shows how an input image is divided into grids. There are many grid cells of equal dimension. Every grid cell will detect objects that appear within them. For example, if an object center appears within a certain grid cell, then this cell will be responsible for detecting it.

B. Bounding Box Regression

A bounding box is an outline that highlights an object in an image.
Every bounding box in the image consists of the following attributes:
Width (bw), Height (bh).
Class (for example, person, car, traffic light, etc.)- This is represented by the letter c.
Bounding box center (bx,by).The bounding box has been represented by a yellow outline.
YOLO uses a single bounding box regression to predict the height, width, center, and class of objects. In the image above, represents the probability of an object appearing in the bounding box.

C. Intersection Over Union (IOU)

Intersection over union (IOU) is a phenomenon in object detection that describes how boxes overlap. YOLO uses IOU to provide an output box that surrounds the objects perfectly.

Each grid cell is responsible for predicting the bounding boxes and their confidence scores. The IOU is equal to 1 if the predicted bounding box is the same as the real box. This mechanism eliminates bounding boxes that are not equal to the real box.

In the image below fig.1, there are two bounding boxes, one in green and the other one in blue. The blue box is the predicted box while the green box is the real box. YOLO ensures that the three bounding boxes are equal[11].

This social distancing detection tool was developed to detect the safety distance between people in public spaces. The deep CNN method and computer vision techniques are employed in this work. Initially, an open-source object detection network based on the YOLO v3 algorithm was used to detect the pedestrian in the video frame. From the detection result, only pedestrian class was used and other object classes are ignored in this application. Hence, the bounding box best fits for each detected pedestrian can be drawn in the image, and these data of detected pedestrians will be used for the distance measurement[6].

For camera setup, the camera is captured at fixed angle as the video frame, and the video frame was treated as perspective view are transformed into a two-dimensional top-down view for more accurate estimation of distance measurement. In this methodology, it is assumed that the pedestrians in the video frame are walking on the same flat plane. Four filmed plane points are selected from frame and then transformed into the top-down view. The location for each pedestrian can be estimated based on the top-down view. The distance between pedestrians can be measured and scaled. Depending on the present minimum distance, any distance less than the acceptable distance between any two individuals will be indicated with red lines that serve as precautionary warnings.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 2, March 2022

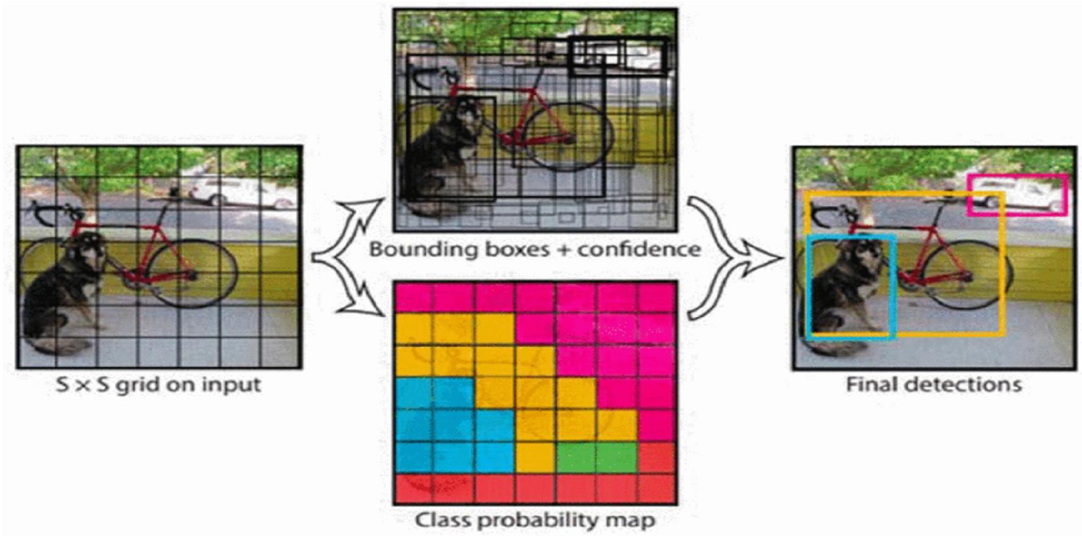


Figure 3.1: The illustration of YOLO model for object detector pipeline.

The work was implemented using the Python programming language. The pipeline of the methodology for the social distancing detection tool is shown in Figure 3.2.

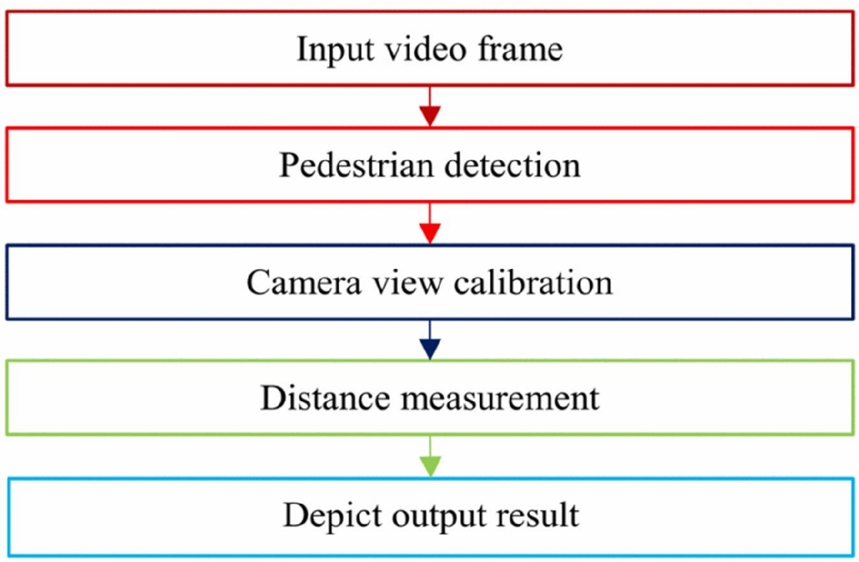


Figure 3.2: Pipeline for social distancing detection.

A. Pedestrian Detection

Deep CNN model was the object detection approach was proposed that mitigated computational complexity issues by formulating the detection with a single regression problem. When it comes to deep learning-based object detection, the YOLO model is considered one of the state-of-the art object detectors which can be demonstrated to provide significant speed advantages will suitable for real-time application. In this work, the YOLO model was adopted for pedestrian detection is shown in Figure 3.1.1.

The YOLO algorithm was considered as an object detection taking a given input image and simultaneously learning bounding box coordinates (tx, ty, tw, th), object confidence and corresponding class label probabilities (P1, P2, …, Pc). The YOLO trained on the COCO dataset which consists of 80 labels including human or pedestrian class. In this work, the only box coordinates, object confidence and pedestrain object class from detection result in the YOLO model were used for pedestrian detection[8].

DOI: 10.48175/IJARSCT-2930

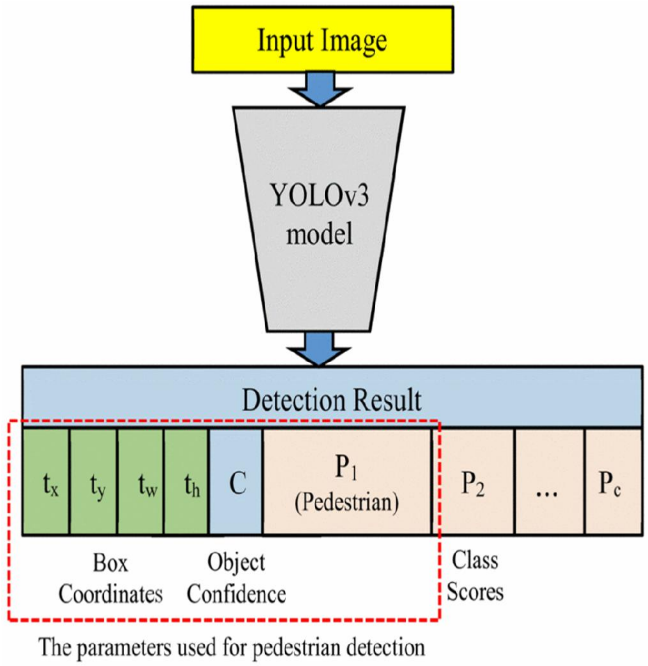


Figure 3.1.1: YOLO model applied for pedestrian detection.

3.3 Camera View Calibration

The region of interest (ROI) of an image focuses on the pedestrian walking street was transformed into a top-down 2D view that contains 480×480 pixels as shown in Figure 3.2.1. Camera view calibration is applied which works by computing the transformation of the perspective view into a top-down view.

In OpenCV, the perspective transformation is a simple camera calibration method which involves selecting four points in the perspective view and mapping them to the corners of a rectangle in the 2D image view. Hence, every person is assumed to be standing on the same level flat plane. The actual distance between pedestrians corresponds to the number of pixels in the top-down view can be estimated[2].

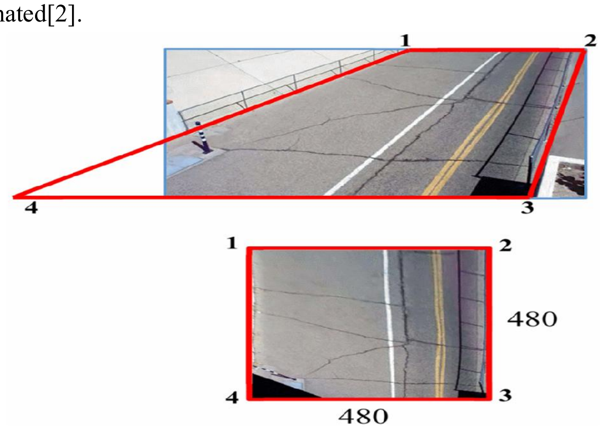


Figure 3.2.1: Sample of original perspective view (top) and top-down view after calibration

3.4 Distance Measurement

In this step of the pipeline, the location of the bounding box for each person (x, y, w, h) in the perspective view is detected and transformed into a top-down view. For each pedestrian, the position in the top-down view is estimated based on the bottom-centre point of the bounding box. The distance between every pedestrian pair can be computed from the top-down view and the distances is scaled by the scaling factor estimated from camera view calibration. Given the position of two

DOI: 10.48175/IJARSCT-2930

pedestrians in an image as (x1, y1) and (x2, y2) respectively, the distance between the two pedestrians, d., can be computed as:

$$d=\sqrt{(x2-x1)^2+(y2-y1)^2}$$

The pair of pedestrians whose distance is below the minimum acceptable distance, t., is marked in red, and the rest is marked in green. A red line is also drawn between the pair of individuals whose distance is below the pre-defined threshold. The bounding box's colour threshold operation, c, can be defined as[9]:

$$c=\begin{cases} \text{red} & d<t \\ \text{green} & d<=t \end{cases}$$

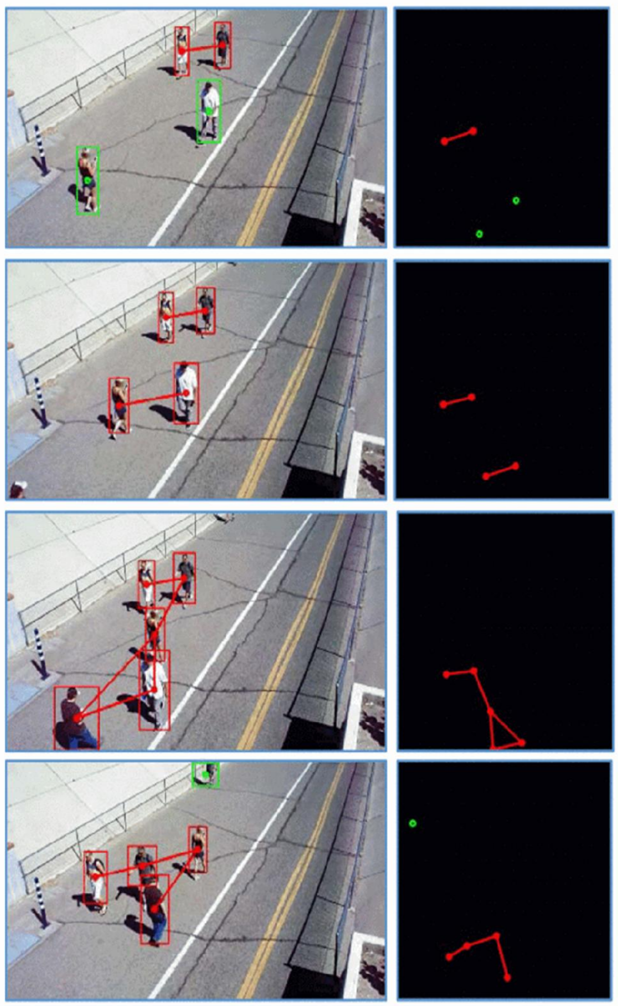


Figure 3.3.1: Social distancing measurements

IV. Software Requirements

- Operating System: Window 7 Or Later
- Language: Python 3.9.7
- Open-cv (Library)
- Python IDE
- PyCharm 2020.3 Or Above Version

IJARSCT

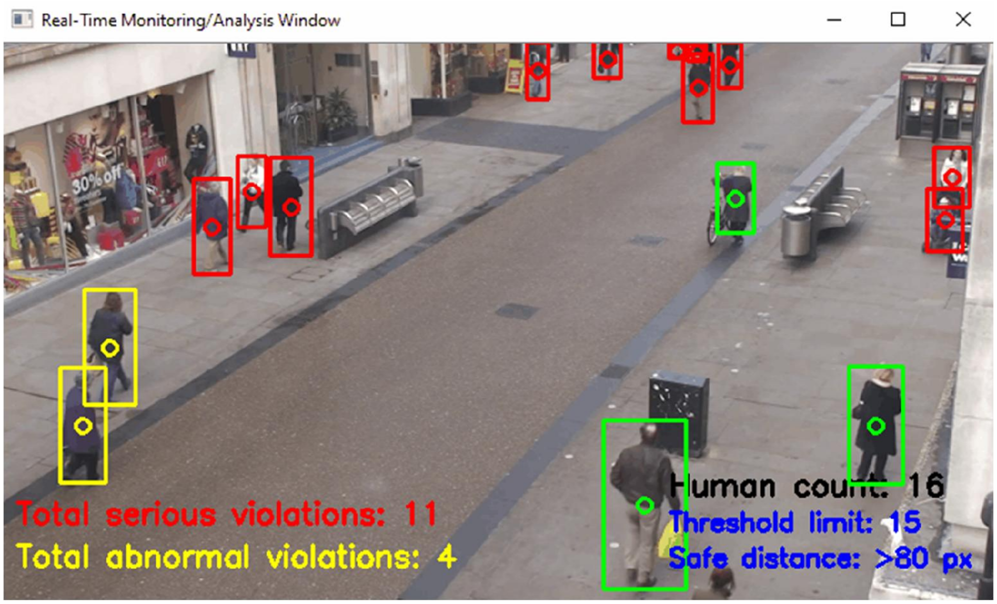
ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 2, March 2022

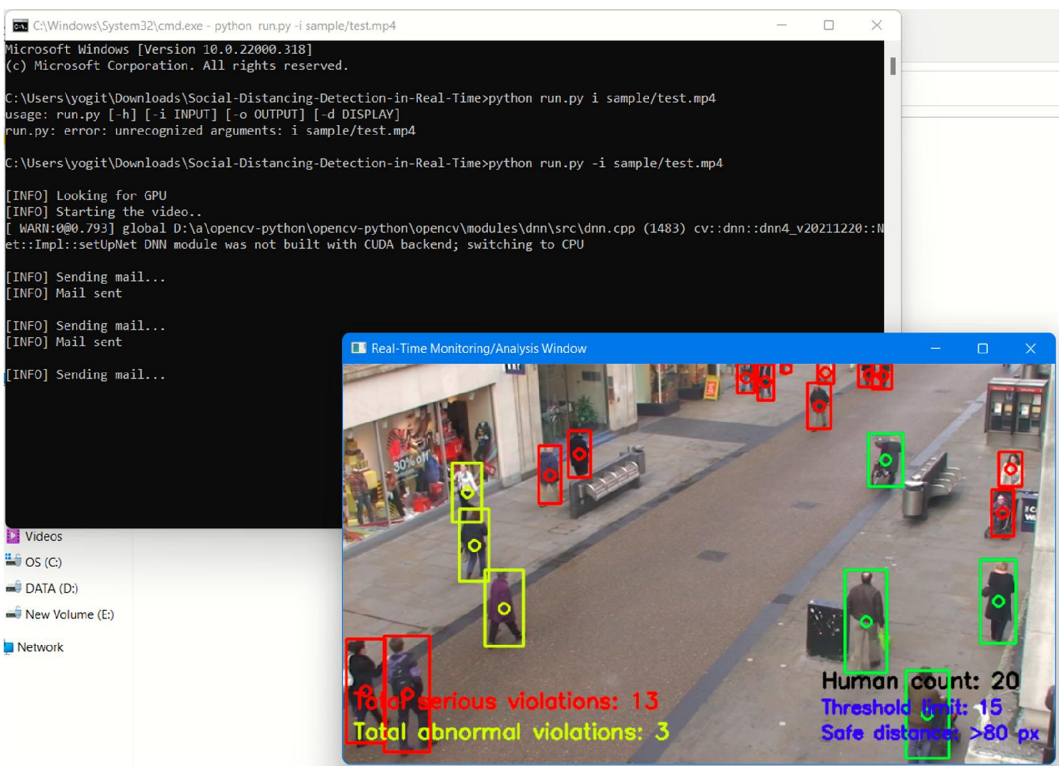
V. Output



Output 1. Social Distance Detection For Each Video Frame

Output 2: Social Distance Detection For Each Image Frame

Output 3. Social Distance Detection in live streaming



Output 4. Sending mail as detecting a distancing violation

VI. Conclusion

In our proposed system, we successfully used live video stream like CCTV, cameras as an input. And finally in output it gives alert (red) mark on person in video, when pedestrians are not maintaining a safe social distance among them. Send a mail to operator or user as a real-time warning. Our goal is to identify whether the person on image or in video stream is having a safe distance or not with the help of Computer vision and Deep learning.

REFERENCES

[1]. Implementation of Mitigation Strategies for Communities with Local COVID-19. Available:https://www.who.int/emergencies/diseases/novel-coronavirus, 2019.

[2]. W. H. Organization, "WHO corona-viruses (COVID-19)," https://www.who.int/emergencies/diseases/novel-coronavirus,2019,2020

[3]. D.T. Nguyen, W. Li and P.O. Ogunbona, "Human detection from images and videos: A survey", Pattern Recognition, 2016.

[4]. J. Redmon and A. Farhadi, Yolov3: An Incremental improvement, retrieved september vol 17,2018.

[5]. T.Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C. L. Zitnick, "Microsoft coco: Common objects in ´ context," in European conference on computer vision. Springer, 2014, pp 740 - 755.

[6]. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580–587.

[7]. Pias, "Object detection and distance measurement," https://github.com/ paul-pias/Object-Detection-and-Distance-Measurement, 2020, [Online; accessed 01-March-2020].

[8]. J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You only look once: Unified real-time object detection", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788, 2016.

[9]. https://ieeexplore.ieee.org/abstract/document/9243478.

[10]. https://github.com/saimj7/Social-Distancing-Detection-in-Real-Time

[11]. https://www.section.io/engineering-education/introduction-to-yolo-algorithm-for-object-detection.