



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Special Issue 2, April 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Monitoring Covid-19 Social Distancing with Person Detection and Tracking via Fine Tuned Yolo V3 Deep sort Techniques

Parvatham Jithendra Reddy, Praveen Reddy Chejeti, A. Annie Michael

U.G. Student, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

Associate Professor, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

ABSTRACT: The deadly spread of the coronavirus disease 2019 (COVID-19) has triggered a worldwide crisis. The global epidemic is nevertheless currently in progress, potentially causing acute breathing problems that have resulted across many deaths. Though vaccines have recently been produced and cannot be made accessible to everyone in a timely basis, social distancing is the only effective method for curbing this pandemic. Empowered by this idea, this article proposes a deep learning-based framework for automating the activity of tracking and measuring distancing through the use of video footage. In the proposed model, humans are distinguished from the background using the YOLO v3 object detection model, and the labelled individuals are tracked using the Deepsort approach and indicate it using anchor boxes and assigned IDs. In addition to this, violation index used to quantize the nonadopting of social distancing protocol. The usage of YOLO object detection model will produce accurate results in surveilling people even in the state obstruction, slight visibility, and brightness variabilities.

KEYWORDS: Social Distancing, Machine Learning, YOLO V3, Deep Learning, CNN, COVID-19, Video Capture, Bounding boxes, People detection

I. INTRODUCTION

The COVID-19 virus is an airborne pathogen that primarily infects people within a close proximity by inhaling small droplets created by speech, sneezing, and coughing. The first coronavirus-related disease, COVID-19, was discovered in Wuhan, China, in late December 2020. WHO declared this to be a pandemic on March 11 after it spread to 114 countries, had 118,000 active cases, and resulted in 4000 fatalities. Around 3,519,901 confirmed cases and 247,630 fatalities were reported as of May 4, 2020, globally. The development of effective drugs and vaccines for this deadly virus is being pursued by numerous healthcare organisations, medical professionals, and scientists, but there has been no reported progress to date. The world community is now under pressure to find new strategies to halt the spread of this contagious virus. All affected nations are under strict orders to implement social distancing, which is hailed as the best spread inhibitor in the current situation.

This study proposes a method to identify social distance among people gathered in any public space, with the goal of supporting and mitigating the coronavirus pandemic with the least possible disruption to economic endeavours. In order to minimise or stop the transmission of COVID-19, it is best practise to use the term "social distance." Reduced contact between potentially infected persons and healthy people is its main goal.

In order to follow social distancing guidelines, it is advised that people keep a minimum distance of 6 feet between each other. Due to the possibility that individuals with mild or no symptoms may unintentionally carry corona infections and spread them to others, a recent study suggests that social isolation is a crucial containment strategy and is necessary to prevent SARSCoV-2. demonstrates that the best way to prevent contagious body interaction and subsequently lower the infection rate is through appropriate social distance. This lower peak may undoubtedly be compatible with the current healthcare infrastructure and assist in providing patients fighting the coronavirus pandemic with better facilities.

Epidemiology is the study of the causes and mechanisms behind the transmission of infectious diseases. Mathematical models are almost always the first choice when researching epidemiological phenomena. The Kermack and McKendrick classic SIR model, which was first developed in 1927, is the ancestor of well almost all models.

The SIR model and its extensions by the deterministic system have been the subject of numerous research projects, and as a result, many researchers have studied stochastic biological systems and epidemic models. The rate and mode of the causative virus's transmission are the most important aspects to take into account when treating or finding ways to stop the virus' spread in the community for infectious diseases like respiratory diseases. There is currently no well-known medicine that can be used to treat COVID19, despite the efforts of numerous medical organisations and pandemic researchers to create vaccines. In order to prevent the spread of infection, precautions are taken worldwide.

The number of infected and recovered individuals, denoted as I and R, respectively, can be used to calculate the social distance parameter, $a(I, R)$, which was recently proposed by Eksin et al. [1]. On April 16, 2020, a company called Landing AI [2] led by one of the most well-known figures in AI, Dr. Andrew Ng, announced the development of an AI tool to track social distance at work.

The company claimed in a article that by examining real-time video streams from the camera, the upcoming tool could determine whether individuals are keeping a safe physical distance from one another. It is further asserted that this tool can be easily integrated with the security cameras already in place at various workplaces to maintain a safe distance between all employees. Three steps—calibration, detection, and measurement—shown in a brief demo are used to keep track of social distance.

In recognition of their timely initiative in this ground-breaking field to support the fight against the COVID -19, Gartner, Inc. named Landing AI as one of the Cool Vendors in AI Core Technologies on April 21, 2020. This is the driving force behind the authors' current work, which aims to evaluate and compare how well-known object detection and tracking schemes track social distance.

II. RELATED WORK

International travel-related nonpharmaceutical interventions (NPIs), such as traveller screening, travel restrictions, and border closures, are frequently incorporated into national influenza pandemic preparedness plans, according to [3]. To find evidence of their effectiveness, we carried out systematic reviews. We discovered 15 studies in total. NPIs may be able to prevent the spread of the influenza virus, according to some studies. There was no research on exit screening, and there was no indication from the evidence that screening of incoming travellers would significantly slow the spread of pandemic influenza. One study found that small Pacific islands were successful in preventing the importation of pandemic influenza during 1918–19 by completely closing their borders. Other studies suggested that travel restrictions could delay the start of local transmission and slow global spread. According to the scant evidence available, NPIs related to international travel would be ineffective in containing a pandemic of influenza and would require a significant investment in terms of funding.

One of the community mitigation strategies that may be suggested during influenza pandemics, according to [4], is social estrangement. By putting people farther apart physically or reducing their tendency to congregate in places where there are lots of people, like workplaces or schools, social distancing can lessen the spread of viruses. To evaluate the evidence that social isolation in non-healthcare settings lessens or slows influenza transmission, we carried out a systematic review.

In [5], In a simple susceptible-infected-recovered (SIR) model, the initial rate at which infected cases increase predicts the outbreak's long-term trajectory. In contrast, people may alter their behaviour and take precautions during outbreaks in the real world. The initial rate of spread and the final case count may no longer be closely related as a result. Here, we test this hypothesis by contrasting the dynamics resulting from a straightforward SIR epidemic model with those resulting from a modified SIR model in which people cut back on contacts in relation to the number of active or cumulative cases.

Despite the initial disease spread rates for both models being nearly identical, dynamics with behaviour change show significantly lower final case counts. We demonstrate how the critical role played by individual behavioural changes in this variation in final size projections. Incorporating behaviour change into iterative forecast models is justified by these results, as well. In order to update models with and without changing behaviour, we therefore suggest using a Kalman filter as part of iterative forecasts. Sequential predictions made using a simple SIR model struggle to account for behaviour change in the ground truth outbreak despite repeated observations, but predictions made using a modified

SIR model are able to do so. These results demonstrate how beneficial it is to include behaviour change in baseline epidemic models.

We examine the fundamental dynamical characteristics of a stochastic SIR epidemic model incorporating media coverage in [6]. In order to investigate the asymptotical stability and global stability of equilibria of deterministic model, we first discuss the positivity and boundedness of solutions of the model within a deterministic environment. Second, we demonstrate that the stochastic model has a singular global positive solution, and that, under specific circumstances, this solution oscillates around the equilibria of the deterministic model. To illustrate our analytical findings, we give some numerical simulations at the end. Since there is no automated system for object-based social distancing monitoring, the government will have to designate someone to keep an eye on people to determine whether or not social distance is being maintained. The issues with the current system include the difficulty in locating those who do not uphold the established social hierarchy and the time and effort required to gather resources and human energy.

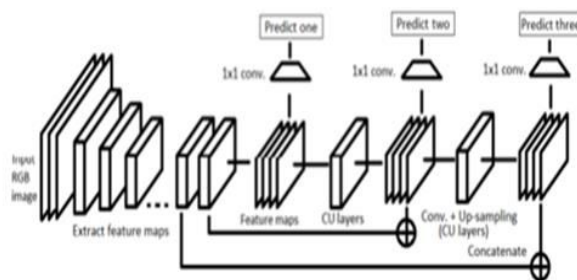
III. METHODOLOGY

In our proposed system, each individual is identified in real-time with the aid of bounding boxes, and we implement an effective real-time deep learning-based framework to automate the process of monitoring social distance via object detection and tracking approaches. This solves the issues in the existing system. When using a pair-wise vectorized approach to compute the closeness property, clusters or groups of people that satisfy it can be found using the generated bounding boxes. Cost effectiveness is a benefit of this system for this purpose. The maintenance and reduction of human efforts are also simple.

A. YOLO for object detection

YOLO [7] is another competitor of SSD for object detection. By examining the image just once, this technique can determine the kind and location of an object. Instead of using classification to assign class probabilities to the anchor boxes, YOLO views the object detection problem as a regression task. Multiple bounding boxes and class probabilities are simultaneously predicted by a single convolutional network in Fig. 1.

Fig 1: Schematic representation of YOLO v3 architecture.



YOLO has primarily three iterations: v1, v2, and v3. GoogleNet's (Inception network) object classification algorithm served as a model for YOLO v1. 24 convolutional layers and 2 fully connected layers make up this network. YOLO v1 only utilises a reduction layer, followed by the Inception modules used by GoogleNet.

Later, the idea of YOLO v2 [8] is put forth with the hope of significantly enhancing accuracy while speeding up the process. Darknet-19, a network with 19 convolutional layers, 5 max pooling layers, and an output softmax layer for object classification, is used as the backbone network in YOLO v2. With notable gains in mAP, FPS, and object classification score, YOLO v2 outperformed its forerunner (YOLO v1).

YOLO v3, on the other hand, performs multi-label classification using logistic classifiers rather than softmax as in YOLO v1 and v2. Darknet-53 was suggested by Redmon et al. as the backbone architecture for YOLO v3 to extract feature maps for classification. Darknet-53, in contrast to Darknet-19, consists of residual blocks (short connections) as well as the upsampling layers for concatenation and additional network depth. To solve the issue of not being able to detect small objects effectively, YOLO v3 generates three predictions at various scales for each spatial location in an image [9]. Calculating the objectness, boundary box regressor, and classification scores allows for the monitoring of each prediction. Loss function: The overall loss function of YOLO v3 consists of localization loss (bounding box regressor), cross entropy and confidence loss for classification score.

A. Deep Sort

Deepsort is a method for tracking specific objects in a video that is based on deep learning [10]. The Deepsort tracking algorithm is used in the current study to identify people in the surveillance footage. In order to predict the associated trajectories of the objects of interest, it uses patterns learned through the detection of objects in the images, which are later combined with temporal information. In order to facilitate further statistical analysis, it keeps track of each object under consideration by mapping unique identifiers. In addition, Deepsort is helpful for addressing related issues like occlusion, multiple viewpoints, non-stationary cameras, and annotating training data.

The Kalman filter and the Hungarian algorithm are used for efficient tracking. The Kalman filter can forecast future positions based on the current position and is used recursively for better association. For association and id attribution, a Hungarian algorithm is used to determine whether an object in the current frame is the same as an object in the previous frame. A Faster RCNN is initially trained to identify people, and for tracking, a linear constant velocity model [11] is used to describe each target in an eight-dimensional space as follows:

$$x = [u, v, \lambda, h, x_c, y_c, \lambda_c, h_c]^T \quad (1)$$

where a is the aspect ratio, h is the height of the image, and (u, v) is the centroid of the bounding box. The other variables are the variables' individual velocities. Later, the linear observation model and the standard Kalman filter are used with constant velocity motion, where the bounding coordinates $(u, v, \lambda, \text{ and } h)$ are taken as direct observations of the object state. The total number of frames are computed for each track k , starting from the most recent successful measurement association a_k . The counter is increased when the Kalman filter predicts positively, and it is later reset to 0 when the track is associated with a measurement. Additionally, those objects are deemed to have left the scene if the identified tracks are older than a predetermined maximum age, and the corresponding track is then removed from the set of tracks. Additionally, new track hypotheses are started for each unidentified track of novel detected objects that cannot be mapped to the existing tracks if there are no tracks available for some detected objects. The new tracks are categorised as indefinite for the first three frames up until a successful measurement mapping is computed. The tracks are removed from the track set if the measurement attempts to map them are unsuccessful.

By taking into account the motion and appearance information and using the Mahalanobis distance calculated between them, the Hungarian algorithm is used to solve the mapping problem between the recently arrived measurements and the predicted Kalman states.

B. Euclidean distance and violation thresholds

The Euclidean formula is used at this point in the project to calculate the distance between each detected centroid pair using the enclosed bounding boxes and the centroid information after individuals are identified in the thermal videos. To calculate real-world metrics, the scaling factor is added to the calculation distance. The Euclidean measurement defines the distance as 6 feet (approximately 180 cm). In order to determine distance assumptions, two distinct predefined minimum social distance violation rules are used. Given this, we set two distinct thresholds for these violations as well as yellow and red colours for the bounding boxes that were detected.

The first threshold violation is defined as "Alert," which is highlighted in yellow, and the second assumption of violation is defined as "Risk," which is highlighted in red. If the gap between the detected bounding boxes is greater than 6 feet, the bounding box's colour is initially initiated as green. The bounding boxes' colour is updated to yellow if the distance between the detected individuals is between 7 and 5 feet (the first threshold value "Alert") and to red if it is less than or equal to 5 feet (the second threshold value "Risk"), indicating that the social distance between them has not been maintained.

D. Work Process

The necessary steps taken to create a framework for tracking social distance are included in this section.

1. To identify and track the person in a video, fine-tune the trained object detection model.
2. The surveillance footage is fed to the trained model. For each identified person, the model creates a set of bounding boxes and an ID.
3. Each individual is linked to a three-dimensional feature space with the coordinates (x, y, d) , where (x, y) represent the centroid coordinates of the bounding box and d represents the depth of the individual as seen by the camera.

$$d = ((2 * 3.14 * 180) / (w + h * 360) * 1000 + 3) \quad (2)$$

where w is the width of the bounding box and h is the height of the bounding box [12]. 4. For the set of bounding boxes, pairwise L2 norm is computed as given by the following equation.

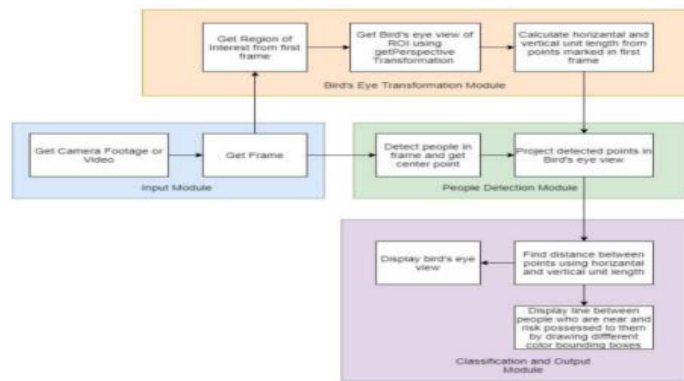
$$\|D\|_2 = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (4)$$

5. After assigning the neighbours for each person who meets the closeness sensitivity, the dense matrix of L2 norm is used. With many trials, the spatial location of the person in a frame, which can be anywhere between (90, 170) pixels, is used to dynamically update the closeness threshold.

6. Any person who satisfies the closeness property is given a neighbour or neighbours, who together form a group and are shown in a different colour coding from other individuals.

7. The emergence of groups suggests a breach in the social distance norm, which is quantified by the following:

- Take into account ng as the total number of groups or clusters found, and np as the total number of people found nearby.



- The violation index, vi= np/ng.

Fig 2: Architecture

The run is the initial and important module in our project after run the project process will start. Monitoring social distance is second module in our project. Here, we are observing the social distance between the people. While monitoring, The persons with red symbol are not following distance, with green symbol are following the social distance.

To produce results with greater accuracy, data is preprocessed and collected via the internet. After that, a model is created using deep learning algorithms after the dataset has been trained. When a user uploads an image, an algorithm is used to extract features from the image and detect the body. Later, we categorise the user's distance.

The suggested framework makes use of the YOLO v3 object detection model to separate people from background objects and the Deep Sort approach to track the people who have been identified using bounding boxes and assigned IDs. Further comparisons between the YOLO v3 model's output and other well-known state-of-the-art models are made in terms of mean average precision (map), frames per second (FPS), and loss values determined by object classification and localization. These models include faster region-based CNN (convolution neural network) and single shot detector (SSD).

The pair wise vectorized L2 norm is then calculated using the three-dimensional feature space created by using the centroid coordinates and bounding box dimensions to identify the clusters of individuals defying the social distance order. The non-adoption of social distance protocol is quantified using the violation index term. According to the results of the experimental analysis, the YOLO v3 with Deep Sort tracking scheme produced the best results for tracking social distance in real time, with a balanced map and high FPS.

The company claimed in a brief article that by examining real-time video streams from the camera, the upcoming tool could determine whether people are keeping a safe physical distance from one another. Additionally, it is stated that this tool can be easily integrated with current security cameras that are already present at various workplaces to maintain a safe distance between all employees. Three steps—calibration, detection, and measurement—shown in a brief demo are used to keep track of social distance. Gartner, Inc. recognised Landing AI as a Cool Vendor in AI Core Technologies on April 21, 2020 in recognition of their timely initiative in this ground-breaking field to support the fight



8th National Conference on Advanced Computing Technologies (NCACT'23)

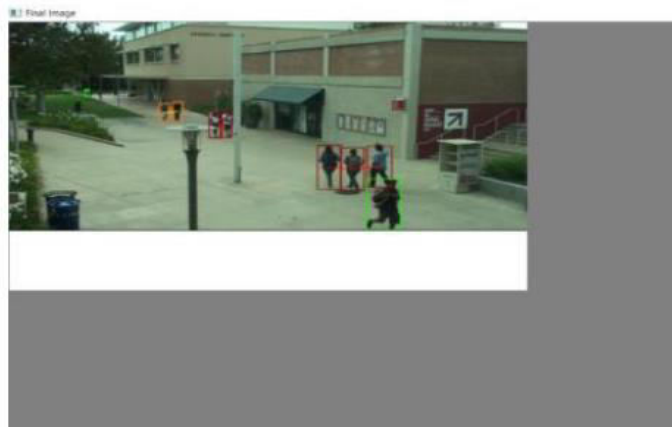
Organized by

Department of CSE, Velammal Engineering College, Chennai, India

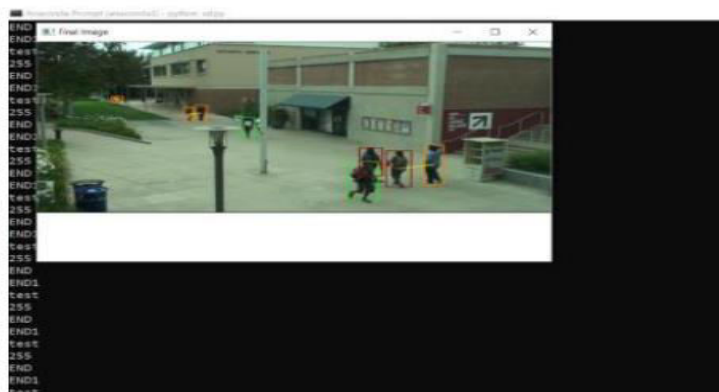
against the COVID-19. As a result, the authors of the current work were motivated to evaluate and compare how well-known object detection and tracking systems tracked social distance.

IV. EXPERIMENTAL RESULTS

The suggested approach contributes to limiting the spread of the coronavirus by offering an automated method for tracking and identifying instances of people breaking social distance norms. The suggested people detector model can identify and follow both stationary and moving individuals in public areas, as well as monitor and uphold COVID-19 and later social distancing standards. This method can be used anywhere that has CCTV cameras of any kind. This makes it easier to categorise areas based on risk, add more restrictions to some areas, allow only limited movement through certain zones, or redesign the environment's scene to make it a safer place. The restrictions have different rules depending on the environment and situation. People can be divided into a variety of categories, such as safe and high-risk.



The proposed methodology generates a processed frame with the identified individuals contained within bounding boxes, as well as simulating statistical analysis showing the total number of social groups displayed by the same colour encoding and a violation index term computed as the proportion of the number of individuals to the number of groups. The frames show 3, 2, 2, and 2.33 as the violation index. For later analysis, the frames with detected violations are saved with the timestamp.



V. CONCLUSION

The use of person detection and tracking techniques, specifically with fine-tuned YOLOv3 and Deep SORT, has shown promising results in monitoring social distancing during the COVID-19 pandemic. By analyzing video footage, these techniques can accurately identify individuals and track their movements, allowing for the detection of violations

of social distancing guidelines. Overall, these techniques can help to mitigate the spread of COVID-19 by identifying and preventing potential hotspots of transmission. However, it is important to note that these technologies must be used responsibly and with consideration for privacy concerns. As with any new technology, continued research and evaluation are necessary to ensure its effectiveness and ethical implementation.

REFERENCES

- [1] K. Paarporn, J. S. Weitz and C. Eksin, "Systematic biases in disease forecasting—the role of behavior change," *Epidemics*, vol. 27, pp. 96–105, 2019.
- [2] P. Alto, "Landing AI Named an April 2020 Cool Vendor in the Gartner Cool Vendors in AI Core Technologies," 2020.
- [3] J. Y. Wong, J. Xiao, M. W. Fong, H. Gao, E. Y. Shiu, S. Ryu, and B. J. Cowling, "Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings social distancing measures," 2020.
- [4] N. Zviedrite, F. Ahmed, and A. Uzicanin, "Effectiveness of workplace social distancing measures in reducing influenza transmission: a systematic review," *BMC public health*, vol. 18, no. 1, p. 518, 2018.
- [5] H. Zhao and M. Zhao, "Asymptotic behavior of global positive solution to a stochastic sir model incorporating media coverage," *Advances in Difference Equations*, vol. 2016, no. 1, pp. 1–17, 2016.
- [6] S. Divvala, R. Girshick, J. Redmon, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.
- [7] A. Farhadi and J. Redmon "Yolo9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.
- [8] A. Sonawane, "Medium : YOLOv3: A Huge Improvement," 2020.
- [9] A. Bewley, D. Paulus, and N. Wojke, "Simple online and realtime tracking with a deep association metric," in *2017 IEEE international conference on image processing (ICIP)*. IEEE, 2017, pp. 3645–3649.
- [10] A. Bewley, and N. Wojke "Deep cosine metric learning for person re-identification," in *2018 IEEE winter conference on applications of computer vision (WACV)*. IEEE, pp.748-756, 2018.



SJIF: 8.423



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INTERNATIONAL CENTRE



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details