JRAEE

Journal of Robotics, Automation, and Electronics Engineering
Vol. 2, No. 1, 2024, pp. 20-29
e-ISSN 3025-3780
https://journal.uny.ac.id/publications/jraee

Design and Implementation of a Student Counting and Monitoring System in a Laboratory Using Human Tracking Method with OpenCV and TensorFlow

Nancy Febriani Taek[a,1], Arya Sony[a,2,*]

[a] Universitas Negeri Yogyakarta

[1] nancyfebriani.2019@student.uny.ac.id; [2] arya.sony@uny.ac.id

[*] Corresponding Author

ARTICLE INFO

Article History
Received 10 Dec. 2023
Revised 28 Jan. 2024
Accepted 5 Mar. 2024

Keywords
Laboratory;
Object Detection;
Human Tracking;
Machine Learning.

ABSTRACT

Laboratories serve as crucial facilities supporting practical activities, with a recommended maximum of 20 students, necessitating periodic monitoring to count the dynamic number of students within. The system utilizes the COCO dataset labeled ”person,” involving an approach with entry and exit preference lines, ID identification implementation, and object detection models YOLO v3 Tiny and Faster R-CNN ResNet50. The main system components, Raspberry Pi 3 Model B+, Raspberry Pi Camera 5 MP (f/1.3), and Raspberry Pi 7-inch Touch Display, are integrated for processing, real-time video recording, and image display functions. Test and evaluation results reveal that YOLO v3 Tiny achieves an 88.24% accuracy for entry counting and 75% for entry-exit counting, with an average processing rate of 4.89 FPS, while Faster R-CNN ResNet50 demonstrates lower accuracy, reaching 70.59% and 45.83%, with an average processing rate of 0.58 FPS.

Laboratorium sebagai fasilitas penting pendukung kegiatan praktik, dengan jumlah siswa maksimum yang direkomendasikan adalah 20 orang, memerlukan pemantauan berkala guna menghitung dinamika jumlah siswa yang berada di dalamnya. Program sistem memanfaatkan dataset COCO yang berlabel ”person” juga pendekatan yang melibatkan garis preferensi masuk dan keluar, penerapan identifikasi ID, dan model deteksi objek YOLO v3 Tiny dan Faster R-CNN ResNet50. Komponen utama sistem, yakni Raspberry Pi 3 Model B+, Raspberry Pi Camera 5MP (f/1.3), dan Raspberry Pi 7-inch Touch Display, saling terintegrasi untuk fungsi pemrosesan, perekaman video real-time, dan penampilan citra. Hasil pengujian dan evaluasi menunjukkan YOLO v3 Tiny berhasil mencapai akurasi 88.24% untuk penghitungan masuk dan 75% untuk penghitungan masuk-keluar, dengan rata-rata pemrosesan 4.89 FPS, sementara Faster R-CNN ResNet50 menunjukkan akurasi yang lebih rendah yaitu 70.59% dan 45.83%, dengan rata-rata pemrosesan 0.58 FPS.

1. Introduction

The laboratory serves as a critical facility supporting practical activities, necessitating attention to the number of students present during its use. Generally, the recommended maximum number of students is 20. This regulation is reinforced by the Decree of the Minister of National Education of the Republic of Indonesia Number 234/U/2000 on Guidelines for the Establishment of Higher Education Institutions, the Higher Education Law Number 12/2012, and Government Regulation Number 4/2014 on the Implementation of Higher Education, which state that the ideal ratio of lecturers to students is 1:20 for Exact Sciences and 1:30 for Social Sciences.

Although considered an important task, tracking the number of students is a complex challenge that requires a solution such as developing a tool for automatically monitoring the number of students in the laboratory by applying digital image processing techniques. One of the challenges of image processing in the context of image applications is computer vision, which involves creating digital images from original images according to human perception [1]. This technology can identify objects and classify them as human or non-human [2]. Subsequently, it can count the number of students in a video frame. However, several implementations still face challenges, particularly related to modeling the multi-object tracking process, as evidenced by research on people counting systems for attendance monitoring [3].

The application of human tracking technology, normally developed by utilizing reference lines on an image spatial area, can effectively monitor the movement and dynamics of students in the laboratory [4]. The object detection models used throughout the detection process and for counting students in the laboratory are YOLO v3 Tiny and Faster R-CNN ResNet50. These models are implemented using OpenCV and TensorFlow, with Raspberry Pi 3 Model B+ serving as the primary processor in the system.

2. Methods

2.1 Raspberry Pi 3 Model B+

The Raspberry Pi is a credit card-sized Single Board Computer (SBC) equipped with an integrated circuit chip that functions as a system-on-chip (SoC), providing general processing capabilities, graphics rendering, and input/output. The Raspberry Pi 3 Model B+ runs on Raspberry Pi OS and serves as the main processor in the system. The appearance of the Raspberry Pi 3 Model B+ is shown in Fig. 1.

Fig. 1. Raspberry Pi 3 model B+

2.2 Mechanical Design of the Device

The main components of the system device include the Raspberry Pi 3 Model B+, Raspberry Pi Camera 5 MP (f/1.3), and Raspberry Pi 7-inch Touch Display, along with supporting components such as the power supply and other voltage sources, designed to be housed in a box container. The dimensions of this device enclosure measure 256 x 180 x 150 mm. The mechanical design of the device can be viewed in Fig. 2.

Journal of Robotics, Automation, and Electronics Engineering
Vol. 2, No. 1, 2024, pp. 20-29
ISSN 3025-3780 (online)
ISSN 3025-4590 (printed)

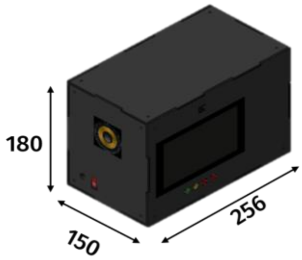

Fig. 2. Mechanical design of the device

2.3 Electronic System Design of the Device

The electronic design layout is formulated to determine the placement of each required component. The Raspberry Pi 3 Model B+ serves as the main processor controlling the system. It can be observed that the Raspberry Pi Camera is connected via the CSI connector port, and the Raspberry Pi Touch Display is connected via the HDMI port. Power is supplied through the Raspberry Pi adapter via the micro-USB power port. The diagram illustrating the Electronic System Design of the Device can be seen in Fig. 3.

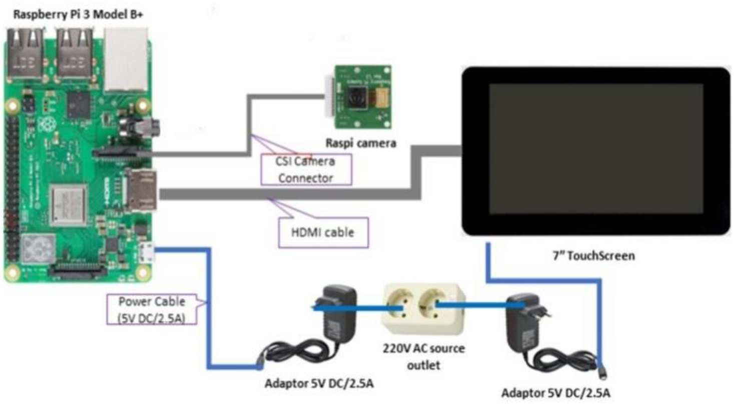

Fig. 3. Electronic system design of the device

2.4 COCO Dataset

The COCO (Common Objects in Context) dataset provides a large collection of data covering various labeled categories, including humans, vehicles, animals, and other domestic objects, as illustrated in Fig. 4. This research specifically focuses on object detection with the label "person". Both object detection models used in this study have been trained with this dataset. In YOLO v3 Tiny, the class ID 0 corresponds to "person" in the COCO names, while in Faster R-CNN ResNet50, the class ID 1 corresponds to "person" in the COCO dataset [5].

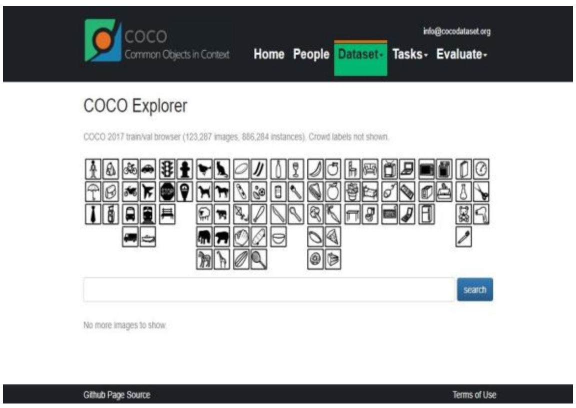

Fig. 4. COCO dataset collection

2.5 Tracking Method

Vertical preference lines are utilized as indicators for the entry and exit paths of detected students within the room, captured by the camera. The system employs two vertical preference lines in the image display: a blue entry line and a red exit line. The monitoring process begins by marking individuals detected by the camera with a center point, followed by tracking their movement direction [6]. When a detected person moves toward and crosses the entry preference line, the system logs it as an entry event; conversely, moving toward and crossing the exit preference line is logged as an exit event [7]. Information regarding the count of entry and exit events generated by the system is displayed on the image. The design layout of the preference lines system is depicted in Fig. 5.

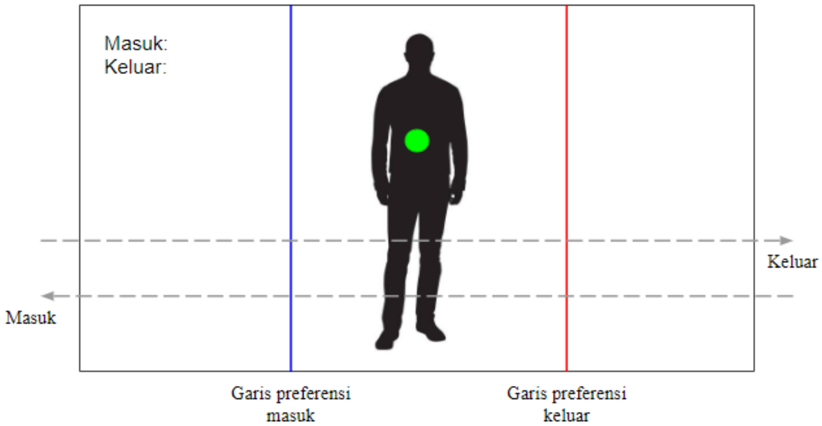

Fig. 5. Design of preference line display

2.6 Data Acquisition and Evaluation

The image preference line tracking method approach is implemented in both YOLO v3 Tiny and Faster R-CNN ResNet50 detection models [8]. Documented data components include a timestamp indicating the current measurement time, Person ID representing a unique identification for each detected student, Entry and exit counts recorded by the system compared to actual counts, and Confidence value indicating the model's prediction certainty that an object in the frame is a human.

Evaluation of the test results also utilizes error percentage (PE) and accuracy metrics, as follows:

$$PE = \frac{|x - x^{\wedge}|}{x} \times 100\% \quad (1)$$

$$Accuracy = 100\% - PE \quad (2)$$

Where,

x = Number of people tested
$x^{\wedge}$ = Number of people recorded

3. Result and Discussion

3.1 Implementation and Hardware Testing

The device enclosure is constructed from 2 mm thick acrylic using cutting techniques, and assembled by connecting each corner with nuts and bolts. Following this, the main system components and supporting elements are mounted, and wiring installation is performed for component interconnection. The device enclosure is also equipped with a 1360 mm tall tripod. The implementation of the mechanical and electronic design results is shown in Fig.6.

Journal of Robotics, Automation, and Electronics Engineering
Vol. 2, No. 1, 2024, pp. 20-29
ISSN 3025-3780 (online)
ISSN 3025-4590 (printed)

Fig. 6. Implementation of device design

The testing was conducted to measure the system's performance, starting with the evaluation of the hardware system performance. The results of the hardware performance testing are presented in Table 1 below.

Table 1. Hardware System Testing

Main Component	Status
Raspberry Pi 3 Model B+	Working
Raspberry Pi Camera 5 MP (f/1.3)	Working
Raspberry Pi 7" Touch Display	Working

The table above shows the test results of the main hardware components used in the development of this student count detection system in the laboratory, namely the Raspberry Pi 3 Model B+, Raspberry Pi Camera 5MP (f/1.3), and Raspberry Pi 7" Touch Display, all of which operate effectively.

3.2 Entry Count Testing

The entry count testing process involves the participation of 17 individuals entering the room regularly, as shown in Fig. 7.

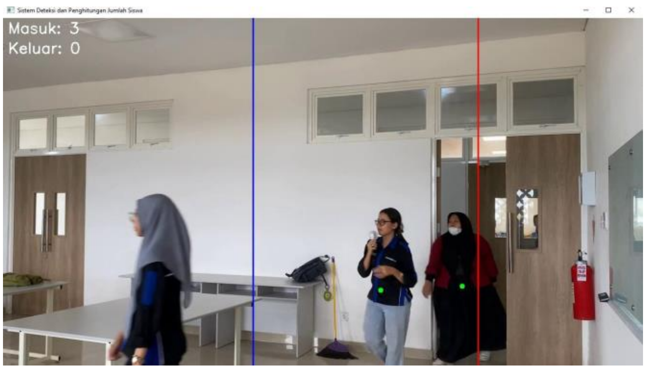

Fig. 7. Entry count testing

The results of this testing are presented in Table 2 for entry count testing using YOLO v3 Tiny, and in Table 2 for entry count testing using Faster R-CNN ResNet50, as follows.

Table 2. Entry Count Testing YOLO v3 Tiny

Timestamp (02/12/2023)	Person ID	Expected Entry	Expected Exit	Recorded Entry	Recorded Exit	Confidence Score
12:03:12	person01	1	0	1	0	91.66%
12:03:15	person02	2	0	2	0	90.18%
12:03:27	person03	3	0	4	0	92.58%
12:03:42	person04	4	0	5	0	91.29%
12:03:51	person05	5	0	6	0	98.61%
12:03:54	person01	6	0	7	0	97.73%
12:04:05	person02	7	0	8	0	88.75%
12:04:15	person03	8	0	9	0	96.80%
12:04:20	person04	9	0	10	0	92.69%
12:04:35	person01	10	0	11	0	87.01%
12:04:40	person02	11	0	12	0	95.01%
12:04:54	person03	12	0	13	0	95.63%
12:05:00	person04	13	0	14	0	91.94%
12:05:40	person01	14	0	15	0	99.09%
12:06:00	person02	15	0	17	0	98.33%
12:06:05	person03	16	0	18	0	97.26%
12:06:28	person04	17	0	19	0	94.94%

The testing results using YOLO v3 Tiny indicate that the ID assignment can correctly reference up to 5 individuals, but there were 2 instances where different individuals were counted redundantly, specifically the 3rd and 15th persons. The average confidence value displayed during testing reached 94.09%.

Table 3. Entry Count Testing Faster R-CNN ResNet50

Timestamp (02/12/2023)	Person ID	Expected Entry	Expected Exit	Recorded Entry	Recorded Exit	Confidence Score
02:13:21	person01	1	0	1	0	91.66%
02:13:33	person02	2	0	2	0	87.98%
02:13:37	person03	3	0	4	0	92.58%
02:13:51	person01	4	0	5	1	91.29%
02:14:00	person02	5	0	6	0	98.61%
02:14:23	person01	6	0	7	1	95.50%
02:14:26	person02	7	0	8	1	88.66%
02:14:45	person01	8	0	9	0	85.69%
02:14:51	person02	9	0	10	0	95.01%
02:15:12	person03	10	0	11	1	91.94%
02:15:17	person01	11	0	12	0	97.59%
02:15:25	person01	12	0	12	0	99.45%
02:15:59	person02	13	0	13	1	95.72%
02:16:13	person01	14	0	14	1	96.93%
02:16:14	person02	15	0	15	0	98.33%
02:16:19	person01	16	0	16	0	96.77%
02:16:40	person01	17	0	16	0	94.80%

The testing results using Faster R-CNN ResNet50 show that ID assignment can correctly reference up to 3 individuals, with one individual counted redundantly, specifically the 3rd person [9]. Additionally, 2 individuals were not counted, specifically the 12th and 17th persons, and there were errors in counting individuals exiting. The average confidence value displayed during testing reached 94.03%.

 ISSN 3025-3780 (online)
ISSN 3025-4590 (printed)

3.3 Entry-Exit Count Testing

The entry-exit count testing process involves the participation of 17 individuals entering and leaving the room regularly, as shown in Fig. 8.

Fig. 8. Entry-exit count testing

The results of this testing are presented in Table 4 for entry-exit count testing using YOLO v3 Tiny, and in Table 5 for entry-exit count testing using Faster R-CNN ResNet50, as follows.

Table 4. Entry-Exit Count Testing YOLO v3 Tiny

Timestamp (02/12/2023)	Person ID	Expected Entry	Expected Exit	Recorded Entry	Recorded Exit	Confidence Score
16:40:51	person01	1	0	1	1	97.55%
16:41:18	person02	1	1	1	2	95.80%
16:41:23	person01	2	1	2	2	97.57%
16:41:52	person02	2	2	2	3	95.54%
16:42:34	person01	3	2	3	3	94.02%
16:44:33	person01	3	3	4	3	96.56%
16:46:57	person01	4	3	5	3	97.51%
16:47:02	person02	5	3	6	4	97.00%
16:47:07	person01	5	4	6	5	97.19%
16:48:00	person02	5	5	7	6	99.18%
16:48:01	person01	6	5	7	6	97.72%
16:48:05	person02	7	5	8	6	95.71%
16:48:11	person03	8	5	9	7	97.45%
16:49:45	person01	8	6	9	8	91.39%
16:49:46	person02	8	7	9	9	97.38%
16:49:50	person03	8	8	10	9	90.76%
16:49:51	person04	8	9	10	10	92.22%
16:50:55	person01	9	9	10	11	98.66%
16:50:59	person01	10	9	11	11	95.30%
16:51:47	person01	10	10	12	12	95.12%
16:51:52	person02	10	11	12	12	96.64%
16:53:38	person01	11	11	15	13	95.53%
16:53:40	person02	11	11	16	13	90.60%
16:53:44	person03	13	11	17	13	97.96%

The testing results using YOLO v3 Tiny indicate that ID assignment can correctly reference up to 4 individuals. The expected counts for entries and exits were 13 and 11 respectively, but the system counted 17 entries and 13 exits. The average confidence value displayed during testing reached 95.85%.

Table 5. Entry-Exit Count Testing Faster R-CNN ResNet50

Timestamp (02/12/2023)	Person ID	Expected Entry	Expected Exit	Recorded Entry	Recorded Exit	Confidence Score
00:47:50	person01	1	0	1	1	97.55%
00:48:17	person01	1	1	1	2	95.80%
00:48:23	person01	2	1	2	2	97.57%
00:48:52	person01	2	2	2	3	95.54%
00:49:28	person01	3	2	3	3	94.02%
00:49:57	person01	3	3	3	4	98.41%
00:51:19	person01	4	3	4	4	98.66%
00:51:23	person01	5	3	6	5	96.51%
00:52:30	person01	5	4	6	5	94.61%
00:53:51	person01	5	5	7	6	95.90%
00:53:54	person01	6	5	8	6	96.77%
00:53:57	person01	7	5	8	7	97.00%
00:54:55	person01	8	5	9	7	99.18%
00:54:56	person01	8	6	9	8	98.93%
00:54:58	person02	8	7	9	9	97.72%
00:55:18	person01	8	8	11	10	96.44%
00:55:26	person01	8	9	11	11	98.24%
00:56:42	person01	9	9	13	12	97.55%
00:57:53	person01	10	9	14	14	99.09%
00:57:54	person01	10	10	16	15	94.47%
00:58:39	person01	10	11	17	16	95.12%
00:58:45	person01	11	11	18	17	96.23%
01:00:18	person02	12	11	19	17	93.24%
01:00:34	person01	13	11	20	17	97.96%

The testing results using YOLO v3 Tiny indicate that ID assignment was not accurately maintained. The expected counts for entries and exits were 13 and 11 respectively, but the system counted 20 entries and 17 exits. The average confidence value displayed during testing reached 96.77%.

3.4 System Performance

The system performance evaluation in counting the number of entries, along with the results obtained from each model, is also detailed in Table 6 as follows.

Table 6. System Performance

Timestamp (02/12/2023)	Person ID	Expected Entry	Expected Exit	Recorded Entry	Recorded Exit
Entry	YOLO v3 Tiny	17+0	19+0	11.76%	88.24%
	Faster R-CNN ResNet50	17+0	16+6	29.41%	70.59%
Entry-Exit	YOLO v3 Tiny	13+11	17+13	25%	75%
	Faster R-CNN ResNet50	13+11	20+17	54.17%	45.83%

Based on the table above, the performance evaluation results of the system in counting the number of student entries and exits show that the YOLO v3 Tiny model has higher accuracy compared to the Faster R-CNN ResNet50 model in both counting schemes [10].

During processing, YOLO v3 Tiny demonstrated better performance with an average FPS measured during entry-counting testing at 4.89 FPS and entry-exit counting at 4.17 FPS. In contrast, Faster R-CNN ResNet50 showed an average FPS measured during entry-counting testing at 0.58 FPS and entry-exit counting at 0.51 FPS.

4. Conclusion

The student monitoring system in the laboratory using OpenCV and TensorFlow-based human tracking method was designed and constructed with the main components: Raspberry Pi 3 Model B+, Raspberry Pi Camera 5MP (f/1.3), and Raspberry Pi 7-inch Touch Display, integrated for processing, real-time video recording, and image display functions. These main components, along with supporting elements, were housed in an acrylic-constructed device enclosure with a thickness of 2mm. The device enclosure dimensions are 256x180x150 mm. The device is also equipped with a 1360 mm tall tripod for support. Connections between the main system components involve linking each port on the Raspberry Pi 3 Model B+ to the Raspberry Pi Camera 5 MP (f/1.3) via the CSI camera port and the Raspberry Pi 7-inch Touch Display via the DSI display port.

Based on the testing and evaluation results, in the entry count scheme, YOLO v3 Tiny successfully detected 19 entry cases out of a total of 17 actual cases, resulting in an error rate of 11.76% and an accuracy of 88.24%. On the other hand, Faster R-CNN ResNet50 detected 16 entry cases out of 17 actual cases with misidentification leading to falsely counted exits, resulting in an error rate of 29.41% and an accuracy of 70.59%. In the entry-exit count scheme, YOLO v3 Tiny detected 30 entry-exit cases out of 24 actual cases, resulting in an error rate of 25% and an accuracy of 75%. Conversely, Faster R-CNN ResNet50 showed less satisfactory results by detecting 37 entry-exit cases out of 24 actual cases, with an error rate of 54.17% and an accuracy of only 45.83%. In both counting schemes, YOLO v3 Tiny demonstrated higher accuracy compared to Faster R-CNN ResNet50. The entry count scheme showed relatively higher accuracy compared to the entry-exit count scheme for both models. Additionally, during processing, YOLO v3 Tiny exhibited better performance with an average FPS of 4.89 FPS, whereas Faster R-CNN ResNet50 only achieved 0.58 FPS. Therefore, YOLO v3 Tiny was selected and implemented for the student monitoring system in the laboratory environment.

In the future, we will Improve the system performance by optimizing hardware resource utilization, especially on devices with limitations such as the Raspberry Pi 3 Model B+, by migrating to devices with higher specifications, such as the latest Raspberry Pi models or other types of processors. The use of more advanced hardware can enhance performance and processing capabilities, particularly in terms of processing speed like increasing FPS, enabling the system to operate more efficiently and responsively.

References

[1] F. A. Hariz, I. N. Yulita, and I. Suryana, "Human Activity Recognition Berdasarkan Tangkapan Webcam Menggunakan Metode Convolutional Neural Network (CNN) Dengan Arsitektur MobileNet," JITSI: Jurnal Ilmiah Teknologi Sistem Informasi, vol. 3, no. 4, pp. 103-115, 2022, doi: 10.30630/jitsi.3.4.97.

[2] M. I. H. Azhar, F. H. K. Zaman, N. M. Tahir, and H. Hashim, "People Tracking System Using DeepSORT," In 2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), 2020, pp. 137-141, doi: 10.1109/ICCSCE50387.2020.9204956.

[3] M. C. Le, M. H. Le, and M. T. Duong, "Vision-Based People Counting for Attendance Monitoring System," In 2020 5th International Conference on Green Technology and Sustainable Development (GTSD) IEEE, Nov. 2020, pp. 349-352, doi: 10.1109/GTSD50082.2020.9303117.

[4] F. Indaryanto, A. Nugroho, and A. F. Suni, "Aplikasi Penghitung Jarak Dan Jumlah Orang Berbasis YOLO Sebagai Protokol Kesehatan COVID-19," Edu Komputika Journal, vol. 8, no. 1, pp. 31-38, 2021, doi: 10.15294/edukomputika.v8i1.47837.

[5] J. Luo, Y. Wang, and Y. Wang, "Real-Time Pedestrian Detection Method Based on Improved YOLOV3," In Journal of Physics: Conference Series, vol. 1453, no. 1, p. 012149, 2020, doi; 10.1088/1742-6596/1453/1/012149.

[6] G. Gallo, F. Di Rienzo, F. Garzelli, P. Ducange, and C. Vallati, “A Smart System for Personal Pro- Tective Equipment Detection in Industrial Environments Based on Deep Learning at the Edge,” IEEE Access, 2022, pp. 110862-110878, doi: 10.1109/ACCESS.2022.3215148.

[7] S. A. Velastin, R. Ferna´ndez, J. E. Espinosa, & A. Bay, “Detecting, Tracking and Counting People Getting On/off a Metropolitan Train Using a Standard Video Camera,” Sensors, vol. 20, no. 21, p. 6251, 2020, doi: 10.3390/s20216251.

[8] W. Liu, X. Du, Q. Geng, J. Li, H. Li, and L. Liu, “Metro Passenger Flow Statistics Based on YOLOV3,” In IOP Conference Series: Materials Science and Engineering, IOP Publishing, vol. 688, no. 4, p. 044025, 2019, doi: 10.1088/1757-899X/688/4/044025.

[9] I. Ahmed, M. Ahmad, A. Ahmad, & G. Jeon, “Top View Multiple People Tracking by Detection Using Deep Sort and YOLOV3 With Transfer Learning: Within 5G Infrastructure,” International Journal of Machine Learning and Cybernetics, vol. 12, 3053-3067, 2021, doi: 10.1007/s13042-020-01220-5.

[10] Y. W. Hsu, T. Y. Wang, and J. W. Perng, “Passenger Flow Counting in Buses Based on Deep Learning Using Surveillance Video,” Optik, vol. 202, p. 163675, 2020, doi: 10.1016/j.ijleo.2019.163675.