Journal of Positive School Psychology
2022, Vol. 6, No. 3, 6670–6679 http://journalppw.com

Minimising The Spread of Covid-19 Using Yolo V3 Algorithm

Dinesh Komarasamy[1]
Assistant Professor (SRG)
dinesh.nova@gmail.com

Rakshitha R[2]
UG Scholar
rakshitharaja2000@gmail.com

Sakthi Sri S V[3]
UG Scholar
svsakthisri@gmail.com

Nandhini N[4]
UG Scholar
nagarajnandh99@gmail.com

Dharani M K[5]
Assistant Professor
dharanianishka@gmail.com

Kowsalya K[6]
Professor
keerthi.kous@gmail.com

[1,2,3,4,5,6] *Department of Computer Science and Engineering, Kongu Engineering College, Erode, India.*

Abstract

Outbreak of Coronavirus disease have dramatically increased throughout the world and forced the world to a global disaster. The government has taken several steps to control the spread of covid-19 such as vaccination, social distancing, quarantine and so on. Researchers introduced various methodologies to ensure social distancing for controlling the spread of covid-19. To increase the accuracy of existing systems and also effectively maintain the social distancing, this work proposes an idea in which the social distancing with covid people only maintained instead of maintaining social distancing with all the people. In the proposed idea, the movement of the people is monitored using the surveillance camera. After capturing the image, the social distance is calculated using Euclidean distance and a tracking technique is used to track the people who are covid positive. The proposed idea will produce more accurate findings and allow you to calculate real measurable units. Therefore, the proposed system will be helpful to identify, count and alert the people who are near to infected people. The proposed idea also identifies the people violates the social distancing protocol and identifies the number of people around the COVID positive patient to minimize the spread of the Corona virus. The proposed system will improve the accuracy by 98% for accurately predicting the social distancing only with the people who are positive

1. INTRODUCTION:

Coronavirus Disease 2019 (COVID-19) is a virus that started in Wuhan, China, and has spread to numerous nations around the globe since December 2019. Therefore, World Health Organization (WHO) declared it as a pandemic sickness on March 11, 2020, after the virus spread to 114 countries, resulting in 4000 deaths and 118,000 active cases. Most of the industries and educational institutes remain closed in a pandemic situation [1]. The COVID-19 spreads among the people easily by touching the infected surface in which the surface contains the covid virus. After touching the infected surface, the person touching eyes or noses can easily transmit the disease. The infection of the COVID-19 is controlled by maintaining social distance. But it is difficult to maintain the social distance among the people who stayed inside an apartment, in public places and so on [2]. Moreover, young kids and children don't obey and follow the social distancing and violate the COVID-19 which results in the spread of COVID-19. Recently, all countries around the world are in a state of lockdown, forcing citizens to stay at home. However, as time passes, people will begin to visit more public places, religious sites, and tourist destinations. Although little success has been observed to date, healthcare organisations, scientists, and medical professionals are hunting for effective vaccinations and medications to combat this deadly illness. The international group is looking for new ways to halt the virus from spreading.

Researchers introduced various methodologies to control the spread of COVID-19. Due to the evolution of technologies, surveillance cameras are used to monitor the movement of people in public places. From the state of the art, the existing systems use surveillance cameras to insist the people in maintaining social distance [3]. In order to identify the person in a surveillance camera, object detection algorithms are applied in various fields such as video surveillance, object tracking, and pedestrian recognition. Deep learning is one of the most recent evolution techniques for processing the video dataset [4]. Deep learning contains several architectures to process the image and video dataset such as Region Proposals Region-based Convolutional Neural Network, Single Shot Detector, You Only Look Once, and so on. Among these, the YOLO algorithm was introduced by Joseph Redmon which has the ability to detect objects with a detection rate of better than 95 percent accuracy [5].

From the state of the art, existing systems effectively detect the social distancing among the people in public places through surveillance cameras. Even though it detects the people in public places, the existing system does not analyse the data and the existing system has only less accuracy. To improve the accuracy and to communicate with the people who are nearby infected people, this work introduces a framework to control the spread of covid-19. Thus, the ultimate goal of the project is to track the social distancing only with the individual who was affected due to covid-19. In order to recognize the person from the surveillance video footage, the framework employs the You Only Look Once Version 3 (YOLOv3) object recognition paradigm. Of which, YOLO3 detects the people who are roaming nearby people infected by coronavirus. With the help of this information, the Government takes several steps such Covid-19 affected patients come out of the Home Quarantine [6].

The following section describes the state of art.

2. LITERATURE SURVEY:

Many researchers offered many concepts for detecting the social distance between persons, using technologies such as R-CNN, YOLO, and YOLOv3. In order to improve the accuracy of the proposed work and due to the evolution of recent technologies, deep learning plays a significant role for finding social distance. To identify the social distance between people, Dushyant Kumar Singh et al. presented a method to train the CNN model which use the Inria image dataset. This method predicts the social distance using real-time surveillance footage. From the surveillance footage, this method detects human tracking and monitoring of social distance using Euclidean distance. It produces a red mark on people who are approaching the allowed limit of 6 feet, with an accuracy of 86%. The fundamental problem is that CNN does not encode the location and orientation of objects, necessitating a vast amount of training data [7]. To compensate for the above disadvantage, Juhi Shahet al., proposed a technique which uses Artificial Intelligence. Artificial Intelligence enforces to maintain social distance by alerting the people who cannot follow the social distance in public places. Even though R-CNN and Fast R-CNN increase the speed of the algorithm, it is only up to a certain extent [8].

Imran Ahmed et al. introduced a method to overcome drawbacks of CNN model which does not encode the position and orientation of objects. To train the faster R-CNN model, this method uses the Microsoft Common Objects in Context (MS-COCO) dataset. The image input is used to detect social separation using Euclidean distance, Transfer Learning, and top view human detection, with a 95% accuracy and a 0.6% false-positive rate. The drawbacks of faster R-CNN is to train the RPN using single image to extract all of the anchors in a 256-bit mini-batch. Because all samples from the same image may be correlated, it may take a long time for the network to reach convergence. [9]. In order to detect social distance in faster rate, Ning Xu et al., proposed a SSD300 model which detects objects. The objects are bounded by a red line to warn the people for who not maintain the default distance. The SSD300 model has high accuracy and recognition speed and there by this suitable for predicting multi objects. Moreover SSD is suitable for identifying small objects in a dataset [10]. Even though SSD detects small objects to a certain extent, it fails to detect tiny objects.

Therefore, YOLO algorithm has evolved for detecting tiny objects. Therefore, Sheshang Degadwala et al. proposed a new method for calculating social distance and raising alerts for those who violate social distance norms to overcome the disadvantage of all samples from a single image being correlated. The MS-COCO and PASCAL-VOL datasets are used to train R-CNN, Single Shot Detector, and YOLO based models to ascertain people and calculate distance between them, with video as input.

Using strategic relapse, YOLO predicts an objectless result for each bounding box. YOLO's biggest flaw is that it has trouble detecting small objects with high precision [11]. To overcome the problem of detecting small objects with high precision, N Pooranam et al., proposed a technique which uses YOLOv3 to detect the objects. To detect the objects and bounded by green and red bounding boxes based on the distance. Moreover, the technique generates a voice alert when people cross the safety limit. The YOLOv3 algorithm predicts more bounding boxes for an object and takes the bounding box with maximum probability to detect the object when compared to the YOLO algorithm [12]. Similarly, a novel methodology was proposed by Misbah Ahmad et al., which employs the Transfer learning methodology followed by YOLO V3 model. The dataset contains 4 films shot using overhead-mounted fisheye cameras, as well as images in the dataset containing annotations for a total of 5,837 frames. In the proposed work the input videos are splitted into training and testing set and are converted into frames. The frames are passed to transfer learning model and further passess to the YOLO V3 model. YOLO V3 model is pre-tranied over COCO dataset. Once the persons are identified in the frames, pairwise distance is calculated among the people. Number of person violating the social distance is noted when the calculated distance is less than the safety distance. The work has a 92% accuracy with transfer learning and a 98% accuracy without transfer learning. It has a 95% total tracking accuracy. The drawback is that this dataset only includes the daily activities of less than five persons [13].

To have more training and testing data, F Toschi et al., used a custom dataset to train the YOLOv3 model. The data set consists of frontal and side view photos of people. The YOLOv3 algorithm is used to detect people in the images and the drone camera helps to capture and monitor persons in public from the side or frontal perspective. The work is also expanded to include facial mask monitoring [14]. Similarly, using YOLO model Krisha Bhambani et al., presented a method to detect person along with mask. In comparison to its competitors, YOLO's identification of masked faces and human subjects is more robust and has a faster detection speed. With a 38 Frames Per Second (FPS) inference speed, the method achieves 94.75% accuracy. As this detection method takes image as an input and predicts the social distance, but failed to consider video [15]. Raghav Mangoo et al., proposed a method using YOLOv3 that considers both image and video inputs. Surveillance video is passed as an input to the model. Red bounding boxes are shown when social distancing is violated among the people. In terms of inference time and frame rate, the YOLO V3 model outperforms the SSD and Faster RCNN models [16]**. Similarly,** Akhil Kumar et al., explained a method that attempts to detect the face mask. This method makes use of a new dataset and is also highly good for detecting public health and social measures. This method was evaluated using eight different YOLO algorithm variations which recognizes people who are wearing masks incorrectly, those who are not wearing masks, and those who are not wearing masks at all. It boasts a 90 percent identification accuracy, a quick recognition speed, and a high precision, and it is appropriate for multi-target recognition. Although individuals are frequently small in surveillance video, the YOLOv3 object detection algorithm properly detects these little things as well [17].

3. PROPOSED WORK:

Researchers introduced various vaccines to control the spread of covid-19. Even though vaccines are available, new strains of virus start spreading the word. Therefore, Social distance has come out as one of the most effective techniques for preventing the spread of covid-19. The rate of spreading varies from time to time, it is necessary to follow the social distancing and isolation of infected people to control the spread of covid-19. From the state of the art and with the knowledge of existing algorithms, social distancing and monitoring of isolated people is achieved using Closed-Circuit Television (CCTV) and drones. Therefore, this work introduces a new idea called minimise the spread of covid-19 using deep learning techniques to control the spread of covid-19. The proposed idea is to track human activities in public locations during the epidemic situation. In this scenario, the social distancing is computed among the people in surveillance footage as well as monitor the people who breaches social distancing. But, it's very difficult to track all the people in public places. Therefore, this work focusses only on tracking the infected people. So, the proposed idea is to track humans using computer vision.

Figure 3.1 depicts the flow diagram of identifying the covid people. Initially, the system is trained using the dataset in which the video sequence is converted into frames. After converting into frames, the images are trained using YOLO V3 with the help of pre-trained COCO dataset. After training the dataset, the proposed system reads the input from the video dataset which is generally captured from the CCTV. The CCTV video footage is transmitted into frames and the people in the frame are detected as a bounded box. It segregates the people from the other frames. After segregation, the proposed work further segregates the image which has more than one person in the dataset. After that's, the proposed system computes the pairwise distance between the people in a dataset. The social distancing among the people is computed using a centroid tracking algorithm. After computing the social distancing, the people are indicated with a green box who are not violating the social distancing. Moreover, the people are indicated with a red bounding box for those who violated the social distance. In addition to that, the proposed work identifies the total number of people who violate social distance and the number of people in a safe region.

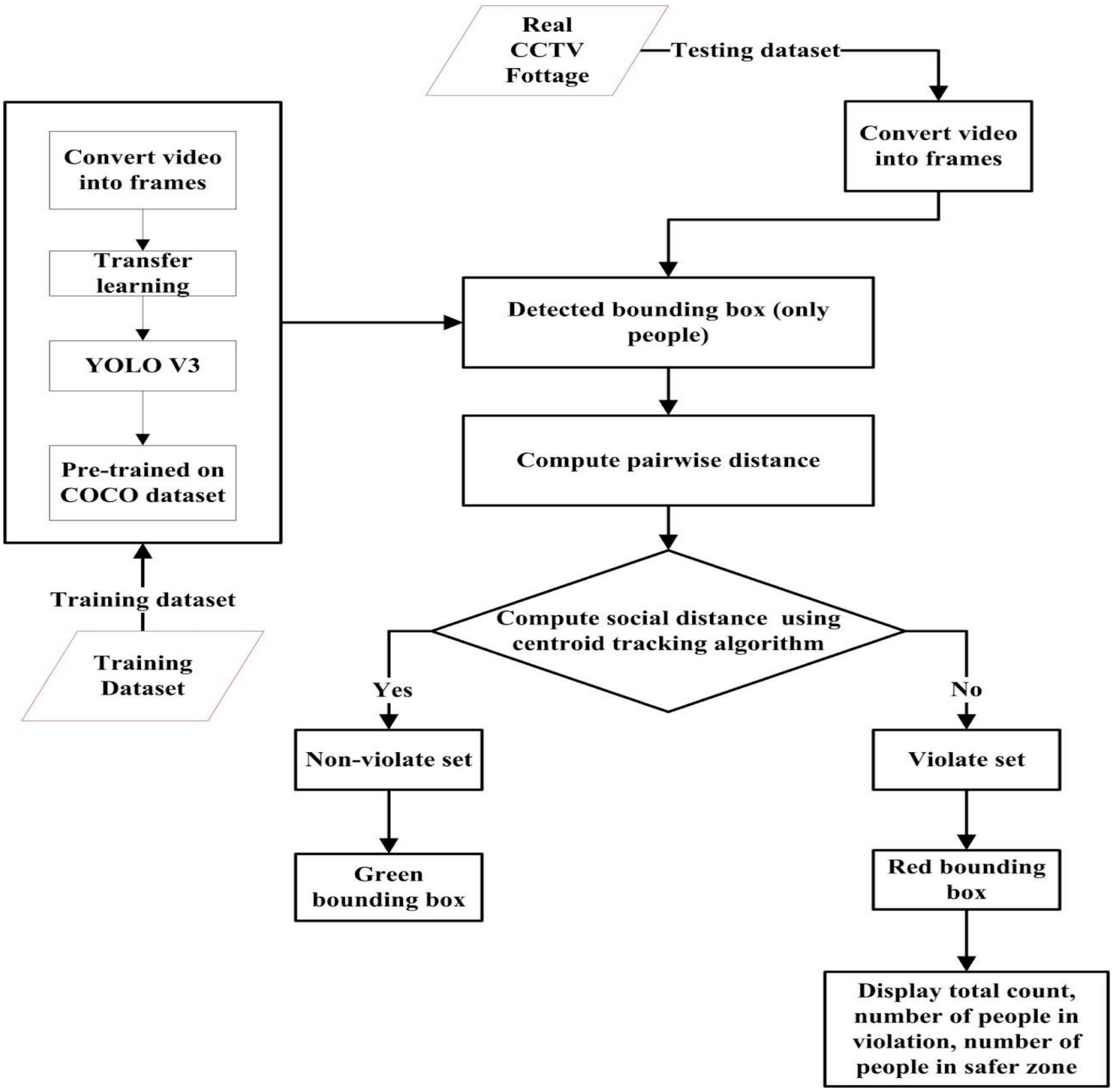

Figure 3.1 represents the flow diagram of identifying the covid people

3.1 YOLO Algorithm

You only look once (YOLO), a single convolutional neural network that differs from previous detectors. The YOLO object detection algorithm is used to identify the human from any object. The frame or an input video is provided with a pre-trained and pre-initialized YOLO object identification model, as well as CNN output layer names, person index to detect only the person in the video frame. In the detection module, the input frame's height and breadth are determined by extracting the frame dimensions. Scale the bounding box coordinates back to the image's size, YOLO really delivers the bounding box's centre (x, y)-coordinates. The top-left corner of the bounding box are derived using the centre of (x, y)- coordinates, and the values of bounding box coordinates, centroids, and confidences of a detected object are then updated. The centroid of an identified person is the centre of the bounding box, which is derived by extracting the centre of the bounding box. Finally, verify that the frame contains at least one detection, extract the bounding box coordinates, and update the result list accordingly. The confidence of each person detection, bounding box of each person, and centroid of each person are returned to the social distance module for further processing and to determine whether or not the people are maintaining a healthy distance.

Figure 3.1.1 Detection of people after Non- Maxima Suppression

Figure 3.1.1 which represents detection of people after Non-Maxima Suppression depicts detection of people after applying Non-Maxima Suppression. To ignore duplicate detections of the same object, non-max suppression is helpful by declining multiple predictions for the same object. By taking a list of predicted boxes for the same image, it only considers the detection of an object which has been detected as duplicate with maximum confidence.

3.2 Social distance module

In order to place the bounding box, a single-stage network is developed for the whole given image and employed in the architecture. Convolution layers are employed for feature extraction, while fully linked layers are used for class predictions in this design. The input frame is divided into a zone of S*S, also known as grid cells, during human identification. The grid cells are connected with bounding box estimation and class probabilities, as well as predicting whether the person's bounding box's centre is in the grid cell or not.

Figure 3.2.1 Detection of human bounding boxes.

Figure 3.2.1 which represents Detection of human bounding boxes, depicts that bounding boxes can be seen around every human and centroid is also marked for each and every person.

A list is created for maintaining the centroid information of all the people. The colour of the bounded box is set to green for the person who are not violating social distance. Otherwise, the colour of the bounded box is set to red and the information is appended with the violation set. The tracking algorithm tracks the persons who violate social distance threshold. The bounding box will move along with the persons in different frames. The bounding box is used to compute the centroid position of the object.

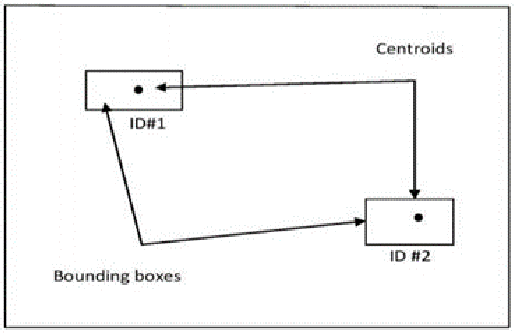

Figure 3.2.2 Centroid tracking

Figure 3.2.2 which represents centroid tracking, depicts how an object's centroid is computed for each consecutive frame using the bounding box notion that one covered earlier. However, assigning a new unique ID for each object detection may obstruct object tracking. To solve this, the algorithm checks whether one can correlate the centroid of the new object with that of an existing object, which one can do by calculating the Euclidean distance between the two objects using the formula in Equation (4.1):

$$d(x,y)=\sqrt{(x_1+y_1)^2+(x_2-y_2)^2} \quad \textbf{(3.1)}$$

It is very difficult to track the person in a real environment. Therefore, the proposed work assigns a ID to the objects detected and the ID remains the same for the object even though the object moves to another frame. The new objects can be registered by drawing bounding boxes around them and assigning new IDs to them. One can determine the centroid of the bounding box and store it in the list after that, and one track the object's movements by calculating the minimum distance using Euclidian's method, as stated before.

Fig 3.2.3 Input frame from video

Figure 3.2.3 illustrate the input frames from video. The frames contain images which are extracted from a video sequence. In the proposed YOLO model contains approximately 45 frames per 1 second video.

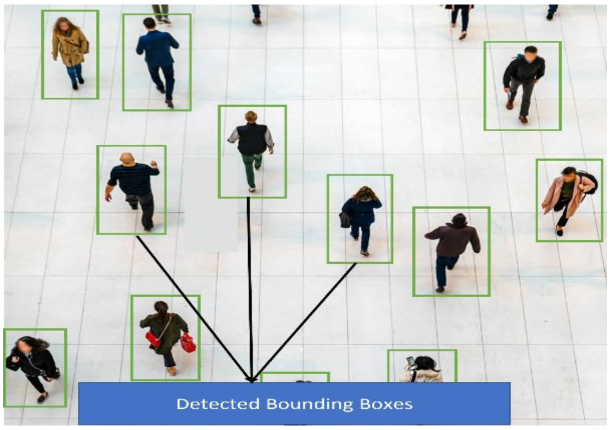

Fig 3.2.4 Bounded boxes using YOLO V3 algorithm.

Figure 3.2.4 shows bounded boxes using YOLO V3 algorithm which provides the bounded boxes only for the humans in a given frame. Moreover, the YOLO V3 algorithm can be applied for real time CCTV footage.

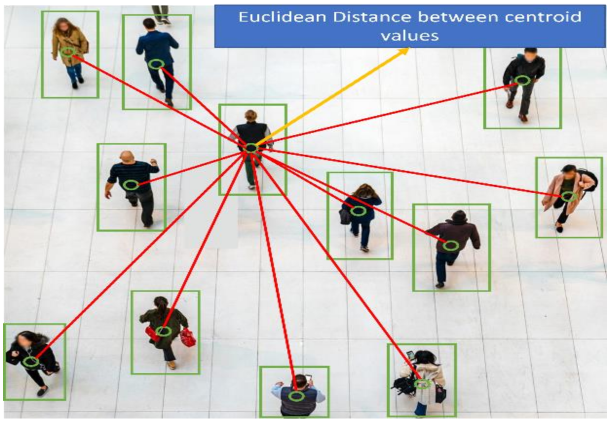

Fig 3.2.5 Represents Euclidean distance between centroid coordinates

Figure 3.2.5 denotes the Euclidean distance between centroid coordinates value (x, y). The centroid value is initially computed for each bounded box. After computing the centroid value, the distance between the bounded boxes is measured by calculating the Euclidean distance between the centroid coordinates. This distance is computed for all the pairs of bounded boxes. Based on the Euclidean distance the colour of the bounded box varies in the image.

3.3 PEOPLE COUNTING MODULE

Figure 3.3.1 Counting people in a frame.

Figure 3.3.1 which represents counting of people in a frame shows the total number of people in a particular frame when passing a video sequence as an input. Using OpenCV, a proposed people counting system based on object detection that automatically detects a human from video was constructed. For object detection, OpenCV provides a powerful tool. For face detection, the Viola-Jones detector was used to recognize frontal human faces in video sequences.

3.4 RESTRICTION MODULE

Figure 3.4.1 Number of social distance violation

As shown in figure 3.4.1, which represents the number of social distance violations, a hall is put up in the limitation module where a group of people can enter with or without preserving social distance. Using a people counting method, the number of persons entering the hall is also calculated. The green zone is defined as one or two persons entering the hall with adequate distance between them, or two or more people with appropriate distance between them. When two or more people completely violate the required distance, they are regarded to be in the Red Zone, and an Alarm will be triggered.

The number of individuals entering and exiting the venue is also taken into consideration by the Restriction module. In order to maintain the specified spacing in the hall, it also sounds an alarm when the number of people exceeds the allocated number of people. When the alarm goes off, it not only alerts the person who is breaking the rules, but it also alerts the complete team that is present in that particular hall, which will help to alert the surroundings.

4. RESULTS AND DISCUSSIONS

4.1 EXPERIMENTAL SETUP:

The proposed work is implemented in a google colab. Colab is a free cloud service based on Jupyter Notebooks which allows to write python code. Colab is one of the best platforms for solving real world problems using Morden techniques [18].

An indoor data collection containing video sequences acquired from an above view is used for social distance monitoring. The data gathering is split into two parts: training and testing, which account for 70% and 30% of the total. The freedom of movement of people is unrestricted throughout the setting. People in the scenario move around freely, with radial distance and camera location influencing their visual appearance. It can be seen from the example frames that a human's physical appearance is not uniform, and that people's heights, stances, and scales differ across the data set. OpenCV is utilised in the implementation.

Figure 4.1. Prediction of Small Objects and showing violations

Figure 5.2 which represents violations shown in a frame of a video input and prediction of small objects and showing violations are some of the sample outputs of the proposed framework. The outcomes of the experiment are divided into two subsections: the first discusses the testing results of the pre-trained model, while the second discusses the detection model's results after performing transfer learning and training on the overhead data set. The model is evaluated against the same video sequences for comparison.

Different video sequences are used to evaluate the testing outcomes. The people in the video sequences are free to move above. Because the model only examines the human (person) class, a pre-trained model can only detect an object that looks like a human. As long as a social distance threshold is maintained, people are identified with green rectangles in sample frames. When more than one individual enters the scenario, the model is likewise put to the test. Following person detection, the distance between each detected bounding box is calculated to determine whether or not the person in the picture breaches the social distance, as illustrated in the sample photos [19]. However, in other circumstances, the person's look changes, resulting in miss detections by the model. The cause for miss detection could be that an individual's look from an overhead view change as the pre-trained model is applied, which could be deceptive for the machine. The results of utilising transfer learning to monitor social distance The transfer learning process is used to improve the detection model's accuracy. The model is also trained with 500 sample frames using an overhead data set. The model is trained with an epoch size of 40 and a batch size of 64.

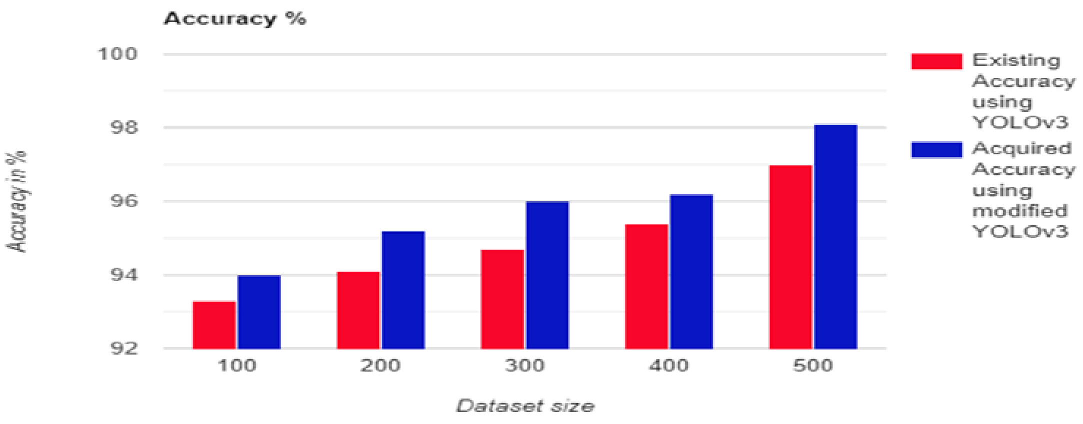

Fig 4.2 Accuracy of model (YOLOv3)

Figure 4.2 Accuracy of model (YOLOv3) of the existing system and acquired accuracy. In the proposed work, the YOLO V3 algorithm is modified by changing and setting Graphics Processing Unit (GPU) as true. As a result, modified YOLO V3 algorithm has the ability to clearly capture the small objects in an image and thereby the bound box is drawn for all the objects in a frame simultaneously. Thereby, the proposed work increases the accuracy to 98% by effectively detecting all the small objects in a frame. The false detection rate of many deep learning models is exceptionally low, at 0.7–0.4 % without any training, demonstrating deep learning effectiveness.

5. CONCLUSION AND FUTURE WORK

After the COVID-19 epidemic, the need for self-responsibility became clear. The scenario would primarily focus on embracing and following the safeguards and rules that the WHO has enforced, with the individual taking full responsibility for himself rather than the government. Because COVID 19 spreads by close contact with infected people, social distancing would surely be the most significant aspect. An effective method for supervising huge crowds is critical. The main objective of the proposed work is to track the human beings who violate social distance. The proposed work modifies the YOLOv3 algorithm by setting the GPU as true. Since the GPU is set to true, the proposed work can clearly visualise the small objects [20]. The proposed work clearly captured the smaller objects, the accuracy of the proposed work is increased to 98% compared to the existing algorithms.

Reference

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

2. Gabutti, G., d’Anchera, E., Sandri, F., Savio, M. and Stefanati, A., “Coronavirus: update related to the current outbreak of COVID-19”. *Infectious diseases and therapy*, *9*(2), pp.241-253, 2020

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

4. Pooranam, N., Sushma, P.P., Sruthi, S. and Sri, D.K., “A Safety Measuring Tool to Maintain Social Distancing On COVID-19 Using Deep Learning Approach”, In *Journal*

of Physics: Conference Series, Vol. 1916, No. 1, p. 012122, IOP Publishing, 2020.

5. Du, J., “Understanding of object detection based on CNN family and YOLO”, In Journal of Physics: Conference Series, Vol. 1004, No. 1, p. 012029, IOP Publishing, 2018.
6. Zhao, L. and Wan, Y., “A New Deep Learning Architecture for Person Detection”, In 2019 IEEE 5th International Conference on Computer and Communications (ICCC), pp. 2118-2122, IEEE, 2019.
7. Ansari, M. and Singh, D.K, “Monitoring social distancing through human detection for preventing/reducing COVID spread”, *International Journal of Information Technology*, *13*(3), pp.1255-1264, 2021.
8. Shah, J., Chandaliya, M., Bhuta, H. and Kanani, P., “Social Distancing Detection Using Computer Vision”. In *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 1359-1365). IEEE, 2021.
9. Ahmed, I., Ahmad, M. and Jeon, G., “Social distance monitoring framework using deep learning architecture to control infection transmission of COVID-19 pandemic”, *Sustainable cities and society*, *69*, p.102777, 2021
10. Qin, J. and Xu, N., “Reaserch and implementation of social distancing monitoring technology based on SSD”, *Procedia Computer Science*, *183*, pp.768-775,2021.
11. Degadwala, S., Vyas, D., Dave, H. and Mahajan, A., ” Visual Social Distance Alert System Using Computer Vision & Deep Learning”. In *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 1512-1516). IEEE, 2020.
12. Pooranam, N., Sushma, P.P., Sruthi, S. and Sri, D.K., “A Safety Measuring Tool to Maintain Social Distancing On COVID-19 Using Deep Learning Approach”. In *Journal of Physics: Conference Series* (Vol. 1916, No. 1, p. 012122). IOP Publishing, 2021.
13. Ahmed, I., Ahmad, M., Rodrigues, J.J., Jeon, G. and Din, S., “A deep learning-based social distance monitoring framework for COVID-19”. *Sustainable Cities and Society*, *65*, p.102571, 2021.
14. Pouw, C.A., Toschi, F., van Schadewijk, F. and Corbetta, A., “Monitoring physical distancing for crowd management: Real-time trajectory and group analysis”. *PloS one*, *15*(10), p.e0240963, 2020.
15. Bhambani, K., Jain, T. and Sultanpure, K.A., “Real-time face mask and social distancing violation detection system using yolo”. In *2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC)* (pp. 1-6). IEEE, 2020.
16. Magoo, R., Singh, H., Jindal, N., Hooda, N. and Rana, P.S., “Deep learning-based bird eye view social distancing monitoring using surveillance video for curbing the COVID-19 spread”. *Neural Computing and Applications*, *33*(22), pp.15807-15814, 2021.
17. Kumar, A., Kalia, A., Verma, K., Sharma, A. and Kaushal, M., “Scaling up face masks detection with YOLO on a novel dataset”. *Optik*, *239*, p.166744, 2021.
18. Alves, F.R.V. and Vieira, R.P.M., “The Newton fractal’s Leonardo sequence study with the Google Colab”. *International Electronic Journal of Mathematics Education*, *15*(2), p.em0575, 2019.
19. 1He, Y., Zhu, C., Wang, J., Savvides, M. and Zhang, X., “Bounding box regression with uncertainty for accurate object detection”. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 2888-2897), 2019.
20. Kłusek, A., Topa, P., Wąs, J. and Lubaś, R., “An implementation of the social distances model using multi-GPU systems”. *The International Journal of High Performance Computing Applications*, *32*(4), pp.482-495, 2018.
21.