



HAL
open science

Live Social Spacing Tracker Based on Domain Detection

Divyank Agarwal, Monark Mehta, Nirmala Paramanandham

► **To cite this version:**

Divyank Agarwal, Monark Mehta, Nirmala Paramanandham. Live Social Spacing Tracker Based on Domain Detection. 5th International Conference on Computational Intelligence in Data Science (ICCIDS), Mar 2022, Virtual, India. pp.115-123, 10.1007/978-3-031-16364-7_9 . hal-04381301

HAL Id: hal-04381301

<https://inria.hal.science/hal-04381301v1>

Submitted on 9 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Live Social Spacing Tracker Based on Domain Detection

Divyank Agarwal¹, Monark Mehta¹, Nirmala Paramanandham^{1*}

¹ Vellore Institute of Technology, Chennai, INDIA

divyank199602@gmail.com, monark3293m@gmail.com, nirmalavp.ece@gmail.com*

Abstract. Corona virus Disease- 2019 (COVID-19) is caused by infection with the severe acute respiratory syndrome (SARS) coronavirus. In this COVID 19 pandemic, the virus spread bringing the whole world to downfall. This disease can spread through the slightest touch, breathing the same air or using same basic things like clothes, hairs, combs, etc. These viruses can live for hours even without a host body. To prevent the spread, the world was put on lock-down, and people were constrained to their homes, but human life cannot go on without interaction. We need a better way of preventing the spread of this type of disease. Hence a detector is proposed for measuring the social distance between the people. This social distancing detector can track the people who are not following social distancing norms, then that person can be tracked down or the person is marked in red and triggers the warning. In this paper we will be using Computer Neural Network (CNN) to process our images and videos because CNN is a type of artificial neural network that is used in recognition of large pixel videos or images. The proposed technique can be the perfect way to help the person while people carry out their daily tasks. From the evaluation parameters it is proven that the proposed technique yields better results when compared to the state-of-the-art techniques.

Keywords: COVID-19, person detection, social distancing.

1 Introduction

Coronavirus and its variants like delta, omicron and IHU cause many deaths in all over the world. This virus does not show symptoms based on age. Anyone can be affected by this virus at any age. In this current pandemic, coronavirus cases are increasing day by day. Each country and state has been affected by this virus which has caused the demolition of many families. This virus can be spread over an infected person's mouth or nose to the recipient's mouth or nose. This could happen when the infected person sneezes, cough, speak, sing or breathe [1].

World Health Organization (WHO) have considered the implementation of public health and social measures (PHSM) which can result in decreasing the number of deaths and spreading of coronavirus. PHSM is not a rule of law; it has been implemented based on the cause of spreading coronavirus. It includes social distancing, avoiding crowded places, and cleaning hands regularly using any antibacterial soaps or sanitizers [2].

Coronavirus can enter to our body when the people are not maintaining social distance from the affected person. To avoid this type of disease, the best practice can be following the social distancing guidelines. Consider our own country India, which is the second most populated country after China. Due to more population, the public places are becoming more crowded. In turn, it is increasing the COVID-19 cases rapidly.

In this paper, a live social distancing tracker tool is proposed. It will detect the people who are not following the social distancing rules. This can be done by computing the distance between the nearby people.

The technique has been specially utilized for the more crowded area like markets, bus stands, and other public areas as the probability of spreading COVID-19 is more in these particular areas. In addition to that the proposed technique can also be used for any type of restricted area or property like a gas station where entry of any outsider is prohibited; this will detect the person and alert the security in that area.

The following sections discuss the related works, proposed work, results and discussions and the conclusions.

2 Research Background

Yibing ma et al [3], have proposed a fire alarm with person detection and thermal camera. This system uses the YOLOv3-tiny algorithm for detecting a person. It also checks the temperature of the environment, if the temperature is high in the region, then it automatically starts detecting the people and alerting them.

In [4], the authors have proposed a technique for the detection of people from a distance of several miles using a high-quality camera lens. They have used a two-stage approach for detecting a person and distinguish between them using a distanced captured video. A convolution neural network (CNN) is used for distinguishing a person carrying a bundle or a person carrying a long arm. Several methods have been proposed for background subtraction to robustly detect the people in thermal video across different environmental conditions. They have used thermal cameras which can detect the images even at the night, this is done by capturing the thermal radiation reflected from objects which are captured by the camera, and hence the person is detected even in low light conditions. In the developed technique, the shadow of the person will not be considered in the detection process unless the person is stationary for a long time [5]. Video surveillance has become a necessary part in most of the places such as malls, railway station, bus station and no entry places. Detecting a person in these areas has a major role. In [6], a simple and fast real-time person tracking system was developed based on the SiamMask network for video surveillance. This system uses deep learning-based algorithms. Initially the person is detected using SiamMask technique then additional training is done using transfer learning for obtaining the accuracy of 95%. The technique [7] has proposed for multispectral pedestrian detection in the field of computer vision. The authors have detected the pedestrian by leveraging

RGB-thermal pairs with convolutional neural networks. Further, they have studied the effect of noise on the output by creating a clean version of ground-truth annotations. Current industries required a minimum separation of workers from the machine which is a good practice in industries for the safety of workers [8]. It was performed using a sensor-based approach and this system uses a background model for representing an aspect of the environment. Ideally, the safety zone surrounds by a green mark and the danger zone is surrounded by a red mark. However, people have not considered robots and if they enter in the red region then the system alerts the workers. Siezen et al [9] have implemented a technique using magnetic fields for keeping the track of social distancing between people. They made a wearable device, it measures the distance between people by the strength of the coil present in. If the distance is less than two meters then the wearable device alerts the person to follow the social distancing guidelines. Nowadays many digital companies involve Bluetooth technology which they have declared the most effective technique for tracking persons in contact. Scott et al [10], proposed a system that tracks the people based on Bluetooth in smartphone. It calculates the distance between the smartphones and then track the social distancing, but this study may not be feasible in all cases. Yadhav et al [12], have conducted a mean accuracy and maximum validation performance comparison between different models for detecting the objects. Lin et al [13], have performed an experiment on using the dataset of Common Object in Context (COCO) and the obtained results are shown in Table 1.

Table 1. Mean average Precision score results of different pre-trained model

Model	Minimum validation	Mean accuracy
MobileNet SSD	19.3	19
Resnet-101 Fast R-CNN	31	30.9
Resnet-101 Fast R-CNN	33.2	33
Inception Resnet Fast R-CNN	34.7	34.2
MobileNet R-FCN	13.8	13.4

Yadhav et al [12] have discussed the different dataset for analyzing Execution time and number of object detected in each frame for the different dataset and the highest accuracy between the different model and datasets is shown in Table.2.

Table.2. Comparison work for different pre trained model

Model name	Execution time(s)	Highest Accuracy	Object Detected
SSD Mobilenet V1 COCO	219.58	94%	2
SSD Inception V2 COCO	298.22	97%	2
Inception V2COCO Faster RCNN	420.71	99%	3

Resenet V2 Atrous Mask RCNN Incep- tion	6008.02	99%	5
--	---------	-----	---

On the basis of the results obtained, the data show that the MobileNet SSD has a faster recognition rate and this model runs faster for object recognition. Hence, the proposed work chooses the MobileNet SSD as a recognition model.

3 Proposed Work

In this work, a live spatial detection technique is proposed. Initially, the live video will be converted to frames. The people standing nearby will be detected in each and every frame by the proposed technique. When the people are not maintaining the social distancing, they will be surrounded by 'Red mark' or else they will be surrounded by 'Green mark'. The method will be achieved by calculating the center point of the people and a rectangular box type border will be created around them and if the box overlaps the other nearby box, then the alarm or warning will be given to the people. This algorithm can be divided into four functions where each function has its definitive algorithms and properties.

Check function: In this function, the person class which is identified by the mobile net SSD is bounded by a box. This box may get overlapped to prevent the infinite bounding box which is due to the simultaneous change in the position of the object. After the boxes are bounded, the centroid of the bounded box is calculated and it is given a unique ID and the same is done with other objects after the centroid is calculated.

Image Process function: It is the major and most significant function as this function is responsible for the generation of input and output files which are in the form of mp4 formats. This function takes the input file and splits it into many frames which are then processed individually and then the output frames are joined to create an output video in the same video format.

Input parameters: Here, the algorithm is initialized, and input is given as 'filename' while the output file is generated and stored in the .mp4 format as 'opname'. The input video is in form of mp4, or live feed and the output video also can be downloaded in the same format. Centroid rotation is used for finding the distance between a person. When the separation distance is less than the set limit, then the person is shown with red mark. The person finder function is used for finding and identifying the person of the figure in the taken videos.

Main function: this is the final part of the algorithm where the output processed frames generated by the algorithm are joined once again to create an output mp4 file. This output file contains the time stamp which tells the speed and duration of the video, the .time() functions calculate the duration of the video and process which can be

used for finding the efficiency of the proposed algorithm. As the frames generation and object class identification go simultaneously, this process time depends greatly on the hardware used, speed of the processor, and also the duration of video like in my example – a 3-second video took almost 150 sec to be processed with the social distancing tool.

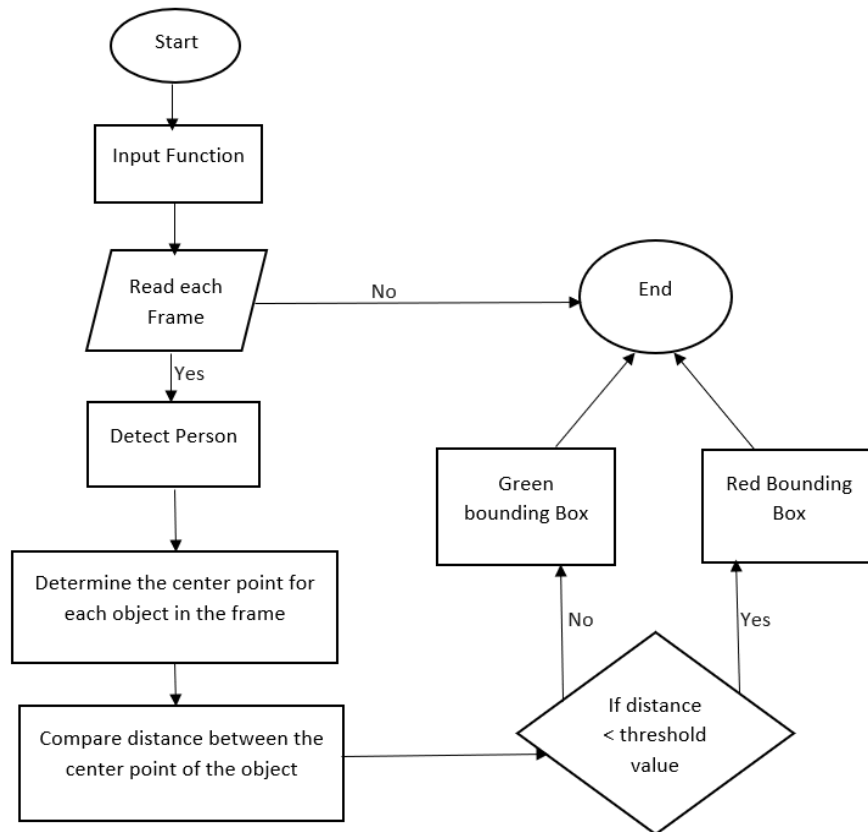


Fig.1. Architecture for social distancing model

A. Object Detection Model

In this paper, the machine learning and person identification tool is used with deep learning algorithms. While there are many other algorithms with higher accuracy, the Mobile net SSD came out to be the fastest, with the least hardware requirements it has a detection rate of 87.7% with a ratio of 80:20, training to test data [11].

B. Determine person location

In determining the person and identification of the location of the object, some of the important factors are kept in mind. For example, the camera should be at a high point where it can capture large and stable video with minimum contact and disturbance. The camera quality should be above average so that at least the person in close prox-

imity can be seen and identified. It should also have some kind of protection or covering for preventing the outer disturbance to be captured as video.

C. Calculate the center point in the bounding box

The person bounded by the bounding box is first given an id and then the centroid of the person is calculated by using the formulae:

$$C(x, y) = \frac{(xmin+xmax), (ymax+ymin)}{2} \quad (1)$$

Each of the minimum and maximum values for the corresponding width, $xmin+xmax$ and height $ymax+ymin$, of the bounding box will be used to calculate the center point of the bounding box.

D. Distance between bounding box

After the bounding box has been calculated and the centroid has been determined, the distance between two centroids are calculated using the equation (2)

$$D = ((xmax - xmin)^2 + (ymax - ymin)^2)^{1/2} \quad (2)$$

This distance is the maximum distance between two objects or persons. It is compared with two meters which is the distance to be maintained for preventing the spread of COVID-19, thus the person can be then marked according to them as a green or red box.

3.1 Algorithm

Step 1: Give the video (mp4) or live feed as an input.

Step 2: The video is divided into frames.

Step 3: Each frame is processed, and the person can be detected in each frame.

Step 4: A bounding box is created and the distance between each bounding box is calculated.

Step 5: The value is compared with the minimum distance.

Step 6: The box is either displayed red or green based on the measured distance.

Step 7: Output processed frames are then converted to a video.

4. Results and discussions

In this work, python 3 and openCV libraries are used for image processing techniques. For calculation, two videos are taken on bridge and a 4-way road. Camera recordings are processed, while some random frames are taken from the output video to create a proper dataset. We have taken four different benchmark dataset as Self video, Bridge video, 4-Lane video and PETS-2009 video whose accuracy has been listed in table 3. This dataset has been analyzed based on the multitude i.e., if this system can detect the person.

To measure the accuracy of the system, values for true positive (TP)- number of persons correctly identified who are following the social distancing norms, true negative (TN)- number of persons identified who are not following the social distancing norms, false positive (FP)- number of persons correctly identified and not following the social distancing norms, false negative (FN)- number of persons not correctly identified and following the social distancing norms.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

The accuracy is then taken from these environments to give the final results and comparison with the different models and algorithms.

Table.3. Accuracy of the proposed algorithm.

NO	VIDEO	TP	TN	FP	FN	ACCURACY
1	SELF-video	2	2	0	0	100%
2	BRIDGE	6	9	2	1	83%
3	4- LANE	10	16	10	4	65%
4	PETS-2009	0	7	1	2	70%

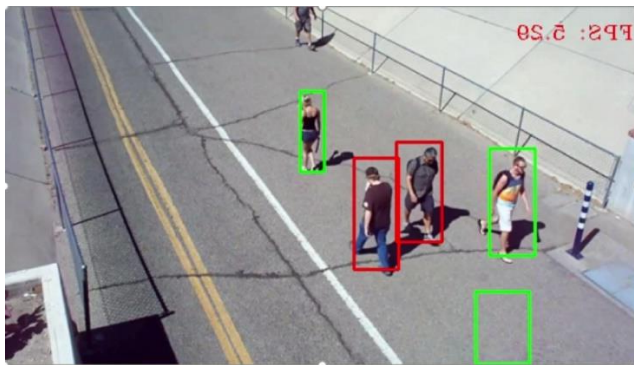


Fig.2. BRIDGE Output

Based on the findings and results, it can be proven that the proposed technique yields better results when compared to other algorithms.

The obtained results for different dataset are tabulated in Table 3. For the self-video, the proposed technique achieves 100 % accuracy. But when the other datasets are considered, the accuracy is getting decreased because of the increased density of crowd.



Fig.3. PATS-2009 Output

Figure.3 and 4 shows the model detection for the dataset PATS-2009 and 4- lane. The detection model accuracy is reduced when the person is too close because both the persons are identified as a single person. The technique achieves very high efficiency in tracking static or less crowded places. The algorithm is implemented in Intel I5, 7th generation. The computational complexity is reduced and accuracy can be improved by removing the redundant frames. After analyzing all four dataset we understood that in figure 1 and 2 has less crowd so the efficiency of the system is 100% but when coming to figure 3 and 4 this dataset contains more crowd so the system is unable to detect person who are very close to each other it detects them as a single person so this can be a major issue of the existing system.



Fig.4. Four Lane Output

5. Conclusion and Future Work

The major precaution to prevent COVID- 19 and its variant is maintaining social distance and avoiding physical contact of the person who is affected from the disease. Viral transmission rates will be increased because of non-compliance with these rules. To achieve the desired functionalities, proposed technique is implemented using Python and the OpenCV library. The first component detects social distancing infractions, while the second feature detects violations of entering restricted locations. The accuracy of both characteristics has been verified. The proposed technique has met all its objectives and it can be proven from the obtained results. In future, we are planned to remove the redundant frames in the dataset which will improve the accuracy and reduces the computational complexity.

References

1. World Health Organization, "Coronavirus Disease 2019," Coronavirus disease (COVID-19) pandemic, 2021. https://www.who.int/health-topics/coronavirus#tab=tab_1
2. World Health Organization, "Considerations for implementing and adjusting public health and social measures in the context of COVID- 19," Interim guidance 14 June 2021.
3. Yibing Ma, Xuetao Feng, Jile Jiao, Zhongdong Peng, Shenger Qian, Hui Xue and Hua Li, "Smart Fire Alarm System with Person Detection and Thermal Camera," International Conference on Computational Science, vol. 12143, pp. 353–366, 2020.
4. H. Wei, M. Laszewski, N. Kehtarnavaz, "Deep Learning-Based Person Detection and Classification for Far Field Video Surveillance," Department of Electrical & Computer Engineering, University of Texas at Dallas, Published November 2018.
5. James W. Davis and Vinay Sharma, "Robust Background-Subtraction for Person Detection in Thermal Imagery," Dept. of Computer Science and Engineering, Ohio State University Columbus, OH 43210 USA.
6. Imran Ahmed and Gwanggil Jeon, "A real-time person tracking system based on SimMask network for intelligent video surveillance," Journal of Real-Time Image Processing, pp. 1803–1814 (2021).
7. Chengyang Li, Dan Song, Ruofeng Tong, Min Tang, "Multispectral Pedestrian Detection via Simultaneous Detection and Segmentation," State Key Lab of CAD&CG, Zhejiang University Hangzhou, China Published 14 August 2018.
8. Paul Rybski, Peter Anderson-Sprecher, Daniel Huber, Chris Niess, Reid Simmons, "Sensor fusion for human safety in industrial workcells," IEEE/RSJ International

Conference on Intelligent Robots and Systems, Published 24 December 2012.

9. Sizhen Bian, Bo Zhou, Hymalai Bello, Paul Lukowicz, "A Wearable Magnetic Field Based Proximity Sensing System for Monitoring COVID-19 Social Distancing," German Research Center for ArtificialIntelligence Kaiserslautern, Germany 2020.
10. Scott McLachlan, Peter Lucas, Kudakwashe Dube, Graham A Hitman, Magda Osman, Evangelia Kyrimi, Martin Neil, Norman E Fenton, "Bluetooth Smartphone Apps: Are they the most private and effective solution for COVID-19 contact tracing?".
11. Wattanapong Kurdthongmee, "A comparative study of the effectiveness of using popular DNN object detection algorithms for pith detection in cross-sectional images of parawood," School of Engineering and Technology, Walailak University 222 Thaibury, Thasala, Nakornsithammarat, 80160, Thailand- February 2020.
12. N. Yadav and U. Binay, "Comparative Study of Object Detection Algorithms," Int. Res. J. Eng. Technol., pp. 586–591, 2017, [Online]. Available: www.irjet.net.
13. T. Y. Lin et al., "Microsoft COCO: Common objects in context," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8693 LNCS, no. PART 5, pp. 740–755, 2014, doi: 10.1007/978-3-319-10602-1_48.