Origin-Destination Estimation System Utilizing Thermosensor Network

Kazuhito Hori
Ritsumeikan University
Graduate School of Information Science and Engineering
Osaka, Japan

Hideaki Miyaji
Ritsumeikan University
College of Information Science and Engineering
Osaka, Japan

Hiroshi Yamamoto
Ritsumeikan University
College of Information Science and Engineering
Osaka, Japan

***Abstract*—Public transportation, utilizing buses, plays a crucial role in urban mobility. However, in recent years, the challenge of maintaining routes has become severe, particularly in low-population areas where many bus lines struggle to maintain profitability. In this situation, the Origin-Destination (OD) estimation that tracks the bus stop where each passenger boarded and alighted is essential not only for optimizing routes and enhancing service quality but also for justifying the continued operation of these vital services. However, existing OD estimation methods primarily rely on RGB cameras, which pose significant privacy concerns due to the direct recording of the faces and actions of passengers. On the other hand, while the use of LiDAR for observing the passengers inside the bus vehicle offers superior privacy protection, its high power consumption and prohibitive installation costs present substantial barriers to widespread deployment in power-limited environments. To address these challenges, this study proposes a novel system for monitoring passenger movements inside bus vehicles. This system utilizes low-power thermosensors placed at multiple locations throughout the vehicle and analyzes the acquired thermal images for tracking movement of each passenger insidel the vehicle by checking the time periods when the same passenger passed in front of different thermosensors. Furthermore, the system employs an autoencoder, which is a machine learning method, to evaluate the similarity of thermal images captured by different thermosensor. An autoencoder model is constructed for each passenger based on thermal images acquired by a thermosensor installed at the boarding entrance. By inputting thermal images collected by thermosensors placed at other locations into each model and deriving similarity metrics, the system identifies which passenger corresponds to the collected image. A demonstration experiment is conducted in an assumed area inside the bus vehicle to confirm the efficary of the proposed system. By analyzing thermal images using an autoencoder, we demonstrate a capability to identify the same passenger during boarding and alighting, with a precision of 0.62 and a recall of 0.56.**

***Index Terms*—Thermal Images, Passenger Re-identification, Autoencoder, OD Estimation, Privacy Protection.**

I. Introduction

The Origin–Destination (OD) estimation method quantifies the actual flow of passengers, things, and information from an origin to a destination. The method serves as a crucial tool for assessing the utilization of public transportation systems, including buses. The OD data which records the boarding and alighting stops for each passenger is an indispensable information for understanding the operational status of public transportation, optimizing routes, and improving services.

Existing technologies for the OD estimation often involve installing RGB cameras inside bus vehicles to analyze collected videos, thereby estimating passenger density and movement. However, this approach raises serious privacy concerns because the faces and actions of users are directly recorded. Instances of legal action seeking damages against camera installers have indeed been reported [1].

Furthermore, existing studies address the improvement of real-time property of the analysis or the reduction of power comsumption for RGB cameras within bus vehicles [2][3]. On the other hand, the observation systems utilizing LiDAR (Light Detection and Ranging) for observing condisions inside the bus vehicles are proposed [4]. The LiDAR measures the three-dimensional state of individuals and objects within the observation range by projecting infrared lasers in multiple directions. This approach helps to protect privacy by avoiding the direct identification of individuals. However, the LiDAR-based system suffers from high power consumption and difficulties in maintaining a long-term power supply within bus vehicles.

Therefore, this study proposes of a novel system designed to estimate the OD data for each passenger within buses. In the proposed system, thermosensors capable of obtaining two-dimensional thermal images are installed at multiple locations inside the bus. By analyzing and comparing the temperature images obtained from these sensors, the system aims to precisely determine when the same individual passes through various points. Furthermore, to efficiently process temperature images from an unspecified number of users, the system leverages a trained autoencoder model. This model is trained only on the thermal images for each passenger so as to correctly reconstruct the input thermal images. Consequently, when a thermal image is input, the model belonging to the corresponding passenger yields a low reconstruction error, whereas models belonging to others yield a significantly higher error, allowing the system to uniquely identify the individual.

II. Related Work

A. Research on OD Estimation of Bus Passengers Using RGB Cameras

Yamashita et al. propose an OD estimation method based on passenger tracking using RGB cameras inside bus vehicles

[5]. This method involves applying a YOLOR object detection model to images captured by multiple cameras installed at boarding and alighting entrances of the bus vehicle to detect location of each passenger. Subsequently, an MPNTrack tracking model is utilized to estimate trajectories of passengers, which are then used to match boarding and alighting passengers for the OD estimation.

Additionally, Hadisurya et al. propose an estimation method of passenger trajectories by analyzing video captured by cameras installed inside bus vehicles [6]. Their method applies a YOLOv3 object detection model to detect multiple individuals within the vehicle and estimate moving trajectories of passengers, which is utilized for congestion problems in public transportation in large cities.

However, a significant drawback of these camera-based methods is the potential for privacy issues of bus passengers, because the cameras clearly record appeerences of passengers.

B. Research on Passenger Tracking Using LiDAR

Ukyo et al. propose a method for detecting individuals walking on roads and estimating their trajectories by analyzing point-cloud data obtained by multiple 3D LiDARs [7]. This method achieves robust passenger tracking by applying a Kalman filter to on wide-ranging 3D point-cloud data obtained from multiple LiDAR. The Kalman filter continuously estimates and updates state variables such as the detected postion and height of each passenger across successive frames. This approach is anticipated to be an effective method for measuring passenger flow in high-traffic areas such as public and commercial facilities. This process plays an indispensable role in robustly tracking the same individual despite irregular motion, noise, and occlusions. However, a limitation of this method is the lack of consideration for maintaining long-term power supply for LiDAR devices, in environments without external power sources.

C. Research on Autoencoder-based machine learning model

Autoencoder is a type of unsupervised machine learning method based on neural networks [8][9]. The machine learning model of the autoencoder is trained so that the output data is as close as possible to the input data. Here, the architecture of the autoencoder is designed that the dimensionality of the input data is first reduced to the latent vector and then restored to reconstruct the output data with the same dimensionality as the input data. During training, the weights of the neural network are adjusted to minimize the discrepancy between the input and the reconstructed output data. Through this process, the model is trained to extract the feature vector which is the most essential information required for reconstruction of the input data. The reconstruction error between the input and output data can be used to specify a category of the input data. A low reconstruction error indicates that the new input data belongs to the same category as the training data.

By utilizing the autoencoder-based machine learning model, Tobari et al. propose a novel network anomaly detection method to address security challenges in cloud networks [10]. In this system, the autoencoder learns normal network traffic data for training. Furthermore, this method primarily utilizes the reconstruction error as a criterion for anomaly detection. The reconstruction error is defined as the quantitative metric representing the discrepancy between the input data and the reconstructed output data of the autoencoder. As the reconstruction error, the Mean Squared Error (MSE), the Structural Similarity Index Measure (SSIM), or a composite score derived from them are utilized. Specifically, if the reconstruction error of the input data exceeds a predefined threshold, the data is classified as anomalous.

D. Purpose of Our Proposed System

Existing for estimating passengers' trajectories and OD data in public transportation mainly rely on RGB cameras or LiDAR sensors. However, both approaches present significant challenges for practical and long-term deployment in real-world environments. Specifically, camera-based methods inherently have privacy risks of passengers. Conversely, while LiDAR-based methods offer reduced privacy concerns, the high power consumption makes them unsuitable for long-term operation in bus vehicles where a stable and external power source is difficult to ensure. Therefore, there is a critical need for a system that can accurately observe the passengers while ensuring privacy and supporting energy efficient, long-term operation.

To address these limitations, we propose a new observation system that estimates the passage of the same passengers through various points within a bus vehicle, including boarding and alighting entrances. This is achieved by utilizing thermosensors that consume less power than LiDAR, and by analyzing the thermal images measured by these sensors. In the proposed system, thermal images are analyzed using an autoencoder to identify the images corresponding with the same passenger. Specifically, an autoencoder model is individually constructed for each passenger based on thermal images acquired by a thermosensor installed at the boarding entrance. Subsequently, thermal images collected by thermosensors placed at other locations are input into each of these models. Based on the derived similarity metrics, the system identifies the model corresponding to the passenger on the collected image. To ensure long-term operation of the system in bus vehicles where providing a stable power supply is difficult, the thermosensors are activated only when the passengers are likely to move. For instance, the activation occurs when the bus stops at a bus stop.

III. Proposed Thermal Signature-based In-Vehicle Passenger Observation System

A. System Overview

The overall architecture of the proposed system is illustrated in Fig. 1. As shown in this figure, the system comprises sensor nodes installed inside the bus vehicle for passenger detection and a server responsible for data collection, analysis, and visualization.

提案システムの全体像

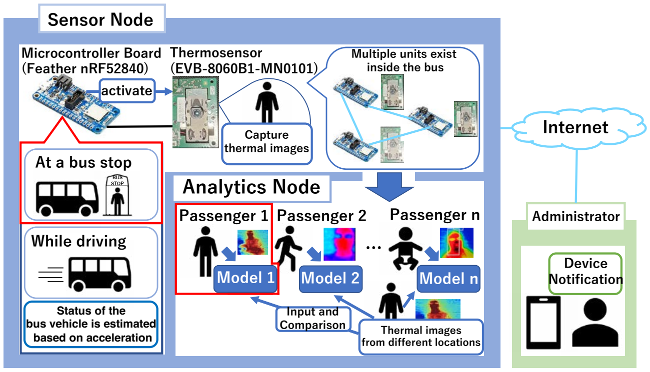

Fig. 1. Overall of Proposed Observation System.

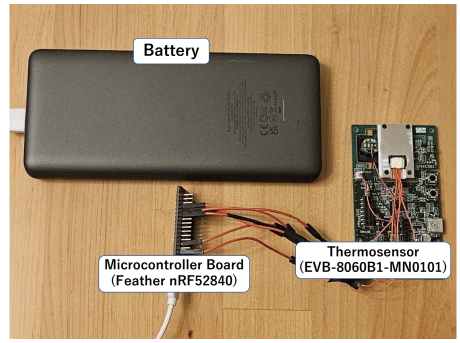

Fig. 2. Device Configuration of Sensor Node.

To reduce the power consumption, the sensor nodes activate their thermosensors only under limited conditions such as when the bus vehicle stops at the bus stop, to obtain the temperature images of the target objects at various points within the vehicle. Specifically, a machine learning model corresponding to each passenger is created based on images acquired at the boarding entrance. The system utilizes similarity metrics, which are quantitative measures representing the degree of difference between the input image and the reconstructed output image, for passenger identification. Subsequently, images captured at other locations are input into each model, and the corresponding passenger is identified by comparing multiple similarity metrics. Finally, these estimation results are transmitted to administrator terminals or other devices to estimate OD data for each passenger.

B. *Configuration of sensor node.*

Fig. 2 illustrates the hardware configuration of the sensor nodes within the proposed system. Each sensor node observes thermal images inside the bus vehicle and performs preprocessing on these observed thermal images. The preprocessing involves background removal through binarization to reduce the load of machine learning processing. To perform the preprocessing, the sensor node consists of a microcontroller board (Feather nRF52840 Sense) equipped with a built-in motion sensor and supporting BLE communication. This board is connected to a thermosensor (Mitsubishi Electric EVB-8060B1-MN0101), and multiple sensor nodes are strategically placed at various points throughout the bus vehicle.

Furthermore, to reduce the overall power consumption, the sensor nodes activate their thermosensors only when detecting the stop of the vehicle by using the motion sensor embedded with the board. The main specifications of the thermosensors used in this study are listed in Tab. I. In the proposed system, preprocessing is applied to the thermal image captured by the sensor node during boarding of each passenger on the bus vehicle. As a result of the preprocessing, the pixels corresponding to the foreground (passenger) are extracted and transmitted via BLE communication to a small computer (Raspberry Pi 4) for analysis by the autoencoder-based machine learning model.

TABLE I
MAIN SPECIFICATIONS OF THERMOSENSOR (MITSUBISHI ELECTRIC EVB-8060B1-MN0101).

Number of pixels	80×60 pixels°
Frame rate	4fps / 8fps
Supply voltage	3.3V
SPI communication frequency	2.1MHz
Measurement range of temperature	from -5°C to 60°C

C. *Thermal Image Acquisition and Preprocessing*

In the proposed system, the sensor node judges the vehicle to be stationary when both the measured acceleration in the direction of travel and that in the vertical direction fall below a predetermined threshold, and then activates the thermosensor to measure the thermal images. The thermosensor captures high-resolution thermal images, with each frame containing 4,800 pixels.

To achieve real-time and highly accurate analysis, the system applies Otsu's binarization as a preprocessing step. The binarization is a process that sets a threshold for pixel values of the thermal image and pixels values above the threshold are set to 1 and all other pixels are set to 0. The Otsu's binarization automatically determines an optimal threshold for dividing the pixel values in an image into two groups (classes) so that the intra-class variance is minimized and the inter-class variance is maximized. The overview of the binarization is shown in Fig. 3.

Consequently, this approach which applies the binarized image as a mask for the thermal image enables robust foreground extraction that is resilient to noise. An overview of the process for extracting pixels corresponding to a passenger from the thermal image is shown in Fig. 4. To further improve estimation accuracy, after the foreground extraction of the thermal images, the passenger region is cropped to tightly enclose the foreground object. This cropped region is then resized to the required input dimension while preserving its aspect ratio.

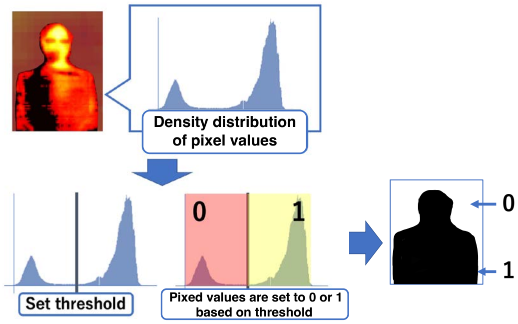

Fig. 3. Overview of Otsu's Binarization.

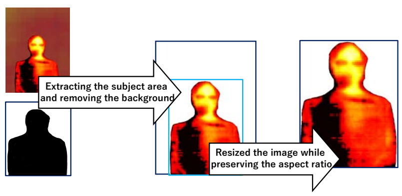

Fig. 4. Foreground Extraction from Temperature Images.

D. Alignment of Dimensions for Data Augmentation

The thermosensor utilized in this system generates thermal images with a resolution of 80×60 pixels. Conversely, the convolutional neural network (CNN) architecture of the encoder is designed with the requirement that the width and height of the input images are divisible by a power of 2. Specifically, since four consecutive 2×2 pooling layers are applied in the proposed system, the size of the input image should be a multiple of 16. However, since the resolution of the thermosensor is 80×60 pixels, a preprocessing step is introduced to extend the input images to 80×64 pixels. This is achieved by appending four columns of pixels to the right edge of the original thermal image with 80×60 pixel and filling all of these new pixels with zero, thereby satisfying the requirement of being a multiple of 16 for the input layer of encoder. The basic concept of the adjustment of the input image is shown in Fig. 5.

E. Construction of Autoencoder-based Model Corresponding with Each Passengerr

The system utilizes an autoencoder for constructing a machine-learning model corresponding with each passenger. The model is trained using thermal images of individual passengers as training data and then determines whether the passenger appearing in an input image corresponds to the passenger represented in the training data. In this study, the autoencoder is a convolutional autoencoder (CAE) designed with multiple convolutional layers, pooling layers, batch normalization layers, and Leaky ReLU activation functions.

The encoder section compresses the input thermal image of 80×64 pixels into a latent space of 5×4 pixels through a sequence of four convolutional and pooling layers. The number of filters in these convolutional layers is progressively

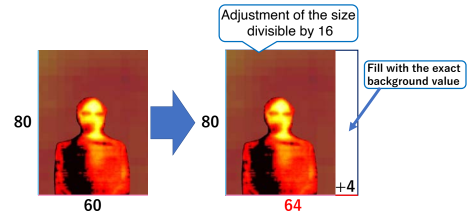

Fig. 5. Overview of Resolution Adjustment for Input Image.

reduced from 256 to 128, 64, and 32. Leaky ReLU serves as the activation function, with the exception of the output layer, which uses a Sigmoid function. The decoder section is composed of four deconvolutional layers and upsampling layers to reconstruct the image from the latent space. The output is then cropped to 80×60 pixels to match the dimension of the original image with the reconstructed image. The detailed architecture of the autoencoder is presented in Tab. II, and the overall processing workflow is shown in Fig. 6.

The model is trained using a composite loss function that combines Mean Squared Error (MSE) and the Structural Similarity Index (SSIM). MSE is a simple metric that squares and averages the pixel-wise difference, making it highly sensitive to image intensity differences. In contrast, SSIM captures perceptual quality differences by evaluating the similarity of image luminance, contrast, and structure based on the human visual system. This loss function is specifically designed to prioritize the reconstruction of a shape of passengers by assigning a higher weight to foreground pixels and a lower weight to background pixels. The cost function for this model is defined as

$$\mathcal{L}_{\text{Combined}} = \alpha \cdot \mathcal{L}_{\text{WeightedMSE}} + \beta \cdot \mathcal{L}_{\text{SSIM}} \tag{1}$$

with

$$\mathcal{L}_{\text{WeightedMSE}} = \sum_{i,j} \left(W_{i,j} \cdot (y_{i,j} - \hat{y}_{i,j})^2 \right) \tag{2}$$

$$\mathcal{L}_{\text{SSIM}} = 1.0 - \text{SSIM}(y, \hat{y}) \tag{3}$$

where $\mathcal{L}_{\text{Combined}}$ is the composite loss function, $\mathcal{L}_{\text{WeightedMSE}}$ is the weighted Mean Squared Error, and $\mathcal{L}_{\text{SSIM}}$ is the Structural Similarity Index loss. Furthermore, y and $\hat{y}$ are the true and reconstructed pixel values, respectively, and W is the weight matrix assigned to the pixels, defined as

$$W_{i,j} = \begin{cases} 1.0 & \text{if } y_{i,j} \neq 0 \quad \text{(foreground pixel)} \\ 0.07 & \text{if } y_{i,j} = 0 \quad \text{(background pixel)} \end{cases} \tag{4}$$

and α and β are the weights controlling the balance between the two loss components.

F. Passenger Identification Based on Reconstruction Error

First, time-series data of thermal images collected when a specific passenger passes through the boarding entrance are used as training data to train a model corresponding to that

TABLE II
ARCHITECTURE OF PROPOSED AUTOENCODER-BASED MODEL..

Layer Type	Layer Name	Output Shape
Input Layer	-	80×64×1
Encoder		
Conv2D + BN + LeakyReLU	enc_conv1	80×64×256
MaxPooling2D	enc_pool1	40×32×256
Conv2D + BN + LeakyReLU	enc_conv2	40×32×128
MaxPooling2D	enc_pool2	20×16×128
Conv2D + BN + LeakyReLU	enc_conv3	20×16×64
MaxPooling2D	enc_pool3	10×8×64
Conv2D + BN + LeakyReLU	enc_conv4	10×8×32
MaxPooling2D	enc_pool4	5×4×32
Bottleneck (Latent Space)		
Conv2D + BN + LeakyReLU	encoded_layer	5×4×8
Decoder		
Conv2D + BN + LeakyReLU	dec_deconv4	10×8×32
Conv2D + BN + LeakyReLU	dec_deconv3	20×16×64
Conv2D + BN + LeakyReLU	dec_deconv2	40×32×128
Conv2D + BN + LeakyReLU	dec_deconv1	80×64×256
Conv2D	output_final	80×64×1

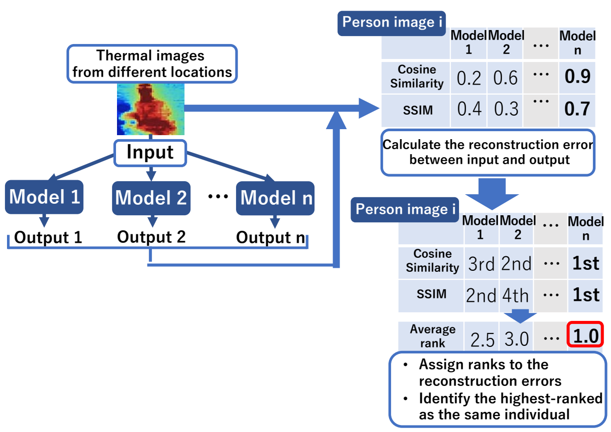

Fig. 6. Passenger Identification Utilizing Autoencoder-based Model.

passenger. Subsequently, thermosensors installed at various locations, such as the alighting exit, acquire thermal images. These images are then input into the models corresponding to each passenger, and the reconstruction error is calculated for each model. To calculate the reconstruction error between the input image and the output reconstructed image, we adopt a composite evaluation method rather than relying on a single metric. This method combines two representative metrics, MSE and SSIM, as the same as the loss function. The MSE focuses on quantifying the magnitude of the physical error by assessing the strict disparity between corresponding pixel values. In contrast, the SSIM provides a more visually relevant assessment by considering the structural, luminance, and contrast components of the images.

After deriving the metrics, the system decides a "successful identification" as a case where a model achieved the highest average rank of in these two metrics. The passenger corresponding to this model is then identified as the same passenger

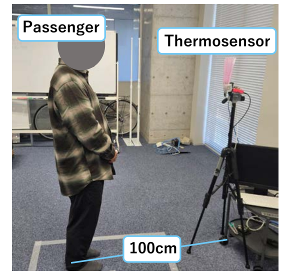

Fig. 7. Positional Relation of Thermosensor and Passenger.

as the one in the input thermal image.

IV. EXPERIMENTAL EVALUATION

A. Experimental Setup

To clarify the accuracy of estimating the same passenger by the proposed system, we conduct the experimental evaluation. In this evaluation, we focus on the identification accuracy of the system. The performance assessment is conducted based on a suite of metrics derived from a confusion matrix, namely Precision and Recall for each passenger.

The experimental environment is established in an indoor environment at the Ritsumeikan University, Osaka Ibaraki Campus, simulating the interior of a bus vehicle. For six participants of the experiment, thermal images are captured from a distance of 100 cm. The positional relations of the thermosensor and the participant is illustrated in Fig. 7.

First, a machine learning model is trained for each participant using 30 thermal images, assuming the "boarding" phase. Subsequently, 30 new thermal images for each participant are captured and input into their respective trained models. To evaluate the effectiveness of our proposed method, the reconstruction error between the input images and the output of each model is measured using not only the combined evaluation metric defined in Section III-F but also each metric (i.e., MSE, SSIM).

B. Experimental Result

For the newly obtained dataset including 30 thermal images for each participants, the performance of the proposed system is evaluated. Tables III, IV, and V show the confusion matrices for the three metrics (i.e., MSE, SSIM, and the Composite Metric), detailing the identification results. Table VI and VII show the Precision and Recall by the single metrics and the Composite Metric.

As shown in Tables III and VI, when using the MSE, the system completely makes false identification for both Participants 2 and 6. On the other hand, the use of SSIM successfully mitigates the failure of Participant 2 as shown in Tab. IV because the SSIM can capture structural similarity that the MSE fails.

Nevertheless, the use of SSIM alone exhibits limitations that the precision for Participant 4 decreases to 0.50 as shown in Tab. VI. This suggests that the SSIM can make a false misrecognition when structural features are highly similar. As

TABLE III
CONFUSION MATRIX (MSE).

Evaluation Data	Training Data					
	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Participant 6
Participant 1	4	0	0	0	**26**	0
Participant 2	0	0	0	0	**30**	0
Participant 3	0	0	5	0	**25**	0
Participant 4	0	0	0	**30**	0	0
Participant 5	0	0	0	0	**30**	0
Participant 6	0	**30**	0	0	0	0

TABLE IV
CONFUSION MATRIX (SSIM).

Evaluation Data	Training Data					
	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Participant 6
Participant 1	4	0	0	0	**26**	0
Participant 2	0	**28**	0	0	2	0
Participant 3	0	0	**29**	0	1	0
Participant 4	0	0	0	**30**	0	0
Participant 5	1	0	3	0	**26**	0
Participant 6	0	1	0	**29**	0	0

TABLE V
CONFUSION MATRIX (COMPOSITE METRIC).

Evaluation Data	Training Data					
	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Participant 6
Participant 1	4	0	0	0	**26**	0
Participant 2	0	**28**	0	0	2	0
Participant 3	0	0	**24**	0	6	0
Participant 4	0	0	0	**30**	0	0
Participant 5	1	0	0	0	**29**	0
Participant 6	0	**30**	0	0	0	0

TABLE VI
PRECISION AND RECALL (SINGLE METRIC).

Participant ID	MSE		SSIM	
	Precision	Recall	Precision	Recall
1	**0.80**	**0.13**	**0.80**	**0.13**
2	**0.00**	**0.00**	**0.96**	**0.93**
3	**1.00**	**0.16**	**0.90**	**0.96**
4	**1.00**	**1.00**	**0.50**	**1.00**
5	**0.27**	**1.00**	**0.47**	**0.86**
6	**N/A**	**0.00**	**N/A**	**0.00**

TABLE VII
PRECISION AND RECALL (COMPOSITE METRIC).

Participant ID	Precision	Recall
1	**0.80**	**0.13**
2	**0.48**	**0.93**
3	**1.00**	**0.80**
4	**1.00**	**1.00**
5	**0.46**	**0.96**
6	**N/A**	**0.00**

a result, the use of single metric lacks the robustness for generalized identification because each metric is optimized for specific factors.

In contrast, the Composite Metric achieves a more robust result by integrating MSE and SSIM to mutually consider both of intensity and structure features as shown in Tabs. V and VII. Especially, the use of the composite metric markedly increases the recall for Participant 2 to 0.93 while achieving a high precision of 1.00 for Participant 4.

V. CONCLUSION

In this study, we proposed a system for tracking the locations of each passenger inside a bus vehicle. This system involves constructing a sensor network that collects thermal images on various locations inside the vehicle using low-power microcontrollers and thermosensors, and subsequently analyzing the obtained thermal images using machine learning for identifying the images corresponding with the same passenger. Through demonstration experiments, we confirmed that an autoencoder can be effectively utilized to identify thermal images of identify during boarding and alighting, enabling the tracking the same passenger inside the vehicle.

To address challenges such as model overfitting and insufficient generalization performance, we will consider reviewing the internal architecture of the autoencoder. Furthermore, the future work will involve comprehensive verification using real-world data from operating buses under a wide variety of conditions to rigorously evaluate the robustness of model in practical environments.

ACKNOWLEDGMENT

This work is supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number JP24K20774, JP24K02916, and JP23K28078.

REFERENCES

[1] Courts of Japan, “Judgment delivered on November 5, 2015, Original received the same day by court clerk”, https://www.courts.go.jp /app/hanrei_jp/search, accessed October 14, 2025.

[2] J. Zhang, D. Shen, L. Tu, et al., “A Real-Time Passenger Flow Estimation and Prediction Method for Urban Bus Transit Systems”, IEEE Transactions on Intelligent Transportation Systems, vol. 18, pp. 1-11, Apr. 2017.

[3] D. Liciotti, A. Cenci, E. Frontoni, A. Mancini, and P. Zingaretti, “An Intelligent RGB-D Video System for Bus Passenger Counting”, Intelligent Autonomous Systems 14, pp. 473-484, Feb. 2017.

[4] IEE Smart Sensing Solutions, “LiDAS – Life Detection Assistance System”, https://iee-sensing.com/app/uploads/2021/08/iee-lidas-occupant-detection-sensing-solutions-factsheet.pdf, accessed September 26, 2025.

[5] M. Yamashita, T. Nishiura, A. Kubota, S. Yokoyama, H. Kawamura, K. Yuminosaki, and Y. Sato, “Development of a Method for Estimating Passenger OD and Congestion on a Route Bus using Deep Learning”, Software Symposium 2021, pp. 54-64, June 2021.

[6] N Hadisurya, et al, “Passengers Monitoring Models at Jakarta Bus Stops Using the YOLOv3”, IEEE International Conference on Artificial Intelligence, Blockchain, Cloud Computing, and Data Analytics (ICoABCD), pp. 219–224, Aug. 2024.

[7] R. Ukyo, T. Amano, T. Hiromori, H. Yamaguchi, and M. Moriya, “Implementation and Evaluation of a Pedestrian Tracking Method using Large-Scale 3D Point Clouds from Multiple LiDARs”, in IPSJ SIG Technical Report, Mobile Computing and Ubiquitous Systems (MBL), vol. 101, no. 28, pp. 1-7, Nov. 2021.

[8] W. Wang, Y. Huang, Y. Wang, and L. Wang, “Generalized Autoencoder: A Neural Network Framework for Dimensionality Reduction” , IEEE Conference on Computer Vision and Pattern Recognition, pp. 490-497, June 2014.

[9] MathWorks, “What is Autoencoder?”, https://jp.mathworks.com/ discovery/autoencoder.html, accessed September 25.

[10] H. Torabi, S. L. Mirtaheri, and S. Greco, “Practical autoencoder based anomaly detection by using vector reconstruction error”, Cybersecurity, vol. 6, no. 1, pp. 1, Jan. 2023.