J. Electrical Systems 20-3 (2024): 7703-7714

[1] Shuying Liu
[2] Vladimir Y.Mariano

Passenger Flow Prediction of Urban Rail Transit based on Dual- Channel and Multi-Factor

Journal of Electrical Systems

Abstract: - At present, in the process of subway passenger flow forecasting, there are often problems such as single forecasting model and incomplete construction features, which lead to the phenomenon of low forecasting accuracy. To solve these problems, this paper proposes a subway passenger flow prediction model based on dual-channel and multi-factor (DCMF-CNN-LSTM). Based on the comprehensive consideration of the influence of time characteristics, space characteristics and external characteristics on subway passenger flow, the model adopts dual-channel experimental route to predict subway passenger flow. First of all, based on the analysis of the passenger flow between adjacent stations, the passenger flow of the same station in the same time period, peak hours, weather and other factors on the passenger flow, the spatial characteristics, time characteristics and external characteristics of subway passenger flow models are established. Secondly, a dual-channel experimental route was established. Channel 1 was composed of CNN+LSTM+Attention model to extract the spatial and external features affecting passenger flow and complete the passenger flow prediction influenced by spatial and external factors; channel 2 was composed of LSTM model to extract the time features affecting passenger flow and complete the passenger flow prediction influenced by time factors. Finally, the random forest algorithm is used to fusion the prediction results. The experimental results show that the proposed prediction method based on dual-channel and multi-factorhas the lowest prediction error and higher practicability compared with other prediction methods.

Keywords: Dual channel, Multi-factor, Urban Rail Transit, Passenger Flow Prediction

I. INTRODUCTION

With the development strategy of "accelerating the development of railways and urban rail transit" [1] being proposed and implemented, urban rail transit plays an even more important role in my country's urban transportation development and construction planning. How to improve the efficiency of transportation and improve service levels is a concern of every urban rail transit operator. At the same time, making reasonable and efficient driving plans and decisions during the operation process requires the support of passenger flow data. Therefore, accurate passenger flow forecasting is particularly important. In the urban rail transit AFC system, a large amount of historical data has been accumulated. Through in-depth mining of these massive data sets, future passenger flow can be predicted. In recent years, with the joint efforts of domestic and foreign researchers, subway passenger flow forecasting has gradually matured and achieved fruitful results. At present, subway passenger flow prediction can be divided into the following types: prediction methods based on statistical principles, prediction methods based on deep learning, and prediction methods combining multiple methods.

Zhang Chunhui [2] et al. used the Kalman filter model to achieve real-time traffic flow prediction. Li Lihui et al. [3] used random forest regression algorithm to predict short-term passenger flow of high-speed railway.Zhao Peng et al. used ARIMA model to complete the research of urban rail transit inbound passenger flow prediction [4]; Huang Xiaohui used cuckoo algorithm to optimize wavelet neural network and applied it to short-term traffic flow prediction. By learning nonlinear features in passenger flow, the prediction accuracy was improved [5]. In 2015, Ma Xiaolei et al. applied the short-duration memory neural network to the prediction of traffic speed with high accuracy [6]. In 2019, Li Ruoyi carried out optimization on the basis of LSTM model and integrated spatio-temporal characteristics to complete the prediction of OD short-term passenger flow of urban rail transit [7]. In 2023, K. Zhang et al. proposed a short-time traffic flow prediction model integrating graph convolutional network and bidirectional long and short term neural network, which reduced the prediction error by 12.24% and 13.20% compared with the classical algorithm [8]. In 2023, Y. Sun proposed a prediction model combining attention mechanism and graph convolutional neural network. Simulation results show that the average displacement error and final deviation error of this model are significantly reduced [9]. Feng Biyu combined CNN model and LSTM model to overcome the defects of a single prediction model and improve the prediction

[1] *Corresponding author: Shuying Liu [1]College of Computing and Information Technologies, National University, Philippines
[2]College of Big Data and Artificial Intelligence, An Hui Xin Hua University, China
[2] Vladimir Y.Mariano College of Computing and Information Technologies, National University, Philippines

accuracy [10]. Luo Xianglong proposed a short-term traffic flow prediction model based on KNN and LSTM, which improved the prediction accuracy by 12.28% [11].

In the above researches, Zhang Chunhui, Yao Zhisheng, Ma et al. only considered the impact of time series data on passenger flow; Although Li Ruoyi et al. integrated the impact of time and space factors on passenger flow, they did not consider the comprehensive impact of external characteristics such as weather and holidays on subway passenger flow. Therefore, this paper proposes a prediction model dual-channel and multi-factor, which fully considers the time factors, space factors and external factors affecting passenger flow. At the same time, this paper also selects the AFC data of Line 1 in a city to complete the simulation experiment of the prediction model with the time granularity of 1 hour. Through the subway passenger flow prediction model constructed in this paper, accurate and reliable forecast data can be obtained, which can not only help the subway operation department to make more reasonable travel plans, optimize the configuration of passenger service, and warn the peak passenger flow in the station, but also provide passengers with a safer and more comfortable riding environment.

II. Related Work

A. CNN

Convolutional Neural Network (CNN) evolved from neural networks [12], it is a classical model in deep learning and one of the most representative and popular neural networks in the field of deep learning [13]. Figure 1 is network architecture .

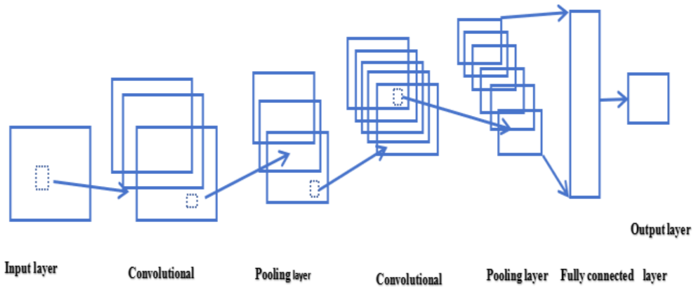

Figure 1.CNN neural network structure diagram

B. LSTM

LSTM [14] is an improved recurrent neural network, which overcomes the problem of gradient disappearance in RNN and also solves the problem of unmanageable long-term dependence [15], [19].

The memory unit of LSTM mainly consists of three gates symbolizing information, which controls the transmission of neuron information[16], [23]. Figure 2 is its network structure .

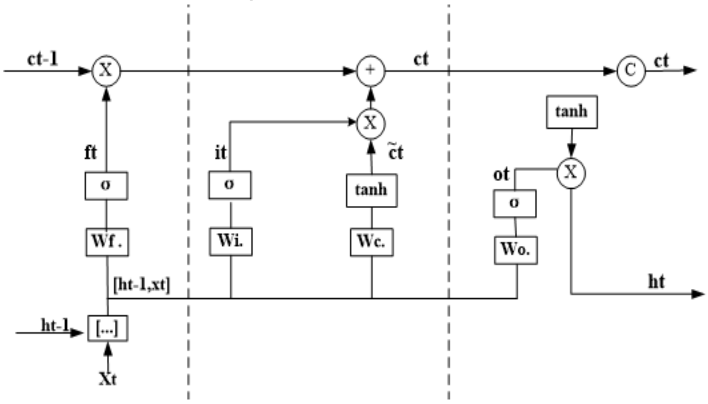

Figure 2. network structure of LSTM

III. Subway Passenger Flow Prediction Problem Formulation and Feature Modeling

A. Expression of subway passenger flow prediction problem

This article analyzes the historical passenger flow data and external factors (such as weather, holidays, etc.) of Metro Line 1 in a certain city and establishes a prediction model to predict passenger flow. The passenger flow at time t is predicted through the historical data of the previous t-n time steps. It can be expressed as:

$$Y_t^{prc} = F(X_{t-n}, X_{t-n-1}, X_{t-n-2} \dots\dots, X_{t-n}, X_{t-1}), (E_1, E_2, \dots\dots, E_n) \qquad (1)$$

Where, Y_t^{prc} is the passenger flow ; F(.) is the fusion function; $(X_{t-n}, X_{t-n-1}, X_{t-n-2} \dots\dots, X_{t-n}, X_{t-1})$ are multiple historical data that affect passenger flow; $(E_1, E_2, \dots\dots, E_n)$ are n external factors that affect passenger flow.

1）Temporal feature modeling

Through the analysis of historical passenger flow, it can be found that the subway passenger flow has obvious continuity and periodicity in time characteristics [17]. Therefore, through information mining of historical passenger flow data, this paper extracts daily cycle data and time cycle data that have an impact on passenger flow respectively, and constructs the time characteristics that affect passenger flow.

a) Daily cycle characteristics

The daily cycle characteristics of subway passenger flow can be expressed as: the passenger flow at a station at time t is affected by the passenger flow at the same time at the station n days before. Its expression is:

$$Y'_{dt} = f\left(X'_{(d-n)t},\ X'_{(d-n-1)t},\ X'_{(d-n-i)t},\ \dots\dots,\ X'_{1t}\right) \quad (2)$$

Where, Y'_{dt} is the passenger flow of the site at time t on a certain day that is affected by the daily cycle; $X'_{(d-n-i)t}$ is the passenger traffic volume at the station i days prior to time t; f (.) is the mapping function between them.

As shown in formula 2, the passenger flow at time t of this period is affected by the passenger flow at the same time in the previous n days. Figure 3 is daily data.

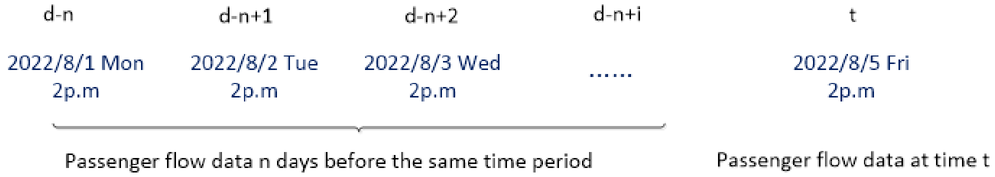

Figure 3.Daily period data

b) Time cycle characteristics

Temporal periodic features of subway passenger flow can be expressed as: The passenger flow is influenced by the passenger flow in the preceding n time periods of the same day. Its expression is:

$$Y''_{dt} = f(X''_{t-n},\ X''_{t-n-1},\ \dots\dots X''_{t-i},\ \dots\dots,\ X''_{t}) \quad (3)$$

Where, Y''_{dt} is the passenger flow at station t influenced by temporal cyclical factors, X''_{t-i} is the passenger flow at station t in the previous i time slots. As shown in formula 3, the passenger flow of this period is affected by n consecutive time steps before time t. Figure 4 is time preiod data.

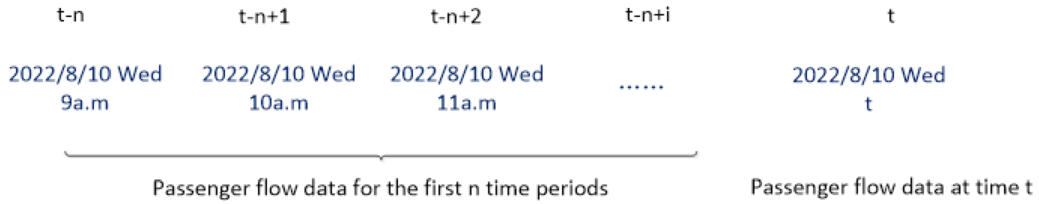

Figure 4. Time period data

2）External feature modeling

The spatial characteristics of the passenger flow of a station can be expressed as follows: the passenger flow of a station is not only affected by the historical passenger flow of the same station, but also affected by the passenger flow of other stations adjacent to this station. Its expression is:

$$Y'''_{dt} = f(X'''_{t-n},\ X'''_{t-n+1},\ \dots\dots X'''_{t-n+i},\ \dots\dots,\ X'''_{t}) \quad (4)$$

is the passenger flow of this station at time t affected by neighboring stations; X'''_{t-n+i} is the traffic of the previous i stations at t time. As shown in formula 4, the passenger flow at time t is affected by the passenger flow at time t of n stations adjacent to this station. AS shown in Figure 5, the base AS-i represents the i -th site adjacent to this site.

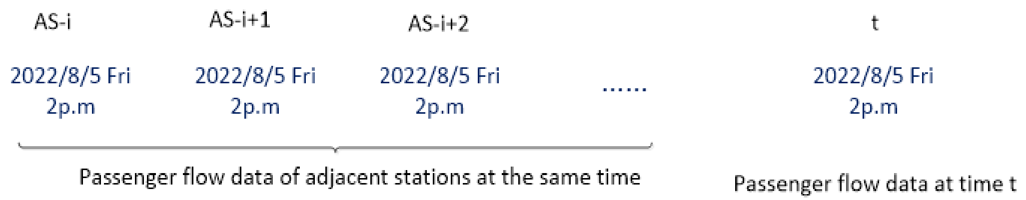

Figure 5.Spatial characteristics of passenger flow

3）Modeling of external features

Subway passenger flow is not only affected by historical passenger flow, but also by external factors such as peak hours and weather, which can be expressed as:

$E_t = f(R_t, W_t)$ (5)

Where, E_t Indicates the passenger flow of the current site at time t, R_t indicates whether the current site is in peak hour, and W_t indicates the weather information.

IV. PASSENGER FLOW PREDICTION MODEL BASED ON DUAL-CHANNEL AND MULTI-FACTOR

This paper predicts subway passenger flow by comprehensively considering the impact of the daily cycle characteristics, time cycle characteristics and external characteristics of the time series on passenger flow. The simulation experiment of this article consists of two parts. The first part is to complete the prediction of multi-factor subway passenger flow through dual channels, and obtain the predicted values Y_1' and Y_2' respectively. Among them, channel 1 mines information on spatial factors and external factors that passenger flow, and channel 2 mines information on time factors that passenger flow.

The experimental steps for channel 1 are as follows:

Step1: Select the time period data of the first n moments of m stations adjacent to the predicted station, construct an m*n matrix, and use it as the input of CNN to extract the spatial features that affect passenger flow.

Step2: After the output of the CNN model and the external factors affecting passenger flow enter LSTM+Attention, the predicted value Y_1' is obtained.

Channel 2 realizes the mining of daily cycle characteristic information. Using the historical data of the site n days before the same time as the input of LSTM, the predicted value Y_2'can be obtained.

The second part uses the random forest algorithm to fuse the two predicted values$Y_1'$and Y_2' obtained in the first part, and then obtains the final prediction result.

The research technical route of the simulation experiment is shown in Figure 6.

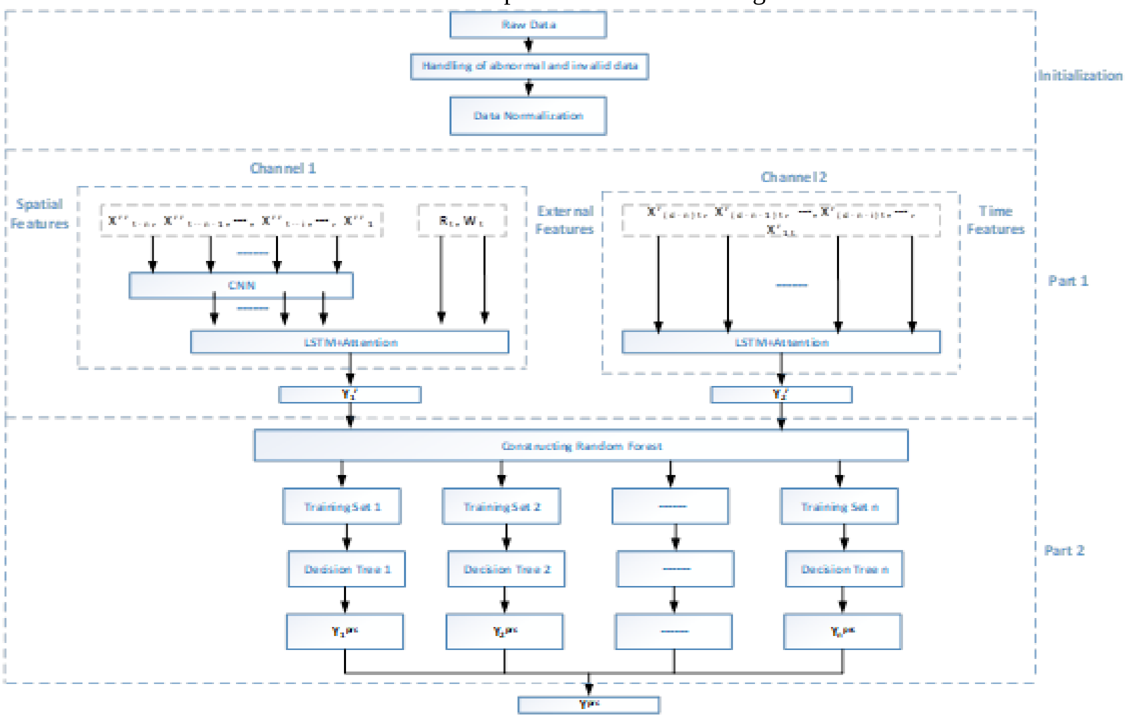

Figure 6.Simulation experiment technology framework diagram

A. Experimental data sets

The data set comes from the Metro Line 1 in a city in Henan Province in 2022. The time span is from January 2, 2022 to December 17, 2022, a total of 349 days of AFC data. The simulation experiment of this article is to predict passenger flow in units of 1 hour. Table I shows the hourly data of some sites on June 15, 2022.

TABLE I Part of the original data

Date	Xiliu Lake	Erqi Square	…	Zheng zhou East	…	City Sports Center
6 a.m	576	1542	……	346	…	864
7 a.m	1632	2846	……	1421	…	2001
8 a.m	2229	3076	……	2728	…	2350
9 a.m	1770	2516	……	2431	…	1932
…	…	…	……	…	…	…
8 p.m	781	2056	……	1184	…	1395
9 p.m	593	1712	……	570	…	1521
10 p.m	221	326	……	176	…	536

This line has a total of 20 passenger pick-up and drop-off stations. After statistical analysis of the passenger flow data of these 20 stations, 3 stations with large passenger flow gaps (Erqi Square, Municipal Sports Center, and West Third Ring Road) were selected for passenger flow analysis. Forecast, Figure 7 is the passenger flow distribution map of these three stations in one day. As can be seen from the figure below, Erqi Square and Municipal Sports Center are single-peak sites, while the West Third Ring Road is a double-peak site. Moreover, the passenger flow of the three stations is quite different.

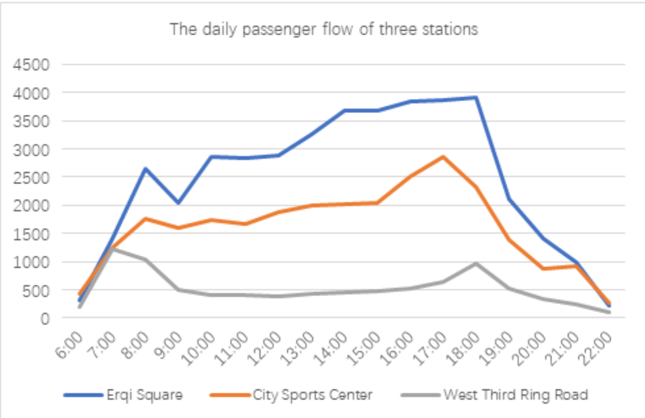

Figure 7.Passenger flow diagram of three stations in one day

B. Data preprocessing

Because during the collection process of original data, the collected data may be incomplete due to AFC system failure, card swiping failure, etc., and even abnormal data may appear [18]. These phenomena will lead to a reduction in the prediction accuracy of the model. In order to improve the prediction accuracy, this paper preprocesses the original data in the first line before training the prediction model.

1） Abnormal data and invalid data processing

First, detect whether there are site anomalies, time anomalies, data duplication, etc. in the data set [19]. For missing data in abnormal data, regression method, Lagrangian interpolation method, average method, maximum likelihood method, etc. are generally used for data interpolation [20], [21]. This simulation experiment is filled with the average method. The formula can be expressed as:

$$X_t=\frac{X_{t1}+X_{t2}\ \ \ldots\ldots+X_{t6}+X_{t7}}{7} \quad (6)$$

Where, X_t the average passenger flow at time t, and X_{ti} is the passenger flow for the previous i days at the same time. In the experiment, the average passenger flow at the same time in the previous seven days was used for interpolation.

2） Data normalization processing

If there are singular sample data in the original data, the training time of the prediction model will increase, and even convergence may not occur. At the same time, the prediction accuracy of the model will also be reduced [22]. Normalization processing can limit the data to the range [-1, 1], thereby reducing the adverse consequences caused by singular sample data [23]. This article uses the Min-Max Scaling method for normalization, and its formula is as follows:

$$X^n=\frac{X_t^d-\mathrm{Min}(X^d)}{\mathrm{Max}(X^d)-\mathrm{Min}(X^d)} \qquad (7)$$

Where, X^n is the normalized data, X_t^d is the data after difference.

After normalizing the data set, the corresponding predicted values can be obtained through model training, and then the data is denormalized to obtain the true predicted values. The denormalization formula is as follows:

$$X^p = \mathrm{Min}(X^p) + X^n[\mathrm{Max}(x) + \mathrm{Min}(X^p)] \qquad (8)$$

After the data is standardized, the processed data set needs to be split into three parts: training set, verification set and test set, with a split ratio of 6:2:2. The experiment selects the data of the past 39 weeks from January 3, 2022 to October 5, 2022 as the training set and verification set, and the data of the past 10 weeks from October 5, 2022 to December 16, 2022 as the test set. The prediction model proposed in this paper is trained and tested.

C. Model evaluation indicators

The quality of the prediction model needs to be evaluated through prediction errors. Prediction error refers to the deviation between the predicted value and the true value. The smaller the deviation, the smaller the error, indicating the higher the prediction accuracy of the model [24]. In this experiment, the average absolute error and the root mean square error are selected as the model performance evaluation indicators.

1) MAE

MAE [25] is the average value of the absolute value of the deviation between all individual predicted values and the true value[26]. In the MAE, the deviations are converted into absolute values, so there will be no phenomenon of mutual cancellation due to differences in the sign of the deviations. The calculation formula is as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i^p - y_i^a| \qquad (9)$$

2) RMSE

RMSE is standard error [27]. It represents a discrete program of predicted values[28], and is also used to measure the deviation between predicted values and true values. Its calculation formula is as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i^p-y_i^a)^2}{n}} \qquad (10)$$

Where, y_i^p is predicted value , y_i^a is true value.

D. Model evaluation indicators

1) Optimizing LSTM model parameters

Because there are many parameters in the LSTM model and have a great impact on the prediction accuracy, this article first conducts model parameter optimization experiments. At the same time, in order to improve the efficiency of the model and reduce training time, some parameters are fixed and some parameters are adjusted using the grid during the optimization process. Among them, the parameters involved in the grid adjustment are: input step size, hidden layers, hidden layer neurons and batch size. The remaining parameters are fixed values. The fixed parameter values are shown in table II:

TABLE II Fixed parameter value table in LSTM model

Parameters	loss function	Optimizer	Learning rate	Activation function	Number of iterations
Value	MAE	Adam	0.001	Relu	100

a) Input step size

The accuracy of the prediction model is related to the historical data it inputs. If the input step size is small, the model cannot learn the time series characteristics of the data well, resulting in a low prediction accuracy of the model; if the input step size is too large, the model will not only It will increase the computational intensity of the model by obtaining redundant information in historical data, and also cause the prediction accuracy to decrease. In order to verify the impact of the input step size on the prediction accuracy of the model, in addition to the fixed parameters taking the values in table II, on the basis of setting the hidden layer to 1, the batch size to 30, and the number of hidden layer neurons to 32, the historical data The input step sizes are set to 1, 2, 3... for model training.

TABLE III Training result table with different input step sizes

Site	West Third Ring Road		Erqi Square		City Sports Center	
Enter step size	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**
1	80.55	108.13	282.26	356.40	252.49	321.26
2	71.21	112.74	191.09	257.69	191.52	252.08
3	**61.61**	**95.735**	**153.27**	**215.75**	**178.38**	**238.40**
5	76.68	116.05	185.45	239.72	187.59	250.09
7	70.19	109.89	209.46	264.28	196.08	262.60
9	77.22	108.97	216.24	278.38	240.74	310.12

Figure 8 shows the MAE and RMSE values corresponding to three sites at different input step sizes.

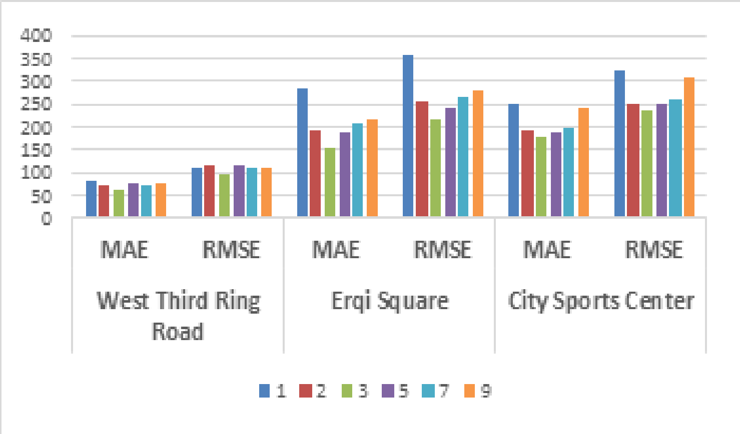

Figure 8.Optimization result diagram with different step lengths

It can be seen from the training results that when the input step size of the three sites is 3, the model has the best prediction effect and the values of MAE and RMSE are the smallest.

b) Number of hidden layers

The number of hidden layers in the LSTM model is also called the number of structures. The relationship between the number of hidden layers of the model and its performance is not a simple linear relationship [29]. Theoretically, the more hidden layers there are, the better the model can capture the complex, long-distance dependencies in historical data. However, when there are too many hidden layers, it may cause overfitting of the model during training, thereby reducing the performance of the model. In order to verify the impact of the number of hidden layers on the prediction accuracy of the model, this article designs 3 different numbers of hidden layers for model training. Among them, the input step size is 3, the batch size is 30, the number of hidden layer neurons is 32, and the remaining fixed parameters take the values in table I.

TABLE IV Training results of different hidden layers

Site	West Third Ring Road		Erqi Square		City Sports Center	
hidden layers	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**
1	**66.16**	**96.390**	**195.18**	**273.9**	**172.46**	**172.46**
2	67.24	98.111	218.52	314.46	196.75	196.75
3	73.89	122.33	223.37	307.324	174.116	174.116

Figure 9 shows the corresponding MAE and RMSE values of the three sites in different hidden layers.

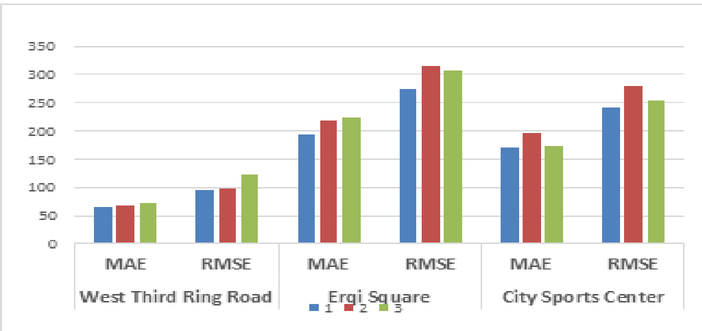

Figure 9.Optimization results of different hidden layers

It can be seen from the training results that when the hidden layer of this prediction model is 1, the MAE and RMSE results of the three sites are optimal.

c) Number of hidden layer neurons

The number of hidden layer neurons will have an important impact on the performance and learning ability of the model [30]. When the number of neurons is too large, although it can increase the expression ability and complexity of the model, it will also increase the training time of the model and consume more system resources; if the number of neurons is too small, it will lead to the model's imitation. The combined ability is insufficient and cannot accurately capture the relationship between input data, resulting in a decrease in the prediction accuracy of the model. In order to verify the impact of the number of hidden layer neurons on the prediction accuracy of the model, this paper designs different numbers of hidden layer neurons for model training. Among them, the input step size is set to 3, the batch size is 30, the number of hidden layers is 1, the fixed parameters are still the values in table I.

TABLE V Training results table with different numbers of neurons

Site	West Third Ring Road		Erqi Square		City Sports Center	
number of neurons	MAE	RMSE	MAE	RMSE	MAE	RMSE
16	90.41	141.76	243.97	301.61	247.58	321.03
32	75.52	117.95	217.73	273.00	221.42	284.24
64	**68.95**	**104.02**	**171.65**	**221.56**	**186.49**	**245.05**
128	77.93	108.32	188.94	266.89	207.03	278.77
256	75.64	104.79	192.03	273.71	204.77	275.27

Figure10 shows the MAE and RMSE values corresponding to different neuron numbers at the three sites.

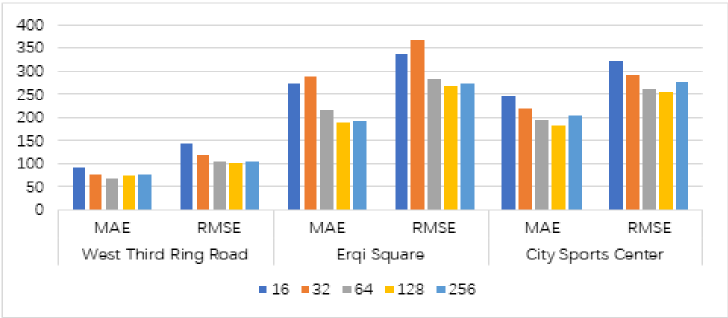

Figure 10.Optimization results for different numbers of neurons

When the number of hidden neurons is 64, the prediction performance of the model is optimal.

d) Batch size

Batch size refers to the number of samples processed at one time when the model is trained. When the batch setting is larger, although the training speed of the model can be improved, it may also lead to overfitting of the model; if the batch setting is small, although the generalization ability of the model can be improved, it will also increase the model training time [31], [32]. The fixed parameters are still the values in table I.

TABLE VI Batch size optimization results table

Site	West Third Ring Road		Erqi Square		City Sports Center	
Batch size	MAE	RMSE	MAE	RMSE	MAE	RMSE
10	65.45	100.69	164.84	215.02	199.48	266.24
20	78.55	107.25	164.71	220.89	182.91	249.37
30	**62.95**	**95.85**	**148.39**	**204.46**	**172.41**	**231.68**
40	69.64	105.28	192.41	261.18	188.22	245.46
50	65.13	101.97	210.71	290.68	185.82	246.53
60	66.35	100.99	193.83	262.94	178.68	244.53

Figure 11 shows the MAE and RMSE values corresponding to three sites at different batch sizes.

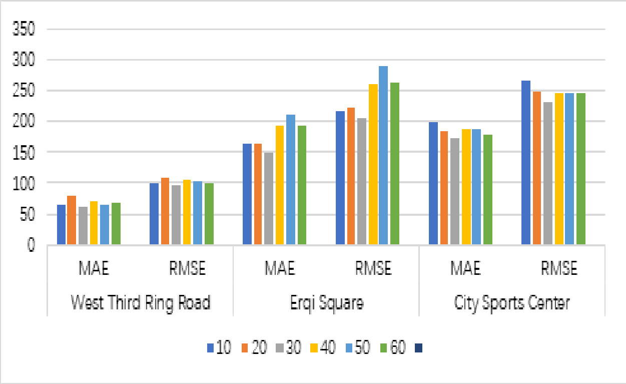

Figure 11.Optimization results for different batch sizes.

When the batch size is 30, the model has the best prediction performance.

In summary, The optimal parameter combination for LSTM model: input step size of 3, batch size of 30, number of hidden neurons of 64, and number of hidden layers of 1.

2) CNN model parameter optimization

In the construction process of the CNN model, there are many hyperparameters involved, such as learning rate, training times, batchsize, convolution kernel size and pooling method, etc. Reasonable adjustment of hyperparameters can improve the performance and generalization ability of the model [33], [34]. This article will focus on the impact of the following two parameters on the model.

a) Convolution kernel size

The convolutional layer implements feature extraction from the input data. Theoretically, the larger the convolution kernel, the more comprehensive the features extracted, which can better capture the characteristics of the input data, thus improving the accuracy of the model. However, large convolution kernels often bring large amounts of parameters and calculations, which can also cause model performance to decline; while smaller convolution kernels can extract more detailed features, they may lose some global information. Therefore, choosing an appropriate convolution kernel size is particularly important to improve the accuracy of the model.

In order to verify the impact of the convolution kernel size on model prediction, this paper sets the convolution kernel to 1×1, 3×3, 5×5, and 7×7 in sequence for model training.

TABLE VII Convolution kernel size optimization results table

Site	West Third Ring Road		Erqi Square		City Sports Center	
Pooling method	MAE	RMSE	MAE	RMSE	MAE	RMSE
Maximum pooling	**67.15**	**97.65**	**229.42**	**337.65**	**191.74**	**258.87**
Average pooling	97.65	106.34	233.65	346.62	193.01	265.78

Figure 12 shows the results corresponding to the three sites.

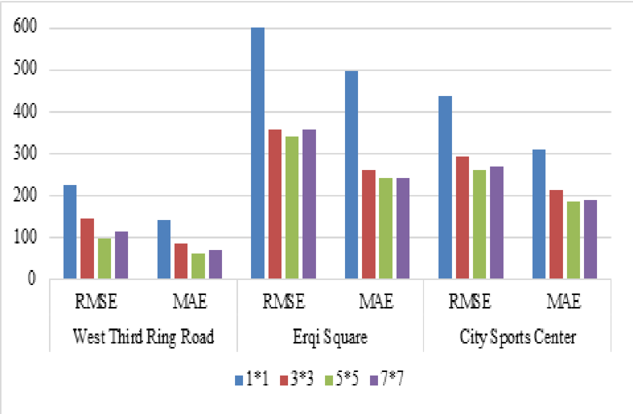

Figure 12.Optimization results of different convolution kernel sizes.

When the convolution kernel is 5×5, the MAE and RMSE results of the three sites are optimal.

b) Pooling method

In the CNN model, pooling can reduce the dimension of features, and the maximum pooling function or the average pooling function is usually used to obtain new, smaller-dimensional features. Among them, max pooling can retain more of the most significant features in the input data, while average pooling can retain more of the global information of the input data. Different pooling operations will affect the performance of the model. Table VIII shows the training results of maximum pooling and average pooling.

TABLE VIII Pooling Optimization Results Table

Site	West Third Ring Road		Erqi Square		City Sports Center	
Pooling method	MAE	RMSE	MAE	RMSE	MAE	RMSE
Maximum pooling	**67.15**	**97.65**	**229.42**	**337.65**	**191.74**	**258.87**
Average pooling	97.65	106.34	233.65	346.62	193.01	265.78

From the above results, we can see that there is not much difference between the two pooling methods. This is because the purpose of pooling is to reduce the dimensionality of the data, and the experimental data used in this article considers the mutual influence between adjacent stations on the same line, but does not consider the influence between stations on cross lines. Therefore, the experimental data Dimensions are smaller. To sum up, the optimal parameter combination of the CNN model was finally determined to be: the convolution layer is 1, the number of convolution kernels is 25, the number of iterations is 100, the convolution kernel is 5×5, and the maximum value pooling method.

V. MODEL PERFORMANCE EVALUATION

In this paper, three stations with large passenger flow differences are selected for simulation experiments to verify the performance and generalization ability of the dual-channel multi-factorprediction model (DCMF-CNN-LSTM). At the same time, in order to further verify the validity of the dual-channel multi-factorprediction model, this paper compares it with the single-channel multi-factorprediction model (SCMF-CNN-LSTM) and other basic models (GRN,LSTM,CNN). Table IX shows the prediction results of different models.

TABLE IX Results of Different Prediction Models

Prediction model	West Third Ring Road		Erqi Square		City Sports Center	
	MAE	**RSME**	**MAE**	**RSME**	**MAE**	**RSME**
D_CNN_LSTM	**9.61**	**22.72**	**19.16**	**26.76**	**19.29**	**37.05**
S_CNN_LSTM	38.14	53.23	32.79	47.72	46.33	57.17
LSTM	64.92	97.99	167.12	228.92	177.44	239.23
CNN	80.89	82.90	290.30	286.51	222.99	226.29

From the above results, the prediction model proposed in this article has the best prediction performance.Compared with the single-channel and multi-factor prediction model, MAE and RMSE of the dual-channel and multi-factorprediction model are reduced by 28.5% and 30.5% respectively when applied to Xisanhuan Station with small passenger flow. However, in Erqi Square station with large passenger flow, MAE and RMSE of dual-channel and multi-factor prediction model decreased by 13.6% and 20.9% respectively. From the experimental results, it can be seen that the dual channel multi factor prediction model proposed in this article has achieved good prediction results for passenger flow at stations of different orders of magnitude. At the same time, it also indicates that the model has strong adaptability, stability, and good generalization ability.

VI. CONCLUSION

In view of the periodic and nonlinear characteristics of subway passenger flow, and on the basis of fully considering the time, space and external factors that affect passenger flow, this paper proposes a dual-channel and multi-factor passenger flow prediction method, constructs a dual-channel and multi-factor subway passenger flow prediction model, and conducts simulation experiments on different sets of stations. Finally, the results of this model are compared with other models, and the results show that:

(1) For the prediction model of dual-channel and multi-factor, because the parameters of the basic model have a great influence on the prediction results of the model, after several optimization experiments, the optimal parameter combination of the LSTM model is selected as follows: input step size is 3, hidden layers is 1, neurons is 64, and batch size is 30;The optimal parameter combination of CNN model is: convolution layer is 1,

convolution nuclei is 25, iterations is 100, and convolution kernel is 5×5.In the final dual-channel multi-factorprediction model, MAE and RMSE decreased significantly compared with other parameter combinations.

(2) Compared with SCMF-CNN-LSTM, GRN, LSTM and CNN, the two-channel multi-factorprediction model showed significant advantages in predicting subway passenger flow, with the MAE decreased by at least 13.6% and the RMSE decreased by at least 20.95%. This result fully proves the effectiveness of the forecast model designed in this paper in the subway passenger flow prediction, showing its excellent forecasting performance.

(3) Through simulation and comparison experiments, this paper verified the applicability of the proposed dual-channel and multi-factorprediction model in subway passenger flow prediction, and also improved the prediction accuracy.However, due to the limited experimental data set in this paper, the impact of spatial factors between different stations and the historical time factors of the same station on passenger flow on only one line was not taken into account.At the same time, in terms of external factors, we only consider the impact of weather and holidays on passenger flow, but do not take into account the geographical location of the station, special dates, train number distribution frequency and other influences on passenger flow.Therefore, in the subsequent research, I will further improve the data set, integrate more factors affecting subway passenger flow, select more reasonable and effective prediction methods, further improve the performance, generalization ability and prediction accuracy of the prediction model, and enhance the fitting ability of the model.

ACKNOWLEDGMENT

1. Anhui Province Higher Education Natural Science Research Project.NO:KJ2021A1159

2. 2023 Anhui Province College Student Innovation and Entrepreneurship Project.NO:202312216027

3.2021 National first-class undergraduate major construction point: Software Engineering Major.NO:2020ylzy01

REFERENCES

[1] Yao Junfeng, He Rui, Shi Tongtong, Wang Ping, Zhao Xiangmo. A review of traffic flow prediction methods based on machine learning [J]. Journal of Traffic and Transportation Engineering 2023, 34(03), pp. 44-67.

[2] Zhang Chunhui, Song Rui, Sun Yang. Short term passenger flow prediction at bus stops based on Kalman filtering [J]. Transportation System Engineering and Information, 2011,11(04), pp. 154-159.

[3] Li Lihui, Zhu Jiansheng, Qiang Lixia, et al. Short term passenger flow prediction of high-speed railways based on random forest regression algorithm [J]. Railway Transport and Economy, 2017, 39 (09), pp.12-16.

[4] Zhao Peng, Li Lu. Prediction of Urban Rail Transit Station Arrival Volume Based on ARIMA Model [J]. Journal of Chongqing Jiaotong University (Natural Science Edition), 2020, 39(01), pp. 40-44.

[5] Huang Xiaohui. Research on Short term Traffic Flow Prediction of Wavelet Neural Network Based on cuckoo algorithm [D]. Southwest Jiaotong University, 2017.

[6] Xiaolei Ma, Zhimin Tao, Yinhai Wang, Haiyang Yu,Yunpeng Wang. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data [J]. Transportation Research Part C, 2015. pp. 1423-1432.

[7] Li Ruoyi. Short term prediction of OD passenger flow in urban rail transit systems based on an improved spatiotemporal LSTM model [D]. Beijing: Beijing Jiaotong University, 2019.

[8] K. Zhang, H. Liu and S. Clarke.GATFormer: A Graph-based Transformer for Long-Term Forecasting of Traffic Overcrowding [J]. 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC), Bilbao, Spain, 2023, pp. 1629-1635.

[9] Y. Sun, X. Zhou, Z. Yang, W. Wang and Q. Shi.Flight Trajectory Prediction Method Based on Attentional Graph Convolutional Network [J]. 2023 3rd International Symposium on Computer Technology and Information Science (ISCTIS), Chengdu, China, 2023, pp. 127-132.

[10] Feng Biyu. Research on Short term Passenger Flow Prediction of Urban Rail Transit Based on CNN-LSTM Combination Model [D]. Nanchang: East China Jiaotong University, 2020.

[11] Luo Xianglong, Li Danyang, Yang Yu, et al. Short-term traffic flow prediction based on KNN-LSTM [J]. Journal of Beijing University of Technology, 2018, 44(12), pp.1521-1527.

[12] Jinlei Zhang;Feng Chen;Zhiyong Cui;Yinan Guo;Yadi Zhu.Deep Learning Architecture for Short-Term Passenger Flow Forecasting in Urban Rail Transit[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, pp. 1221-1227.

[13] Zhang Wenjuan, Yang Haozhe, Zhang Bin et al. Short term passenger flow prediction model for urban rail transit considering multi time scale characteristics [J]. Transportation System Engineering and Information, 2022, 22 (06), pp. 212-223.

[14] SUN B, SUN T, ZHANG Y J, JIAO P P. Urban traffic flow online prediction based on multi-component attention mechanism [J]. IET Intelligent Transport Systems, 2020, 14(10), pp. 1249-1258.

[15] DAI Junming, CAO Yang, SHEN Qingin, et al. Traffic flow prediction based on multi-spatial-temporal graph convolu- tional network [J]. Application Research of Computers, 2022, 39(3), pp. 780-784.

[16] Feng Siyun, Shi Zhenquan,Cao Yang. A prediction model for urban road network traffic speed based on global spatiotemporal characteristics [J].Computer Engineering, 2022, 48 (5), pp. 112-117 (inChinese).

[17] LU S Q, ZHANG Q Y, CHEN G S, SENG D W. A com-bined method for short-term traffic flow prediction based on recurrent neural network [J]. Alexandria Engineering Joumal , 2021 ,60(1), pp. 87-94.

[18] Yao Xiangming, Zhao Peng, Yu Dandan. Short term passenger flow OD estimation model for urban rail transit network [J]. Transportation System Engineering and Information, 2015, 15 (02), pp. 149-155+162.

[19] Li Shuqing. Research on subway passenger flow prediction based on improved LSTM model with attention mechanism [J]. Industrial Control Computer, 2023,36 (11), pp. 124-125+128.

[20] Kim Y,Wang Peng,Mihaylova L. Scalable Learning of Structural Recurrent Neural Networks for Short Term Traffic Prediction [J].IEEESensorsJournal,2019,19 (23), pp. 11359-11366.

[21] Zhang,Feng Chen,Zhiyong Cui,Yinan Guo,Yadi Zhu.Deep Learning Architecture for Short-Term Passenger Flow Forecasting in Urban Rail Transit[J]. Jinlei IEEE Transactions on Intelligent Transportation Systems, 2020. 31 (10), pp.669-682.

[22] Yang Xin,Xue Qiuchi,Yang Xingxing,Yin Haodong,Qu Yunchao,Li Xiang,Wu Jianjun.A novel prediction model for the inbound passenger flow of urban rail transit[J]. Information Sciences,2021 ,36 (11), pp.258-263.

[23] Li Jie, Peng Qiyuan, Wen Chao. Research on Short term Passenger Flow Prediction of High Speed Railways Based on LSTM Deep Neural Network [J]. Systems Engineering Theory and Practice, 2021, 41 (10), pp.2669-2682.

[24] XIONGTing,QIYong,ZHANG Weibin,etal.Short term traffic flow forecasting model based on temporal-spatial correlation [J].ComputerEngineeringandDesign,2019,40(2), pp.501-507.

[25] R. Liu, Y. Wang, H. Zhou and Z. Qian.Short-Term Passenger Flow Prediction Based on Wavelet Transform and Kernel Extreme Learning Machine[J]. IEEE Access, vol. 7, 2019,pp. 1525-1534,

[26] LU Wenqi,GU Yuanli,CHENLun.Short-term traffic flow forecasting of urban expressway based on spatial and temporal traffic characteristics [J].ComputerSimulation,2018,35 (9), pp. 136-140 .

[27] DU Zhiming,LI Zhuang,GuoLimin.Short-termtraffic flow prediction based on ACGRU model [J].ComputerEngineering and Applications,2020,56 (21), pp. 260-265.

[28] Haiying Li,Yitang Wang,Xinyue Xu,Lingqiao Qin,Hanyu Zhang.Short-term passenger flow prediction under passenger flow control using a dynamic radial basis function network[J]. Applied Soft Computing Journal,2019. 46 (21), pp.1 260-1265.

[29] Lei Yi,Liu Mingsheng,Xie Yunchi,Nie Xuefang. A traffic flow prediction model based on attention gated graph convolution [J]. Journal of Physics: Conference Series. 2023. 34 (06) , pp.400-405.

[30] Y. Danfeng and W. Jing.Subway Passenger Flow Forecasting With Multi-Station and External Factors [J]. IEEE Access, vol. 7, 2019 , pp. 5415-5423.

[31] Haydari A, Yilmaz Y. Deep reinforcement learning for intelligent transportation systems: A survey [J]. IEEE Trans on Intelligent Transportation Systems, 2022, 23 (1) , pp.11-32.

[32] LI J, PENG Q Y, WEN C. Short term passenger flow prediction of high speed railway based on LSTM deep neural network[J]. Systems Engineering— Theory & Practice, 2021, 41(10). pp. 2669-2682.

[33] GU Zhenyu, CHEN Cong, ZHENG Jiajia, et al. Traffic flow prediction based on STG-CRNN [J]. Control and Deci- sion, 2022, 37(3) , pp.645-653.

[34] P. Li et al..IG-Net: An Interaction Graph Network Model for Metro Passenger Flow Forecasting [J]. IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 4, pp. 4147-4157.

[35] Y. Zhao, Z. Ma, Y. Yang, W. Jiang and X. Jiang.Short-Term Passenger Flow Prediction With Decomposition in Urban Railway Systems [J]. IEEE Access, vol. 8, pp. 2020,1876-1886.