A hybrid clustering–regression approach for predicting passenger congestion in a carriage at a subway platform

Juhyeon Kwak[a,b], Donggyun Ku[c], Joonsik Jo[d], S.C.Wong[e,f], Seunghyeon Lee[a], Seungjae Lee[a,*]

[a] *Department of Transportation Engineering, University of Seoul, Seoul, Republic of Korea*
[b] *Department of Smart City, University of Seoul, Seoul, Republic of Korea*
[c] *Mobility Research Division, Gyeonggi Research Institute, Gyeonggi, Republic of Korea*
[d] *Division of Urban Transportation Research, Seoul Institute, Seoul, Republic of Korea*
[e] *Department of Civil Engineering, The University of Hong Kong, Hong Kong, China*
[f] *Guangdong - Hong Kong - Macau Joint Laboratory for Smart Cities, Hong Kong, China*

Abstract

This paper describes a novel integrated deep-learning framework that uses accessibility and time-varying patronage demand data to forecast the passenger congestion levels of individual carriages at a subway platform. The forecasting task involved the following challenges: 1) preprocessing spatiotemporal multivariate patronage data, 2) defining the effects of accessibility at platforms and time-series passenger demand on carriage congestion levels, and 3) designing an integrated deep-learning framework to manage heterogeneous spatiotemporal data. To address these challenges, an integrated deep-learning mechanism, namely a Conv-LSTM, was developed, which consisted of a convolutional neural network and long short-term memory (LSTM) framework to manage spatial and temporal features, respectively. Multidimensional datasets for testing and training the Conv-LSTM framework were collected from line one of the metropolitan subway systems in Busan, Korea. These datasets comprised 1) accessibility data corresponding to the entrance and exit locations at a subway platform relative to a carriage, 2) time-varying passenger demand data for a station, and 3) time-varying congestion data for a carriage. The performance of the Conv-LSTM framework was compared with those of other deep-learning approaches, namely a recurrent neural network, an LSTM, and a gated recurrent unit. The Conv-LSTM framework outperformed the other deep-learning approaches on the test dataset. This research can promote the application of deep-learning algorithms for addressing the challenges associated with handling spatiotemporal multivariate datasets and defining the relationships between congestion levels, accessibility, and passenger demand patterns for a platform in a subway station.

Keywords: Passenger congestion, Crowd safety, Artificial intelligence, Platform accessibility, Multivariate spatiotemporal data

1. Introduction

Underground railway systems, also known as subway systems, transport numerous passengers in metropolitan areas (Yang et al., 2018) worldwide. These systems have been established in major cities because, although they have enormous construction and operational costs, they also have efficient transport capacity and the ability to alleviate road traffic congestion (Nasri & Zhang, 2014). Efforts have been made to improve the operational and management strategies of railway systems to maximize the advantage of mass transit in urban transport systems. Initial research has been focused on the synchronization of timetabling optimization methods to minimize passengers' waiting times (Wu et al., 2015). Zhang et al. (2018) designed

*Corresponding author

Email addresses: juhyeon3266@uos.ac.kr (Juhyeon Kwak), dgku@gri.re.kr (Donggyun Ku), jjs971018@si.re.kr (Joonsik Jo), hhecwsc@hku.hk (S.C.Wong), seunghyeon.lee@uos.ac.kr (Seunghyeon Lee), sjlee@uos.ac.kr (Seungjae Lee)

timetables to minimize the total passenger travel time while considering train operations and passenger boarding and alighting processes. Additionally, several fare systems have been devised to establish efficient operational strategies for railway systems (Tang et al., 2020). Due to concerns regarding the spread of coronavirus disease 2019, people have recently avoided dense public transportation environments (Ku et al., 2021b). Moreover, the continual occurrence of accidents caused by crowding has highlighted the importance of crowd safety (Sharma et al., 2023). Consequently, the need for passenger congestion management in public transportation systems has emerged (Kim et al., 2015). In Korea, operational and public safety is often threatened by congestion in carriages, especially during peak hours. Therefore, safety personnel are deployed for congestion management. A unique characteristic of a subway system is the presence of individual carriages, which have different degrees of congestion (Yao et al., 2021). Although carriages are connected by narrow passages, for safety reasons, passengers are discouraged from moving between carriages. Inefficient and imbalanced space usage among carriages of a train can decrease the operational efficiency of railway systems by decreasing the passenger capacity of subway trains and increasing the risks associated with overcrowded platforms and carriages, particularly during peak hours.

Considering these aspects, the objective of this study was to use the accessibility and time-varying patronage demand data for a subway platform to predict the congestion in individual subway carriages at a specific time of day. To this end, first, a bi-level deep-learning architecture is established that combines unsupervised and supervised learning methods to improve predictive performance while considering the spatiotemporal characteristics of platforms at each station. In the first level, an unsupervised learning method is used to cluster the stations according to the temporal characteristics of the passenger behavior for a day. In the second level, a supervised learning method is used to predict the congestion levels for individual carriages at each clustered station. The accessibility of a platform is incorporated as a weight parameter in the supervised learning process to predict the congestion level. Second, a convolutional neural network (CNN) and a long short-term memory (LSTM) are integrated to simultaneously manage spatial and time-series features for the demand. The CNN extracts spatial features corresponding to the accessibility, defined by the locations of carriages and platforms at each station, whereas the LSTM manages the time-series features for the demand, comprising hourly, daily, and annual demand patterns for the subway line. The challenges in realizing this task include 1) preprocessing spatiotemporal multivariate real-world data, 2) defining the effectiveness of the accessibility at platforms, and 3) constructing a bi-level deep-learning framework that can manage time-series data collected from different spatial conditions that are intuitively related.

This paper includes six sections that outline the approaches devised to address the aforementioned challenges. Section 2 provides a comprehensive review of the relevant literature that identifies the key characteristics of congestion in subway systems and appropriate prediction methods. Section 3 describes the data used in our analysis and clarifies the characteristics of congestion in a carriage. Section 4 describes the methodological framework and approaches devised for predicting congestion levels in individual carriages. Section 5 uses field data collected from Korea to evaluate the performance of the devised deep-learning mechanism. Section 6 discusses the findings, presents the concluding remarks, and provides recommendations for future research directions.

2. Literature review

This section presents a review of the studies related to subway congestion, accessibility, and deep-learning-based demand and congestion prediction. First, the studies related to subway congestion and accessibility are reviewed to identify the congestion characteristics of subway systems and the problem of space imbalance caused by accessibility. Moreover, research on deep-learning-based demand and congestion prediction is reviewed, focusing on novel deep-learning algorithms and their applications to real-world datasets. By examining these studies, the research gaps are identified to be addressed by our subway congestion forecasting method. Overall, this section provides a comprehensive overview of the relevant literature and highlights the key findings and insights that inform our research on subway congestion and the application of deep-learning algorithms for forecasting.

2.1. Subway congestion and accessibility

Researchers have attempted to identify various factors that influence congestion in subway systems. Jiang et al. (2018) observed that the coordinated passenger inflow control strategy is influenced by three main elements: real-time passenger demand, arrival and departure times, and capacity. Zhu et al. (2018) noted that the journey times for passengers during peak hours are higher than those during off-peak hours, as passengers are often left behind. Wei & Chen (2012) developed a hybrid forecasting approach that integrated empirical mode decomposition and backpropagation neural networks to predict short-term passenger flow in metro systems. The model defined temporal factors, such as the day of the week, period of the day, and type of day (weekday or weekend), as inputs. Li et al. (2017) introduced a novel multiscale radial basis function network for forecasting irregular fluctuations in subway passenger flows. Their approach outperformed prevailing computational intelligence methods for non-regular demand forecasting by at least 30 min. Zhu et al. (2021) investigated the factors influencing station selection behaviors of passengers and predicted the behaviors using three variables: walking score, public transportation accessibility level, and service and facility indices. Furthermore, several solutions have been devised to relieve passenger congestion in subway systems. Zou et al. (2018) presented a method based on station inflow control to resolve the recurrent congestion problem in subways during peak hours. The method mitigated congestion by determining passenger flow distributions in a subway network using a traffic assignment model without capacity constraints and by controlling the bottleneck through a feedback-based bottleneck elimination strategy. Ding et al. (2021) found that the congestion in a subway train varies by carriage and developed a method to enhance the uniformity of passenger distribution across carriages.

Moreover, numerous studies have been conducted to examine the factors responsible for uneven carriage congestion and their implications. Li et al. (2019) investigated the impact of group behavior on a platform when boarding a subway and found that passenger behavior patterns depend on platform structure. Kuipers et al. (2021) investigated the effect of passenger-related factors on train dwell time and showed that this time was significantly affected by clustered boarding, which was closely related to the boarding door position. Among studies on subway passenger boarding and alighting process mechanisms, Lee et al. (2018) found that waiting passengers tend to cluster near the platform entrance, suggesting that the proximity to the entry point likely affects passenger distribution. Oliveira et al. (2019) demonstrated the effect of the riding concentration phenomenon on overall railway efficiency through video recordings. They noted that passengers tend to ride on carriages within 24 m of a platform entrance. Fang et al. (2019) analyzed the passenger distribution in the London subway using the weight data for each carriage and showed that 44% of passengers chose their carriage based on the accessibility of the destination station during the morning peak. Liu et al. (2016) emphasized the importance of passengers' waiting distribution and developed a model that incorporated the distance from a platform entrance (as a significant influencing factor) to model passenger distribution at the platform before train arrival.

Deep-learning techniques have recently been applied for predicting subway passenger flow patterns. Liu et al. (2019) devised a deep-learning architecture named deep passenger flow to predict inbound and outbound passenger flows in subway systems. Similarly, Yang et al. (2020) presented a practical and hierarchical passenger flow estimation framework, which used various passenger flow variables in a multilayer hierarchical flow network based on deep learning. Their model successfully predicted the time-varying origin–destination (OD) matrix, passenger departure rates, and travel time of passengers with greater accuracy than the existing dynamic OD estimation methods. Chen et al. (2021) used a Conv-LSTM model to reflect the waiting time of passengers who wished to board the next train. This model yielded excellent predictions, especially for station passenger flow congestion, by extracting the spatial and temporal characteristics of passenger flow congestion.

The above-described comprehensive review identifies research gaps related to subway congestion characteristics. The congestion of subways exhibits a time-series variation that is closely related to the distribution of passenger demand on a platform. As shown in Figure 1, passengers tend to select a carriage near the platform entrance or prefer a carriage closer to the exit at their arrival station, resulting in varying levels of congestion between carriages. Thus, by accurately predicting the subway congestion level, the spatiotemporal characteristics of passenger behaviors at a platform can be understood. In addition, our review indicates that deep-learning strategies can effectively reflect the spatiotemporal characteristics of congestion.

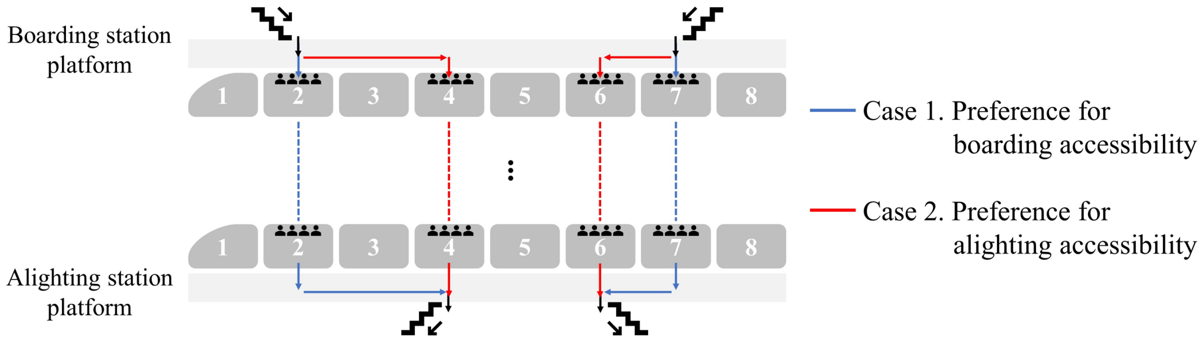

Figure 1: Travel types of passengers according to accessibility preference for different carriages.

2.2. Deep-learning-based prediction

Prediction of travel demand and traffic congestion is a key task in transportation modeling and analysis. With the rapid expansion of big urban data, the types of data being used for such analyses have increased (Batty, 2016). Such data diversification has ushered in new possibilities for data utilization in various sectors, including transportation (Chen et al., 2016). Before the development of deep neural networks, prediction models were typically based on machine learning algorithms. For instance, Myung et al. (2011) predicted travel time using the k-nearest-neighbor (KNN) method, whereas Rashidi et al. (2014) modeled bus dwell time using the decision tree method. In addition, Chiabaut & Faitout (2021) investigated real-time estimation techniques for traffic conditions and travel time on French highways using principal component analysis (PCA) and clustering methods.

Although traditional machine learning models have been used for solving simple and well-constructed problems, they are likely ineffective in addressing complicated problems involving real-world applications, such as predicting traffic phenomena involving multiple variables and situations. With recent advances in hardware and the increased availability of big data, deep-learning algorithms with unique architectures (such as CNNs) have been developed (LeCun et al., 1989). Deep-learning algorithms can consider various variables that have not been previously considered, and their high accuracy makes them attractive for application in the transportation field. Recurrent neural networks (RNNs), which were developed to learn data with time-series characteristics (Rumelhart et al., 1986), have been applied in various transportation-related studies, such as for predicting locations with high demands for taxis (Ku et al., 2021a). However, RNNs are prone to long-term depth and vanishing (or exploding) gradient problems. LSTM frameworks were developed to address these problems. For instance, Mohanty et al. (2020) used an LSTM neural network to predict the traffic congestion in neighborhoods within a region. The gated recurrent unit (GRU) was devised as a simplified LSTM variant having fewer parameters, which can facilitate the training process (Cho et al., 2014). Ham et al. (2021) developed the encoder–RNN–decoder framework using an LSTM and a GRU to create a spatiotemporal demand-prediction model for e-scooter sharing services. The mean squared errors (MSEs) of the LSTM-based and GRU-based models were found to be similar, and both models could efficiently solve spatiotemporal data problems. Finally, Bogaerts et al. (2020) devised a deep neural network architecture combining convolutional layers and LSTM cells. This model could simultaneously extract the spatial and temporal features of traffic, rendering it useful for predicting traffic phenomena.

Moreover, researchers have combined learning methods to obtain enhanced predictive performance. For instance, Antoniou et al. (2013) established an approach that integrates neural networks and KNN clustering methods to predict traffic conditions and can improve the accuracy of existing traffic estimation and prediction models. Wang et al. (2018) used a preprocessing technique that involved PCA and clustering, which was followed by prediction using a multicell neural network. The prediction accuracy was considerably improved by the preprocessing step. Similarly, Chen et al. (2017) developed a fuel consumption estimation model and devised a clustering-based regression model that considered data characteristics, thereby improving estimation performance. Gerum et al. (2019) devised a hybrid prediction method to predict defects in railways. This approach clustered the features of the rail segments using the k-means clustering method and obtained predictions via a regression model. Meanwhile, predicting spatiotemporal data often involves combining different types of deep learning models to leverage their respective strengths. Conv-LSTM combines CNN and RNN, allowing it to capture both spatial and temporal features effectively. Qu et al. (2023) proposed a flight delay prediction model based on Conv-LSTM which combines CNNs to capture spatial features and LSTM networks for temporal features. Spatiotemporal Graph Convolutional Network (STGCN) merges Graph Neural Network (GNN) with temporal convolutions. Yu et al. (2017) utilized STGCN for traffic flow prediction in large urban road networks, successfully capturing both the spatial dependencies between road segments and their temporal evolution. GWNN (Graph Wavelet Neural Network) integrates wavelet transforms with GNN. Zhang & Ma (2023) applied GWNN to traffic speed prediction, demonstrating its ability to capture spatial correlations in road networks at different scales and improve prediction accuracy.

The above-described review of research on traffic prediction highlights that deep-learning algorithms can predict multidimensional nonlinear relationships, thereby enhancing the performance of traffic environment prediction. Specifically, CNN and RNN algorithms can be used to extract the spatial and temporal features of traffic, respectively. Using a bi-level model that combines unsupervised and supervised learning methods can help enhance the accuracy of traffic prediction.

2.3. Research gaps and contribution

When only real-time data is available, travelers can make decisions only after arriving at the platform. However, by utilizing the short-term and long-term forecasting results from the prediction models, it becomes possible to make decisions in advance, starting from the actual departure and arrival points of a trip. If congestion is anticipated, travelers can choose less congested times and adjust their travel schedule accordingly, contributing to congestion mitigation. Accurate future congestion predictions enable proactive measures, such as selecting optimal travel times to avoid peak congestion periods, thereby improving overall travel efficiency and comfort. This emphasizes the importance of developing and implementing robust forecasting models for transportation systems. In this study, spatiotemporal features of a subway are extracted and used for subway congestion prediction. To this end, two approaches are developed. First, a novel congestion prediction framework that combines unsupervised and supervised learning methods is introduced. Second, a CNN and an RNN are combined to establish a time-series prediction model that incorporates spatial characteristics. Our study addresses the abovementioned research gaps and enhances the accuracy of congestion prediction. Specifically, our framework facilitates congestion prediction in specific subway carriages, thereby enabling the development of strategies for enhancing the convenience of subway users and efficiency of subway operations.

3. Data description

The proposed methods are validated using field data collected in Korea. First, a statistical analysis is performed to examine the data characteristics. Second, the relationship between the accessibility of carriage doors on platforms and spatial imbalance by analyzing the statistical significance of the accessibility variable is investigated. Finally, the performance of the proposed method in predicting carriage congestion at different times of the day is evaluated. Table 1 defines the indices, parameters, and variables used in this study.

Index and variables	
$d \in D$	stations
$k \in K$	carriages
$t \in T$	time steps
$l \in L$	layers in the CNN framework
$m \in M$	rows in a kernel in the CNN framework
$n \in N$	columns in a kernel in the CNN framework
$p \in P$	rows in the input layer in the CNN framework
$q \in Q$	columns in the input layer in the CNN framework
P_{full}	pressure of a full carriage
P_0	pressure of an empty carriage
$P_{dk}^{(t)}$	pressure of the k-th carriage at the d-th station at time t
$c_{dk}^{(t)} \in C$	congestion rate of the k-th carriage at the d-th station at time t
a_{dk}	accessibility of the k-th carriage at the d-th station
$enter$	numbers of passengers entering subway stations
$exit$	numbers of passengers exiting subway stations
tmp	temperature (°C)
$precip$	precipitation (mm)
$wind$	wind speed (km/h)
hum	humidity (%)
sun	amount of sunlight (W/m^2)
$cloud$	cloud cover (%)
$peak$	indicates whether the time period is peak (1) or off-peak (0)
$covid$	number of confirmed COVID-19 cases
Parameters and variables of the unsupervised learning method	
μ_i	center of the i-th cluster
$S_i \in S$	set of points belonging to i-th cluster
Parameters and variables of the supervised learning method	
$x^{(t)}$	input of the RNN at time t
$h^{(t)}$	hidden state at time t
u	weight of an input vector
w	weight of a hidden state
b	bias of the RNN
$y_{pq}^{l} \in Y$	input for row p and column q in vector Y^l
h_{pq}^{l}	hidden state in row p and column q in the l-th layer
u_{pq}	weight in row p and column q of an input layer
$w_{mn}^{l} \in W$	weight in row m and column n in vector W^l
b^l	bias in the l-th layer

Table 1: Parameters and variables used for modeling and estimation.

3.1. Data collection and preprocessing

Passenger data sets were collected from January 2020 to March 2021 to investigate the passenger congestion level in each carriage on Busan Line 1. The dependent variable is the congestion in a carriage, which is calculated for each train consisting of eight carriages. The passenger congestion is indicated by the pressure measured by the air-spring sensor of the train. Specifically, the congestion rate is calculated by dividing the current pressure by the pressure when the carriage is full, as shown in Equation 1. The unit of pressure is kilopascals (kPa), and the calculated rate of passenger congestion is described as a percentage (%).

$$c_{dk}^{(t)} = \frac{P_{dk}^{(t)} - P_0}{P_{full} - P_0} \tag{1}$$

Figure 2 illustrates the unbalanced distribution of passenger congestion among carriages in 2020, and Table 2 presents the statistical characteristics of the congestion levels. The results indicate that the least and highest congestion was observed in the fifth and eighth carriages, respectively.

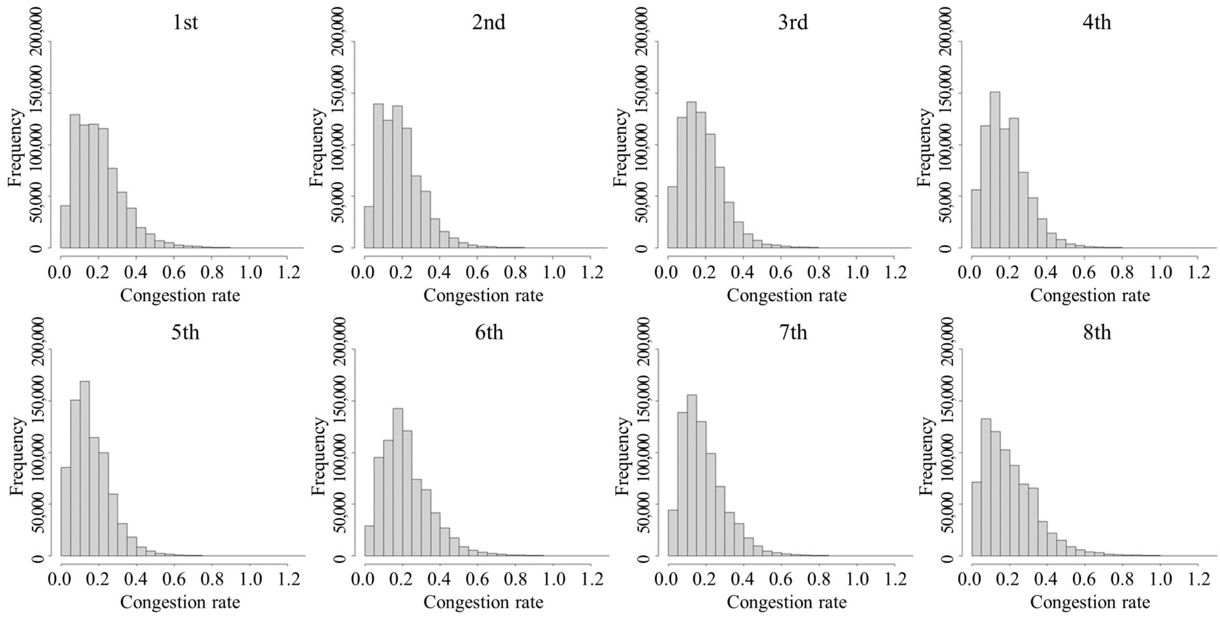

Figure 2: Histogram of rates of passenger congestion in different carriages.

	1st	2nd	3rd	4th	5th	6th	7th	8th	Total
1st Qu. (%)	10.77	10.96	10.26	10.23	8.05	12.50	10.39	9.52	10.23
Median (%)	18.46	17.57	16.67	17.05	13.79	19.48	15.85	17.46	17.05
Mean (%)	20.56	19.18	18.15	18.62	16.04	22.10	18.61	20.31	19.20
3rd Qu. (%)	27.69	25.68	24.36	25.00	21.84	28.95	24.68	26.98	25.97
Max. (%)	120.00	104.05	102.56	91.01	86.21	111.11	103.90	126.98	126.98
S.D (%)	12.64	11.39	11.29	11.05	10.25	12.83	11.66	14.00	12.06
Accessibility	2.250	2.000	2.300	2.950	2.875	2.175	1.825	2.075	2.306

Table 2: Statistics of rates of passenger congestion in a carriage.

The independent variables used in this study are the demand, accessibility by carriage, weather, and other events, owing to the following reasons. The numbers of passengers entering and exiting subway stations are the variables most directly related to the congestion level of a subway. Weather conditions affect the demand for public transportation (Wei et al., 2018). Thus, variables that indicate the weather conditions for a day, such as precipitation, temperature, and wind speed, are included in the analysis. Additionally, because the demand for public transportation varies significantly between peak and off-peak times, a dummy variable is introduced to assign a weight to the peak time.

The accessibility variable is used as a weighted variable for allocating the demands for different carriages at each subway station. Because data normalization is crucial for improving the prediction accuracy of deep-learning models, each variable is normalized using the Z-score normalization method (Patro & Sahu, 2015).

3.2. Relationship between accessibility and congestion level

A statistical test method is used to investigate the relationship between accessibility and carriage congestion. Specifically, statistical tests are conducted to determine any differences in congestion between groups based on seven accessibility levels. Figure 3 shows the box plot of the distribution of congestion by accessibility. The groups with accessibility rankings from 1 to 4 exhibit a similar distribution, and the distribution decreases from the group ranked 5. Although the groups with accessibility rankings of 1 to 4 exhibit similar distributions in the interquartile range, the distributions of values located above the third quartile are different. These results suggest that the accessibility from a platform entrance to a carriage door considerably affects the level of passenger congestion, although the effect is not proportionate and increases above a specific level. That is, subway passengers prefer to board using a door that is close to their platform.

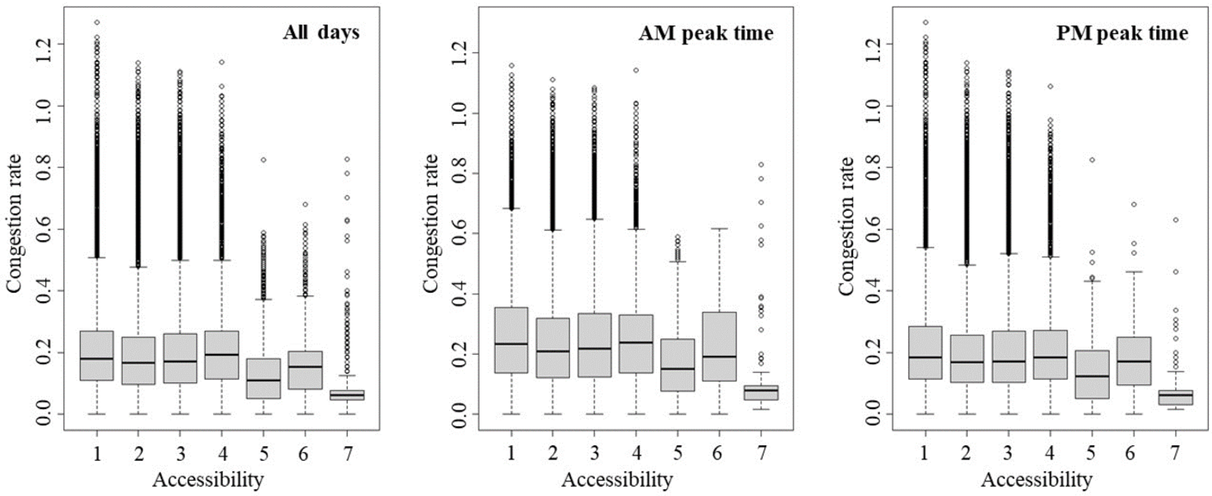

Figure 3: Distributions of congestion rates for different carriages.

4. Methodology

This section describes our methodological framework and the main concepts for analyzing the spatiotemporal characteristics of passenger behavior at a platform and predicting the carriage congestion levels. Figure 4 illustrates the complete framework. Data pertaining to the rates of passenger congestion for each carriage and variables related to the subway passenger demand are collected. Temporal and spatiotemporal approaches that reflect the time-series spatiotemporal characteristics of subway congestion are developed. Section 4.1 describes the temporal approach that categorizes stations according to the time-series characteristics through unsupervised learning. A deep-learning-based regression model is established to predict the level of carriage congestion for each clustered station. Section 4.2 describes the spatiotemporal approach that reflects the adjacent subway station and carriage characteristics. Section 4.3 explains the strategy used for generating predictions from the preprocessed data, i.e., the deep-learning model that is used to generate predictions from extracted characteristics.

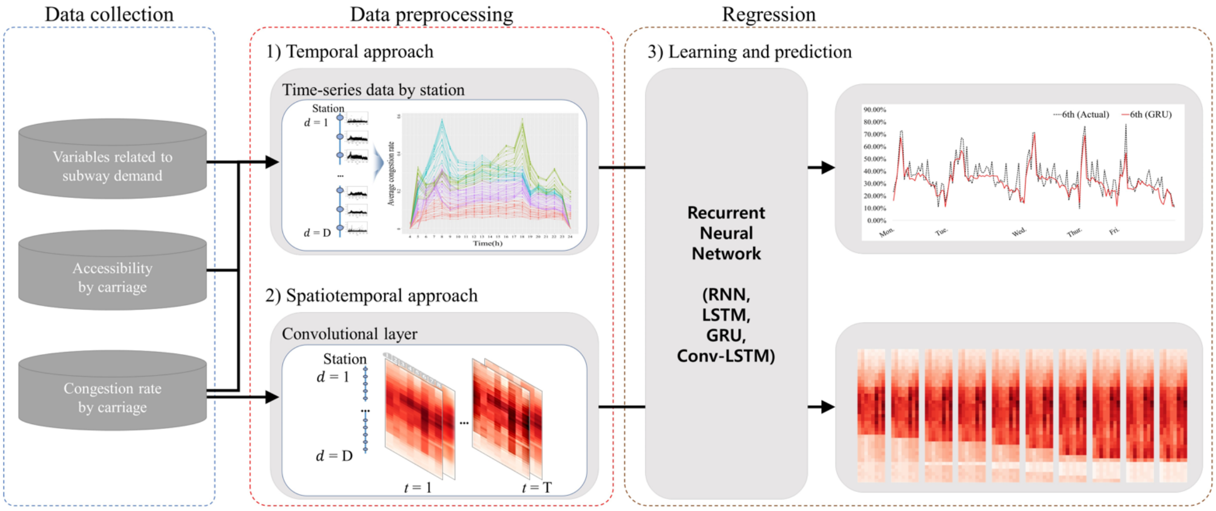

Figure 4: Process flow of our method, from data collection to modeling of the congestion level in a carriage.

4.1. *Temporal approach*

A clustering method is used to examine the temporal characteristics of passenger congestion in an individual carriage in subway systems. Figure 5 shows the framework of the bi-level deep-learning method. In the first stage, time-series characteristics of passenger congestion in each carriage through clustering are extracted. In the second stage, a cluster-specific regression model is established to predict the level of passenger congestion in a carriage corresponding to a subway line.

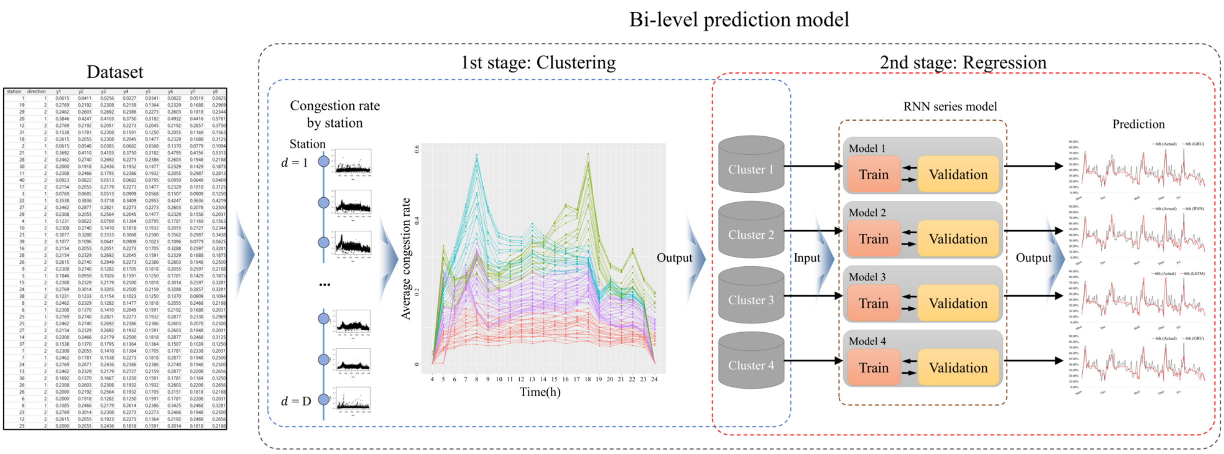

Figure 5: Data processing in the temporal approach.

Let μ_i be the center of the i-th cluster, and S_i be the set of points belonging to this cluster. The overall variance is V. The objective of the algorithm is to find S_i that minimizes V, which can be expressed as

$$\underset{S}{argmin}\, V = \underset{S_1,S_2,\ldots,S_k}{\operatorname{argmin}} \sum_{i=1}^{k} \sum_{x_j \in S_i} |x_j - \mu_i|^2 \tag{2}$$

k data are randomly extracted from the dataset and assigned to μ_i. The distance of each datapoint to μ_i is calculated, and the datapoints most similar to the cluster centroid are identified. New datapoints with the highest similarity are assigned to the cluster. Subsequently, the centroid of the cluster is recalculated, i.e., the centroid is recalculated based on the reallocated clusters. This process is repeated until the cluster to which each datapoint belongs remains unchanged. This process can be mathematically expressed as in Equations 3∼4.

$$S_i^{(t)} = x_p : \left| x_p - \mu_i^{(t)} \right|^2 \leq \left| x_p - \mu_j^{(t)} \right|^2 \forall j, 1 \leq j \leq k \tag{3}$$

$$\mu_i^{(t+1)} = \frac{1}{\left| S_i^{(t)} \right|} \sum_{x_j \in S_i^t} x_j \tag{4}$$

Clustering techniques are used to investigate the predictive performance of stations with similar temporal characteristics of passenger congestion in a carriage. Specifically, the average congestion rate for each time of the day is used as a variable for clustering. The k-means clustering algorithm is used to determine the optimal number of clusters, k, required to differentiate clusters.

4.2. Spatiotemporal approach

The spatiotemporal approach simultaneously considers the spatial and temporal characteristics of the daily passenger demand patterns. To this end, a Conv-LSTM model (Figure 6) is designed, motivated by the work of Shi et al. (2015). Convolutional layers are used to extract the spatial properties of carriages, and an RNN is used to capture temporal characteristics.

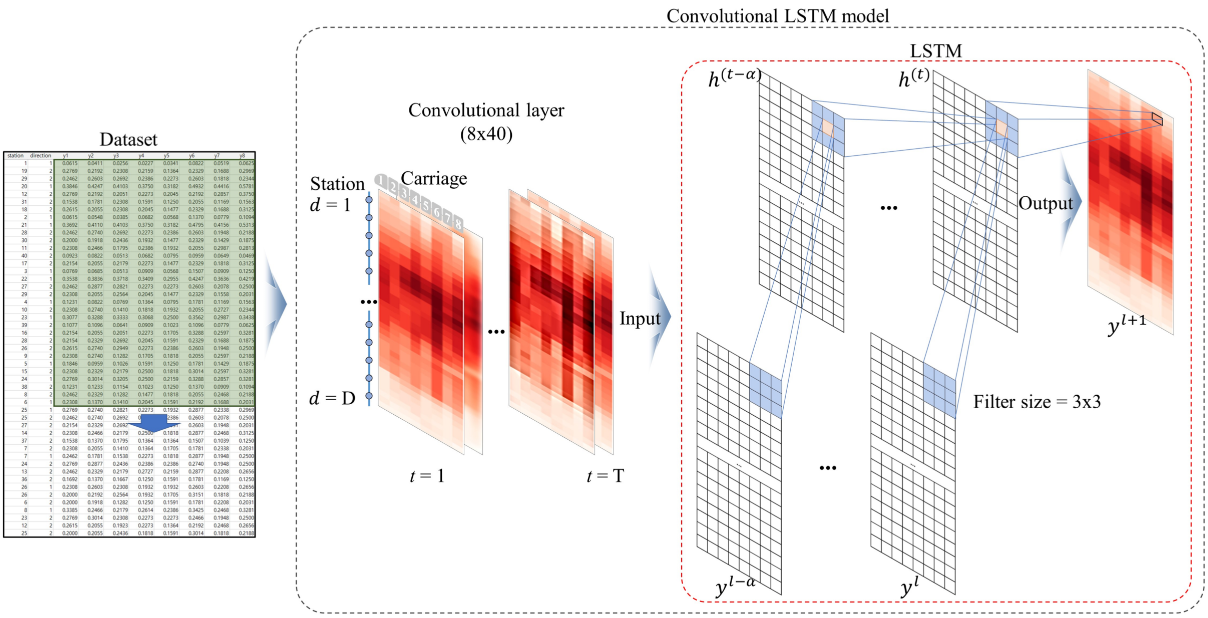

Figure 6: Data processing in the spatiotemporal approach.

The design principle of the CNN is the same as that of a standard neural network, but with different connection patterns between neurons in adjacent layers. In contrast to a multilayer perceptron, in which each node is fully connected to a node in the previous layer, the neurons in a CNN are connected only to small areas in the previous layer for exploiting spatially local correlations. By leveraging these concepts, two-dimensional spatial properties can be captured within a deep-learning framework. For instance, Dabiri & Heaslip (2018) used CNNs to extract the spatial characteristics of Global Positioning System trajectories to infer the means of travel. Zheng et al. (2020) devised a Conv-LSTM model for short-term traffic flow prediction. To enhance the predictive performance of carriage congestion, a method for extracting spatial and short-term temporal features is developed based on the Conv-LSTM network.

The proposed Conv-LSTM model is calibrated and validated using the congestion levels of all carriages and stations from the data collected from the Korean subway line. Specifically, the model is used to predict the congestion map of 40 stations × 8 carriages, which can be interpreted as a digital image that changes over time. To reflect the headway of the actual train, the interval is set as 5 min. Figure 7 shows an example of the 2-hour morning peak data, where the x-axis represents eight carriages in each frame, and the y-axis represents 40 stations. The progression of each frame is observed to propagate the congestion level, and the primary boarding and alighting stations are identified.

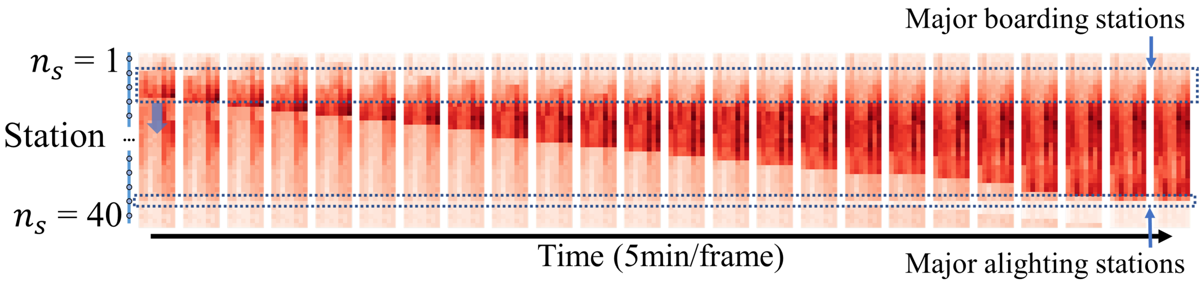

Figure 7: Sample of the convolutional layer.

4.3. Regression with deep learning

A deep-learning model using a neural network with two hidden layers is constructed to calculate the congestion in each carriage, where each output node represents one carriage. The loss function is configured using the MSE, and the Adam optimizer is used. To select the most appropriate RNN layer, the performances of three models, namely SimpleRNN, the LSTM, and the GRU, are compared with that of the Conv-LSTM model. The LSTM model was originally developed to address the problem of long-term dependency and gradient vanishing associated with the SimpleRNN model (Hochreiter & Schmidhuber, 1997). The input and output activation functions are sigmoid, and the activation function of the gate is a hyperbolic tangent. The GRU model is a variant of the LSTM model that combines a forget gate and an input gate into a single update gate and regulates the flow of information through a reset gate. The GRU model has fewer parameters to learn than the traditional LSTM model and is known to perform better when the data volume is low (Rosenblatt, 1958). The Conv-RNN model was originally developed for image-based sequence recognition and improved performance by rotating each part of an image and returning the results. Shi et al. (2016) performed optical character recognition using convolutional and recurrent layers. Lee et al. (2019) combined a CNN and an LSTM to model car-following behaviors on multi-lane motorways. Their multi-lane stochastic optimal velocity model could solve unpredictable fluctuations in the vehicle speed. Liu et al. (2017) constructed a Conv-LSTM module for short-term traffic flow prediction and effectively extracted spatiotemporal features. Their method achieved a high prediction accuracy.

5. Results of a case study on real-world subway data

Our framework is evaluated using field data collected in Korea. Specifically, the temporal and spatiotemporal methods are applied to real data to generate predictions, and their performances are compared. Finally, the applicability of our framework is demonstrated.

5.1. Temporal approach

The k-means clustering algorithm is used to identify the time-series characteristics among stations. Forty stations in two directions are analyzed, and the optimal number of clusters is determined using the Elbow

Method. The Elbow Method involves incrementally increasing the number of clusters and calculating the Total Within Sum of Squares (WSS) for each cluster configuration, which is then plotted on a graph. In Figure 8, the WSS decreases sharply as the number of clusters increases but the rate of decrease becomes gradual after four clusters. This point where the decrease slows down is referred to as the 'elbow,' and thus, four is chosen as the optimal number of clusters. When analyzing the average congestion by cluster according to the time zone, the clustering results are presented in Figure 9. Cluster 1 includes non-crowded stations; Cluster 2 contains stations in normal conditions; and Clusters 3 and 4 include stations with AM- and PM-peak characteristics, respectively. Specifically, the stations in the downward and upward directions belong to Clusters 3 and 4, respectively. Each station exhibits differences in time-series fluctuations, which can affect the predictive performance. Therefore, the predictive performance of each station is evaluated by classifying it based on the cluster it belongs to.

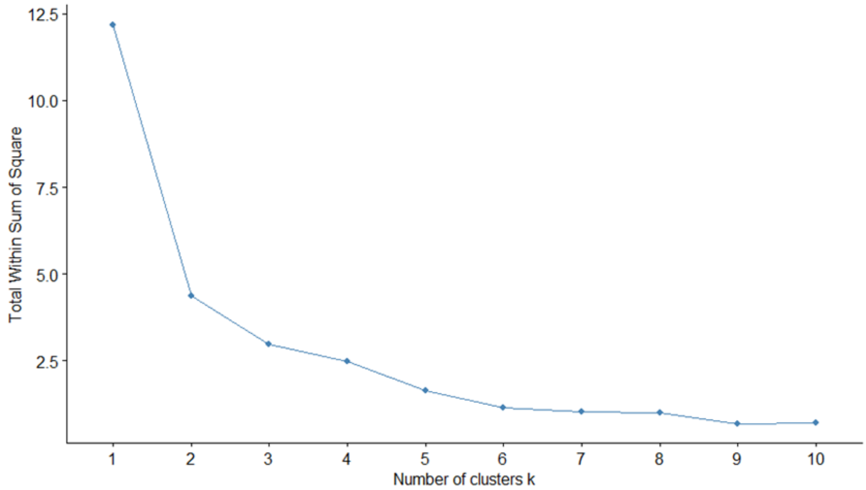

Figure 8: The relationship between the number of clusters and the total within sum of squares.

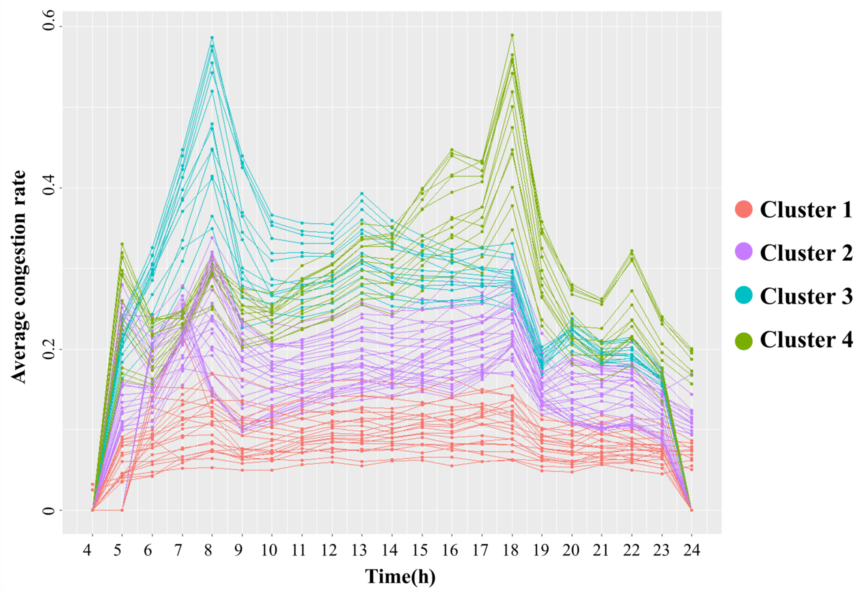

Figure 9: Clustering results.

Next, a typical RNN series model is used for learning and prediction. The congestion levels in the fifth and sixth carriages are the highest, as presented in Table 2. Figure 10 shows the actual congestion and predicted congestion (obtained by each model) for the Seomyeon station from January 4 to January 8, 2021. Seomyeon station, which belongs to Cluster 3, has the highest number of passengers on Busan Line 1 and is crowded during the AM peak. Although all of the models exhibit prediction errors, they accurately predict the congestion trends, especially the explosive increase in congestion during peak hours. In terms of the congestion distribution by carriage, the fifth and sixth carriages exhibit the greatest difference, even though they are adjacent. The prediction results for the fifth and sixth carriages are shown in Figure 10. The RNN model yields a constant prediction for all days of the week, leading to an overestimation of the peak on Tuesday morning when the maximum congestion rate decreases. In contrast, the other two models adapt to specific situations and yield predictions based on time-series variations. Therefore, the Tuesday morning peak time, which is overestimated by the RNN model, is accurately predicted by these models.

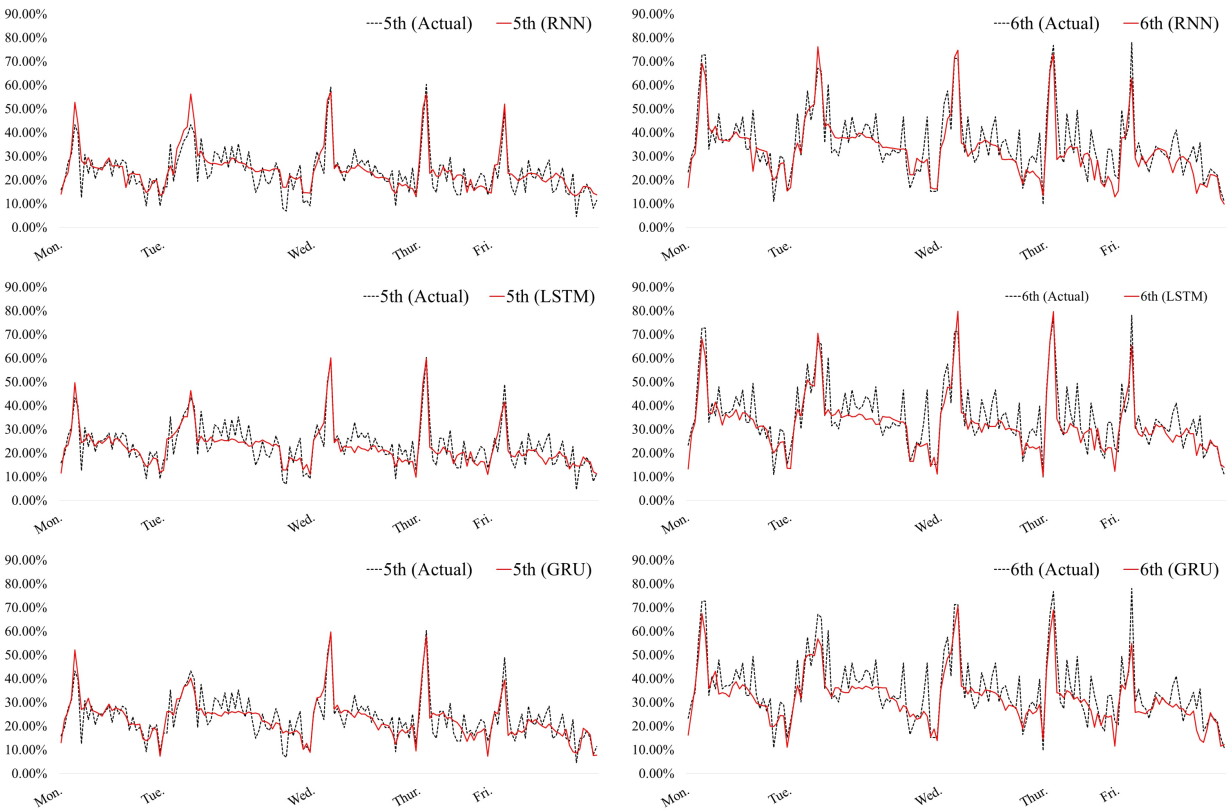

Figure 10: Comparison of prediction performance by each model (left: 5th carriage, right: 6th carriage).

The subway system accommodates a large number of users, and congestion problems arise during peak hours. Effective subway capacity management requires accurate prediction of congestion levels, which is challenging, especially in the presence of non-typical patterns. Considering these aspects, LSTM or GRU models that can handle time-series variations are promising tools for real-world prediction applications. However, our empirical results indicate that GRU models tend to underestimate congestion situations. The LSTM model, which yields more accurate congestion predictions than the GRU model, is thus the optimal model for predicting congestion levels.

5.2. Spatiotemporal approach

The Conv-LSTM model demonstrates good performance in predicting the level of carriage congestion, as it accounts for the influence of adjacent cells. Specifically, the x-axis reflects the effects of the left and right carriages, and the y-axis reflects the influence of the station location (downstream or upstream). The CNN filter size is 3 × 3. Each cell in the prediction frame represents the congestion effect of two front and rear stations and two left and right carriages. The Conv-LSTM model predicts the next frame by exploring the time-series changes between each frame in the video. Figure 11 presents the results of predicting carriage congestion levels for an hour. The Conv-LSTM model demonstrates excellent performance in predicting the level of carriage congestion across 40 stations and eight carriages. The Conv-LSTM model also accurately predicts the situations. Based on the input data, the model identified the location of stations and carriages where pedestrians were crowded. The prediction results express the congestion level of each carriage differently even at the same station. The congestion of the next frame is predicted based on the previous situation. The results show congestion spreading to the next station. However, as the prediction time increases, the prediction accuracy decreases.

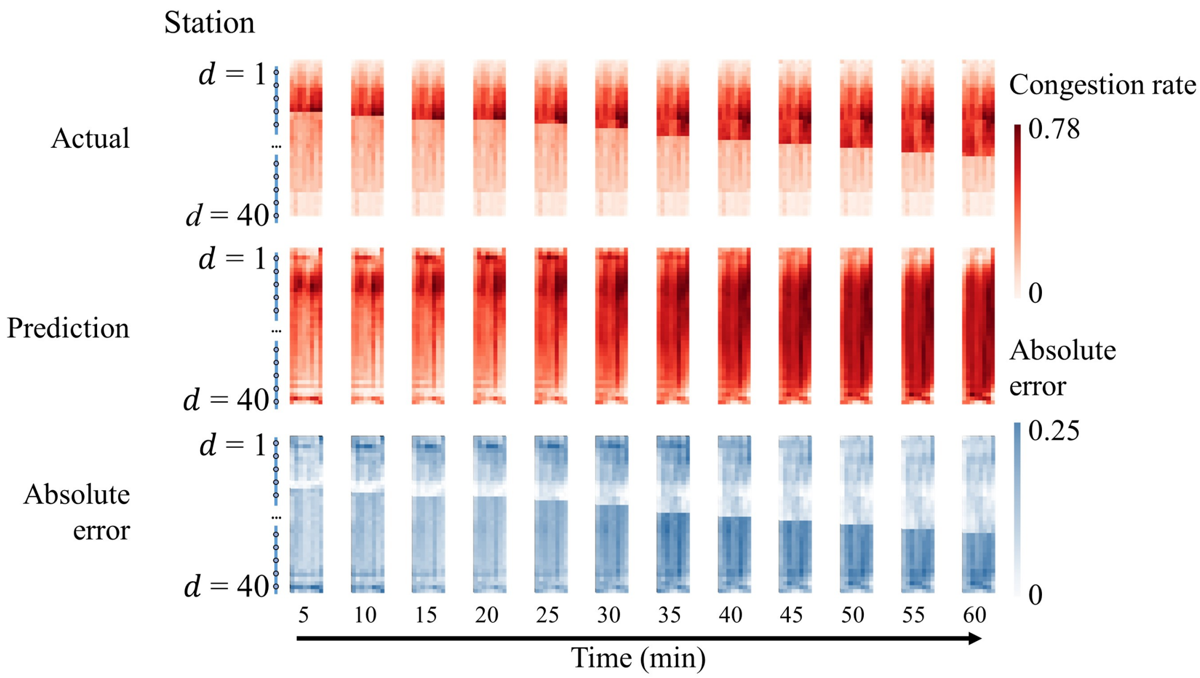

Figure 11: Results of prediction using the Conv-LSTM models.

5.3. *Comparison of the performances of different models*

The deep-learning model is trained on data collected from January to December 2020. The forecasting performance of the model is evaluated on data from January to March 2021. The metrics used to evaluate the model performance were the root MSE (RMSE), symmetric mean absolute percentage error (SMAPE), correlation coefficient, and median absolute percentage error (MDAPE). The RMSE measures the average error in prediction. The SMAPE measures the average error rate and is typically used to complement mean absolute percentage errors, which cannot be defined if the actual value is zero. Because subway congestion data often includes values close to or equal to zero, the SMAPE is used as the metric in this study. The correlation coefficient is used to evaluate whether the prediction tendencies are consistent with the time-series fluctuations of the actual data. The MDAPE is a statistical measure used to assess the accuracy of predictive models, particularly in the context of time series forecasting or regression analysis. It is an extension of the Mean Absolute Percentage Error (MAPE) but uses the median instead of the mean to provide a more robust measure against outliers.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\widehat{y_i} - y_i)^2}{n}} \tag{5}$$

$$\text{SMAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\frac{|\widehat{y_i} - y_i|}{(|\widehat{y_i}| + |y_i|)/2} \tag{6}$$

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \tag{7}$$

$$\text{MDAPE} = median(\left|\frac{y_i - \widehat{y_i}}{y_i}\right| \times 100) \tag{8}$$

All training and prediction of the deep learning models were conducted in a consistent hardware and software environment consisting of an Intel Xeon Gold 6240R CPU, 768 GB RAM, Nvidia GeForce RTX 3090, and Windows 10 OS, to prevent discrepancies caused by differences in experimental conditions. Additionally, the versions of Python and its libraries have been made compatible with the hardware in use. The Keras deep learning framework was employed to train, validate, and test real-time subway carriage congestion predictions, with GPU used to speed up the training process. All deep learning networks were trained with 200 epochs, a batch size of 10. In addition, a reduced learning rate callback (ReduceLROnPlateau) is used to dynamically adjust the learning rate.

The predictive performance of all deep learning models is presented in Table 3. Additionally, a baseline is provided for comparing the scenarios with and without the application of deep learning models by using a Naïve forecast, where the next time point was predicted using the previous value. The results indicate that the LSTM model achieves the highest predictive performance among the RNN series models. The Conv-LSTM model outperforms the other models. The predictive performance is low for stations in Cluster 1. This low performance is attributable to the fact that compared with the stations in other clusters, these stations have lower pedestrian volumes, and thus the RNN series models cannot effectively capture their time-series patterns. The predictive performance is adequate for the remaining clusters with distinct time-series patterns. In addition, the predictive performance of spatiotemporal GNN models such as STGCN and GWNN is also investigated to verify the overall performance of RNN-based deep learning models. The analysis results show that while the performance metrics of spatiotemporal GNN models were generally better than those of single RNN models, they were inferior to those of the Conv-LSTM model. It appears that while the spatiotemporal GNN models can handle time-series data, it struggle to effectively capture sequential characteristics as well as the Conv-LSTM, making it difficult to model time-series patterns accurately. Additionally, GNN-based models involve complex node-edge graph computations (such as those for road networks), which makes them relatively less effective than Conv-LSTM in predicting the congestion of subway cars on a single line used in this study.

Model		Cluster	RMSE	SMAPE	Correlation coefficient	MDAPE
baseline model	Naïve forecast	cluster1	0.315	48.92%	0.358	35.73%
		cluster2	0.221	31.93%	0.460	22.40%
		cluster3	0.120	28.67%	0.526	21.34%
		cluster4	0.240	35.73%	0.426	24.73%
		Total	0.215	35.18%	0.461	25.57%
RNN-based model	RNN	cluster1	0.239	45.28%	0.133	32.30%
		cluster2	0.139	26.88%	0.444	19.19%
		cluster3	0.089	21.62%	0.714	15.55%
		cluster4	0.141	33.38%	0.239	23.43%
		Total	0.143	33.64%	0.516	21.93%
	LSTM	cluster1	0.057	39.27%	0.503	28.44%
		cluster2	0.103	24.88%	0.593	16.73%
		cluster3	0.087	21.49%	0.721	14.73%
		cluster4	0.066	26.16%	0.561	17.90%
		Total	0.075	28.95%	0.807	19.78%
	GRU	cluster1	0.084	40.24%	0.354	29.29%
		cluster2	0.111	25.96%	0.583	18.38%
		cluster3	0.090	21.93%	0.696	15.67%
		cluster4	0.087	30.01%	0.395	22.30%
		Total	0.087	30.86%	0.757	20.46%
	Conv-LSTM	cluster1	0.026	26.39%	0.704	18.15%
		cluster2	0.032	5.06%	0.897	2.40%
		cluster3	0.012	4.43%	0.956	1.79%
		cluster4	0.012	5.18%	0.967	3.10%
		Total	0.017	10.89%	0.968	6.47%
GNN-based model	STGCN	cluster1	0.063	33.06%	0.662	24.56%
		cluster2	0.074	15.19%	0.722	10.01%
		cluster3	0.049	11.72%	0.761	7.34%
		cluster4	0.037	17.85%	0.757	12.81%
		Total	0.054	21.17%	0.843	14.06%
	GWNN	cluster1	0.065	36.72%	0.632	26.11%
		cluster2	0.078	15.51%	0.662	10.26%
		cluster3	0.051	12.55%	0.748	7.72%
		cluster4	0.042	19.02%	0.680	12.96%
		Total	0.058	22.81%	0.816	14.44%

Table 3: Comparison of predictive performances.

To rigorously evaluate the predictive performance of the Conv-LSTM model, the Diebold-Mariano (DM) test is conducted. This test is particularly effective for comparing the accuracy of two competing forecasting models and is widely recognized for its robustness in handling different types of forecast errors (Diebold & Mariano, 2002). The Conv-LSTM model was compared with several baseline models, including RNN, LSTM, GRU, STGCN, and GWNN. The DM test was applied to the forecast errors from each pair of models, using the MSE as the loss function. The DM test statistic was calculated for each model pair, and the corresponding p-values were obtained to assess the statistical significance of the differences in predictive accuracy. The results are summarized in Table 4. The DM test results indicate that the Conv-LSTM model significantly outperforms all other models, in terms of forecast accuracy of the total dataset, with p-values well below the 1% significance level. This demonstrates the superior ability of the Conv-LSTM model to capture simple spatial patterns of the subway network, leading to more accurate predictions. These findings are consistent with previous studies that highlight the advantages of combining convolutional operations with LSTM architectures for spatiotemporal data.

	RNN		LSTM		GRU		STGCN		GWNN	
	DM	P-value	DM	P-value	DM	P-value	DM	P-value	DM	P-value
cluster1	-184.683	0.003	-117.361	0.006	-141.779	0.004	-80.869	0.009	-89.025	0.007
cluster2	-10.717	0.059	-34.699	0.023	-45.263	0.015	-12.795	0.050	-15.401	0.041
cluster3	-92.258	0.007	-61.557	0.012	-83.537	0.008	-18.789	0.036	-23.519	0.034
cluster4	-34.604	0.023	-5.400	0.117	-8.152	0.083	-29.494	0.032	-35.601	0.021
Total	-114.236	0.006	-81.436	0.009	-94.224	0.007	-29.722	0.031	-34.237	0.024

Table 4: Summary of DM test.

6. Conclusion

The objective of this study is to predict passenger congestion levels in individual carriages in subway systems using accessibility and time-varying patronage demand data for subway platforms. A bi-level deep-learning architecture that incorporates unsupervised and supervised learning methods is used to enhance predictive performance while considering the spatiotemporal characteristics of each station.

Accurate prediction of passenger congestion levels is important for subway users and operators. The availability of such information can potentially influence user behavior by allowing users to check the congestion level in each carriage at each station in their regular travel route and select a boarding space that suits their preferences. This information can also allow users with flexible preferences to occupy available boarding spaces, thereby easing congestion levels at particular boarding locations. Moreover, accurately predicting passenger congestion can allow operators to prepare counterstrategies in advance for scenarios in which excessive imbalances are expected.

In the real-world analysis, the Conv-LSTM model outperforms other methods. This study highlights the feasibility of predicting subway congestion using congestion-related variables, without measuring actual congestion values. Moreover, the clusters with clear time-series patterns correspond to high predictive performances. The predictive performance for stations belonging to Cluster 1 (which has fewer pedestrians than stations in other clusters) can be improved if the time-series pattern is clarified.

Overall, a novel method is devised for predicting subway congestion that can be applied in real-world situations. This method uses accessibility and time-varying patronage demand data for subway platforms and adopts a bi-level deep-learning architecture that incorporates unsupervised and supervised learning methods while considering the spatiotemporal characteristics of passenger congestion in each carriage in each station. The results suggest that the method can yield information that would be valuable to subway users and operators. This study highlights the feasibility of predicting subway congestion using congestion-related variables without measuring actual congestion.

Acknowledgement

The first author was supported by the Korea Ministry of Land, Infrastructure, and Transport (MOLIT) as an Innovative Talent Education Program for Smart City. The Second author was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022R1A6A3A01087328). The fourth author was supported by the Guangdong - Hong Kong - Macau Joint Laboratory Program of the 2020 Guangdong New Innovative Strategic Research Fund, Guangdong Science and Technology Department (Project No.: 2020B1212030009) and Francis S Y Bong Professorship in Engineering from the University of Hong Kong. The fifth author was supported by the 2023 Research Fund from the University of Seoul. The last author was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2019K1A4A7A03112460).

References

Antoniou, C., Koutsopoulos, H. N., & Yannis, G. (2013). Dynamic data-driven local traffic state estimation and prediction. *Transportation Research Part C: Emerging Technologies*, *34*, 89–107.

Batty, M. (2016). Big data and the city. *Built Environment*, *42*, 321–337.
Bogaerts, T., Masegosa, A. D., Angarita-Zapata, J. S., Onieva, E., & Hellinckx, P. (2020). A graph cnn-lstm neural network for short and long-term traffic forecasting based on trajectory data. *Transportation Research Part C: Emerging Technologies*, *112*, 62–77.
Chen, C., Ma, J., Susilo, Y., Liu, Y., & Wang, M. (2016). The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation research part C: emerging technologies*, *68*, 285–299.
Chen, W., Li, Z., Liu, C., & Ai, Y. (2021). A deep learning model with conv-lstm networks for subway passenger congestion delay prediction. *Journal of Advanced Transportation*, *2021*, 1–10.
Chen, Y., Zhu, L., Gonder, J., Young, S., & Walkowicz, K. (2017). Data-driven fuel consumption estimation: A multivariate adaptive regression spline approach. *Transportation Research Part C: Emerging Technologies*, *83*, 134–145.
Chiabaut, N., & Faitout, R. (2021). Traffic congestion and travel time prediction based on historical congestion maps and identification of consensual days. *Transportation Research Part C: Emerging Technologies*, *124*, 102920.
Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, .
Dabiri, S., & Heaslip, K. (2018). Inferring transportation modes from gps trajectories using a convolutional neural network. *Transportation research part C: emerging technologies*, *86*, 360–371.
Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, *20*, 134–144.
Ding, H., Di, Y., Zheng, X., Liu, K., Zhang, W., & Zheng, L. (2021). Passenger arrival distribution model and riding guidance on an urban rail transit platform. *Physica A: Statistical Mechanics and its Applications*, *571*, 125847.
Fang, J., Fujiyama, T., & Wong, H. (2019). Modelling passenger distribution on metro platforms based on passengers' choices for boarding cars. *Transportation Planning and Technology*, *42*, 442–458.
Gerum, P. C. L., Altay, A., & Baykal-Gürsoy, M. (2019). Data-driven predictive maintenance scheduling policies for railways. *Transportation Research Part C: Emerging Technologies*, *107*, 137–154.
Ham, S. W., Cho, J.-H., Park, S., & Kim, D.-K. (2021). Spatiotemporal demand prediction model for e-scooter sharing services with latent feature and deep learning. *Transportation research record*, *2675*, 34–43.
Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*, 1735–1780.
Jiang, Z., Fan, W., Liu, W., Zhu, B., & Gu, J. (2018). Reinforcement learning approach for coordinated passenger inflow control of urban rail transit in peak hours. *Transportation Research Part C: Emerging Technologies*, *88*, 1–16.
Kim, K., Kim, D., Noh, K.-S., & Lee, J.-Y. (2015). An exploratory study on improvement method of the subway congestion based big data convergence. *Journal of Digital Convergence*, *13*, 35–42.
Ku, D., Na, S., Kim, J., & Lee, S. (2021a). Real-time taxi demand prediction using recurrent neural network. In *Proceedings of the Institution of Civil Engineers-Municipal Engineer* (pp. 75–87). Thomas Telford Ltd volume 174.
Ku, D., Yeon, C., Lee, S., Lee, K., Hwang, K., Li, Y. C., & Wong, S. C. (2021b). Safe traveling in public transport amid covid-19. *Science advances*, *7*, eabg3691.
Kuipers, R. A., Palmqvist, C.-W., Olsson, N. O., & Winslott Hiselius, L. (2021). The passenger's influence on dwell times at station platforms: a literature review. *Transport Reviews*, *41*, 721–741.
LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, *1*, 541–551.
Lee, J., Yoo, S., Kim, H., & Chung, Y. (2018). The spatial and temporal variation in passenger service rate and its impact on train dwell time: A time-series clustering approach using dynamic time warping. *International Journal of Sustainable Transportation*, *12*, 725–736.
Lee, S., Ngoduy, D., & Keyvan-Ekbatani, M. (2019). Integrated deep learning and stochastic car-following model for traffic dynamics on multi-lane freeways. *Transportation research part C: emerging technologies*, *106*, 360–377.
Li, C.-Y., Yang, R.-Y., & Xu, G.-m. (2019). Impacts of group behavior on boarding process at the platform of high speed railway station. *Physica A: Statistical Mechanics and its Applications*, *535*, 122247.
Li, Y., Wang, X., Sun, S., Ma, X., & Lu, G. (2017). Forecasting short-term subway passenger flow under special events scenarios using multiscale radial basis function networks. *Transportation Research Part C: Emerging Technologies*, *77*, 306–328.
Liu, Y., Liu, Z., & Jia, R. (2019). Deeppf: A deep learning based architecture for metro passenger flow prediction. *Transportation Research Part C: Emerging Technologies*, *101*, 18–34.
Liu, Y., Zheng, H., Feng, X., & Chen, Z. (2017). Short-term traffic flow prediction with conv-lstm. In *2017 9th International Conference on Wireless Communications and Signal Processing (WCSP)* (pp. 1–6). IEEE.
Liu, Z., Li, D., & Wang, X. (2016). Passenger distribution and waiting position selection model on metro station platform. In *2016 International Conference on Civil, Structure and Environmental Engineering* (pp. 35–41). Atlantis Press.
Mohanty, S., Pozdnukhov, A., & Cassidy, M. (2020). Region-wide congestion prediction and control using deep learning. *Transportation Research Part C: Emerging Technologies*, *116*, 102624.
Myung, J., Kim, D.-K., Kho, S.-Y., & Park, C.-H. (2011). Travel time prediction using k nearest neighbor method with combined data from vehicle detector system and automatic toll collection system. *Transportation Research Record*, *2256*, 51–59.
Nasri, A., & Zhang, L. (2014). The analysis of transit-oriented development (tod) in washington, dc and baltimore metropolitan areas. *Transport policy*, *32*, 172–179.
Oliveira, L. C., Fox, C., Birrell, S., & Cain, R. (2019). Analysing passengers' behaviours when boarding trains to improve rail infrastructure and technology. *Robotics and Computer-Integrated Manufacturing*, *57*, 282–291.
Patro, S., & Sahu, K. K. (2015). Normalization: A preprocessing stage. *arXiv preprint arXiv:1503.06462*, .
Qu, J., Xiao, M., Yang, L., & Xie, W. (2023). Flight delay regression prediction model based on att-conv-lstm. *Entropy*, *25*,

770.
Rashidi, S., Ranjitkar, P., & Hadas, Y. (2014). Modeling bus dwell time with decision tree-based methods. *Transportation Research Record*, *2418*, 74–83.
Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, *65*, 386.
Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, *323*, 533–536.
Sharma, A., McCloskey, B., Hui, D. S., Rambia, A., Zumla, A., Traore, T., Shafi, S., El-Kafrawy, S. A., Azhar, E. I., Zumla, A. et al. (2023). Global mass gathering events and deaths due to crowd surge, stampedes, crush and physical injuries-lessons from the seoul halloween and other disasters. *Travel medicine and infectious disease*, *52*.
Shi, B., Bai, X., & Yao, C. (2016). An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE transactions on pattern analysis and machine intelligence*, *39*, 2298–2304.
Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., & Woo, W.-c. (2015). Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, *28*, 1–9.
Tang, Y., Yang, H., Wang, B., Huang, J., & Bai, Y. (2020). A pareto-improving and revenue-neutral scheme to manage mass transit congestion with heterogeneous commuters. *Transportation Research Part C: Emerging Technologies*, *113*, 245–259.
Wang, Z., Liang, M., & Delahaye, D. (2018). A hybrid machine learning model for short-term estimated time of arrival prediction in terminal manoeuvring area. *Transportation Research Part C: Emerging Technologies*, *95*, 280–294.
Wei, M., Corcoran, J., Sigler, T., & Liu, Y. (2018). Modeling the influence of weather on transit ridership: A case study from brisbane, australia. *Transportation Research Record*, *2672*, 505–510.
Wei, Y., & Chen, M.-C. (2012). Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. *Transportation Research Part C: Emerging Technologies*, *21*, 148–162.
Wu, J., Liu, M., Sun, H., Li, T., Gao, Z., & Wang, D. Z. (2015). Equity-based timetable synchronization optimization in urban subway network. *Transportation Research Part C: Emerging Technologies*, *51*, 1–18.
Yang, J., Chen, S., Qin, P., Lu, F., & Liu, A. A. (2018). The effect of subway expansions on vehicle congestion: Evidence from beijing. *Journal of Environmental Economics and Management*, *88*, 114–133.
Yang, Y., Liu, J., Shang, P., Xu, X., & Chen, X. (2020). Dynamic origin-destination matrix estimation based on urban rail transit afc data: deep optimization framework with forward passing and backpropagation techniques. *Journal of Advanced Transportation*, *2020*, 1–16.
Yao, E., Hong, J., Pan, L., Li, B., Yang, Y., & Guo, D. (2021). Forecasting passenger flow distribution on holidays for urban rail transit based on destination choice behavior analysis. *Journal of Advanced Transportation*, *2021*, 1–13.
Yu, B., Yin, H., & Zhu, Z. (2017). Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. *arXiv preprint arXiv:1709.04875*, .
Zhang, L., & Ma, J. (2023). A spatiotemporal graph wavelet neural network for traffic flow prediction. *Journal of Information and Intelligence*, .
Zhang, T., Li, D., & Qiao, Y. (2018). Comprehensive optimization of urban rail transit timetable by minimizing total travel times under time-dependent passenger demand and congested conditions. *Applied Mathematical Modelling*, *58*, 421–446.
Zheng, H., Lin, F., Feng, X., & Chen, Y. (2020). A hybrid deep learning model with attention-based conv-lstm networks for short-term traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, *22*, 6910–6920.
Zhu, Y., Koutsopoulos, H. N., & Wilson, N. H. (2018). Inferring left behind passengers in congested metro systems from automated data. *Transportation Research Part C: Emerging Technologies*, *94*, 323–337.
Zhu, Z., He, Y., Guo, X., Zhang, Y., & Chen, J. (2021). Improved huff model for estimating urban rail transit station catchment areas considering station choices. *Journal of Advanced Transportation*, *2021*, 1–11.
Zou, Q., Yao, X., Zhao, P., Dou, F., & Yang, T. (2018). Managing recurrent congestion of subway network in peak hours with station inflow control. *Journal of Advanced Transportation*, *2018*, 1–16.