Short-term Bus Passenger Flow Forecast Based on CNN-BiLSTM

Chaohua Wu[a], Xingzu Qi[b]

Shenzhen Urban Transport Planning Center Co.,Ltd, Shenzhen, China

[a] wch@sutpc.com , [b] qixingzu@126.com

Abstract. Effective prediction of urban bus passenger flow is critical for improving urban bus operation efficiency and optimizing the bus network. However, there are some issues with predicting urban bus passenger flow at the moment, such as lack of single eigenvalue consideration and insufficient research depth. In order to improve the short-term prediction accuracy of urban bus passenger flow, this paper proposed a deep learning prediction model that is based on CNN-BiLSTM. Based on historical data of urban bus passenger flow, this paper analyzes the dependence of bus credit card data, clusters the travel feature of different groups of people, and analyzes the dependence of bus passengers. Simultaneously, external factors of passenger flow, such as rainfall, weather condition, traffic flow state, and date, are introduced to build the bus passenger flow prediction feature matrix, and the correlation analysis of the characteristic matrix structure is performed to optimize the matrix structure. Finally, the optimized passenger flow characteristic matrix is fed into the CNN-BiLSTM deep learning model for prediction, and the results are compared to the LSTM, CNN and CNN-LSTM models. The results shown that the root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) of the the CNN-BiLSTM deep learning model are lower than those of other models, and the prediction accuracy is the highest. Meanwhile, this method has a good generalization effect and can improve deep learning prediction accuracy.

Keywords: intelligent transportation; passenger flow prediction; deep learning; bus passenger flow.

1. Introduction

Urban public transportation is an important component of the city's traffic system, which attracting a large number of travel demands. With the development of metro and urban transportation, there is a mismatch between supply and demand in the urban bus network. Therefore, urban ground bus passenger flow forecasting is critical for developing scientific and reasonable network schemes and station planning. At present, there have been many research results on urban traffic flow prediction, but there have been few studies on bus passenger flow prediction. Most of the passenger flow prediction methods are machine learning and statistics, such as Kalman filter model[1], empirical mode decomposition[2] and adaptive enhancement model[3]. This kind of models are difficult to reflect the irregular state of bus passenger flow. In order to explore the deep feature of passenger flow, deep learning models has made progress in short-term prediction in recent years. Deep learning model[4] establishes relationships with the feature structure of multidimensional data and analyze the complex features between the data. And through a lot of deep learning to make up the traditional prediction model, which are not well understood of eigenvalues. Sohani et al[5]. used the BiLSTM model to forecast the passenger flow of bus stations by historical data, but did not consider the weather, environment, and other factors. Li et al.[6] analyzed the dynamic rule of subway passenger flow based on Internet personal behavioral information and passenger flow data. And the subway passenger flow is predicted by LSTM model. The limitation is that, when the passenger flow is low, the differences between individual behavioral information will lead to large relative errors. Wang et al.[7] built a short-term bus passenger flow prediction model based on the real-time correlation degree of stations. But the prediction model requires a large amount of calculation, and the model architecture cannot realize parallel calculation. The calculation efficiency is low.

In summary, in the short-term forecast of urban bus passenger flow, the forecast accuracy depends on the ability to extract and process passenger flow feature. However, in terms of feature matrix extraction, convolutional neural networks (CNN) and convolutional long-term and short-term networks (BiLSTM) have distinct advantages. The two-dimensional matrix convolution operation of CNN and BiLSTM model can handle the intricate interrelations in high-dimensional fused data, as opposed to the deep learning model, which can only capture a single feature vector. Therefore, this paper combines the CNN and BiLSTM models to construct a CNN-BiLSTM fusion prediction model. The passenger flow dependence influencing factors are added, and the constructed passenger flow characteristic matrix is optimized. The results show that the model in this paper has higher prediction accuracy when compared to other deep learning models, which reflects the CNN-BiLSTM model's prediction ability for complex multidimensional features.

This paper's technical path is as follows:

The passenger flow characteristic matrix X is formed by combining external factors based on the analysis of passenger flow card data and travel characteristic factors. The characteristic matrix structure is optimized and transformed into model input data Xt = [x1, x2, x3,···,xn]. xn is the NTH eigenvector at time t. The short-term prediction model of bus passenger flow is built using CNN-BiLSTM.

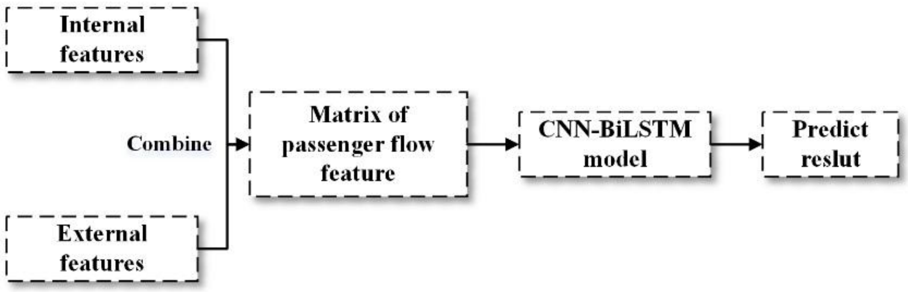


Fig. 1 The framework of CNN-BiLSTM

2. Constructing feature matrix

2.1 External features

Many external variables influence bus passenger flow, the most important of which are as follows:

(1) Peak/Off-peak hours. The peak hours for urban traffic are 7:00-9:00 and 17:00-19:00. Off-peak hours are 6:00-7:00, 9:00-17:00, 17:00-24:00. In this paper, the feature value of passenger travel during peak periods is set to 1, while it is set to 0 during off-peak periods.

(2) Weather conditions. The weather has a significant impact on bus passenger movement. This paper highlights the weather features of passenger travel based on precipitation levels. The characteristics of precipitation are classified into four groups: precipitation less than 10mm has a feature value of 1, precipitation between 10 and 25mm has a feature value of 2, precipitation between 25 and 50mm has a feature value of 3, and precipitation beyond 50mm has a feature value of 4.

(3) Holidays/Working Days. Holidays and working days are crucial variables influencing bus passenger flow, with working days having amount passenger flow than holidays. In this study, the eigenvalue of working days is fixed to 1 and that of holidays to 2.

(4) Attraction around the bus stop. The public facilities around the bus stop, such as residential, commercial, and office buildings, are appealing to the bus passenger flow. This paper selects the number of residential facilities, commercial facilities, educational facilities and leisure places within 500 meters around the site as the attraction of the bus stop.

2.2 Internal features

Internal influencing factors are mainly those related to passengers, including the sorts of bus cards used by passengers and the level of passenger reliance. The following are the particular categories:

(1) Type of bus card. Different types of bus cards have different travel features. The types of experimental data used in this paper mainly include ordinary cards, elderly cards and QR codes.

(2) The level of passenger reliance. Passengers will take a predetermined bus route, which will mostly represent daily commuting and go to and from school. This type of feature is defined as the level of passenger reliance. The number of people who use the bus in a week and the number of consecutive days on the same route in a week are used as assessment indices. Reflect the degree of reliance on the bus, and use the Apriori model to examine the relationship between the two dimensions. Calculate the support degree, confidence degree, and promotion degree of each indication and group the findings into three categories. Figure 2 shows the correlation analysis results of bus passengers' card data in a week. The level of passenger reliance on public transportation can be categorized into three categories: low, medium, and high, with a significant variation in each cluster center value. It demonstrates that the Apriori correlation algorithm is effective in processing bus card data from passenger flows.

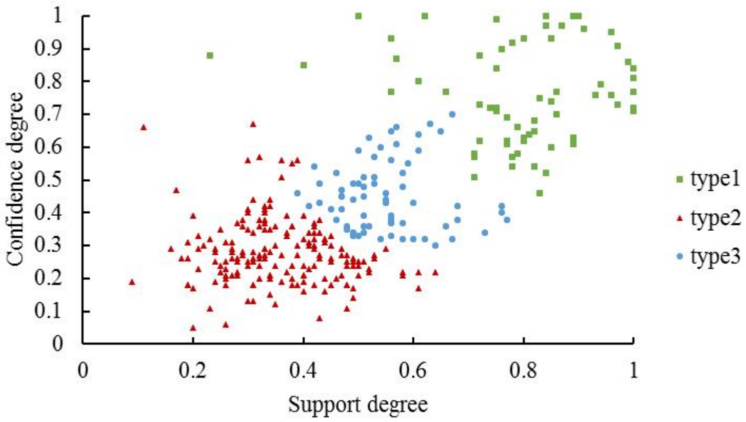


Fig. 2 Classification of passenger reliance T.

2.3 Construction of feature matrix

The external and internal features of bus passengers are constructed into a feature matrix, as illustrated in Table 1.

Table 1. Matrix of passenger flow feature

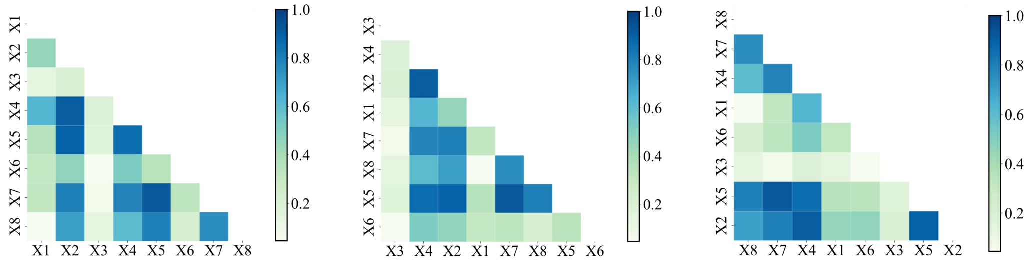
External eigenvector		Internal eigenvector	
Eigenvector	Meaning	Eigenvector	Meaning
x1	Weather conditions	x5	Passenger flow in the past
x2	Peak/Off-peak hours	x6	Boarding time
x3	Holidays/Working Days	x7	Type of bus card
x4	Attraction around the bus stop	x8	The level of passenger reliance

Mutual information (MI) is used in this work to calculate the correlation between feature vectors. The degree of correlation between two variables, which are determined by two discrete random variables, is referred to as MI. The MI correlation between distinct eigenvectors is determined using bus card data in this paper using formula (1).

$$I(X;Y)=\iint_{XY} p(x,y)\log(\frac{p(x,y)}{p(x)p(y)})dxdy \tag{1}$$

Where $p(x,y)$ denotes the joint probability density function of X and Y, $I(X,Y)$ denotes MI.

This paper's passenger flow feature matrix is similar to a two-dimensional picture matrix. The eigenvalues of each pixel are retrieved for prediction using a convolution operation on the passenger flow characteristic matrix. Because the structure of the feature matrix affects the deep learning model's capacity to extract and predict feature information, the coupling between feature vectors has a significant impact on model accuracy. As shown in Figure 2, the deeper the hue, the higher the correlation between the eigenvector and the distinctive matrices with varied structures display different matrix correlations. To enhance the model's prediction effect, this study analyzes the prediction results of characteristic matrices of different structures in the prediction stage and picks the structure with the best prediction effect as the prediction result to increase the model's forecast accuracy.



(a) matrix structure 1 (b) matrix structure 2 (c) matrix structure 3

Fig. 3 Correlation heat maps of feature matrix

3. Model description

3.1 CNN-BiLSTM model

This paper proposes a CNN-BiLSTM model for predicting bus passenger flo. Figure 4 shows the technical path. The model is made up of an input layer, a CNN layer, a BiLSTM layer, a complete connection layer, and an output layer. CNN layer is made up of convolution layer and pooling layer. After normalization, the bus passenger flow characteristic matrix is fed to the CNN layer, and a new characteristic matrix is created by extracting the local characteristic matrix, which is then input to the BiLSTM layer. The data is then sent into the dropout layer through to prevent over-fitting, and ultimately the output result is achieved.

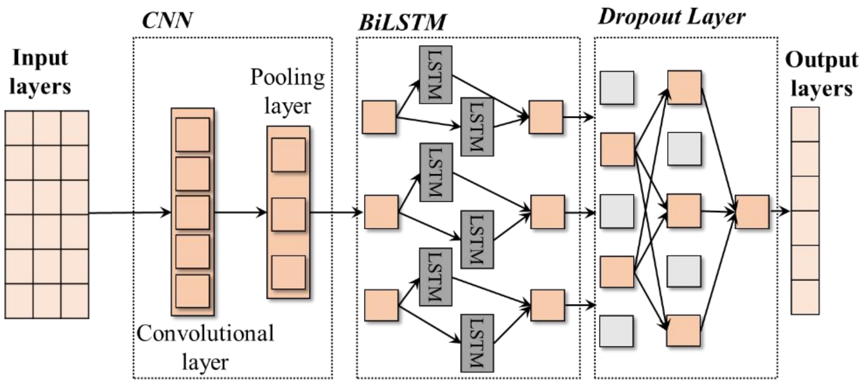


Fig. 4 Model structure of CNN-BiLSTM

4. Experiment

4.1 Experimental data

The data of No.151 bus in Foshan in June 2019 is studied and validated in this paper. The bus route has 23 stations and a total length of 13.5 kilometers. In Foshan city, the bus route runs through various urban functional areas like as residential regions, commercial districts, schools, hospitals, subway stations, and so on. Because the passenger flow and traffic circumstances are different, it has a good representation. Pre-process the data from the passenger flow. The operating hours of the bus are from 6:00 to 23:00. The passenger flow is measured at statistical intervals of 15 minutes and a day has 72 passenger flow data. The data from the first three weeks is utilized as the training set, while the data from the last week is used as the test set. The experimental setup is built using anaconda python3.5, tensorflow = 1.0.1, gpu=GTX750Ti.

4.2 Parameter Setting

In this model, the kernel size of CNN layer is set to 2, the activation function is Relu, the hidden state vector dimension of BiLSTM is set to 200, and the activation function is Relu. At prevent over-fitting, the Dropout rate parameter is set to 0.1. The optimizer of the model uses Adam optimizer, with 50 iterations, 64 batches and 0.001 learning rate.

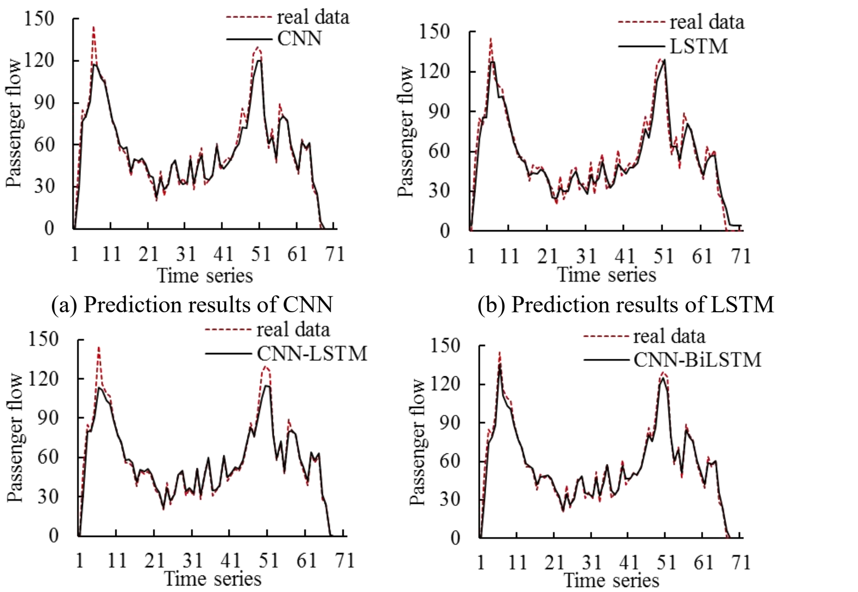
To validate the prediction impact of this model, it is compared to three additional deep learning models: CNN, LSTM, and CNN-LSTM. The prediction evaluation indexes are Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), and the evaluation indexes are as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_i - \hat{y})^2} \tag{2}$$

$$MAE = \frac{1}{n}\sum_{i-1}^{n}\left|\frac{y_i - \hat{y}}{y_i}\right| \tag{3}$$

$$MAPE = \sum_{t-1}^{N}\left|\frac{y_i - \hat{y}}{y_i}\right| \times \frac{100}{n} \tag{4}$$

Where *yi* is the observed values of bus passenger flow; $\hat{y}$ is the predicted values of bus passenger flow; *n* is the sample size of bus passenger flow. The smaller the RMSE, MAE, and MAPE values, the lower the forecast error of bus passenger flow. The prediction performance indicators and predictions of each model are compared in the table below to measure the accuracy of the CNN-BiLSTM prediction model in this article.



(a) Prediction results of CNN (b) Prediction results of LSTM

(c) Prediction results of CNN-LSTM (d) Prediction results of CNN-BiLSTM

Fig.5 Comparison of prediction results

Table 2. Comparison of prediction results in different methods

Model	MAE	RMSE	MAPE(%)
CNN	10.29	12.65	21.35
LSTM	8.87	9.73	18.94
CNN-LSTM	5.12	6.43	12.48
CNN-BiLSTM	3.98	5.87	11.91

Table 2 and the prediction results show that the prediction accuracy of CNN and LSTM is lower than that of CNN-LSTM, which is because CNN-LSTM combines the benefits of CNN in extracting local features and LSTM in collecting contextual data characteristics, and the prediction accuracy is improved. The CNN-BiLSTM model proposed in this paper has lower RMSE, MAE, and MAPE values than the CNN-LSTM method. The study showed that CNN-BiLSTM can effectively deal with the time and geographical features of bus passenger flow, as well as further research the connection between complicated feature matrices, improving prediction accuracy.

5. Conclusion and futureworks

In this paper, bus passenger flow data is classified into internal and external factors, and passenger flow dependent factors are introduced to define the features of bus passenger flow. The organization structure of the passenger flow characteristic matrix is optimized, and the passenger flow forecast accuracy is increased, by developing the characteristic matrix of internal and external components. The results of the experiments reveal that:

(1) The CNN-BiLSTM-based deep learning model can successfully deal with the complicated spatiotemporal feature matrix, collect information between feature vectors, and increase prediction accuracy.

(2) The structure of the bus passenger flow characteristic matrix has a significant impact on the passenger flow forecast effect, and the arrangement structure of the characteristic matrix has an optimal structure, which may increase the passenger flow forecast effect even further.

(3) This paper's study framework of bus passenger flow has a good generalization impact. In the future, we will do more research on the internal and external influencing elements of passenger flow, such as enhancing the influencing factors of transfer passenger flow for the subway-bus link line. Improve prediction accuracy even further.

References

[1] Lingru Cai, Zhanchang Zhang, Junjie Yang, Yidan Yu, Teng Zhou and Jing Qin,"A noise-immune Kalman filter for short-term traffic flow forecasting,". Physica A: Statistical Mechanics and its Applications, 536(C)(2019).

[2] Ma Yingying and Jin Xuezhen,"Short-term Traffic Flow Forecast Method Based on EEMD-Wavelet Threshold, " Journal of Chongqing Jiaotong University(Natural Science), 41(06), 22-29(2022).

[3] Li Shubin, Kong Xiangke, Li Qingtong et al,"Short-term traffic flow prediction with PSR-XGBoost considering chaotic characteristics," Journal of Southeast University (English Edition), 38(01),92-96(2022).

[4] Qi Xingzu, "Freeway Short-Term Traffic Flow Forecasting Based on MI-CEEMDAN-ADABOOST[J]. Journal of Highway and Transportation Research and Development," 39(06),136-143(2022).

[5] Thieu Nguyen, Giang Nguyen and Binh Minh Nguyen, "EO-CNN: An Enhanced CNN Model Trained by Equilibrium Optimization for Traffic Transportation Prediction," Procedia Computer Science, 176:800-809(2020).

[6] Hongtao Li, Kun Jin, Shaolong Sun et al,"Metro passenger flow forecasting though multi-source time-series fusion: An ensemble deep learning approach". Applied Soft Computing, 120(5), 108644(2022).

[7] Wang Fu-jian, Yu Jia-hao, Zhao Jin-huan et al, "Short-term Public Traffic Passenger Volume Forecasting Method Based on Real-time Relevance of Stations," Journal of Transportation Systems Engineering and Information Technology, 21(06), 131-144(2021).