

## *Retraction*

# **Retracted: A Multiscale and High-Precision LSTM-GASVR Short-Term Traffic Flow Prediction Model**

### **Complexity**

Received 19 December 2023; Accepted 19 December 2023; Published 20 December 2023

Copyright © 2023 Complexity. 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.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] J. Zhou, H. Chang, X. Cheng, and X. Zhao, "A Multiscale and High-Precision LSTM-GASVR Short-Term Traffic Flow Prediction Model," *Complexity*, vol. 2020, Article ID 1434080, 17 pages, 2020.

## Research Article

# A Multiscale and High-Precision LSTM-GASVR Short-Term Traffic Flow Prediction Model

Jingmei Zhou <sup>1</sup>, Hui Chang <sup>2</sup>, Xin Cheng <sup>2</sup>, and Xiangmo Zhao <sup>2</sup>

<sup>1</sup>School of Electronic and Control Engineering, Chang'an University, Xi'an 710064, China

<sup>2</sup>School of Information Engineering, Chang'an University, Xi'an 710064, China

Correspondence should be addressed to Jingmei Zhou; [jmzhou@chd.edu.cn](mailto:jmzhou@chd.edu.cn), Hui Chang; [480796959@qq.com](mailto:480796959@qq.com), and Xin Cheng; [xincheng@chd.edu.cn](mailto:xincheng@chd.edu.cn)

Received 28 March 2020; Revised 23 April 2020; Accepted 28 April 2020; Published 17 June 2020

Guest Editor: Zhihan Lv

Copyright © 2020 Jingmei Zhou 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.

Short-term traffic flow has the characteristics of complex, changeable, strong timeliness, and so on. So the traditional prediction algorithm is difficult to meet its high real-time and accuracy requirements. In this paper, a multiscale and high-precision LSTM-GASVR short-term traffic flow prediction algorithm is proposed. This method uses 15 min traffic flow data of the first 16 sections as input and completes the data preprocessing operation through reconstruction, normalization, and rising dimension by working day factor; establishing the prediction model based on the long- and short-term memory network (LSTM) and inverse normalization; and proposing the GA-SVR model to optimize the prediction results, so as to realize the real-time high-precision prediction of traffic flow. The prediction experiment is carried out according to the charge data of a toll station in Xi'an, Shaanxi Province, from May 2018 to May 2019. The comparison and analysis of various algorithms show that the prediction algorithm proposed in this paper is 20% higher than the LSTM, GRU, CNN, SAE, ARIMA, and SVR, and the  $R^2$  can reach 0.982, the *explanatory variance* is 0.982, and the MAPE is 0.118. The proposed traffic flow prediction algorithm provides strong support for traffic managers to judge the state of the road network to control traffic and guide traffic flow.

## 1. Introduction

High-precision and short-term traffic flow forecast shows the future trend of traffic development [1]. For intelligent traffic management, road network planning provides future traffic flow data [2]. It is not only helpful to alleviate traffic congestion, but also important for autonomous vehicles [3]. Traffic flow prediction refers to forecast traffic flow in the next period based on several historical traffic flow data [4]. A prediction one to several hours ahead of schedule is often called a short-term prediction [5, 6]. As a hot theme in Intelligent Transportation [7–9], there were lots of relating research studies about short-term traffic flow prediction. Wang et al. [10] proposed a hybrid short-term traffic speed prediction framework based on empirical mode by decomposition (EMD) and autoregressive integrated moving average (ARIMA) and achieved short-term traffic flow prediction of expressway in different scenarios, but there are

great differences in the prediction effect for different types of vehicles. Vasantha Kumar and Vanajakshi [11] developed a SARIMA short-term traffic flow prediction model under constrained data. The model needed high stable data and was weak at generalization. A two-step prediction method based on stochastic differential equation (SDE) was implemented in reference [12], which improves the prediction accuracy of a periodic data, but neglects the prediction effect of periodic traffic flow. Salamanis et al. [13] studied a density-based clustering method which can accurately predict traffic flow under normal and abnormal conditions. The prediction effect is good, but the time complexity is high and the feasibility is slightly lower. Neuhold et al. [14] established a model to predict traffic flow and an algorithm to optimize lane allocation in front of the toll plaza in Austria. The prediction model and optimization algorithm are not site specific and can be applied to different toll plaza or expressway bottlenecks (such as road engineering, transit, and

highway intersections). But for the data that are one-day or one-hour interval traffic flow, this model cannot fully meet the intelligent traffic management needs. A prediction model based on deep recurrent neural network (RNN) was designed in reference [15]. While the prediction effect is better, there are great differences in the prediction effect of a periodic data. Qu et al. [16] used historical traffic flow data and environmental factor data to predict all-day traffic flow based on deep neural networks, using multilayer supervised learning algorithm to train predictors, mining the potential relationship between traffic flow data and key contextual factors, an intermittent training method for reducing training time. But the prediction effect for smaller traffic flow data is poor. In [17], based on the BP neural network to predict the expressway traffic flow, the incentive function of the model is improved, and to some extent, the prediction effect is improved, but it is not suitable for all highways, and the limitation is great. In reference [18], the traffic flow information acquisition technology and combination model of video image analysis were applied to the traffic flow prediction to reduce the error and time for establishment. But the prediction effect is affected by the accuracy of data acquisition and fluctuates greatly. Luo et al. [7] implemented a CNN-SVR (convolutional neural networks-support vector regression) short-term traffic flow prediction model. The Adam optimization algorithm was used to ensure the completeness of the time-space flow characteristics, which reduces the interference of external factors and can effectively predict the traffic flow. However, because the data are relatively single, the generalization ability of the model is not strong. In reference [19], a stationary short-time traffic flow prediction method was proposed. The support vector machine predicts the traffic flow data after stationary, which solves the influence of the asymmetric distribution of the data on the prediction effect, but caused a bad prediction effect of the extreme point. Wang et al. [20] proposed an improved BP neural network model based on thought evolution algorithm. The accuracy of the prediction model was improved by phase space reconstruction of traffic flow time series using chaotic theory. However, because the model uses less data, lacks credibility, and did not consider the prediction effect of special time periods such as holidays, it is not universal. A traffic flow prediction model based on convolutional neural network is implemented in reference [21], and the paper considered the spatial-temporal correlation of regional traffic flow, which improves the prediction accuracy and stability of the model. Only taxi traffic flow data were studied, which do not fully reflect the road state. Previous neural network models used values from previous days or weeks as inputs to predict future traffic flow data [22, 23], and the real time is poor, which exists the problem of gradient vanishing. To address this, LSTM (Long Short-Term Memory) neural network was used to predict traffic flow [24–27], but the influence of commuting traffic flow on urban traffic state was not considered.

Tidal traffic flow phenomenon is one of the important factors causing congestion in toll stations. Accurate and real-time prediction of short-term traffic flow provides strong support for traffic management departments to guide traffic

flow. And it plays an important role in the information communication system [28]. This paper presents a multi-scale and high-precision LSTM-GASVR (Long Short-Term Memory Genetic Algorithm Support Vector Regression) short-term traffic flow prediction algorithm. Data preprocessing was utilized to filter, normalize, and reconstruct the data and improve the data dimension by workday flags, which effectively improves the convergence speed and prediction accuracy of the model; building prediction model based on the output LSTM, Long Short-Term Memory network, and GA-SVR model was proposed to optimize the prediction results to achieve real-time and high-precision prediction of short-term traffic flow. The proposed traffic flow prediction algorithm improves the prediction accuracy and stability and improves the generality and feasibility of the traffic flow prediction model.

## 2. Short Traffic Flow Data Preprocessing

In order to obtain the appropriate amount of data and the required data format, traffic flow data need to be pre-processed by data merging, expanding the amount of data, and data resampling to obtain time traffic volume. The preprocessing of short-term traffic flow data is divided into the following steps, as shown in Figure 1:

- (i) Data normalization makes the data model training have better convergence
- (ii) Considering the occasional and violent weekend and holiday tidal traffic flow, it is necessary to enhance the data dimension
- (iii) Data reconstruction aims to obtain a specific data format

Step 1: extract the data and merge the data for  $n$  consecutive months as needed to expand the dataset.

Step 2: resample the data as needed. Table 1 shows the charge data of Shaanxi Provincial Toll Station in August 2018; “STARTTIME” is the computer boot time; the “ENTRYLANE” is the entrance lane number; the “ENTRYTIME” indicates the vehicle entry time; and the “VEHICLETYPE” is the vehicle type; 0, 1, and 2, respectively, indicate the uncertain vehicle type, passenger car, and freight car; “VEHICLELICENSE” represents the license plate number. To extract the “ENTRYTIME” column of Table 1, we regard one line of data as a vehicle entering the station. These data are resampled according to 15 min, and the data shown in Table 2 are obtained. The “Value” represents the traffic volume within 15 min with the starting time of “Date.”

Step 3: data normalization. For a better convergence of the prediction results and the better results, the data are normalized by the maximum and minimum value standardization, assuming that the  $x$  is the traffic volume of a certain period of time, and it is normalized as follows:

$$x' = \frac{x - \min}{\max - \min}, \quad (1)$$

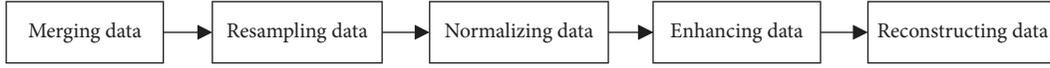


FIGURE 1: Data preprocessing.

TABLE 1: Partial data of a toll station in Shaanxi Province.

STARTTIME	...	ENTRYPANE	ENTRYTIME	VECHILETYPE	...	VECHILELICENSE	...
2018/7/29 19:22:51	...	12	2018/8/1 14:05:05	1	...	(B)S ADT702	...
2018/8/1 11:55:01	...	7	2018/8/1 14:05:05	1	...	(B)E FLG085	...
2018/7/29 19:22:51	...	12	2018/8/1 14:05:02	1	...	(B)S A53H75	...
...	..	...	...	...	..	...	...

TABLE 2: Partial traffic data of 15 min at Qujiang Toll Station, Shaanxi Province.

Date	Value
2018/8/1 8:00	186
2018/8/1 8:15	239
2018/8/1 8:30	300
...	...

TABLE 3: Format of traffic flow forecast training data for  $m$  period before the day.

Sample no.	0	1	...	$n$
Period 0 traffic volume (input)	$v_0^0$	$v_0^1$	...	$v_0^n$
Period 1 traffic volume (input)	$v_1^0$	$v_1^1$	...	$v_1^n$
...	...	...	...	...
Period $m-1$ traffic volume (input)	$v_{m-1}^0$	$v_{m-1}^1$	...	$v_{m-1}^n$
Working day factors	$h_b^0$	$h_b^1$	...	$h_b^n$
Next period traffic volume (output)	$v_m^0$	$v_m^1$	...	$v_m^n$

$x'$  is the values after  $x$  normalization fall, from 0 to 1, and the min and max in equation (1) are the minimum and maximum values in the dataset.

Step 4: dimension promotion. The traffic demand caused by commuter's commute is an important part of urban road traffic volume. Because of the urban planning problems, the traffic flow of commuters has an obvious influence on urban traffic state. The working day factors are divided into three categories according to "week," "weekend," and "holiday."

Step 5: data reconstruction. The data are reconstructed as needed, the data are divided into  $m+2$  row and  $n+1$  column, the first  $m+1$  row of each column is input, and the last row is compared with the predicted value as the real value. The final data structure is shown in Table 3.  $v_i^j$  represents traffic volume in sample  $i$  at period  $j$ .  $h_b^j$  represents the working day factor in sample  $j$ .  $h_b$  represents the working day factor,  $b$  is 1, 2, or 3,  $h_1 = 0.1$ ,  $h_2 = 0.5$ , and  $h_3 = 0.9$ .

### 3. Analysis of Time Series Predictive Assessment Indicators

The prediction results of time series need to evaluate each prediction model through indicators and adjust the parameters, so as to determine the high-precision short-term traffic flow prediction model. Equations (2)–(9) are evaluation indicators of time series prediction [2].  $f_i$  represents the predicted value of sample  $i$ ,  $y_i$  represents the actual value of sample  $i$ ,  $\bar{y}$  represents the average value of the real data, and  $n$  is the sample size.

① MAE (mean absolute error):

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |f_i - y_i|. \quad (2)$$

② MSE (mean square error is one of the most commonly used performance measures for regression tasks):

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (f_i - y_i)^2. \quad (3)$$

③ RMSE (root mean squared error):

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (f_i - y_i)^2}. \quad (4)$$

④ MAPE (mean absolute percentage error) often used to measure prediction accuracy:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - f_i}{y_i} \right|. \quad (5)$$

⑤ Explanatory variance:

The explanatory variance represents the extent to which the regression model explains the variation of the variance of the dependent variable (i.e., fitting effect of the regression model to the true value), which is a common evaluation index for the regression model:

$$\text{Explained\_variance}(y, f) = 1 - \frac{\text{Var}\{y - f\}}{\text{Var}\{y\}}. \quad (6)$$

The explanatory variance is between  $[0, 1]$ , and the closer to 1 indicates the better the prediction effect.

⑥  $R^2$  ( $R$ -squared):

$R^2$  is the ratio of the sum of squared residuals to the sum of the total deviation squares, indicating the degree to which the regression equation can explain the variation of the dependent variable, and  $R^2$  is the fitting effect of the regression equation on the true value.

The sum of squared residuals:

$$SS_{\text{res}} = \sum_{i=1}^n (f_i - y_i)^2. \quad (7)$$

The sum of total deviation squares:

$$SS_{\text{tot}} = \sum_{i=1}^n (y_i - \bar{y})^2, \quad (8)$$

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} = 1 - \frac{\sum_{i=1}^n (f_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (9)$$

The value of  $R^2$  is between 0 and 1. The closer  $R^2$  is to 1, the more accurate the model and the better the regression prediction effect. It is generally considered that the model fit is higher when  $R^2$  exceeds 0.8.

Because ①②③ use the mean error, and the mean error is more sensitive to outliers, if there is a large difference between a certain regression value fitted by the regressor and the true value, it will lead to a large average error, which will have a great influence on the final evaluation value, that is, the mean value is not robust. According to Table 4, it is found that the value of ①②③ increases with the increase of  $y$ , so the reliability is low. By comparing and analyzing the curve fitting effect in Figures 2–13, in combination with the value of evaluation indicators in Tables 4–6, it is found that the evaluation value of ④⑤⑥ is more consistent with the curve fitting effect. By analyzing and comparing the above evaluation indicators, take ④⑤⑥ as the main index to evaluate traffic flow prediction and ①③ as the auxiliary index.

#### 4. Short-Term Traffic Flow Prediction Algorithm Based on LSTM-GASVR

*4.1. Optimize the Model Based on GA-SVR.* Genetic algorithm (GA) is used to optimize the parameters  $(C, \varepsilon, r)$  in the SVR model, where the  $C$  is the penalty coefficient,  $r$  is the kernel function coefficient, and  $\varepsilon$  is the insensitive coefficient, which has the following main steps:

- (1) Constructing the chromosome assemblage  $(C, \varepsilon, r)$  and formulating the fitness calculation function of genetic algorithm
- (2) Determining the parameters of selection, crossover, mutation, and so on in the GA, and setting the iterative termination condition of the algorithm
- (3) Initializing the GA and generating the initialization population
- (4) Calculating individual fitness in chromosome populations
- (5) Generation of next-generation chromosomes through selection, crossover, variation, etc.

TABLE 4: Comparison of short-term traffic flow prediction effect at different time intervals.

Time intervals	$R^2$	Explanatory variance	MAPE	MAE	RMSE
5	0.980	0.980	0.232	7.705	11.532
10	0.979	0.981	0.186	13.408	19.777
15	0.982	0.982	0.160	21.406	27.960
20	0.980	0.980	0.156	28.095	40.394
25	0.974	0.977	0.219	36.871	54.758
30	0.971	0.972	0.168	49.940	69.037

- (6) If the iterative termination condition of the algorithm is satisfied or not, jump to (4).
- (7) Termination of the iteration to determine the optimal parameters  $(C, \varepsilon, r)$ . Figure 14 shows a flowchart of GA solving SVR optimal parameters  $(C, \varepsilon, r)$ .

*4.2. Chromosome Coding.* Genetic method is based on chromosome coding, selection, crossover, and mutation, and other operations of the algorithm effectively depend on the chromosome coding method to some extent. Considering that the SVR model parameter selection itself is a constrained problem, we code the SVR model parameter by real number. That is the gene of each chromosome consists of three decimal float.

*4.3. Population Initialization.* Population number is an important parameter that affects the efficiency and convergence of the algorithm. In this paper, the population size  $\text{pop\_size}$  is set to 50, and the population initialization state is evenly distributed in the solution space. And  $C \in [275, 285]$ ,  $\varepsilon \in [0.4, 0.6]$ , and  $\sigma \in [0.001, 0.003]$ .

*4.4. Fitness Function Selection.* GA is used to find the best parameters of the SVR model, so the fitness function is chosen as average relative error percentage between the prediction results and the actual value, and the fitness function can be designed as follows:

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^n \frac{|y_i - f(x_i)|}{y_i}. \quad (10)$$

Among them, the MAPE is the average relative error percentage between the prediction results and the actual value of the algorithm.  $t_0$  is the initial time of training data,  $y_i$  represents the actual driving speed of the road at the time  $t_0 + i \times 15$  min, and the prediction result of the time  $t_0 + i \times 15$  min is  $f(x_i)$ .

*4.5. Selection Operations.* The selection operation used in this paper is the combination of roulette and elite strategy. The key of the roulette selection method is to produce offspring population by calculating the probability of each individual appearing in the offspring, that is, the greater the probability of individual selection when the adaptation value is higher. The probability of individual appearance is shown in Table 7. Roulette selection algorithm refers to the

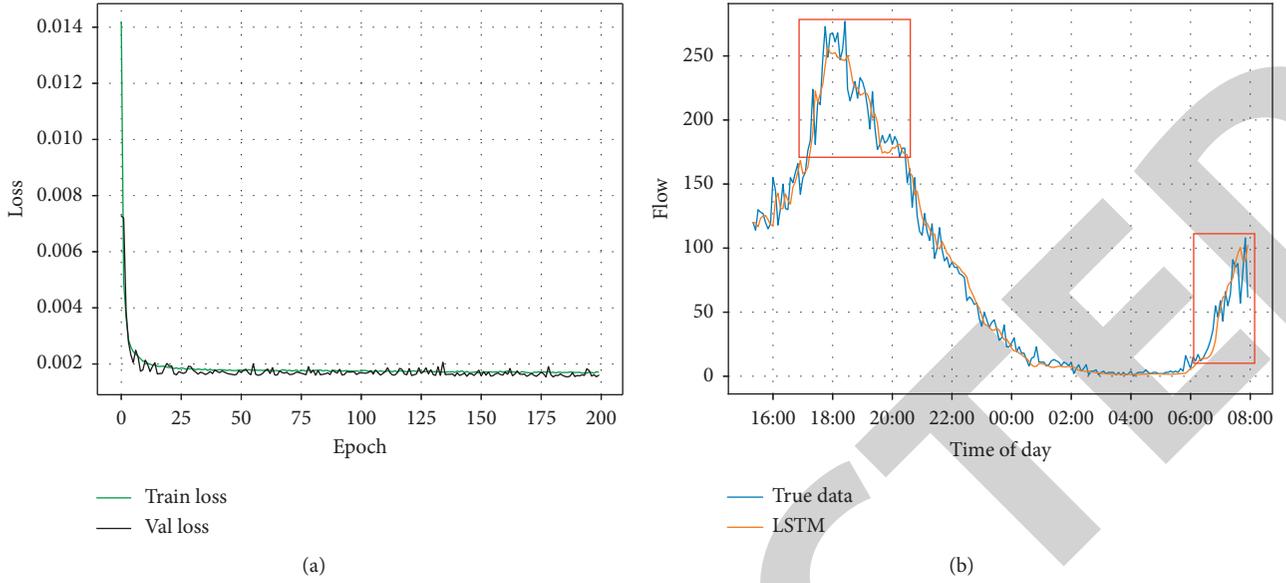


FIGURE 2: 5 min LSTM training and prediction. (a) 5 min LSTM loss function and (b) 5 min LSTM prediction results.

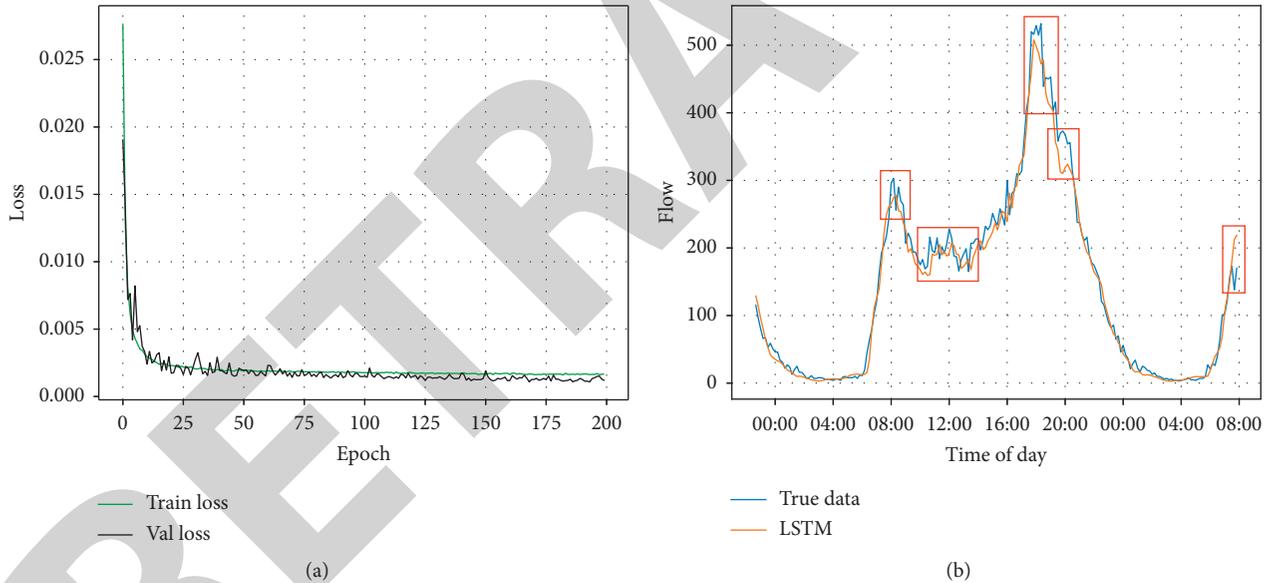


FIGURE 3: 10 min LSTM training and prediction. (a) 10 min LSTM loss function and (b) 10 min LSTM prediction results.

generation of random numbers in the interval  $[0, 1)$ , which falls within a certain probability interval shown in Table 7, which corresponds to individuals inherited to the next generation. For example, the random number generated is 0.3,  $0.3 \in [0.2, 0.45)$ , corresponding to individual 2, so individual 2 is inherited to the next generation. The disadvantage of this method is that it is easy to fall into local optimum when the range of adaptive value interval is small.

The elite strategy is to keep the better individuals in the last generation population and increases the number of the better individuals so as to guarantee the global optimum, but it is easy to fall into the local optimum when the proportion of the better individuals is large. Elite strategies retain better

individuals by copy better individuals into the next generation or cross and mutation better individuals in the population, and we select the latter method.

Therefore, the paper adopts the combination of roulette and elite strategy to keep the better population number in elite strategy proportionally dynamic, so as to get the solution of the population number, as shown in the following equation:

$$\text{GoodPop\_size} = \min\left(\text{Pop\_size}, \frac{k \times \text{current\_Gen}}{\text{Max\_Gen}}\right). \quad (11)$$

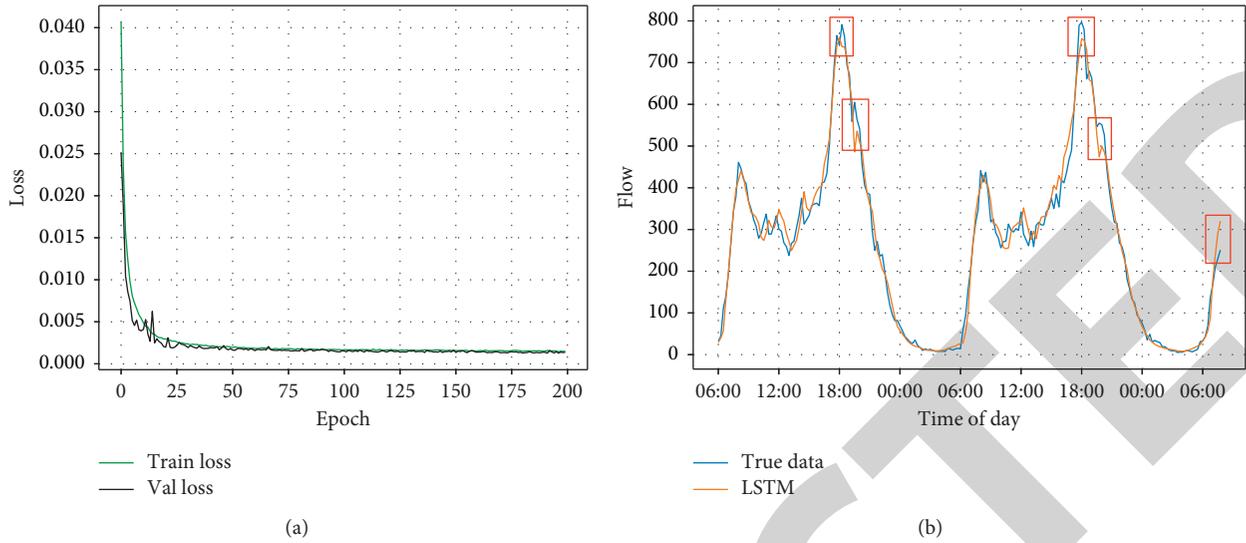


FIGURE 4: 15 min LSTM training and prediction. (a) 15 min LSTM loss function and (b) 15 min LSTM prediction results.

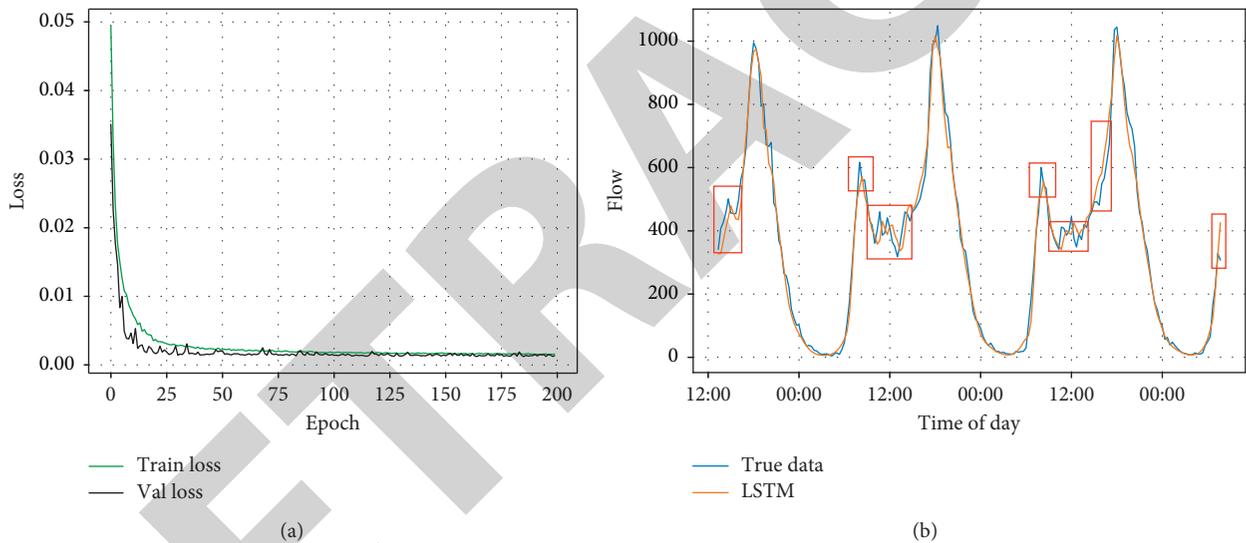


FIGURE 5: 20 min LSTM training and prediction. (a) 20 min LSTM loss function and (b) 20 min LSTM prediction results.

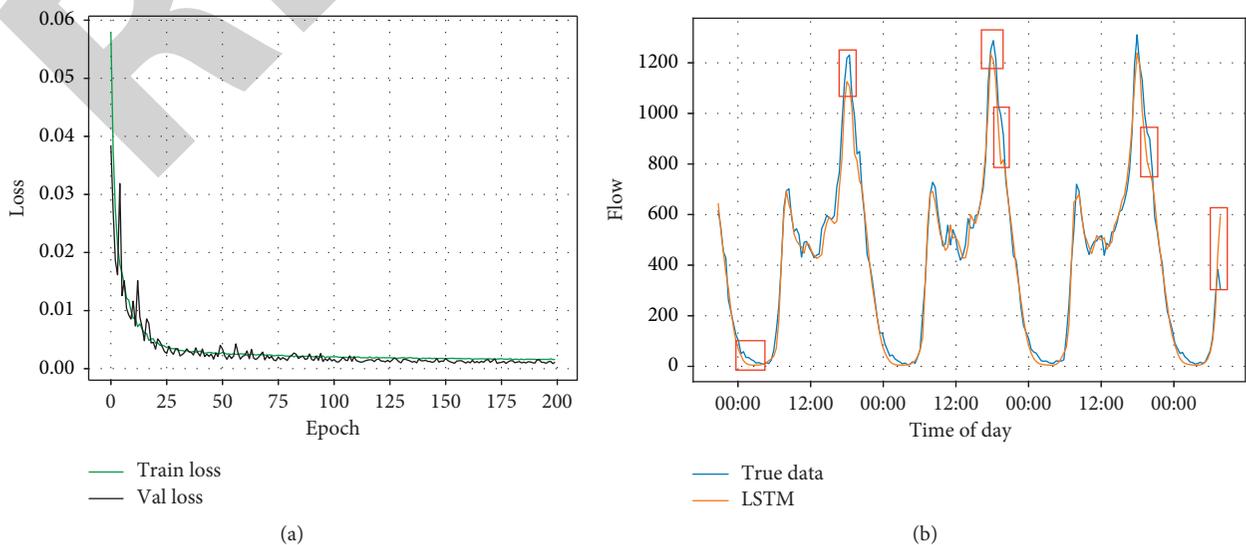


FIGURE 6: 25 min LSTM training and prediction. (a) 25 min LSTM loss function and (b) 25 min LSTM prediction results.

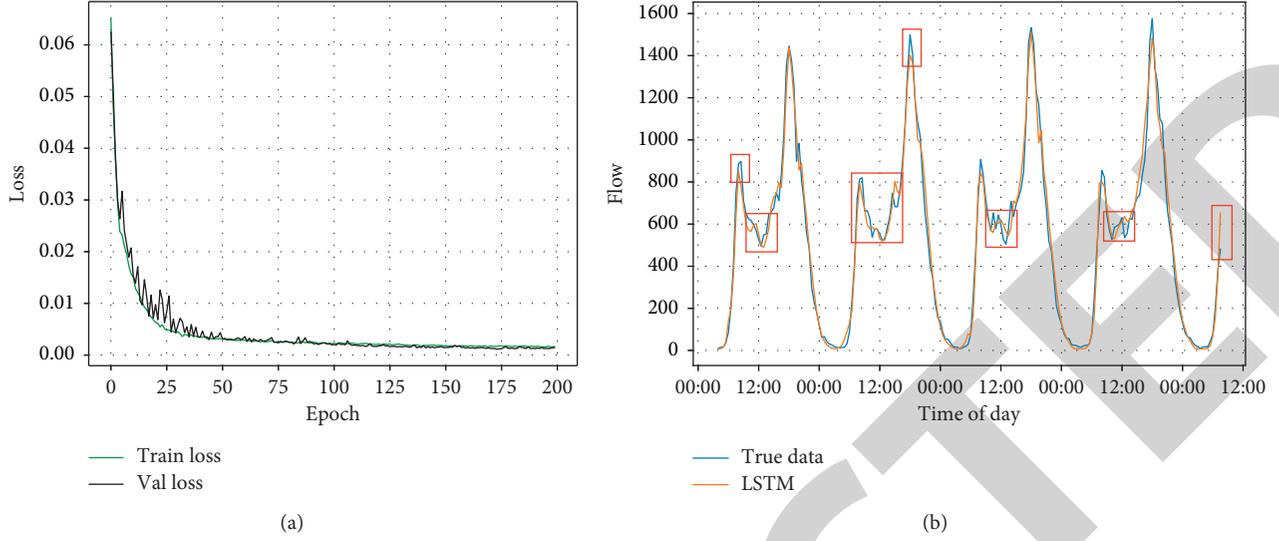


FIGURE 7: 30 min LSTM training and prediction. (a) 30 min LSTM loss function and (b) 30 min LSTM prediction results.

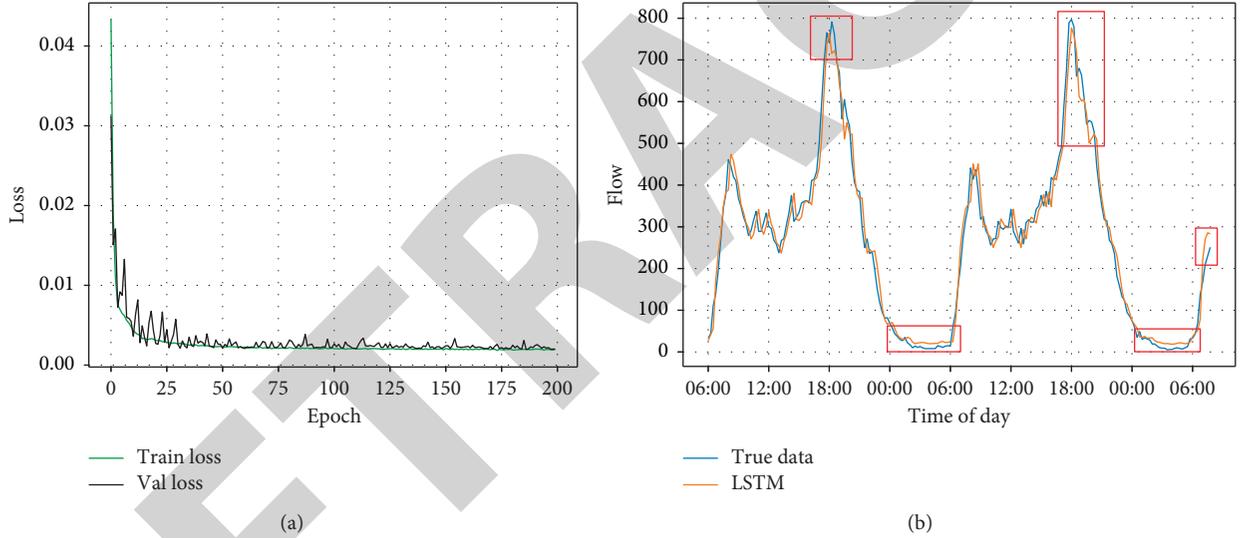


FIGURE 8: 4 steps LSTM training and prediction. (a) 4 steps LSTM loss function and (b) 4 steps LSTM prediction results.

The GoodPop\_size in equation (11) is to retain the number of better individuals, Pop\_size represent the number of populations,  $k$  is the sum of the parent and the offspring at each run time of the algorithm, current\_Gen is the number of runs of the current algorithm, and Max\_Gen is the maximum number of iterations set in advance for the algorithm. Equation (11) provides a corresponding guarantee for the efficiency and optimization performance.

**4.6. Genetic Operator Design.** Genetic operator mainly includes cross operator and mutation operator. Cross operation is one of the main operations of offspring innovation in GA. By simulating the cross inheritance of chromosome weight, the individual genes are exchanged according to certain probability. The mutation operation is to simulate the process of chromosome compilation and recombination

to change some genes in individuals according to certain probability. Cross operator and mutation operator are a way to produce innovation.

**4.6.1. Cross Operator.** During the evolution of natural organisms, the recombination of biological genetic genes is very important, and cross operators cannot be replaced in GA. The arithmetic cross operator is selected in the paper, as shown in equation (12).  $a$  in equation (12) represents the random number between 0 and 1 of the values of cross probability.  $P'_i$  and  $P'_{i+1}$  represent gene value  $P_i$  and  $P_{i+1}$  after cross operator, and  $i$  and  $i + 1$  are gene positions:

$$P'_i = aP_i + (1 - a)P_{i+1}, \quad P'_{i+1} = (1 - a)P_{i+1} + aP_i. \quad (12)$$

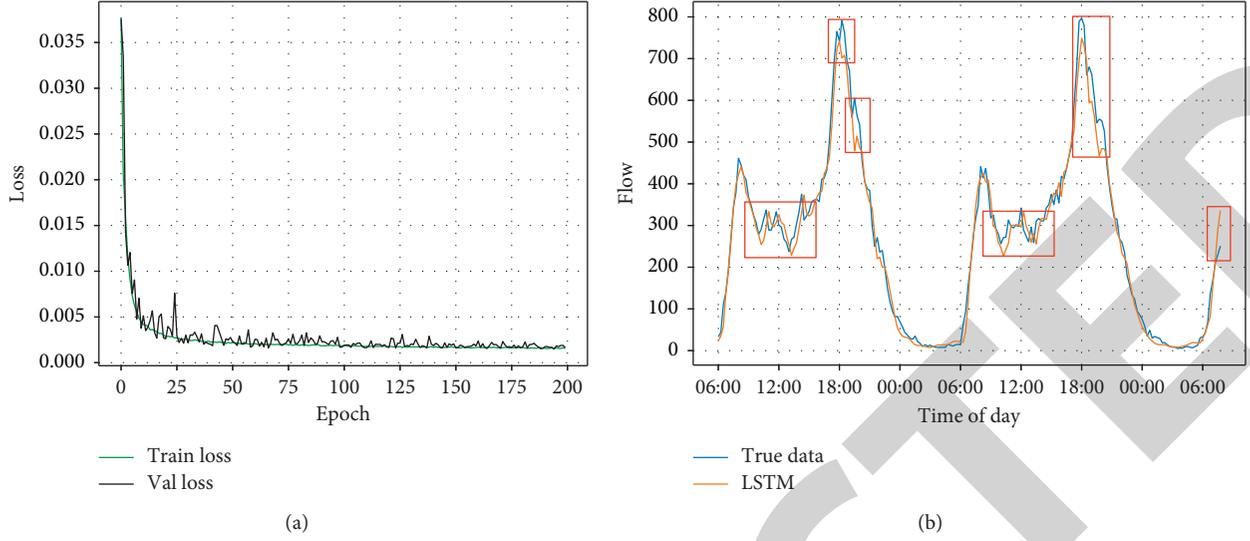


FIGURE 9: 8 steps LSTM training and prediction. (a) 8 steps LSTM loss function and (b) 8 steps LSTM prediction results.

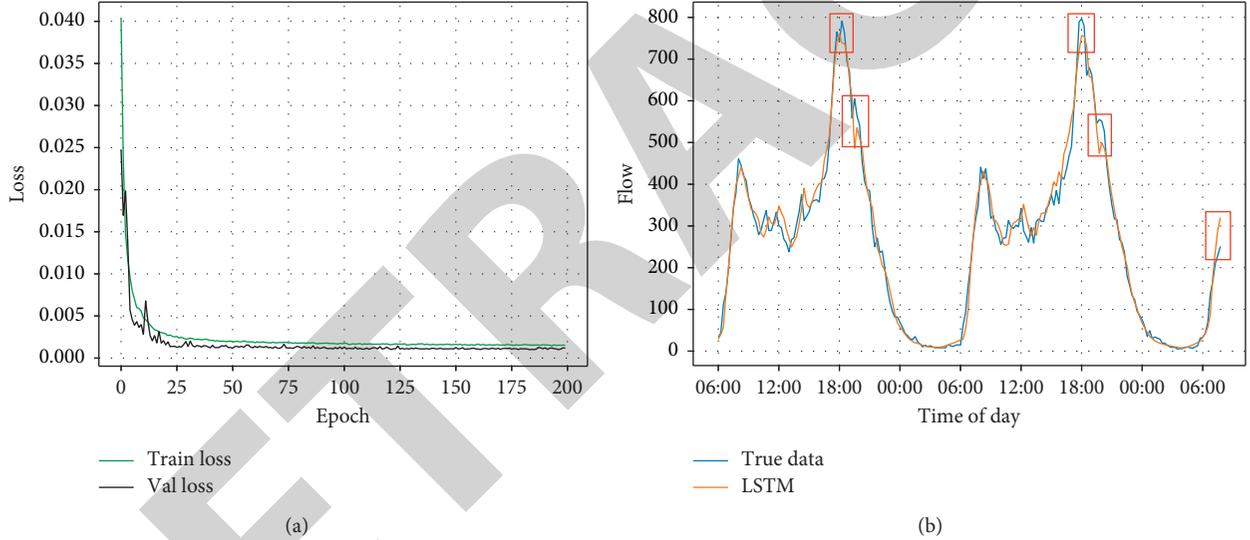


FIGURE 10: 12 steps LSTM training and prediction. (a) 12 steps LSTM loss function and (b) 12 steps LSTM prediction results.

**4.6.2. Mutation Operator.** Mutation operator is the operation of changing individual gene value  $x_j$  in chromosome population, as shown in equation (13). A random function  $\text{Random}(a_i, b_i)$  is used to generate a random number, and  $a_i$  and  $b_i$  are the upper and lower limits of the gene. And a gene position  $j$  in the chromosome is set to the value generated by the random function:

$$x_j = \begin{cases} \text{Random}(a_i, b_i), & i = j, \\ x_j, & i \neq j. \end{cases} \quad (13)$$

**4.7. Stop Sign.** The stop flag is the maximum number of iterations set for the experiment, and the value is set to 200 in the experiment.

During the GA training as shown in Figure 15, the blue boxes indicates where the optimal value goes up and the

optimal value increased to 0.86 after 30 iterations. The red curve in the figure represents the accuracy of the optimal parameter SVR of the prediction model. Because of the elite retention strategy, the green curve gradually approaches red, that is, the average accuracy tends to the optimal value. The SVR model parameters of the final GA tuning parameter ( $C, \epsilon, r$ ) are (276.7, 0.05998, 0.001595).

**4.8. LSTM Prediction of Short-Term Traffic Flow Based on GA-SVR Optimization.** Based on the research and analysis of various algorithms, it is found that the LSTM model prediction evaluation index  $R^2$  is higher, that is, the fitting degree between the predicted data and the real traffic flow data is higher, which is more consistent with the real data trend, and the prediction results are more reliable. GA-SVR prediction model evaluation index  $MAPE$  is smaller, the

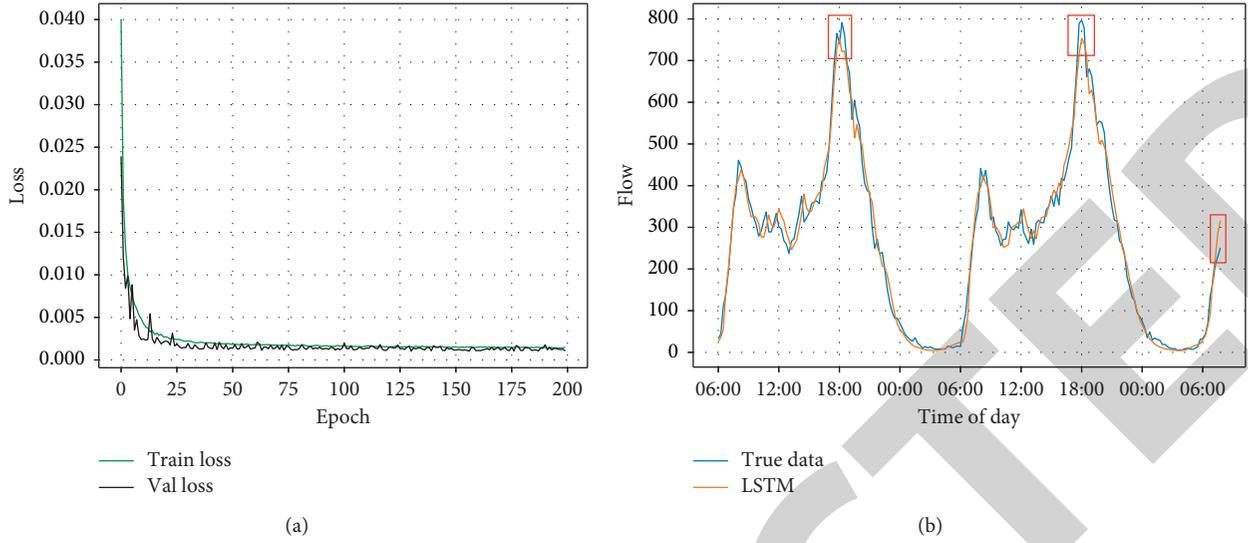


FIGURE 11: 16 steps LSTM training and prediction. (a) 16 steps LSTM loss function and (b) 16 steps LSTM prediction results.

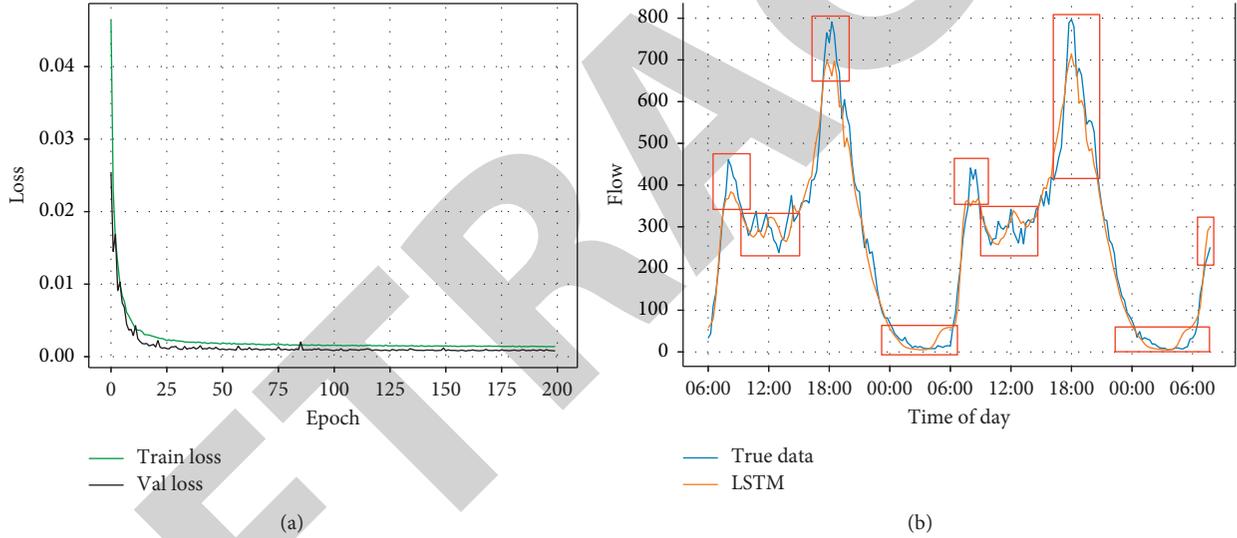


FIGURE 12: 20 steps LSTM training and prediction. (a) 20 steps LSTM loss function and (b) 20 steps LSTM prediction results.

model prediction value is closer to the real traffic flow data numerically, and the error is smaller. To balance the fitting degree and error of short-term traffic flow prediction model, this neural network model LSTM combined with SVR is optimized by GA, and the short-term traffic flow prediction model based on LSTM-GASVR is obtained.

The structure of LSTM-GASVR is shown in Figure 16. The neural network is partly composed of the input layer, two layers of LSTM, dropout layer, and full connection layer. The output of the neural network is processed by the SVR to obtain the output results. Next moment workday flags  $h^m$  input to the first layer contain 64 LSTM networks and get 64 features. The 64 features and  $v_0$ , which is the traffic volume of period 0 to form a new vector, then input to the second layer in the network and get 64 features. These characteristics and traffic flow data of the next time step constitute a new vector input to 64 LSTM networks and get 64 LSTM characteristics.

And so on, the first LSTM layer outputs  $(m + 1) * 64$  data as input of the second LSTM layer. The output from last time step of the second LSTM will input to the dropout layer; finally, the results will be output by the dense layer of the model. The average of the results and the traffic flow at the previous moment inputs to the SVR model and then outputs the final results. "A in Figure 16" is shown in Figure 17, it is LSTM unit structure.  $C_{t-1}$  is the cell state,  $h_{t-1}$  is the output of the last LSTM, and  $X_t$  is the input of LSTM.  $f_b$ ,  $i_b$ , and  $o_t$  are output of forget gate, input gate, and output gate.  $\sigma$  and  $\tanh$  is activation function.  $C_t$  is the new cell state, and  $h_t$  is the output of LSTM.  $t-1$  means the last moment, and  $t$  is the current moment.

The LSTM-GASVR prediction process is shown in Figure 18. The input traffic data are preprocessed using MinMaxScaler normalization. According to the current short-term traffic flow data affected by the previous  $m$  data

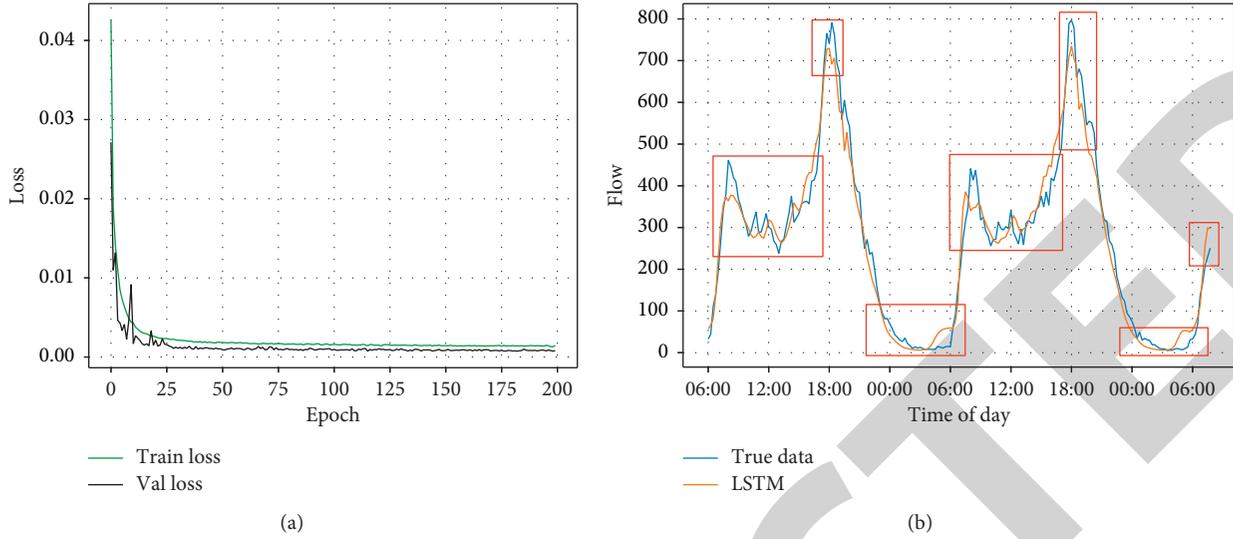


FIGURE 13: 24 steps LSTM training and prediction. (a) 24 steps LSTM loss function and (b) 24 steps LSTM prediction results.

TABLE 5: Comparison of short-term traffic flow prediction effect in different time steps.

Time step	$R^2$	Explanatory variance	MAPE	MAE	RMSE
4	0.976	0.976	0.285	24.134	32.012
8	0.976	0.979	0.177	22.846	31.753
12	0.979	0.980	0.155	22.304	30.140
16	0.983	0.983	0.153	19.758	28.595
20	0.965	0.966	0.329	28.814	38.456
24	0.963	0.963	0.394	30.845	39.793

TABLE 6: Comparison of short-term traffic flow prediction effect of different prediction algorithms.

Algorithm	$R^2$	Explanatory variance	MAPE	MAE	RMSE
LSTM	0.983	0.983	0.153	19.758	28.595
GRU	0.982	0.982	0.206	20.781	27.233
CNN	0.952	0.962	0.293	33.297	45.206
SAE	0.976	0.976	0.374	24.813	31.744
ARIMA	0.975	0.975	0.222	23.313	32.381
GASVR	0.966	0.967	0.144	26.092	37.995
LSTM-GASVR	0.982	0.982	0.118	20.156	28.237

and the workday flag, the array is reconstructed to  $[m + 1 : 1]$ . The data are divided into training set and test set, training set is used to train the model, and test set is used to test the prediction effect of the model. The data input to the neural network composed of LSTM, dropout, and dense, and then training of neural network. And the test data input neural network to make the preliminary prediction. The prediction results are renormalized, and the data are reset to the average of the LSTM model prediction results and the traffic flow at the previous time. The average input SVR model then optimizes the parameters of the model by the GA. After that, SVR model is trained according to the optimal parameters.

Finally, input the test data into the LSTM-GASVR model for prediction, and output the prediction results.

The prediction effect of the LSTM-GASVR model is shown in Figure 19. The time interval of the model is 15 minutes and the time step is 16. The green line in the figure represents the real traffic flow, and the red line represents the LSTM-GASVR model prediction value. It can be seen the prediction results of the LSTM-GASVR model are very close to the real value, which shows slightly higher error of prediction results at peak, but the error is small on the whole.  $R^2$  is 0.982, *explanatory variance* is 0.982, and MAPE is 0.118, and these are evaluation indicators of the model.

## 5. LSTM Prediction Model Parameter Determination

The batch size is 64, the epoch is 200, and the loss function is “rmsprop” as quantitative in the training of the LSTM model. The optimal performance of the model under different sampling time intervals and time steps is analyzed to determine the optimal sampling interval and time steps.

### 5.1. LSTM Prediction Model Sampling Interval Determination.

LSTM network model predicts the short-term traffic flow data with different time interval downsampling; then, we compare and analyze the training speed, loss function value, and the performance of the prediction results under different sampling intervals to determine the most suitable LSTM network model under the sampling interval.

The LSTM, trained with 5 min downsampling data, is shown in Figure 2(a). The loss function reaches a stable value of 0.002 at 25 epochs. As shown in Figure 2(b), the part marked by the red boxes is the larger part of the prediction difference. From Figure 2, the 5 min LSTM model is trained quickly, but the error of prediction results is large.

The LSTM of the trained 10 min downsampling data is shown in Figure 3(a). The loss function reaches a stable value

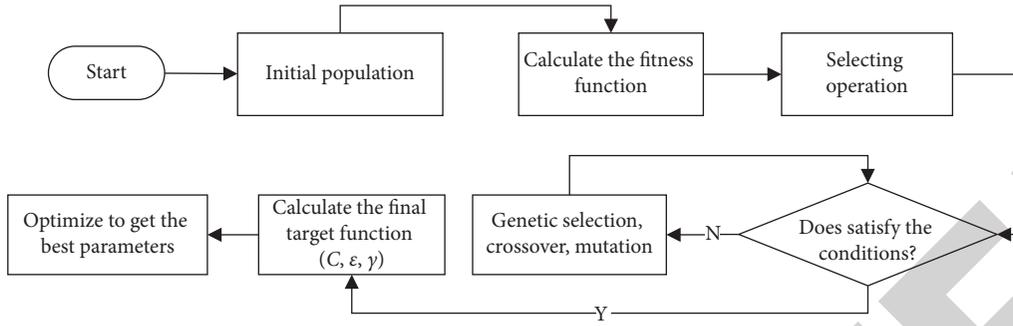


FIGURE 14: Flowchart of genetic algorithm.

TABLE 7: Probability interval.

Individual	1	2	3	4
Probability of individuals	0.2	0.25	0.35	0.2
Probability interval	[0, 0.2)	[0.2, 0.45)	[0.45, 08)	[0.8, 1)

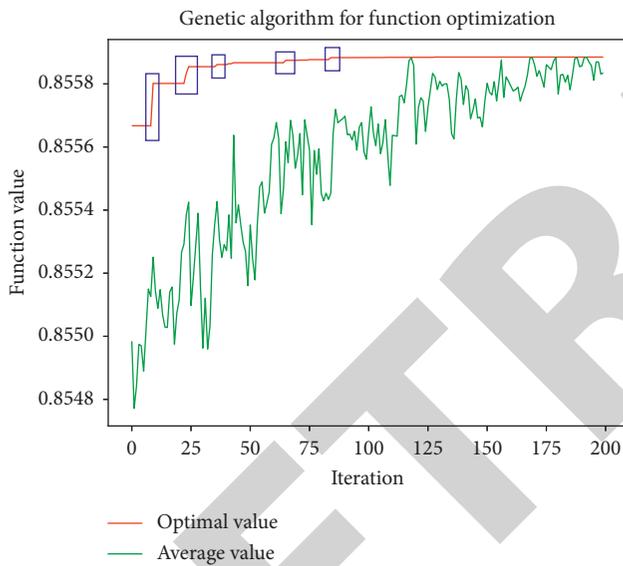


FIGURE 15: GA training.

of 0.002 at 50 epochs. As shown in Figure 3(b), the prediction difference is mostly labeled in the red boxes. From Figure 3, we can see that the LSTM model training speed of 10 min data is slow, and there are some deviations from the prediction results.

The LSTM of the analyzed 15 min downsampling data training is shown in Figure 4(a), and the loss function reaches a stable value of 0.002 at 50 epochs. As shown in Figure 4(b), the part marked in the red boxes is a larger part of the prediction difference. It can be seen from the figure that the 15 min LSTM model has a fast training speed, and the prediction results are generally good, and the errors at the peak and fluctuation are larger.

The LSTM of 20 min downsampling data training is analyzed. As shown in Figure 5(a), the epoch is 75 where the loss function is stable at 0.003. As shown in Figure 5(b), the prediction difference is mostly marked in red boxes. From

Figure 5, the 20 min LSTM model training speed is better, the value of loss function is slightly higher, the prediction results are general, and the prediction error is large at peak and fluctuation.

The LSTM model obtained from the training of 25 min downsampling data is shown in Figure 6(a). The stable value of the loss function is 0.003 at 75 epochs. As shown in Figure 6(b), the part marked by the red boxes is the larger part of the prediction difference. From Figure 6 the 25 min LSTM model training speed is slightly slow, the value of the loss function is slightly higher, the prediction results are general, and the prediction error is large at the peak and the fluctuation.

The LSTM, trained under 30 min downsampling data, is shown in Figure 7(a). The loss function reaches a stable value of 0.003 at 75 epochs. As shown in Figure 7(b), the red boxes indicate that the prediction results showed slightly larger error. From the figure, the 30 min LSTM model training speed of the data is slightly slower, the loss function value is higher, the prediction results are general, and the prediction result error is larger at peak and fluctuation.

Since the difference judgment of the fitting curve is difficult, the prediction results of the LSTM network model with different sampling intervals are again compared and analyzed by various evaluation indexes. As shown in Table 4, the model evaluation trained by the data of 15 min is the best.

## 5.2. LSTM Prediction Model Time Step Determination.

The LSTM model predicts the short-term traffic flow data of 15 min downsampling with different time steps  $n$  (the current moment is affected by the previous  $n$  moment). The most suitable time step is selected by analyzing the stability of the training process, the value of the loss function, and the curve of the fitting degree of the predicted results to the real value.

The LSTM model is trained with 15-minute interval downsampling data at 4 time steps. Its loss function image is

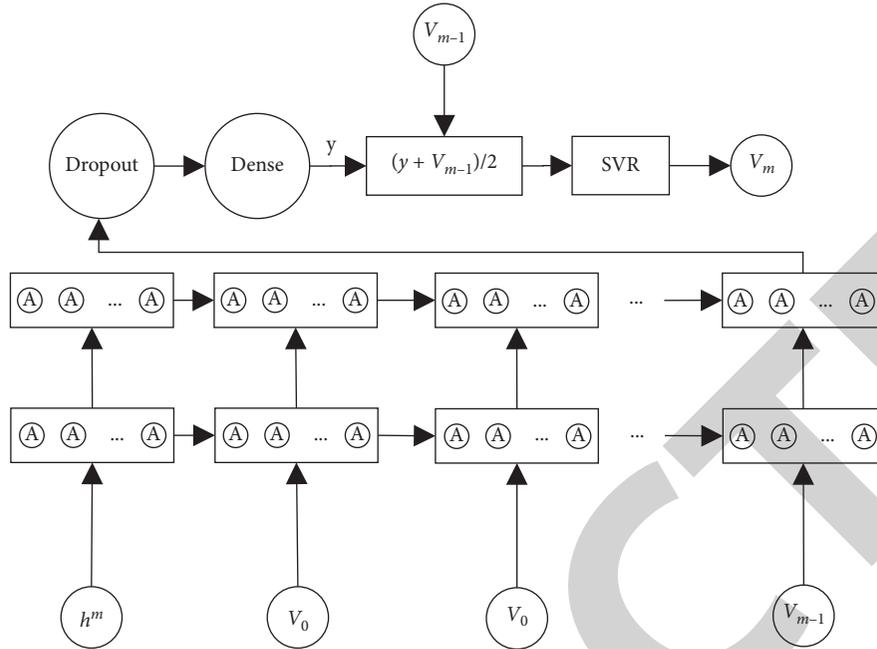


FIGURE 16: LSTM-GASVR structure.

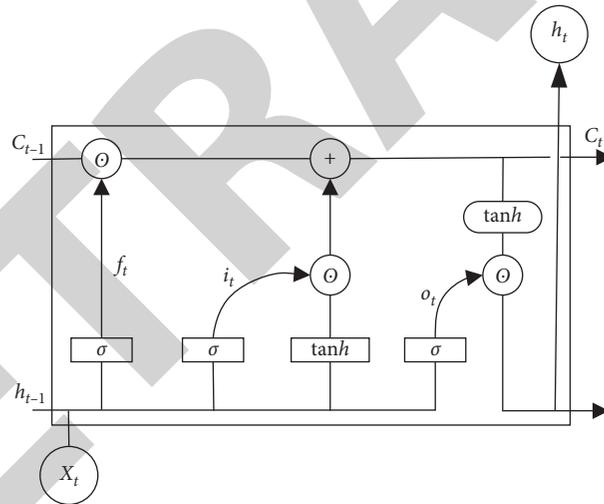


FIGURE 17: The structure of LSTM.

shown in Figure 8(a), and the loss function is stable around 0.002 at 50 epochs. The prediction results are shown in Figure 8(b). The red boxes are marked with a large prediction difference. The training speed of the LSTM model with 4 time steps is slightly slower, the loss function value is higher, the prediction results are general, and the prediction result error at peak is very high.

Analyzing the 15 min of downsampling data trained at 8 time steps LSTM, as shown in Figure 9(a), the loss function is stabilized around 0.003 at 75 epochs. As shown in Figure 9(b), the part marked in the red boxes shows that the prediction difference is large. The training speed of the LSTM model with 8 time steps is slow, the loss function value is slightly larger, the overall prediction results are

general, and the prediction results at peak and fluctuation are obviously poor.

The LSTM model of downsampling data trained at 12 time steps and the loss function are shown in Figure 10(a), the loss function reaches a stable value of 0.002 at 50 epochs, the prediction results are shown in Figure 10(b), and the prediction difference is mostly marked in red boxes. It can be seen from the figure that the LSTM model with 12 time steps has a fast training speed, accurate prediction results, but a large error in the prediction results at the peak and with large fluctuations.

The LSTM model, trained with 15 min interval downsampling data at 16 time steps, is shown in Figure 11(a), and the loss function reaches a stable value of 0.003 at 50 epochs. As shown in Figure 11(b), the red boxes are marked with a

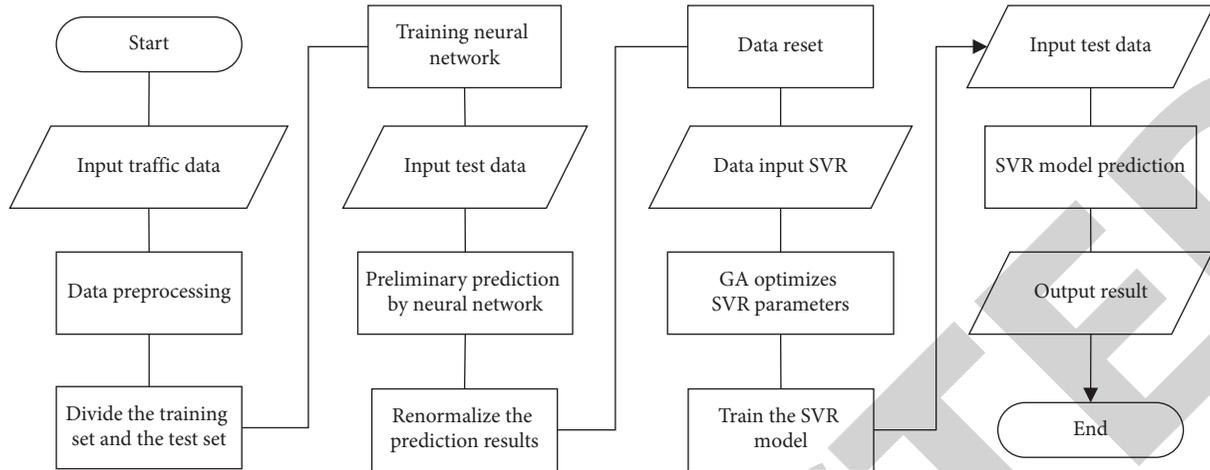


FIGURE 18: LSTM-GA SVR prediction process.

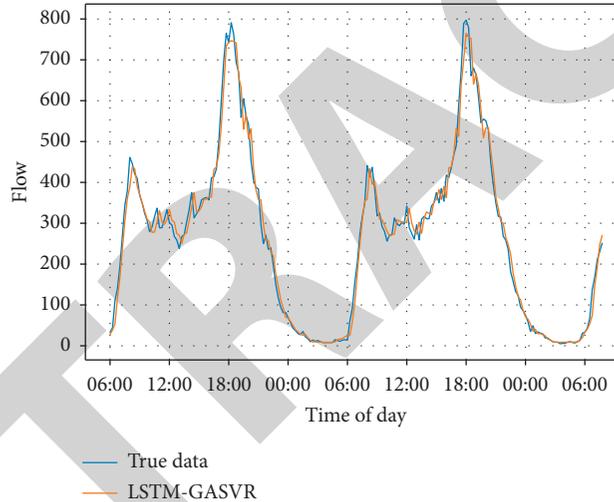


FIGURE 19: LSTM-GASVR prediction results.

large difference in prediction. The LSTM model with 16 time steps shows better training speed, and the loss function value is low. Overall, the prediction results are accurate, but slightly higher error of prediction results at peak.

The LSTM with downsampling interval of 15 min was trained at 20 time steps. As shown in Figure 12(a), the loss function at 50 epochs is stable at around 0.002. As shown in Figure 12(b), the prediction difference of red boxes part is large. The training speed of LSTM model with 20 time steps is general, the prediction results are general, and the error of peak, fluctuation, and valley bottom part is very high.

The LSTM model were trained of the data sampled at intervals of 15 min at 24 time steps, and the loss function is shown in Figure 13(a), reaching a stable value of 0.003 at 50 epochs. As shown in Figure 13(b), the part marked by the red boxes is the larger part of the prediction difference. From the figure, the training speed of LSTM model with 24 time steps is general, and the predicted results are general. The prediction error is large at the peak and the fluctuation, and the prediction results of the valley bottom are very high too.

The LSTM model predicts the short-term traffic flow data of 15 min downsampling with different time steps  $n$  (the current moment is affected by the previous  $n$  moment). Then, we compare the prediction results according to various evaluation indexes. As shown in Table 5, it is found that the prediction effect of 16 time steps is the best.

The optimal performance of the LSTM model under different sampling time intervals and time steps is analyzed, and the sampling interval is finally determined to be 15 min at 16 time steps. The LSTM model training speed is the fastest at this time, the loss function reaches the stable value at 25 epochs and the loss function value is small, and the prediction accuracy is the highest.

## 6. Comparative Analysis of Short-Time Traffic Flow Prediction Model

Through LSTM, GRU (gated recurrent unit), CNN (convolutional neural networks), SAE (stacked autoencoder), ARIMA (auto regressive integrated moving average), SVR,

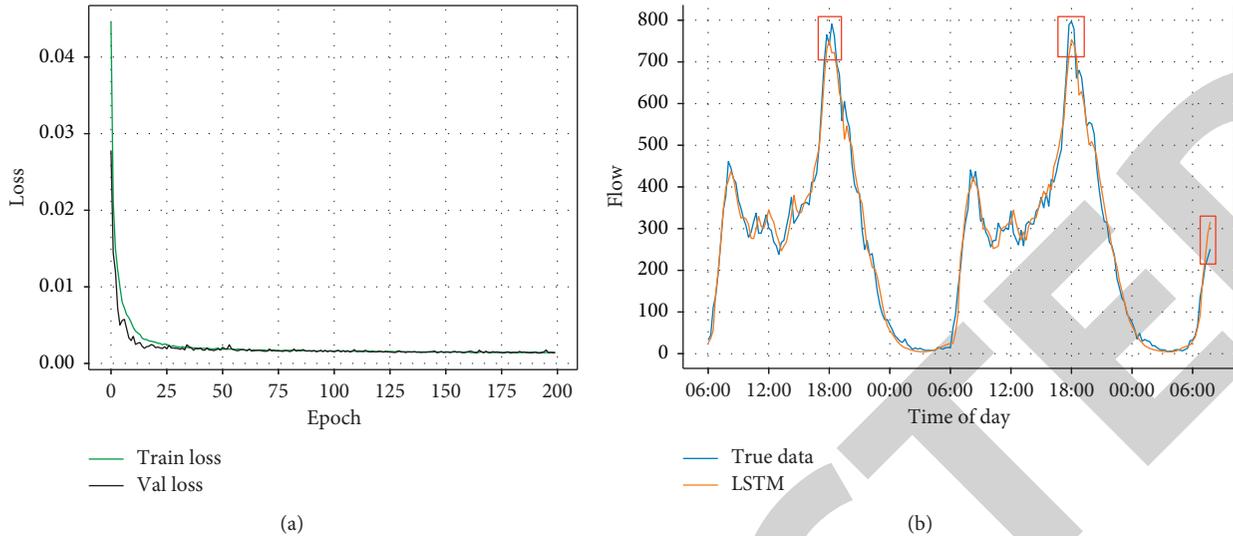


FIGURE 20: LSTM training and prediction. (a) LSTM loss function and (b) LSTM prediction results.

LSTM-GASVR prediction of 15 minutes short time traffic volume, in this paper, the training speed and loss function of LSTM, GRU, CNN, and SAE in training are compared and analyzed, the prediction results of seven algorithms are compared to fit the curve of real value, and the prediction effect is analyzed.

Analyzing the LSTM model trained with the down-sampling interval of 15 min data at the 16 time steps, as shown in Figure 20(a), the loss function reaches a stable value of 0.002 at 25 epochs. As shown in Figure 20(b), the part marked in the red boxes is the larger part of the prediction difference, the training speed of the LSTM model is faster, the prediction results are generally more accurate, and the error of the prediction results at the peak is slightly larger.

The data sampled by 15 min are trained to GRU model at the 16 time steps, and the loss function is stable around 0.002 at 75 epochs as shown in Figure 21(a). As shown in Figure 21(b), the partial prediction difference is marked by the red boxes. From the figure, we can see that the model training speed is general, the prediction accuracy is general, and the prediction result error at the peak is very large.

The 1D CNN model of 15 min downsampling data trained at 16 time steps is analyzed. As shown in Figure 22(a), the 50-epoch training sets get a stable training set loss function value of 0.007, and the validation set loss function stability value of 0.003. As shown in Figure 22(b), the red boxes are marked with a large difference in prediction, the model training speed is general, the loss function value is large, the prediction accuracy is not high, and the prediction error is large.

The SAE model trained with 15 min downsampling data at 16 time steps is analyzed. The loss function reaches a stable value of 0.001 at 100 epochs, as shown in Figure 23(a). As shown in Figure 23(b), the difference in prediction is mostly marked in red boxes, the SAE model training speed is

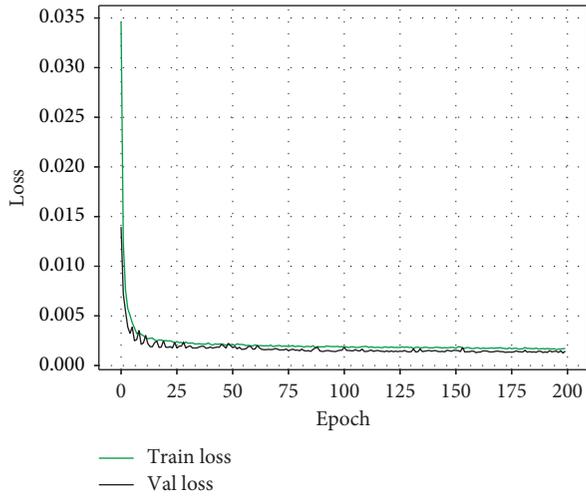
average, the loss function value is small, the prediction accuracy is not high, and the prediction effect is very poor at peak, fluctuation, and turning point.

The ARIMA model and the GASVR model were used to predict 15 min of short-term traffic flow, respectively. The prediction results of the ARIMA model are shown in Figure 24, the red boxes are marked with a large part of the prediction difference, it can be seen that the prediction accuracy of peak value is low, and it takes a long time in the prediction process of the ARIMA model, which does not meet the real-time requirement of short-term traffic flow prediction. As shown in Figure 25, for the prediction results of the GASVR model, the red boxes are marked with a large part of the prediction difference. It can be seen that the prediction accuracy is slightly lower at the peak value, and the GASVR model prediction curve is shifted backward compared with the real value.

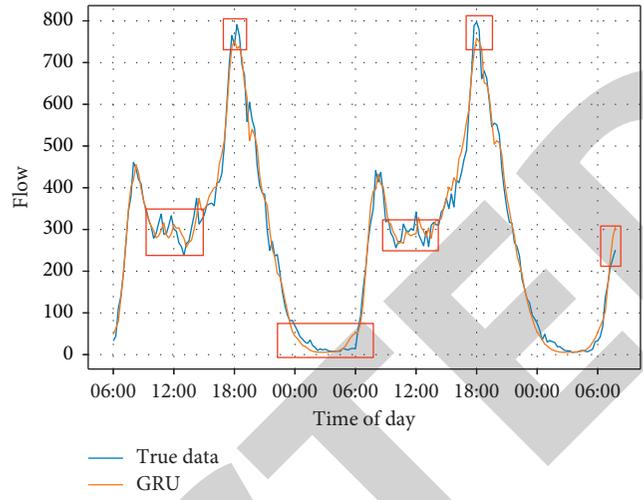
In Figure 26, the prediction results of the LSTM-GASVR model show that the model effectively improves the migration phenomenon of the GASVR and improves the accuracy of the LSTM model. These are some differences were marked by the red boxes, the prediction accuracy of the LSTM-GASVR model is not enough at the peak of traffic flow. But it has a little influence on the results of the algorithm, and the LSTM-GASVR model predicted the most accurate results.

The timeliness of the LSTM-GASVR model is normal compared with other algorithms. The prediction time of the model is 0.003 s longer than GASVR, and the time is 2 s shorter than ARIMA. Compared with the LSTM model, the prediction time of the LSTM-GASVR is 0.001 s longer than that of the LSTM model. In addition, compared with CNN, SAE, and GRU models, the LSTM-GASVR model takes time to predict that traffic flow is same almost.

However, the LSTM-GASVR model's timeliness is normal compared with other algorithms and the accuracy is

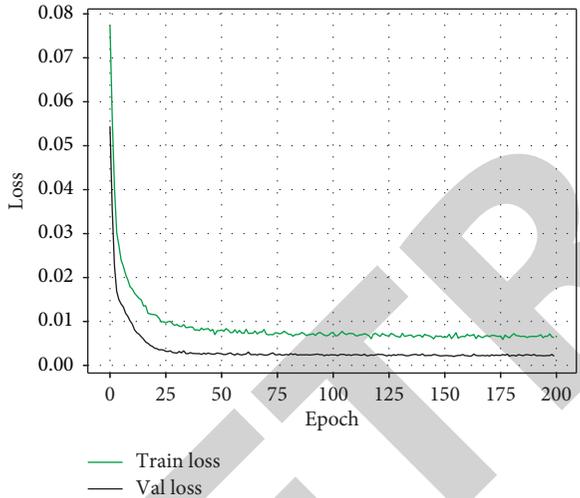


(a)

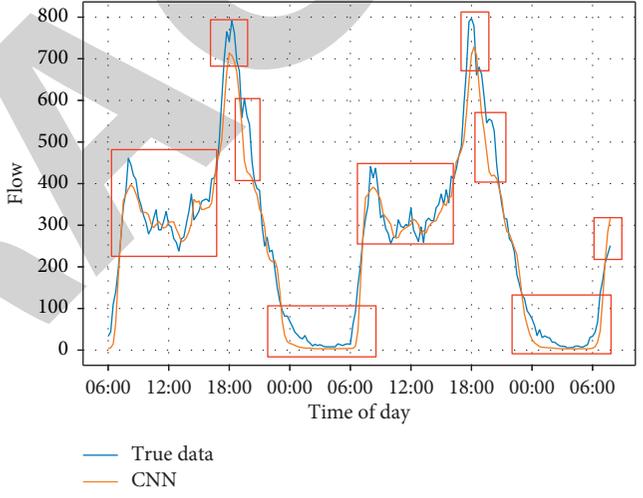


(b)

FIGURE 21: GRU training and prediction. (a) GRU loss function and (b) GRU prediction results.

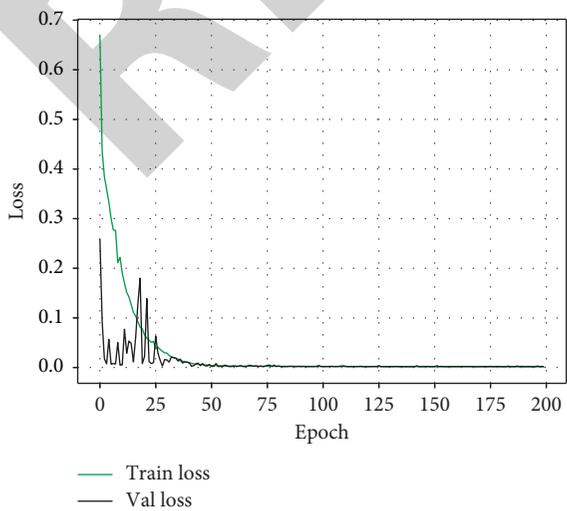


(a)

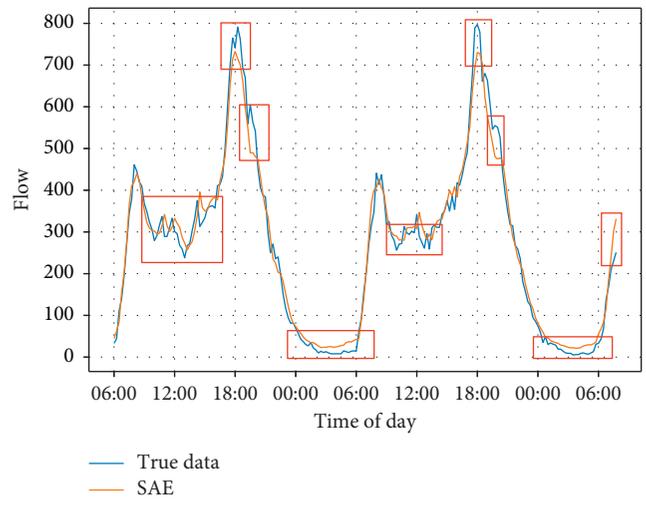


(b)

FIGURE 22: CNN training and prediction. (a) CNN loss function and (b) CNN prediction results.



(a)



(b)

FIGURE 23: SAE training and prediction. (a) SAE loss function and (b) SAE prediction results.

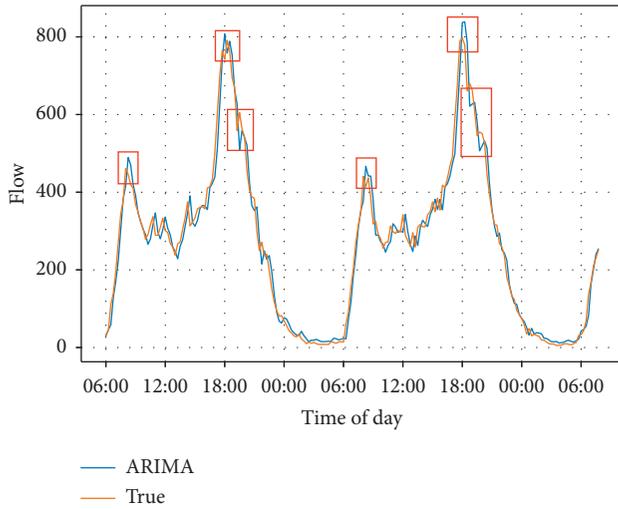


FIGURE 24: ARIMA prediction results.

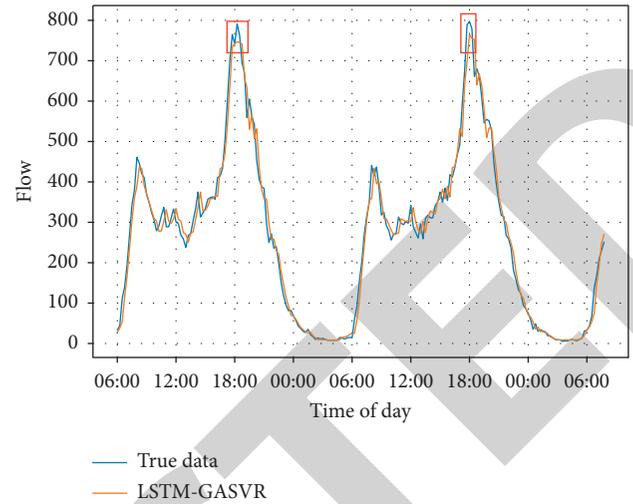


FIGURE 26: LSTM-GASVR prediction results.

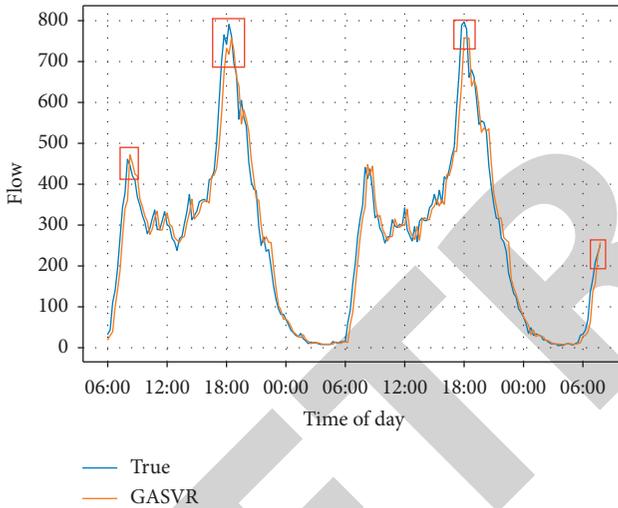


FIGURE 25: GASVR prediction results.

well. According to the above algorithm, the 15-minute downsampling data are predicted, and the prediction results are compared and analyzed according to various evaluation indexes. As shown in Table 6, it is found that the comprehensive prediction effect of the LSTM-GASVR model is the best.

Six prediction algorithms LSTM, GRU, CNN, SAE, GASVR, and LSTM-GASVR are analyzed and compared and the conclusions are summarized in Table 8.

## 7. Conclusion

Based on the charging data from May 2018 to May 2019 at a toll station around Shaanxi Province, the data normalization and reconstruction are adopted. And the working day flag bit is added to enhance the data dimension and to realize the data preprocessing of short-

TABLE 8: Analysis of short-time traffic flow prediction algorithm for different prediction algorithms.

Algorithm	Summary
LSTM	The prediction effect is good and timeliness is general, but there are big differences for different models trained
GRU	Simpler than the LSTM structure, but the traffic flow prediction at the turning point is less effective
CNN	It is suitable for predicting stable traffic flow, and the prediction effect is not good when the volatility is large
SAE	SAE model has large prediction error when traffic flow data is small
ARIMA	ARIMA model prediction error is general and timeliness is poor
GASVR	The prediction effect is good and timeliness is better, but there is a certain shift phenomenon
LSTM-GASVR	The timeliness of the model is normal, and a variety of time interval data prediction effect is better and more stable

term traffic flow. Comparing and analyzing the evaluation indexes of various time series, an effective combination evaluation index of short-term traffic flow prediction is established. Through analyzing the neural network prediction model and other machine learning prediction models and using GA to optimize the SVR model parameters, a short-term traffic flow prediction model based on LSTM-GASVR is proposed. By analyzing and comparing different time intervals in the multiple groups of experimental results, we selected 15 min as the time interval, and the time step is 16. This model is used to predict the short-term traffic flow data, and various prediction models are analyzed by combination index. The LSTM-GASVR model has normal timeliness and the best and stable prediction effect,  $R^2$  is 0.982, *explanatory variance* is 0.982, and *MAPE* is 0.118. The next step is to optimize the prediction accuracy at the peak traffic volume.

## Data Availability

The traffic flow data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Key Research and Development Program of China under Grant 2018YFB1600600, the National Natural Science Funds of China under Grant 51278058, “111 Project on Information of Vehicle-Infrastructure Sensing and ITS” under Grant B14043, Shaanxi Natural Science Basic Research Program under Grants 2019NY-163 and 2020GY-018, Joint Laboratory for Internet of Vehicles, Ministry of Education-China Mobile Communications Corporation under Grant 213024170015, the Special Fund for Basic Scientific Research of Central Colleges, Chang’an University in China under Grants 300102329101 and 300102249101.

## References

- [1] C. Zheng, *A Study on Prediction Model of Urban Road Traffic Flow in Short Time*, University of Science and Technology of China, Hebei, China, 2016, in Chinese.
- [2] B. Lu, S. Qin, and G. Ma, “Forecast of short-term traffic flow based on multi-source traffic data fusion,” *Journal of Chongqing Jiaotong University (Natural Science Edition)*, vol. 38, no. 5, pp. 13–19, 2019, in Chinese.
- [3] A. Miglani and N. Kumar, “Deep learning models for traffic flow prediction in autonomous vehicles: a review, solutions, and challenges,” *Vehicular Communications*, vol. 20, 2019.
- [4] F. Van Wageningen-Kessels, H. van Lint, K. Vuik, and S. Hoogendoorn, “Genealogy of traffic flow models,” *EURO Journal on Transportation and Logistics*, vol. 4, no. 4, pp. 445–473, 2015.
- [5] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, “Short-term traffic forecasting: where we are and where we’re going,” *Transportation Research Part C Emerging Technologies*, vol. 43, no. 1, pp. 3–19, 2014.
- [6] A. Ermagun and D. Levinson, “Spatiotemporal traffic forecasting: review and proposed directions,” *Transport Reviews*, vol. 38, no. 6, pp. 786–814, 2018.
- [7] W. Luo, B. Dong, and Z. Wang, “Short-term traffic flow prediction based on CNN-SVR hybrid deep learning model,” *Transport Systems Engineering and Information*, vol. 17, no. 5, pp. 68–74, 2017, in Chinese.
- [8] J. Tang, X. Chen, Z. Hu, F. Zong, C. Han, and L. Li, “Traffic flow prediction based on combination of support vector machine and data denoising schemes,” *Physica A: Statistical Mechanics and its Applications*, vol. 534, Article ID 120642, 2019.
- [9] L. N. N. Do, H. L. Vu, B. Q. Vo, Z. Liu, and D. Phung, “An effective spatial-temporal attention based neural network for traffic flow prediction,” *Transportation Research Part C*, vol. 108, pp. 12–28, 2019.
- [10] H. Wang, L. Liu, S. Dong, Z. Qian, and H. Wei, “A novel work zone short-term vehicle-type specific traffic speed prediction model through the hybrid EMD-ARIMA framework,” *Transportmetrica B: Transport Dynamics*, vol. 4, no. 3, pp. 159–186, 2016.
- [11] S. Vasantha Kumar and L. Vanajakshi, “Short-term traffic flow prediction using seasonalARIMA model with limited input data,” *European Transport Research Review*, vol. 7, no. 3, 2015.
- [12] Y. Rajabzadeh, A. H. Rezaie, and H. Amindavar, “Short-term traffic flow prediction using time-varying Vasicek model,” *Transportation Research Part C*, vol. 74, pp. 168–181, 2017.
- [13] A. Salamani, G. Margaritis, D. D. Kehagias, and G. Matzoulas, D. Tzovaras, Identifying patterns under both normal and abnormal traffic conditions for short-term traffic prediction,” *Transportation Research Procedia*, vol. 22, pp. 665–674, 2017.
- [14] R. Neuhold, F. Garolla, S. Oliver, and M. Fellendorf, “Predicting and optimizing traffic flow at toll plazas,” *Transportation Research Procedia*, vol. 37, pp. 330–337, 2019.
- [15] R. Liu, F. Hong, C. Lu, and W. Jiang, “Short-term traffic flow prediction based on deep circulation neural network,” *Journal of Physics: Conference Series*, vol. 1176, no. 3, Article ID 032020, 2019.
- [16] L. Qu, W. Li, W. Li, D. Ma, and Y. Wang, “Daily long-term traffic flow forecasting based on a deep neural network,” *Expert Systems With Applications*, vol. 121, pp. 304–312, 2019.
- [17] L. Zhi, *Study on the Prediction of Expressway Traffic Flow Based on BP Neural Network*, Wuhan University of Technology, Wuhan, China, 2014, in Chinese.
- [18] Li Chin-ming, *A Study on the Prediction Technology of Traffic Flow at Crossroads Facing Intelligent Transportation*, Beijing University of Posts and Telecommunications, Beijing, China, 2017, in Chinese.
- [19] J. Kang, Z. Duan, L. Tang et al., “A short-time traffic flow prediction method for smoothing,” *Measurement And Control Technology*, vol. 37, no. 2, pp. 33–37, 2018, in Chinese.
- [20] S. Wang, Y. Gu, Li Meng et al., “Short-term traffic flow prediction of expressway based on chaos theory and MEA-BPNN model,” *Shandong Science*, vol. 32, no. 2, pp. 98–107, 2019, in Chinese.
- [21] J. Xue, *Research and Implementation of Traffic Information Prediction Model Based on Deep Learning*, University of Information Engineering, Strategic Support Force, Henan, China, 2018, in Chinese.
- [22] Z. Leng, J. Gao, Y. Qin, X. Liu, and J. Yin, “Short-term forecasting model of traffic flow based on GRNN,” in *Proceedings of the 25th Chinese Control and Decision Conference (CCD)*, pp. 3816–3820, IEEE, Guiyang, China, 2013.
- [23] Y. Jia, J. Wu, and Y. Du, “Traffic speed prediction using deep learning method,” in *Proceedings of the IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1217–1222, IEEE, Rio de Janeiro, Brazil, 2016.
- [24] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, “Long short-term memory neural network for traffic speed prediction using remote microwave sensor data,” *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, 2015.
- [25] Y. Jia, J. Wu, M. Ben-Akiva, R. Seshadri, and Y. Du, “Rainfall-integrated traffic speed prediction using deep learning method,” *IET Intelligent Transport Systems*, vol. 11, no. 9, pp. 531–536, 2017.
- [26] Y. Wu and H. Tan, “Short-term traffic flow forecasting with spatial-temporal correlation in a hybrid deep learning framework,” 2016, <https://arxiv.org/abs/1612.01022>.
- [27] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu, “LSTM network: a deep learning approach for short-term traffic forecast,” *IET Intelligent Transport Systems*, vol. 11, no. 2, pp. 68–75, 2017.
- [28] W. Wangyang, H. Wu, and H. Ma, “An auto encoder and LSTM-based traffic flow prediction method,” *Sensors*, vol. 19, no. 13, p. 2946, 2019.