



Category: STEM (Science, Technology, Engineering and Mathematics)

ORIGINAL

Enhancing Public Safety: A Real-time Social Distance Monitoring with Computer Vision and Deep Learning

Mejora de la seguridad pública: Una monitorización social a distancia en tiempo real con visión por computador y aprendizaje profundo

Sivakumar Karuppan¹ , Krishnaprasath V T¹ , Pradeep V² , Sruthi S Madhavan¹ 

¹Department of Artificial Intelligence and Data Science, Nehru Institute of Engineering and Technology, Coimbatore, Tamil Nadu, India.

²Department of Information Science and Engineering, Alva's Institute of Engineering and Technology, Moodbidri, Karnataka, India.

Cite as: Karuppan S, Krishnaprasath VT, Pradeep V, Sruthi SM. Enhancing Public Safety: A Real-time Social Distance Monitoring with Computer Vision and Deep Learning. Salud, Ciencia y Tecnología - Serie de Conferencias 2024;3:616. <https://doi.org/10.56294/sctconf2024616>.

Submitted: 08-12-2023

Revised: 19-02-2024

Accepted: 09-03-2024

Published: 10-03-2024

Editor: Dr. William Castillo-González 

ABSTRACT

In spite of the fact that the COVID-19 epidemic has lately afflicted millions of individuals all over the world, the number of people who are being affected is continuing to climb. In response to the ongoing pandemic scenario throughout the world and in an effort to stop the virus from further disseminating, a number of governments have initiated a number of groundbreaking preventative measures. One of the most effective methods for warding off the spread of infectious diseases is maintaining adequate social distance. In the context of a real-time top view environment, the purpose of this study survey is to propose the use of a social distance framework that is built on deep learning architecture as a preventative strategy for maintaining, monitoring, managing, and lowering the amount of physical connection that occurs between individuals. In order to identify people in the photographs, we made use of a number of different deep learning detection models, including R-CNN, Fast R-CNN, Faster-RCNN, YOLO, and SSD. Because of the significant differences between the top and bottom views of a human's appearance, the architecture was trained using the top view human data set. After that, the Euclidean distance is utilised to derive a pair-wise distance estimate between the individuals depicted in a picture. Using the information obtained from a detected bounding box, one may determine where the centre point of a single detected bounding box is located. A violation threshold is constructed, which is determined by the information of a person's distance to a pixel and determines whether or not two people are in breach of social distance.

Keywords: Social Distance; Coronavirus; Disease; Monitoring System; Detection Model and Deep Learning.

RESUMEN

A pesar de que la epidemia de COVID-19 ha afectado últimamente a millones de personas en todo el mundo, el número de afectados sigue aumentando. En respuesta al actual escenario pandémico en todo el mundo y en un esfuerzo por impedir que el virus siga propagándose, varios gobiernos han puesto en marcha una serie de medidas preventivas pioneras. Uno de los métodos más eficaces para evitar la propagación de enfermedades infecciosas es mantener una distancia social adecuada. En el contexto de un entorno de vista cenital en tiempo real, el propósito de esta encuesta de estudio es proponer el uso de un marco de distancia social construido sobre una arquitectura de aprendizaje profundo como estrategia preventiva para mantener, supervisar, gestionar y reducir la cantidad de conexión física que se produce entre los individuos. Con el fin de identificar a las personas en las fotografías, hicimos uso de una serie de diferentes modelos de detección de aprendizaje profundo, incluyendo R-CNN, Fast R-CNN, Faster-RCNN, YOLO, y SSD. Debido a las diferencias significativas entre las vistas superior e inferior de la apariencia humana, la arquitectura se

entrenó utilizando el conjunto de datos humanos de la vista superior. A continuación, se utiliza la distancia euclidiana para obtener una estimación de la distancia por pares entre los individuos representados en una imagen. A partir de la información obtenida de un recuadro delimitador detectado, se puede determinar dónde se encuentra el punto central de un único recuadro delimitador detectado. Se construye un umbral de violación, que se determina a partir de la información de la distancia de una persona a un píxel y determina si dos personas incumplen o no la distancia social.

Palabras clave: Distancia Social; Coronavirus; Enfermedad; Sistema de Monitorización; Modelo de Detección y Deep Learning.

INTRODUCTION

Around 8 million individuals all around the world have been infected with this potentially fatal virus as of right now. Medical experts, scientists, and researchers have been putting in countless hours to create a vaccine and therapy for this potentially fatal condition. The global community is doing research into alternative approaches and processes in the hope of controlling the virus's spread.^(1,2) In light of the current circumstances, maintaining a social distance has been considered as one of the most effective techniques for reducing the global transmission of this illness. It has to do with increasing the amount of distance between individuals in crowded areas while simultaneously decreasing the amount of direct physical touch that occurs between them. By limiting the amount of close physical contact that individuals have with one another, we may bring the curve depicting reported cases down and reduce the likelihood of viral transmission. Maintaining a healthy social distance is essential for those who have a greater likelihood of developing a serious disease as a result of COVID-19.⁽³⁾



Figure 1. Sample images of social distance monitoring position

The likelihood of the virus spreading and the severity of the illness can both be mitigated by the implementation of social surveillance systems. As seen in Figure 1, the early implementation of an adequate social distance mechanism has the potential to play a critical part in minimising the peak of the pandemic and lowering the rate at which infections are spread. It is recognised as one of the most reliable techniques to control the transmission and spread of the severe illness because of the large number of confirmed cases arising each day, which results in a shortage of medical fundamentals.⁽⁴⁾ As can be seen in figure 1, the rate of viral transmission over the whole planet decreases. The graphs make it abundantly evident that the management of social distance is what controls the spread of the virus, as the number of individuals who are ill continues to decrease while the number of those who have recovered continues to rise. The number of infected persons and cases, in addition to the strain placed on healthcare institutions, can be reduced and brought under control by implementing efficient social distancing mechanisms in public spaces. It achieves this by ensuring that the

number of infected cases does not exceed the capacity of the healthcare personnel,⁽⁵⁾ which in turn lowers the risk of people losing their lives to the disease.

The remaining content of the article is organised as described in the following paragraphs. In the second part of this article, we will investigate the many methods that have been developed to calculate social distance. Discussion of the deep learning framework that was developed for tracking social distance may be found in Section 3. The findings of the investigation as well as potential future directions are discussed in Section 4.

Literature Survey

Researchers have made a variety of efforts to develop effective methods for measuring social distance. They used a variety of machine and deep learning-based technologies to track and measure people's social distance. To measure the distance between persons, the authors used a variety of clustering and distance-based algorithms; however, they largely focused on the side or frontal view perspectives.⁽⁶⁾

Many countries have developed technical ways to restrict the spread of covid since the commencement of Covid-19. The server can detect Coronavirus infection in a specific location using GPS integrated in cell phones. Several countries throughout the world have developed and are using such applications. China and Israel looked deep and examined the individuals in the contact lists of the infected people. The Aarogya Setu⁽⁷⁾ is an application developed in India that uses GPS to locate a covid infected patient in a certain region. It calculates the risk of a healthy individual becoming infected if they visit the same site as an infected person and are within 2 metres of each other. Bluetooth signals are used to calculate the distance between an infected person and an individual.

Nicola et al.⁽⁸⁾ investigated the relationship between social distancing strictness and the region's economic state and reported that simple steps from this exercise may be taken to avoid a huge epidemic. As a result, numerous governments have developed remedies based on advanced technology to prevent pandemic loss up to this point. Many advanced governments are using Global Positioning System technology to track the travels of suspects and infected persons. Bradley et al.⁽⁹⁾ provided an overview of numerous developing technologies, including Bluetooth, Wi-Fi, GPS, mobile phones, computer vision, image processing, and machine learning, that played an important role in a variety of conceivable social distance circumstances. Few used drones and other monitoring systems to locate people gatherings.

Kobayashi et al.⁽¹⁰⁾ conducted a systematic evaluation of COVID-19 illness preventive and treatment methods. For COVID-19, Ahmed et al. developed a social distance monitoring system. Bhattacharya et al.⁽¹¹⁾ investigated the United States' position during the epidemic. The study indicated that if decision-makers supported and adopted the social distance practise at an earlier stage, human health would not be damaged at such a high rate. Al-Khazraji⁽²⁰⁾ proposed a smart monitoring physical distances system that can track people's physical distances and provide them with appropriate feedback. The proposed technology counts how many people are in a given region and calculates their distances. The system then sends out warning signals to the person who is not keeping to the social distance. As we described above, various researchers have been undertaken in order to create a better and more effective social distance monitoring system, still lacking in handle the large amount of datasets. In order to tackle this issue, we have to research on the direction of the social distancing monitoring system using deep learning techniques.

Deep Learning based Social Distance Monitoring System

The structure of a social distance monitoring system that is based on deep learning is going to be the topic of discussion in this part. The entirety of the study may be broken down into two distinct components: human detection and social distance monitoring. The IP camera that was installed in the top view recorded movies, which were then transformed into frames, after which the frames were analysed by the human detection module, which makes use of a deep learning architecture to recognise human bounding boxes. In addition, the data from the bounding box is utilised by a social distance monitoring module. This module looks for infractions and reports them to the surveillance unit so that they may be processed.

A comprehensive and in-depth explanation of a deep learning-based social distance monitoring system is presented in figure 2. The entire architecture is comprised of a number of different modules, but the two most important ones are the bounding box detection and the social distance violation modules. The human boundary box is determined by the first module, and whether or not there has been a breach of social distance is evaluated by the second module. Samples for both training and testing are generated through the use of recorded top-view human data collecting.

As shown in figure 3, deep learning paradigm for human detection models are generally categorized as region based detection (two-stage detection approaches) and regression based detection (one stage detection approaches).^(12,13,14,15)

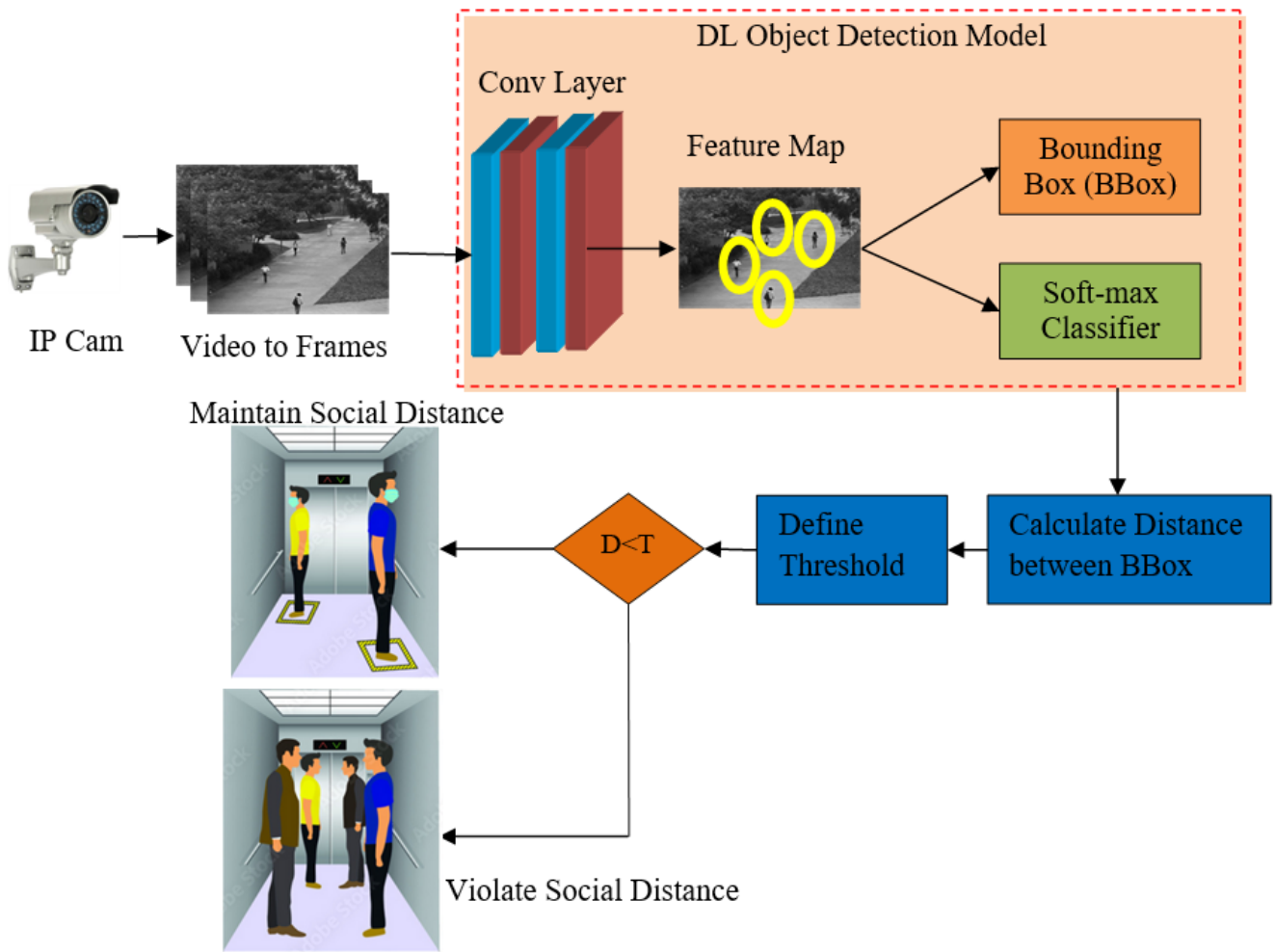


Figure 2. Deep learning based social distance monitoring framework

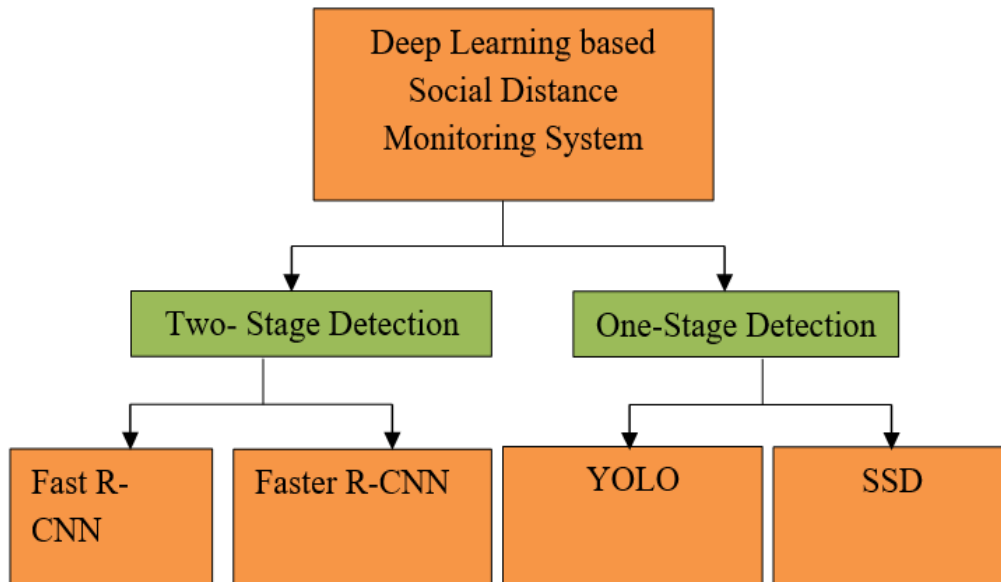


Figure 3. Deep Learning Object Detection Techniques

Human Object Detection

There is a wide variety of deep learning available for human objection detection modules, such as Region based CNN (R-CNN), Fast R-CNN, Faster R-CNN, You Only Look Once (YOLO), and Single Shot Multi-box Detector (SSD).^(16,17,18,19) A deep learning architecture known as Faster RCNN is utilised for the purpose of doing human

detection from the top view perspective. The detecting process may be broken down into two distinct steps.

In the initial step of the process, a Region Proposal Network, also known as an RPN, is utilised to develop region proposals or feature maps for the sample picture. In the first step of the feature mapping process, convolution layers are used to analyse the data. In this stage, a Fast-RCNN architecture, also known as a deep convolutional network, is utilised in order to choose certain object qualities from the many area recommendations. Last but not least, the softmax and bounding box classifiers are put to use in the top view scenario in order to locate the object (person). Convolutional Neural Network (CNN) layers are utilised in order for the RPN to segment the input picture into separate parts. It can create a significant number of different suggestions for orthogonal areas for images of any size.

The visual look of the human body varies depending on the scenario, as was mentioned before; in addition, the RPN is optimised for different size region suggestions as a result of the extremely flexible nature of CNNs. In order to generate RPNs, we used ResNet-50 as the backbone architecture. This design uses a sliding window approach that is applied across the feature maps of the most recent layer. Plain networks are utilised as opposed to more conventional forms of deep learning architectures such as VGG, which frequently rely on convolutional layers for classification rather than completely linked layers with no shortcut or skip connections in between. It does this by utilising what is known as a skip connection, which is also known as a shortcut connection. This produces a bigger network that is able to convert the input of the layer below it to the input of the layer above it without changing or modifying the input.

Social Monitoring Module

The centre point, also known as the centroid of each detected bounding box, is determined after human detection in top view photos, as illustrated in figure 4 with red boxes (a). The bounding box's calculated centroid, whereas figure 4b shows a set of bounding box coordinates with centre point data. The bounding box's centroid is determined as follows:

$$C_{(x,y)} = \frac{x_{min}+x_{max}}{2}, \frac{y_{min}+y_{max}}{2} \quad (1)$$

In equation (1), x represents the bounding box's minimum and maximum width, and y represents the bounding box's minimum and maximum height. Following the computation of the centroid, the distance between each identified centroid is calculated using the Euclidean distance approach. We first calculate the centroid of bounding boxes for each detected bounding box in the image, as shown in figure 4(b); then we measure the distance (represented by red colour lines) between each identified bounding box, as shown in figure 4(c). The distance between the two centroids is calculated as follows:

$$D(c1_{(x,y)}, c2_{(x,y)}) = \sqrt{(x_{max} - x_{min})^2 + (y_{max} - y_{min})^2} \quad (2)$$



Figure 4. Distance measurement example

The estimated distance values are used to determine a T threshold value. This value is used to determine if any two or more people are separated by less than the T pixel threshold. The following equation is used to calculate the threshold T:

$$v = \begin{cases} 1, T \geq D \\ 0, \text{Otherwise} \end{cases} \quad (3)$$

CONCLUSION

One of the corrective steps that may be used to avoid the transmission of the coronavirus illness from public areas is to create social distance. Within the scope of this work, we sought to conduct a review of deep learning-based object detection strategies that are employed for the purpose of keeping the social distance system operational. Human detection and social distance monitoring make up the two most important aspects of the system as a whole. The pretrained Faster-RCNN paradigm is utilised in order to detect humans in the top view position. Transfer learning is a strategy that is used to improve human detection outcomes. This is due to the significant differences between the top and frontal viewpoints. The information on the centroid, also known as the centre point coordinates, is determined by doing calculations with the bounding box data that is supplied by the human detection module. A violation threshold is constructed, which is determined by the information of a person's distance to a pixel and determines whether or not two people are in breach of social distance.

REFERENCES

1. K. Mingis, "Tech pitches in to fight COVID-19 pandemic.", *Computer World*, May 5, 2020. Accessed: Apr. 20, 2020.
2. S. Maharaj and A. Kleczkowski, "Controlling epidemic spread by social distancing: Do it well or not at all," *BMC Public Health*, Vol. 12(1), pp. 679-697, 2012.
3. Al-Khazraji, A.; Nehad, A.E. Smart Monitoring System for Physical Distancing. In *Proceedings of the 2020 Second International Sustainability and Resilience Conference: Technology and Innovation in Building Designs (51154)*, Sakheer, Bahrain, 11-12 November 2020; pp. 1-3.
4. Alrashidi, M. Social Distancing in Indoor Spaces: An Intelligent Guide Based on the Internet of Things: COVID-19 as a Case Study, *Jo. of Computers*, Vol. 9, pp. 80-91, 2020.
5. Neelavathy Pari, S., Vasu, B., and Geetha, A.V., Monitoring Social Distancing by Smart Phone App in the effect of COVID-19. *Glob. J. Comput. Sci. Technol.* Vol. 9, 946-953, 2020.
6. Raju K, Lavanya R, Manikandan S and Srilekha K, "Application of GIS in COVID -19 Monitoring and Surveillance", *International Journal for Research in Applied Science & Engineering Technology*, Volume 8, Issue V, May 2020, ISSN: 2321-9653, <http://doi.org/10.22214/ijraset.2020.5231>
7. Jhunjhunwala, A., Role of Telecom Network to Manage COVID-19 in India: Aarogya Setu. *Trans Indian National Academic Engineering*, Vol. 5, pp. 157-161, 2020. <https://doi.org/10.1007/s41403-020-00109-7>
8. M. Nicola, Z. Alsafi, C. Sohrabi, A. Kerwan, A. Al-Jabir, C. Iosifidis, M. Agha, and R. Agha, The socio-economic implications of the coronavirus and COVID-19 pandemic: a review, *International Journal of Surgery*, vol. 78, pp. 185-193, 2020.
9. S. Bradley, Statistical Analysis of Human Overpopulation and its Impact on Sustainability, *Jo. of Medical Image Analysis*, Vol. 1(8), pp. 1-8, 2018.
10. Manikandan, S., Pasupathy, S., & Hanees, A. L., (2021) "Regression Analysis of Colour Images using Slicer Component Method in Moving Environments", *Quing: International Journal of Innovative Research in Science and Engineering*, 01(01), 01 - 05
11. Bhattacharya, S., Maddikunta, P. K. R., Pham, Q.-V., Gadekallu, T. R., Chowdhary, C. L., Alazab, M., et al. (2020). Sustainable cities and society, 102589.
12. W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu, and A.C. Berg, "SSD: Single Shot multi-box Detector", In: *Proc. of European Conference on Computer Vision*, pp. 21-37, 2016.
13. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object

detection”, In: Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, pp.779-788, 2016.

14. Uijlings, J.; Van de Sande, K.; Gevers, T.; Smeulders, A., “Selective Search for Object Recognition”, *International Journal of Comput. Vis.*, 104, pp.154-171, 2013.

15. R. Girshick, “Fast R-CNN”, In: Proc. of the IEEE International Conference on Computer Vision, pp. 1440-1448, 2015.

16. S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks”, *Advances in Neural Information Processing Systems*, pp.91-99, 2015.

17. W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu and A. Berg. SSD: Single Shot MultiBox Detector. arXiv:1512.02325 (v2), 2016.

18. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You Only Look Once: Unified, Real-Time Object Detection. In: NIPS, 2015.

19. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21-26 July 2017; pp. 6517-6525.

20. Al-Khazraji, A.; Nehad, A.E. Smart Monitoring System for Physical Distancing. In Proceedings of the 2020 Second International Sustainability and Resilience Conference: Technology and Innovation in Building Designs (51154), Sakheer, Bahrain, 11-12 November 2020; pp. 1-3.

FINANCING

The author did not receive funding for the development of this research.

CONFLICT OF INTEREST

No conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Sivakumar Karuppan, Krishnaprasath V T, Pradeep V, Sruthi S Madhavan.

Resarch: Sivakumar Karuppan, Krishnaprasath V T, Pradeep V, Sruthi S Madhavan.

Writing - original draft: Sivakumar Karuppan, Krishnaprasath V T, Pradeep V, Sruthi S Madhavan.

Writing - revision and editing: Sivakumar Karuppan, Krishnaprasath V T, Pradeep V, Sruthi S Madhavan.