

# Research on Forecasting Short-term Passenger Flow of Subway Based on Deep Neural Network

Jinbing Ha, Zhengguang Zhao \*

School of Economics and Management, Nanjing University of Science and Technology, Nanjing, China

\* Corresponding author: 11758751378@qq.com

**Abstract.** With the rapid development of urbanization, a large number of people have poured into cities, which makes the traffic congestion problem increasingly serious. As a stable, efficient, safe and environmentally friendly public transport, subway has gradually become the first choice for residents to travel by bus. The rapid increase in the number of passengers has caused great pressure on the quality of subway operation and service. Therefore, how to accurately predict the passenger flow of the station at a certain time in the future and let the subway operation department make the corresponding operation plan in advance according to the predicted passenger flow is of great significance to improve the service quality of subway operation. Under this background, this paper summarizes the existing short-term passenger flow forecasting methods at home and abroad, and puts forward the application of deep neural network method in the field of short-term passenger flow forecasting of urban rail transit according to the advantages and disadvantages of different forecasting methods. Taking the short-term passenger flow forecast of Nanjing subway as an example, the original credit card data is preprocessed to obtain passenger flow data, and the temporal and spatial distribution law of historical passenger flow is studied, and the LSTM prediction model and CNN-LSTM combined prediction model are established, which improves the prediction accuracy of short-term inbound passenger flow of rail transit.

**Keywords:** Short-term passenger flow forecast, Convolutional neural network, long short-term memory network, Combined model.

## 1. Introduction

In recent years, the rapid development of urbanization has gradually expanded the scale of the city, and a large number of people have poured into the city. At the same time, the rapid increase in the number of road networks and the continuous increase in the number of motor vehicles have made traffic problems one of the most important issues in urban development and management. In order to improve the problem of urban traffic congestion, the construction of urban rail transit represented by the subway came into being.

As the scale of rail transit lines continues to expand, the road network becomes more and more complex, and the problem of unpredictable traffic congestion becomes more serious. In order to alleviate and solve this congestion problem and control the uneven distribution of passenger flow, in addition to increasing efforts to lay new trunk lines, it is also necessary to plan existing lines, use historical passenger flow information, predict short-term passenger flow changes at future stations, and help choose reasonable travel routes to avoid congestion.

In a theoretical sense, the law of passenger flow is the most critical information in the operation of urban rail transit. The development of intelligent transportation provides a convenient channel for obtaining large-scale passenger flow data.

In practical applications, the results of short-term inbound passenger flow prediction can help pedestrians choose reasonable travel routes and avoid traffic congestion. The short-term prediction results of rail transit passenger flow can be used as the data support of the intelligent transportation system to help residents make plans before travel, choose travel routes reasonably, avoid traffic congestion as much as possible, and improve the safety of citizens' travel. On the other hand, it can provide early warning of passenger flow in subway stations for relevant urban transportation management departments, deploy security in advance, and help safe travel in future cities.

## **2. Related theoretical basis**

### **2.1. Recurrent neural network**

RNN, recurrent neural network, also known as recurrent neural network. The RNN neural network is an improvement to the traditional artificial neural network. The layers of the traditional neural network model are all connected, and there is no connection between the neurons of each layer, and the neurons of the hidden layer of the RNN cyclic network are also connected. Interconnected, its neurons not only receive the input information of the sample at the current moment, but also receive the output information of itself at the previous moment [1].

### **2.2. Long short-term memory network**

LSTM, long short-term memory network, is a special form of RNN network. On the basis of RNN, memory characteristics are added to maintain the long-term memory of the neural network, so that the model can also be used well for long-term sequences [2]. For RNN, since its network layer updates information without restriction, the information will become confusing and easy to disappear and change, so there may be a problem of gradient disappearance. The improved LSTM network adds forgetting units and memory units in the hidden layer. When new information is input, the LSTM network will filter out some information for retention and discarding, and save important information in long-term memory. Its internal structure is relatively complex, and information is selectively transmitted through a unique gating unit, which is a more complex cyclic network structure of neurons.

### **2.3. Convolutional neural network**

CNN, convolutional neural network, is a network model composed of convolutional layers, pooling layers and fully connected layers. Each convolution layer consists of several convolution kernels, which can be imagined as neurons in a classic neural network, except that the activation function becomes a convolution operation [3]. The convolution operation has a strict mathematical definition, and its purpose is to extract different features of the input data. Convolutional neural network has the characteristics of local connection and weight sharing, which reduces parameter calculation and speeds up training [4]. This characteristic makes it more suitable for processing two-dimensional data and is often used to process images. Then the corresponding input data is The pixel information of the image. For such input data, one layer of convolutional layer is relatively simple and basic for extracting data features. Networks with more layers are then iteratively trained to extract more complex features from the basic features, which can refine more deep-level features in the data, improve the effect of the model [5]. The CNN structure is relatively simple and can be trained using the backpropagation algorithm, making it an attractive deep neural network model.

## **3. Short-term passenger flow forecasting model based on CNN-LSTM**

### **3.1. Experimental data**

This paper takes Yuantong Subway Station as the prediction target site. It is an interchange station between Metro Line 2 and Metro Line 10. The daily passenger flow is huge and the nonlinear characteristics are obvious. Accurately predicting the short-term passenger flow at Yuantong Station has strong practical significance. When predicting short-term passenger flow, it is necessary to divide the time granularity of passenger flow. If the time granularity is too small, it will affect the accuracy of the prediction; conversely, if the time granularity is too large, the prediction results cannot reflect the short-term passenger flow variety. This article chooses to use 10 minutes as the time granularity to count passenger flow.

This section mainly uses the inbound passenger flow data with a time granularity of 10 minutes at Yuantong Station of Nanjing Line 10 from March 1 to March 14, 2021 as the research object for

simulation prediction. Among them, March 14 as a test set, it is used to ultimately evaluate the prediction effect of the model.

### 3.2. Selection of evaluation index

In order to evaluate the prediction performance of the LSTM-based short-term subway passenger flow prediction model and the CNN-LSTM-based subway short-term passenger flow prediction model, this paper selects the mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE) three error evaluation indicators to measure the prediction accuracy of the passenger flow prediction model. The error calculation formula is as follows.

$$MAE = \frac{1}{K} \sum_{i=1}^K |y_i - \hat{y}_i| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{K} \sum_{i=1}^K (y_i - \hat{y}_i)^2} \quad (2)$$

$$MAPE = \frac{1}{K} \sum_{i=1}^K \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (3)$$

The MAE indicator represents the average value of the absolute value error between the predicted value and the real value, which intuitively reflects the size of the error. The smaller the value of MAE, the closer the predicted value of the model is to the real value. The RMSE indicator represents the average degree of dispersion between the predicted value and the real value, and its indicator is sensitive to outliers. The smaller the RMSE value, the more similar the prediction result of the model to the real situation, and the better the prediction effect of the model. The MAPE index reflects the degree of deviation between the predicted value of passenger flow and the real value. MAPE is more reasonable in evaluating the passenger flow prediction of different models at different time periods and at different sites. The smaller the value of MAPE, it represents the performance of the forecasting model the better.

### 3.3. LSTM model prediction results and analysis

In this section, the ReLU function is selected as the activation function, the square loss function is used as the loss function, Adam is used as the optimization algorithm to optimize the loss function, and Dropout is used to prevent overfitting, and the number of iterations, the number of batches, and the number of hidden layer neurons are respectively 100, 64 and 64, establish the LSTM model, predict the short-term inbound passenger flow data of Yuantong Station, and analyze the prediction results. Draw a fitting diagram between the predicted value of the LSTM model and the actual value of subway passenger flow, as shown in Figure 1.

As can be seen from the figure, the LSTM model can more accurately predict the changes in subway passenger flow in the test set. However, the passenger flow on holidays is relatively scattered and has a certain degree of randomness, which results in the fitting effect of small fluctuations not being particularly good.

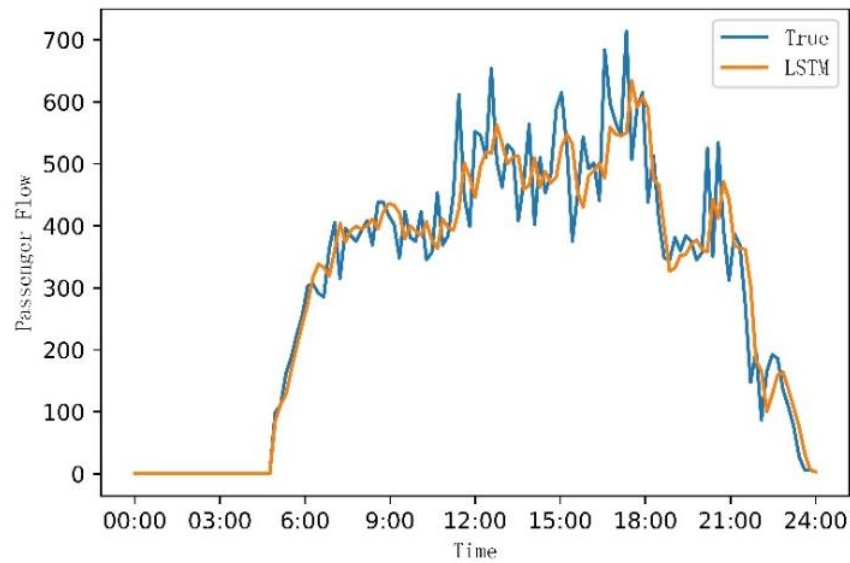


Figure 1. LSTM model passenger flow prediction at Yuantong Station

### 3.4. CNN-LSTM model prediction results and analysis

The passenger flow prediction model in this section is mainly composed of CNN and LSTM neural network. After using CNN to extract the spatial characteristics of passenger flow, an LSTM network layer is built to extract the temporal characteristics of passenger flow. The model selects a convolution kernel with a size of 3\*3, adopts the average pooling method, sets the number of neurons in the CNN layer to 32, the step size is 1, the number of neurons in the LSTM layer is 64, iterates 100 times, and reads neurons at a time The number is 64, and the activation function is selected as ReLU, the loss function is the square loss function, the optimization algorithm is Adam, and the Dropout is 0.2 to avoid the occurrence of overfitting, so as to build the convolutional long-term short-term memory network CNN-LSTM Model. Predict the short-term inbound passenger flow data of Yuantong Station and analyze the prediction results. Draw a fitting diagram between the predicted value of the CNN-LSTM model and the actual value of subway passenger flow, as shown in Figure 2.

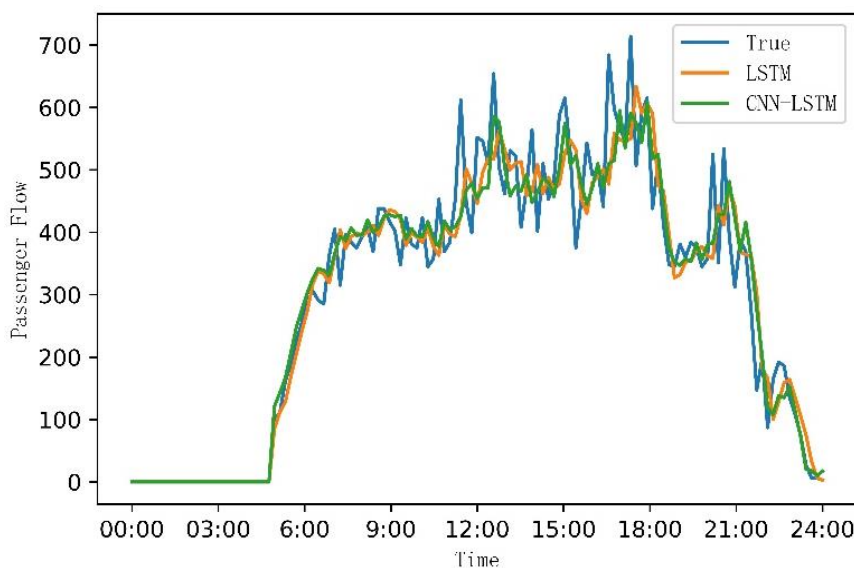


Figure 2. CNN-LSTM model passenger flow prediction at Yuantong Station

As can be seen from the figure, the CNN-LSTM model has a good effect in predicting changes in passenger flow in the test set. Compared with the LSTM prediction model, the CNN-LSTM prediction model fully considers the spatial characteristics of different sites and is more accurate in predicting small fluctuations in passenger flow. It is verified that in the study of short-term prediction

of subway passenger flow, the CNN-LSTM prediction model has better prediction results than the LSTM prediction model.

It can be seen from the table 1 that the CNN-LSTM combination prediction model is the best when comparing the various evaluation indicators of the model, and the error values are the smallest. Its MAE value on the test set is 6.8378, and the RMSE value is 10.0632. The MAPE value is 21.2618. It further proves the superiority of the convolutional long short-term memory network CNN-LSTM model established in this paper in the short-term prediction of local passenger flow.

**Table 1.** Model prediction result evaluation index

Model	MAE	RMSE	MAPE(%)
LSTM	8.4854	15.5693	26.6443
CNN-LSTM	6.8378	10.0632	21.2618

The experimental results show that CNN can effectively extract the spatial distribution characteristics of passenger flow between sites, LSTM can extract the temporal distribution characteristics of passenger flow in the time dimension, and CNN-LSTM can extract passenger flow characteristics from two dimensions of time and space to improve prediction accuracy.

#### 4. Summary

In this paper, using the historical passenger flow data of Nanjing Metro, supported by the powerful data processing ability and nonlinear fitting ability of neural network prediction method, this paper analyzes the time and space distribution characteristics of Nanjing Metro passenger flow, and uses the LSTM model and CNN-LSTM combination model to establish The short-term passenger flow forecasting model was established, and the optimal parameter combination model was obtained after adjusting the model parameter combination many times. Finally, by comparing the prediction performance of different models, the effectiveness of the subway short-term passenger flow prediction model based on the CNN-LSTM combined model is verified.

#### References

- [1] Palm R B. Prediction as a candidate for learning deep hierarchical models of data [J]. Technical University of Denmark, 2012, 5.
- [2] Greff K, Srivastava R K, Koutnik J, et al. LSTM: A search space odyssey [J]. IEEE transactions on neural networks and learning systems, 2017, 28 (10): 2222 - 2232.
- [3] HUBEL D H and WIESEL T N. Integrative action in the cat's lateral geniculate body [J]. The Journal of physiology, 1961, 155: 85 - 98.
- [4] Yann LeCun. Using curvature information to improve back-propagation [J]. Neural Networks, 1988, 1(Supplement-1): 168 - 171.
- [5] Huang W, Song G, Hong H, et al. Deep architecture for traffic flow prediction: deep belief networks with multitask learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 15 (5): 2191 - 2201.