

Article

Short-Time Traffic Forecasting in Tourist Service Areas Based on a CNN and GRU Neural Network

Yan-Qun Yang ¹ , Jie Lin ¹ and Yu-Bin Zheng ^{2,*}¹ College of Civil Engineering, Fuzhou University, Fuzhou 350116, China² The Key Laboratory of Ministry of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai 201804, China

* Correspondence: 1711037@tongji.edu.cn

Abstract: The continuous development of highway construction projects has prompted the function of service areas to be improved day by day. A traditional service area gradually transforms from a single traffic service mode to a complex traffic service mode. The continuous enrichment and perfection of the service area's function makes the surrounding highway network more attractive, which leads to a sudden increase in highway traffic volume in a short period of time. In order to better improve the service level of a tourist service area by predicting the short-term traffic volume of the toll station around the tourist service area, this paper proposes a model combining a convolutional neural network and a gated recurrent unit (CNN plus GRU) to solve the problem of short-term traffic volume prediction. The data from 17 toll stations of the Yu'an Expressway in Guizhou Province were selected for the experiment to test the prediction effect of the CNN plus GRU-based model. The experimental results show that the prediction accuracy, the MAE and RMSE, are 1.8101 and 2.7021, respectively, for the toll stations with lower traffic volumes, and 3.820 and 5.172, respectively, for the toll stations with higher traffic volumes. Compared with a single model, the model's prediction accuracy is improved, to different degrees. Therefore, the use of a convolutional neural network operation is better when the total traffic volume is low, considering the algorithm's time and error. When using the combined convolutional neural network and gated recurrent unit model and when the total traffic volume is high, the algorithm error is significantly reduced and the prediction results are better.

Keywords: traffic engineering; short-term traffic volume forecast; CNN plus GRU neural network; tourism service area; spatiotemporal characteristics; spatial symmetry



Citation: Yang, Y.-Q.; Lin, J.; Zheng, Y.-B. Short-Time Traffic Forecasting in Tourist Service Areas Based on a CNN and GRU Neural Network. *Appl. Sci.* **2022**, *12*, 9114. <https://doi.org/10.3390/app12189114>

Academic Editors: Xue Yang and Luliang Tang

Received: 24 August 2022

Accepted: 7 September 2022

Published: 10 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, with the rapid development of expressway construction projects, the functions of expressways have been continuously upgraded and expanded. As an important carrier for the integrated development of expressway transportation, a service area plays an important role in the expressway's service quality window. In the future, with the gradual popularization of toll-free flows without stations, the expressway service areas will seamlessly connect with the surrounding scenic spots. Further, some expressway service areas will thus become important gateways to scenic spots. Therefore, the function of the traditional service area needs to change from a single traffic service to a new situation comprised of a compound traffic service. It must also take into account the development, according to local conditions, and absorb local characteristics, as well as consider the development of new business models (introduction attractions, feeders, leisure, and entertainment). Finally, a new tourism service area based on a service system of "quick entry and good tour" is formed.

This type of service area will also become the new direction in the development of future service areas.

Since “the Several Opinions on Promoting the Integrated Development of Transport and Tourism” were jointly promulgated by all government departments in February 2017, all regions have actively responded. For example, Qinghai Province has created a new situation of “Transportation + Tourism” integration by taking advantage of its tourism resources [1]. The research on tourism service areas includes all aspects, such as the study by Zhang Lu [2] which obtained the characteristic portrait, overall layout, function setting, spatial form, human needs, regional culture, and place order of the tourism service areas by selecting typical cases. Zhou Chang [3] put forward planning and design methods suitable for mountainous expressway tourism service areas. Yun Bai et al. proposed a modeling framework combining tourism planning and transport network design for developing synergies between transport and tourism, which can be used for the site selection of tourism service areas [4]. Wang Dan and others put forward the idea of a “place spirit” in the environmental planning and design of tourist service areas [5,6]. Yang Z. expounded the classification of tourist highway service areas and key points of layout, summarizing the functions and facilities of different service areas and putting forward the requirements of landscape planning and designing service areas [7]. At present, the planning and construction of unconstructed tourism service areas are studied, and existing service areas are transformed into tourism service areas for environmental construction. Nowadays, for a tourist service area is in the stages of construction or trial operation, the effect of traffic growth has not been obvious. With the promotion, continuous development, and function improvement of tourism service areas, the attraction to the surrounding expressway network is increasing, and the traffic flow of the expressway is increasing rapidly. Most of the self-driving traffic has brought daily maintenance and operation pressures to the expressway. However, if the capacity of the highway cannot meet the growing traffic demands, there will be a lower sense of a tourism experience and a longer travel time. This makes a tourism service area lose its meaning of leisure. Therefore, the timely and accurate prediction of the entrance flow of each toll station and the accurate grasp of the high-speed traffic situation in a tourist service area are conducive to providing a well-functioning expressway management plan.

There are various prediction models for traffic volume. A model can be divided into parametric models and non-parametric models, according to the model’s structure. In addition, according to different training methods and methods of deep learning, a model can be further divided into generative deep structures, discriminative deep structures, and hybrid deep structures [8]. Very fruitful research results have been achieved in different development periods. The current research direction is mainly derived from a parametric statistical model to a non-parametric model, gradually, and then the combination model appears. Some scholars using the parametric model have used the growth curve to predict new rail transit passenger flow [9]. Among the parametric models, a time-series is the most common, including various time-series models and their variants [10–12]. The advantages of using parametric models for short-term traffic forecasting are mainly focused on the simplicity of the models and their ease of interpretation. However, this type of model does not reflect the non-linear characteristics of traffic volume. Because, in reality, the characteristics of traffic volume are mostly non-linear, error is larger. Given the limitations of parametric models, non-parametric models provide good solutions. For example, Liu L. used a support vector regression algorithm (SVR) to predict urban traffic flow. SVR has good results in predicting small-sample, multidimensional data, with non-linear processing and local minimums. Compared with BP, LSTM, and ARIMA, SVR is better than other methods. It has excellent generalization ability, which avoids some problems caused by over-fitting, and it can accurately predict urban short-term traffic flow [13]. Li Z. used three different variants of an RNN to construct a model to experiment on traffic data and found that a recurrent neural network has a good prediction effect on specific road traffic flow, while LSTM has a less-good prediction error [14]. Toan T.D. combined the KNN method with SVM and found that the KNN method helped to significantly reduce the training scale of an SVM, thus speeding up the training process without affecting prediction performance [15].

Weibin Zhang used a combination of an algorithm (STFSA) and a CNN deep learning framework for a short-term traffic flow prediction two-dimensional matrix. The prediction accuracy was better than an SVR, SARIMA, KNN, ANN, or CNN single model and a combined model (STFSA plus ANN) [16]. Dai G scholars adopted the same idea to combine an STFSA and a GRU for a short-term traffic flow prediction model. Compared with the CNN and GRU models, it was found that the combined model was significantly better than the single model in terms of accuracy and stability [17]. Balachandran Vijayalakshmi constructed a CNN-LSTM multi-step prediction model by combining feature data with an attention mechanism. The results showed that the prediction model provided nearly 99% accuracy, and it applied to peak and non-peak hours, as well as working days and non-working days [18]. Zhang W. proposed three hybrid depth-time models (CL-CN-G, CL-CNG, and G-CN-CL models) based on CNN, GRU, and ConvLSTM models to address the lack of existing studies regarding the prediction of characteristic traffic flows such as holidays and severe weather. The case studies showed that the CL-CN-G, CL-CNG, and G-CN-CL models were accurate and effective in all cases, and the G-CN-CL model performed best [19].

At present, most of the data on highway traffic volume prediction mainly comes from relevant checkpoints [20]. Some scholars also have used the import and export data of toll stations [21]. Among the prediction models after extraction of the spatio-temporal features of traffic volume data, CNN models are commonly used to extract the spatial characteristics of traffic flow because of their high interpretability [16]. GRU models are commonly used to extract the temporal characteristics of traffic volume due to their ability to effectively solve problems such as gradient disappearance or explosion during temporal extraction [22]. Therefore, the two can be combined to build CNN and GRU neural networks to extract the spatial and temporal features of traffic flow.

With the increasing popularity of self-driving tourism, highways have become an indispensable part of tourism development. It is very important to forecast and analyze the traffic volume of a highway in time and study its tourism utilization potential. Although a number of studies have been carried out on highway traffic forecasting at home and abroad, there is a lack of research on the service traffic forecasting of highway tourist areas. In summary, this paper uses CNN convolution to extract spatial information and a GRU to extract long-term sequence information to predict the traffic volume at the entrance of a toll station in a tourist service area. The results can provide a scientific basis for optimizing the layout of expressway transportation networks and reshaping service areas for tourism.

2. Introduction to the Network Model

2.1. Recurrent Neural Network Based on a GRU Gating Mechanism

A recurrent neural network (RNN) is a neural network that contains a recurrent structure in a hidden layer. Through this connection, an RNN can store current input information, as well as previous time step information. Due to the sharing of RNN weights, if a sequence is long enough to multiply each output value of the previous hidden layer by the same weight during propagation, it will be difficult to pass information from an earlier time step to a later time step. It is easy to cause gradient disappearance and gradient explosion during propagation. In order to overcome this defect, long short-term memory (LSTM) was proposed by Hochreiter S and Schmidhuber J in 1997 and has been widely used. The key of LSTM is the cell state and 'gate' structure. The LSTM hidden layer cell structure is composed of an input gate, a forgetting gate, and an output gate.

A gated recurrent unit neural network, as proposed by Cho, is a variant of an LSTM. A GRU is a type of gating mechanism in an RNN. The purpose of a GRU is the same as an LSTM, and both to overcome the problem of gradient disappearance or the explosion of the RNN during propagation, retaining as much information as possible for long sequences. A GRU is as excellent as an LSTM in many sequence tasks, but its parameters are less than those of an LSTM and include only one "reset gate" and one "update gate". Therefore, a GRU model saves computation time and shows equally good results for some tasks. The

“update gate” of a GRU replaces the forgetting gate and input gate of an LSTM. It merges the cell state and hidden state, which makes a GRU simpler than an LSTM model by these transformations. Figure 1 shows the internal structure of a GRU’s basic neuron.

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \tag{1}$$

$$\tilde{h}_t = \tanh(W_h[r_t * h_{t-1}, x_t]) \tag{2}$$

$$z_t = \sigma(W_z[h_{t-1}, x_t]) \tag{3}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \tag{4}$$

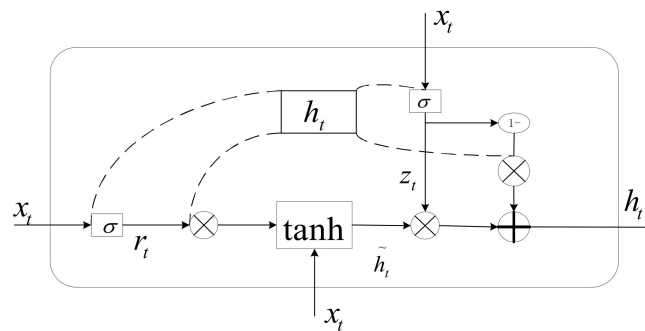


Figure 1. GRU structure diagram.

Here, x_t is the input at the current moment, h_{t-1} is the hidden state at the previous moment, and h_t is the hidden state calculated at the current moment. When calculating the hidden state at the current moment, the model will first calculate a candidate state (\tilde{h}_t). The value of the reset gate is taken into account when calculating the candidate states, and the formula for the reset gate is shown in (1). If the reset gate is close to 0, the current candidate ignores the previous hidden state and uses the current input to calculate it. This can effectively cause the hidden state to discard any irrelevant information found in the future. The candidate state calculation formula is shown in (2), where * represents the term-by-term multiplication. After the candidate values are calculated, the update gate controls how much information from the previous hidden state can be passed to the current hidden state. This is similar to the memory unit of an LSTM, which allows a GRU to remember long-term information. The update gate calculation formula is shown in (3). Finally, it is possible to calculate the hidden state at the current moment.

2.2. Convolutional Neural Network

A convolutional neural network (CNN) has a unique advantage in processing data with a grid-like structure. For example, a convolutional network has an excellent performance in image classification, face recognition, and audio retrieval, all of which are mainly used in computer vision [23]. A CNN mainly consist of a convolutional layer, a pooling layer, and a fully connected layer. The convolution layer is composed of several convolution kernels, and the local features of the samples are extracted layer by layer through convolution calculation. The pooling layer achieves down-sampling primarily by reducing the pooling function of the data dimension. This not only improves the computational efficiency, but also maintains the important information in the sample data. The fully connected layer is a fully connected operation on the feature matrix extracted after the continuous convolution and pooling layers to expand the complexity of the network, and, finally, the output of the network is obtained. A CNN reduces a model’s number of parameters and increases the training speed because of its local connectivity and their shared weights. This feature is more suitable for processing two-dimensional data.

The attraction of tourist-oriented service areas to passenger flow is mainly reflected in the entrance of each toll station on the highway. The data of highway traffic volume are

periodic and time-varying. For a feature matrix constructed by traffic volume, it can be input to a 1D or 2D convolutional neural network to extract features. For example, in a 1D CNN, a feature matrix of 512×8 is extracted by 32 convolution kernels of size 2 and step 1. The output matrix is 511×32 , as shown in Figure 2. The features can be extracted by setting different sizes and steps, according to different requirements.

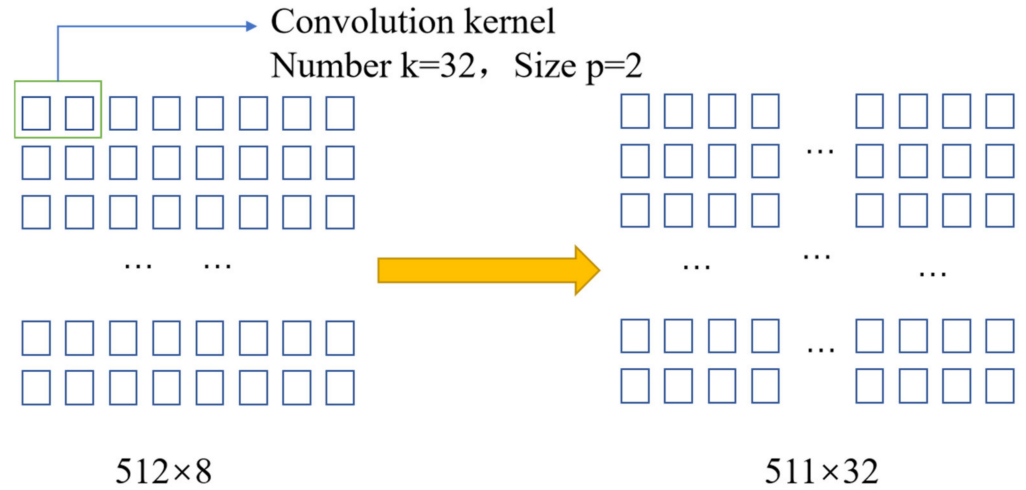


Figure 2. 1D convolution diagram.

3. Model Construction

3.1. Structure of the Model

The CNN and GRU models were combined to extract the temporal and spatial characteristics of the entrance flow of toll stations in an expressway service area by using the spatial perception performance of a CNN and the time-memory performance of a GRU. The short-term prediction of the entrance flow of a toll station is realized. The process of the CNN plus GRU highway service area traffic prediction is as follows:

Firstly, the time-series data of the toll station entrance traffic distribution before and after the tourist service area space is divided and normalized. Next, the data is imported into a CNN network to extract the spatial features of the traffic flow. Then, the data are input to a GRU layer to extract the temporal features of the data matrix. Finally, the inverse normalization is processed to obtain the predicted values. The structure of the model is shown in Figure 3.

The data imported into the neural network is the normalized spatio-temporal data matrix. The convolutional layer uses several filters to convolve the spatial location information in the input data to extract features. The filter-extracted data matrix is then fed into the GRU layer. Finally, the traffic flow at the entrance of the toll station at the next moment is obtained through the processing of the fully connected layer.

3.2. Input Data Based on Spatio-Temporal Features

Convolutional neural networks (CNNs) are characterized by local perception and weight sharing. Therefore, this paper uses a CNN to extract the location-related spatial features of passenger flow data and construct a spatio-temporal matrix containing location and time information. The rows of the matrix represent the number of entrances into a toll station over the past n moments. The columns of the matrix represent the traffic at the entrance of different toll stations at a certain moment. The temporal matrix is:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_m \end{bmatrix} = \begin{bmatrix} x_1(t-n) & x_1(t-n-1) & \dots & x_1(t-1) \\ x_2(t-n) & x_2(t-n-1) & \dots & x_2(t-1) \\ x_3(t-n) & x_3(t-n-1) & \dots & x_3(t-1) \\ \dots & \dots & \dots & \dots \\ x_m(t-n) & x_m(t-n-1) & \dots & x_m(t-1) \end{bmatrix} \quad (5)$$

where x_i in row i is all the flows at station i for the past n moments and $x_i(t - n)$ is the passenger flow at station i until n moments.

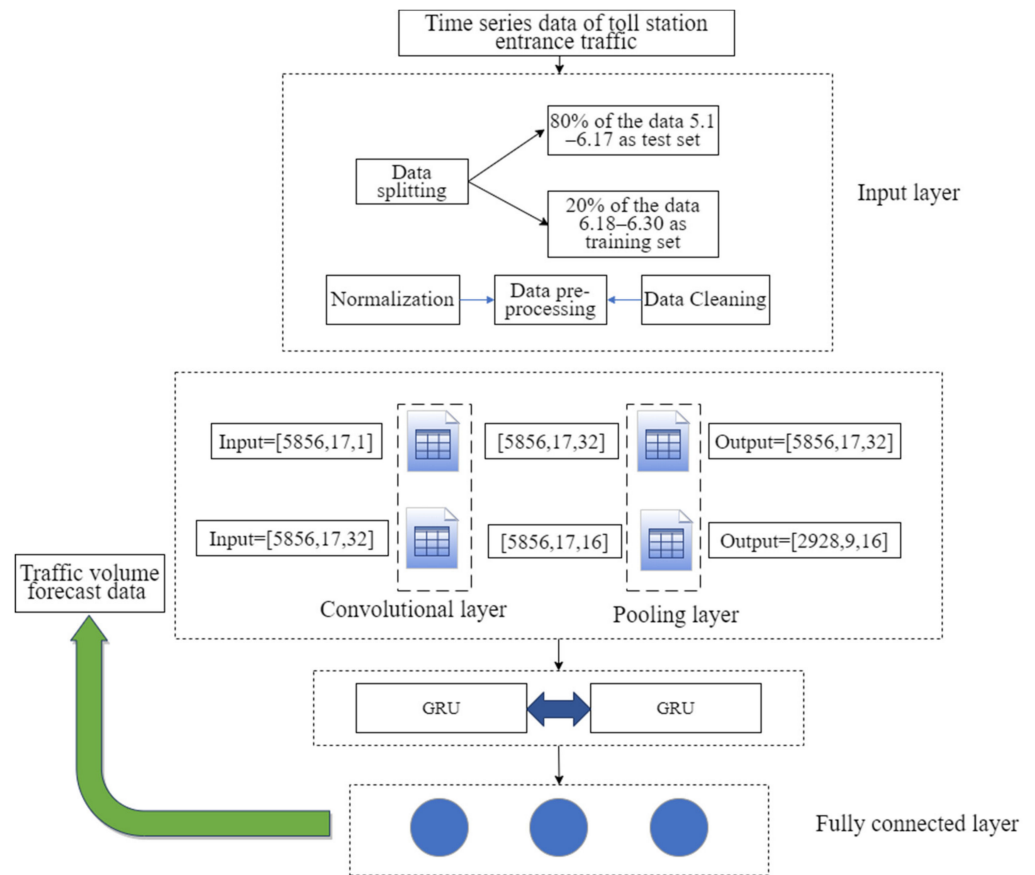


Figure 3. CNN-GRU model framework diagram.

3.3. Data Pre-Processing

Traffic flow data is the basis for traffic prediction research. High-quality data can effectively ensure better training results and prediction accuracy. Some characteristics differ greatly in magnitude, which will expand the effects of large-scale features on the model such that the model depends on these features, worsening the model’s performance. At the same time, the difference in the orders of magnitude will slow down the model training, and so the data is normalized before being fed into the network model.

$$x_i^* = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{6}$$

where x_i is the data in the same row vector at a moment in the input data, x_{\min} is the minimum value in the row vector, and x_{\max} is the maximum value in the row vector.

3.4. Model Parameters and Comparison Scheme

In this paper, we refer to past representative deep learning methods and training parameters studies, and B stands for Batch Size, which indicates the number of samples input into a neural network at one time for training. A model’s size affects the degree of optimization and its speed, which also directly affect the memory usage. The number of epochs is represented by E, indicating that all data inputs in a network perform forward propagation and back propagation calculation processes. Drawing on previous studies [24], an epoch takes 50 s, with excellent results when using a CNN with an LSTM or a GRU hybrid model. Since different settings of the number of training samples can cause overfitting of the data, B is set to 64. It is possible to guarantee the training speed while obtaining

a relatively low training and testing error. Optimizer uses the Adam gradient descent algorithm. It uses gradient moment estimation to dynamically adjust the learning rate of the parameters. Using bias correction, the iterative learning rate is within a certain range and the parameters are relatively stable. This is conducive to fully extracting traffic flow characteristics [25]. To test the validity of the proposed model, it is preferable to use a learning rate equal to 0.01.

According to the characteristics of traffic volume extraction by a convolutional neural network, this paper intends to set up two convolutional layers. The convolution kernel size is 1×1 , the step size is 1, the number of convolution kernels is 32 and 16, respectively, and the number of pooling layers is 2. The maximum pooling is used. The activation function is the rule function. For the gated circulation unit layer, a GRU is applied for extracting the traffic volume time characteristics. The GRU sets up 2 layers with 20 and 10 neurons in each layer, the time step is taken as 8, and 8 historical data are used to predict the next moment's traffic volume. The specific data flow is shown in Figure 4.

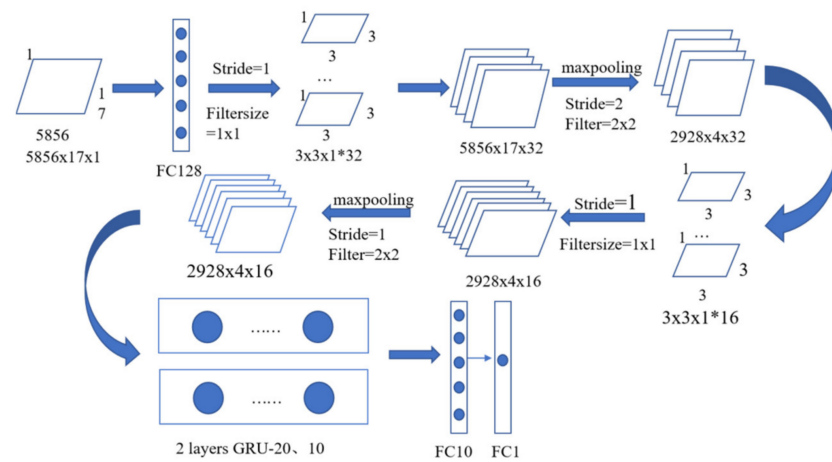


Figure 4. Flowchart of the data.

3.5. Evaluation Indicators

In order to evaluate the performance of the model, this paper uses two error evaluation indexes—mean absolute error (MAE) and root mean square error (RMSE)—to evaluate the prediction accuracy of the highway tourism service area traffic forecasting model. The error calculation formula is as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n \left| \hat{q}_i - q_i \right| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (q_i - \hat{q}_i)^2} \tag{8}$$

where \hat{q}_i is the predicted value, q_i is the true value, and n is the sample size.

4. Example Demonstration

4.1. Data Source and Description

The toll booth entrance flow data obtained from the toll booth are divided by 22 aspects. Different aspects of the data can be selected for processing, depending on the content of the study. This paper mainly uses the entrance station number and entrance date and time for data processing. Data pre-processing was performed using a total of 61 days of data from May–June 2017 for highways in Guizhou Province. First of all, the data from the toll station will be imported into Excel for filtering and statistics. The final structure of the data is divided into 96 data sets for each 15 min time span. The 17 toll stations (TS) of the S62 Yu’an Expressway were selected. The data were finally counted into a 5856×17 matrix.

4.2. Spatial Correlation Analysis of Toll Stations

When studying the traffic volume of toll stations, there is often a certain correlation between the traffic volume of the front and the rear of the toll stations. The correlation of the traffic volume at each toll station is analyzed by Pearson correlation coefficient. The Pearson correlation coefficient formula is shown in (9):

$$\rho_{x,y} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}} \quad (9)$$

Pearson correlation analysis is carried out on the data of toll stations of the Yu 'an Expressway. Table 1 shows the Pearson correlation coefficients between the toll stations. According to the calculation, the average value of Pearson correlation coefficient among the toll stations of the Yu 'an Expressway is 0.91. The overall correlation is strong. Except for the weak correlation between toll station 5 and the other toll station gates, the other toll station gates have a strong correlation with each other. The Pearson correlation coefficient of adjacent toll gates is higher.

Table 1. Pearson correlation coefficient.

	TS 1	TS 2	TS 3	TS 4	TS 5	TS 6	TS 7	TS 8	TS 9	TS 10	TS 11	TS 12	TS 13	TS 14	TS 15	TS 16	TS 17
TS 1	1	0.989	0.983	0.984	0.371	0.958	0.97	0.957	0.925	0.95	0.93	0.987	0.968	0.97	0.978	0.942	0.937
TS 2	0.989	1	0.962	0.972	0.448	0.964	0.976	0.975	0.947	0.961	0.923	0.99	0.948	0.956	0.956	0.932	0.936
TS 3	0.983	0.962	1	0.94	0.286	0.902	0.914	0.893	0.863	0.895	0.895	0.952	0.922	0.927	0.941	0.877	0.874
TS 4	0.984	0.972	0.94	1	0.425	0.976	0.986	0.973	0.949	0.969	0.943	0.984	0.989	0.981	0.988	0.98	0.975
TS 5	0.371	0.448	0.286	0.425	1	0.535	0.489	0.482	0.679	0.597	0.623	0.504	0.403	0.43	0.371	0.513	0.589
TS 6	0.958	0.964	0.902	0.976	0.535	1	0.986	0.973	0.977	0.996	0.953	0.982	0.976	0.987	0.973	0.981	0.978
TS 7	0.97	0.976	0.914	0.986	0.489	0.986	1	0.993	0.968	0.980	0.938	0.990	0.984	0.979	0.970	0.978	0.979
TS 8	0.957	0.975	0.893	0.973	0.482	0.973	0.993	1	0.958	0.966	0.903	0.976	0.961	0.956	0.949	0.957	0.960
TS 9	0.925	0.947	0.863	0.949	0.679	0.977	0.968	0.958	1	0.991	0.968	0.972	0.936	0.944	0.924	0.961	0.982
TS 10	0.95	0.961	0.895	0.969	0.597	0.996	0.980	0.966	0.991	1	0.965	0.982	0.963	0.973	0.956	0.974	0.983
TS 11	0.93	0.923	0.895	0.943	0.623	0.953	0.938	0.903	0.968	0.965	1	0.962	0.939	0.950	0.939	0.961	0.972
TS 12	0.987	0.99	0.952	0.984	0.504	0.982	0.990	0.976	0.972	0.982	0.962	1	0.973	0.977	0.971	0.966	0.972
TS 13	0.968	0.948	0.922	0.989	0.403	0.976	0.984	0.961	0.936	0.963	0.939	0.973	1	0.991	0.988	0.985	0.974
TS 14	0.97	0.956	0.927	0.981	0.43	0.987	0.979	0.956	0.944	0.973	0.95	0.977	0.991	1	0.993	0.984	0.967
TS 15	0.978	0.956	0.941	0.988	0.371	0.973	0.970	0.949	0.924	0.956	0.939	0.971	0.988	0.993	1	0.978	0.956
TS 16	0.942	0.932	0.877	0.98	0.513	0.981	0.978	0.957	0.961	0.974	0.961	0.966	0.985	0.984	0.978	1	0.99
TS 17	0.937	0.936	0.874	0.975	0.589	0.978	0.979	0.960	0.982	0.983	0.972	0.972	0.974	0.967	0.956	0.99	1

4.3. Experimental Environment and Parameter Configuration

The paper builds a neural network model using the Keras module, an advanced API for deep learning within TensorFlow, developed by Google. The development environment is a pycharm for traffic prediction model building and training. The experimental environment is an HP All-in-One with a 2.5GHz, Inter(R) core (TM) i5-7400T with 8 GB RAM and an NVIDIA GTX 930MX graphics card as the primary configuration platform.

The models selected in this paper are representative deep learning methods and training parameters in traffic volume prediction. The numbers of each hidden layer and neuron are shown in Table 2, where B, E, Optimizer, and Adam are selected according to those described in Section 3.4, respectively.

Table 2. Example of toll station flow data.

Model	Name of the Parameter	Parameter Value
GRU	Hidden Layer B/E/Optimizer	GRU(20) + GRU(10) 64/50/Adam
LSTM	Hidden Layer B/E/Optimizer	LSTM(20) + LSTM(10) 64/50/Adam
1DCNN	Hidden Layer B/E/Optimizer	Dense(128) + Conv1d(32) + Maxpooling1d + Conv1d(16) + Maxpooling + Dense(10) 64/50/Adam
1DCNN + GRU	Hidden Layer B/E/Optimizer	Dense(128) + Conv1d(32) + Maxpooling1d + Conv1d(16) + Maxpooling + GRU(20) + GRU(10) + Dense(10) 64/50/Adam

4.4. Model Evaluation Results

The samples are fed into the LSTM, GRU, 1DCNN, and 1DCNN plus GRU models for training. The loss functions of the training and validation sets are obtained by training, as shown in Figure 5, where (a) shows that the loss function of the training set of each model gradually decreases and eventually stabilizes with the increase in the number of iterations. The loss function of the validation set in (b) gradually decreases with the increasing number of iterations. Although all fluctuate, the 1DCNN model fluctuates the most. Thus, it can be concluded that the 1DCNN model applied to the data of this thesis has a large error.

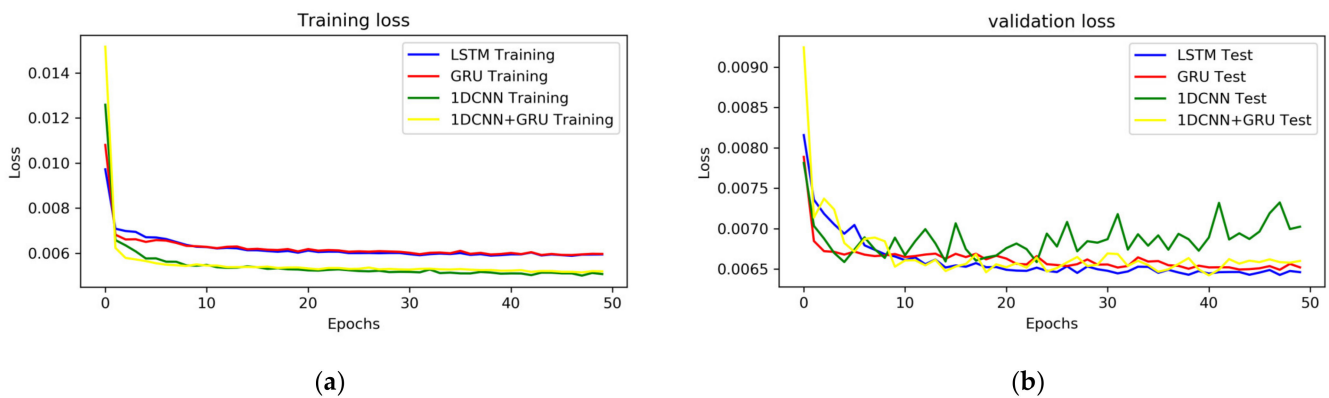


Figure 5. Loss function diagram. (a) Training set loss function. (b) Validation set loss function.

In this paper, a 1DCNN plus GRU model considering temporal and spatial characteristics, a GRU and LSTM model considering temporal characteristics, and a 1DCNN model considering spatial characteristics were set up for experimental comparison. The results can be used to verify the highway toll station space–time characteristics of the impacts of traffic. A comparison of Tables 3 and 4 reveals that the 1DCNN plus GRU model with spatio-temporal characteristics considered was 3.94% lower in the MAE, 1.12% lower in the RMSE, and 16.05 s longer in computing time than the LSTM model without spatio-temporal characteristics for predicting less traffic at toll station 17. The MAE was reduced by 4.10%, the RMSE was reduced by 1.57%, and the computing time was increased by 15.75 s compared to the GRU model for predicting traffic at toll station 17.

Table 3. Toll station 17 model comparison results.

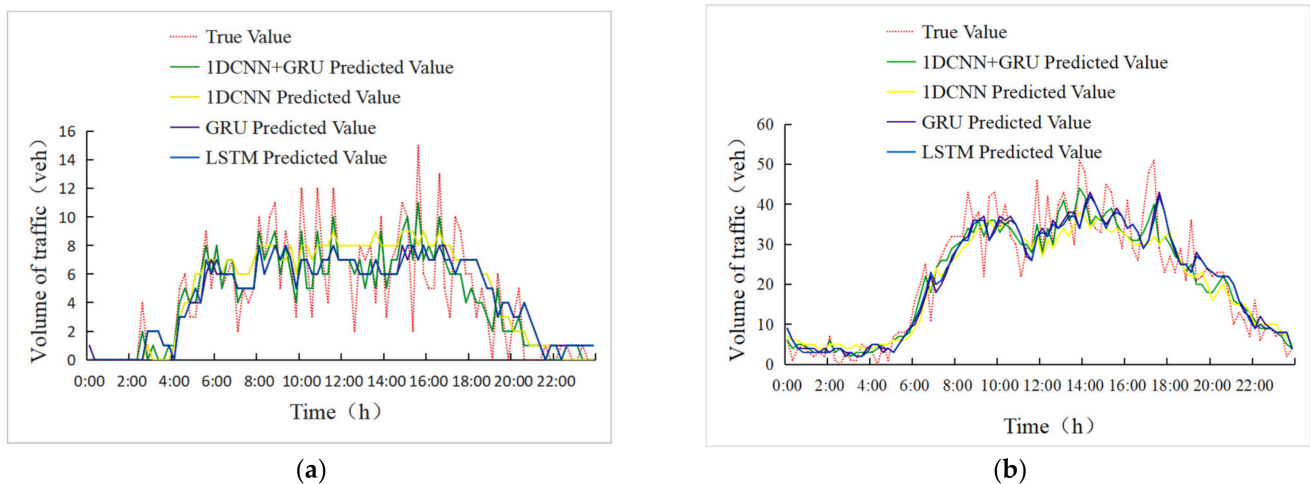
Model	MAE	RMSE	Algorithm Time/s
LSTM	1.8843	2.7327	43.53
GRU	1.8875	2.7452	43.93
1DCNN	1.9193	2.7754	22.29
1DCNN + GRU	1.8101	2.7021	59.68

Table 4. Toll station 3 model comparison results.

Model	MAE	RMSE	Algorithm Time/s
LSTM	4.445	5.930	43.21
GRU	4.484	5.985	43.91
1DCNN	4.426	5.677	22.61
1DCNN + GRU	3.820	5.172	58.55

A comparison of Tables 3 and 4 reveals that the 1DCNN plus GRU model with spatio-temporal characteristics considered had a 14.06% reduction in the MAE, 12.78% reduction in the RMSE, and 15.34 s increase in computing time compared to the LSTM model that did not consider spatio-temporal characteristics in predicting higher traffic volumes at toll booth 3. In comparison with the GRU model, the MAE decreased by 14.81%, the RMSE decreased by 13.58%, and the computing time increased by 14.64 s when predicting toll station 17.

The prediction accuracy is relatively low due to the small total volume of traffic at toll station 17 and the large fluctuation of real traffic data, as shown in Figure 6a. By comparing the LSTM and GRU models that only considered the temporal characteristics, we can find that the 1DCNN plus GRU can explore the spatial and temporal characteristics of toll station traffic more deeply and reduce the prediction error. The total volume of traffic at toll station 3 is relatively large. The model prediction effect is obviously improved.

**Figure 6.** Predictive model comparison results. (a) Toll station 17 forecast results. (b) Toll station 3 forecast results.

5. Conclusions

In response to the development of the concept of “transportation + tourism” vigorously promoted by the state in recent years, traffic volume is a crucial factor for tourist service areas. Traffic forecasting is a direct quantitative indicator that directly affects the benefits provided by tourism service areas. Therefore, this paper proposes to apply a theoretical model based on the combination of a convolutional neural network and a gated circulation unit to predict the traffic flow of toll station entrances before and after the spatial location of a tourist service area. A one-dimensional convolutional neural network can extract the spatial dimensional features of traffic volume at the entrance of a toll station. The gated cycle unit learns the traffic volume data in the time axis. By combining the two, the spatial and temporal characteristics of traffic flow at the entrance of a toll station can be fully explored to finally obtain the traffic volume prediction results for a targeted time. As the complexity of the model increases, the prediction error decreases. When the computing time increases, the model has better prediction results. Considering the time and error of the algorithm, the convolutional neural network model works well when the total volume

of traffic is low. When the total volume of traffic is high, the algorithm error is considered to be significantly lower, and the use of the combined convolutional neural network and gated recurrent cell model has good prediction results.

As there are few cases of “traffic + tourism”-type service areas, it is difficult to offer full credit to the role of neural networks in predicting the traffic volume of tourist-type service areas on highways. The May and June 2017 data from 17 of the toll stations of the S62 Yu’an Expressway in Guizhou Province are proposed as a hypothetical case study of the tourist-oriented service area of the Sky Bridge. The subsequent study will obtain information on the traffic volume of the toll stations before and after the spatial location of tourist service areas, and further demonstrate the impacts of setting up tourist-oriented service areas on the additional traffic volume. At the same time, the subsequent study will pay attention to human factors to make the model more applicable.

Author Contributions: Conceptualization, Y.-Q.Y.; Investigation, Y.-B.Z.; Methodology, J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from Guizhou Expressway Group and are available from the authors with the permission of Guizhou Expressway Group.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Han, Y.; Ma, L.; Zhang, Z.; Han, R. Discussion about building tourism—based service areas in Qinghai province. *Qinghai Transp. Sci. Technol.* **2018**, *4*, 5–7.
2. Zhang, L. *A Study on the Design of Tourist-oriented Service Areas on Highways*; China Architecture Design & Research Group: Beijing, China, 2017.
3. Chang, Z.; Wei, H. Analysis on the Planning and Design of Tourism Service Area of Mountainous Expressway. *Technol. Highw. Transp.* **2021**, *37*, 133–138.
4. Bai, Y.; Kou, X.; An, S.; Ouyang, Y. Integrated Planning of Tourism Investment and Transportation Network Design. *Transp. Res. Rec. J. Transp. Res. Board* **2014**, *2467*, 91–100. [[CrossRef](#)]
5. Wang, D.; Huang, S.; Geming; Jiang, H. Research on construction technology of visual environment for highway tourism service area. *Technol. Highw. Transp.* **2018**, *14*, 322–324.
6. Wang, D. Study on the Construction Technology of Visual Environment in Highway Tourism Service Area. In Proceedings of the IOP Conference Series: Earth and Environmental Science, Raipur, India, 25–26 February 2022; IOP Publishing: Bristol, UK, 2020; Volume 453, p. 012022.
7. Yang, Z.; Jin, F.; Shou, J.; Yang, Y. Landscape planning and design of tourist highway service area under the background of all-for-one tourism system. In Proceedings of the IOP Conference Series: Earth and Environmental Science, Raipur, India, 25–26 February 2022; IOP Publishing: Bristol, UK, 2019; Volume 304, p. 032092.
8. Liang, D.; Yang, M.; Chao, Q.; Yun, M.; Jin-Ming, L. Survey on Short-term Traffic Flow Forecasting Based on Deep Learning. *Comput. Sci.* **2019**, *046*, 39–47.
9. Zhang, J.; Lv, H.; Liu, X.; Wang, F.; Xie, L. Prediction of Induced Passenger Flow in Urban Rail Transit Under TOD Mode. *J. Transp. Eng. Inf.* **2017**, *15*, 76–82.
10. Alves, M.; Cordeiro, R. Effective and unburdensome forecast of highway traffic flow with adaptive computing. *Knowl. Based Syst.* **2021**, *212*, 106603. [[CrossRef](#)]
11. Lu, S.; Zhang, Q.; Chen, G.; Seng, D. A combined method for short-term traffic flow prediction based on recurrent neural network. *Alex. Eng. J.* **2021**, *60*, 87–94. [[CrossRef](#)]
12. Dissanayake, B.; Hemachandra, O.; Lakshitha, N.; Haputhanthri, D.; Wijayasiri, A. A Comparison of ARIMAX, VAR and LSTM on Multivariate Short-Term Traffic Volume Forecasting. In Proceedings of the Conference of Open Innovations Association, Moscow, Russia, 25–29 January 2021; FRUCT: Helsinki, Finland, 2021; pp. 564–570.
13. Liu, L. A Short-Term Traffic Flow Prediction Method Based on SVR. In Proceedings of the 2021 2nd International Conference on Urban Engineering and Management Science (ICUEMS), Sanya, China, 29–31 January 2021; pp. 1–4.

14. Li, Z.; Li, C.; Cui, X.; Zhang, Z. Short-term Traffic Flow Prediction Based on Recurrent Neural Network. In Proceedings of the 2021 International Conference on Computer Communication and Artificial Intelligence (CCAI), Guangzhou, China, 7–9 May 2021; pp. 81–85.
15. Toan, T.D.; Truong, V.-H. Support vector machine for short-term traffic flow prediction and improvement of its model training using nearest neighbor approach. *Transp. Res. Rec.* **2021**, *2675*, 362–373. [[CrossRef](#)]
16. Zhang, W.; Yu, Y.; Qi, Y.; Shu, F.; Wang, Y. Short-term traffic flow prediction based on spatio-temporal analysis and CNN deep learning. *Transp. A Transp. Sci.* **2019**, *15*, 1688–1711. [[CrossRef](#)]
17. Dai, G.; Ma, C.; Xu, X. Short-term traffic flow prediction method for urban road sections based on space–time analysis and GRU. *IEEE Access* **2019**, *7*, 143025–143035. [[CrossRef](#)]
18. Vijayalakshmi, B.; Ramar, K.; Jhanjhi, N.; Verma, S.; Kaliappan, M.; Vijayalakshmi, K.; Vimal, S.; Kavita; Ghosh, U. An attention-based deep learning model for traffic flow prediction using spatiotemporal features towards sustainable smart city. *Int. J. Commun. Syst.* **2021**, *34*, e4609. [[CrossRef](#)]
19. Zhang, W.; Yao, R.; Du, X.; Ye, J. Hybrid Deep Spatio-Temporal Models for Traffic Flow Prediction on Holidays and Under Adverse Weather. *IEEE Access* **2021**, *9*, 157165–157181. [[CrossRef](#)]
20. Wen, H.; Zhang, D.; Siyuan, L.U. Application of GA-LSTM model in highway traffic flow prediction. *J. Harbin Inst. Technol.* **2019**, *51*, 81–87.
21. Shi, T.; Yuan, W.; Wang, P.; Zhao, X. Regional Traffic Flow Prediction on multiple Spatial Distributed Toll Gate in a City Cycle. In Proceedings of the 2021 IEEE 16th Conference on Industrial Electronics and Applications (ICIEA), Chengdu, China, 1–4 August 2021; pp. 94–99.
22. Weerakody, P.B.; Wong, K.W.; Wang, G.; Ela, W. A Review of Irregular Time Series Data Handling with Gated Recurrent Neural Networks. *Neurocomputing* **2021**, *441*, 161–178. [[CrossRef](#)]
23. Rawat, W.; Wang, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput.* **2017**, *29*, 2352–2449. [[CrossRef](#)] [[PubMed](#)]
24. Aloraifan, D.; Ahmad, I.; Alrashed, E. Deep learning based network traffic matrix prediction. *Int. J. Intell. Netw.* **2021**, *2*, 46–56. [[CrossRef](#)]
25. Huang, H.B.; Huang, X.R.; Li, R.X.; Lim, T.C.; Ding, W.P. Sound quality prediction of vehicle interior noise using deep belief networks. *Appl. Acoust.* **2016**, *113*, 149–161. [[CrossRef](#)]