Physica A 570 (2021) 125838

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Short-term demand forecasting for online car-hailing using ConvLSTM networks

Xijin Lu, Changxi Ma [*], Yihuan Qiao

School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China

ARTICLE INFO

Article history:
Received 8 December 2020
Received in revised form 1 February 2021
Available online 10 February 2021

Keywords:
Online car-hailing demand
Short-term demand forecasting
Deep learning
Conv-LSTM
Spatiotemporal feature

ABSTRACT

This paper used the previous data of online car-hailing orders in Haikou provided by Didi Chuxing GAIA Initiative to predict the short-term demand for online car-hailing service. This paper contains two main steps. The first step is about converting online car-hailing demand for images that contains spatiotemporal feature of online car-hailing orders. This paper draws a picture every 72 min from 2017/5/8 to 2017/8/8, with a total of 1,840 binary vector images. The second step is to employ the deep learning method of a Conv-LSTM neutral network to the image for online car-hailing demand prediction. Conv-LSTM has excellent image prediction properties, so it is ideal for predicting such binary vector figures with spatiotemporal information. After learning the first 1460 images, the last 380 images were simulated, predicted and tested. Finally, the simulation results of the last 20 images were taken as the effect of the model. Our result shows that the Conv-LSTM neutral network can train the model with a reasonable output and is suitable for short-term forecasting network ride-hailing demand forecast with spatiotemporal feature information. By comparing five different training session times, it can be seen that when the number of training session reaches 30, the model reaches an optimum. Reasonable prediction results can provide data support for vehicle dispatching and distribution, solve problems such as energy waste and traffic congestion caused by asymmetric supply and demand, and maximize the interests of passengers, drivers and ride-hailing platforms.

1. Introduction

The transportation system is the basic condition for urban economic and social development. Among the numerous public travel modes, online car-hailing is favored by people by virtue of its convenient and reliable service, and has gradually become one of the important travel modes. However, due to the volatility and randomness of passengers' travel demands, drivers of online ride-hailing vehicles also have great blindness in searching for passengers on the road, which leads to a great contradiction between passenger demand and online ride-hailing services.

In recent years, with the maturity of perception technology and Intelligent Traffic System(ITS), a variety of big data and urban computing concepts have emerged. The popularity of ride-hailing apps like Didi Chuxing and Uber has enabled us to continuously collect GPS data and massive data on ride-hailing demand. This provides us with more accurate dynamic information of traffic than traditional traffic survey and traffic statistics, making accurate and effective short-term demand forecasting possible.

* Corresponding author.
E-mail address: machangxi@mail.lzjtu.cn (C. Ma).

https://doi.org/10.1016/j.physa.2021.125838
0378-4371/

Short-time traffic forecasting methods are mainly divided into four categories: forecasting based on statistical analysis [1], forecasting based on nonlinear theory, intelligent forecasting and combined forecasting model [2]. What are frequently used in data-driven forecasting is the two different approaches which are statistics and neural networks [3].

Statistic approaches refer to the model with fixed structure and parameters [4], the most widely used methods is ARIMA model, which assumes that the traffic forecasting is a stationary process where the mean, variance and auto-correlation are unchanged [5–7]. Statistic methods also include Kalman filters [8,9] and Markov chain [10]. These approaches generally focus on finding the inner characteristics of traffic flow by linear mathematic representations. It is known from previous studies that the traditional time series prediction methods perform well in the stable and linear time series prediction, but often does poorly in the nonlinear, unstable and heteroscedasticity time series prediction. At this time, many prediction methods based on nonlinear theory and intelligent prediction have been widely concerned. The prediction model based on the nonlinear theory mainly excavates the nonlinear characteristics of short-term online car-hailing demand to reflect the random characteristics of urban traffic. There are three representative theoretical methods: wavelet analysis method [11], chaos theory [12] and mutation theory [13]. The intelligent prediction model mainly adopts the "black box" mode to automatically summarize historical data rules for prediction, and mainly includes non-parametric regression prediction [14], neural network model [15–17], Support Vector Machine (SVM) regression prediction [18] and other types [19–21]. Combined forecasting model refers to the comprehensive use of two or more models for forecasting. For example, the ARIMA time series model can combine with neural network [22], and ARIMA can also combined with SVM [23]. Some scholars try to introduce more advanced models which in other fields in the short-term online car-hailing demand forecasting, such as the predicted cloud model [24].

Deep learning method is successfully used in fields such as computer vision [25] and nature language processing [26], which encourage researchers to utilize deep learning method for traffic prediction problems. Part of current papers are about long-term traffic flow forecasts [27], but most of them are about short-term traffic forecasts [28–32]. On the basis of these short-term traffic flow prediction papers, some experts combined a lot of external factors, such as accidents and weather, into the neutral network model to build a more accurate model [33]. Online car-hailing demand forecasting is also traffic forecasting, their goal is to predict a traffic related number for a certain location with a particular designed timestamp. A lot of papers on online car-railing demand are based on one-dimensional data [34–37]. However, traffic flow datasets have both time dimension and space dimension [38], most models cannot fully learn the spatial and temporal characteristics of traffic flow under the limited computational conditions. This makes it particularly difficult to study the car-hailing demand with the consideration of time and space. If the ride-hailing platform can predict the passenger travel demand of different regions in a short time, it can effectively solve the supply and demand contradiction caused by the information asymmetry between passengers and drivers, and realize the reasonable configuration of the ride-hailing vehicles in space and time through online scheduling and distribution.

Previous studies have considered the demand information of online ride-hailing as one-dimensional data, ignoring its two-dimensional nature. However, for the two-dimensional data, there is a lack of suitable models to predict and analyze. This paper uses the previous online car-hailing information to predict the car-hailing demand in the next hour. The goal of prediction is to solve a problem that contains spatial information and time series structure. From this point of view, the research object of this paper is a space–time series prediction problem.

Based on the defects in previous papers, we add our own innovations. The contributions of this paper are 3-fold: (1) The online ride-hailing order information with spatiotemporal feature information is processed into pictures. (2) The Conv-LSTM neutral network is introduced to predict images. (3) A real case was studied to demonstrate the feasibility of this neutral network in dealing with short-term forecasting of online car-hailing demand.

The remainder of this paper is organized as follows: Section 2 introduces a grid-based transportation network representation approach for converting a historical online car-hailing demand into a series of images. And the Conv-LSTM neutral network is employed to capture the spatiotemporal traffic features. In Section 3, historical data of online ride-hailing in Haikou is employed to test the effectiveness of the proposed method. To evaluate the performance of Conv-LSTM, and compare five different training session times. In Section 4, at the end of this paper, the conclusions are presented and future studies are discussed.

2. Methods

A two-step methodology, the first step is about converting online car-hailing demand for images that represent time and space dimensions of a city traffic network as two dimensions of an image. The second step is to employ the deep learning method of a Conv-LSTM neutral network to the image for online car-hailing demand prediction.

2.1. Converting online car-hailing data to images

Online car-hailing has the functions of real-time reservation, selection of the pick-up location and end of the journey after arriving at the destination, so the management platform has accurate travel information, such as longitude and latitude coordinates, journey duration, etc. To predict online car-hailing demand across the entire city, the city must first be divided into $I \times J$ regions. Each grid represents a traffic area, and the cumulative online car-hailing demand in t time intervals is calculated. This can form a space–time matrix B_{t}. The common ways of division are regular region division

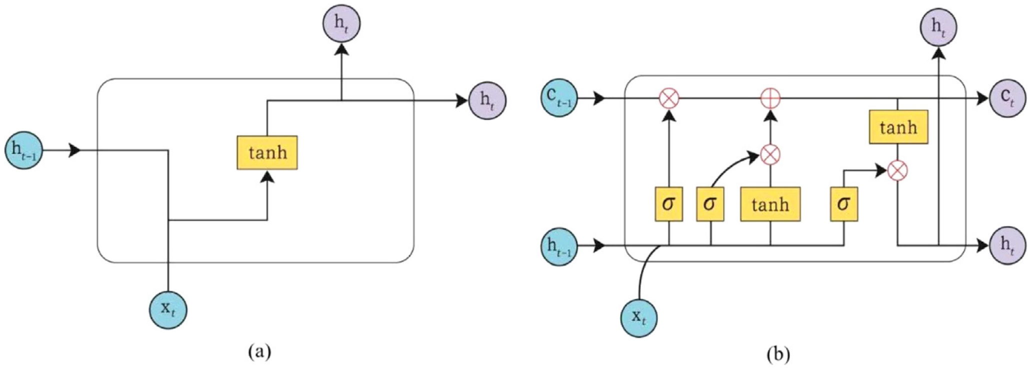

Fig. 1. (a) RNN model (b) FC-LSTM model.

and irregular region division. The regular area division usually selects a rectangular area in the city, and then divides the rectangular area into several grids of equal size according to the latitude, longitude or distance. The size of the grid is determined according to the application requirements. Take the demand forecasting problem as an example. If the online car-hailing platform requires a higher spatial granularity for demand forecasting, a smaller grid size should be selected, and vice versa. Despite the small-scale grid division describes the demand more finely in terms of spatial granularity, this will lead to an increase in the total number of grids and increase the computational complexity, and online car-hailing demand within a small area fluctuates greatly, which makes prediction more difficult. On the contrary, although the description granularity of the large-size grid is coarser, the total amount of the grid is smaller, so the calculation complexity is smaller; and the online car-hailing demand in a large range will show higher regularity, and the prediction difficulty will therefore be reduced. In actual application, the choice of partitioning parameters needs to weigh the effects of description granularity, prediction difficulty, and computational complexity.

Mathematically, we describe the time–space matrix by:

$$B_t = \begin{bmatrix} X_{lon_1,lat_j} & X_{lon_2,lat_j} & \cdots & X_{lon_i,lat_j} \\ X_{lon_1,lat_{j-1}} & X_{lon_2,lat_{j-1}} & \cdots & X_{lon_i,lat_{j-1}} \\ \vdots & \vdots & \ddots & \vdots \\ X_{lon_1,lat_1} & X_{lon_2,lat_1} & \cdots & X_{lon_i,lat_1} \end{bmatrix} \tag{1}$$

where, every $X_{\mathrm{lon_i},lat_j}$ in the matrix represents the online car-hailing demand in the region of $grid(i,j)$ at time t, lon_i lat_j are the latitude and longitude coordinates of x respectively.

2.2. FC-LSTM model

Recurrent Neural Network (RNN) is a network model that evolved from a multilayer recurrent neural network, and differs from Artificial Neural Network (ANN) in that RNN neural networks have a more complex compositional structure, RNN is a model in which cyclic feedback appears on the network. The RNN neural network can "memorize" the learned historical data information through the hidden layer and apply it to the calculation of the current output to achieve the effect of sequential processing and modeling. It shows high performance in accuracy rate and trial scenario.

The Long Short-Term Memory (LSTM) is an improvement of the RNN. The Long Short-Term Memory is also the processing of the network forward calculation. It can also be abbreviated to FC-LSTM. FC-LSTM has a great improvement in the modeling ability of time series. It can not only be applied to speech recognition, language processing and machine translation like the RNN, and also has a significant effect on processing and predicting time series events with long time intervals. FC-LSTM uses a gradient descent network learning method. The key information in the past is retained through forward propagation. After the output value is compared with the expected value and the cross entropy error is obtained, the back propagation algorithm is used to feedback and move forward layer by layer, calculate the gradient of the weight of each layer and modify the weight of the cyclic neuron. In this way, the connection weights between different neurons in the final neural network are adjusted to the optimal.

The main difference between FC-LSTM and RNN is that FC-LSTM adds a new unit structure for judging whether information should be recorded at the moment. This judging unit is generally called a cell. There is only one tanh or ReLU layer in the hidden unit of each step of RNN, while there are four layers in each loop of LSTM, namely 3 Sigmoid layers and 1 tanh layer. The difference between RNN and LSTM is shown in Fig. 1.

There are three kinds of gate controllers in each cell in the FC-LSTM: input gate, forget gate, and output gate. The gate here is actually a full connection layer, the input is a standardized vector, and the output is a vector value between 0 and 1. The FC-LSTM Cell diagram is shown in Fig. 2.

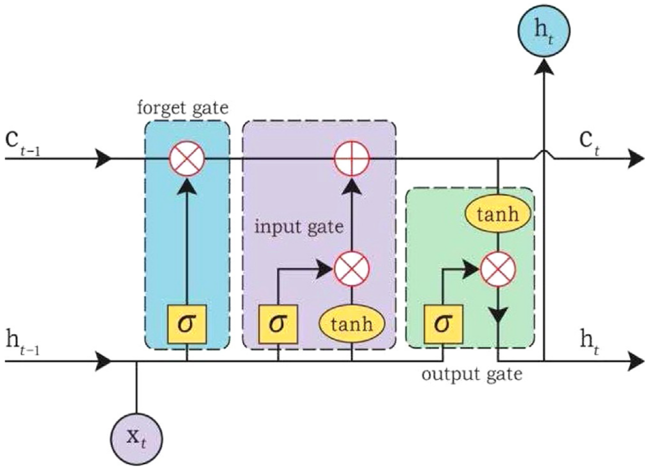

Fig. 2. FC-LSTM Cell diagram.

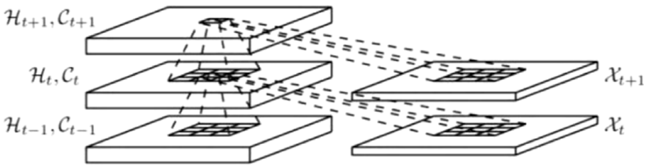

Fig. 3. Inner structure of Conv-LSTM [38].

The function of the gate is to achieve the "filtering" effect by multiplying the output vector of the gate with the vector that needs to be controlled. From the output range of the gate, it can be seen that when the output is 0, the multiplication result of other vectors is 0, which is equivalent to this no information will flow to the next moment. When the output is 1, the result of multiplying with other vectors will not change. At this time, all information flows to the next moment. So the state of the door is always half-open and half-closed. And the relationship between these doors can be expressed by the following equations.

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf} \circ c_{t-1} + b_f) \tag{2}$$

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci} \circ c_{t-1} + b_i) \tag{3}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{4}$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co} \circ c_t + b_o) \tag{5}$$

$$h_t = o_t \circ \tanh(c_t) \tag{6}$$

The above equation is the change process of the unit state at each moment in the LSTM, which is also the processing of the network forward calculation.

2.3. Conv-LSTM

FC-LSTM can handle temporal data well, but it will bring redundancy to spatial data, because spatial data has strong local characteristics, but FC-LSTM cannot describe such local characteristics. Conv-LSTM was born out of a precipitation prediction problem, the problem of predicting precipitation [39]. The precipitation distribution map of the first few hours is given, and the precipitation distribution of the next few hours is predicted. This is done by replacing the input-to-state and state-to-state portions of FC-LSTM from feed-forward computation to convolutional computation. The internal structure of Conv-LSTM is shown as Fig. 3, cell diagram is shown in Fig. 4.

The principle of Conv-LSTM can be expressed by the following formula:

$$i_t = \sigma(W_{xi} * x_t + W_{hi} * h_{t-1} + W_{ci} \circ c_{t-1} + b_i) \tag{7}$$

$$f_t = \sigma(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} \circ c_{t-1} + b_f) \tag{8}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tanh(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c) \tag{9}$$

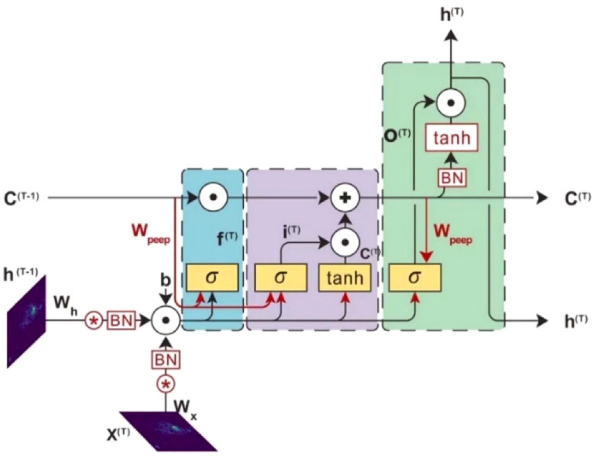

Fig. 4. Conv-LSTM Cell diagram.

$$o_t = \sigma(W_{xo} * x_t + W_{ho} * h_{t-1} + W_{co} \circ c_t + b_o) \tag{10}$$

$$h_t = o_t \circ \tanh(c_t) \tag{11}$$

where $*$ means convolution, x, c, h, i, f, o are all three-dimensional tensors, and their last two dimensions represent the spatial information of rows and columns. We can imagine Conv-LSTM as models which work on the eigenvectors of two-dimensional grids. It can predict the time–space features of the central grid based on the time–space features of the points around them.

3. Experiments

3.1. Data process

This paper selects the online car-hailing data of Haikou to prove the good performance of the model. The longitude and latitude ranges of Haikou are 110.2078°–110.4995° and 19.9291°–20.07°. Data used in this paper came from Didi Chuxing GAIA Initiative [40]. The data set was collected by Didi in Haikou from 2017/5/1–2017/10/31. The data included daily order data of Haikou within the above time range, including the longitude and latitude of the starting and finishing points of orders, order type, travel category and number of passengers' order attribute data. All data involving personal information has been anonymized. The data used in this article is from 0:00 to 23:59 every day on 2017/5/8–2017/8/8. The data is processed using Tensorflow and KERAS. The names and types of data are shown in Table 1.

Haikou is divided into 40*40 grids, and the actual size of each grid is [0.432 km*0.216 km]. The darker the color of the grid, the higher the number of ride-hailing needs. The space–time matrix B_t was constructed through the above data set and Eq. (1). The partition diagram is shown in Fig. 5.

In order to deal with the problem of online car-hailing demand forecasting containing both time and space information at the same time, it can be done by saving the time and space information in a binary vector diagram. Aggregate the online car-hailing demand data within 72 min to generate a picture by using the method of Section 2.1, and obtain a total of 1840 pictures. Here are six pictures to illustrate. Fig. 6 is the cumulative online car-hailing demand from 00:00 to 01:12 every day from 2017/5/8 to 2017/5/13.

The basic process for all operations in this section is shown in Fig. 7.

3.2. Online car-hailing demand prediction

The structure and parameters for ConvLSTM neural network are shown in Fig. 8.

We use Rooted Mean Error (RMSE), Mean Absolute Error (MAE) and R^2 to evaluate our algorithm, which are defined as follows:

$$RMSE = \sqrt{\frac{1}{\xi}\sum_{i=1}^{\xi}(\hat{y}_{t+1}^{i} - y_{t+1}^{i})^2} \tag{12}$$

$$MAE = \frac{1}{\xi}\sum_{i=1}^{\xi}\left|\hat{y}_{t+1}^{i} - y_{t+1}^{i}\right| \tag{13}$$

Table 1
Data field description.

Field name	Data type	Note
order_id	The order ID	String type and desensitized
type	The order time	0=Real-time order, 1=Appointment order
start_dest_distance	The estimated road distance between departure and destination	The estimated road distance between departure and destination
arrive_time	The time when the driver clicks 'arrive'	The time when the driver clicks 'arrive'
departure_time	The departure time	If it is a real-time order, the departure time (departure_time) has the same meaning as the time when the driver clicks 'start billing' (begin_charge_time); if it is an appointment order, it refers to the departure time filled in by the passenger
pre_total_fee	Estimated price	Estimated price based on the starting point and destination filled in by the passenger
normal_time	The time it will take to complete the entire order	minute
dest_lng	Longitude of the destination	The longitude of the destination
dest_lat	Latitude of the destination	The latitude of the destination
starting_lng	Longitude of the starting point	The longitude of the starting point
starting_lat	Latitude of the starting point	The latitude of the starting point

Note: Desensitization is performed on the points of single households, and the starting point and destination are drifted to streets.

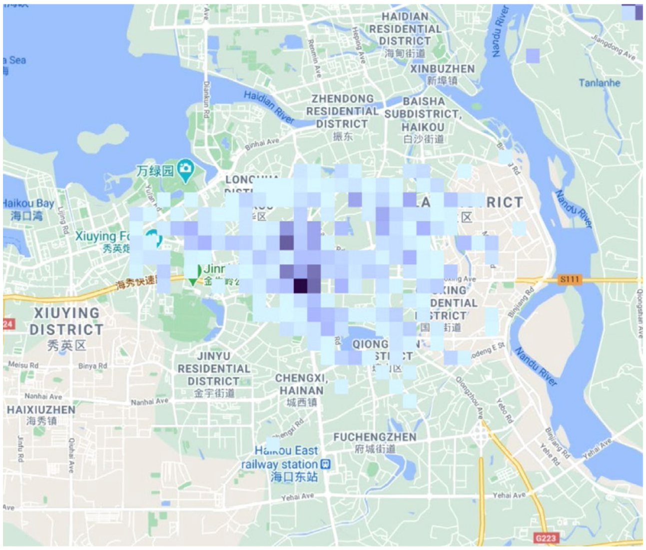

Fig. 5. Schematic diagram of grid division.

$$R^2 = 1 - \frac{\sum_{i=1}^{\xi}(\hat{y}^{i}_{t+1} - y^{i}_{t+1})^2}{\sum_{i=1}^{\xi}(\bar{y}^{i}_{t+1} - y^{i}_{t+1})^2} \tag{14}$$

where $\hat{y}^{i}_{t+1}$ and y^{i}_{t+1} mean the prediction value and real value of region i for the time interval $t+1$, and where ξ is total number of samples.

MAE is the sum of the absolute value of the difference between the target value and the predicted value, which can be used to measure the distance between the predicted value and the ground truth. But it cannot tell whether the predicted value of the model is smaller or larger than the ground truth. RMSE takes the square root of the mean square error to

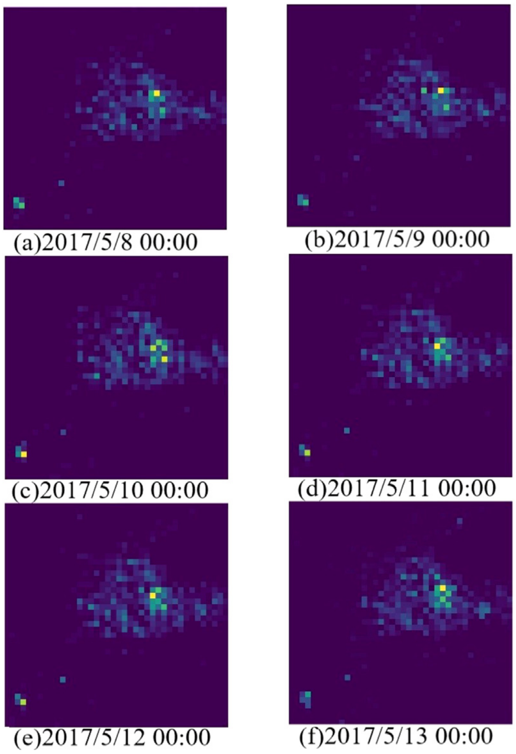

Fig. 6. Cumulative online car-hailing demand.

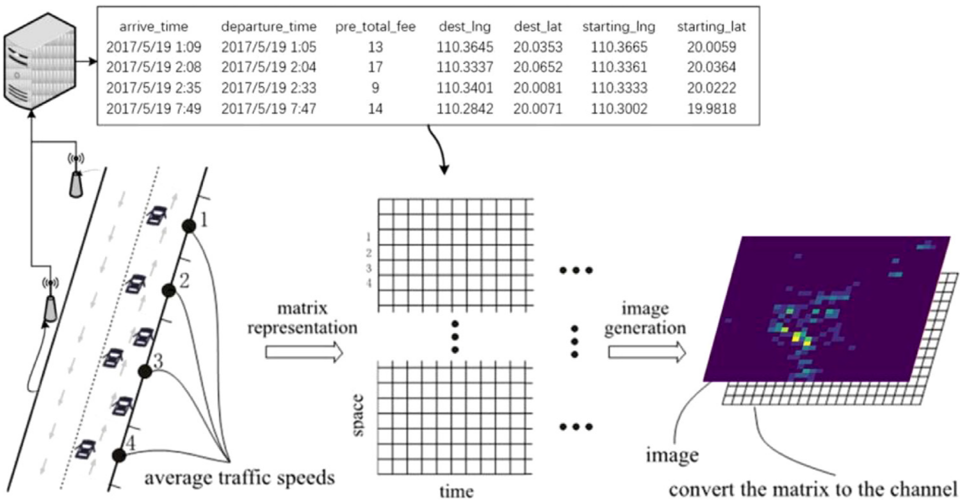

Fig. 7. Online car-hailing data is converted to binary vector illustration.

make the dimensions consistent, which is meaningful for description and representation. These two indexes represent the errors between predicted values and actual values. So the smaller the indexes, the better the model. R^2 is coefficient of determination, which is known as the best measure of linear regression method. If we use the same algorithm model to solve different problems, indexes such as MSE and RMSE cannot reflect the pros and cons of this model for different problems, because of the different dimensions of data sets, and we cannot judge which model is more suitable for the problem. If R^2 is 1, it is a perfect fit, and if R^2 is 0, it is consistent with the baseline. The bigger R^2, the better. If R^2 is less than 0, which means the numerator is greater than the denominator, the training model is going to produce more errors than the mean.

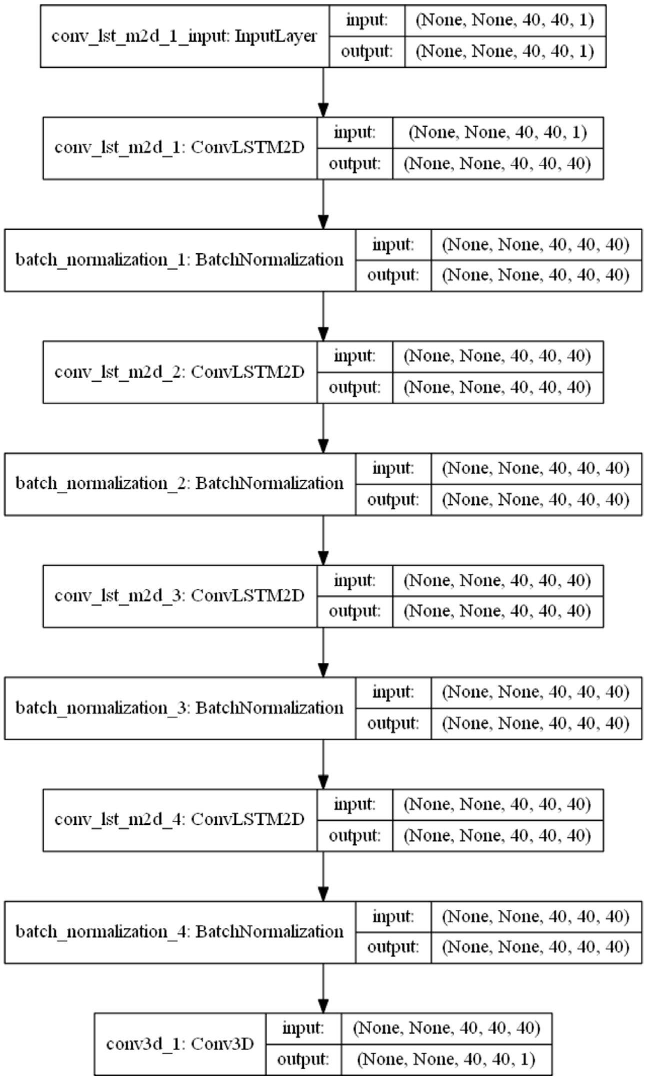

Fig. 8. The structure of ConvLSTM neural network.

Table 2
Indexes with five different training times.

	RMSE	MAE	R^2
Training 10 times	0.040262	0.012452	−0.075871
Training 15 times	0.078489	0.061749	−3.088673
Training 20 times	0.108834	0.041873	−6.861245
Training 25 times	0.041119	0.033520	−0.122140
Training 30 times	0.027319	0.014061	0.504668

The article divides 92 days into two parts, with the first 73 days as training and the last 19 days as testing. The first 10 frames of each day are used as input, and the last 10 frames are used as output. August 8 is used as a comparison to check the accuracy of the prediction. By comparing different training times, it can be found that training 30 times is the most ideal. When the number is lower than 30, the model underfits; when the number is higher than 30, the model overfits. Fig. 9 shows the output results of different training times in the period from 12:00 to 18:00.

Fig. 10 represents the changes of Training loss and Training Mae with the increase of Training times.

When the number of session reaches 30, RMSE and MAE are at their lowest, and R^2 is positive and close to 1. Comparing the indexes with five different training times, it can be seen that training 30 times is the best fit. All the indexes are in Table 2.

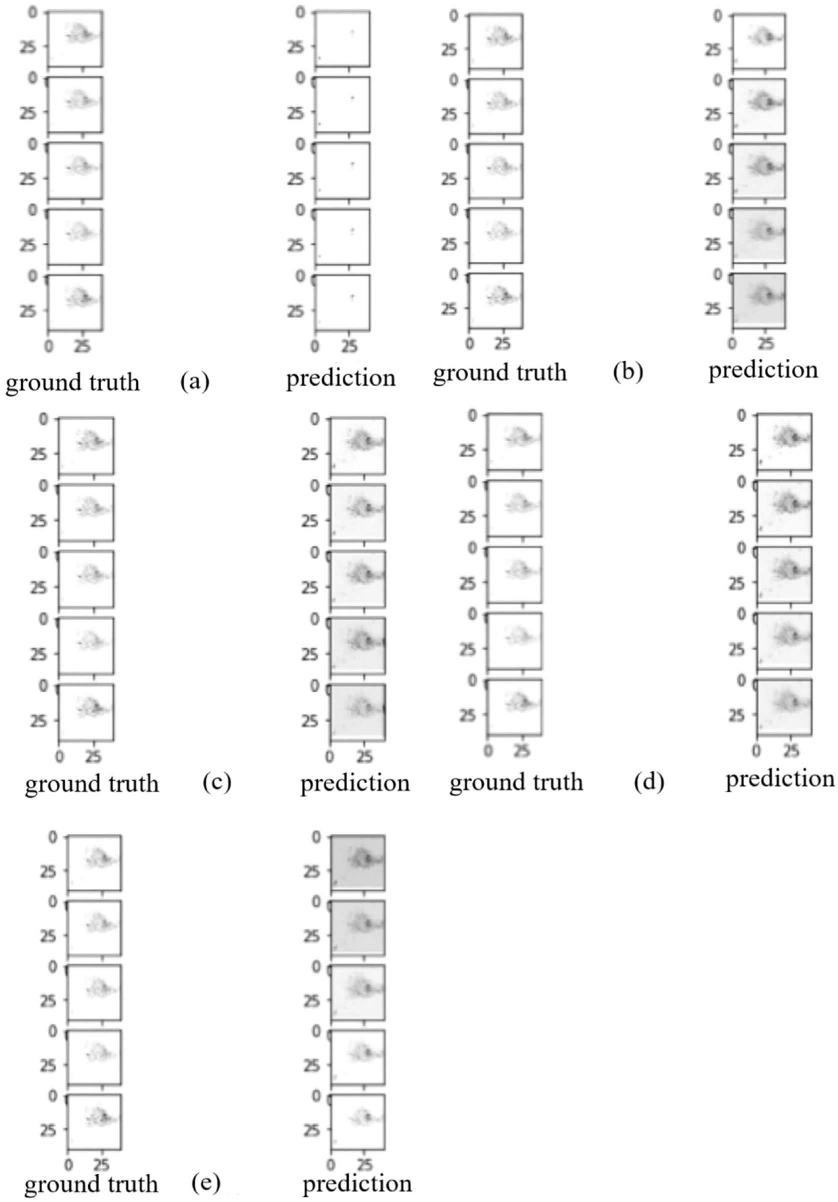

Fig. 9. (a) The number of training sessions is 10. (b) The number of training sessions is 15. (c) The number of training sessions is 20. (d) The number of training sessions is 25. (e) The number of training sessions is 30.

4. Conclusions

This paper proposes an image-based online car-hailing demand prediction method that can extract spatiotemporal traffic-related features to generate images and thus, forecast the short-term online car-hailing demand. The paper contains two main steps. The first step is about converting online car-hailing demand for images that contains spatiotemporal feature of online car-hailing orders. The second step is to employ the deep learning method of a Conv-LSTM neutral network to the image for online car-hailing demand prediction. Conv-LSTM has excellent image prediction properties, so it is ideal for predicting such binary vector figures with spatiotemporal information.

Generally speaking, the paper draws three conclusions:

(a) The spatial and temporal information of online car-hailing demand can be presented in the form of pictures.

(b) By using the historical online car-hailing data of Haikou to conduct experiments, it can be seen that the CONV-LSTM model has a good effect in predicting the binary vector graph, which proves that this model can solve the problem of the online car-hailing demand.

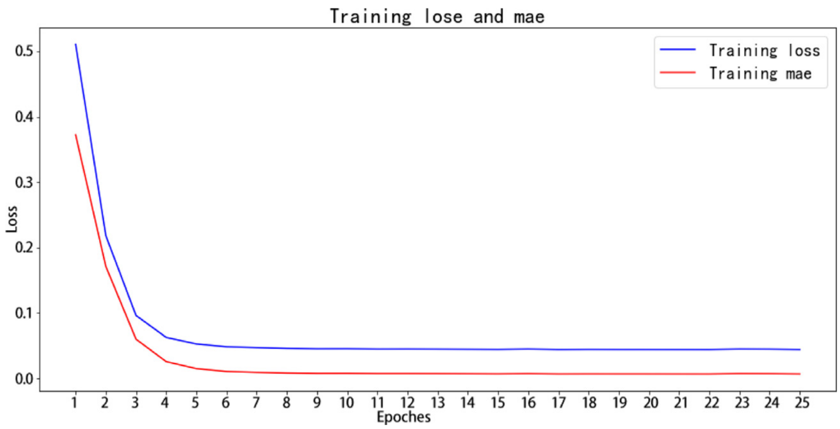

Fig. 10. Training loss and Training MAE.

(c) This paper also points out that when the number of training reaches 30, the model fitting degree is optimal.

From the perspective of managers, reasonable prediction of demand distribution can reduce the waste of road resources and reduce traffic congestion. From an online car-hailing driver's point of view, reasonable prediction can reduce empty-loaded rate and increase profits. From the point of view of citizen passengers, this can save taxi waiting time and living costs.

This paper only divides Haikou into 40*40 grids in proportion, with different road network densities in different grids. In the next paper, it can be divided according to the road network to improve the accuracy.

CRediT authorship contribution statement

Xijin Lu: Conceptualization, Methodology, Writing - original draft. **Changxi Ma:** Writing - review & editing. **Yihuan Qiao:** Visualization, Software.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by the National Natural Science Foundation of China Grant No. 52062027 and 71861023, the Program of Humanities and Social Science of Education Ministry of China Grant No. 18YJC630118, and Foundation of A Hundred Youth Talents Training Program of Lanzhou Jiaotong University.

References

[1] W. Wu, S. Jiang, R. Liu, Economic development demographic characteristics road network and traffic accidents in Zhongshan, China: gradient boosting decision tree model, Transp. A Transp. Sci. 16 (2020) 359.
[2] W. Wu, P. Li, W. Jin, Predicting peak load of bus routes with supply optimization and scaled Shepard interpolation: A newsvendor model, Transp. Res. E 142 (2020) 102041.
[3] M.G. Karlaftis, E.I. Vlahogianni, Statistical methods versus neural networks in transportation research: Differences, similarities and some insights, Transp. Res. C 19 (2011) 387.
[4] B.L. Smith, B.M. Williams, R.K. Oswald, Comparison of parametric and nonparametric models for traffic flow forecasting, Transport. Res. Part C 10 (2002) 303.
[5] B.M. Williams, L.A. Hoel, Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: a theoretical basis and empirical results, Transport. Eng. 129 (2003) 664.
[6] S.R. Chandra, H. Al-Deek, Prediction of freeway traffic speeds and volumes using vector autoregressive models, J. Intell. Transport. Syst. Technol. Plann. Oper. 13 (2009) 53.
[7] B.L. Smith, B.M. Williams, R.K. Oswald, Comparison of parametric and nonparametric models for traffic flow forecasting, Transport. Res. C 10 (2002) 303.
[8] I. Okutani, Y.J. Stephanedes, Dynamic prediction of traffic volume through kalman filtering theory, Transport. Res. B 18 (1984) 1.
[9] H. Chen, S. Grant-Muller, Use of sequential learning for short-term traffic flow forecasting, Transport. Res. C 9 (2001) 319.
[10] Y. Qi, S. Ishak, A hidden markov model for short term prediction of traffic conditions on freeways, Transport. Res. C 43 (2014) 95.
[11] X. Jiang, H. Adeli, Dynamic wavelet neural network model for traffic flow forecasting, J. Transp. Eng. 131 (2005) 771.
[12] J.E. Disbro, M. Frame, Traffic flow theory and chaotic behavior, Transp. Res. Rec. J. Transp. Res. Board 1225 (1989) 109.
[13] G.J. Forbs, F.L. Hall, The applicability of catastrophe theory in modeling freeway traffic operations, Transp. Res. A 24 (1990) 335.

[14] G.A. Davis, N.L. Nihan, Nonparametric regression and short-term freeway traffic forecasting, J. Transp. Eng. 117 (1991) 178.
[15] B.L. Smith, M.J. Demetsky, Short- term traffic flow prediction: neural network approach, Transp. Res. Rec. 1453 (1994) 98.
[16] M.S. Dougherty, M.R. Cobbett, Short-term inter-urban traffic forecasts using neural networks, Int. J. Forecast. 13 (1997) 21.
[17] H. Dia, An object-oriented neural network approach to short-term traffic forecasting, European J. Oper. Res. 131 (2001) 253.
[18] X. Li, Y. Luo, Y. Zhi, Y. Zhang, Chaos support vector machine traffic prediction based on Heuristic algorithm, Comput. Eng. 13 (2011) 163–165.
[19] W. Wu, R. Liu, W. Jin, Simulation-based robust optimization of limited-stop bus service with vehicle overtaking and dynamics: A response surface methodology, Transp. Res. E 130 (2019) 61.
[20] W. Wu, R. Liu, W. Jin, Stochastic bus schedule coordination considering demand assignment and rerouting of passengers, Transp. Res. B 121 (2019) 275.
[21] W. Wu, Y. Xia, W. Jin, Predicting bus passenger flow and prioritizing influential factors using multi-source data: Scaled stacking gradient boosting decision trees, IEEE Trans. Intell. Transp. Syst. (2020) http://dx.doi.org/10.1109/TITS.2020.3035647.
[22] M.V.D. Voort, M. Dougherty, S. Watson, Combining kohonen maps with arima time series models to forecast traffic flow, Transport. Res. C 4 (1996) 307–318.
[23] M. Tan, Y. Li, J. Xu, A hybrid ARIMA and SVM model for traffic flow prediction based on wavelet denoising, J. Highw, Transp. Res. Dev. 26 (2009) 127–132, 138.
[24] Q. Liu, J. Xu, Research on prediction of the short-term traffic flow based on cloud model, Comput. Eng. Desig. 33 (2012) 1953.
[25] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inf. Process. Syst. 25 (2012) 1.
[26] I. Sutskever, O. Vinyals, Q.V. Le, Sequence to sequence learning with neural networks, Adv. Neural Inf. Process. Syst. 9 (2014) 1.
[27] Z. He, C. Chow, J. Zhang, STCNN: A spatio-temporal convolutional neural network for long-term traffic prediction, in: 2019 20th IEEE International Conference on Mobile Data Management (MDM), IEEE, 2019.
[28] X. Zhao, D. Zhang, K. Zhang, A spatial–temporal framework including traffic diffusion for short-term traffic prediction, in: ICCAE 2020: 2020 12th International Conference on Computer and Automation Engineering, 2020.
[29] Y. Lu, J. Yan, Automatic lip reading using convolution neural network and bidirectional long short-term memory, Modern Phys. Lett. B 34 (2020) 01.
[30] X. Ran, Z. Shan, Y. Shi, Short-term travel time prediction: A spatiotemporal deep learning approach, Modern Phys. Lett. B 18 (2019) 1087.
[31] H. Yu, Z. Wu, S. Wang, Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks, Sensors 17 (2017) 1501.
[32] Y. Qiao, Y. Wang, C. Ma, Modern Phys. Lett. B (2020) 2150042,
[33] X. Ma, Z. Tao, Y. Wang, Long short-term memory neural network for traffic speed prediction using remote microwave sensor data, Res. Part C 54 (2015) 187.
[34] Z. Qiu, L. Liu, G. Li, Taxi origin–destination demand prediction with contextualized spaticial-temporal network, in: 2019 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2019.
[35] L. Xu, Y. Guo, Short-term forecasting model of demand for network booking taxi based on GWO-LSTM, Autom. Instrum. 35 (2020) 86.
[36] J. Ke, H. Zheng, H. Yang, Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach, Transp. Res. C 85 (2017) 591.
[37] Y. Lin, N. Zou, Short-term prediction model of taxi passenger demand basedon operation systems, J. Northeast. Univ. Nat. Sci. 37 (2016) 1235.
[38] S. An, L. Xu, G. Chen, A new car-following model on complex road considering driver's characteristics, Modern Phys. Lett. B 34 (2020) 16.
[39] X. Shi, Z. Chen, H. Wang, Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, NIPS, 2015.
[40] Didi Chuxing, https://outreach.didichuxing.com/app-vue/HaiKou?.