

Research Article

Short-Term Passenger Flow Prediction of Urban Rail Transit Based on SDS-SSA-LSTM

Haijun Li^(b),¹ Yongpeng Zhao,¹ Changxi Ma^(b),¹ Ke Wang,¹ Xiaoting Huang,¹ and Wentao Zhang²

¹School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China ²Operation Branch of Xi'an Rail Transit Group Co., Ltd., Xi'an 710000, China

Correspondence should be addressed to Haijun Li; 264160@qq.com

Received 29 June 2022; Revised 26 August 2022; Accepted 9 September 2022; Published 21 September 2022

Academic Editor: Zhenzhou Yuan

Copyright © 2022 Haijun Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Predicting rail transit passenger flow is crucial for modifying the metro schedule. To increase prediction accuracy, a model is proposed that combines long short-term memory (LSTM) with single spectrum analysis (SSA). Firstly, a stepwise decomposition sampling (SDS) strategy based on SSA progressive decomposition is proposed as a solution to the data leaking issue in traditional sequence decomposition. Then, based on this strategy, the passenger flow time series with complex features is decomposed into a relatively single trend and fluctuation component. Finally, the LSTM network is employed to perform short-term predictions on each component separately. The predicted value of each component is accumulated to obtain the original passenger flow' predicted result. The example shows that, compared with the single LSTM and other hybrid models, the proposed method offers a greater overall prediction accuracy in the experimental days, and the method has specific applicability.

1. Introduction

Subway system has the advantages of large capacity, high speed, and high reliability. It has attracted more and more attention both academically and practically [1]. However, while the public is pursuing the high-quality metro, various problems such as congestion and transfer inconvenience must be solved.

The metro system's short-term passenger flow (STPF) prediction is a crucial component in managing public transportation. By releasing the predicted STPF and congestion status of bus routes to the public, passengers can adjust travel mode, route selection, or departure time in advance to better understand travel mode. For the government and operators, the accurate STPF forecast can effectively control and assess the system status and implement response measures when emergencies or special events occur [2]. Hence, the accurate STPF forecast is of great significance.

Summing up, in order to predict the STPF more accurately, it is necessary to consider whether the prediction

results can reflect the actual online prediction effect and effectively avoid the problem of data leakage. In this research, a prediction model based on SDS-SSA-LSTM is proposed. The SDS strategy can effectively avoid the prediction model from involving future data in the testing phase and solve the data leakage problem of the traditional ODS strategy. Using the SSA method, the trend and fluctuation components in the passenger flow sequence with mixed features can be effectively extracted. Combined with the ability of the LSTM network to transmit long-term dependencies of sequence data, the accurate prediction of STPF can be achieved.

2. Literature Review

For unstable traffic flow, at present, wavelet decomposition (WD) [3, 4] and variational mode decomposition (VMD) [5] have been studied for rail transit STPF prediction. However, the above studies all adopt the overall decomposition sampling (ODS) sampling technique to process the data



FIGURE 1: The input and output samples.



FIGURE 2: LSTM model structure diagram.

samples; that is, the passenger flow time series is decomposed into subsequences. Nevertheless, the future data is assumed to be known in the decomposition process, and there is a problem of data leakage. Therefore, the model in online prediction has poor practicability and unreliability.

The sampling technology based on stepwise decomposition sampling (SDS) [6] can effectively avoid the data leakage problem in traditional ODS technology. EMD cannot be used for SDS because the number of modal decompositions is uncontrollable [7]. EEMD and CEEMD have their algorithm principle [8, 9]. With the update of the edge value of the time series to be decomposed, each decomposed component will change significantly. When they are sampled by SDS, the time sequence of the training samples is uncontrollable. It is difficult to determine the wavelet basis and decomposition level of WD. The trend items directly decomposed by VMD [10] will produce significant errors, which will make prediction difficult. SSA is a nonlinear time sequence analysis method [11], which can effectively extract the trend and volatility in the original sequence. There is no need to select a priori basis function. With the update of the time series to be decomposed, the sequence characteristics of the history of each decomposed component can be almost wholly preserved [12]. Zhou et al. [13] combined SSA with AdaBoost weighted limit learning machine. The original data was divided into three parts for the construction of SSA: trend, periodicity, and residue. Shuai et al. [14] show that the complex characteristics of regularity and randomness of traffic flow can be captured by the decomposed traffic flow prediction hybrid model. Therefore, with the advantage of the SSA decomposition method, this paper decomposes the mixed passenger flow series into trend series and fluctuation series with relatively single characteristics, which is easy for the model to learn.

In terms of the STPF prediction, traditional prediction models mainly include the autoregressive integrated moving average model [15], gray model [16], and Kalman filtering model [17]. The traditional time series prediction model has weak generalization ability, its performance and application will be limited, and it is challenging to meet the requirements of processing massive data in practice. In recent years, a large amount of data generated by various city sources has allowed us to understand better the essence of hidden dynamics in the traffic system and significantly changed the way of predicting passenger traffic. Due to higher computing capabilities, there is various computing intelligence and data-driven technology to create opportunities for the use of big data to broaden the practical application of the intelligent transportation system (ITS) [18].

Advanced intelligent algorithms have attracted the attention of many scholars due to their superior structure and performance. Li et al. [19] combined the characteristics of segmented passenger flow data with a BP neural network to predict segmented passenger flow. Tsai et al. [20] proposed a new neural network which proves to produce satisfactory results in forecasting STPF. Deep learning can capture complex nonlinear relationships in big data, which significantly improves predictive accuracy. In terms of the STPF prediction, the most commonly employed models are CNN [21-24] and RNN [25], which have shown their excellent ability to extract space-time structure information. LSTM [26–28] as a variant of RNN can automatically learn the dependencies of long-term sequences and capture the connection relationship between the steps. Through experiments, Zhao et al., Guo et al., Du et al., Jiang et al., and Hao et al. [29-33] proved that LSTM has certain advantages in capturing the regularity of STPF.

The sampling strategy of SSA stepwise decomposition (SDS) can be applied to avoid the data leakage in the traditional method of sequence decomposition. LSTM neural network is suitable for predicting each component by learning historical data. The SDS-SSA-LSTM hybrid model can provide a new method for the accurate predicting of STPF. In this paper, we attempt to decompose the time series with complex characteristics through the SDS strategy based on SSA, so as to extract the internal characteristics of complex passenger flow data, and build a LSTM memory network model to forecast the short-term metro passenger flow.

3. Methodology

3.1. Singular Spectrum Analysis (SSA). SSA is a method of nonlinear time-series analysis. By constructing the trajectory matrix of the time series and decomposing and reconstructing it, the trend and fluctuation components in the original series can be effectively extracted. The specific calculation steps of this algorithm are as follows:



FIGURE 3: The process of the STPF prediction based on SDS-SSA-LSTM.

(1) Embedding. According to the set embedding dimension $L(L \le n/2)$, the original time series $X = (x_1, x_1, \dots, x_n)$ is transformed into trajectory matrix Y. The formula is as follows:

$$Y_{L\times K} = \left(y_{ij}\right)_{i,j=1}^{L,K} = \begin{pmatrix} x_1 & x_2 & \cdots & x_{n-L+1} \\ x_2 & x_3 & \cdots & x_{n-L+2} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \cdots & x_n \end{pmatrix}, \qquad (1)$$

where K = n - L + 1

(2) Singular value decomposition. Calculate the eigenvalues λ₁, λ₂, ..., λ_L(λ₁ ≥ λ₂≥····≥λ_L) and corresponding eigenvectors U₁, U₂, ..., U_L of matrix YY^T. d is the number of nonzero eigenvalues; define V_i = Y^TU_i√λ_i; the singular value decomposition of the trajectory matrix Y can be written as follows:

$$Y = Y_1 + Y_2 + \dots + Y_d,$$
 (2)

where $Y_i = \sqrt{\lambda_i} U_i V_i^T$, $1 \le i \le d$; U_i is the *i*th singular value of the trajectory matrix *Y* corresponding to the left singular vector; and V_i is the *i*th singular value of the trajectory matrix *Y* corresponding to the right singular vector. The contribution degree of each matrix Y_i to the trajectory matrix *Y* is related to the eigenvalue λ_i , and its contribution

rate η can be defined as follows:

$$\eta = \frac{\lambda_i}{\sum \lambda} \tag{3}$$

(3) Diagonal averaging. Each submatrix Y_i is reduced to a time series with n length. Suppose that Y is a matrix with the same shape as L×K. Let L* = min (L, K) and K* = max (L, K). If L < K, y_{ij}^{*} = y_{ji}; otherwise, the restructured series Z_c = (z_{c1}, z_{c2},..., z_{cn}) is defined as follows:

$$z_{cl} = \begin{cases} \frac{1}{l+1} \sum_{m=1}^{l+1} y_{m,l-m+2}^{*}, & 1 \le l \le L^{*}, \\ \frac{1}{L^{*}} \sum_{m=1}^{L^{*}} y_{m,l-m+2}^{*}, & L^{*} \le l \le K^{*}, \\ \frac{1}{n-1} \sum_{m=l-K^{*}+2}^{l-K^{*}+1} y_{m,l-m+2}^{*}, & K^{*} \le l \le n \end{cases}$$

$$(4)$$

(4) Grouping. The singular vector corresponding to the largest eigenvalue represents the largest change trend of the original sequence, and the singular vector corresponding to the smaller eigenvalue reflects the fluctuation components of the original sequence. By setting the contribution rate threshold η_e (generally taken $\eta_e \ge 80\%$) and then gradually



FIGURE 4: Periodic analysis of sampling passenger flow.

reconstructing the subsequence in the order of eigenvalue from large to small, when it appears $\sum \lambda_i / \sum \lambda \ge \eta_e$, the obtained sequence is the trend component of the original sequence, and the remaining subsequence is reconstructed into the fluctuation component

3.2. Sampling Strategy Based on Stepwise Decomposition. Let time series $TS = \{x_t | t = 1, \dots, T + K\}$, where T + K is the length of time series TS. Use x_1 to x_T as training data and the rest as test data. The specific stepwise decomposition steps are as follows:

- (1) Decompose the time series $x_{1:L}$ into subsequences $x_{1:L+1}^{i2}$
- (2) Decompose the time series x_{1:L+1} with length L + 1 into subsequence xⁱ²_{1:L+1}. Then, the input and output of the first sample of the *i*th subsequence come from xⁱ¹_{1:L} and xⁱ²_{1:L+1}, respectively, as shown in Figure 1
- (3) Continue to decompose the next time series of length L and L + 1 according to step 1 and step 2 until all training datasets are obtained
- (4) Establish a prediction model for each subsequence, respectively
- (5) In the prediction stage, the time series of historical L length is decomposed, and each subsequence is predicted according to the trained model

(6) Accumulate the predicted values of each subsequence to obtain the predicted result of the original sequence

The application of the above SDS strategy completely avoids the data leakage problem in traditional ODS, and the model based on this strategy conforms to the actual online prediction scenario.

3.3. Long Short-Term Memory Network (LSTM). Compared with traditional neural networks, RNN can capture the temporal regularity of sequence data more effectively. LSTM was originally a variant of the traditional recursive neural network. It has strong memory ability when dealing with time-series prediction and is widely used in time-series prediction scenarios with long time interval and delay. The basic structure of the LSTM unit is shown in Figure 2.

LSTM adds more memory modules to the neural network to selectively let information pass through, with each door having a different function.

The forget gate f_t decides which news to discard by integrating the information of the last moment and the news of the present moment. The formula is as follows:

$$f_{t} = \sigma \left(w_{fx} x_{t} + w_{fh} h_{t-1} + w_{fc} C_{t-1} + b_{f} \right), \tag{5}$$

$$\sigma(z) = \frac{1}{1 + e^{-z}},\tag{6}$$



(b) Volatility component training samples

FIGURE 5: Continued.



FIGURE 5: The training samples and test samples of the passenger flow sequence data 1.

where $\sigma(\cdot)$ is the sigmoid activation function, x_t the input at t, h_t the output at t, h_{t-1} the output at t - 1, C_t the candidate vectors at time t, w_{fx} , w_{fh} , w_{fc} weight coefficient of the forgetting gate, and b_f bias of forgetting gate.

An input gate i_t is the opposite of a forget gate f_t . This section decides the news to enter the cell state based on the input threshold. The output of the input gate at t is as follows:

$$i_{t} = \sigma(w_{ix}x_{i} + w_{ih}h_{t-1} + w_{ic}C_{t-1} + b_{i}),$$

$$v_{t} = \tanh(w_{cx}x_{t} + w_{ch}h_{t-1} + b_{o}),$$

$$\tanh(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}},$$
(7)

where w_{ix} , w_{ih} , w_{ic} are the weight coefficients of input gate, b_i is bias of input gate, w_{cx} , w_{ch} are the weighting coefficients and bias of candidate vectors, b_o is the bias of candidate vectors, tanh (·) is the hyperbolic tangent activation function, and v_t is the updated value of candidate vectors.

The output gate o_t determines what information will be output. No information can pass through the output gate except what is required. So the output at time *t* is as follows:

$$o_t = \sigma(w_{ox} + w_{oh}h_{t-1} + w_{oc}C_{t-1} + b_o),$$

$$h_t = o_t \cdot \tanh(C_t),$$

$$C_t = C_{t-1} \cdot f_t + i_t v_t,$$
(8)



FIGURE 6: Continued.



(d) Volatility component test sample

FIGURE 6: Training samples and test samples of passenger flow sequence data 2.

		•	
Model parameters	SDS	SSA	LSTM
Number of neurons	_	_	32
Learning rate	—	—	0.001
Look back	—	—	3
Epochs	—	—	50
Window length	—	20	_
Contribution threshold	—	80%	—
Length of decomposition	50	—	_

TABLE 1: SDS-SSA-LSTM model parameters.

Note: "—" indicates that the parameter in this row does not exist in this column model.

where o_t is the output gate and w_{ox} , w_{oh} , w_{oc} are the weighting coefficients of the output gate.

3.4. STPF Prediction Model. Figure 3 is the process of the SSA-LSTM model based on the SDS strategy, and here are the modeling steps:

Step 1. Original data processing. The original rail transit passenger flow's preprocessed time series is divided into a training set and a test set.

Step 2. Decomposition of original passenger flow data. SDS for the training set based on the SSA method until the



FIGURE 7: Prediction results of sequence data 1.



FIGURE 8: Prediction results of sequence data 2.

Data	Model	Evaluation index		
		RMSE	MSE	MAE
Data 1	ARIMA	4.484	20.102	0.163
	LSTM	4.340	18.838	2.642
	GRU	4.469	19.970	2.776
	DWT-LSTM	3.367	11.335	2.008
	EMD-LSTM	3.693	13.637	2.335
	VMD-LSTM	4.142	17.160	2.693
	SDS-SSA-LSTM	2.040	4.161	1.280
Data 2	ARIMA	8.450	71.403	0.229
	LSTM	5.958	35.496	4.872
	GRU	6.059	36.712	4.960
	DWT-LSTM	5.939	35.274	4.636
	EMD-LSTM	5.919	35.040	5.010
	VMD-LSTM	5.902	34.830	4.823
	SDS-SSA-LSTM	4.320	18.661	1.838

TABLE 2: Comparison of model prediction performance.

complete training sample of the trend component and the fluctuation component is obtained. The trend series represents the broader trend of original data. The fluctuation series represents the periodic of original data. *Step 3.* Establish prediction model. The LSTM prediction model in view of the deep learning is established for the training set in each subsequence, respectively.

Step 4. Perform SSA stepwise decomposition on the test set, and input it into the LSTM to generate the prediction results of each component.

Step 5. Accumulate the predicted values of each component. Get the STPF forecast results.

4. Experiments and Results

4.1. Data Analysis. The original data is selected from the daily passenger flow data of Xi'an Metro Line 3 from January 1, 2017, to January 24, 2020, to analyze the established model. The sampling interval is one day, and a total of 1119 groups of sample points are obtained.

Passenger flow data has a time correlation. The closer the time is , the higher the correlation between the speed. With the increase of time, the correlation weakens. In addition, the passenger flow data has an obvious trend in the short term and periodic in the long term. To analyze the temporal characteristics of passenger flow data, some datasets are selected for visual analysis, as shown in Figure 4.

The weekly trend changes in Figure 4 are very similar, which shows the passenger flow data has periodicity of

prominence. Therefore, in this study, the passenger flow data is grouped into two samples, from Monday to Thursday (data 1) and from Friday to Sunday (data 2), and the two sets of datasets are, respectively, predicted to verify the model.

4.2. Passenger Flow Decomposition

4.2.1. SSA Decomposition. The SSA is applied to the historical time-series for SDS. Figures 5 and 6 show the decomposition results of training samples and test samples of the two datasets, respectively.

4.2.2. Data Normalization. To improve the data processing efficiency of the subsequent prediction model, it is necessary to normalize each subsequence. In this paper, MinMax is selected for normalization.

4.3. STPF Forecast

4.3.1. Model Parameter Setting. In this paper, given the daily passenger flow data of Xi'an Metro Line 3 from January 1, 2017, to January 24, 2020, the short-term passenger flow in the future period is calculated. After many tests, the parameters of the SDS-SSA-LSTM model are set in the example as shown in Table 1.

4.3.2. Forecast Result. To test the performance of the model in the STPF prediction, ARIMA, LSTM, GRU single model, and DWT-LSTM, EMD-LSTM, VMD-LSTM, SDS-SSA-LSTM, and other hybrid models are used to predict the two datasets. The prediction results are shown in Figures 7 and 8.

4.4. Performance Comparison. In this paper, three single models and four mixed depth models commonly used in the existing literature are selected to compare the prediction accuracy. To evaluate the effectiveness and accuracy of the SDS-SSA-LSTM model, the mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE) were used to evaluate the model. Table 2 shows the error comparison results of the prediction model.

According to the error results in Table 2, it can be seen that

- (1) in the classic single-model prediction, the LSTM model shows better prediction effect on both datasets, and RMSE and MSE indicators are the lowest. However, the traditional statistical method has low prediction accuracy for passenger flow data with complex influencing factors
- (2) in terms of predicting STPF, the hybrid model with decomposition and fusion strategies performs better than the single model. To some extent, the prediction accuracy has increased
- (3) the prediction performance of the DWT-LSTM, EMD-LSTM, and VMD-LSTM models using ODS strategy is lower than that of the SSA-LSTM model based on SDS strategy
- (4) ODS strategy assumes that the future data is known from the beginning, but in practical application, the

future data needs to be predicted. Therefore, the prediction results obtained by the model based on the ODS strategy cannot reflect the actual online prediction effect. The model based on the SDS strategy completely avoids the leakage of future data, and its prediction results can guide online prediction

5. Conclusion

This research suggests a prediction methodology based on SDS-SSA-LSTM to better precisely predict the passenger flow sequence of rail transit. The SDS strategy can effectively avoid the future data involved in the prediction model in the testing phase and solve the data leakage problem of the traditional ODS strategy. Using the SSA method, the passenger flow sequence with mixed features is decomposed into trend sequence and fluctuation sequence with relatively single features, which is easy for model learning. Through example verification, the proposed model established in this paper performs well in the STPF prediction, and this method reveals great application prospects.

In the future, developing a visual STPF forecasting platform is the next research focus, and developing a prediction model based on decomposition needs to be combined with SDS strategy. In addition, the cumulative error caused by the increased in modeling quantity is a challenge in achieving the STPF prediction.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the National Natural Science Foundation of China Grant Nos. 52062027 and 71861023, the "Double First Class" Major Research Programs, Gansu Education Department (No. GSSYLXM-04), and Foundation of A Hundred Youth Talents Training Program of Lanzhou Jiaotong University.

References

- Y. Liu, Z. Y. Liu, and R. Jia, "DeepPF: a deep learning based architecture for metro passenger flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 18– 34, 2019.
- [2] Y. Yang, Z. Yuan, and R. Meng, "Exploring traffic crash occurrence mechanism toward cross-area freeways via an improved data mining approach," *Journal of Transportation Engineering*, *Part A: Systems*, vol. 148, no. 9, article 04022052, 2022.
- [3] R. Liu, Y. Wang, H. Zhou, and Z. Qian, "Short-term passenger flow prediction based on wavelet transform and kernel extreme learning machine," *IEEE Access*, vol. 7, pp. 158025– 158034, 2019.

- [4] X. Yang, Q. Xue, X. Yang et al., "A novel prediction model for the inbound passenger flow of urban rail transit," *Information Sciences*, vol. 566, pp. 347–363, 2021.
- [5] D. Liang, J. Xu, S. Li, and C. Sun, "Short-term passenger flow prediction of rail transit based on VMD-LSTM neural network combination model," in 2020 Chinese Control And Decision Conference (CCDC), pp. 5131–5136, Hefei, China, 2020.
- [6] Z. Qian, Y. Pei, H. Zareipour, and N. Chen, "A review and discussion of decomposition-based hybrid models for wind energy forecasting applications," *Applied Energy*, vol. 235, pp. 939–953, 2019.
- [7] W. Fang, S. Huang, K. Ren et al., "Examining the applicability of different sampling techniques in the development of decomposition-based streamflow forecasting models," *Journal* of Hydrology, vol. 568, pp. 534–550, 2019.
- [8] T. Wang, M. Zhang, Q. Yu, and H. Zhang, "Comparing the applications of EMD and EEMD on time-frequency analysis of seismic signal," *Journal of Applied Geophysics*, vol. 83, pp. 29–34, 2012.
- [9] P. Flandrin, E. Torres, and M. A. Colominas, "A complete ensemble empirical mode decomposition with adaptive noise," in 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp. 4144–4147, Prague, Czech Republic, 2011.
- [10] P. Zhang, P. Qi, R. Y. Zhang, M. Y. Shao, and C. R. Li, "Dissolved gas prediction in transformer oil based on empirical wavelet transform and gradient boosting radial basis," *Power System Technology*, vol. 45, no. 9, pp. 3745–3754, 2021.
- [11] F. H. Wen, J. H. Xiao, Z. F. He, and X. Gong, "Stock price prediction based on SSA and SVM," *Procedia Computer Science*, vol. 31, pp. 625–631, 2014.
- [12] J. Wu, S. Xiang, C. Yan, X. D. Wu, J. T. Ma, and F. S. Liu, "Multi-step prediction of super-short-term wind power based on singular spectrum analysis," *Renewable Energy Resources*, vol. 39, no. 11, pp. 1548–1555, 2021.
- [13] W. Zhou, W. Wang, and D. Zhao, "Passenger flow forecasting in metro transfer station based on the combination of singular spectrum analysis and AdaBoost-weighted extreme learning machine," *Sensors*, vol. 20, no. 12, 2020.
- [14] C. Shuai, Z. Pan, L. Gao, and H. Gao, "Short-term traffic flow prediction of expressway: a hybrid method based on singular spectrum analysis decomposition," *Advances in Civil Engineering*, vol. 2021, Article ID 4313970, 10 pages, 2021.
- [15] Y. Chen, Y. Sha, X. Zhu, and X. Zhang, "Prediction of Shanghai Metro Line 16 Passenger Flow Based on Time Series Analysis," *Operations Research and Fuzziology*, vol. 6, no. 1, pp. 15–26, 2016.
- [16] L. M. Hou and G. F. Ma, "Forecast of railway passenger traffic based on a grey linear regression combined model," *Computer Simulation*, vol. 28, no. 7, pp. 1–3, 2011.
- [17] C. Zhang, R. Song, and Y. Sun, "Kalman filter-based shortterm passenger flow forecasting on bus stop," *Journal of Transportation Systems Engineering and Information Technology*, vol. 11, no. 4, pp. 154–159, 2011.
- [18] Z. Cui, R. Ke, Z. Pu, and Y. Wang, "Deep bidirectional and unidirectional LSTM recurrent neural network for networkwide traffic speed prediction," 2018, https://arxiv.org/abs/ 1801.02143.
- [19] Q. Li, Y. Qin, Z. Y. Wang et al., "The research of urban rail transit sectional passenger flow prediction method," *Journal* of *Intelligent Learning Systems and Applications*, vol. 5, no. 4, pp. 227–231, 2013.

- [20] T. H. Tsai, C. K. Lee, and C. H. Wei, "Neural network based temporal feature models for short-term railway passenger demand forecasting," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3728–3736, 2009.
- [21] G. Liu, Z. Yin, Y. Jia, and Y. Xie, "Passenger flow estimation based on convolutional neural network in public transportation system," *Knowledge-Based Systems*, vol. 123, pp. 102– 115, 2017.
- [22] Y. Wang, Y. Qin, J. Guo, Z. Gao, and L. Jia, "Multi-point short-term prediction of station passenger flow based on temporal multi-graph convolutional network," *Physica A: Statistical Mechanics and its Applications*, vol. 604, article 127959, 2022.
- [23] T. Bogaerts, A. D. Masegosa, J. S. Angarita-Zapata, E. Onieva, and P. Hellinckx, "A graph CNN-LSTM neural network for short and long-term traffic forecasting based on trajectory data," *Transportation Research Part C: Emerging Technologies*, vol. 112, pp. 62–77, 2020.
- [24] J. Wang, W. Wang, X. Liu, W. Yu, X. Li, and P. Sun, "Traffic prediction based on auto spatiotemporal multi-graph adversarial neural network," *Physica A: Statistical Mechanics and its Applications*, vol. 590, p. 126736, 2022.
- [25] Z. N. Zou, H. Peng, L. Liu et al., "Deep convolutional mesh RNN for urban traffic passenger flows prediction," in 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ ATC/CBDCom/IOP/SCI), pp. 1305–1310, Guangzhou, China, 2018.
- [26] C. Ma, G. Dai, and J. Zhou, "Short-term traffic flow prediction for urban road sections based on time series analysis and LSTM_BILSTM method," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, 2022.
- [27] C. Zeng, C. Ma, K. Wang, and Z. Cui, "Predicting vacant parking space availability: a DWT-Bi-LSTM model," *Physica A: Statistical Mechanics and its Applications*, vol. 599, p. 127498, 2022.
- [28] J. Zhang, F. Chen, Z. Cui, Y. Guo, and Y. Zhu, "Deep learning architecture for short-term passenger flow forecasting in urban rail transit," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 7004–7014, 2021.
- [29] Y. Y. Zhao, L. Xia, and X. G. Jiang, "Short-term metro passenger flow prediction based on EMD-LSTM," *Journal of Traffic and Transportation Engineering*, vol. 20, no. 4, pp. 194–204, 2020.
- [30] J. Guo, Z. Xie, Y. Qin, L. Jia, and Y. Wang, "Short-term abnormal passenger flow prediction based on the fusion of SVR and LSTM," *IEEE Access*, vol. 7, no. 7, pp. 42946–42955, 2019.
- [31] B. Du, H. Peng, S. Wang et al., "Deep irregular convolutional residual LSTM for urban traffic passenger flows prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 3, pp. 972–985, 2019.
- [32] W. Jiang, Z. Ma, and H. N. Koutsopoulos, "Deep learning for short-term origin-destination passenger flow prediction under partial observability in urban railway systems," *Neural Computing and Applications*, vol. 34, no. 6, pp. 4813–4830, 2022.
- [33] S. Hao, D. H. Lee, and D. Zhao, "Sequence to sequence learning with attention mechanism for short-term passenger flow prediction in large-scale metro system," *Transportation Research Part C: Emerging Technologies*, vol. 107, pp. 287– 300, 2019.