Estimating congestion in train cars by using BLE Signals

Eigo Taya*†, Yuji Kanamitsu*†, Koki Tachibana*,
Yugo Nakamura‡, Yuki Matsuda*†§, Hirohiko Suwa*†, Keiichi Yasumoto*†

* *Nara Institute of Science and Technology*, Nara, Japan,
Email: {taya.eigo.td3, kanamitsu.yuji.ky4, tachibana.koki.ti0, yukimat, h-suwa, yasumoto}@is.naist.jp.
† *RIKEN Center for Advanced Intelligence Project AIP*, Tokyo, Japan.
‡ *Kyushu University*, Fukuoka, Japan, Email: yugo@arakawa-lab.com.
§ *JST PRESTO*, Tokyo, Japan.

***Abstract*—In many of the world's major cities, commuter trains provide vital transportation support and thus play an essential role in our daily lives. Therefore, it has become necessary to estimate the degree of congestion in each train car, both to improve passenger comfort levels and, more recently, to prevent worsening the COVID-19 pandemic infection rate. However, it is difficult to estimate the degree of congestion within a train without violating passenger privacy. The same issues are true for busses, which is noteworthy because we have previously developed and evaluated a system that can estimate the degree of congestion within a bus while protecting passenger privacy by using Bluetooth Low Energy (BLE) signals. In this paper, we report on our efforts to extend that system to railway use, which were conducted on actual trains in cooperation with Kintetsu Railway Co., Ltd. During this trial, we collected BLE signals and used the data to estimate congestion levels in each car using an ML regression model. The results show that the mean absolute error (MAE) and the mean absolute percentage error (MAPE) could be estimated at accuracy levels of 5.56 and 0.27, respectively.**

***Index Terms*—BLE, Train, Public transportation, Congestion estimation, Machine learning**

I. Introduction

In many of the world's major cities, commuter trains provide vital transportation support for commuters and thus play an essential role in our daily lives. However, commuter rush hours and weather-related passenger surges can cause train congestion levels to rise past those where comfort is compromised. Because of this, it is well known that many passengers decide which train they will take based not only on its on-time operation and fare but also on its comfort level [1]. Accordingly, train operators often collect and provide information on congestion levels to prevent users from abandoning trains due to their lack of comfort. Another factor requiring consideration is the COVID-19 pandemic, which has been ongoing since November 2019 and was still causing severe economic, sanitary, and social problems for human society as of December 2021 [2]. However, even though people are striving to reduce contact with others to avoid spreading the infection, they still use trains to sustain daily activities such as commuting to work or shopping, so knowing train congestion levels in advance will permit them to schedule journeys and use trains with less risk of spreading the infection. This is a particularly urgent problem since trains are among the most widely used public transportation systems and tend to be the most congested at specific times. Our research group has previously proposed a congestion Bluetooth Low Energy (BLE) estimation system for local busses that protects passenger privacy [3]. In this system, BLE signals from the mobile devices of bus passengers are first detected, after which the Bluetooth Device Address (BD Address) included in the BLE signal, and the received signal strength indication (RSSI), are processed. In our method, passenger numbers are defined as the number of addresses satisfying the threshold value for the obtained BLE signal RSSI and a dataset that combines BLE data and bus-specific information. That study also discussed a method for estimating passenger numbers on busses using a machine learning (ML) model.

In this paper, we report on adapting the above system to trains by estimating the degree of congestion in each train car. To accomplish this, we first carried out a data collection experiment on operational trains in cooperation with Kintetsu Railway Co., during which the passenger numbers estimated from the collected data showed a threshold-based estimation result with a mean absolute error (MAE) and a mean absolute percentage error (MAPE) of 7.27 and of 0.27, respectively. In contrast, the ML-based estimation resulted in a MAE of 5.56 and a MAPE of 0.27. These results indicate that the estimation accuracy levels improved for the car, especially at high passenger levels.

The contributions of this research are as follows:

1) We constructed a device that collects BLE signals emitted by passenger mobile terminals and collected BLE data on actual trains in operation.
2) We constructed an ML-based model to estimate the congestion in train car and found that the model was sufficiently accurate for use in actual operation.

II. Related Work

There are currently numerous research types and a wide variety of different approaches to congestion estimates [4]–[12]. In this section, we introduce research on estimating passenger numbers and BLE signal-based congestion estimations, both of which are particularly relevant to our study.

A. Estimating passenger numbers

Camera-based systems have been proposed as the simplest way to estimate passenger numbers [13]–[15]. For example, Song *et al.*proposed a system that counts passenger numbers using surveillance camera video footage [13]. However, their camera-based estimation system records passenger faces and thus could potentially be used to track individuals. This is an important issue because it is not desirable for a public transportation system, such as a train system, to routinely collect information that may violate passenger privacy.

Several other studies have proposed using Wi-Fi signals to estimate congestion levels [16]–[19]. For example, Handte *et al.*proposed estimating passenger numbers on a Wi-Fi-equipped bus by counting the number of passenger mobile terminal media access control (MAC) addresses connected to an access point. In this system, the error for of the estimated number of passengers was 5.1, but the accuracy decreased when the bus became crowded. In addition, estimating congestion using Wi-Fi signals collects the MAC address for each passenger, which may also be used to identify individuals.

Methods aimed at detecting passengers using infrared sensors have also been proposed [20]–[24]. For example, to detect passenger numbers at an airport, Bauer *et al.*proposed combining an infrared sensor with a mat-type pressure sensor [20]. However, infrared sensors count people by detecting changes in signal intensity levels, so the estimation accuracy of that system tends to be poor when people loiter near the sensor or when a large number of people rapidly pass the sensor. Such situations frequently occur in the case of trains (e.g., leaning against the doorway, getting on and off a train at stations), which means that infrared sensors are not suitable for use under those conditions.

B. Estimating congestion using BLE signals

With the widespread use of smartphones, many studies have proposed congestion-based estimation schemes using wireless communication technologies such as BLE [25]–[29]. For example, Umeki *et al.*proposed a system to estimate the congestion levels in sightseeing spots by focusing on RSSI intensity levels, which vary significantly depending on the number of people present [27]. In that study, they installed a BLE signal broadcaster and a receiver at a sightseeing spot and estimated the degree of congestion by observing and evaluating the RSSI intensity distribution at three levels: ”low,” ”medium,” and ”high.”

Separately, Weppner *et al.*proposed a method for estimating crowd densities that works by aggregating the number of mobile signals detected from nearby user-carried BLE terminals moving in a monitored environment [28], [30]. Their method appropriately aggregates the data obtained from user devices, and estimates crowd density levels without the need to install new sensors. However, the above methods require a mechanism to encourage user involvement since the estimation accuracy depends primarily on the number of users participating in the sensing process.

C. Positioning of this research

As described above, there are numerous studies on estimating passenger numbers, but only a few both ensure privacy protection and can be practically used in transportation systems. In contrast, our research group has previously proposed a BLE signal-based congestion estimation system for busses that can more thoroughly protect passenger privacy [3]. Therefore, this study extends the application of that system to trains and reports on our attempt to estimate railway passenger numbers without compromising their privacy.

III. Data Collection System

A. System overview

A photograph of the sensing device used in our implemented system is shown in Figure 1. As can be seen, the system is implemented on a Raspberry Pi single-board computer and consists of a Bluetooth dongle, a global positioning system (GPS) module, and a network module to enable Long-Term Evolution (LTE) communication.

1) Bluetooth dongle: To protect passenger privacy while estimating congestion levels on public transport systems, our system detects the BLE signals generated by passenger mobile terminals. BLE is a power-saving communication standard among Bluetooth-standard short-range wireless communication devices. When active, each BLE device continuously broadcasts its availability data in an attempt to connect with other BLE devices. The dongle attached to the system collects BD addresses and RSSI signal strengths from the signals sent by the passenger BLE devices in order to identify the devices. However, since BD addresses change randomly at regular intervals, passenger privacy is protected.

2) GPS module: To estimate the congestion level, it is necessary to identify at which station sections the system has collected BLE signals. In theory, since public transportation systems operate according to established timetables, it is possible to map BLE signals to the station sections using the time information. However, as public transport is often delayed by external factors such as weather and accidents, it is challenging to identify BLE signals related to station sections using only time information. Therefore, the system uses a GPS module to obtain the train's position, so that it can map BLE signals to station sections even if the transport service is delayed.

3) Network module: Since many public transportation systems do not have network access points like Wi-Fi in every car, we use a network module that can be attached via a Universal Serial Bus (USB) connector to monitor the system remotely and collect real-time data.

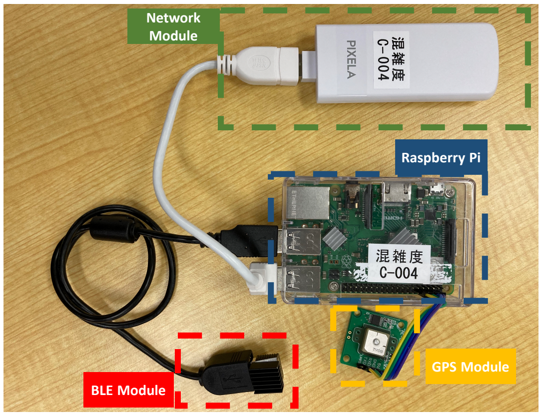


Fig. 1. System sensing device.

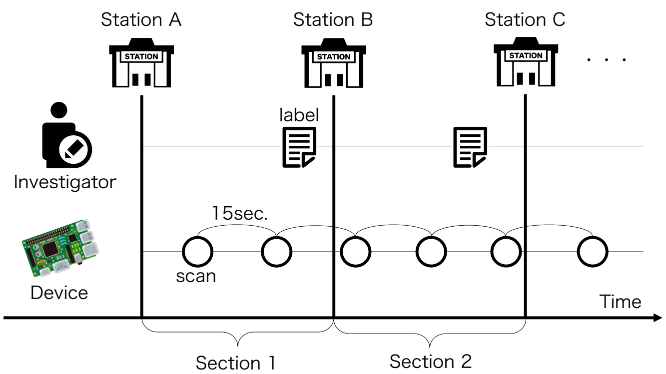


Fig. 2. Sensing process.

B. Assumption

The system collects the BLE signal, which requires the passenger's mobile device to turn on Bluetooth while moving. However, Bluetooth is becoming more widespread with the development of BLE devices such as wireless earphones and smartwatches and contact tracing technology. We, therefore, assume that the system can collect a sufficient number of BLE signals to estimate the number of passengers.

C. Sensing process

In our system, we estimate passenger numbers on the cars while the train moves from station to station(section) in order to provide an indicator of their congestion levels. An overview of the data sensing and labeling process is shown in Figure 2, where it can be seen that the system detects BLE signals from the surrounding area every 15 seconds and sends that data to the data server. The sent data includes timestamp, latitude, longitude, and BLE data collected from several terminals (Figure 3). In our experiment, an investigator carrying a sensing device boarded a car of an operating train and manually recorded passenger numbers while the device simultaneously collected sensing data (Table I).


```
{
    "timestamp": 1628662870,
    "lat": 34.693779,
    "lng": 135.782786,
    "ble": [
          {
             "addr": "72:12:4f:67:44:92",
             "rssi": -54
          },
          {
             "addr": "1a:b0:4f:e5:94:e4",
             "rssi": -87
          },
      ],
}
```


Fig. 3. Scan data excerpt.

TABLE I
LABELING DATA EXCERPT.

Time	Start	End	Type	Order	Num
15:13	Nara	Shinomiya	Local	3	33
15:15	Shinomiya	Yamatosaidaiji	Local	3	38
15:21	Yamatosaidaiji	Gakuenmae	Express	5	47
15:25	Gakuenmae	Ikoma	Express	5	49
15:32	Ikoma	Ishikiri	Express	5	38
15:36	Ishikiri	Fuse	Express	5	35

D. Feature extraction

Since the BLE data collected by the system in a single scan may contain signals from terminals outside the car, we aggregate all the data scanned in each station section and build two features. First, we assume that the same BD address can be detected multiple times within a station section when the BLE terminal and system sensing device are in the same car. Hence, the frequency F of signals generated by a single BD terminal is calculated as follows:

$$F(\%) = \frac{n_{detected}}{N_{scan}} \times 100 \tag{1}$$

where N_{scan} is the detection count in a station section and $n_{detected}$ is the number of times that the same BD address is detected. In addition, the average of the RSSI S_{mean} in station sections is calculated as follows.

$$S_{mean} = \frac{1}{n_{detected}} \sum_{i=1}^{n_{detected}} S(i) \tag{2}$$

where S(i) is the RSSI of the i-th detection in a station section. Next, we adopt the BD address count above a certain threshold as a station section feature, using the values obtained by the above equations 1, 2.

IV. Estimation and Evaluation

A. Experiment environment

We then collected BLE data on trains in operation using the system described in station section III. More specifically, we collected data from 381 station sections (with overlaps) in Nara Prefecture. There were two restrictions to our actual train experiment: (1) no power supply could be used for the sensing

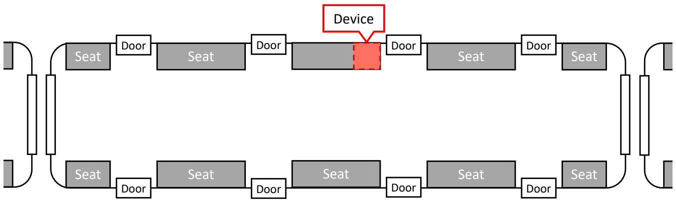


Fig. 4. Device and investigator location in the cars.

device, and (2) no sensing device could actually be installed in the trains. As explained above, our investigators carried a sensing device equipped with a mobile battery onto the car to collect data while simultaneously visually checking passenger numbers in order to obtain the actual passenger counts. This experimental environment ensured that the sensing device was located at the same position as the investigator who sat in the middle of the car where its interior could be observed easily, as shown in Figure 4.

B. Evaluation methodology

We estimate passenger numbers in each station section with the data collected in the IV-A section. Two methods are used for the estimation: one is based on the BLE signal threshold, and the other is based on an ML regressor.

1) Threshold estimation: As described in the III-D section, we calculate the average RSSI (S_{mean}), and occurrence frequency (F) for all BD addresses detected in each station section. Then, BD addresses above a particular value are set as valid addresses, and the total number of these addresses is set as the estimated value.

2) ML regressor estimation: Table II shows the ML features used. Here, we consider the average RSSI or occurrence frequency of a BD address to be essential indicators that indicate the address is present in the car. Therefore, the number of addresses at each incremental threshold is included in the BLE features. We also include the number of scans in the BLE features since the BD address count for each threshold is considered to affect the number of scans a device performs within a station section. Additionally, since information such as train operation times and types: e.g., local, limited express, or express, is considered vital for estimating congestion levels, such as commuting rush hour traffic and transit times, information specific to the train being observed is included in the train features.

Next, we constructed a train dataset (TD) with BLE features and a dataset containing all features (TD$^+$) and then used them to compare accuracy levels with and without train information. Each model estimates passenger numbers in one station section for these datasets using the Random Forest (RF) and LightGBM (LGBM) regression models. The estimation results were then evaluated by leave-one-out cross-validation, and the Optuna Framework was used to adjust the hyperparameters of each model [31].

TABLE II
FEATURE LIST.

Domain	Feature
BLE	BD address count [1]
	Number of scans
Train	Departure hour
	Operation type
	Station id
	Order of cars

[1] Threshold is set S_{means}:-30 to -100 or higher, F:0 to 100or higher

TABLE III
RESULT OF THRESHOLD ESTIMATION.

Method	MAE	MAPE
ALL BD Addresses	283.98	13.67
$S_{mean} \geq -50$	9.29	0.33
$S_{mean} \geq -55 \cap F \geq 40$	**7.27**	**0.27**

C. Result

1) Threshold Estimation: First, we explain the results when passenger numbers are defined by the total BD address count obtained in each station section. The graph in Figure 5 plots the total BD address count (Estimated value) and passenger numbers (True value) on the vertical and horizontal axes, respectively. Note that the blue line in the graph shows the ideal state when there is no difference between the true and estimated values. This graph shows that the total BD address count is significantly larger than the actual number of passengers. Accordingly, we then calculated the absolute error between each station section's true and estimated values and then averaged them over the total number of station sections to obtain the MAE and MAPE values, which were 283.98, and 13.67, respectively (Table III).

Next, we describe the estimation results obtained by setting thresholds for the average RSSI (S_{mean}) and occurrence frequency (F) of BD addresses. Table III shows the MAE and MAPE values when the appropriate thresholds are set, while Figure 6 shows a graph plotting the true and estimated values (number of valid BD addresses). Here, the threshold is set to an empirical value. We also show a graph plotting the true and estimated values for the number of valid BD addresses in Figure 6. From Table III, it can be seen that estimating passenger numbers with threshold results in smaller MAE and MAPE values, and thus more accurate estimations. Compared to Figure 5, it can be seen that the estimation is much closer to the true value. In contrast, the estimation accuracy is low in the station section where the true value is more than 50 people.

2) Estimating ML regressor estimation: Table IV shows cross-validated MAE and MAPE values. Here, it can be seen that the model with TD$^+$ achieves the best performance, with RF MAE and MAPE values of 5.85 and 0.28, respectively, and LGBM MAE and MAPE values of 5.56, of 0.27, respectively. The graph plotting the true and estimated values obtained by

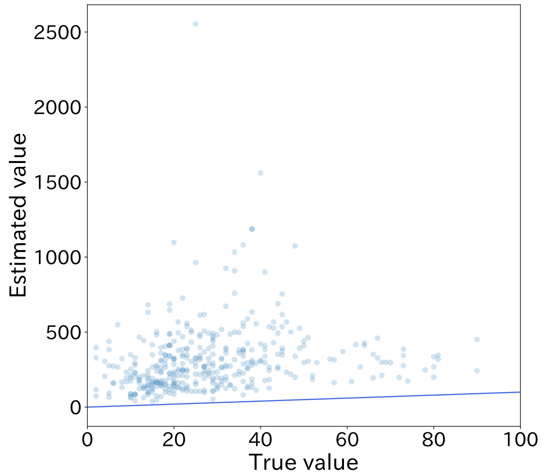
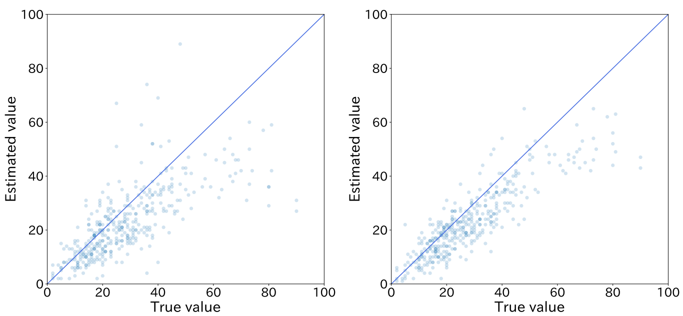


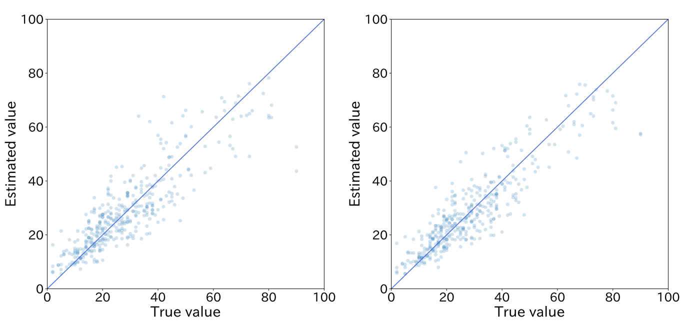
Fig. 5. Measured values (BD address counts) versus true values (passenger counts) obtained from raw data.



(a) $S_{mean} \geq -50$ (b) $S_{mean} \geq -55 \cap F \geq 40$

Fig. 6. Values estimated by LGBM (BD address count) versus true values (passenger count).

LGBM using TD$^+$ shown in Figure 7 references the feature importance values in Figure 8. Here, it can be seen that the occurrence frequency feature (add_fre_x) makes a particularly important contribution to the estimation.



(a) TD (b) TD$^+$

Fig. 7. Values estimated by LGBM (number of BD addresses) versus true values (number of passengers)

TABLE IV
PERFORMANCE OF EACH MODEL FOR EACH DATASET

	TD		TD$^+$	
Model	MAE	MAPE	MAE	MAPE
RF	5.96	0.28	5.85	0.28
LGBM	6.07	0.29	**5.56**	**0.27**

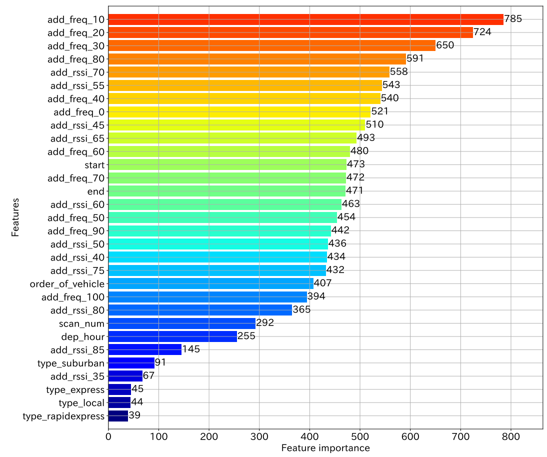


Fig. 8. Feature importance value.

V. DISCUSSION

A. Estimation threshold

Figure 5 shows that the total BD address count detected in the station section is much larger than the true value. This may be due to the detection of BLE devices outside the car and the fact that the BD address changes over time. This indicates that by setting appropriate thresholds for the average RSSI and the occurrence frequency of BLE signals, we can significantly improve estimation accuracy levels, as shown in Figure 6. On the other hand, as shown in Figure 6 (b), the estimated value is much lower than the true value for the interval where the true value is more than 50 people. This is believed to be due to the characteristics of BLE signals, which are attenuated significantly when they encounter obstructions. Therefore, in a crowded environment, the signals are attenuated by human bodies, and it is assumed that the signals from some terminals cannot be detected. In particular, since the devices were situated near the center of the car in our experiment, they experienced difficulties detecting the terminals of passengers at the car ends.

B. ML regressor estimation

In this study, we improved the accuracy of the model by including train-specific features in addition to the BLE-related features used to estimate passenger numbers per car. The cross-validation results show that the best-performing model has a MAE of 5.56 and a MAPE of 0.27. Comparing the results with the threshold estimation in Figure 6, it can be seen that the accuracy of our model has been improved, especially for the station section where the true value is greater than 50 individuals. The feature importance values in the model with the best accuracy, shown in Figure 8, indicate that there are high values related to relatively small frequency values of 10-30%. When detecting BLE terminal signals, we found that the device could detect the signal from a nearby terminal with a high threshold while more distant terminal signals could only be detected with low thresholds. Therefore, the small frequency value is considered to have contributed significantly to signal determination for distant terminals.

VI. Conclusion

This paper reports on the collection of BLE data from operating trains as part of efforts to estimate passenger numbers while protecting their privacy. To deal with the varying number of devices in the train and passengers carrying multiple BLE devices, we applied an estimation method that sets thresholds for BLE data and uses an ML model to estimate passenger numbers. The results show that passenger numbers can be estimated with an MAE accuracy of about 5.56. Our future work will include the collection of additional data. Our experimental environment only includes data from the hours of 13:00 to 19:00, which do not include the morning rush hour. Accordingly, it will be necessary to collect data under various conditions in order to achieve a high level of accuracy before our system can be put into practical use. We also made careful efforts to protect the privacy of passengers by using BLE addresses, which are better than video images for producing congestion estimations. However, since a detailed analysis of the addresses in the BLE data could be used to identify individuals, it will be necessary to design a more privacy-aware system that discards the addresses as soon as the feature construction process has been completed.

Acknowledgment

This study was funded in part by the Japan Science and Technology Agency (JST)'s Strategic Basic Research Program, Precursory Research for Embryonic Science and Technology (PRESTO) grant number JPMJPR2039. We would like to thank Kintetsu Railway Co., Ltd. for their cooperation in the data collection process of this study.

References

[1] D. van Lierop, M. G. Badami, and A. M. El-Geneidy, "What influences satisfaction and loyalty in public transport? a review of the literature," *Transport Reviews*, vol. 38, no. 1, pp. 52–72, 2018.

[2] W. H. Organization, "Coronavirus disease (covid-19) pandemic." https://www.who.int/emergencies/diseases/novel-coronavirus-2019, accessed: 2021-12-01.

[3] Y. Kanamitsu, E. Taya, K. Tachibana, Y. Nakamura, Y. Matsuda, H. Suwa, and K. Yasumoto, "Estimating congestion in a fixed-route bus by using ble signals," *Sensors*, vol. 22, no. 3, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/3/881

[4] L. Del Pizzo, P. Foggia, A. Greco, G. Percannella, and M. Vento, "A versatile and effective method for counting people on either rgb or depth overhead cameras," in *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE, 2015, pp. 1–6.

[5] S. D. Pore and B. Momin, "Bidirectional people counting system in video surveillance," in *2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. IEEE, 2016, pp. 724–727.

[6] P. G. Kannan, S. P. Venkatagiri, M. C. Chan, A. L. Ananda, and L.-S. Peh, "Low cost crowd counting using audio tones," in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, 2012, pp. 155–168.

[7] A. B. Chan, Z.-S. J. Liang, and N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008, pp. 1–7.

[8] T. Nishimura, T. Higuchi, H. Yamaguchi, and T. Higashino, "Detecting smoothness of pedestrian flows by participatory sensing with mobile phones," in *Proceedings of the 2014 ACM International Symposium on Wearable Computers*, ser. ISWC '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 15–18. [Online]. Available: https://doi.org/10.1145/2634317.2642869

[9] Y. Wada, T. Higuchi, H. Yamaguchi, and T. Higashino, "Accurate positioning of mobile phones in a crowd using laser range scanners," in *2013 IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2013, pp. 430–435.

[10] M. Fu, P. Xu, X. Li, Q. Liu, M. Ye, and C. Zhu, "Fast crowd density estimation with convolutional neural networks," *Engineering Applications of Artificial Intelligence*, vol. 43, pp. 81–88, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0952197615000871

[11] G. Li, C.-J. Chen, W. Peng, and C. Yi, "Estimating crowd flow and crowd density from cellular data for mass rapid transit," 2017.

[12] K. Nagao, D. Yanagisawa, and K. Nishinari, "Estimation of crowd density applying wavelet transform and machine learning," *Physica A: Statistical Mechanics and its Applications*, vol. 510, no. C, pp. 145–163, 2018. [Online]. Available: https://EconPapers.repec.org/RePEc:eee:phsmap:v:510:y:2018:i:c:p:145-163

[13] H. Song, X. Liu, X. Zhang, and J. Hu, "Real-time monitoring for crowd counting using video surveillance and gis," in *2012 2nd International Conference on Remote Sensing, Environment and Transportation Engineering*. IEEE, 2012, pp. 1–4.

[14] H. Takayanagi, T. Sano, and H. Watanabe, "A study on the pedestrian occupied territory in the crossing flow: The analysis with pedestrian territory model," *Journal of Architecture and Planning (Transactions of AIJ)*, vol. 66, no. 549, pp. 185–191, 2001.

[15] H. Su, H. Yang, and S. Zheng, "The large-scale crowd density estimation based on effective region feature extraction method," in *Asian Conference on Computer Vision*. Springer, 2010, pp. 302–313.

[16] M. Handte, S. Foell, S. Wagner, G. Kortuem, and P. J. Marrón, "An internet-of-things enabled connected navigation system for urban bus riders," *IEEE internet of things journal*, vol. 3, no. 5, pp. 735–744, 2016.

[17] S. Jackson, A. Lesani, and L. Miranda-Moreno, "Toward a wifi system for traffic monitoring in different transportation facilities," in *Transportation Research Board 93rd Annual Meeting*, 2014.

[18] A. Hidayat, S. Terabe, and H. Yaginuma, "Wifi scanner technologies for obtaining travel data about circulator bus passengers: case study in obuse, nagano prefecture, japan," *Transportation Research Record*, vol. 2672, no. 45, pp. 45–54, 2018.

[19] ——, "Estimating bus passenger volume based on a wifi scanner survey," *Transportation Research Interdisciplinary Perspectives*, vol. 6, p. 100142, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2590198220300531

[20] D. Bauer, M. Ray, and S. Seer, "Simple sensors used for measuring service times and counting pedestrians: strengths and weaknesses," *Transportation research record*, vol. 2214, no. 1, pp. 77–84, 2011.

[21] H.-b. Qian and H. Han, "The applications and methods of pedestrian automated detection," in *2010 International Conference on Measuring Technology and Mechatronics Automation*, vol. 3, 2010, pp. 806–809.

[22] H. Yang, K. Ozbay, and A. Bartin, "Investigating the performance of automatic counting sensors for pedestrian traffic data collection," in *Proceedings of the 12th World Conference on Transport Research*, vol. 1115, 2010.

[23] R. Greene-Roesel, M. C. Diogenes, D. R. Ragland, and L. A. Lindau, "Effectiveness of a commercially available automated pedestrian counting device in urban environments: Comparison with manual counts," *UC Berkeley: Safe Transportation Research and Education Center*, 2008.

[24] U. Scheunert, H. Cramer, B. Fardi, and G. Wanielik, "Multi sensor based tracking of pedestrians: a survey of suitable movement models," in *IEEE Intelligent Vehicles Symposium, 2004*, 2004, pp. 774–778.

[25] L. Schauer, M. Werner, and P. Marcus, "Estimating crowd densities and pedestrian flows using wi-fi and bluetooth," in *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, ser. MOBIQUITOUS '14. Brussels, BEL: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2014, p. 171–177. [Online]. Available: https://doi.org/10.4108/icst.mobiquitous.2014.257870

[26] A. Lesani and L. Miranda-Moreno, "Development and testing of a real-time wifi-bluetooth system for pedestrian network monitoring, classification, and data extrapolation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 4, pp. 1484–1496, 2018.

[27] K. Umeki, Y. Nakamura, M. Fujimoto, Y. Arakawa, and K. Yasumoto, "Real-time congestion estimation in sightseeing spots with ble devices," in *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 2018, pp. 430–432.

[28] J. Weppner and P. Lukowicz, "Bluetooth based collaborative crowd density estimation with mobile phones," in *2013 IEEE international conference on pervasive computing and communications (PerCom)*. IEEE, 2013, pp. 193–200.

[29] V. Kostakos, T. Camacho, and C. Mantero, "Towards proximity-based passenger sensing on public transport buses," *Personal and ubiquitous computing*, vol. 17, no. 8, pp. 1807–1816, 2013.

[30] J. Weppner and P. Lukowicz, "Collaborative crowd density estimation with mobile phones," *Proc. of ACM PhoneSense*, 2011.

[31] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, ser. KDD '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 2623–2631.