International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

# A Novel Approach to Detect Social Distancing Among People in College Campus

**Jenita Hermina J*[1], N. Shunmuga Karpagam[2], P. Deepika[3], D. Santhi Jeslet[4], Dinesh Komarasamy[5]**

**Submitted: 07/02/2022 Accepted : 04/03/2022**

***Abstract:*** The world has shrunk due to the COVID-19. This virus took many lives of people. Hence the government proposed a lockdown and closed the educational institutions to decrease the death rate. The cases are controlled to some extent, but still the spread of the virus is not yet controlled. Though lockdown has been released and the educational institutions are being slowly opened with some conditions like wearing masks, maintaining social distance in the public areas and avoiding contact with the infected persons. But the people are not maintaining social distance in the colleges and schools, which can cause risk. Hence it is mandatory to build a model which identifies the people who cross the safe limit of distance. This work focuses on determining the social distance in the educational institutions mainly in the colleges. Generally, most of the colleges have Closed Circuit Television (CCTV) cameras fixed on the campus, in order to determine the social distance among the people, the videos are taken from those CCTV cameras. The people are detected from the video clips using the You Only Look Once Version 3 (YOLO V3) algorithm which is trained on the Common Objects in Context (COCO) dataset. And the safety distance among the people is found using Euclidean distance. The minimum threshold value is fixed to the safe distance suggested by World Health Organization (WHO). Moreover, a warning is sent to the corresponding coordinators through Short Message Service (SMS) when the people are not in the safer social distance.

***Keywords:*** *COVID-19, social distance, YOLO V3 algorithm, SMS alerts*



## 1. Introduction

COVID-19 is a type of virus that causes cold, fever, headache and for some people it causes severe problems like shortness of breathing, lung damage which in turn leads to death. WHO declares that the virus spreads from people to people more easily by the droplets from the nose and mouth. The spread is mainly due to the physical contact with the people who has affected by COVID-19 [6]. Due to this, the death rate has been kept on increasing in the past years. Therefore, government announced a lockdown for some period of time. Now the death rate decreased to a certain level, but still the scariness of the COVID-19 does not drop. The Government released the lockdown with some conditions like social distancing and covering the faces with a mask so that the spread can be kept in control [8]. With the precautions followed the Government allowed the college people and school people are started their regular flow of going to the colleges and schools. But it is difficult to maintain the social distance in the colleges. Because of this, the widespread of COVID-19 may occur again in a huge amount. Therefore monitoring the social distance among the college is more important.

To monitor the people, the existing systems by different researchers use different techniques to detect the people safety movement using Computer Vision, sensors and deep learning methods. However, the accuracy, time complexity and the real-time working of the system show fewer scopes [7] [10]. To improve the accuracy and minimize the time complexity, YOLO V3 algorithm is used in the proposed work. Furthermore, to reduce the time complexity, the extracted videos are converted into frames with a time interval of 1 minute. Moreover, the work utilizes different kinds of videos in different locations and timings.

The proposed work uses the video captured from the CCTV that is fixed on the college campus. With the video collected, the person detection is carried out using YOLO V3 algorithm. YOLO V3 uses the Convolutional Neural Network (CNN) to detect various objects in the video or images by extracting the features, which are trained on COCO Dataset. COCO dataset contains 80 different object categories, which help to detect the object in the video. Moreover, the distance between the people is found by calculating the Euclidean distance. The alert is sent to the corresponding faculty in-charge through SMS. The following section describes the related work followed by the proposed work.

## 2. Related Work

Researchers who worked on detecting the social distance from the video are using various techniques like Internet of Things (IoT), Computer Vision and Deep Learning techniques. Afiq Harith

[1] *Department of Computer Science and Engineering, Er.Perumal Manimekalai College of Engineering, Hosur, India, ORCID ID:0000-0003-0894-923X*
[2] *Department of Computer Science and Engineering, Er.Perumal Manimekalai College of Engineering, Hosur, India ORCID ID:0000-0002-9100-102X*
[3] *Assistant Professor, Department of Electrical and Electronics Engineering, Er.Perumal Manimekalai College of Engineering, Hosur, India, ORCID ID:0000-0002-4351-3194*
[4] *Department of Computer Science, M.G.R. College, Hosur, India, ORCID ID:0000-0003-2232-8878*
[5] *Department of Computer Science and Engineering, Kongu Engineering College, Perundurai, India, ORCID ID:0000-0002-3614-5530*
* *Corresponding Author Email: jenihermina@gmail.com*

Ahamad et al. introduced a method “Person Detection for Social Distancing and Safety Violation Alert based on Segmented ROI” to detect social distance among the person using MobileNet Single Shot Multibox Detector (SSD) algorithm. The distance between the people is found using the central points of the persons. This method also finds the people who are in the restricted area and gives the alert based on Region Of Interest (ROI). Though the work gives good accuracy for indoor videos, the accuracy for outdoor videos is less [2]. In order to acquire more accuracy, Suresh K et al. proposed a solution “Social Distance Identification Using Optimized Faster Region-Based Convolutional Neural Network” which detects the social distance of the person using Optimized Faster Region-Based Convolutional Neural Network (OFRCNN) which is a modified version of CNN. COCO dataset is used for training the model. The dataset is split into 6:1:3 ratio for training, validation and testing respectively. The person with the not safer social distance is marked with the red bounded box. The overall accuracy of the solution is 97%. Though the accuracy is better, the time for training the model is substantially longer [3].

In order to make the faster and more accurate training, Sergio Saponara et al. proposed the work “Implementing a real-time, AI-based, people detection and social distancing measuring system for Covid-19” to detect the safe social distancing among the people. A video is obtained from the thermal camera and YOLOv2 algorithm is used to detect the people from the video. Euclidean distance is used to find the distance between the people. The proposed work is embedded with the hardware device called NVIDIA device [5]. Some researchers used the newer version of YOLO to predict more bounding boxes. To support the above statement, Yew Cheong Hou et al. proposed a work called “Social Distancing Detection with Deep Learning Model” to detect the social distance from the video using YOLO V3 algorithm. The algorithm is trained using COCO dataset and is used to detect the pedestrian who walks in public place. The distance between the pedestrian is found using Euclidean distance. The pedestrian with less social distance is marked with the red bounding box [1]. Similarly, Sheshang Degadwala et al. introduced the idea “Visual Social Distance Alert System Using Computer Vision & Deep Learning” which uses YOLOv3 algorithm to detect the person in the video which is trained in both COCO and PASCAL-VOC datasets. And Euclidean distance is used to calculate the distance between the people. The idea uses both online and offline videos to detect social distance [4].

From the state of the art, researchers introduced various algorithms to identify social distance among people. But the accuracy in detecting social distance is less. Moreover, the approaches didn’t use the different videos taken at different times like in the morning and evening. In addition to that, the alert system to the respective in-charge or higher official is not maintained. Thus, the proposed work solves all the short comes of the existing methods.

## 3. Proposed Work

Social detection among the people in the college is important to control the widespread of COVID-19. The proposed work detects the social distance among the people using YOLO V3 algorithm. There are different phases like video capturing, converting the videos to frame, preprocessing, person detection, distance calculation and sending alerts to SMS. Below shows the pseudocode for detecting social distance. Input for the algorithm is Video captured from CCTV. And the output of the algorithm is detecting people in the frame, predicting social distance among the people and sending SMS to the particular in-charge.

```
Start
   Capture Video from CCTV, Convert into frames
   for each frames:
          resize the frame
          train YOLOV3 against COCO dataset to detect people
          identify the mid point of each people
          if count of people >=2
                    calculate the Euclidean distance
                    if distance < 2meter
                         bound the people with red color
                         send alert message to the coordinator
                    else
                         bound the people with green color
                    End if
          End if
   End for
End
```

### 3.1. Working of Proposed Work

The working of the proposed work is represented in the form of a flow diagram. There are different phases namely video capturing, preprocessing, person detection, detection of social distance and sending SMS. Fig. 1 shows the overall flow of the proposed work. The below section explains each module in the proposed work.

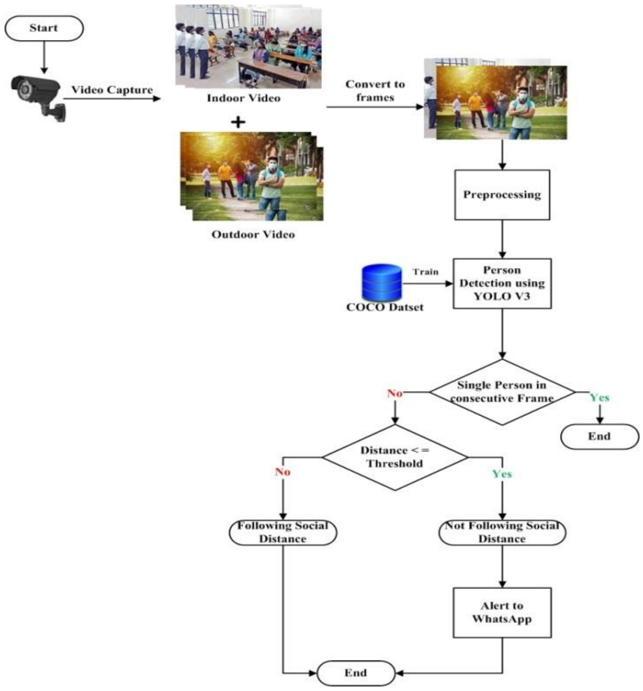


**Fig. 1.** Overall flow of proposed work

### 3.2. Video capturing and converting to frames

The video is captured from CCTV cameras which are fixed in the various areas in the college campus. The work utilizes the videos gathered from both the indoor and outdoor areas inside the college campus. As the video is a large number of consecutive frames, the video is converted into individual frames. For every 1 minute, the frame is extracted from the video captured to reduce the time complexity of the model and to reduce the overloading of the model.

### 3.3. Preprocessing

Once the videos are converted into frames, the preprocessing technique of resizing is applied to the collection of frames.

Preprocessing is done as the videos are captured from various sources and all the frames might not be in the same size. In order to acquire a good accuracy of detecting a person from the frame, the frames are resized before it is fed to the YOLO V3 algorithm. The width and height of the frame is converted to 416 and 416 respectively as YOLO V3 expects the input in the dimension of 416 X 416 X 3.

### 3.4. Usage of COCO Dataset

COCO dataset consists of 80 different object categories with 330K images. Generally, the COCO dataset is used for object detection. COCO dataset follows a JSON file format that has key-value pairs. It is a collection of "info", "licenses", "images", "categories", "annotations". Among these "info" gives high-level information about the dataset like description, version, Uniform Resource Locator (URL), year, contributor and created date. "licenses" is a list that contains licenses of the images that are available in the dataset. "images" gives the details about the images in the dataset. The details like licenses, file name, coco URL, height, width, captured date, Flickr URL and id. "categories" gives the category, subcategory, key points and skeleton that the image resides to. Here key points refer to the parts of the objects, skeleton refers to the connection between the keypoints. "annotations" give detailed information about the objects in the images. The annotations contain id, category id, image id, is crowd, segmentation, area and bounding box. Here isCrowd contains 0 or 1, 0 means a single object in the image; 1 means a group of objects. Segmentation shows the list of points that define the object which is the combination of x and y coordinates. Whereas area determines the total number of pixels of the object in the image. The bounding box shows the box around the image which is represented by x and y coordinates. The below fig. 2 shows the image which is taken from the coco dataset with the category "Person" along with the segmentation [9].

**Fig. 2.** Overall flow of proposed work

### 3.5. Person detection using YOLO V3

The preprocessed frames are fed into the YOLO V3 algorithm, with the input frame forward pass of the YOLO V3 object detector is performed. YOLO V3 consists of two parts the first part for feature extraction and the other part for feature detection. For feature extraction, Darknet-53 is used. Darknet-53 is a neural network that consists of 53 convolution layers. Along with Darknet-53, another 53 layers are combined to make the object detection more accurate. Hence totally YOLO V3 has 106 layers through which the pre-processed frames pass to detect the objects. The network is constructed using 3x3 and 1x1 convolution layers in that order, preceded by a skip connection. Skip connection skips some of the layers and feeds the output of the layer as input to the next layer to eliminate the degradation problem that occurs in a deep neural network. The degradation problem refers to the degradation of the performance of the model when training the deep convo networks as the number of layers increases. The next part in YOLO V3 is feature detection, the detection takes place in the three layers. For detection of smaller objects occurs at layer 82, detection of medium objects occurs at layer 94 and detection of larger objects occurs at layer 106 at 3 scales. With the input of 416 X 416 X 3, the detection occurs in the scales of 13 X 13, 26 X 26 and 52 X 52. YOLO V3 detector is generally similar to CNN, which has 106 layers.

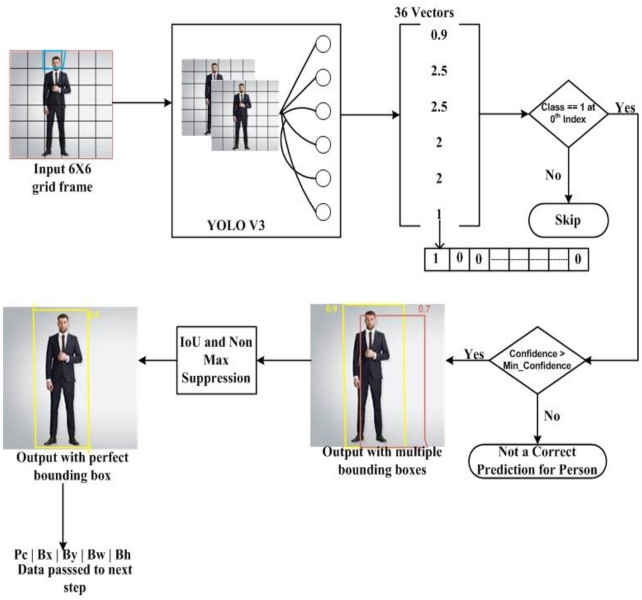


**Fig. 3.** Working of YOLO V3 model

Fig. 3 depicts the overall working model of YOLO V3 for object detection. Since the work focuses on detecting the social distance between the person, the objective of YOLO V3 is to detect the person in the given frames. This is achieved by training the YOLO V3 against the COCO dataset. The frames are divided into anchor boxes and in turn, return the vector with the information about the anchor box. Detection occurs for each anchor box. Once the feature extraction and detection is done over the anchor boxes of the frame, the output of the model gives the information about the anchor box in the form of a vector. The vector is used to determine the object in the frame along with the bounding box. The vector contains the parameters like {$P_c$, $B_x$, $B_y$, $B_h$, $B_w$, C}. $P_c$ represents the probability of confidence which means how confident the model is about the object identified in the particular grid. $B_x$, $B_y$ denotes the centroid where x, y are the coordinates of the anchor box. $B_h$, $B_w$ denotes the height and width of the bounding boxes. In order to find the bounding boxes, the centroid of the object is found by the below formula:

$$Centroid(X,Y) = {}^{Width}/_{2}, {}^{Height}/_{2} \quad (1)$$

Where,

$Width = top\ left\ point + top\ left\ point$

$Height = top\ right\ point + bottom\ right\ point$

C represents the Class that has the values 0 and 1. The class gets the value 1 when an object is detected in the anchor box. The value is 0 when there are no objects in the anchor box. 0 indicates that the particular anchor box does not contain an object related to the COCO dataset. Whereas, 1 indicates that any of the objects is detected which is available in the COCO dataset. C is generally a list with a size of 80, as the YOLO V3 is trained over the COCO dataset which has 80 different categories. According to the feature learned from the model, based on the category the object resides, the list gets the value 1.

From fig. 3, the confidence probability is 0.9 for the particular grid. The centroid points are (2.5, 2.5) and the bounding point height and width are (2, 2). The list of classes shows the first position of the list as 1 and the rest is 0. Since the feature identified is the feature related to the person, and the person category in COCO dataset is represented in the category id 1. The minimum confidence threshold is fixed as 0.3, and it is compared with the confidence obtained from the model to ensure that the detected object is a person. Then the multiple bounding boxes for the detected object are drawn along with the ground truth bounding box. A Ground truth bounding box is obtained from the training of the model. Whereas bounding boxes are the predicted boxes from the trained model. To get the accurate bounding box by avoiding the overlap of bounding boxes the techniques like Intersection over Union (IoU) and Non-Maximum Suppression (NMS) is used. The IoU threshold is fixed as 0.5. NMS removes all the minimum confidence prediction from the maximum confidence of the prediction. The maximum confident acquired from the prediction $P_s$ is compared with all other predictions $P_o$. IoU of $P_o$ and $P_s$ is calculated using equation 2.

$$IoU(BB1, BB2) = \frac{Intersection\ (BB1, BB2)}{Union(BB1, BB2)} \quad (2)$$

Once the IoU is calculated, the calculated IoU is compared with the minimum IoU threshold. If the calculated IoU is greater than IoU threshold remove the prediction from Po. This removes the bounding box that overlaps with the ground truth bounding box.

### 3.6. Detecting Social Distance

Once the YOLO V3 detects the person in the image, our work checks for the number of the person detected in each frame. The social distance is calculated only when there is more than 1 person in the frame. With the help of the centroid captured in the result of YOLO V3, Euclidean distance is calculated between all pairs of the centroids. Equation 3 shows the Euclidean distance.

$$Euclidean\ Distance = \sqrt{(X2 - X1)^2 + (Y2 - Y1)^2} \quad (3)$$

Where (X1, Y1) & (X2, Y2) are the Centroid of two persons

With the Euclidean distance calculated, the calculated distance is compared with the safer distance suggested by WHO. As of now, the safer distance is 2 meters. The frames in which the people are not in the safer distance; that is if the calculated Euclidean distance is less than the safer distance, then the people are bounded with red color in the frame. Moreover, to instruct the people to maintain the social distance, the alert is sent to the respective coordinators through SMS. The message contains the frame in which the people exceeded the safety distance. So that it is helpful to take further actions to avoid spreading COVID-19. The in-charge faculties need to create an account in the Twilio services, which is an online application. Twilio helps to send an SMS to the registered mobile number.

## 4. Experimental Results

### 4.1. Metric

mean Average Precision (mAP) is used to calculate the accuracy of the object detection in YOLO V3. Fig. 4 represents the pictorial view of the mAP metric of object detection.

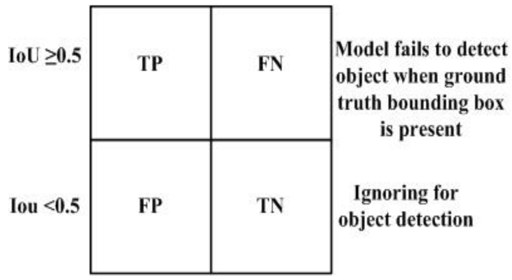


**Fig. 4.** mAP metric for object detection

Based on the IoU threshold value 0.5 in the proposed work, the metric is defined as:

**True Positive (TP):** When calculated IoU is greater than or equal to IoU threshold

**False Negative (FN):** When model fails to detect object when ground truth bounding box is present

**False Positive (FP):** When calculated IoU is less than IoU threshold

**True Negative (TN):** When model does not predict the object, this occurs commonly in the blank or empty grids where there is no object. So, the term TN is ignored for object detection.

Equation 4 and 5 shows the formula to calculate Precision and Recall values.

$$Precision = {}^{TP}/_{(TP + FP)} \quad (4)$$

$$Recall = {}^{TP}/_{(TP + FN)} \quad (5)$$

Precision and Recall are used as a measurement for evaluating the performance of the model.

### 4.2. Result Analysis

Fig. 5 to 12 shows the input video frame and its respective output after identifying the social distance. Fig. 5 shows the video taken in the evening in the outdoor location. Because the lighting in some portions of the place in the video is poor, it is difficult to identify individuals. Fig. 7 shows the video taken in the outdoor in the morning. As the lighting is good and bright the model predicts the person and finds the social distance accurately.

**Fig. 5.** Input image taken outside the college campus in evening


**Fig. 6.** Output after calculating social distance among the people outside college campus in evening

**Fig. 7.** Input image taken outside the college campus in morning

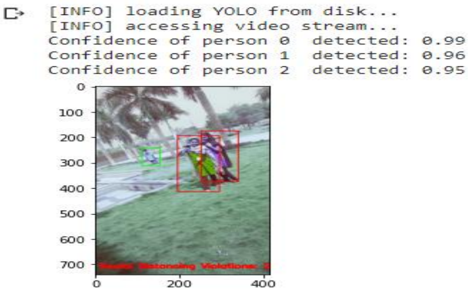


**Fig. 8.** Output after calculating social distance among the people outside college campus in morning

Fig. 9 shows the video captured inside the classroom and the students are sitting closer without maintain the social distance. Hence the model bounded the students with red bounding box accurately.

**Fig. 9.** Input image taken inside the classroom

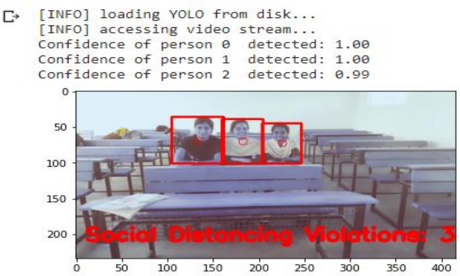


**Fig. 10.** Output after calculating social distance among the people inside classroom

In Fig. 11 as the students maintained the social distance, the model gives the green bounding box.

**Fig. 11.** Input image taken inside the college campus

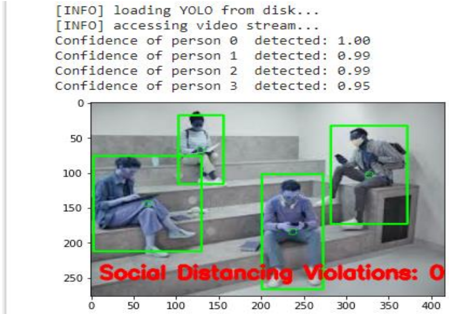


**Fig. 12.** Output after calculating social distance among the people inside college campus

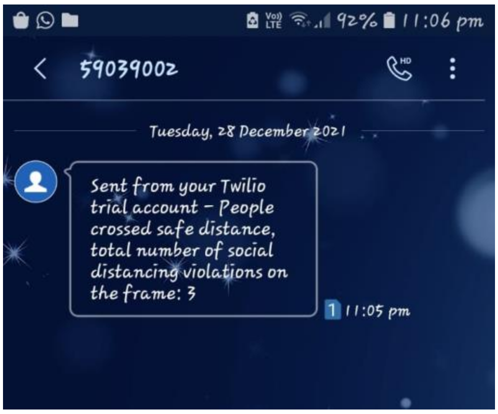


**Fig. 13.** SMS alert sent to the in-charge

Fig. 13 illustrates the SMS sent to the in-charge when the student in the class violates the safe distance recommended by WHO. The SMS sent according to the Figure 10, as the students are closer to each other without following social distance.

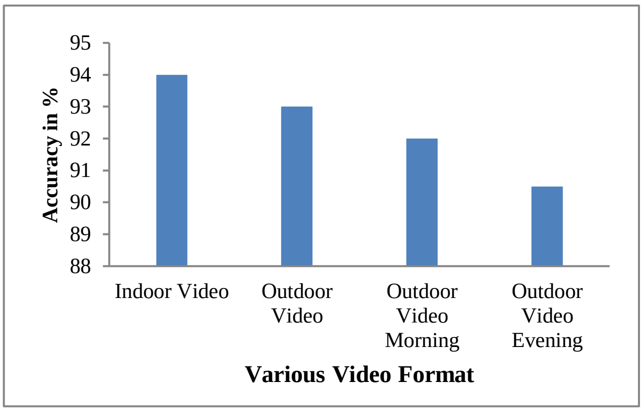


**Fig. 14.** Comparison of video taken in different areas and at different time

Fig. 14 shows the comparison chart of YOLO V3 algorithm in detecting social distancing of the college people in the different areas like indoor, outdoor, outdoor (morning) and outdoor (evening). Also, the chart shows the comparison of outdoor videos at different time namely morning and evening. This shows that the accuracy of indoor video is higher than the outdoor videos as the outdoor video has some quality loss. The quality is less because of the lighting source. Similarly, in the outdoor video taken at the morning gives relatively high accuracy than the video taken at evening because the quality of the video depends on the lighting source.

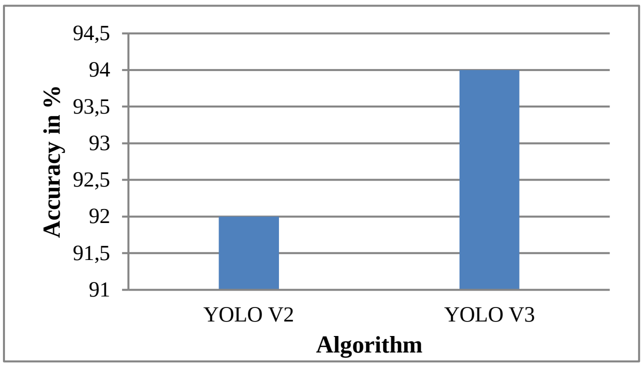


**Fig 15.** Comparison of different algorithm

Fig. 15 depicts the comparison between the two same algorithms of different version YOLO Version 2 and YOLO Version 3. This shows that YOLO V3 gives the more accurate result than YOLO V2 as it fails to detect smaller object hence the accuracy reduces. But in YOLO V3 even the small object is detected thus the accuracy is greater than YOLO V2.

## 5. Conclusion and Future Work

One of the most significant precautions in avoiding physical contact that could contribute to the spread of coronavirus is social separation. Viral transmission rates will be increased as a result of non-compliance with these rules. Python and the OpenCV library are used to create a system that detects the people who exceed the safety limits and also sends the alert or warning message to the respective coordinators through SMS. The work shows better results in terms of identifying the person and calculating the distance among the person from the video surveillance in both indoor and outdoor areas. Since covering the face with a mask also plays a vital role in preventing the attack of COVID-19, the work will extend to detect the face mask covering along with the social distancing.

## 6. References


[1] Y.C. Hou, M.Z. Baharuddin, S. Yussof and S. Dzulkifly, “Social distancing detection with deep learning model”, 8th International Conference on Information Technology and Multimedia (ICIMU), IEEE, 2020, pp. 334-338.

[2] A.H. Ahamad, N. Zaini and M.F.A. Latip, “Person Detection for Social Distancing and Safety Violation Alert based on Segmented ROI”, in 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), IEEE, 2020, pp. 113-118.

[3] K. Suresh, S. Bhuvan and M.B. Palangappa “Social Distance Identification Using Opti-mized Faster Region-Based Convolutional Neural Network”, 5th International Conference on Computing Methodologies and Communication (ICCMC), IEEE, 2021, pp. 753-760.

[4] S. Degadwala, D. Vyas, H. Dave and A. Mahajan, “Visual Social Distance Alert System Using Computer Vision & Deep Learning”, in 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), IEEE, 2020, pp. 1512-1516.

[5] S. Saponara, A. Elhanashi and A. Gagliardi, “Implementing a real-time, AI-based, people detection and social distancing measuring system for Covid-19”, *Journal of Real-Time Image Processing*, pp.1-11, 2021.

[6] I. Chakraborty and P. Maity, “COVID-19 outbreak: Migration, effects on society, global environment and prevention”, *Science of the Total Environment*, 728, p.138882, 2020.

[7] S.V. Shetty, K. Anand, S. Pooja, K.A. Punnya and M. Priyanka, “Social Distancing and Face Mask Detection using Deep Learning Models: A Survey”, in Asian Conference on Innovation in Technology (ASIANCON), IEEE, 2021, pp. 1-6.

[8] M. Qian and J. Jiang, “COVID-19 and social distancing”, *Journal of Public Health*, pp.1-3, 2020.

[9] https://cocodataset.org/#explore?id=156375, last accessed: 25.01.2021

[10] S. Srivastava, A.V. Divekar, C. Anilkumar, I. Naik, V. Kulkarni and V. Pattabiraman, “Comparative analysis of deep learning image detection algorithms”, *Journal of Big Data*, Vol. 8, No.1, pp.1-27, 2021