Preprints.org

Article

A Study on Micro-Mobility Data Collection System Using Incremental Learning Techniques

Eunkyu Lee , Taeseong Kim , Junaid Khan , Umar Zaman , Jaebin Ku , Kyungsup Kim *

Article

A Study on Micro-Mobility Data Collection System Using Incremental Learning Techniques

Eunkyu Lee [1], Taeseong Kim [2], Junaid Khan [3], Umar Jaman [1], Jaebin Ku [1] and Kyungsup Kim [1,*]

[1] Department of Computer Engineering, Chungnam National University, Daejeon, Republic of Korea
[2] Onekick, Co., Ltd, Daejeon, South Korea
[3] Department of Environmental IT Engineering, Chungnam National University, Daejeon, South Korea
[*] Correspondence: sclkim@cnu.ac.kr; Tel.: +82-42-821-5440

Abstract: This study addressed the need for efficient data collection in micro-mobility using low-end hardware, focusing on developing a system capable of handling large real-time data streams. We proposed a novel data collection system that leverages incremental learning techniques, enabling AI-based services for micro-mobility passengers through smartphones. Our approach utilizes 3-axis acceleration sensors integrated into micro-mobility devices, collecting motion data in real-time. This data is processed through an incremental learning pre-processing server, employing the Broad Learning System algorithm, to ensure robust performance in noisy, real-world environments. The use of incremental learning is pivotal, allowing our system to adapt continuously to new data, enhancing the accuracy and relevance of the AI services provided to passengers. This research not only advances the technical capabilities in micro-mobility data handling but also has the potential to enhance the overall passenger experience through improved AI-based services.

Keywords: Incremental learning; Micro-mobility; Data collection system; AI-based services; Broad Learning System algorithm

1. Introduction

With the development and popularization of artificial intelligence technology, several intelligent service products utilizing machine learning have been launched. Micro-mobility is emerging as a highly convenient next-generation eco-friendly mode of transportation, with increasing demand in the related market. Micro-mobility refers to a variety of personal, small-scale, short-distance transportation methods that facilitate quick and convenient movement in urban and city areas. The term 'micro-mobility' is often used interchangeably with 'Smart Mobility' and 'Personal Mobility', and typically includes electric bicycles, electric scooters, shared bicycles, and personal electric vehicles [1].

Currently, with the increasing demand for micro-mobility, vast amounts of data such as departure location, arrival location, usage time, and distance traveled are being accumulated in real-time. However, this data is not being processed into meaningful information [2]. Moreover, the data to ensure the stability of suspension and steering systems that facilitate the direction change and balance of micro-mobility, and data that can help with stable braking of micro-mobility, which frequently comes to sudden stops, are not being collected. Given the nature of micro-mobility, the hardware-based protective gear for rider safety is inherently limited. Therefore, the development of software-based safety systems is necessary. For this, securing motion data of micro-mobility through various sensors is essential. Additionally, to provide safety-related services by learning from the collected vast amount of data, the construction of expensive artificial intelligence learning servers is generally required. However, in the case of learning with big data, even high-specification cluster servers may not be able to learn due to insufficient main memory, or it may take an unacceptably long time. To solve these problems, a universal learning method that can reflect only the changed elements without capacity limitations is needed. Therefore, this paper aims to develop a distributed incremental learning service system that can secure and learn large volumes of micro-mobility motion data using low-end hardware and minimal traffic, providing reliable AI services for the safety of passengers in a network-distributed incremental learning service system.

2. Related Work

2.1. Micro-mobility

Recent advancements in micro-mobility research have led to a variety of innovative approaches aimed at enhancing safety, efficiency, and operational understanding in this rapidly evolving field. These studies, each with its unique focus and methodology, collectively contribute to a deeper comprehension of micro-mobility systems and their integration into urban environments. From leveraging advanced imaging techniques for vehicle classification to employing sophisticated machine learning algorithms for operational optimization, the research landscape is witnessing a significant shift towards data-driven, technology-centric solutions. This shift is not only redefining how micro-mobility services are managed and utilized but also paving the way for more sustainable and safe urban transportation networks. The following paragraphs detail some of the key contributions in this area, highlighting the diverse methods and objectives of recent studies.

Stahl et al. [3] unveiled an innovative strategy for micro-mobility vehicle classification, employing thermal-infrared imagery. The method stands out by merging image and contour features with neuromorphic processing, leading to improved accuracy in identifying specific vehicle types. This advancement offers a unique perspective in the field of micro-mobility vehicle classification, diverging from previous studies and laying essential groundwork for future research directions. Hakan et al. [4] applied machine learning algorithms to calculate delivery times and energy costs for postal services using e-scooters. This approach incorporated a range of models such as Random Forest, Gradient Boosting, k-Nearest Neighbors, and Neural Networks, focusing on the precise prediction of crucial operational metrics. Ma et al. [5] proposed a method for predicting the riding habits of micro-mobility users and identifying factors contributing to accidents using GPS and accelerometer data from micro-mobility devices, aiming to enhance rider safety. Panagiotis et al. [6] introduced an innovative approach for examining micro-mobility routing behaviors in urban road networks. Focusing on Athens, Greece's road infrastructure, their work delves into aspects such as bike lane availability, sidewalk width, speed limits, pavement conditions, and pedestrian crossing locations. A key aspect of this initiative is evaluating perceived safety, achieved by integrating photographic and spatial data to enhance and expand the dataset. The approach, which characterizes the road environment using specific variables, employs the Dijkstra algorithm to identify the most efficient routes for various transportation modes, including e-scooters and walking. Highlighting the importance of traffic regulations and the accessibility of different transport modes, this initiative underscores essential factors in urban planning. Hakan et al. [7] delves into the factors influencing accidents among e-scooter users and assesses the probability of accidents during e-scooter travels. Analyzing data from e-scooter drivers across 15 cities in Turkey, gathered via a rental application, the approach involves statistical analysis and machine learning algorithms to estimate accident probabilities. The analysis categorizes data into dependent and independent variables, with nine independent variables encompassing both numeric and text data types related to drivers' demographics and driving status. The impact of various factors on the e-scooter driving status is examined using regression analysis and standardized effect graphs, contributing significantly to the understanding of safety in micro-mobility sharing systems. These studies propose various method related to safety based on data collected from various sensors attached to micro-mobility, but there has been no research utilizing reliable incremental learning in an environment where large volumes of data are continuously collected.

2.2. Incremental learning

With the increasing amount of data that can be collected in various fields, including micro-mobility, the importance of incremental learning is growing, not only for training models with large batches of data at once but also for continuously improving model performance in environments were data increases in real-time.

Marin et al. [8] proposed a web-based remote robotic system capable of recognizing objects and incrementally learning from new objects. This system emphasized the ability to effectively recognize

new objects in real-time by using an object recognition algorithm and continuously improving performance with an incremental learning algorithm. Chang et al [9] present a novel glaucoma detection system. This system utilizes a Generative Adversarial Network model to generate realistic Optical Coherence Tomography images from fundus images. This method aids doctors in determining whether patients require further screening, thereby increasing the chances of early detection and treatment of glaucoma. The system's effectiveness is demonstrated through experiments showing a high cosine similarity between generated and real Optical Coherence Tomography images. The performance of the proposed method surpasses that of autoencoders, Deep Convolutional Generative Adversarial Network, and U-Net. Additionally, the use of transfer learning and incremental training methods significantly improves the accuracy of glaucoma classification, highlighting the system's potential in various medical fields. Opelt et al. [10] proposed a method of incrementally learning using visual alphabet representations by sharing boundaries and spatial configurations between object categories. Lai et al [11] proposed an innovative filter to address physiological tremors in telerobot systems. This filter utilizes a Convolutional Neural Network to efficiently extract time series features, replacing the Sparse Autoencoder in the original Broad Learning System. The structure incorporates lateral connections to obtain in-depth feature information from multiple layers. An incremental learning algorithm is also integrated, allowing real-time network updates and improved accuracy through the addition of new nodes. The filter's efficiency in forecasting tremors suggests potential applications in medical areas requiring high accuracy. The work also contemplates future enhancements by combining Long Short-Term Memory with Broad Learning System to address sequence correlation problems. Park et al. [12] presented a new approach for image classification. This method involves collecting backdrop and sprite images from the Scratch programming environment and placing sprites randomly on these backdrops. The images are resized and positioned to avoid overlap and ensure visibility. The dataset includes merged images with varying numbers of sprites placed on each backdrop. The goal is to identify the sprites placed on the backdrops, despite the presence of other objects. This new dataset, fundamentally different from existing ones, is ideal for evaluating one-shot class-incremental approaches. The approach outlined in the paper aims to meet several requirements, including class-incremental and decremental learning, effective multi-label classification with limited data, handling unconventional image combinations, and ensuring the model's lightweight nature for efficient running on low-performance computers. The effectiveness of this approach is demonstrated through various experiments, highlighting its potential in diverse application domains. Chen et al. [13] proposed a hybrid incremental learning approach based on Support Vector Machine for quickly and efficiently recognizing new objects and new classes, aiming to recognize objects held in hand in real-time. Weng et al. [14] proposed a method for incrementally learning and recognizing faces in videos. The proposed method handled the problem of continuous video stream data where faces appear, disappear, and new faces are introduced over time, using deep learning and incremental Principal Component Analysis algorithms, and secured performance by incrementally updating the model as new faces are detected while maintaining the ability to recognize previously learned faces.

In the context of micro-mobility, the application of incremental learning techniques is becoming increasingly crucial for the development of efficient data collection systems. Recognizing the dynamic nature of micro-mobility data, which includes variables such as location, speed, and trip duration, a system that can adapt and learn incrementally is essential. This approach allows for the continuous integration of new data, enhancing the system's ability to make more accurate predictions and decisions in real-time. For instance, a micro-mobility data collection system using incremental learning can progressively improve route recommendations, traffic pattern analysis, and maintenance schedules based on the constantly evolving data. Such a system could also adapt to seasonal changes, special events, or urban development, ensuring that micro-mobility services remain efficient and responsive to the needs of users. The potential of incremental learning in this field lies in its ability to handle large streams of data efficiently, making it a valuable tool for cities looking to optimize micro-mobility services and improve urban transportation systems.

3. Micro-mobility Data Collection System Using Incremental Learning Techniques

3.1. Incremental learning pre-processing Server

The role of the incremental learning preprocessing server is to fetch only new data, excluding data already learned in the big data cluster. It parses unstructured raw data, optimizes it for incremental learning in a smartphone environment, and then compresses and transmits it as an archive. Figure 1 shows the configuration of the incremental learning preprocessing server.
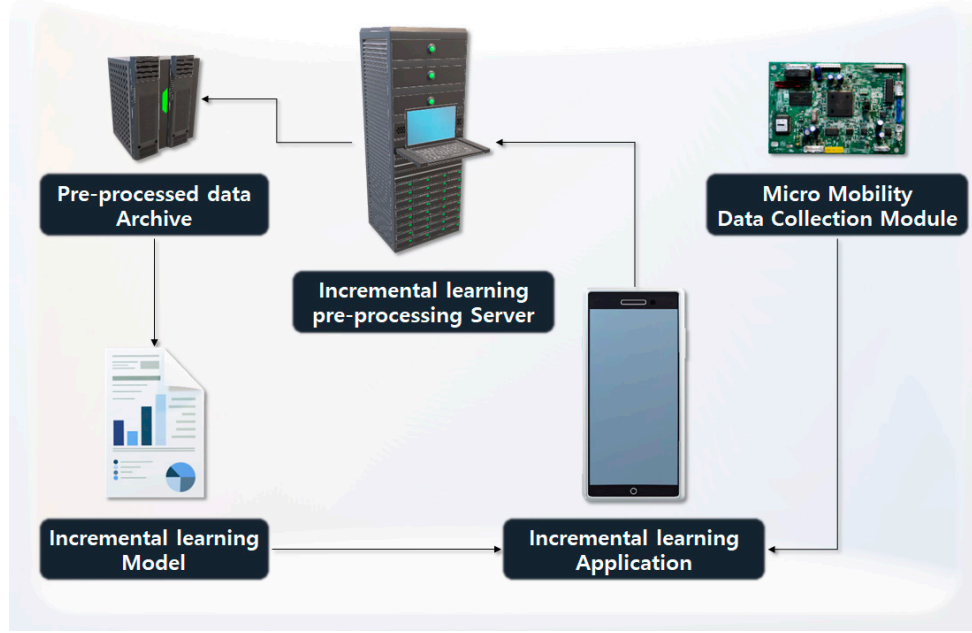
Figure 1. Proposed System Architecture.

To implement this incremental learning preprocessing server, the Broad Learning System (BLS) algorithm was used to handle data with noise in the network. The BLS algorithm is designed to be effective and efficient in incremental learning without relying on the design methods of deep learning models. It outperforms traditional deep learning models in handling online and incremental learning tasks, as it learns from new data without losing information from previous data [15]. Figure 2 shows the structure of the BLS algorithm.
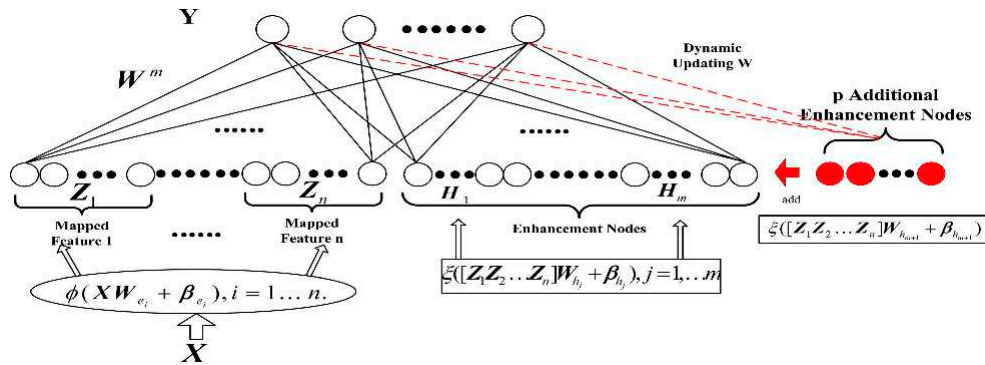

Figure 2. BLS Algorithm network structure.

3.2. Incremental learning Application

Although the specifications of smartphones have significantly improved recently, it is still inconceivable to learn big data in one go. Moreover, requesting queries to a big data cluster can be challenging when dealing with over 100,000 training data requests.

The incremental learning app can download data that has been preprocessed into a format suitable for learning, in an archived form, and then learn on its own. The incremental learning algorithm is designed to operate in a typical smartphone environment and collects motion data including the speed and acceleration in 3-axis directions, as well as basic environmental data such as temperature and humidity when using micro-mobility. Figure 3 shows the application menu and measurement screen developed for Android. Figure 4 displays the performance, manage device and device test screen.
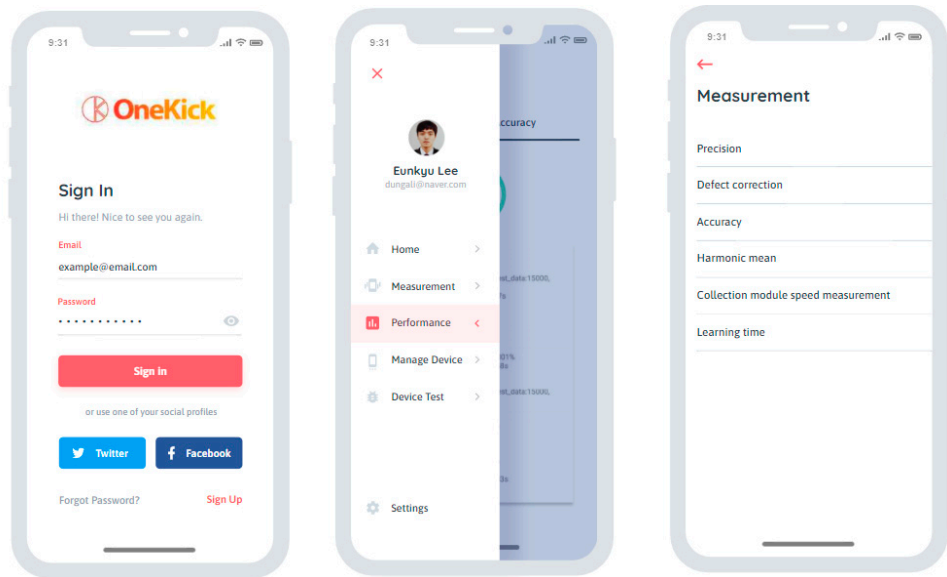
Figure 3. Application menu measurement screen.
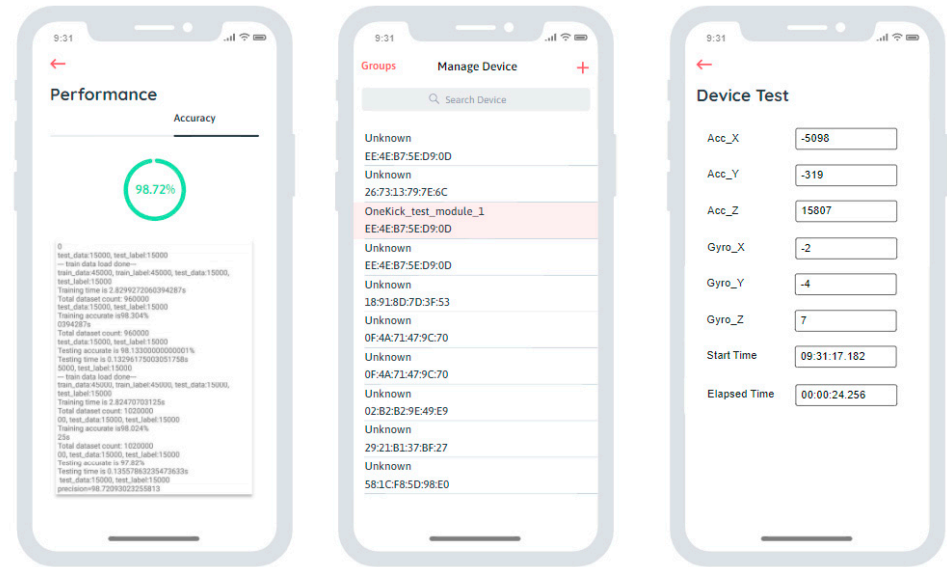
Figure 3. Performance, Manage Device and Device Test screen.

3.3. Data Collection Module

The learning information collection module is mounted on micro-mobility and collects data upon request from the incremental learning app. The collected data is then transmitted to the incremental learning app via Bluetooth. The power supply method for the module can be considered as self-generating, rechargeable, or battery replaceable. In this study, a USB-based charging method was applied. Additionally, to determine if the module is functioning correctly, an LED was installed to differentiate by color, and a Sleep/Active mode was implemented to optimize battery consumption. Figure 5 shows the design of the learning information collection module prototype.
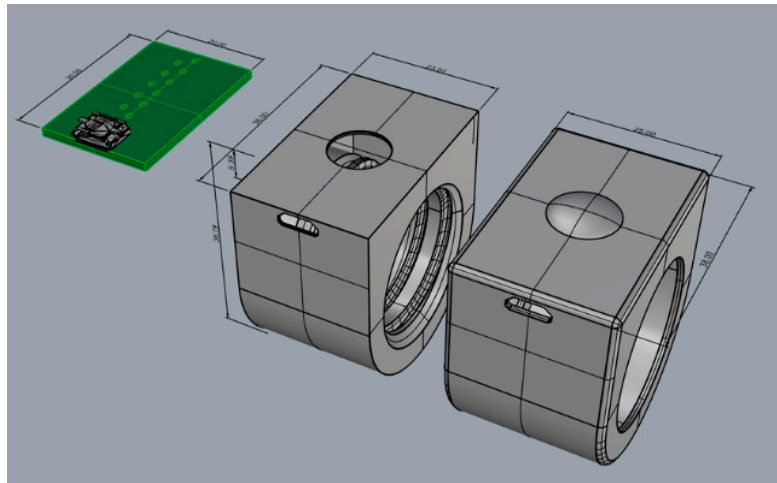

Figure 5. Proposed System Architecture.

Although the learning information collection module is primarily mounted on mobility devices, it is frequently detached for charging. Therefore, a Velcro-type mounting interface was designed and implemented for easy detachment. Figure 6 shows the Velcro-type design for detaching the learning information collection module.
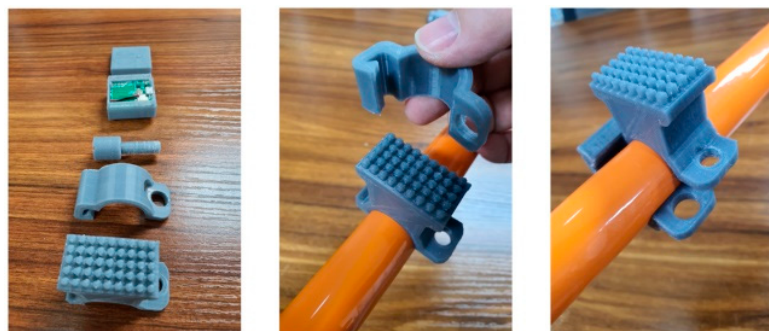

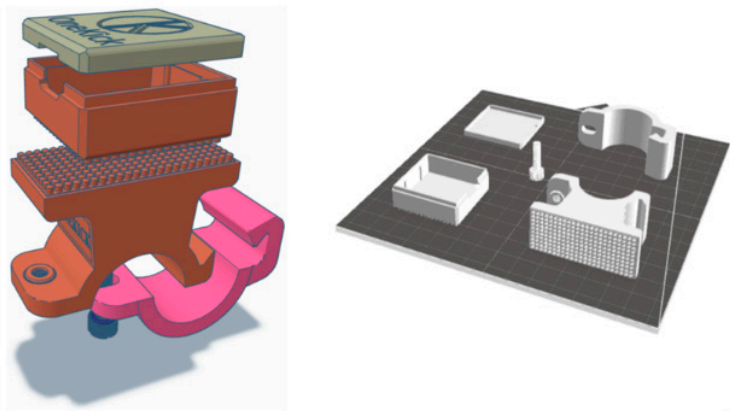
Figure 6. Velcro type design of Data Collection Module.

Figure 7. Final Design of Data Collection Module.

4. Verification of Incremental learning System

The verification of the incremental learning system was conducted by collecting 60,000 micro-mobility operation data to check whether the system continues to improve and adapt while maintaining the accuracy of previous learning when new data is presented. Table 1 shows the format of the micro-mobility operation learning dataset, and Table 2 describes each parameter of the learning model.

Table 1. Specification of micro-mobility data set.

Item	Type	Abbreviation
X-axis acceleration	double	Acc_x
Y-axis acceleration	double	Acc_y
Z-axis acceleration	double	Acc_z
Operational status	int	State_moving
Duration	long	Elasped_time_ms

Table 2. Parameter for training model.

Item	Value	Remark
N1	10	Number of nodes in each window
N2	10	Number of feature mapping layers
N3	500	Number of enhance layers
L	5	Number of incremental steps
M1	50	Number of enhance layers to add
s	0.8	Shrinkage factor
C	2E-30	Regularization coefficient

The evaluation metrics used were Precision, Recall, and Accuracy, which are commonly used to assess the performance of machine learning algorithms. Precision measures the ratio of actual correct positive predictions, which is the proportion of items correctly identified as true among those classified as true. Recall is the ratio of items correctly predicted as true among the actual true items. Accuracy is the proportion of cases where the model's prediction and the actual value are the same out of the total number of data. Table 3 is the confusion matrix, and Equations (1) to (3) are the formulas for calculating Precision, Recall, and Accuracy, respectively.

Table 3. Confusion Matrix.

Predicted / Actual	Negative	Positive
Negative	TN (True Negative)	FP (False Negative)
Positive	FN (False Negative)	TP (True Positive)

$$\text{Precision} = \frac{TP}{FP+TP} \tag{1}$$

$$\text{Recall} = \frac{TP}{FN+TP} \tag{2}$$

$$\text{Accuracy} = \frac{TN+TP}{TN+FP+FN+TP} \tag{3}$$

The ratio of data needed for training to data needed for validation was divided as 0.75:0.25, and training was conducted on a total of 1.02 million data points by repeating the collected 60,000 micro-mobility operation data 17 times. As a large amount of data continuously accumulated in this study, to correct for accuracy degradation due to overfitting, the model was simplified using singular values to avoid overfitting. The validation results initially maintained an accuracy of 97%~98% but dropped slightly to around 96% as the training data increased. Figure 7 shows the changes in training accuracy and validation accuracy over the course of the training, and Table 4 presents the final Precision, Recall, and Accuracy of the model.
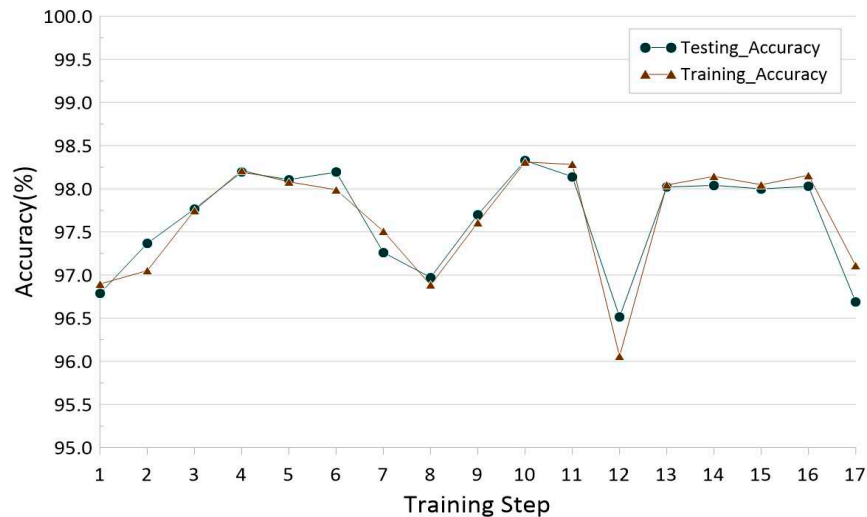
Figure 7. Training and Valid Accuracy.

Table 4. Parameter for training model.

Item	Value
Precision	97.9%
Recall	95.7%
Accuracy	96.8%

6. Conclusion

In this paper, we developed a system capable of learning from a large volume of micro-mobility data through incremental learning. The system consists of an incremental learning preprocessing server, an incremental learning app, and a learning information collection module. The incremental learning preprocessing server uses the BLS algorithm to ensure performance in online and

incremental learning tasks without losing information from previous data. The incremental learning app communicates with the learning information collection module via Bluetooth and downloads data in an archived form from the incremental learning preprocessing server to proceed with learning. The learning information collection module collects data upon request from the incremental learning app and gathers motion data including temperature, humidity, and the speed and acceleration in 3-axis directions, attached to micro-mobility.

The validation results of this system showed a slight decrease in accuracy as the data increased compared to the initial stages of learning. However, when training with a total of 1.02 million data points, the system demonstrated an accuracy of over 97%. Therefore, it is believed that this system can provide AI services for the safety of passengers by securing and learning a large volume of micro-mobility motion data, even with low-end hardware and minimal traffic.

7. Patents

The application of 'Device for collecting micro-mobility data and system for collecting Micro Mobility data and providing artificial intelligence service comprising the device(10-2023-0026527)' is approved by the Korean Intellectual Property Office.

References

1. Kalašová, A.; Čulík, K. The Micromobility Tendencies of People and Their Transport Behavior. *Applied Sciences*. **2023**, 13(19), 10559. DOI : 10.3390/app131910559
2. Fishman, E. Bikeshare: A review of recent literature. *Transp. Rev.* **2016**, 36, 92–113. DOI : 10.1080/01441647.2015.1033036.
3. Stahl, B.; Apfelbeck, J.; Lange, R. Classification of Micromobility Vehicles in Thermal-Infrared Images Based on Combined Image and Contour Features Using Neuromorphic Processing. *Applied Sciences*. **2023**, 13(6), 3795. DOI : 10.3390/app13063795.
4. İnaç, H.; Ayözen, Y.; Atalan, A.; Dönmez, C.Ç. Estimation of Postal Service Delivery Time and Energy Cost with E-Scooter by Machine Learning Algorithms. *Appl. Sci.* **2022**, 12, 12266. DOI : 10.3390/app122312266.
5. Ma, Q.; Yang, H.; Mayhue, A.; Sun, Y.; Huang, Z.; Ma, Y. E-Scooter safety: The riding risk analysis based on mobile sensing data. *Accid. Anal. Prev.* **2021**, 151, 105954. DOI : 10.1016/j.aap.2020.105954.
6. Tzouras, P.G.; Mitropoulos, L.; Koliou, K.; Stavropoulou, E.; Karolemeas, C.; Antoniou, E.; Karaloulis, A.; Mitropoulos, K.; Vlahogianni, E.I.; Kepaptsoglou, K. Describing Micro-Mobility First/Last-Mile Routing Behavior in Urban Road Networks through a Novel Modeling Approach. *Sustainability*. **2023**, 15, 3095. DOI : 10.3390/su15043095.
7. İnaç, H. Micro-Mobility Sharing System Accident Case Analysis by Statistical Machine Learning Algorithms. *Sustainability*. **2023**, 15(3), 2097. DOI : 10.3390/su15032097.
8. Marin, Roque, J; Salvador Sánchez; Pedro J. Sanz. Object recognition and incremental learning algorithms for a web-based telerobotic system. *Proceedings 2002 IEEE International Conference on Robotics and Automation*. **2002**, Vol. 3. IEEE, DOI : 10.1109/ROBOT.2002.1013643.
9. Chang, C.W.; Chang, C.Y.; Lin, Y.Y.; Su, W.W.; Chen, H.S.L. A glaucoma detection system based on generative adversarial network and incremental learning. *Appl. Sci.* **2023**, 13, 2195. DOI : 10.3390/s23208368.
10. Opelt, Andreas, Axel Pinz, and Andrew Zisserman. Incremental learning of object detectors using a visual shape alphabet. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. **2006**, Vol. 1. IEEE, DOI : 10.1109/CVPR.2006.153.

11. Lai, G.; Liu, W.; Yang, W.; Zhang, Y. A Convolutional Neural Network-Based Broad Incremental Learning Filter for Attenuating Physiological Tremors in Telerobot Systems. Appl. Sci. 2023, 13, 890. DOI : 10.3390/app13020890.

12. Park, Y.; Shin, Y. Applying Object Detection and Embedding Techniques to One-Shot Class-Incremental Multi-Label Image Classification. Appl. Sci. 2023, 13, 10468. DOI : 10.3390/app131810468.

13. Chen, C.; Min, W.; Li, X.; Jiang, S. Hybrid incremental learning of new data and new classes for hand-held object recognition. *J. Vis. Commun. Image Represent.* **2019**, 58, 138–148. DOI : 10.1016/j.jvcir.2018.11.009.

14. Weng, J.; Evans, C. H.; Hwang, W. S. An incremental learning method for face recognition under continuous video stream. *In Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, **2000**. pp. 251-256. IEEE. DOI : 10.1109/AFGR.2000.840643.

15. Chen, C.P.; Liu, Z. Broad learning system: An effective and efficient incremental learning system without the need for deep architecture. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, 29, 10–24. DOI : 10.1109/TNNLS.2017.2716952.