Irvine, FRR, Natalia, F, Sudirman, S and Ko, CS

Automated Physical Distancing Monitoring Using YOLOV3

ICIC Express Letters
Part B: Applications
Volume **14**, Number **8**, August **2023** ISSN 2185-2766
pp. 869–876

AUTOMATED PHYSICAL DISTANCING MONITORING USING YOLOV3

Flavius Ryaas Rasyid Irvine[1], Friska Natalia[1,*], Sud Sudirman[2] and Chang Seong Ko[3]

[1]Department of Information Systems
Universitas Multimedia Nusantara
Scientia Boulevard, Gading Serpong, Tangerang, Banten 15811, Indonesia
flavius.rasyid@student.umn.ac.id; *Corresponding author: friska.natalia@umn.ac.id

[2]School of Computer Science and Mathematics
Liverpool John Moores University
Liverpool L3 3AF, UK
s.sudirman@ljmu.ac.uk

[3]Department of Industrial and Management Engineering
Kyungsung University
309, Suyeong-ro, Nam-gu, Busan 48434, Korea
csko@ks.ac.kr

Received December 2022; accepted March 2023

Abstract. *Physical distancing has been practiced and proven to be part of a solution to reduce the spread of COVID-19 during this pandemic. The method, when implemented together with other COVID-19 protocols such as face mask wearing, maintaining personal hygiene, and mass mobility limitation is very effective in reducing the airborne virus infection rate. As more and more countries and communities are returning to normal life during this pandemic, the enforcement of COVID-19 rules will need to be more automated to make it as least intrusive as possible. In this paper, we designed an automated physical distancing monitoring system using the YOLOv3 object detection library to detect people in the video frames of the system's camera and determine the physical distance between them if more than one person is detected. The system has been implemented on our campus and has been shown to be sufficiently accurate in achieving those tasks.*
Keywords: COVID-19 mitigation, Person-object detection, Physical distance monitoring, Rapid application development, YOLOv3

1. **Introduction.** The COVID-19 pandemic has been going on for almost three years. The virus was first discovered in December 2019 in Wuhan City, Hubei Province, China [1]. On 30 January 2020, the World Health Organization (WHO) declared a public health emergency of international concern, and on 11 March 2020, it declared COVID-19 as a pandemic. In the span of a few weeks, the weekly global confirmed cases reached over 51 thousand. Many experts believe that the actual numbers are significantly higher due to the difficult and lengthy testing process to confirm the disease at that time. Governments across the world race to slow or break the spread of COVID by applying mandatory restrictions on public movement and gathering, in addition to promoting better public health measures such as wearing face masks and washing hands.

According to the latest figure from the Center for Systems Science and Engineering, John Hopkins University Coronavirus Resource Center, the virus has so far infected over 633 million people and claimed over 6.6 million lives worldwide [2]. The development of effective and safe COVID-19 vaccines had gone a long way in reducing the number of deaths due to COVID-19 infections and other health complications. The weekly death rate

DOI: 10.24507/icicelb.14.08.869

due to COVID-19 is now averaging around 10.6 thousand compared to over 100 thousand at the peak of the pandemic [2]. This encourages many countries to slowly return to life back before the pandemic. The speed and extent that life can return to the pre-pandemic condition vary from country to country. Some countries, such as the UK and many other Western countries, have forgone all COVID-19 restrictions completely while others, such as China and several other East Asian countries, still maintain strict pandemic rules and regulations. Indonesia, as with the majority of other countries, adopts a more moderate approach where most people have returned to work in their offices as before but with a strict COVID-19 protocol in place. Returning to life back to the pre-pandemic condition without having the pandemic to officially end is termed as 'living with COVID'. It is a situation where we have to accept some level of risk of catching the disease without sacrificing much social and economic progress.

The approach that Indonesia adopts involves the use of face masks in public places, maintaining minimum physical distance from other people who are not from the same household, and extensive use of the official COVID-19 mobile application called PeduliLindungi [3]. The enforcement of this regulation in a workplace is the responsibility of the company and is regulated by law. In practice, enforcement is very hard to implement due to the impracticality of employing a large number of security guards to check on people every so often. It is therefore important for companies to invest in a more economical and sustainable solution to the problem to ensure more long-term compliance with the law.

Using technology, in particular computer vision technology, is one of the best solutions to achieve this. Computer vision is a field of artificial intelligence that develops algorithms to allow computers to obtain high-level understanding from digital images or videos. Traditionally, machine learning algorithms are used to process image features extracted from the images or videos. Machine learning has been well-studied in the literature and has been used to solve a wide range of computing tasks including non-destructive testing in engineering [4], digital watermarking for copyright protection [5], and tourism data analytics [6]. In recent years, a subset of machine learning called deep learning gains a lot of traction in the research community. Deep learning uses artificial neural networks that have a high number of processing layers which are used to extract progressively higher-level features from the input data. This is made possible because of the increase in computing resources and power which then makes the solutions to computer vision problems more accurate and efficient. Deep learning has been used in various applications such as sign language recognition [7], organ size measurement in medical images [8,9], and pest detection and classification in agriculture [10,11].

Deep learning has also been used in several computer systems that aim to reduce the spread of COVID such as a face mask detection system in [12]. In this paper, Chowdary et al. use an InceptionV3 model that has been pre-trained on a subset of the ImageNet dataset [13] containing over a million images and a thousand categories. A transfer learning method is applied by using the weight values of the pre-trained model as the initial weight values of the new model before retraining the new model on a new dataset called the Simulated Masked Face Dataset [14]. Due to the size limitation of the new dataset, an image augmentation technique is adopted to generate more training images for better generalization of the model. The model is reported to achieve very high accuracy when tested on the images from the dataset. Our investigation on this subject area found that the vast majority of approaches, as reported in [15] and [16], concentrate on the development of the algorithm on still image datasets as opposed to real-world implementations on video frames taken in real time.

To solve the problem of real-time physical distance monitoring, we propose to use an off-the-shelves object detection library and tailor it to work on video frames obtained from a web camera. This paper details the design and development of our system that monitors people and calculates the distance between them in real time. We report the system's performance by describing its outputs in several example scenarios. The paper is organized as follows. Section 2 describes the methodology of the study. In Section 3, we describe the implementation detail, define the testing scenarios, and present the experimental results before we provide the summary of our findings in Section 4.

2. **Methodology.** In this study, we use YOLOv3 [17] to perform real-time object detection in videos. YOLO, which is an acronym for You Only Look Once, is a state-of-the-art deep learning model and algorithm that has become the main method of detecting objects in real time. YOLO has many versions, and the current version is YOLOv7; however, each version of YOLO has its own advantages and disadvantages. We use YOLOv3 in this study because we found that this version of YOLO is more suited to the specification of the hardware that we used while at the same time meeting the functional requirements of the system we need to build. According to a study by Dixit et al. [18], YOLOv3 is faster and more precise than Faster R-CNN [19], which allows it to perform person-object detection and physical distance estimation in real time. YOLOv3 is designed to be fast by prioritizing speed at the expense of accuracy, especially in the background region [18].

Figure 1 shows an example result of object detection using YOLO. The picture shows a relatively crowded area of a street. Several classes are shown to have been detected including person, backpack, dog, bench, and chair. The YOLO algorithm can group and detect many objects in an image whose location is marked with a bounding box.

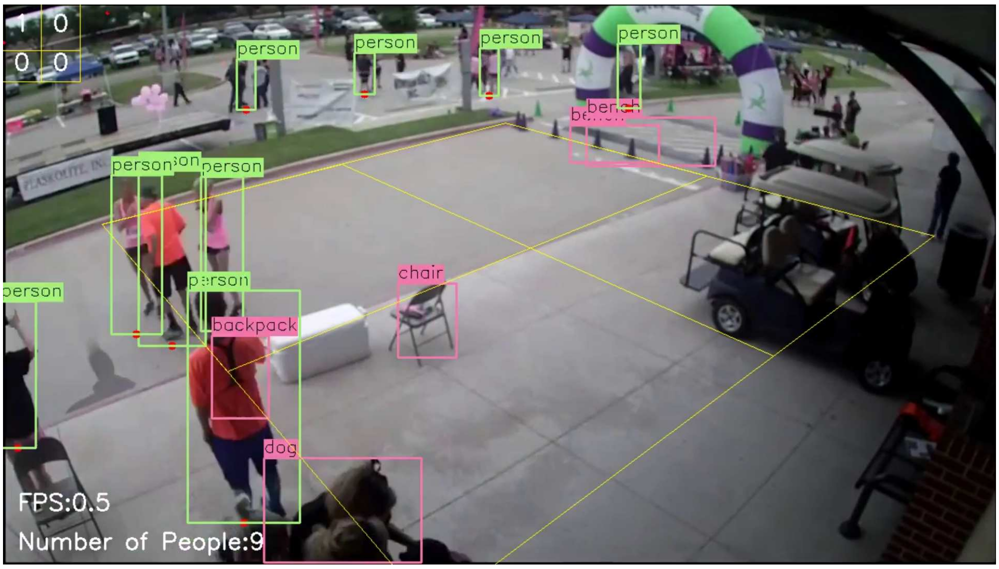

Figure 1. An example of object detection using YOLO

The box size and location for each detected object are determined based on the probability value calculated by YOLO and refined by the bounding box regression technique [20]. Each bounding box can be expressed using a set of four numbers marking two parameters of the box. These can be the two numbers x_c and y_c marking the center coordinate of the box plus p_w and p_h marking the width and height of the box in x and y directions, respectively as illustrated in Figure 2.

Our physical distancing monitoring system is designed using the Rapid Application Development (RAD) method due to its compact iteration to achieve our development goals [21]. The method has four main stages which are depicted in Figure 3, and they are

1) **Requirements Planning** identifies the purpose of the system development by researching the current problem, defining the requirements for the project, and finalizing the requirements.

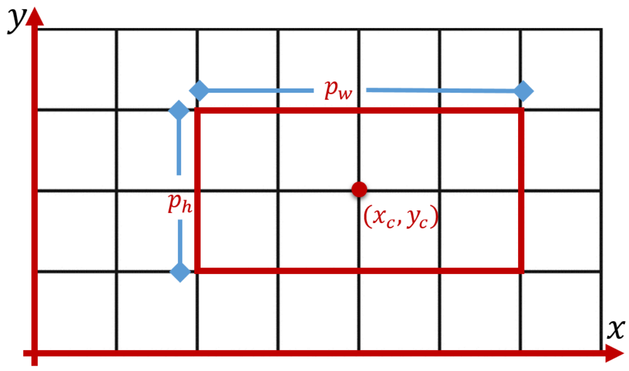

FIGURE 2. Illustration of a bounding box and its parameters

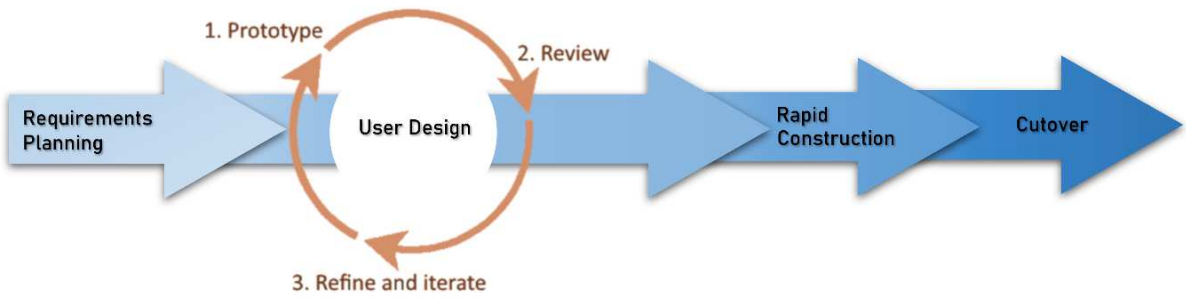

FIGURE 3. Overview of the rapid application development method

2) **User Design** is an iterative process of prototyping, review, and system refinement to ensure the product meets the system requirement.
3) **Rapid Construction** converts the prototype from the design phase and converts them into the working model. This phase involves coding, unit integration, and system testing with a scenario that has been prepared in advance.
4) **Cutover** is the stage of launching and deploying the system after the rapid construction stage.

3. **Implementation, Experimental Results, and Discussion.** The YOLOv3 part of the system is implemented in Python using Anaconda and Jupyter Notebook as the development environment. The YOLOv3 model has already been pre-trained, and to use it in our application we need to tune the model using the images we collected from a webcam. This process takes a short amount of time since the object that we want to detect already belongs to one of the categories that YOLOv3 can already detect, i.e., person. The tuning step helps ensure that the YOLOv3 algorithm adapts to the characteristics of the images captured by the webcam. These include the lighting, viewpoint, and object scale.

The system is deployed inside the campus library of the Universitas Multimedia Nusantara using a personal computer as shown in Figure 4. The computer is equipped with a webcam and has been installed with all the software needed to run the system. The webcam faces the library entrance in order to monitor and capture as much people traffic as possible.

During runtime, we set the system to only detect a person-object from each frame in the video feed. When a single person object is detected, a green bounding box will be drawn in the location of the object. When two or more person objects are detected, the distance between them will be estimated. This is then used by the system to categorize the situation into three categories, namely SAFE, RISKY, and UNSAFE:

1) **SAFE category.** If the distance between pair of person-objects is more than two meters, the two objects will be drawn by green bounding boxes.

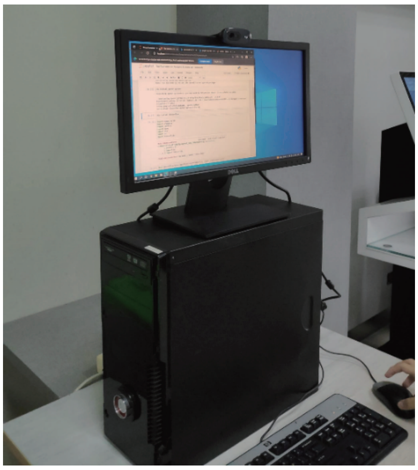

FIGURE 4. A picture of the installed physical distance monitoring system

2) **RISKY category.** If the distance between pair of person-objects is more than one meter but less than two meters, the two objects will be drawn by orange bounding boxes. A yellow line will be drawn from the center of one box to the other.
3) **UNSAFE category.** If the distance between pair of person-objects is less than one meter, the two objects will be drawn by red bounding boxes. A red line will be drawn from the center of one box to the other.

A user acceptance test was carried out to assess the system's ability to correctly categorize the scene into the above categories in several scenarios. The scenarios have been designed to represent many possible cases that take account of the number of people in the frame, the level of occlusion, and their distance to the camera and each other. There are six scenarios to be considered and they are described in Table 1.

TABLE 1. Description of the six scenarios of the user acceptance test

Scenario	Description	Expected outcome
a	A single person who is close to the camera.	A single green bounding box
b	Three people who are at different distances from the camera and all observe the physical distancing rule.	Three green bounding boxes
c	Two people who are far away from the camera and both observe the physical distancing rule.	Two green bounding boxes
d	Two people who are far away from the camera and both observe the physical distancing rule but are less than two meters apart.	Two orange bounding boxes and a yellow line
e	Two people who are far away from the camera and breaking the physical distancing rule.	Two red bounding boxes and a red line
f	Three people who are far away from the camera. Two break the physical distancing rule while the other observes the rule.	Two red bounding boxes and a red line and one green bounding box

The results of the test are shown in Figure 5. Figure 5(a) shows a single GREEN bounding box around the one detected person object. Figure 5(b) shows three GREEN bounding boxes around the three detected person objects at varying distances to the camera, confirming they are all SAFE. Figure 5(c) shows two GREEN bounding boxes around the two detected person objects that are far from the camera, confirming both are SAFE. Figure 5(d) shows two ORANGE bounding boxes and a yellow line between the two detected person objects that are far from the camera, confirming both are RISKY.

(a) (b) (c) (d) (e) (f)

FIGURE 5. Example outputs of the physical distance monitoring system. (a), (b), and (c) show a SAFE situation, (d) shows a RISKY situation, (e) shows an UNSAFE situation, and (f) shows a mix of SAFE and UNSAFE situations (color online).

Figure 5(e) shows two RED bounding boxes and a red line between the two detected person objects that are far from the camera, confirming both are UNSAFE. Figure 5(f) shows two RED bounding boxes and a red line between two detected person objects and one GREEN bounding box around one detected person object, confirming a mix of SAFE and UNSAFE scenarios.

4. **Conclusion.** We present the development of an automated physical distance monitoring system that can help companies comply with government regulations to reduce the spread of COVID-19. The system uses YOLOv3 as a backbone in detecting a person-object in a video frame and calculating the distance between them. The developed system

has been deployed in the campus library of Universitas Multimedia Nusantara. User acceptance testing has been conducted and based on the results, we show that the system can satisfactorily detect person-objects and determine if they observe or break the rule. This system can be adapted to other locations and in different settings and conditions. This will reduce the need of having many patrols around the building. In the future, the system can be improved by providing better visible and audible warnings so people can be made aware if they are breaking the rule of physical distancing.

REFERENCES

[1] Joint WHO-China Study, *WHO-Convened Global Study of Origins of SARS-CoV-2*, 2021.

[2] Center for Systems Science and Engineering (CSSE), *COVID-19 Dashboard*, Johns Hopkins University Coronavirus Resource Center, https://coronavirus.jhu.edu/map.html, Accessed on 09-Nov-2022.

[3] M. Khadapi, D. Riana, A. Arfian, E. Rahmawati et al., Public acceptance of PeduliLindungi application in the acceleration of corona virus (COVID-19) handling, *Journal of Physics: Conference Series*, vol.1641, no.1, 12026, 2020.

[4] S. Sudirman, F. Natalia, A. Sophian and A. Ashraf, Pulsed eddy current signal processing using wavelet scattering and Gaussian process regression for fast and accurate ferromagnetic material thickness measurement, *Alexandria Eng. J.*, vol.61, no.12, pp.11239-11250, 2022.

[5] C. Song, S. Sudirman, M. Merabti, P. Xiao, S. Sudirman and M. Merabti, Region adaptive digital image watermarking system using DWT-SVD algorithm, *Proc. of NASA/ESA Conference on Adaptive Hardware and Systems*, pp.196-201, 2014.

[6] S. Monica, F. Natalia and S. Sudirman, Clustering tourism object in Bali Province using K-means and X-means clustering algorithm, *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, pp.1462-1467, 2018.

[7] R. Cui, H. Liu and C. Zhang, A deep neural framework for continuous sign language recognition by iterative training, *IEEE Trans. Multimed.*, vol.21, no.7, pp.1880-1891, 2019.

[8] F. Natalia, H. Meidia, N. Afriliana, J. C. Young, R. E. Yunus, M. Al-Jumaily, A. Al-Kafri and S. Sudirman, Automated measurement of anteroposterior diameter and foraminal widths in MRI images for lumbar spinal stenosis diagnosis, *PLoS One*, vol.15, no.11, pp.1-27, 2020.

[9] F. Natalia, J. C. Young, N. Afriliana, H. Meidia, R. E. Yunus and S. Sudirman, Automated selection of mid-height intervertebral disc slice in traverse lumbar spine MRI using a combination of deep learning feature and machine learning classifier, *PLoS One*, vol.17, no.1, e0261659, 2022.

[10] L. Liu, R. Wang, C. Xie, P. Yang, F. Wang, S. Sudirman and W. Liu, PestNet: An end-to-end deep learning approach for large-scale multi-class pest detection and classification, *IEEE Access*, vol.7, pp.45301-45312, 2019.

[11] L. Liu, C. Xie, R. Wang, P. Yang, S. Sudirman, J. Zhang, R. Li and F. Wang, Deep learning based automatic multi-class wild pest monitoring approach using hybrid global and local activated features, *IEEE Trans. Ind. Informatics*, 2020.

[12] G. J. Chowdary, N. S. Punn, S. K. Sonbhadra and S. Agarwal, Face mask detection using transfer learning of InceptionV3, *International Conference on Big Data Analytics*, pp.81-90, 2020.

[13] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and F.-F. Li, ImageNet: A large-scale hierarchical image database, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR2009)*, pp.248-255, 2009.

[14] P. Bhandary, *Simulated Masked Face Dataset*, GitHub, https://github.com/prajnasb/observations, Accessed on 09-Nov-2022.

[15] S. Balaji, B. Balamurugan, T. A. Kumar, R. Rajmohan and P. P. Kumar, A brief survey on AI based face mask detection system for public places, *Irish Interdiscip. J. Sci. & Res.*, 2021.

[16] S. V. Militante and N. V. Dionisio, Deep learning implementation of facemask and physical distancing detection with alarm systems, *2020 3rd International Conference on Vocational Education and Electrical Engineering (ICVEE)*, pp.1-5, 2020.

[17] J. Redmon and A. Farhadi, YOLOv3: An incremental improvement, *arXiv.org*, arXiv: 1804.02767, 2018.

[18] K. G. S. Dixit, M. G. Chadaga, S. S. Savalgimath, G. R. Rakshith and M. R. N. Kumar, Evaluation and evolution of object detection techniques YOLO and R-CNN, *Int. J. Recent Technol. Eng.*, vol.8, DOI: 10.35940/ijrte.B1154.0782S319, 2019.

[19] S. Ren, K. He, R. Girshick and J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, *Adv. Neural Inf. Process. Syst.*, vol.28, 2015.

[20] S. Lee, S. Kwak and M. Cho, Universal bounding box regression and its applications, *Asian Conference on Computer Vision*, vol.8, pp.373-387, 2018.
[21] P. Beynon-Davies, C. Carne, H. Mackay and D. Tudhope, Rapid application development (RAD): An empirical review, *Eur. J. Inf. Syst.*, vol.8, no.3, pp.211-223, 1999.