International Journal of Technology 13(6) 1241-1250 (2022)
Received April 2022 / Revised August 2022 / Accepted August 2022

International Journal of Technology

http://ijtech.eng.ui.ac.id

Smart Retail Monitoring System using Intel OpenVINO Toolkit

Iskandar Zulkarnain Jafriz[1], Sarina Mansor[1*]

[1]*Faculty of Engineering, Multimedia University, 63100 Cyberjaya, Malaysia*

Abstract. In the era of Covid-19 infection, enforcing social distance is essential for confined areas such as shopping retails and classrooms. Human workforce is used to ensure the safety measure rules are adhered. However, a better technique to enforce social distancing regulations is to use an automated system that counts and detects people and measures the social distance. This work proposes an innovative retail monitoring system based on the Intel Distribution of Open VINO toolkit. The system uses deep learning techniques and trained models to automatically count the number of individuals, the number of persons entering and exiting premises, and the distance between each person to ensure social distancing. Five experiments were conducted to evaluate the efficiency and accuracy of the system.

Keywords: Intel OpenVINO; Machine learning; Monitoring system; Smart retail; Social distancing

1. Introduction

Several new practices and preventive measures were introduced during the Covid-19 pandemic (Berawi et al., 2020; Baroroh & Agarwal, 2022; Romadlon et al., 2022). One of them is social distancing, defined as keeping a safe gap between persons who are not in the same bubble. An infected person coughing, sneezing, or speaking may infect the next person closest to them (Gupta, 2020).

Currently, workers such as security guards and retail workers are used to checking the temperatures of their customers, counting the number of people on the premises and allocating every customer with their time limits. As humans, doing multiple tasks at once is an exhausting chore as it is difficult for us to comprehend the speed and precision required to perform those tasks. Using people to count and keep track of the time limits is an imprecise and ineffective way to control people in a space especially during peak times. Providing an automated people counter system is a more efficient way to count, detect people, and ensure social distancing.

Few products offer automated people counter system for retails stores and shopping malls, such as "FootfallCam" (FootfallCam, 2022) and SensMax (2022). However, most existing solutions require high-end hardware and software. This paper presents an intelligent retail monitoring system developed using Intel Distribution of OpenVINO toolkit. The system can run on a single-based computer (SBC) to reduce the cost of implementation. It has features to perform line monitoring, capacity limit and social distance detection.

*Corresponding author's email: sarina.mansor@mmu.edu.my, Tel.: 603-83125323
doi: 10.14716/ijtech.v13i6.5872

2. Relates Works

2.1. Line Monitoring System

A line monitoring system considered as a crowd control system. It counts people in a specific area, frame, or queue. Several applications could utilise the line monitoring system. One popular application is queue monitoring at the convenience store. Due to excessively long and poorly managed payment queues, many retail customers abandon their purchases. To solve this problem, an autonomous queue monitoring system based on computer vision has been proposed (Viriyavisuthisakul et al., 2017).

Another interesting application is monitoring vehicles at traffic signals. It has been used to improve traffic signals (Yao et al., 2013). The system keeps track of time by measuring the vehicle's length (Cai et al., 2010). The system spotted the vehicle in real-time because the camera was at the crossroads. The image was then processed using image processing methods. The bank's queue system is also considered to provide efficient customer service, where two infrared sensors are used for real-time detection of queue at the entrance and exit (Gimba et al., 2020).

2.2. People Counting System

A people counter system is a system that is designed to count the number of people that passes through a specific designated area. Previously, people would usually count the number of people in a particular area by hand. In this modern time, various technologies have been created to make counting people smoother, faster, and more accurate. Some of the new technology designed to count people are infrared sensors, using cameras, thermal sensors, Wi-Fi sensors, and many more (Brown, 2019; Hughes, 2021). Every technology has its advantages and limitations, but it all comes down to functionality and usability. For example, Arief-Ang et al. (2018) developed a method of counting and detecting people using carbon-dioxide sensors. The cost of implementation is low, but the accuracy depends on the carbon-dioxide concentrations, which can be fluctuated due to external factors.

Another work by Chang et al. (2018) detected people in a restricted area based on Wi-Fi signals. The information obtained from Wi-Fi channel state was analysed using Deep Neural Network to estimate the number of people. No additional setup was needed, but this method works for indoor environment only. The closest related work to the proposed system in the vision-based people detection system developed by Parthornratt et al. (2016). This system was deployed on Raspberry Pi board and using a Pi camera to capture images. The face detection algorithm was used to detect and count the number of people passing by. Some limitations of the system include customer must face the camera with a minimum distance of 3cm, no head covering, long processing time, and Raspberry Pi speed. Our project proposes to overcome these limitations. The preliminary work for this project has been presented in 2020 (Aslam, 2021), where a low-cost people detector has been developed. The result compared two libraries performance: OpenCV and OpenVINO. It was found out that the system with OpenVINO utilisation is faster at performing inferences and more suitable for real-time applications. Therefore, the proposed intelligent retail monitoring system will use the OpenVINO toolkit.

3. Methodology

3.1. System Architecture

Figure 1 shows the overall system architecture for the smart retail monitoring system. The system requires a PC, a monitor and a webcam. It uses Ubuntu 18.04 operating system and an Intel Distribution of OpenVINO installed on the PC. The uniqueness of the proposed system as compared to the existing ones is the utilization of Intel OpenVINO toolkit (Intel, 2020d). This toolkit has enormous number of pre-installed deep-learning models, which could speed-up the inference stage. The OpenVINO version that was used in this project was version 2020.3. The proposed system is the combination of three single projects developed by Intel, which are: Capacity Limit (Intel, 2020a), Social Distance (Intel, 2020b) and Line Monitoring (Intel, 2020c).

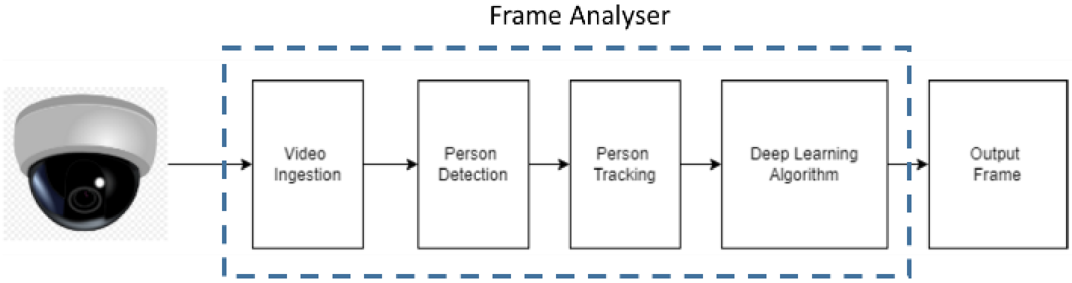

Figure 1 Overall system architecture

The system works by taking in an input video/footage, which could be a video or even live camera footage from the webcam. The OpenVINO software processes the input with its models and inference engine. To begin, the system ingests video from a file and function it frame by frame. Individuals in the frame are recognized using a pre-trained Deep Neural Network (DNN) model (Intel, 2021). To track people, the system will use a second pre-trained DNN to extract their features (Intel, 2020e). The deep learning algorithm depends on the components to be implemented, such as capacity limit, social distance, and line monitoring. The following subsections describe the implementation of these feature in detail.

3.1.1. Capacity Limit System

This is Intel's retail capacity limit application which counts people entering and exiting the store. Virtual lines are drawn in entrance and exit areas to serve as 'virtual gates' (Intel, 2020a). Figure 2 shows the block diagram of Capacity Limit system.

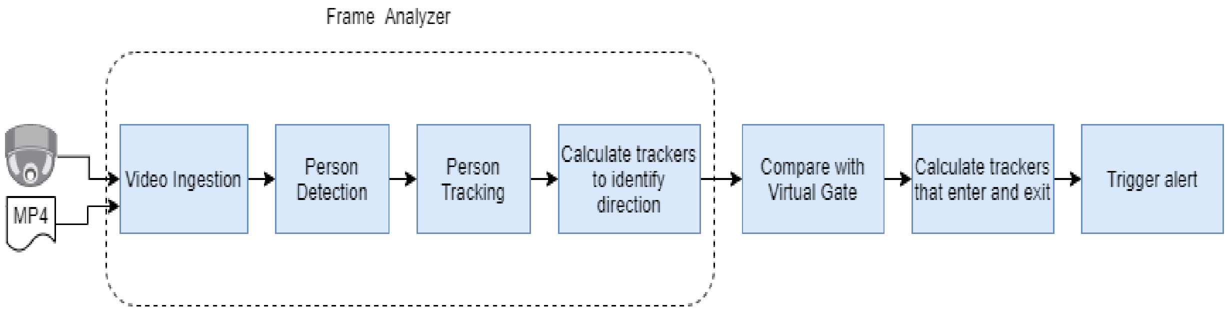

Figure 2 Block Diagram of Capacity Limit System (Intel, 2020a)

After person detection and tracking, the system will determine whether the people crossed any predetermined virtual gates based on the output frame's coordinates and whether the virtual gates indicated one direction or the other. Finally, the person counter's output is updated based on entry and exit data. If a certain number of people crossing the entry line exceeds the threshold, the system will trigger an alert that will pop up a warning at the output frame.

3.1.2. Social Distancing System

This reference implementation demonstrates a retail social distance application that identifies and measures the distance between two persons in a retail setting. If the distance between the two points is less than a value previously specified by the user, an alarm is triggered (Intel, 2020b). Figure 3 shows the block diagram of the social distance system.

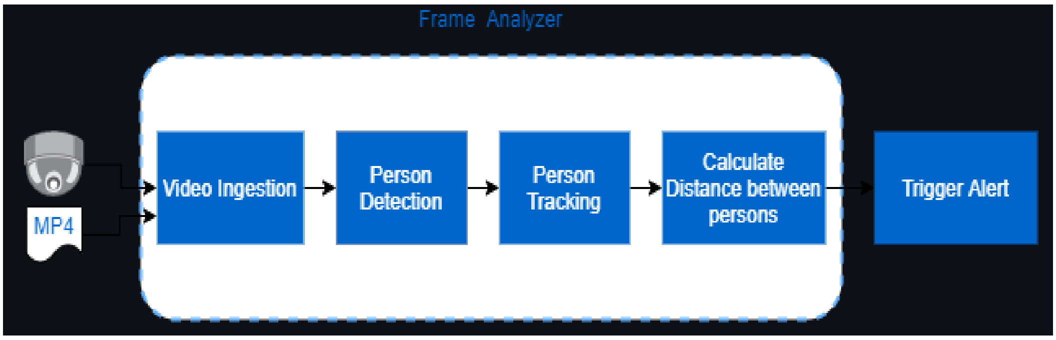

Figure 3 Block Diagram of Social Distancing System (Intel, 2020b)

The system processes the video frame by frame until the stream is complete. A DNN model will detect people in the frame of interest, and another DNN model will extract characteristics from them so they can be tracked. It then calculates the distance between two identified people based on their position, size, and viewpoint to see if the minimal social distance threshold has been exceeded.

3.1.3. Line Monitoring System

This reference implementation demonstrates a retail application that counts the number of people who are waiting in a retail store's waiting queue. The number of persons in a line is estimated by the program’s algorithm by performing an intersection between the people who have been identified in the frame (Intel, 2020c).

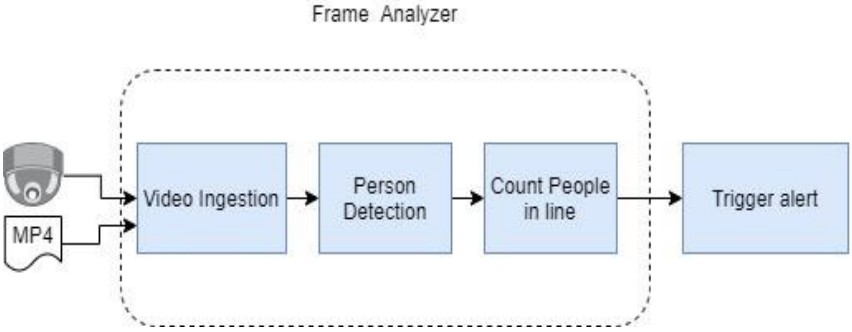

Figure 4 Block Diagram of Line Monitoring System (Intel, 2020c)

Figure 4 shows the block diagram of Line Monitoring system. This method has been integrated with the Capacity Limit System to provide a more efficient and user-friendly system for retails. This system detects individuals waiting outside the retail establishment prior entering.

3.2. Pre-trained Models

3.2.1. Pre-trained Model 1: People-Detection-Retail-0013

The people-detection-retail-0013 is a pre-trained model developed by Intel for person detection application. This model was obtained from Open Model Zoo and can be downloaded using their model downloader. It has 88.62 percent accuracy, uses the Caffe framework, and supports occluded pedestrians, among other features. This model uses the FP32 format, which is a single-precision floating-point format. This model uses a MobileNetV2-like backbone with depth-wise convolutions to reduce the number of calculations required for the 3x3 convolution block (Intel, 2021).

3.2.2. Pre-trained Model 2: People-Reidentification-Retail-0030

Intel also developed pre-trained models for reidentification purpose. In this work, the people-reidentification-retail-0300 model is used. This model was obtained from Open Model Zoo and can be downloaded using their model downloader. As input, it takes a fullbody image and outputs an embedding vector that can be used to compare two images using cosine distance. The model is built on the OmniScaleNet backbone for rapid inference. A single reidentification head extracted from the 1/16 scale feature map generates a 512-float embedding vector (Intel, 2020e).

4. Results and Discussion

The system was assessed with five different experiments to evaluate its efficiency. The experiment focuses on analysing the accuracy of detecting people, the calculation for social distancing, and the counting for number of people entering and exiting a particular premise. There were five volunteers involved in the experiment. All the findings from the experiment are shown and discussed in Section 4.1. Section 4.2 evaluates the overall system performance with 6 videos in different environments.

4.1. Experimental Results

4.1.1. Experiment 1: Angle and Distance Test

This test is designed to determine the optimal camera angle and distance for detecting people. The camera is tested in three different locations. The first camera setup is close-range, which is at the body level of participants. The second camera set-up is medium range, which provides a "bird's-eye" view of the entire specified area. Finally, the camera was positioned at the top of the area, pointing downward, providing an overhead perspective of the people in that location. This experiment is tested using the Social Distancing System.

Figure 5 The First Camera Setup (Close-Range, Body-Level)

The output frame of the first camera setup is shown in Figure 5. As illustrated in the figure, all people can be identified except one. A green bounding box for two people is drawn, as indicated by the arrow in Figure 5. This demonstrates that the system cannot detect the individual who is overlapping with another individual in front of them.

Figure 6 The Second Camera Setup (Medium-Range, Bird's Eyed View)

The output frame of the second camera arrangement is shown in Figure 6. It is positioned above the volunteer's head and a short distance from the first arrangement. As can be seen, all the volunteers are visible from the camera's perspective, were detected by the system, and were all drawn by red and yellow bounding boxes.

Figure 7 The Third Camera Setup (Long-Ranged, Overhead View)

Finally, the camera is positioned on top of the building in Figure 7 to provide an overhead perspective of the volunteers. Only one green bounding box was generated for one of the five volunteers in the frame, and one of them was not spotted by the system. In summary, the optimal camera placement is at a medium range, providing a "bird's-eye" view of the entire specified area. This is because it will have a clear picture of people without overlapping them.

4.1.2. Experiment 2: People Counting System Efficiency Test

This test is used to determine the efficiency of the people counter system. The experiment is tested with the Line Monitoring system. The experiment is set up by having volunteers walk inside the testing area one by one. The results of the test are shown in Figure 8. The system is not able to detect a person who was the furthest away from the camera and outside the detecting region of the system (as pointed by arrow in Figure 8).

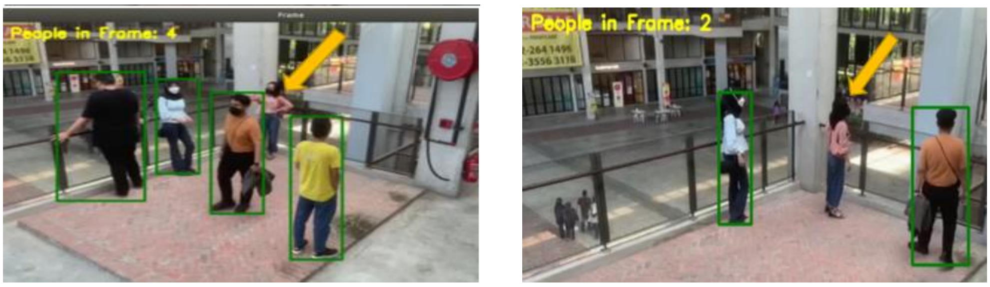

Figure 8 People Counting System

4.1.3. Experiment 3: Social Distancing System Efficiency Test

The social distancing efficiency test aim is to evaluate the Social Distance system's reliability. The experiment is designed to see whether the system recognizes instances of social distancing violation between volunteers. The study begins with two people being physically close, followed by three, and lastly five. This section presents and discusses the experiment's outcomes. The threshold for the social distancing violation to trigger is approximately 1m.

Figure 9 demonstrates that the system recognizes both individuals and acknowledges that they have violated the social distancing rule when they come close because the distance between them is less than the predefined threshold.

Figure 9 Two people in the Social Distancing System

Figure 10 below shows five people in the social distance system. The efficiency test results showed that the system works nearly flawlessly, as it captured all of the participants in the frame and detected them all violating the social distance limit of 1 meter. When the social distance rule is violated, an alert will be displayed on the output frame.

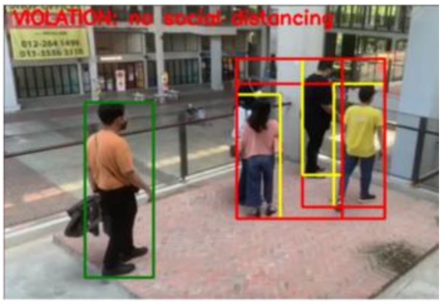

Figure 10 Five people in the Social Distancing System

<u>4.1.4. Experiment 4: Capacity Limit System Efficiency Test</u>

It was planned for this test to analyse the system's capacity limit for persons entering and exiting the premises. By combining the Capacity Limit and Line Monitoring systems, the experiment was designed to determine whether the system is functioning well and accomplishing its intended goal. A two-person admission threshold was established to ensure that the system complied with the requirement. This section illustrates and discusses the results of this experiment.

(a) (b)

Figure 11 Five people in the Social Distancing System

As presented in Figure 11(a), the system identified two individuals crossing the yellow line that indicates they are entering the premise. However, even though one of them struck the line completely, the output of people counting was not yet added up until the frame where he disappeared. The fourth and last individual to enter the premises is seen in Figure 11(b). As can be seen, once the people counter reach 3, the threshold is breached and a warning appears on top of the output frame. Additionally, the system's terminal publishes a text indicating that individuals are entering the premises and "VIOLATION: TOO MANY PEOPLE" when the threshold is triggered.

4.1.5. Experiment 5: Lighting Test

The lighting test is the last assessment of the system. This is to determine whether the system can capture and detect a person under various lighting conditions. The experiment was conducted in the testing room with the brightest, dimmest, and nearly total darkness. The results of this experiment are presented and covered in this section.

(a) (b)

Figure 12 (a) Brightest Lighting and (b) Dimmest Lighting

As shown in Figure 12, the system detects the subject in the testing room's brightest and darkest ambient light. The lighting test demonstrated that the system is capable of detecting humans in a variety of lighting conditions. This means that the system may be utilized both during the day and at night in somewhat dark areas, particularly outside the premises.

4.2. System Performance

Recorded videos were run through the proposed smart retail monitoring system to evaluate its efficiency. Six videos in different environments were used as shown in Table 1.

Table 1 Video Dataset

Video	Description
SC001	Video recorded in an open space outside of a retail shop consists of 5 volunteers
SC002	A recording of people walking in an open space from (Rosebrock, 2020)
LM001	Video recorded in an open space outside of a retail shop consists of 5 volunteers
LM002	Video recorded in front of a retail shop consists of people queuing up before entering the premises
LM003	A recording of people queuing up before entering a mall from (Gunning, 2020)
CL001	Video recorded in front of a classroom which consists of 4 people entering and exiting

Table 2 shows the accuracy of each video when using the smart retail monitoring system. The overall system performance is good, especially for video with good lighting and camera setup. SC002 video has lower accuracy for the social distance system because the camera position is placed at the top-view, which is not the optimum location. LM002 and LM003 videos have low resolution for the line monitoring system, and the camera is positioned at close range (body level). These factors have reduced the accuracy of the system. For capacity limit, CL001 has achieved 100% accuracy.

Table 2 Accuracy of Video Dataset

System		Result	Accuracy
Social Distance	SC001	5 detections, 5 Red and Green Bounding-Boxes	80%
	SC002	4 violations of social distance detected from 6 groups	66.67%
Line Monitoring	LM001	All people detected in the frame	100%
	LM002	6 out of 9 people waiting in line detected	66.67%
	LM003	5 out of 9 people waiting in line detected	55.55%
Capacity Limit	CL001	4 people counted crossing in and out yellow line	100%

5. Conclusions

An intelligent retail monitoring system using Intel Distribution of OpenVINO™ Toolkit has been developed for this project. The system managed to count the number of persons within the frame, the number of people entering and exiting the premises, and to determine the distance between them for social distancing. The system is adequate to be implemented in retails to comply with the SOP and prevent the spread of coronavirus. For future works, other deep learning methods can be explored especially for the capacity limit system. In addition, a graphical user interface can be developed to integrate the three systems altogether.

References

Arief-Ang, I.B., Hamilton, M., Salim, F.D., 2018. RUP: Large Room Utilisation Prediction with Carbon Dioxide Sensor. *Pervasive and Mobile Computing*, Volume 46, pp. 49–72

Aslam, M., 2021. *People Counter and Detection System Using Intel Up Squared Board*. Final Year Project Report, Faculty of Engineering Multimedia University

Baroroh, D.K., Agarwal, A., 2022. Immersive Technologies in Indonesia Faces "New Normal" COVID-19. *International Journal of Technology*. Volume 13(3), pp. 633–642

Berawi, M.A., Suwartha, N., Kusrini, E., Yuwono, A.H., Harwahyu, R., Setiawan, E.A., Yatmo, Y.A., Atmodiwirjo, P., Zagloel, Y.T., Suryanegara, M., Putra, N., Budiyanto, M.A., Whulanza, Y., 2020. Tackling the COVID-19 Pandemic: Managing the Cause, Spread, and Impact. *International Journal of Technology*. Volume 11(2), pp. 209–214

Brown, B., 2019. Pros and Cons of 5 Different People Counting Sensors. Available online at https://www.getdor.com/blog/2019/11/19/people-counting-door-counter-sensors/, Accessed on January 18, 2021

Cai, Y., Zhang, W., Wang, H., 2010. Measurement of Vehicle Queue Length Based on Video Processing in Intelligent Traffic Signal Control System. *In: I*nternational Conference on Measuring Technology and Mechatronics Automation, Volume 2, pp. 615–618

Chang, R.Y., Liu, S. J., Cheng, Y.K., 2018. Device-Free Indoor Localization Using Wi-Fi Channel State Information for Internet of Things. *IEEE Global Communications Conference*, pp. 1–7

FootfallCam, 2022. World's No. 1 People Counting System. Available online at https://www.footfallcam.com/en/, Accessed on April 30, 2022

Gimba, U.A., Okoronkwo, C.D., Yusuf, M., Musa, A.S., Ali, M.S., 2020. Queue Monitoring System for Bank. Dutse Journal of Pure and Applied Sciences, Volume 6(2), pp. 269–276

Gunning, E., 2020. People queuing at Prestons Liquor stores. Available online at https://www.youtube.com/watch?v=3EoaAlAImV0, Accessed on April 25, 2022

Gupta, S., 2020. Social Distancing. In: *Social Analysis and the COVID-19 Crisis*, Routledge

India, pp. 19–31
Hughes, J., 2021. People Counting Using WIFI and Bluetooth. Available online at https://www.manxtechgroup.com/people-counting-using-wifi-and-bluetooth/, Accessed on September 24, 2021
Intel, 2020a. Capacity Limit Python by Intel. Available online at https://github.com/intel-iot-devkit/capacity-limit, Accessed on September 24, 2021
Intel, 2020b. Social Distance Python by Intel. Available online at https://github.com/intel-iot-devkit/social-distance, Accessed on September 24, 2021
Intel, 2020c. Line Monitoring by Intel. Available online at https://github.com/intel-iot-devkit/line-monitoring, Accessed on September 24, 2021
Intel, 2020d. Intel Distribution of OpenVINO Toolkit. Available online at https://www.intel.com/content/www/us/en/developer/tools/openvino-toolkit/overview.html, Accessed on September 26, 2021
Intel, 2020e. Person-Reidentification-Retail-0300 OpenVINO Toolkit. Available online at https://docs.openvino.ai/2020.3/_models_intel_person_reidentification_retail_0300_description_person_reidentification_retail_0300.html, Accessed on September 30, 2021
Intel, 2021. Person-Detection-Retail-0013: OpenVINO Toolkit. Available online at https://docs.openvino.ai/2021.2/omz_models_intel_person_detection_retail_0013_description_person_detection_retail_0013.html, Accesed on September 30, 2021
Parthornratt, T., Burapanonte, N., Gunjarueg, W., 2016. People Identification and Counting System Using Raspberry Pi (AU-Picc: Raspberry Pi Customer Counter). *In*: International Conference on Electronics, Information, and Communications (ICEIC), pp. 1–5
Romadlon, F., Lestiana, F., Putri, N.A., 2022. An Exploration of Personal Decision as Mediating Effect between Passenger Concern and Airport Service Information During COVID-19 Outbreak. *International Journal of Technology*. Volume 13(3), pp. 664–676
Rosebrock, A., 2020. OpenCV Social Distancing Detector. Available online at https://pyimagesearch.com/2020/06/01/opencv-social-distancing-detector/, Accessed on August 29, 2021
SensMax, 2022. People Counting System for Retail Stores. Available Online at: https://sensmax.eu/solutions/people-counting-system-for-retail-stores/, Accessed on April 30, 2022
Viriyavisuthisakul, S., Sanguansat, P., Toriumi, S., Hayashi, M., Yamasaki, T., 2017. Automatic Queue Monitoring in Store using a Low-cost IoT Sensing Platform. *In*: IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW), pp. 53–54
Yao, Y., Wang, K., Xiong, G., 2013. Embedded Technology and Algorithm for Video-based Vehicle Queue Length Detection. *In*: IEEE International Conference on Service Operations and Logistics and Informatics, pp. 45–50