Short-Term Passenger Flow Prediction in Urban Rail Transit based on Hybrid Deep Learning Models

Yayi Kang *

Department of Transportation, Nanjing University of Aeronautics and Astronautics, Nanjing, 210000, China

* Corresponding Author Email: yayi.kang@outlook.com

Abstract. Urban rail transit systems are essential for efficient, reliable, and environmentally friendly transportation in modern cities. Correct and effective short-term passenger flow prediction is crucial for optimizing operational efficiency, enhancing service quality, and ensuring passenger safety. Traditional prediction methods often fail to capture the complex, non-linear, and dynamic patterns in urban rail transit systems. This study uses a hybrid deep learning model combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to predict short-term passenger flow in urban rail transit. By integrating the strengths of CNN in capturing spatial features and LSTM in modeling temporal dependencies, the hybrid model aims to improve prediction accuracy. Using historical data from Hangzhou Metro, the study demonstrates the model's effectiveness in predicting passenger flow, which reasonably has a good fit. Additionally, models with different convolutional layers show different performances. These improved predictions offer valuable insights for transit authorities, enabling them to make more informed decisions regarding train scheduling, resource allocation, and emergency response planning. By anticipating passenger demand more accurately, authorities can optimize the deployment of trains, reduce waiting times, enhance passenger comfort, and improve overall service reliability.

Keywords: CNN-LSTM model; urban rail transit; passenger flow prediction.

1. Introduction

Urban rail transit systems are critical components of modern metropolitan transportation networks, providing efficient, reliable, and environmentally friendly means of mass transit. With the rapid urbanization and continuous growth of urban populations, accurately predicting short-term passenger flow has become essential for optimizing operational efficiency, enhancing service quality, and ensuring passenger safety. Besides, as the need for using new lines in various cities is constantly growing, it is inevitable that already used rail stations have variation in the passenger flow, so the passenger flow changes of existing rail stations have become one of the important contents of demand forecasting [1]. The passenger flow distribution in metro networks is imbalanced, especially during peak hours [2]. Effective passenger flow prediction enables transit authorities to make informed decisions regarding train scheduling, resource allocation, and emergency response planning.

Passenger flow monitoring is primarily done through video surveillance to identify and manage crowd states in real time. However, this approach makes it difficult to predict flow accurately and take timely measures. Zhou et al. in their study on forecasting short-term passenger flow in a multi-level rail transit network highlight the superior accuracy of short-term predictions compared to longer-term ones, due to the reduced variability and higher reliability of immediate data [3]. Therefore, rapid and accurate prediction based on time-series data is crucial [4]. In early stages, many prediction of passenger flow used statistical methods. Ni et al. proposed a mixed seasonal integrated autoregressive moving average model for predicting the short-term passenger flow based on the data of the New York metro [5]. Parametric models are accurate at predicting stationary passenger flows. The shortages are that traditional passenger flow prediction methods, such as statistical models and classical machine learning approaches, often struggle to capture the complex, non-linear, and dynamic patterns inherent in urban rail transit systems. These methods may fail to adequately address the influence of various factors such as time of day, weather, special events, and station-specific characteristics. Consequently, there is a pressing need for more sophisticated models that can accurately capture and predict flow dynamics in real time.

In recent years, deep learning models have demonstrated remarkable success in various domains, including image recognition, natural language processing, and time series forecasting. For instance, the convolutional neural network (CNN), which has fixed-scale input and can eliminate a large number of similar unimportant pixel-neuron connections, is widely used in the field of image recognition, while Recurrent Neural Networks (RNNs) and Long Short-Term Memory networks (LSTMs) are proficient in modeling temporal dependencies [6]. For addressing the imbalanced data set problem, Ma et al. did the research in which the LSTM network was used to forecast both normal and emergency passenger flow of metro transportation [7]. A novel model called the multi-feature fusion graph convolutional network (MFGCN) was proposed and applied by Wu et al. on the experiments on Nanning Metro Line 1 passenger flow datasets [8]. The LSTM model with wavelet denoising was proved accurate by Zhao et al. in demonstrating applicability for short-term rail transit forecasting and practical significance [9].

Specifically, hybrid deep learning models, which combine multiple neural network architectures, have shown significant potential in addressing complex prediction tasks by leveraging the strengths of different network types. The hybrid model's ability to capture intricate patterns in data may lead to more accurate prediction, such as better performances in handling complex data, feature extraction and temporal modeling. Han et al. proposed a hybrid, optimized LSTM network based on Nesterov accelerated adaptive moment estimation (Nadam) and the stochastic gradient descent algorithm (SGD) [10]. Wang et al. compared the prediction result of the single Kernel Extreme Learning Machine (KELM) method, and the hybrid model based on Wavelet Transform (WT) and Back Propagation (BP). The result showed that the W-KELM model had good prediction accuracy [11]. This study applies a hybrid deep learning model for short-term passenger flow prediction in urban rail transit systems, integrating CNN and LSTM to exploit both spatial and temporal features of passenger flow data by using the database of Hangzhou subway traffic. By harnessing the hybrid deep learning models, this research aims to advance the state-of-the-art in passenger flow prediction, providing urban rail transit authorities with a reliable tool to manage and optimize their transit operations effectively.

2. Methodology

2.1. Data Source

The original data used in this paper is the historical data of Hangzhou Metro from January 1, 2019, to January 15, 2019, published in the Tianchi competition. The original data provides the information on the passengers of 81 stations of Hangzhou Metro throughout the day, including passengers' swiping records, swiping time, boarding status, boarding stations, and consumption records for three subway lines. The data contains records from 3 subway lines and 81 stations. There are over 100 million card swiping records. This study selects the station with the highest passenger flow and uses its 15 days' data for short-term prediction. In order to make sure that these data can be applied in the research, it is necessary to process the original data. Thus, the efficiencies can be improved and more reliable results can be obtained.

2.2. Data Process

The smart card data can accurately and comprehensively record the time of entering and leaving for each passenger, thus the authority of data can be guaranteed. Table 1 shows the full names, interpretations and examples of the three variables that are crucial used in the research after statistically analyzing traffic every five minutes.

Table 1. Explanation of variables

Full Name	Explanation	Example
Time	The time for passengers to swipe their cards	2019-1-1 02:00
stationID	Subway Station Number	15
Passenger Count	Passenger flow within 5 minutes	68

2.3. Model Architecture Design

2.3.1. CNN layers

A Convolutional Neural Network (CNN) is a specialized type of deep learning algorithm mainly designed for tasks that necessitate object recognition, including image classification, detection, and segmentation. CNNs are employed in a variety of practical scenarios, CNNs are primarily used for their ability to automatically and efficiently extract spatial features from raw data. The schematic diagram of CNN structure is shown in Figure 1. They are particularly effective in identifying local patterns through convolutional operations.

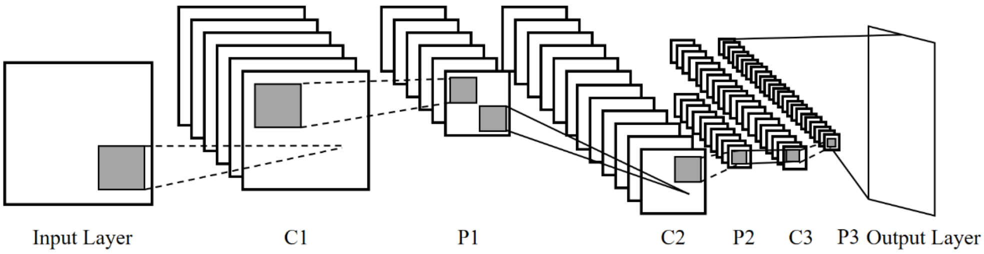

Fig. 1 The Schematic diagram of CNN structure (Picture credit: Original)

2.3.2. LSTM

LSTMs are designed to capture long-term dependencies and temporal patterns in sequential data. They are effective in dealing with time-series data where the order of events is crucial. LSTM cells are capable of maintaining and updating long-term state information through their gating mechanisms (input gate, forget gate, and output gate). This enables the model to learn temporal dependencies in the passenger flow data. The working principle of LSTM is shown in Figure 2.

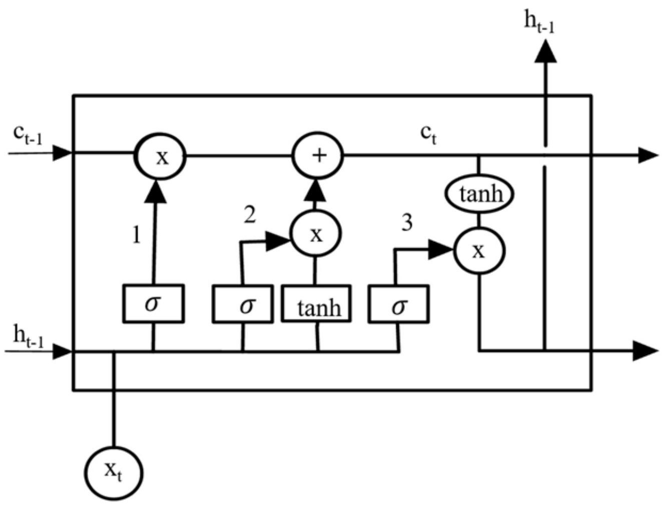

Fig. 2 The working principle of LSTM

2.3.3. CNN-LSTM

The proposed CNN-LSTM model leverages the strengths of both CNN and LSTM networks to effectively capture spatial and temporal features in the passenger flow data. Each layer in the CNN-LSTM model contributes to understanding different aspects of the data. Convolutional layers highlight spatial dependencies, while LSTM layers emphasize temporal relationships. This layer-wise decomposition can aid in model interpretability and debugging. By stacking multiple layers, CNNs and LSTMs can learn hierarchical features at different levels of abstraction, improving the model's ability to make accurate predictions. The architecture is highly adaptable and can be fine-tuned for different contexts and datasets. It can be applied to various urban rail transit systems with different passenger flow characteristics.

3. Results and Discussion

3.1. Model Sizing

In this study, a CNN-LSTM model is employed to predict short-term passenger flow for metro stations over a 15-day period, using historical data from January 1, 2019, to January 15, 2019. This study first compares the average passenger flow of 81 stations over the past 15 days (Figure 3). The data was filtered to focus on a single station (StationID:15) with the highest average daily passenger flow, identified through preliminary analysis. The time series data was then resampled into 5-minute intervals and normalized to facilitate model training.

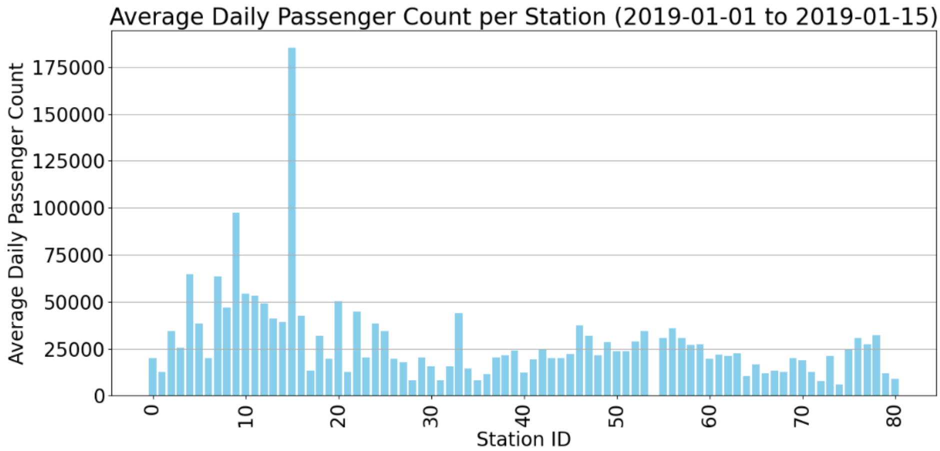

Fig. 3 Average Daily Passenger Count Station (2019-01-02 to 2019-01-15)

After processing the card swiping records of the selected station every five minutes for 15 days, the result reveals that the figure illustrates both single-peaked and double-peaked patterns. The passenger counts for station 15 is shown in the Figure 4. This is because the time can influence people's travel patterns heavily.

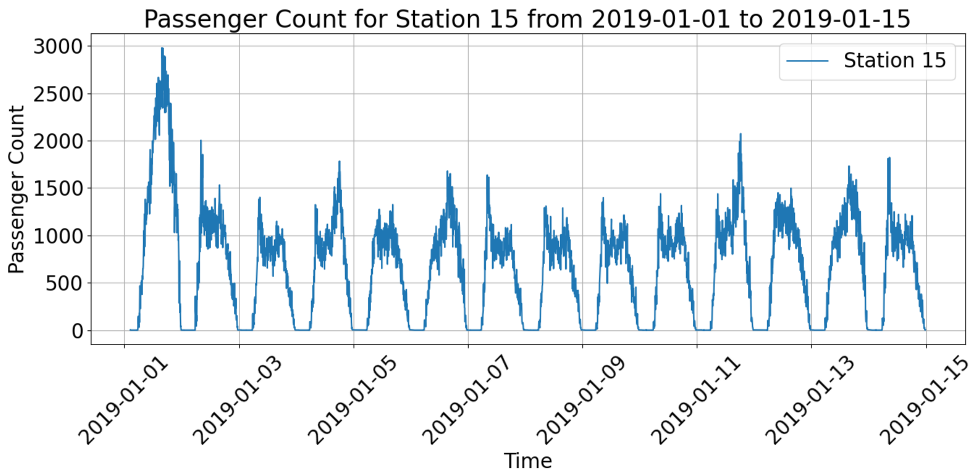

Fig. 4 Passenger Count Station 15 from 2019-01-01 to 2019-01-15

Here, two typical data with different features from January 1st, 2019 and January 4th, 2019 were selected for detailed display (Figure 5, Figure 6). Figure 5 shows a clear unimodal pattern, which happens to be New Year's Day in China. Events such as sports games, concerts, or holidays can lead to temporary spikes in passenger flow, contributing to single or double peak patterns. Seasonal changes, such as school vacations or public holidays, can also impact daily travel patterns. Differently, metro stations located in commercial areas often see double peaks during weekdays, aligning with work hours. Stations in residential areas might also exhibit this pattern, as residents commute to work in the morning and return in the evening.

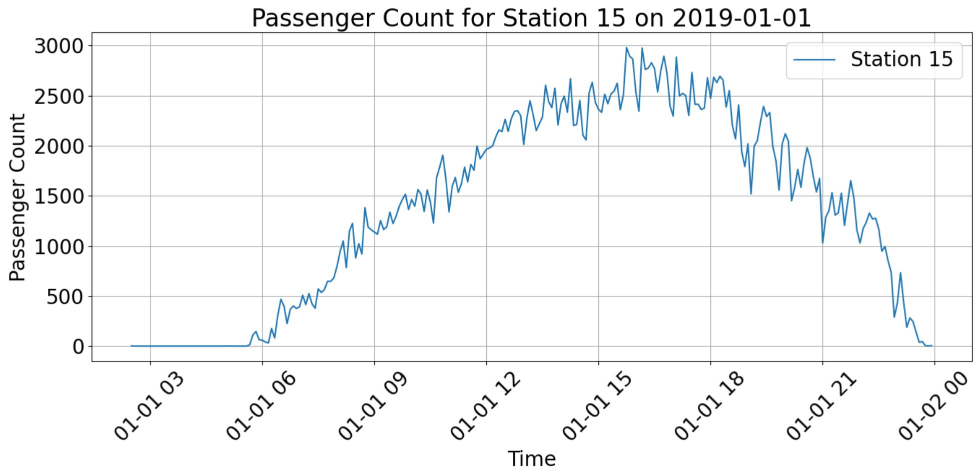

Fig. 5 Passenger Count for Station 15 on 2019-01-01

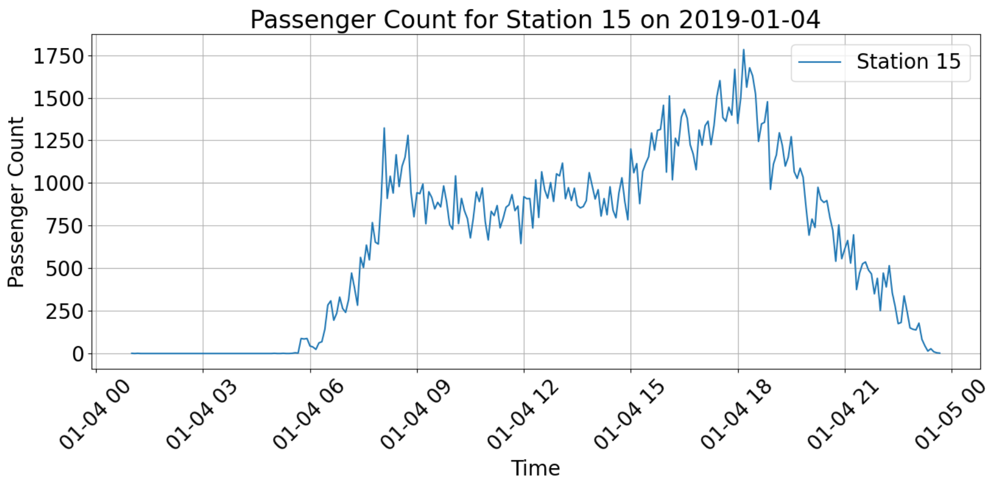

Fig. 6 Passenger Count for Station 15 on 2019-01-04

3.2. Prediction Results

The CNN-LSTM model was designed to capture both spatial and temporal features of the passenger flow data. The convolutional layers extracted spatial patterns, while the LSTM layers handled temporal dependencies. The study splits the data into training and testing sets, with the first 80% used for training and the remaining 20% for testing.

The trained model is applied to predict passenger flow for the last three days (January 13 to January 15, 2019). The predicted values are then compared to the actual passenger counts, and the results showed a reasonably good fit, demonstrating the model's capability to catch the potential trends in the data. The predicted and actual passenger counts for the testing period were plotted together. The Figure 7 clearly showed the alignment between the predicted results and the actual results, with the predicted curve closely following the actual passenger flow trend. This provided a visual confirmation of the model's predictive performance.

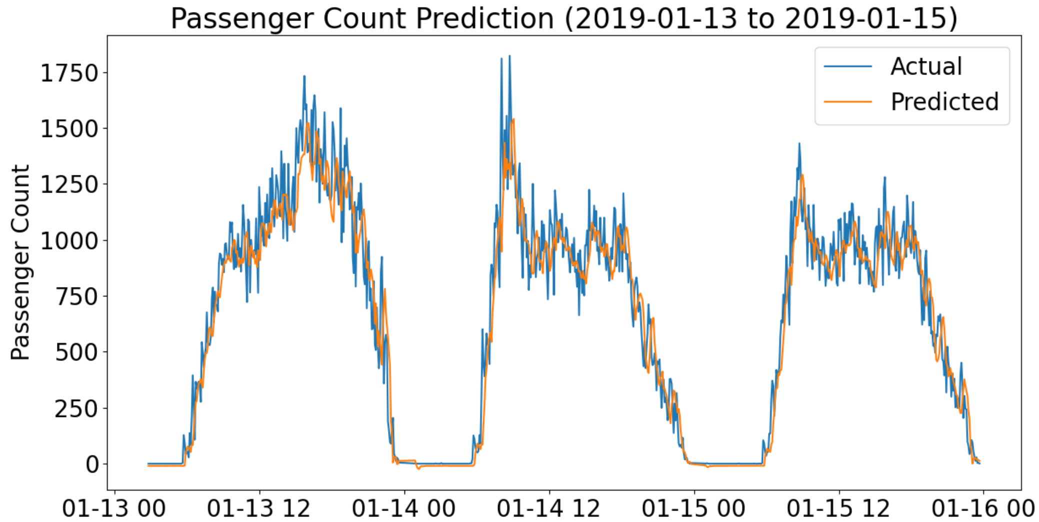

Fig. 7 Passenger Count Prediction (2019-01-13 to 2019-01-15)

The given results are given by using one-layer convolutional layer. Due to the purpose of finding a suitable model for predicting, the parameter of the number of layers has been changed to compare the fitting effect with the original results. Here, the number of convolutional layers is increased to two. The comparison chart of two prediction results is shown in Figure 8.

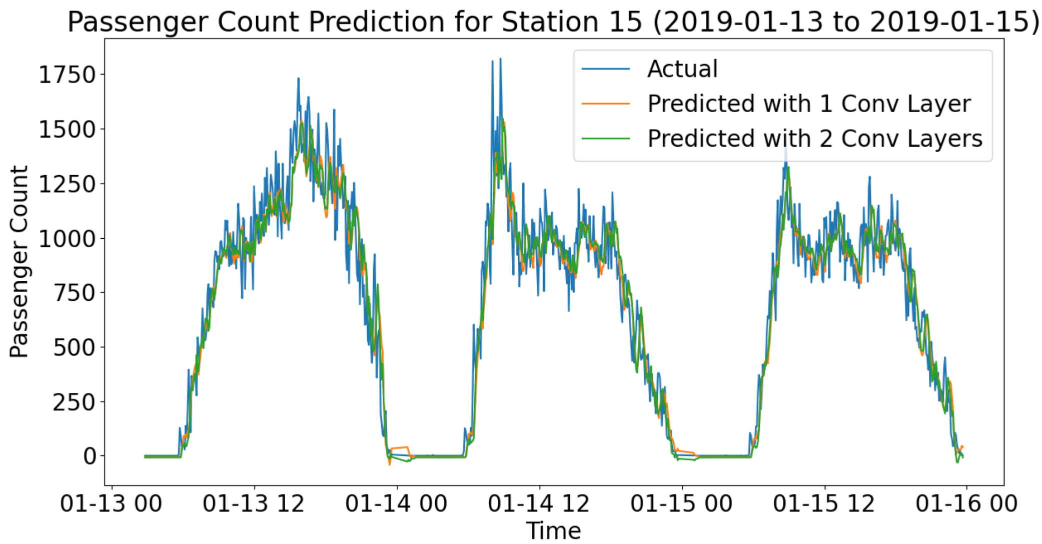

Fig. 8 Comparison of Two Prediction

To better analyze the degree of fitting in the form of data, this research uses RMSE (root-mean-square error) to evaluate. Smaller value means better prediction. The evaluation of the two models, each with a different number of convolutional layers, provides insights into their respective performance in predicting passenger flow. Specifically, the RMSE value of the model with 2 convolutional layers is 129.108, while the RMSE value of the model with 1 convolutional layer is 131.109.

The model with two convolutional layers achieved a slightly lower RMSD compared to the model with one convolutional layer, indicating a marginal improvement in predictive accuracy. This suggests that the additional layer helped in capturing more nuanced patterns in the data. While the two-layer model has slightly better performance, the improvement is not substantial. This highlights the importance of balancing model complexity with the actual performance gains. In many practical scenarios, the simpler one-layer model might be preferred due to its lower computational requirements and faster training time.

The CNN-LSTM model performed well in predicting short-term passenger flow, as evidenced by the close alignment between the predicted and actual passenger counts. The combination of CNN and LSTM layers allowed the model to effectively capture both spatial and temporal patterns in the data. The convolutional layers were particularly useful in identifying local patterns in the time series data, while the LSTM layers captured the sequential dependencies.

The prediction results highlighted several key patterns. For instance, the model accurately predicted peak hours and lower traffic periods, which are critical for operational planning in metro systems. The ability to forecast such fluctuations in passenger flow can help in resource allocation, scheduling, and managing peak loads more efficiently.

Despite the promising results, there were instances where the model's predictions deviated from the actual values. These deviations could be attributed to various factors, such as anomalies in the data, external events affecting passenger flow, or limitations in the model's ability to generalize beyond the training data.

4. Conclusion

The CNN-LSTM model effectively captures both spatial and temporal features in passenger flow, demonstrating superior performance in short-term prediction tasks for urban rail transit. The model was tested on historical data from Hangzhou Metro, showing a close alignment between predicted and actual passenger counts, particularly during peak hours and lower traffic periods. This improved predictive accuracy can aid urban rail transit authorities in optimizing train schedules, allocating resources efficiently, and managing peak loads. The study also found that adding an extra convolutional layer marginally improved prediction accuracy. Meanwhile, attention should be paid to the potential computational complexity and overfitting issues caused by multi-layer models. Despite its promising results, the model's occasional deviations suggest a need for further enhancements, such as incorporating additional features like weather conditions and special events. Future work should focus on refining the model to better generalize beyond the training data and address external factors influencing passenger flow.

References

[1] Wu Qianqian, Sun Lin, Jia Furong, Liu Yi, Zhan Zichen. Research on Dynamic Passenger Flow Distribution of Rail Transit Based on Multi-Dimensional Euclidean Distance. In Advances in Smart Vehicular Technology, Transportation, Communication and Applications, 2022.

[2] Zhao Shuzhi, Tonghe Ni, Yang Wang, Xiang-Tao Gao. A New Approach to the Prediction of Passenger Flow in a Transit System. Computers & Mathematics with Applications, 2011.

[3] Zhou Wenzhong, Gao Chunhai, Tang Tao. Parallel Interactive Attention Network for Short-Term Origin–Destination Prediction in Urban Rail Transit. Applied Sciences, 2024.

[4] Tan Yifan, Liu Haixu, Pu Yun, Wu Xuemei, Jiao Yubo. Passenger Flow Prediction of Integrated Passenger Terminal Based on K-Means–GRNN. Journal of Advanced Transportation, 2021.

[5] Ming Ni, He Qing, Gao Jing. Forecasting the Subway Passenger Flow Under Event Occurrences with Social Media. IEEE Transactions on Intelligent Transportation Systems, 2017.

[6] Wang Jiaxuan, Wang Rui, Zeng Xin. Short-term Passenger Flow Forecasting using CEEMDAN Meshed CNN-LSTM-attention Model Under Wireless Sensor Network. IET Communications, 2022.

[7] Ma Jingye,Zeng Xin, Xue Xiaoping, Deng Ranran. Metro Emergency Passenger Flow Prediction on Transfer Learning and LSTM Model. Applied Sciences, 2022.

[8] Wu Jinxin, Li Xianwang, He Deqiang, Li Qin, Xiang Weibin. Learning Spatial-Temporal Dynamics and Interactivity for Short-Term Passenger Flow Prediction in Urban Rail Transit. Applied Intelligence (Dordrecht, Netherlands), 2023.

[9] Zhao Qingliang, Feng Xiaobin, Zhang Liwen, Wang Yiduo. Research on Short-Term Passenger Flow Prediction of LSTM Rail Transit Based on Wavelet Denoising. Mathematics (Basel), 2023.

[10] Han Yong, Wang Cheng, Ren Yibin, Wang Shukang, Zheng Huangcheng, Chen Ge. Short-Term Prediction of Bus Passenger Flow Based on a Hybrid Optimized LSTM Network. ISPRS International Journal of Geo-Information, 2019.

[11] Liu Ruijian, Wang Yuhan, Zhou Hong, Qian Zeqiang. Short-Term Passenger Flow Prediction Based on Wavelet Transform and Kernel Extreme Learning Machine. IEEE Access, 2019.