Short-Time Inbound Passenger Flow Prediction of Urban Rail Transit Based on STL-HEOA-BiLSTM

Lizhong Zhu, Xinfeng Yang, Dongliang Wang and Xianglong Huo

Abstract—**Accurate short-time passenger volume prediction can guarantee the efficient scheduling command of urban rail transit. However, the short-time passenger flow of rail transit has the characteristics of nonlinearity and high randomness. In order to improve the prediction accuracy of short-time passenger volume, the Seasonal-Trend decomposition using Loess (STL) method and Human Evolutionary Optimization Algorithm (HEOA) is employed to optimize the bi-directional long short-term memory neural network (BiLSTM). Thus, a combined STL-HEOA-BiLSTM prediction model is proposed. Firstly, the inbound passenger volume along the urban rail transit is classified according to the Pearson correlation number. Secondly, the STL algorithm decomposes different types of short-time passenger flow data into Trend component (T_t), Seasonal component (S_t) and Residual component (R_t). Thirdly, the HEOA optimizes the various types of hyper-parameters of the BiLSTM model. Finally, the optimized BiLSTM model predicts T_t, S_t and R_t individually, and the final prediction value is obtained based on the combination of the three predictions. Three evaluation metrics are used to assess the results and quantify the effectiveness of the combined model. The example analysis shows that the prediction accuracy of the combined STL-HEOA-BiLSTM model surpasses that of the other six common and combined models in forecasting short-duration passenger flow. This experimental result shows the effectiveness, accuracy and applicability of the STL-HEOA-BiLSTM model proposed in this paper. It is demonstrated that the proposed model has a reference value for urban rail transit operators.**

Index Terms—**urban rail transit, short-time passenger flow, passenger flow prediction, composition model**

I. Introduction

Urban rail transport offers benefits such as high passenger capacity, punctuality, economy, and environmental friendliness. As urban areas grow in size and population density increases, it has gradually become the preferred travelling tool for urban residents. In the future, the urban rail transit line network and passenger flow will continue to grow continuously. This phenomenon will increase urban rail transport companies' operational and scheduling difficulties. Therefore, it is significant to accurately predict short-term travel demand for rail transit lines based on historical passenger volume patterns. This behavior will help guide the rational allocation of line operation organization plans and improve the level of service for urban rail travel.

Nowadays, many experts and scholars has researched the problem of passenger volume prediction of urban rail transport. In terms of traditional prediction, Li Jie et al.[1] identified the number of passenger volume prediction steps as the primary factor affecting the Guangzhou-Zhuhai Intercity Line section. By analyzing the time series of the SARIMA model and comparing it with four other forecasting models, it is proved that the SARIMA model can be used for multi-step passenger volume prediction. Chen et al.[2] used Kalman filtering method for OD matrix prediction based on smart card data of Beijing public transport. The research findings indicate that this method can reduce the MAPE of the prediction results to 13.05%. This result is an important reference for urban rail transit enterprise operators.

In terms of machine learning-based prediction, Yuchuan Du et al.[3] examined the volatility of road traffic OD flow based on highway license plate recognition technology and performed a cluster analysis of OD pairs. Woo[4] proposed a kind of framework for OD prediction based on data-driven networks by improving the K-NN algorithm. Jing Liu et al.[5] constructed three OD traffic prediction methods. The study findings show that the prediction method is still effective even in the case of high complexity and insufficient data volume. Shengyue Fang[6] constructed a short-time metro passenger flow prediction model based on XGBoost by analyzing the characteristics of various factors of passenger volume in the Dalian metro. Dawei Chen[7] combined rail passenger volume data with rainfall weather time distribution. And the passenger volume data is predicted by the LSTM model. Ruoyi Li et al.[8] proposed an improved LSTM model after qualitatively analyzing the factors affecting passenger volume. The study results show that the model's prediction accuracy surpasses other methods across all time intervals.

In terms of combined model prediction, Xi Jiang et al.[9] combined machine learning and recursive Bayesian networks to develop a dynamic estimation model for OD on urban rail transit networks. The results indicate that this model effectively estimates OD online. Yang et al.[10] proposed an

Manuscript received January 9, 2024; revised April 25, 2024.

This work was supported in part by Double-First Class Major Research Programs, Educational Department of Gansu Province (No. GSSYLXM-04), Procurement of Technical Services for Intelligent Dispatching System of Guoneng Shuohuang Railway (A3. ZW122291).

Lizhong Zhu is a postgraduate student at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail: 17339901027@163.com).

Xinfeng Yang is a professor at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (Corresponding author, e-mail: xinfengyang@mail.lzjtu.cn).

Dongliang Wang is a senior engineer of China Railway Special Cargo Logistics Co., Ltd.Guangzhou Branch, Guangzhou 510800, China. (e-mail: 13802428695@163.com)

Xianglong Huo is a railway officer trainee of MTR Corporation (Shenzhen) Limited, ShenZhen 518109, China. (e-mail:xlhuo@mtrsz.com.cn)

ELF-LSTM recurrent neural network model suitable for short-time OD prediction in processing time series, leveraging the strengths of LSTM network in analyzing time series data. Xiaoyun Hou et al.[11] integrated models such as MLR, KNN, XGBoost and LSTM to predict short-time OD passenger volume under normalized and non-normalized conditions. Comparing the prediction results reveals that the integrated model outperforms the individual sub-models. Yanjun Huang et al.[12] combined the LSTM model with the K-NN model. The LSTM model accurately predicts inbound passenger volume in real time, and the K-NN model accurately predicts inbound passenger volume destinations in real time. Zhang et al.[13] constructed a combined RF-BiLSTM prediction model by combining the BiLSTM model with the RF algorithm for forecasting short-time passenger volume.

Besides, it is worth mentioning that the application of neural networks is pervasive. Jia et al.[14] achieves face detection by improving convolutional neural networks. Meanwhile, Felipe et al.[15] has designed a novel and robust controller based on neural networks.

Since the traditional LSTM network is sensitive to the initial weight threshold, it is easy to get trapped in a local optimum during training and unable to find a specific solution to the problem. Therefore, a short-time passenger volume combination prediction model by STL-HEOA BiLSTM is proposed.

Among them, the HEAO is a meta-heuristic algorithm proposed by Junbo Lian et al.[16]. It is influenced by the survival and evolutionary processes in human evolution, thus developing a global optimization search algorithm. It is characterized by high efficiency, accuracy and simplicity. Good results have been obtained in path optimization, fault diagnosis, performance detection and parameter optimization. However, the algorithm has been less applied in passenger flow prediction. Therefore, a comprehensive short-time passenger volume prediction method by the STL decomposition method, HEOA and BiLSTM network is proposed in this paper.

The initial step involves identifying the types of passenger flow distribution at various stations. The second step involves decomposing the historical passenger inbound data of different categories at different stations. Following this, the weights and thresholds of each hyper-parameter of the BiLSTM model are optimized by the HEOA. Then, the optimized BiLSTM neural network forecasts the inbound passenger volume data. Ultimately, the results of the predicted data are summarized, compared and analyzed.

The rest of the paper is structured as follows. Section II analyzes the weekly change rule of urban rail traffic. In Section III, the STL time series decomposition method, HEOA and BiLSTM neural network model are introduced. Besides, the STL-HEOA-BiLSTM combination prediction model is constructed accordingly. In Section IV, the combined STL-HEOA-BiLSTM prediction model is used to predict passenger flow by analyzing real cases. Besides, the error comparison analysis is carried out using several types of commonly used passenger flow prediction models. Finally, in Section V, the results of this paper are summarized. The details are shown in Fig. 1.

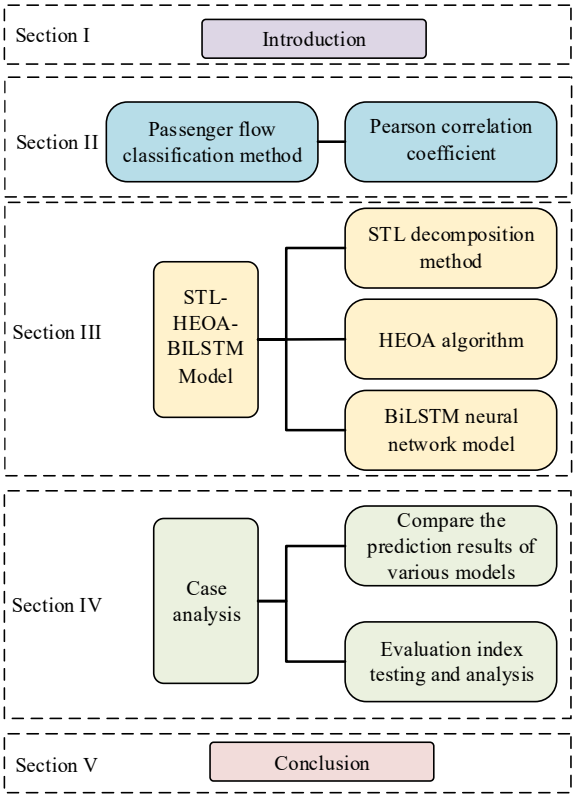


Fig. 1. Research framework diagram

II. Passenger Flow Classification Method

Due to the influence of working days and non-working days, the short-time inbound volume of urban rail transit stations tends to show different time sequence characteristics of passenger volume in one week. At the same time, the degree of correlation between passenger flows is high under the same time sequence characteristics. The changing trends of station passenger flow sequences often differ under different time sequence characteristics.

Various factors influence the short-time passenger inbound volume of urban rail transit. Among them, time series characteristics are one of the most important influencing factors. The Pearson coefficient is often utilized to measure the correlation level between two variables. Therefore, the Pearson coefficient is selected to analyze the degree of correlation between time series characteristics and short-time passenger inbound volume. This behavior facilitates quantifying the degree of influence of time series characteristics on short-time passenger inbound volumes. The Pearson correlation coefficient is introduced to filter the time series features that strongly correlate with short-time passenger inbound volumes. The initial passenger volume time series are analyzed and classified into types, thus effectively improving the prediction accuracy. Equation (1) presents the formula for the Pearson coefficient.

$$R_{XY} = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}}. \quad (1)$$

In equation (1), R_{XY} is the Pearson's correlation coefficient, N is the length of the sample. X_i and Y_i are the actual sample values, $\bar{X}$ and $\bar{Y}$ are the mean values of the actual sample.

The range of R_{XY} is [-1,1]. R_{XY} is positive for positive correlation and negative for negative correlation. R_{XY} is closer to 1, which means a higher correlation.

III. Model Construction and Evaluation Indicators

If the original short-time inbound volume sequence data of metro stations are directly used for passenger flow prediction, the noise and random fluctuation of the initial passenger volume data sequence will interfere with the passenger flow prediction. Therefore, the STL method first decomposes the initial passenger volume time sequence to reduce noise interference. Then, the combined HEOA-BiLSTM prediction model is used to predict passenger flow.

A. STL method

The STL method[17] relies on robust locally weighted regression for data decomposition. It can decompose time series data, like the passenger volume of urban rail transit. The STL method consists of two processes: inner loop and outer loop.

The inner loop is mainly responsible for decomposing the T_t and S_t. The outer loop is primarily accountable for assigning weights to the next inner loop based on the remaining data to minimize the influence of outliers. For the input data, the data interval of the sample is intercepted. Besides, regression analysis is performed using weighted least squares. The local regression model is obtained by making the values near the estimation points relatively large weights. This process is repeated until a regression curve is obtained. The inbound passenger volume exhibits periodic patterns over time. The STL method decomposes the input inbound passenger flow sequence to get the inbound passenger volume data's T_t, S_t and R_t. Equation (2) is the computational formula for its model decomposition algorithm:

$$Y_t = T_t + S_t + R_t, t = 1, 2, \cdots, N. \tag{2}$$

B. BiLSTM

The LSTM model [7] is a Recurrent Neural Network (RNN) model variant. The LSTM model can effectively solve the gradient vanishing and gradient explosion problems in RNN. The LSTM model differs from the RNN model in that it incorporates a memory unit C_t, so the model has a memory function. As shown in Fig. 2, this memory unit has three gates: forgetting, input and output gate.

The forgetting gate decides whether to keep or forget the information previously stored in the memory unit. The input gate computes the current input state of the unit by the last output and current input. The output gate determines the output value of the hidden layer in the current state.

The specific process of updating the state of each gate and cell in the LSTM model is shown in equations (3)-(8):

1) Oblivion gate:

$$f_t = \sigma \cdot (w_f \cdot [h_{t-1}, x_t] + b_f). \tag{3}$$

2) Input gate:

$$i_t = \sigma \cdot (w_i \cdot [h_{t-1}, x_t] + b_i). \tag{4}$$

3) Unit status:

$$\tilde{C}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c), \tag{5}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t. \tag{6}$$

4) Output gate:

$$O_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o), \tag{7}$$

$$h_t = O_t \cdot \tanh(C_t). \tag{8}$$

In equations (3)-(8), C_{t-1} represents the cell state from the previous moment, while C_t represents the cell state from the current moment. $\tilde{C}_t$ is the updated value of the cell state from the current moment (the alternative value of C_t), w is the weight coefficients matrix of each gate, b is the bias of each gate, x_t is the input value from the current moment, h_{t-1} is the value of the hidden layer passed from the previous moment, h_t is the output value of the hidden layer from the current moment.

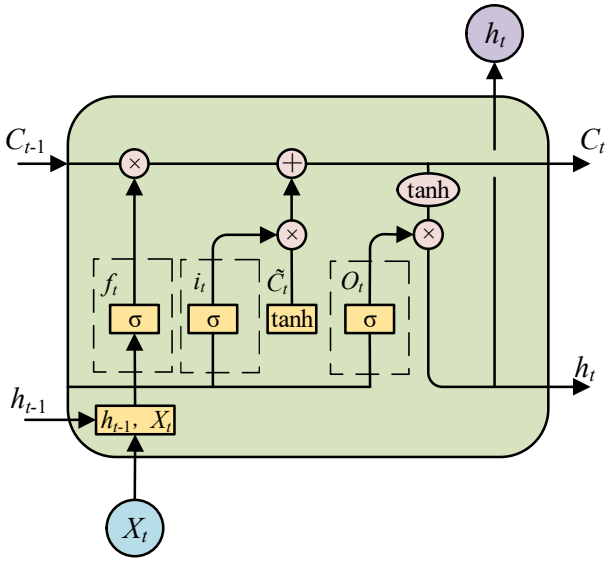


Fig. 2. Structure of LSTM cellular unit

The BiLSTM neural network[18] incorporates forward and reverse data feature information by utilizing the LSTM structure with a reverse input sequence, as shown in Fig. 3.

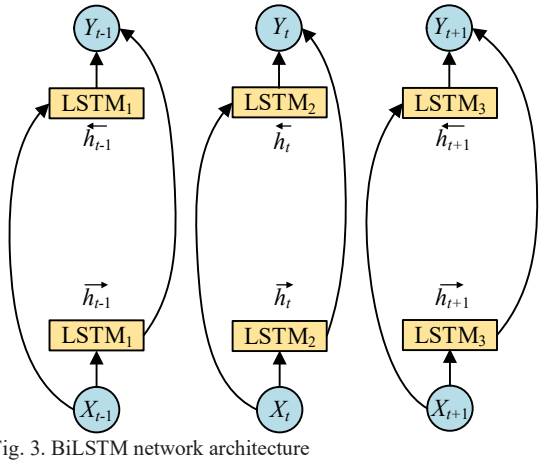


Fig. 3. BiLSTM network architecture

The BiLSTM model combines the weighted predictions from both directions to obtain the final prediction value. This feature allows the model to adapt more quickly and efficiently to fluctuations in the data. The mathematical expressions are shown in equations (9)-(11):

$$\overrightarrow{h_t} = f(w_1 x + w_2 \overrightarrow{h_{t-1}}), \tag{9}$$

$$\overleftarrow{h_t} = f(w_3 x + w_5 \overleftarrow{h_{t+1}}), \tag{10}$$

$$O_t^{'} = g(w_4\overrightarrow{h_t} + w_6\overleftarrow{h_t}). \quad (11)$$

In equations (9)-(11), $\overrightarrow{h_t}$ is the forward output of the hidden layer from the moment of t-1, $\overleftarrow{h_t}$ is the reverse output of the hidden layer from the moment of t+1, $O_t^{'}$ is the final output, f is the LSTM unit function, g is the ReLU function. Besides, $w_1 \sim w_6$ are the corresponding weights.

C. Human Evolutionary Optimization Algorithm

The Human Evolutionary Optimization Algorithm (HEOA) [16] is a meta-heuristic algorithm that simulates human survival and evolutionary processes. The algorithm is rooted in natural selection and human adaptation to highlight the human capacity to discover the best solutions in intricate settings. The HEOA divides the human evolution into two main stages: the stage of human inquiry and the stage of human growth.

In the stage of human inquiry, the algorithm is initialized using Logical Chaos Mapping to mimic human curiosity and exploration. In the human growth phase, the algorithm further explores and develops solutions by using jumping strategies. These strategies are designed to simulate human adaptability and innovation in complex environments.

In order to replicate the tumultuous early phase of human evolution, the HEOA establishes the overall human size as N, Max_{iter} as the maximum number of iterations, and the search space boundaries as lower bound (lb) and upper bound (ub). The formula for initializing Logistic Chaos Mapping can be expressed as equation (12):

$$x_i = \alpha \cdot x_{i-1} \cdot (1 - x_{i-1}), 0 \le x_0 \le 1, i = 1,..., N, \alpha = 4. \quad (12)$$

In equations (12), x_i is the value at iteration i and x_{i-1} is the previous iteration value. Equation (13) is meant to map the chaotic value x_i to the total search space:

$$X_i^0 = lb + (ub - lb) \cdot x_i. \quad (13)$$

After that, individual humans began to explore. This evolutionary process is the stage of human inquiry when the number of iterations is less than or equal to one-fourth of the maximum. During the stage of human inquiry, individuals tend to use a consistent search strategy when encountering unexplored areas with limited knowledge. The specific process is shown in equation (14):

$$\begin{aligned} X_i^{t+1} &= (X_{mean}^t - X_{best}) \cdot floor(\frac{rand}{f_{jump}}) \cdot f_{jump} \\ &+ X_{best} \cdot (1 - \frac{t}{Max_{iter}}) + \beta \cdot Levy(dim). \end{aligned} \quad (14)$$

In equation (14), t denotes the current iteration number, dim indicates the dimension of the problem's dimension and the number of variables, X_i^{t+1} denotes the position of subsequent updates, X_{best} represents the most optimal position currently explored, X^t_{mean} represents the population's average position, $floor$ denotes rounding down, $Levy$ denotes the $Levy$ distribution, f_{jump} is a jump coefficient, $rand$ is a random number in [0,1], and β is the adaptive function.

The expression for the average position average position X^t_{mean} is shown in equation (15):

$$X_{mean}^t = \frac{1}{N}\sum_{k=1}^{N} X_k^t. \quad (15)$$

Equation (16) is utilized to compute the value of β by considering the number of iterations and the weight of the current position. This adaptive tuning of parameters plays a crucial role in achieving stability and convergence of the algorithm. By adjusting the parameters in this manner, the algorithm's global search capability is improved, enhancing overall performance.

$$\beta = 0.2(1 - \frac{t}{Max_{iter}}) \cdot (X_i^t - X_{mean}^t). \quad (16)$$

$Levy$ distribution has strong noise resistance and can effectively deal with missing data or anomalies[19]. Using $Levy$ distribution to simulate the behavioral characteristics of human individuals during exploration is essential for improving the stability and accuracy of the algorithm. The expression of $Levy$ distribution is shown in equation (17):

$$\begin{cases} Levy(D) = \dfrac{\mu \cdot \sigma}{|v|^{\frac{1}{\gamma}}} \\ \mu \sim N(0, D) \\ v \sim N(0, D) \\ \sigma = \left(\dfrac{\Gamma(1+\gamma) \cdot \sin(\frac{\pi\gamma}{2})}{\Gamma(\frac{1+\gamma}{2}) \cdot \gamma \cdot 2^{\frac{\gamma-1}{2}}}\right)^{\frac{1}{\gamma}} \end{cases}. \quad (17)$$

The jump strategy is a vital exploration method. It can help human individuals disperse to new positions within the search space, thereby avoiding being trapped in local optimal solutions. In addition, the jumping strategy is able to memorize specific local information. This behavior ensures that the search area is increased without decreasing the accuracy and efficiency of the search. f_{jump} is the jump coefficient, meaning the quantification of the degree of the jump strategy. The expression is shown in equation (18):

$$f_{jump} = \frac{(lb(1) - ub(1))}{\delta}, \delta \in [100, 2000]. \quad (18)$$

This evolutionary process is defined as the stage of human growth when the number of iterations exceeds one-fourth of the maximum. At the stage of human growth, humans in society are classified into four roles by the algorithm: leader, explorer, follower and loser.

In the evolutionary process, human individuals with fitness values in the top 40 per cent are designated as leaders. During the stage of human growth, leaders are responsible for exploring the human growth space that has already gained advantages. The specific evolutionary process is shown in equation (19):

$$X_i^{t+1} = \begin{cases} \omega \cdot X_i^t \cdot \exp(\dfrac{-t}{rand \cdot Max_{iter}}), R < A \\ \omega \cdot X_i^t + Rn \cdot ones(1, dim), R \ge A \end{cases}. \quad (19)$$

In equation (19), Rn denotes a random number following a normal distribution. $ones(1, dim)$ represents the generation of a unit row vector containing dim elements. R is a random number in [0,1]. A is the evaluation value that indicates a judgment of complexity, with a value of 0.6. ω is the coefficient of the easiness of knowledge acquisition, which decreases as development progresses. The calculation of ω *is* shown in equation (20):

$$\omega = 0.2\cos(\frac{\pi}{2}\cdot(1-\frac{t}{Max_{iter}})). \quad (20)$$

In the evolutionary process, human individuals with fitness values in the top 40 to 80 per cent are designated as explorers. During the stage of human growth, these explorers are responsible for entering the unknown space of human growth for exploration. The specific process is shown in equation (21):

$$X_i^{t+1} = Rn\cdot\exp(\frac{X_{worst}^t - X_i^t}{i^2}). \quad (21)$$

In equation (21), X^t_{worst} denotes the individual in the human society with the lowest fitness value at iteration t.

In the evolutionary process, human individuals with fitness values in the last 10 to 20 per cent are designated as followers. During the stage of human growth, these followers adhere to and follow leaders' leadership. Equation (22) illustrates the specific process:

$$X_i^{t+1} = X_i^t + \omega\cdot Rd\cdot(X_{best}^t - X_i^t). \quad (22)$$

In equation (22), X^t_{best} represents the location of the individual with the highest fitness value in society at iteration t, and Rd represents a random number in the range [1, dim].

In the evolutionary process, human individuals with fitness values in the bottom 10 per cent are designated as losers. During the stage of human growth, these losers are eliminated because they cannot adapt to the environment. The corresponding population gap will be filled through reproduction in an environment conducive to human growth. The specific process of population restoration and replenishment is shown in equation (23):

$$X_i^{t+1} = X_{best} + (X_{best} - X_i^t)\cdot Rn. \quad (23)$$

D. Predictive Model Evaluation Indicators

In order to confirm the validity of prediction method and the accuracy of forecast results, the forecast errors are evaluated with commonly used statistical indicators. Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Correlation Coefficient (R-squared, R^2) are chosen as evaluation indices to quantify the prediction results. The specific expressions of the three evaluation indicators are shown in equations (24)-(26):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}, \quad (24)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|, \quad (25)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(\bar{y}_i - y_i)^2}. \quad (26)$$

E. Prediction Model Based on STL-HEOA-BiLSTM

The complexity of urban rail transit's original passenger volume time series can disrupt neural network predictions. In addition, the neural network hyper-parameters determined solely based on experience can also seriously affect the precision of the passenger flow prediction model. Therefore, the STL algorithm decomposes passenger flow time series data and is combined with the HEOA to dynamically solve the optimal values of some hyper-parameters(number of hidden units, maximum training period, initial learning rate, $L2$ parameter) in the BiLSTM neural network. Finally, a combined STL-HEOA-BiLSTM model is constructed to accurately predict short-time inbound passenger volume in urban rail transit.

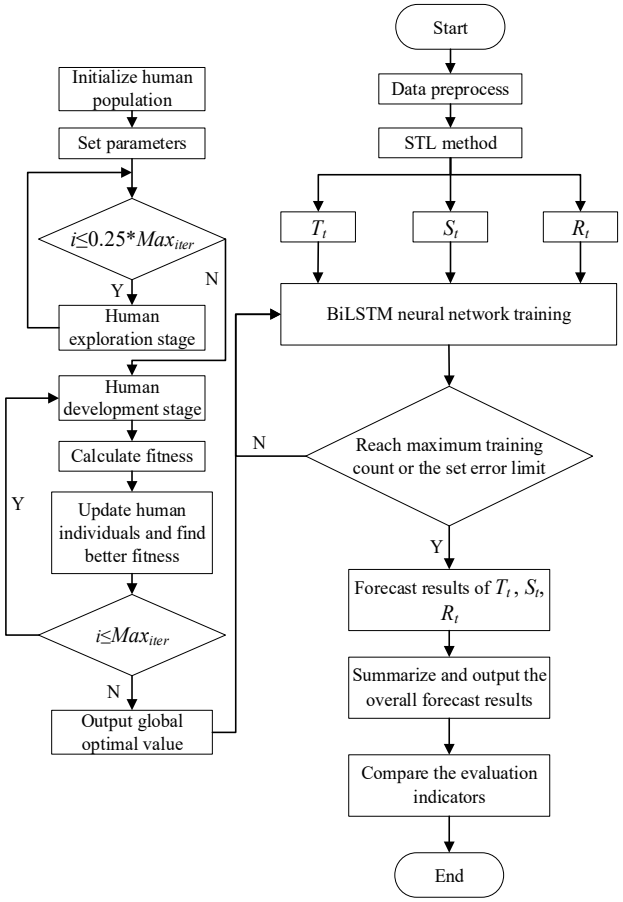


Fig. 4. STL-HEOA-BiLSTM model prediction flowchart

The Fig. 4. illustrates the specific prediction process, which involves the following steps.

Step 1: Data decomposition. The original time series data of inbound passenger flow is decomposed using the STL algorithm to obtain its corresponding T_t, S_t and R_t. These components will serve as the original input data for the BiLSTM.

Step 2: Building a BiLSTM neural network. Initialize the parameters, including the batch size, the number of hidden layers and hidden layer neurons in the BiLSTM.

Step 3: Randomly initialize the parameters in the HEOA. Some hyper-parameters in the BiLSTM network are used as the optimization object. The parameters of initialized population size N, the maximum iterations Max_{iter} and the particle dimension dim of the algorithm are set up.

Step 4: From equations (12)-(23), the stage of human inquiry and stage of human growth are entered successively according to the number of iterations. According to fitness values, the number and position of human leaders, explorers, followers and losers are iteratively updated.

Step 5: Construct the STL-HEOA-BiLSTM prediction model. The HEOA algorithm is utilized to find the optimal hyper-parameter values in the BiLSTM neural network in order to construct the combined prediction model. If the iteration termination condition is met, the optimal values of

the hyper-parameters in the BiLSTM neural network and the final prediction results are output. If it is unsatisfied, make t=t+1 and repeat Steps 3 to 5.

Step 6: Evaluate the prediction results. Evaluation metrics are utilized to assess the predictive performance of the STL-HEOA-BiLSTM combination model. Accordingly, the evaluation results are compared and analyzed with the remaining common and combined models.

IV. Example Analyzes

The raw data of passenger flow inbound at the stations along Hangzhou Metro Line 1 from 2 to 25 January 2019 is selected as the analysis case. The extensive passenger flow data and complex changing rules make it a suitable research subject.

A. Classification of passenger flow

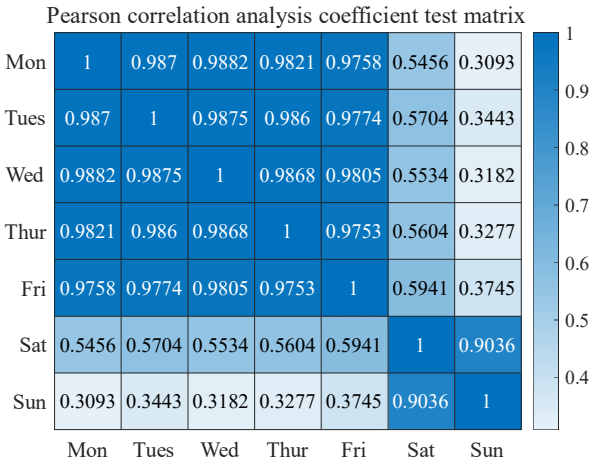
In order to investigate the correlation between the average daily inbound volume in one week, the Pearson correlation number is chosen to carry out correlation analysis on the inbound passenger volume data. On this basis, the inbound passenger volume is classified according to different time series under different time series characteristics of passenger volume. The results show that the time sequence characteristics of passenger flow at stations along Hangzhou Metro Line 1 can be categorized into two main groups. In order to save space, two typical stations, Station 13 and Station 18, are selected for the correlation analysis presentation. The specific results are presented in Table I and Fig. 5.

TABLE I
Classification of Station Passenger Flow Type

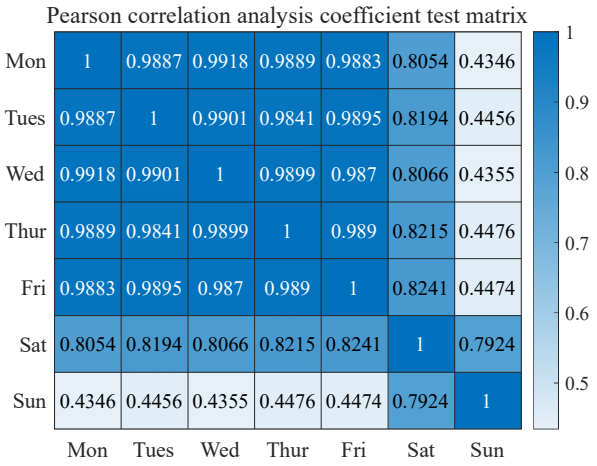
Stations	Characteristics of passenger flow		Passenger flow type
2-9, 10-14, 16, 17, 19, 20, 24-27, 33	Category 1	Monday to Friday	1
		Saturday to Sunday	2
1, 15, 18, 21-23, 28-32	Category 2	Monday to Friday	3
		Saturdays	4
		Sunday	5

From Table I and Fig. 5, the average daily inbound volume at the station displays distinct timing characteristics regarding passenger flow. In concrete terms, the Pearson correlation coefficients of average daily passenger flow under the first type of temporal characteristics are greater than 0.9 among Monday, Tuesday, Wednesday, Thursday, and Friday, as well as between Saturday and Sunday. Furthermore, the Pearson correlation coefficients for both Saturdays and Sundays in relation to Monday through Friday are less than 0.6. The Pearson coefficient greater than 0.9 is set as a significant correlation. Thus, it is concluded that there are two distinct types of passenger flow in the station during the week: the first type occurs from Monday to Friday, and the second type occurs from Saturday to Sunday based on first category time sequence characteristics.

Similarly, the Pearson correlation coefficients of average daily passenger flow under the second type of temporal characteristics are greater than 0.9 among Monday, Tuesday, Wednesday, Thursday, and Friday. In addition, the Pearson correlation coefficients for Saturday and Sunday with the rest of the days, including each other, are less than 0.85. The Pearson correlation coefficient of less than 0.85 is set as a non-significant correlation. Meanwhile, it is concluded that there are three distinct types of passenger flow in the stations during the week: the third type occurs from Monday to Friday, the fourth type on Saturdays, and the fifth type on Sunday based on second category time sequence characteristics.



a) Correlation between average daily inbound traffic over a one-week period at Station 13



b) Correlation between average daily inbound traffic over a one-week period at Station 18

Fig. 5. Correlation plot between average daily inbound volumes over a week at typical stations

B. Passenger flow data decomposition

The passenger volume data of all stations is counted in 15-min intervals daily. The operating hours will be 06:00 to 23:30. Hangzhou Metro Line 1 has 33 stations. Therefore, there is a total of 55440 passenger flow data. In this case, the ratio of the training set to the test set is 9:1.

From Table I, the inbound passenger volume of each station along Hangzhou Metro Line 1 is divided into five categories. The STL method decomposes the data of different types of inbound passenger volume. To save space, two types of typical stations, stations 13 and 18, are selected as the results to be displayed, as shown in Figs 6 to 10. Among them, $T(t)$ reflects the general trend of passenger flow in a cycle, $S(t)$ reflects the fluctuation of passenger flow in one day, and $R(t)$ reflects the overall randomness of passenger flow.

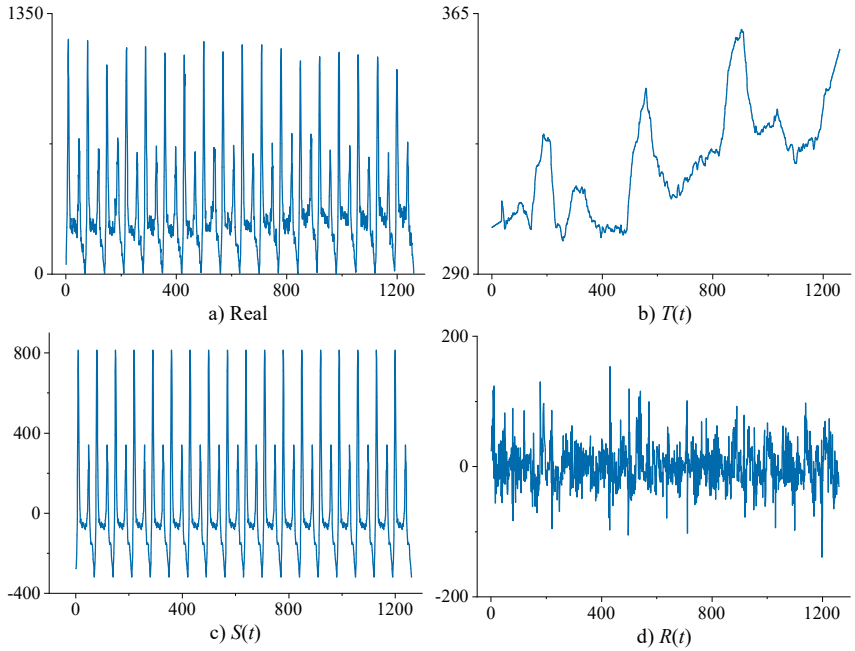


Fig. 6. STL decomposition result for the passenger flow type 1 at Station 13

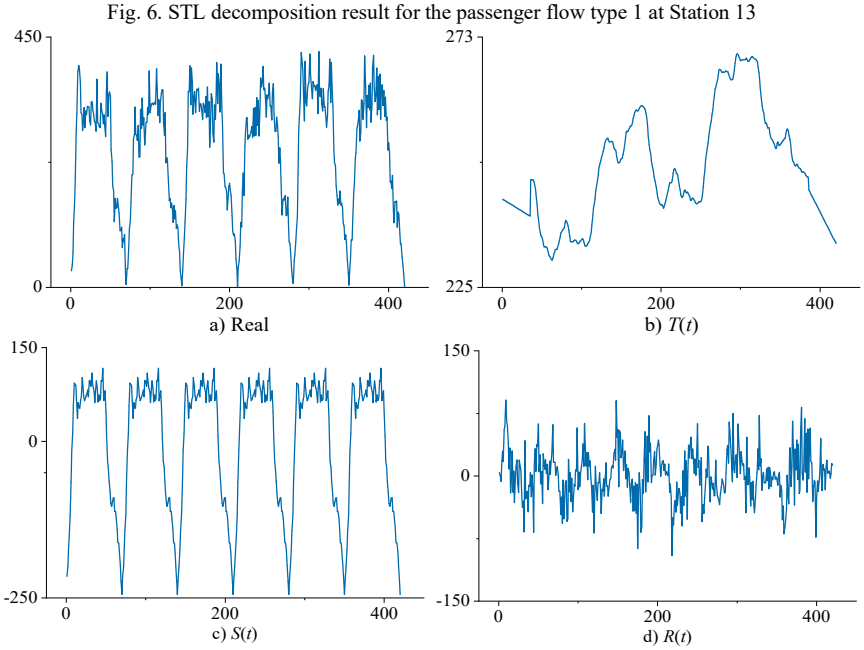


Fig. 7. STL decomposition result for the passenger flow type 2 at Station 13

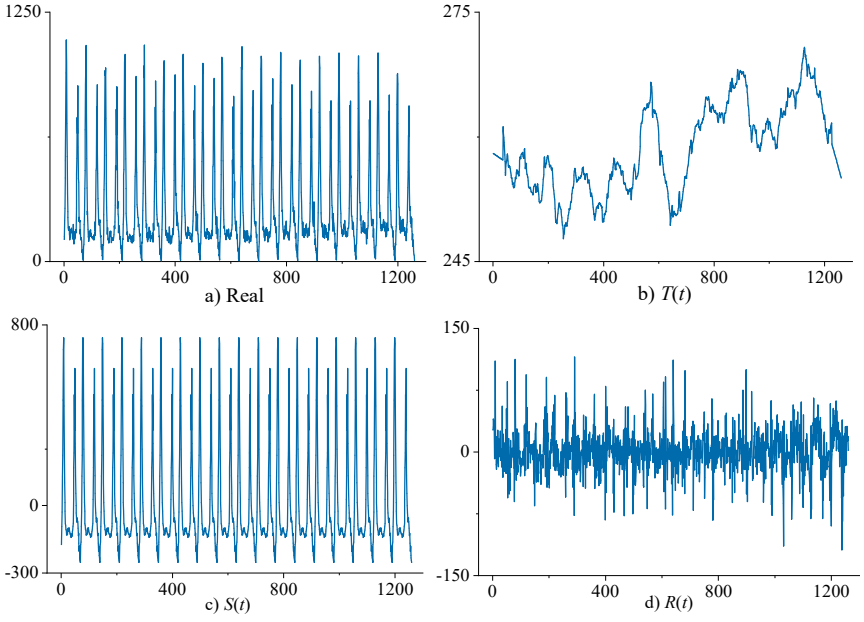


Fig. 8. STL decomposition result for the passenger flow type 3 at Station 18

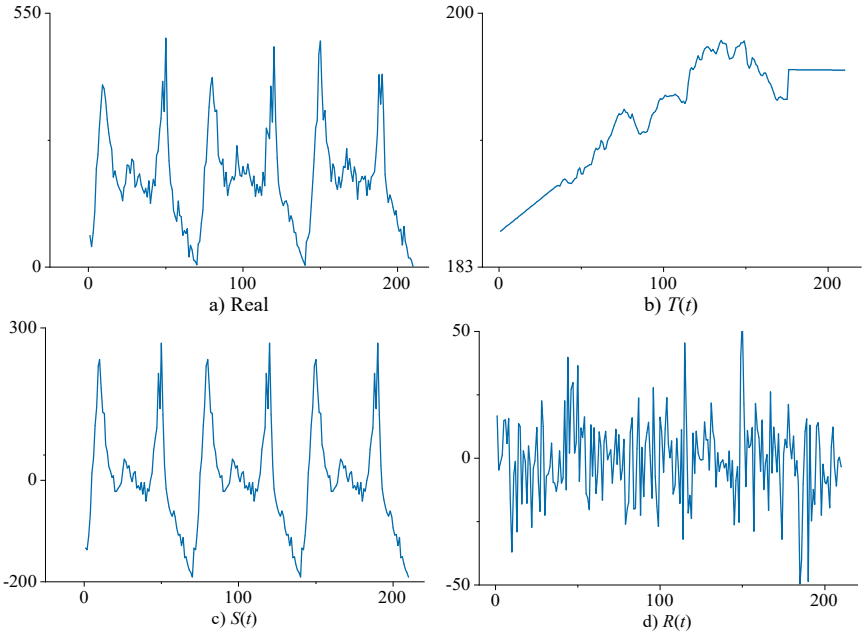


Fig. 9. STL decomposition result for the passenger flow type 4 at Station 18

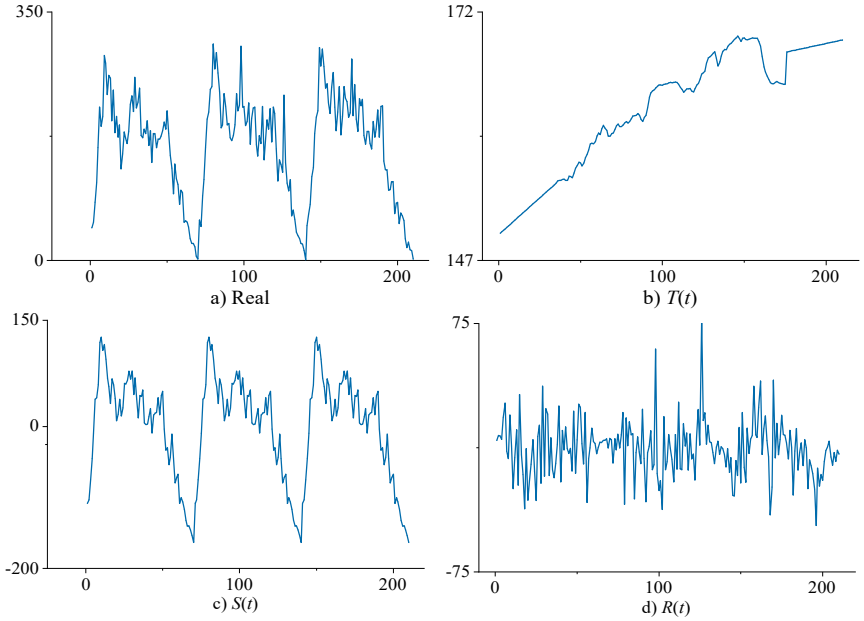
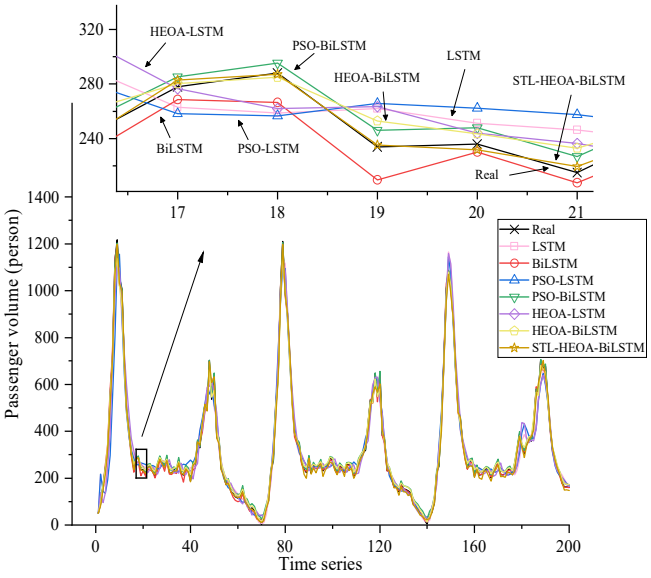


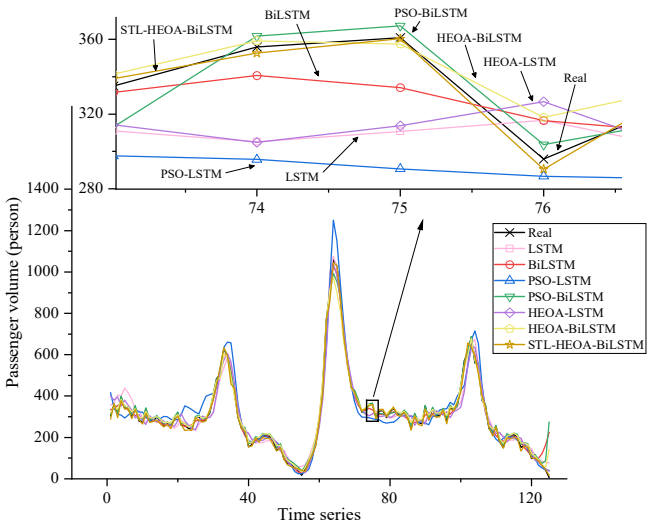
Fig. 10. STL decomposition result for the passenger flow type 5 at Station 18

C. *Analysis of forecast results*

Various types of inbound passenger flow at Station 13 and Station 18 are chosen to assess the prediction accuracy of the models. The STL-HEOA-BiLSTM model predicts different types of inbound passenger flows at station 13 and station 18, respectively. Furthermore, the prediction results are compared with those of LSTM, PSO-LSTM, BiLSTM, PSO-BiLSTM, HEOA-LSTM and HEOA-BiLSTM models. In this case, the LSTM and BiLSTM models have one neuron for input and output layers. Adam optimizer is selected for training. Both the number of iterations and hidden units are 50. The respective algorithms iteratively optimize each hyper-parameter in the PSO-LSTM, PSO-BiLSTM, HEOA-LSTM and HEOA-BiLSTM models. Besides, the remaining parameters are consistent with those in the LSTM models. The first 200 data points of training set and all of test set data are selected for presentation to make the results more transparent and intuitive and to save space. The results are depicted in Fig. 11 to Fig. 15.

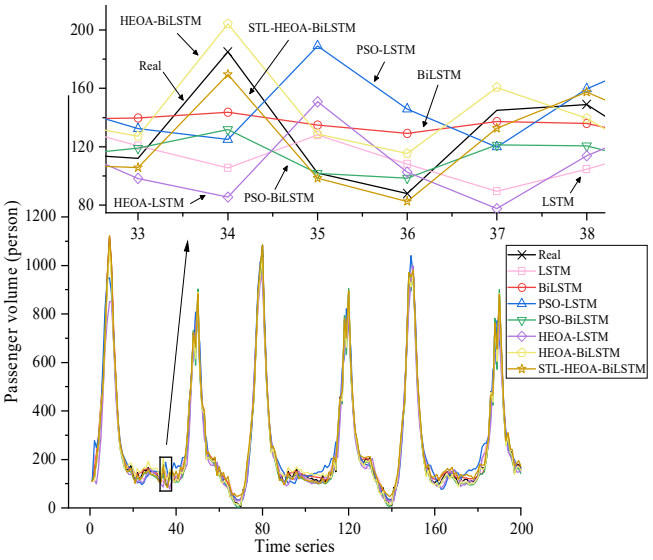


(a) Training set

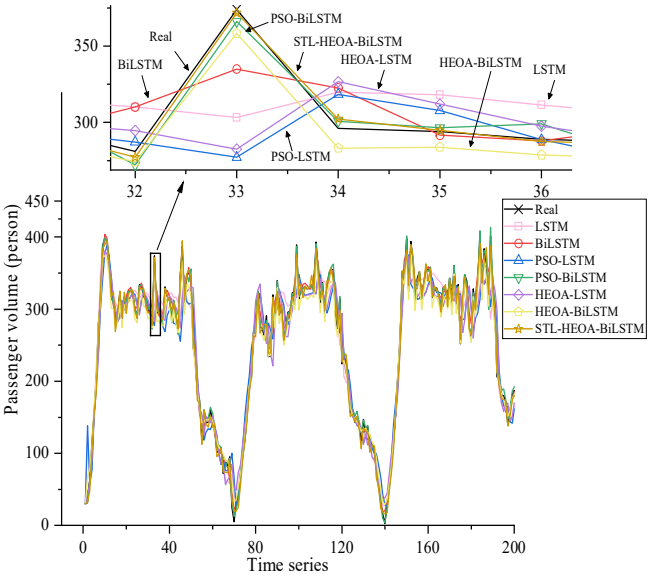


(b) Test set

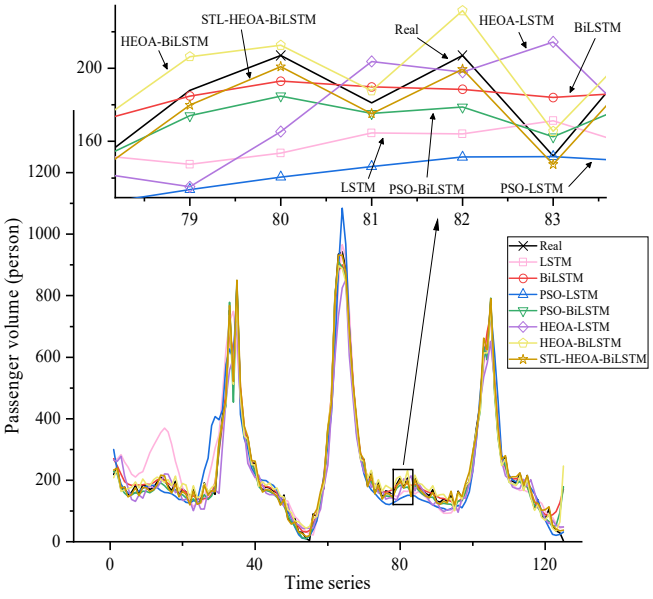
Fig. 11. Comparison of forecast results for site 13 Type 1 passenger flow



(a) Training set

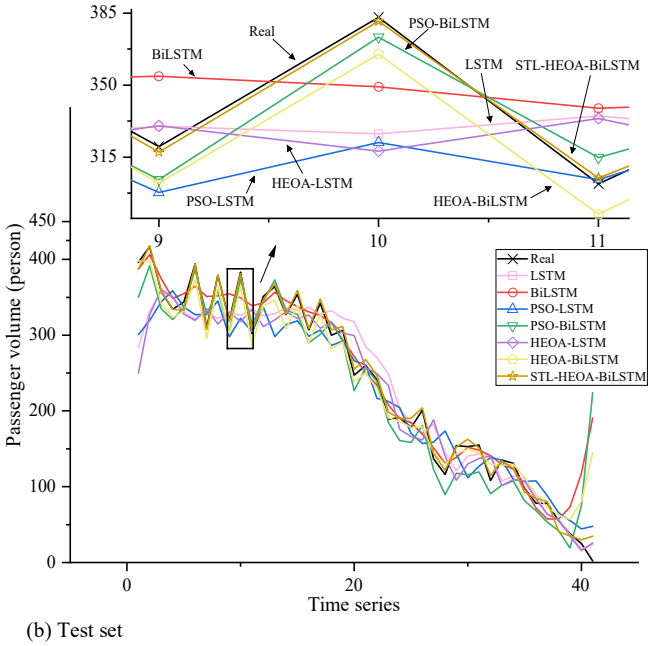


(a) Training set



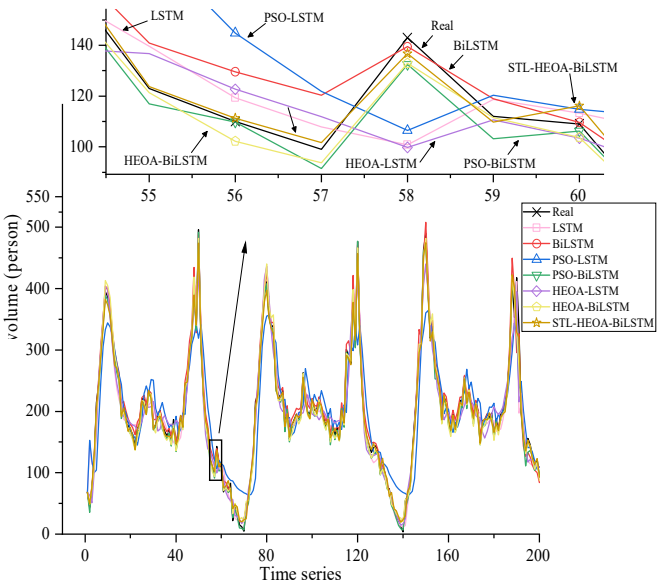
(b) Test set

Fig. 13. Comparison of forecast results for site 18 Type 3 passenger flow

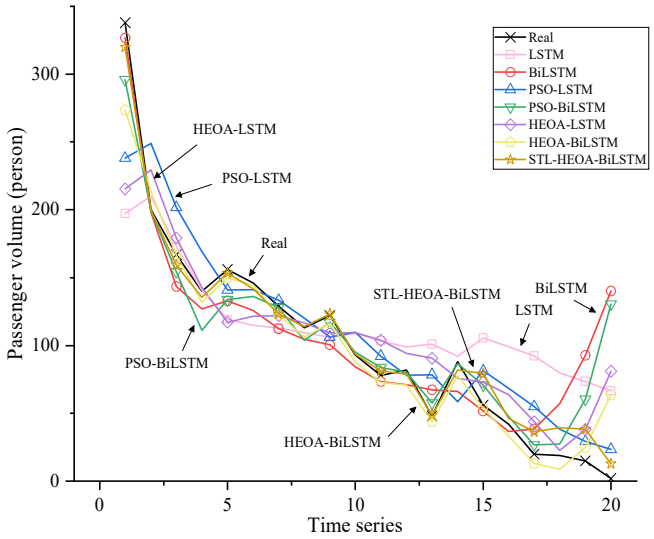


(b) Test set

Fig. 12. Comparison of forecast results for site 13 Type 2 passenger flow

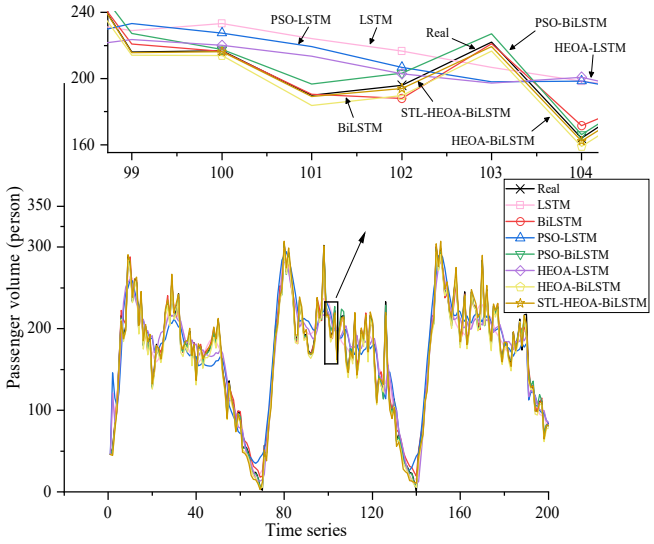


(a) Training set

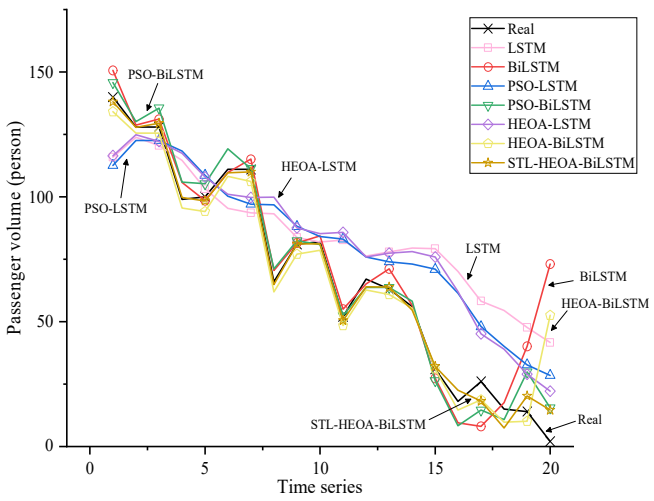


(b) Test set

Fig. 14. Comparison of forecast results for site 18 Type 4 passenger flow



(a) Training set



(b) Test set

Fig. 15. Comparison of forecast results for site 18 Type 5 passenger flow

In order to assess the accuracy of the combined model, RMSE, MAE and R^2 are selected to calculate, compare and analyze the prediction errors of each model. The specific results are presented in Table II, Table III, Fig. 16 and Fig. 17. Among them, the line and bar graph sections indicate the values of R^2 and RMSE in Fig. 16 and Fig. 17, respectively.

TABLE II
EVALUATION INDEX VALUE OF EACH MODEL PREDICTION RESULT IN THE TRAINING SET

Passenger flow type	predictive modelling	RMSE	MAE	R^2
1	LSTM	43.0018	30.0795	0.9680
	BiLSTM	19.0542	14.1300	0.9937
	PSO-LSTM	41.8102	30.8066	0.9697
	HEOA-LSTM	38.2645	28.0691	0.9746
	PSO-BiLSTM	18.4649	14.2395	0.9941
	HEOA-BiLSTM	11.4484	10.2611	0.9977
	STL-HEOA-BiLSTM	**9.9579**	**7.4882**	**0.9983**
2	LSTM	31.7593	23.9486	0.9145
	BiLSTM	15.9102	12.3731	0.9785
	PSO-LSTM	31.3109	24.0502	0.9169
	HEOA-LSTM	27.8794	21.1859	0.9341
	PSO-BiLSTM	13.6645	11.8589	0.9842
	HEOA-BiLSTM	9.2646	7.3147	0.9927
	STL-HEOA-BiLSTM	**6.4946**	**4.4096**	**0.9964**
3	LSTM	57.9592	39.8583	0.9418
	BiLSTM	25.0905	17.9633	0.9891
	PSO-LSTM	50.0435	33.6327	0.9566
	HEOA-LSTM	47.4381	34.0555	0.9610
	PSO-BiLSTM	20.0807	17.5174	0.9930
	HEOA-BiLSTM	18.8976	15.5033	0.9938
	STL-HEOA-BiLSTM	**10.9599**	**9.1262**	**0.9979**
4	LSTM	45.6549	33.5681	0.7880
	BiLSTM	14.5913	10.0192	0.9786
	PSO-LSTM	26.5719	19.2978	0.9289
	HEOA-LSTM	24.8234	19.4993	0.9379
	PSO-BiLSTM	10.7312	6.6887	0.9884
	HEOA-BiLSTM	10.9578	7.5144	0.9879
	STL-HEOA-BiLSTM	**9.0012**	**6.1190**	**0.9918**
5	LSTM	28.2225	22.2504	0.8359
	BiLSTM	7.1058	5.4504	0.9896
	PSO-LSTM	25.7518	19.4163	0.8634
	HEOA-LSTM	22.6225	17.7491	0.8946
	PSO-BiLSTM	6.3489	5.6336	0.9917
	HEOA-BiLSTM	5.9491	4.9790	0.9927
	STL-HEOA-BiLSTM	**3.7733**	**2.5313**	**0.9971**

From Table II, Fig. 15 and Fig. 16, it can be seen that the BiLSTM model predicts better than the unidirectional LSTM model in the training set. This phenomenon is due to the fact that BiLSTM model has a superior bidirectional way of processing time series data.

Besides, a single model can only learn the features of a single model and cannot combine the advantages of other models. This drawback results in low prediction accuracy.

As shown in Table II, the prediction errors of the PSO-BiLSTM model are generally smaller than those of the BiLSTM model under the same type of passenger flow in the training set. At the same time, the prediction precision of the STL-HEOA-BiLSTM model is generally higher than that of the HEOA-BiLSTM model under the same conditions in the training set.

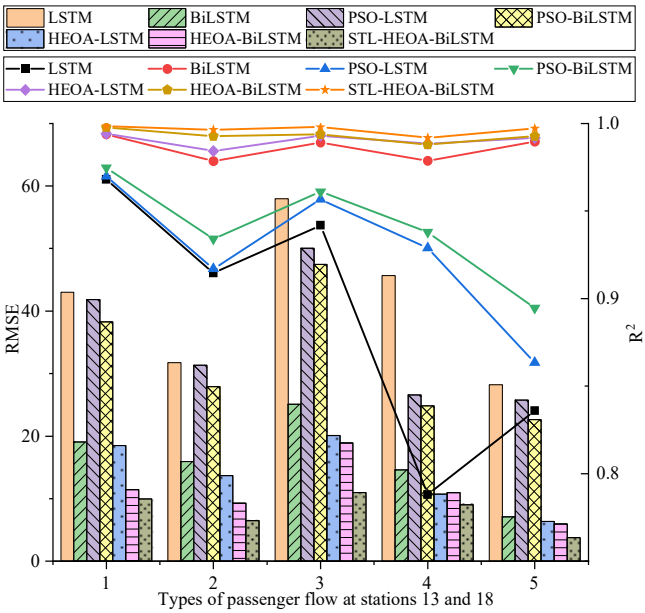


Fig. 16. Tabular chart of RMSE and R^2 evaluation indicator values for each model prediction result in the training set

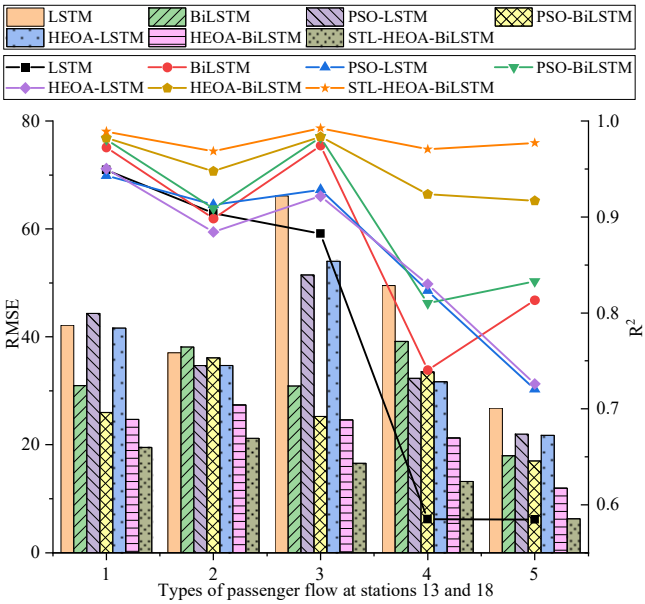


Fig. 17. Tabular chart of RMSE and R^2 evaluation indicator values for each model prediction result in the test set

Therefore, it can be concluded that among the commonly used prediction models, the prediction accuracy of combined prediction models tends to be better than that of single prediction models. In addition, different combinatorial prediction models have different prediction accuracies. This phenomenon is related to the types of algorithms used in the combined prediction models.

As can be seen from Table II, Table III, Fig. 16 and Fig. 17, the STL-HEOA-BiLSTM combination model constructed in this paper outperforms the common models and the combination model in all the indexes, both in the training set and the test set. In other words, the prediction errors of the STL-HEOA-BiLSTM model for different types of inbound passenger flows are minimal in both the training and test sets. For example, for passenger flow type 1, the RMSE value of the combined STL-HEOA-BiLSTM model is 9.9579, the MAE value is 7.4882, and the R^2 value is 0.9983 in the training set. Besides, for the passenger flow type 5, the RMSE value of the combined STL-HEOA-BiLSTM model is 6.2794, the MAE value is 5.1373, and the R^2 value is 0.9772 in the test set. In conclusion, the prediction accuracy of the combined STL-HEOA-BiLSTM model is often higher than that of other models.

TABLE III
EVALUATION INDEX VALUE OF PREDICTION RESULT IN THE TEST SET

Passenger flow type	predictive modelling	RMSE	MAE	R^2
1	LSTM	42.1614	32.5850	0.9491
	BiLSTM	30.9646	20.3998	0.9725
	PSO-LSTM	44.3313	35.0920	0.9431
	HEOA-LSTM	41.6500	31.5547	0.9503
	PSO-BiLSTM	25.9810	10.7903	0.9807
	HEOA-BiLSTM	24.7099	15.5563	0.9825
	STL-HEOA-BiLSTM	**19.5087**	**13.7091**	**0.9891**
2	LSTM	37.0895	28.0254	0.9039
	BiLSTM	38.1345	22.6100	0.8984
	PSO-LSTM	34.6552	27.3591	0.9127
	HEOA-LSTM	34.6880	26.3895	0.8844
	PSO-BiLSTM	36.1036	21.5939	0.9089
	HEOA-BiLSTM	27.3717	16.4976	0.9477
	STL-HEOA-BiLSTM	**21.1886**	**10.6051**	**0.9686**
3	LSTM	66.1309	46.6022	0.8826
	BiLSTM	30.9049	21.7842	0.9744
	PSO-LSTM	51.4637	35.4813	0.9285
	HEOA-LSTM	53.9968	32.5258	0.9217
	PSO-BiLSTM	25.2360	18.1061	0.9829
	HEOA-BiLSTM	24.6122	16.8870	0.9837
	STL-HEOA-BiLSTM	**16.5744**	**14.0582**	**0.9926**
4	LSTM	49.5181	36.9472	0.5849
	BiLSTM	39.1566	24.3665	0.7404
	PSO-LSTM	32.3292	24.8209	0.8230
	HEOA-LSTM	31.6827	18.3682	0.8301
	PSO-BiLSTM	33.5017	17.5084	0.8100
	HEOA-BiLSTM	21.2477	12.4591	0.9236
	STL-HEOA-BiLSTM	**13.1855**	**9.9638**	**0.9706**
5	LSTM	26.7775	22.0628	0.5847
	BiLSTM	17.9579	8.9538	0.8132
	PSO-LSTM	21.9794	18.7159	0.7202
	HEOA-LSTM	21.7547	17.5062	0.7259
	PSO-BiLSTM	16.9817	6.1988	0.8330
	HEOA-BiLSTM	11.9753	6.0328	0.9169
	STL-HEOA-BiLSTM	**6.2794**	**5.1373**	**0.9772**

To represent the prediction results more intuitively and succinctly, the STL-HEOA-BiLSTM combined model is merged with the various types of passenger flow prediction results for Station 13 and Station 18. In addition, the prediction results are compared with the original passenger flow data, illustrated in Fig. 18 and Fig. 19.

From Fig. 20 that all the stations have RMSE values below 35 and MAE values below 20. 33.3% of the stations have RMSE values between 20-30, 36.3% have RMSE values between 10-20, and 21.2% have RMSE values between 3-10. Similarly, 24.2% of the stations have MAE values between

10-15, 39.3% have RMSE values between 5-10 and 27.2% have RMSE values between 1-5. This result shows that the STL-HEOA-BiLSTM model predictions have less deviation from the actual data. Therefore, the STL-HEOA-BiLSTM model is practical and applicable.

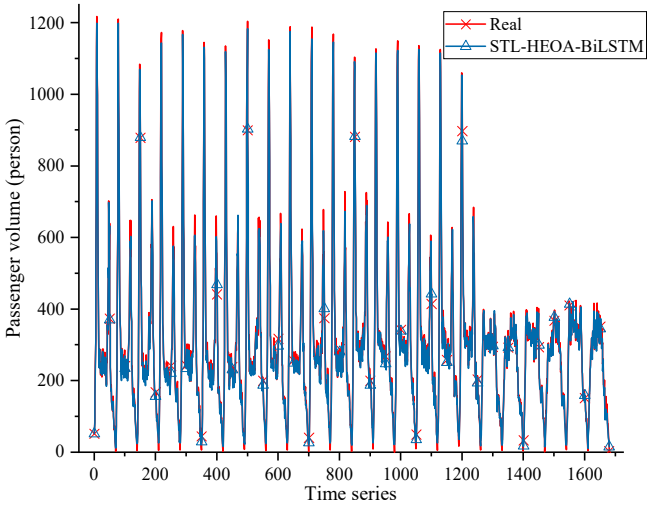


Fig. 18. Station 13 combined results of forecasts for each type of passenger flow

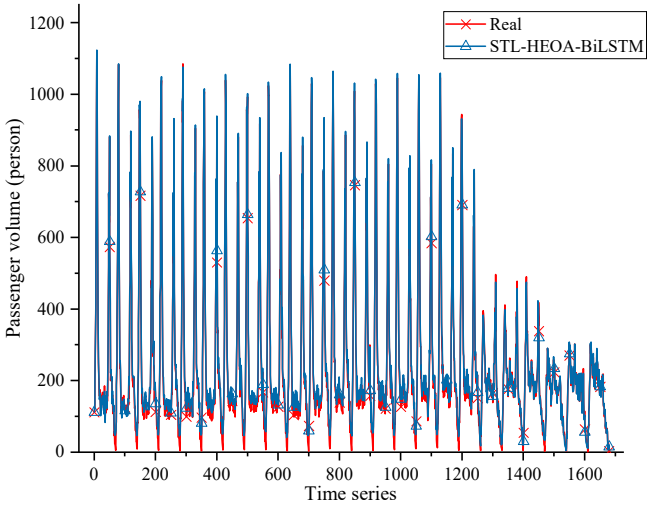


Fig. 19. Station 18 combined results of forecasts for each type of passenger flow

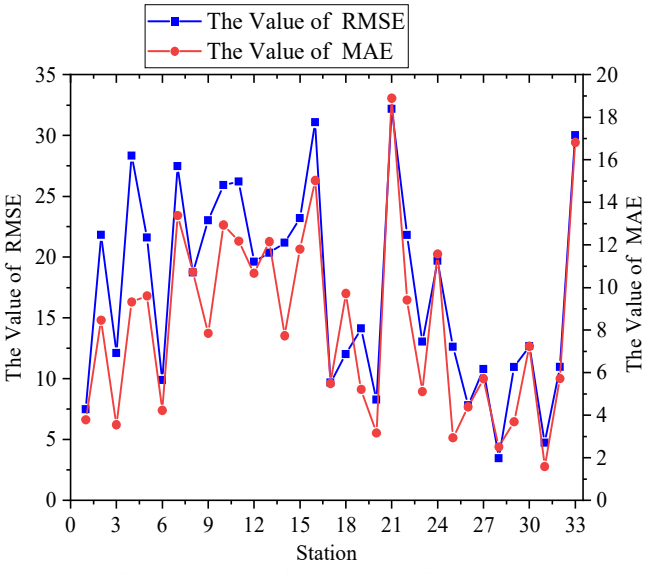


Fig. 20. Prediction errors of each station on the line

V. Conclusion

To enhance the accuracy and effectiveness of short-time passenger flow prediction for urban rail transit, a combined STL-HEOA-BiLSTM prediction model using the Pearson correlation degree analysis is proposed. Hangzhou Metro Line 1 is chosen as a study case. Firstly, the inbound passenger volume at each station along the urban rail transit in a week is classified into five categories based on the Pearson correlation number and time series characteristics. Secondly, the STL method decomposes each category's raw passenger flow data into T_t, S_t and R_t. By the STL method, the unstable passenger volume disturbances are mitigated to enhance the accuracy of predicting passenger volume. Then, the HEOA optimizes the hyper-parameters of the BiLSTM model. In addition, the optimized BiLSTM model is applied to predict T_t, S_t and R_t, respectively. Finally, to evaluate the prediction effect of the models, the six types of prediction models are selected for the error test and comparative analysis of the forecasted results. The above models are LSTM, BiLSTM, PSO-LSTM, PSO-BiLSTM, HEOA-LSTM, and HEOA-BiLSTM. The study's results are as follows:

(1) Based on the Pearson correlation number, the inbound passenger flow at each station is classified by time series characteristics. This behavior facilitates the refinement of the raw passenger flow data, leading to improved prediction accuracy.

(2) Based on the STL time series decomposition method, the classified raw passenger flow data are decomposed. This behavior dramatically improves the accuracy of the prediction model.

(3) The STL-HEOA-BiLSTM combined model presented in this paper outperforms the remaining six common or combined models in predicting all types of passenger volume data. Besides, the values of all prediction error indicators of the STL-HEOA-BiLSTM model are minimal. This result indicates the effectiveness of the STL-HEOA-BiLSTM combined model.

In a word, the proposed combined STL-HEOA-BiLSTM prediction model in this paper enhances the effectiveness of short-time passenger flow prediction in urban rail transit. The study's results offer an efficient method for predicting passenger volume in urban rail transit enterprises. The next research can focus on the spatial characteristics of short-time passenger volume. In this way, the effectiveness of the classification of passenger flow characteristics and prediction accuracy can be further strengthened to achieve dynamic prediction.

References

[1] Jie Li, Qiyuan Peng, Yuxiang Yang, "Passenger flow prediction of Guangzhou Zhuhai intercity railway based on SARIMA model," *Journal of Southwest Jiaotong University*, vol. 55, no. 1, pp. 41-51, 2020.

[2] Xumei Chen, Shuxia Guo, Lei Yu, et al. "Short-term forecasting of transit route OD matrix with smart card data," presented at the 2011 International IEEE Conference on Intelligent Transportation Systems, USA, Washington.

[3] Yuchuan Du, Ganzhe Chen, Xiaopeng Zhou, etc, "Research on OD prediction method based on Bayesian estimation," presented at the 2014 9th China Intelligent Transportation Annual Conference, China, Guangdong.

[4] Soomin Woo, Tak Sehyun, Yeo Hwasoo, "Data-Driven Prediction Methodology of Origin-Destination Demand in Large Network for Real-Time Services," *Transportation Research Record Journal of the Transportation Research Board*, vol. 2567, no. 1, pp. 47-56, 2016.

[5] Jing Liu, "Research on Dynamic OD Prediction Method of Urban Road Network Based on ANPR Data," M.S. thesis, Dept. Transportation. Eng., Southwest Jiaotong University., Sichuan, China, 2019.

[6] Shengyue Fang, "Research on Short term Passenger Flow Prediction of Subway Based on XGBoost," M.S. thesis, Dept. Transportation. Eng., Dalian Maritime University., Liaoning, China, 2022.

[7] Dawei Chen, "Research on passenger flow prediction of rail transit under rainy weather," M.S. thesis, Dept. Transportation. Eng., East China Jiaotong University., Jiangxi, China, 2022.

[8] Ruoyi Li, "Short term prediction of OD passenger flow in urban rail transit systems based on an improved spatiotemporal LSTM model," M.S. thesis, Dept. Transportation. Eng., Beijing Jiaotong University., Beijing, China, 2019.

[9] Xi Jiang, Feifan Jia, Jiaping Feng, "Online dynamic estimation of passenger flow OD in urban rail network based on AFC data," *Transportation System Engineering and Information*, vol. 18, no. 5, pp. 129-135, 2018.

[10] Yang D, Chen K, Yang M, et al, "Urban rail transit passenger flow forecast based on LSTM with enhanced long-term features," *IET Intelligent Transport Systems*, vol. 13, no. 10, pp. 1475-1482, 2019.

[11] Xiaoyun Hou, Liping Shao, Jing Li, et al, "Short term passenger flow origin and destination prediction of urban rail transit based on deep learning," *Urban Rail Transit Research*, vol. 23, no. 1, pp. 55-58, 2020.

[12] Yanjun Huang, "Research on Real time Prediction Method for OD of Passenger Flow in Urban Rail Transit Network," M.S. thesis, Dept. Transportation. Eng., Beijing Jiaotong University., Beijing, China, 2021.

[13] Zhang B, Yang X, Zhang Y, et al, "Short-Term Inbound Passenger Flow Prediction of Urban Rail Transit Based on RF-BiLSTM," *Engineering Letters*, vol. 31, no. 2, pp. 665-673, 2023.

[14] Siyu Jia, Ying Tian, "Face Detection Based on Improved Multi-task Cascaded Convolutional Neural Networks," *IAENG International Journal of Computer Science*, vol. 51, no. 2, pp67-74, 2024.

[15] Felipe Osorio-Arteaga, and Eduardo Giraldo, "Adaptive Neural Network Identification for Robust Multivariable Systems," *IAENG International Journal of Applied Mathematics*, vol. 54, no. 1, pp68-76, 2024.

[16] Junbo L, Guohua H, "Human Evolutionary Optimisation Algorithm," *Expert Systems with Applications*, vol. 241, pp. 1603-1611, 2024.

[17] Jiajing Li, Ning Zhang, Longhui Wen, et al, "Analysis and prediction method of short-term passenger flow changes at subway stations," *Urban Rail Transit Research*, vol. 26, no. 11, pp. 36-42, 2023.

[18] XuDong Li, JieSheng Wang, WenKuo Hao, and XiaoRui Zhao, "Bidirectional Long-term Short-term Memory Classification Method of Diabetes Datasets," *Engineering Letters*, vol. 30, no.4, pp. 1603-1611, 2022.

[19] Menglong Cao, Junlin Ma, "Research on Improving Locust Optimization Algorithm in Fuzzy Neural Network PID Control," *Electronic Measurement Technology*, vol. 45, no.20, pp. 74-80, 2022.