SOCIAL DISTANCE MONITORING USING A LOW-COST 3D SENSOR

Amran Hossen [1]

[1]Universiti Brunei Darussalam

October 31, 2023

Abstract

With its devastating spread, the ongoing COVID-19 coronavirus pandemic has caused devastation worldwide. Due to the lack of successful restorative medications as well as the shortage of vaccinations against the virus, the communities have been left highly vulnerable. While a handful of countries have vaccinated the majority of their populations, for many countries, the virus still presents a big challenge. Social distancing is considered to be an effective preventative measure against the transmission of the pandemic virus, and virus propagation can be considerably curbed by preventing physical contact between individuals. The objective of this work is, therefore, to provide a depth image-based and cost-effective method for social distance monitoring. A widely used human detection algorithm from depth images has been used to estimate body joint position in real-time. To approximate social distance violations between individuals, the distance between individuals can be estimated, and then, compared to a predefined threshold. Outcomes of the work indicate that the proposed method can successfully identify individuals who violate the social distancing rules, with over 98% detection accuracy. The results are significant, as it can be implemented to assist enforcement agencies to ensure that social distancing rule is abided by the population, to limit the spread of COVID-19 among the population.

SOCIAL DISTANCE MONITORING USING A LOW-COST 3D SENSOR

M. A. Hossen[1], P. E. Abas*[1]

[1] Faculty of Integrated Technologies, Universiti Brunei Darussalam,
Brunei Darussalam.
(E-mail: 20h8453@ubd.edu.bn,
emeroylariffion.abas@ubd.edu.bn)

ABSTRACT

With its devastating spread, the ongoing COVID-19 coronavirus pandemic has caused devastation worldwide. Due to the lack of successful restorative medications as well as the shortage of vaccinations against the virus, the communities have been left highly vulnerable. While a handful of countries have vaccinated the majority of their populations, for many countries, the virus still presents a big challenge. Social distancing is considered to be an effective preventative measure against the transmission of the pandemic virus, and virus propagation can be considerably curbed by preventing physical contact between individuals. The objective of this work is, therefore, to provide a depth image-based and cost-effective method for social distance monitoring. A widely used human detection algorithm from depth images has been used to estimate body joint position in real-time. To approximate social distance violations between individuals, the distance between individuals can be estimated, and then, compared to a predefined threshold. Outcomes of the work indicate that the proposed method can successfully identify individuals who violate the social distancing rules, with over 98% detection accuracy. The results are significant, as it can be implemented to assist enforcement agencies to ensure that social distancing rule is abided by the population, to limit the spread of COVID-19 among the population.

Key words: Social distance monitoring, Kinect, COVID-19, Person detection

INTRODUCTION

The novel coronavirus (2019-nCoV) or COVID-19, which first appeared in Wuhan, China, in December 2019, has affected most countries around the world. On March 11, 2020, the World Health Organization (WHO) has declared it as a pandemic disease, after it had spread across 114 countries, resulting in over 4,000 confirmed deaths and 118,000 active cases [1]. 147 million people have been confirmed to be infected with COVID-19, with over 3 million deaths, as of April 25, 2021 [2]. Various healthcare institutes, scientists, and medical practitioners have been looking for effective vaccines and drugs to combat this deadly virus, with certain progress reported to date. Simultaneously, the international community has also been searching for new ways to stop the epidemic from spreading. At an early stage of an outbreak of a unknown virus for which there is no particular drug or vaccine that is available, conventional public-health procedures, such as isolation, quarantine, social distancing, and community containment play vital roles [3][4]; with the main objectives of such public-health measures, are to disrupt transmission, and prevent human-to-human spread of infection via physical separation of people. The virus is mostly transmitted by individuals, who are in close proximity to one another over an extended period of time. When an infected person sneezes, coughs, or speaks, the virus circulates through the air as tiny droplets from their nose or mouth, and infects the surrounding people. The droplets then travel across the respiratory system to the lungs, where the virus begins to destroy the lung cells. According to recent research, people who exhibit no symptoms, but yet have been infected with the virus, play a big role in the spread of the virus [5][6][7]. As such, even though no symptom is present, it is best to wear a mask and maintain a distance of at least 6 feet from one another [8] [9], [10], to contain the spread of the virus.

Social distancing refers to strategies for preventing the transmission of an infection, by limiting humans physical contact, particularly, at public spaces such as shopping centres, parks, colleges, universities, medical centres, airports, and workplaces, as well as avoiding crowded gatherings, and keeping a safe distance among people. Social distancing is crucial, especially for those people who are at an elevated risk of severe illness from COVID-19. The distribution of infections, and severity of the disease can be significantly reduced by lowering the chance of virus transmission from an infectious person to a healthy person via social distancing. According to research from Harvard University, if social distancing had been implemented during the earlier stages of the pandemic, it should play a vital role in reducing the viral transmission and stopping the pandemic outbreak from reaching its peak [11]. Figure 1**.** illustrates the objectives of social distancing measures in minimizing and delaying the peak of the epidemic, and facilitating healthcare capability [12]; with social distancing shown to reduce the number of infected patients, and the pressure on healthcare institutions. Consequently, it also reduces mortality rates by ensuring that the number of affected patients does not exceed the capacity of public healthcare [13].

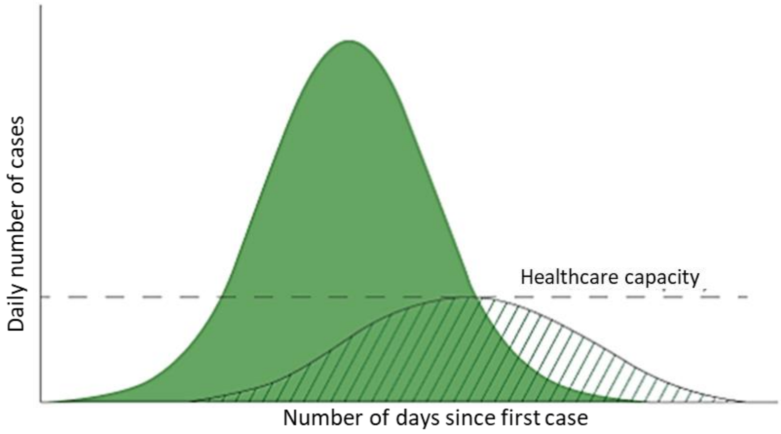

Figure 1. The goals of social distancing measures to decrease and defer the peak of the epidemic and safeguard healthcare facilities [12]

In this paper, a low-cost social monitoring method has been presented. The method uses the locations of human body joints, which have been extracted from depth images captured using a Kinect device, to calculate distances between individuals. These distances are then compared to the social distancing threshold, to provide an effective method for social distance monitoring. The paper is structured as follows: Section 2 reviews the relevant literature review, Section 3 outlines the proposed method, whilst results and discussions are presented in Section 4. Finally, Section 5 concludes the paper.

LITERATURE REVIEW

Over the last few decades, computer vision, machine learning, and deep learning have demonstrated great potential in solving a variety of real-world challenges. With the latest development of different machine learning algorithms, there has been significant progress in addressing previously challenging tasks, including object recognition [14][15]. Human detection is regarded as a problem in the object detection domain, and is frequently utilized by researchers [9], [16]–[19] for identifying people and measuring distances between people through video scenes, as seen in Figure 2. The majority of techniques used in human detection are built using frontal or side view video sequences. On the other hand, other researchers have also suggested that using overhead positioning of cameras would give a more accurate distance measurement, broader area coverage for a large scene, and reduce occlusion problems to a certain degree [9]. Clustering and distance-based techniques have been used to calculate the distances between individuals, which involve appropriate camera calibration, ground plane estimation, image to real-world mapping, and assignment of pixels to distance in quantifiable units, as seen in Figure 2 (a), (b) and (c). However, RGB image-based approaches are generally affected by light conditions, partial occlusion, background complexity, and accuracy problems in performing pixel to distance measurement, as well as the computational cost associated with the processing of RGB image-based approaches.

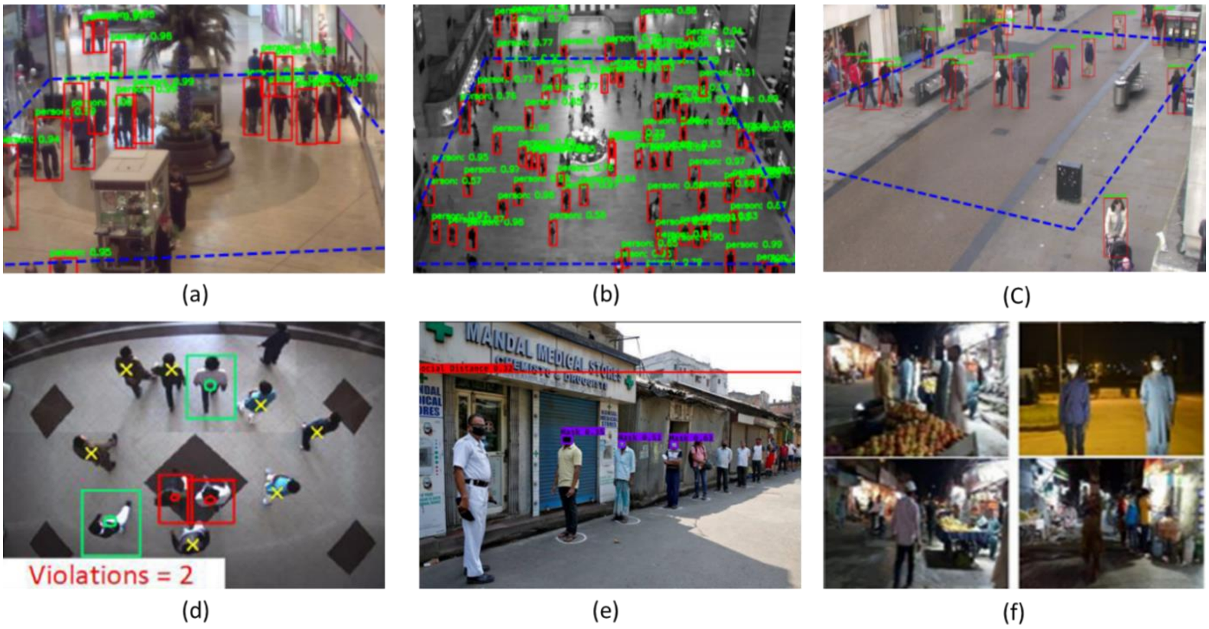

Figure 2. Images from the literature that have been used for social distance monitoring. (a), (b) and (c) [18], (d) [9], (e) [16], (f) [19]

The release of low-cost 3D sensors, such as Microsoft Kinect and Asus Xtion live pro, has somehow reduced these challenges, by offering both RGB image and depth image streams via a single affordable product [20]. Despite being primarily aimed at the entertainment industry, the Kinect has also sparked significant interest from the vision and robotics communities, due to its wide range of potential applications [21]. Skeleton joint positions estimation method using Kinect depth images is more reliable, and has opened opportunities in human-centric computer vision tasks, with the method rapidly adopted by the research communities to address complex tasks, such as human activity recognition [22]–[27]. Shotton et al. [28] have developed a robust technique for estimating 3D locations of skeletal joints effortlessly and precisely from a single Kinect depth image. This has been subsequently adopted in this paper, for human detection and body joint position estimation. A local mode-finding method based on a mean change with a weighted Gaussian kernel is used to compute confidence-scored 3D location for the estimation of body joints. The huge amount of heterogeneous training dataset that is available, enables the classifier to approximate body parts regardless of posture, body shape, lighting, or clothes, with some accuracy. With this technique, the challenges of social distance measurement and monitoring can be addressed relatively easy, in comparison to the use of RGB image on its own. Consequently, a depth camera-based social distance monitoring method is presented in this paper.

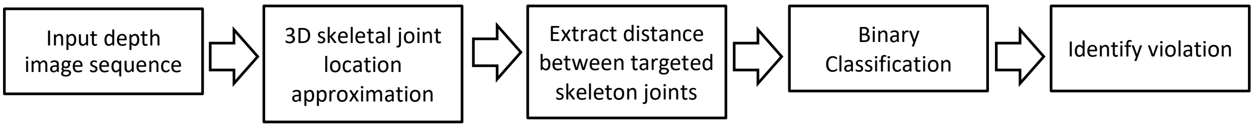

Figure 3. Overview of the proposed social distance monitoring method

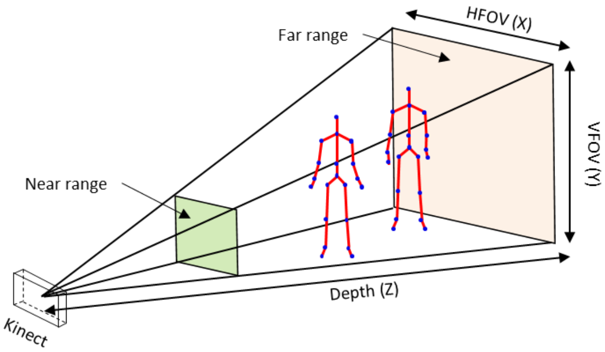

Figure 4. Visualization of Kinect's field view

METHODOLOGY

Figure 3 shows the overview of the proposed social distance monitoring method, which consists of four basic steps for the monitoring of social distance violations. It starts with the capture of depth frames, which are then used to identify the locations of the human body joints, particularly, the centre shoulder joints, in the second step. The distances between the centre shoulder joints of tracked individuals are then calculated, and compared with a threshold value for social distance monitoring, to give binary classifications on social distancing violators. Kinect's fields of view are depicted in Figure 4**Error! Reference source not found.**, where the z-axis represents the distance from the sensor, whilst the x-axis and y-axis are its horizontal field of view (HFOV) and vertical field of view (VFOV). Mechanisms employed for body joint estimation and social distance monitoring are presented in the sections below.

3.1 Body Joint Position Approximation: The human body is an interconnected structure of rigid parts linked together by different joints [29]. Joint positions are used to provide a compact depiction of a person in this context. The method proposed by Shotton et al. [28] has been used to extract 3D body joint locations from a depth frame by applying an object recognition system, with body segments of the human body classified on per-pixel classification results. Depth image to body joint identification can be visualised in Figure 5(a). Figure 5(b) shows the twenty joint skeleton system, obtained using the method. While these skeletal joint locations are crucial to construct posture representation and human activity recognition, in this paper, only one joint position, namely the centre shoulder joint (joint 3), for every detected individual, is utilised to monitor social distance.

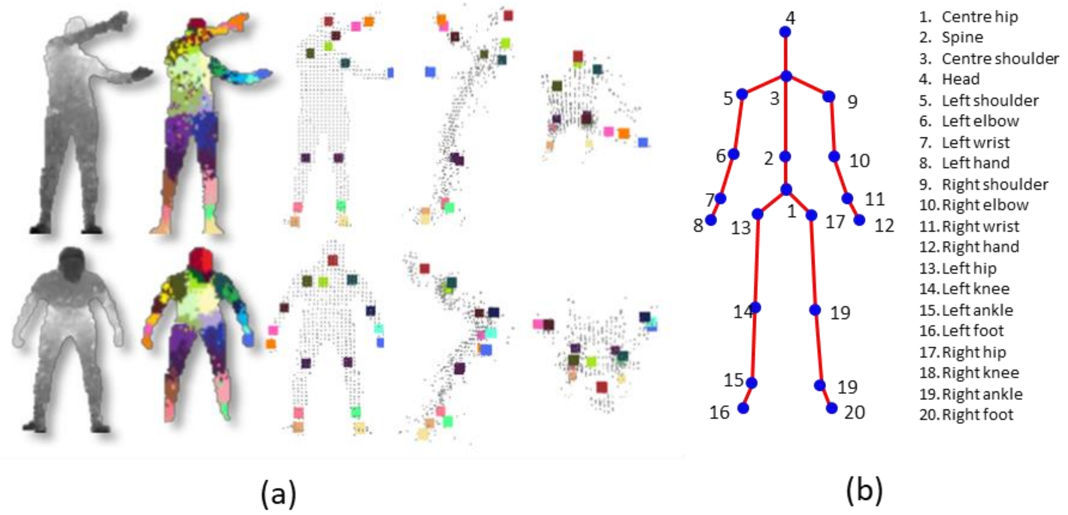

Figure 5. (a) Depth image to body joints estimation [28]. (b) Twenty joint skeleton system

3.2 Social Distance Monitoring: Figure 6(a) shows the depth image, whilst Figure 6(b) illustrates the transformed top view of the derived skeletons from the depth image. To visualize the process of social distance monitoring, a straight line is drawn between the centre shoulder joints as shown in Figure 6(c), whereby if the pair of detected individuals are in close proximity to each other, with distance less than the minimum social distance requirement, the system would detect a social distancing violation and vice versa.

A unique identifier is assigned for each detected individual, based on its first appearance in the sensor's field of view, and then, at every subsequent frame in the stream, the coordinates of the individual are calculated based on the centre shoulder joint. Given that N represents the number of individuals detected in a frame, the position of individual P_i in the image can be represented in the three-dimensional space as $(x_i, y_i, z_i) \in \mathbb{R}^3 : i \in \{1, \ldots\ldots, N\}$.

For every pair of detected individuals, Euclidean distance between centre shoulder joints can be determined (as in Figure 6(c)), where the distance $D(i,j)$ between detected individual P_i with coordinates (x_i, y_i, z_i) and P_j with coordinates (x_j, y_j, z_j), is given by:

$$D(i,j) = \sqrt{\left(x_j - x_i\right)^2 + \left(y_j - y_i\right)^2 + \left(z_j - z_i\right)^2} \qquad i,j \in \{1, \ldots\ldots, N\} \tag{1}$$

The distances between all the N individuals can be populated in $N \times N$ distance matrix $\mathbf{D}$. A computed distance $D(i,j)$, which is less than a given social distance threshold value D_{Th} i.e. $D(i,j) < D_{Th}$, indicates that individual P_i is too close to the individual P_j, and hence, is violating the social distancing rule. On the other hand, any other value of $D(i,j)$ i.e. $D(i,j) \geq D_{Th}$ indicates no violation of the social distancing rule. However, $D(i,j) = D(j,i)$ since the distance matrix $\mathbf{D}$ is symmetric, and $D(i,i)$ is the distance of individual P_i from himself, and as such, need to be taken into account when determining unique social distancing violations. An $N \times N$ violation matrix $\mathbf{V}$, which is a strictly upper triangular matrix, can be used to populate unique social distancing violations at every frame:

$$V(i,j) = \begin{cases} D(i,j) < D_{Th} & , i < j \\ 0 & i \geq j \end{cases} \tag{2}$$

$V(i,j) = 1$ indicates social distancing violations between individuals i and j, whilst $V(i,j) = 0$ indicates no social distancing violations. The number of social distancing violations per frame can then be determined as:

$$V_{Total} = \sum_{i=1}^{N} \sum_{j=1}^{N} V(i,j) \tag{3}$$

with the maximum number of unique violations V_{max} between individuals per frame, given by [30]:

$$V_{max} = \frac{N \times (N-1)}{2} \tag{4}$$

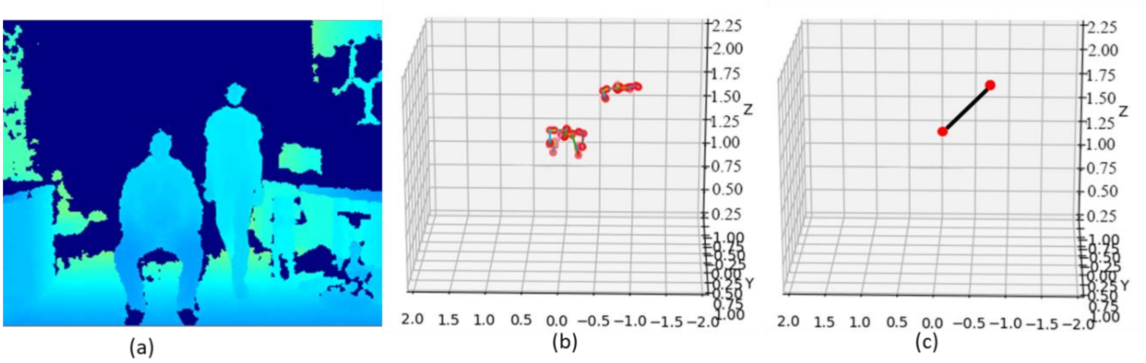

Figure 6. (a) Depth image (b) Skeleton joints derived from depth image (c) Centre shoulder position shown in X-Z plane.

RESULT & DISCUSSION

To test the proposed social distance monitoring method, an indoor dataset containing depth sequences derived from a 3rd person view has been used. The dataset was captured with 30 frames per second (FPS) frame rate for both RGB and depth images using Kinect v1, which has an effective depth range of 0.7–6 meters or 2.3–20 feet. Consequently, the experiments had been conducted within the range. Throughout the scene, there were no restrictions on the movement of the participants, with the dataset captured in varying lighting conditions, and the presence of everyday objects. A total of 11,182 frames were used for evaluation, which has been categorised into four groups, with sample images shown in Figure 7. For the first and second groups of images, all individuals were walking without, and whilst maintaining social distancing rule, respectively. The third group contains images of individuals having group discussions whilst maintaining social distancing rule, and finally, the fourth group contains images of individuals having group discussion; with both social distancing violations and non-violations. D_{Th} has been set at 1.5 meters, which is within the range (1-2 meters) recommended by the Ministry of Health, Brunei Darussalam [31].

Instances of social distance monitoring are depicted in Figure 8. Four persons were detected with no violations in Figure 8(a). In Figure 8(b), two violations were recorded, and had been highlighted in red, with the scene also containing two non-violations marked in green. Similarly, 3 violations were recorded in Figure 8(c).

Figure 7. RGB images (row 1) and corresponding depth images (row 2)

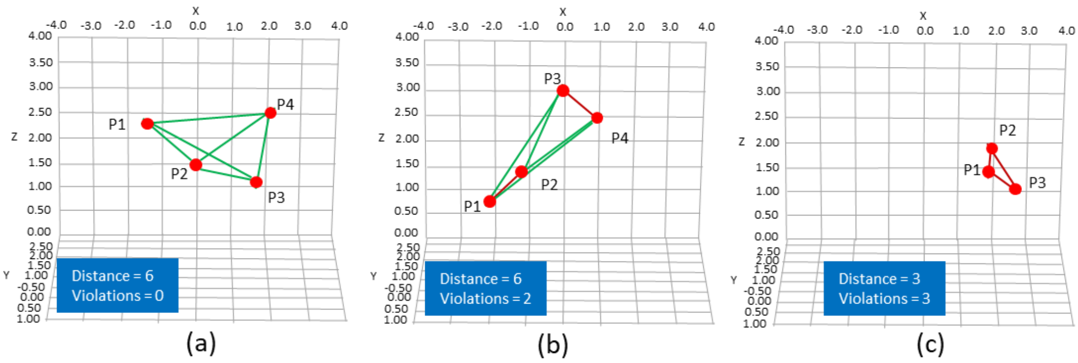

Figure 8. Social distance monitoring results

Table 1 summarises the number of frames used for inspections and distances examined for analysis, from the experiments, categorised according to the two activities: walking and group discussion. A total of 11,331 distances were monitored for walking activity, composed of 4,834 non-violations and 6,497 violations. On the other hand, 40,908 distances were monitored for group-discussion activity, composed of 25,446 non-violations and 15,462 violations. Confusion matrices for the proposed social distance monitoring method, for the two activities, are shown in Figure 9. A sequence of walking containing both violation and non-violation is depicted in Figure 9(a). The result for walking activity in Figure 9(a), demonstrates that whilst walking, some distances have been falsely classified, most probably due to obstruction of the shoulder centre joints. 2% and 3% of social distance violations, and non-violations, respectively, have been wrongly classified. This is in contrast to the result for the group discussion activity shown in Figure 9(b), which shows that the proposed method is able to accurately differentiate social distance violations and non-violations. Overall detection scores for each group of violation and non-violation activities are shown in Figure 10.

Table 1. Summary of frames and distance observed.

	Frames observed	Distance monitored
Walking (non-violation)	2,123	4,834
Walking (with violation)	2,241	6,497
Group discussion (non-violation)	4,241	25,446

Group discussion (with violation)	2,577	15,462

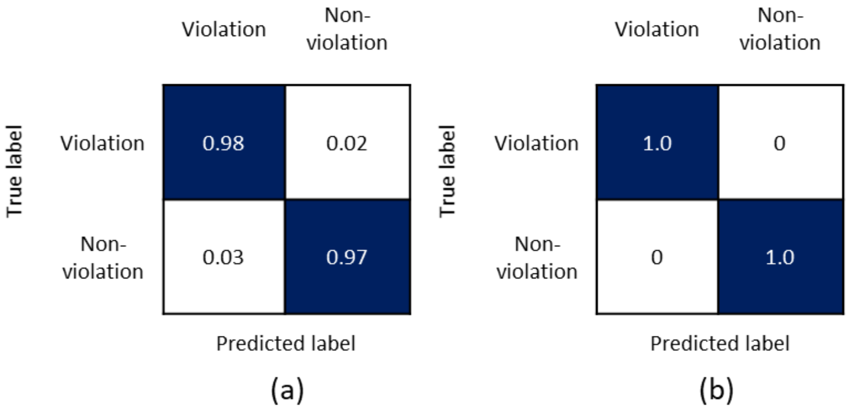

Figure 9. (a) Confusion matrix (walking) (b) Confusion matrix (group discussion)

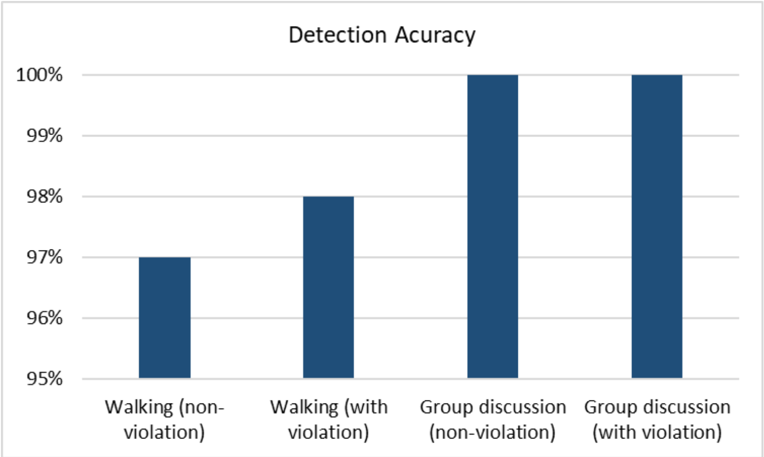

Figure 10. Overall detection accuracy

CONCLUSION

Social distancing has been shown to be an effective prevention against the spread of the Covid-19 virus, and consequently, many governments have adopted it as part of the social distancing rules, especially in public places. However, enforcement of the social distancing rules has been proven to be difficult. In this paper, a low-cost 3D sensor-based social distance monitoring method has been presented. A widely used depth image-based method has been used to derive skeleton joint positions of detected individuals, with the centre shoulder joints utilised for the calculation of Euclidean distances between individuals. These distances are then compared with a predefined social distance threshold, to ascertain abidance to the social distancing rule. Experimental results have shown that the presented method can effectively identify social distance violations, with an accuracy of 98% in identifying social distancing violations on the recorded dataset with varying light conditions. The proposed method utilizing the depth camera does not require camera calibration or ground plane estimation, making it an ideal choice for real-world needs, and can be used by health enforcement agencies in ensuring that the social distancing rule is abided by the populations, especially in public spaces.

Acknowledgement: This work is funded by the UBD Bursary Award.

REFERENCES

1. 'Coronavirus Disease (COVID-19) - events as they happen'. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen. [Accessed: 25-Apr-2021].
2. 'WHO Coronavirus (COVID-19) Dashboard'. [Online]. Available: https://covid19.who.int/. [Accessed: 25-Apr-2021].
3. P. Caley, D. J. Philp, and K. McCracken, 'Quantifying social distancing arising from pandemic influenza', *J. R. Soc. Interface*, vol. 5, no. 23, pp. 631–639, 2008.
4. A. Wilder-Smith and D. O. Freedman, 'Isolation, quarantine, social distancing and community containment: Pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak', *J. Travel Med.*, vol. 27, no. 2, pp. 1–4, 2020.
5. M. Gandhi, D. S. Yokoe, and D. V Havlir, 'Asymptomatic Transmission, the Achilles' Heel of Current Strategies to Control Covid-19', *N. Engl. J. Med.*, vol. 382, no. 22, pp. 2158–2160, May 2020.
6. S. M. Moghadas *et al.*, 'The implications of silent transmission for the control of COVID-19 outbreaks', *Proc. Natl. Acad. Sci.*, vol. 117, no. 30, pp. 17513–17515, Jul. 2020.
7. Y. Bai *et al.*, 'Presumed Asymptomatic Carrier Transmission of COVID-19', *JAMA*, vol. 323, no. 14, p. 1406, Apr. 2020.
8. 'How to Protect Yourself & Others'. [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html. [Accessed: 25-Apr-2021].
9. I. Ahmed, M. Ahmad, J. J. P. C. Rodrigues, G. Jeon, and S. Din, 'A deep learning-based social distance monitoring framework for COVID-19', *Sustain. Cities Soc.*, vol. 65, no. November 2020, p. 102571, 2021.
10. Y. Feng, T. Marchal, T. Sperry, and H. Yi, 'Influence of wind and relative humidity on the social distancing effectiveness to prevent COVID-19 airborne transmission: A numerical study', *J. Aerosol Sci.*, vol. 147, no. April, p. 105585, 2020.
11. 'Preventing the spread of the coronavirus'. [Online]. Available: https://www.health.harvard.edu/diseases-and-conditions/preventing-the-spread-of-the-coronavirus. [Accessed: 25-Apr-2021].
12. ECDC, 'Considerations relating to social distancing measures in response to the COVID-19 epidemic', *Eur. Cent. Dis. Prev. Control.*, no. March, pp. 1–10, 2020.
13. H. Dinh Thai *et al.*, 'Enabling and Emerging Technologies for Social Distancing: A Comprehensive Survey'. 2020.
14. L. Jiao *et al.*, 'A Survey of Deep Learning-Based Object Detection', *IEEE Access*, vol. 7, pp. 128837–128868, 2019.
15. L. Liu *et al.*, 'Deep Learning for Generic Object Detection: A Survey', *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 261–318, Oct. 2019.
16. L. Ramadass, S. Arunachalam, and S. Z., 'Applying deep learning algorithm to maintain social distance in public place through drone technology', *Int. J. Pervasive Comput. Commun.*, vol. 16, no. 3, pp. 223–234, Jun. 2020.
17. N. S. Punn, S. K. Sonbhadra, and S. Agarwal, 'COVID-19 Epidemic Analysis using Machine Learning and Deep Learning Algorithms', *medRxiv*, 2020.
18. D. Yang, E. Yurtsever, V. Renganathan, K. A. Redmill, and Ü. Özgüner, 'A Vision-based Social Distancing and Critical Density Detection System for COVID-19', *arXiv*, 2020.
19. A. Rahim, A. Maqbool, and T. Rana, 'Monitoring social distancing under various low light conditions with deep learning and a single motionless time of flight camera', *PLoS One*, vol. 16, no. 2 February, pp. 1–19, 2021.
20. 'The teardown', *Eng. Technol.*, vol. 6, no. 3, pp. 94–95, 2011.
21. J. Giles, 'Inside the race to hack the Kinect', *New Sci.*, vol. 208, no. 2789, pp. 22–23, 2010.
22. J. K. Aggarwal and L. Xia, 'Human activity recognition from 3D data: A review', *Pattern Recognit. Lett.*,

vol. 48, pp. 70–80, Oct. 2014.

23. L. Xia, C. Chen, and J. Aggarwal, 'View Invariant Human Action Recognition Using Histograms of 3D Joints The University of Texas at Austin', *CVPR 2012 HAU3D Work.*, pp. 20–27, 2012.
24. E. Cippitelli, S. Gasparrini, E. Gambi, and S. Spinsante, 'A Human Activity Recognition System Using Skeleton Data from RGBD Sensors', *Comput. Intell. Neurosci.*, vol. 2016, 2016.
25. G. Paraskevopoulos, E. Spyrou, D. Sgouropoulos, T. Giannakopoulos, and P. Mylonas, 'Real-time arm gesture recognition using 3D skeleton joint data', *Algorithms*, vol. 12, no. 5, pp. 1–17, 2019.
26. M. Trăscău, M. Nan, and A. M. Florea, 'Spatio-temporal features in action recoacgnition using 3D skeletal joints', *Sensors (Switzerland)*, vol. 19, no. 2, 2019.
27. C. Zhao, M. Chen, J. Zhao, Q. Wang, and Y. Shen, '3D behavior recognition based on multi-modal deep space-time learning', *Appl. Sci.*, vol. 9, no. 4, 2019.
28. J. Shotton *et al.*, 'Real-Time human pose recognition in parts from single depth images', *Commun. ACM*, vol. 56, no. 1, pp. 116–124, 2013.
29. K. N. An, 'Kinematic analysis of human movement', *Ann. Biomed. Eng.*, vol. 12, no. 6, pp. 585–597, Nov. 1984.
30. J. Dattorro, 'Euclidean Distance Matrix', *Convex Optim. Euclidian Distance Geom.*, no. 886, pp. 385–485, 2001.
31. 'GUIDELINES ON SOCIAL GATHERINGS AND EVENTS, MEETINGS AND TRAINING SESSIONS'. [Online]. Available: http://www.kesihatan.gov.bn/Shared Documents/2019 ncov/guidelines/mass gathering guidelines/GUIDELINES ON SOCIAL GATHERINGS, MEETINGS AND TRAINING (updated 28 April 2020).pdf. [Accessed: 25-Apr-2021].