2019 International Workshop on Smart Info-Media Systems in Asia (SISA 2019), Sep. 4-6, 2019

Short-term Traffic Flow Prediction Based on Recurrent Neural Network

Bohan Huang, Hongye Yang and Yun Bai
Inner Mongolia University Of Technology, China

***Abstract*—With the increasing traffic congestion problem, the development of intelligent transportation systems has become an important measure for countries to solve traffic problems. As an important part of the intelligent transportation system, short-term traffic flow prediction plays an important role in pedestrian traffic control and traffic control and traffic guidance. The three algorithms of LSTM(Long Short-Term Memory), GRU(Gated Recurrent Unit) and BRNN(Bi-directional RNN) in the recurrent neural network are selected for prediction and compared with real data. The results show that the recurrent neural network is a feasible traffic flow prediction model.**

I. INTELLIGENT TRANSPORTATION SYSTEM

With the rapid development of the economy, the number of motor vehicles has increased year by year, and the resulting traffic congestion has become a major factor plaguing urban development. The development of intelligent transportation systems is an important measure to solve urban traffic problems. Intelligent Traffic System (ITS) is an effective combination of information technology, computer technology, data communication technology, sensor technology, electronic control technology, automatic control theory, operations research, artificial intelligence, etc. for transportation and traffic control. And vehicle manufacturing, strengthen the connection between vehicles, roads, and users, thus forming an integrated transportation system that ensures safety, improves efficiency, improves the environment, and saves energy [1].

The development of intelligent transportation in China has gone through the following stages [4]:

1. Starting stage (before 2000)

In the 1970s and 1980s, China mainly carried out some basic research on urban traffic signal control. In the 1990s, domestic first-tier cities such as Beijing, Shanghai, and Shenzhen introduced advanced foreign technologies and carry out innovative research on the basis of learning. During the "10th Five-Year Plan" period, intelligent transportation has made some breakthroughs in key technologies, and established demonstration points for electronic toll collection systems and traffic management systems; China's transportation system has entered the stage of promotion, application and improvement, but with advanced countries abroad. Compared with the overall technology and application level, there is still a big gap, and the effect of solving the increasingly serious contradiction between supply and demand of transportation is limited.

2. Substantial construction phase (2000-2005)

During the period from the 10th Five-Year Plan to the 12th Five-Year Plan, China’s investment in Intelligent Transportation Systems (ITS) has gradually increased. Among them, the investment in the “Tenth Five-Year Plan” on ITS projects reached 1.5 billion yuan, and it has reached the “Twelfth Five-Year Plan”. During the period, the planned total investment of 100 billion yuan was devoted to the intelligent transportation system, and the investment increased greatly during the period. China's intelligent transportation companies have also developed, investing a large amount of money for the development, production and popularization of intelligent transportation. These have created favorable conditions for the development of intelligent transportation.

3. High-speed development stage (2005-present)

The urban intelligent transportation industry is now in the growth stage. The layout and framework construction of intelligent transportation systems in first-tier cities such as Beijing, Shanghai and Guangzhou have been initially completed, and more intensive deployment of related systems and equipment has begun. In addition, the large-scale international events such as the Beijing Olympic Games, Shanghai World Expo, Guangzhou Asian Games, and Shenzhen Universiade have also greatly promoted the growth of investment in urban intelligent transportation systems.

In recent years, with the improvement of storage devices and computer performance, neural networks have re-entered people's vision, and research on deep learning and deep neural networks has been greatly developed. Short-term traffic flow prediction, as a key technology in ITS applications such as traffic control and traffic guidance, has always been a research hotspot in the field of intelligent transportation. How to apply deep neural networks to short-term traffic flow prediction has become a new research trend.

II. NEURAL NETWORK

According to different model structures, neural networks mainly include Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Generative Adversarial Networks (GAN), etc. The main application areas of the network are:

TABLE I
TYPE USES OF COMMON MODEL

CNN	RNN	GAN
Behavior recognition	Speech Recognition	3D video
Image recognition and processing	Speech modeling	Image generation
Scene classification	Time series prediction	Image PS

Recurrent neural networks, because of their time memory function, are often used in many fields such as language recognition, intelligent dialogue, and time series prediction. Take language recognition as an example: When predicting the next word of a sentence, you usually need to use the previous word. At this time, the performance of RNN is better than that of traditional neural network. Since the traffic flow data is time-related time series data, it is suitable to use a recurrent neural network as a prediction algorithm.

III. RECURRENT NEURAL NETWORK

In the RNN, the output of the sequence is not only related to the current input but also to the previous hidden layer output. RNN has temporal memory and strong nonlinear learning ability compared to other models because it has a loopback structure compared to other models, as shown in Fig. 1:

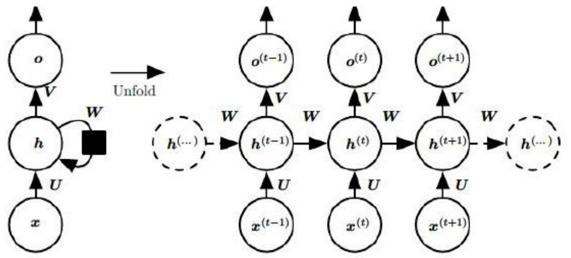

Fig. 1 RNN structure.

x_{t-1} is the input of the previous moment, w_{t-1} is the loss function of the previous moment, h_{t-1} is the output of the hidden layer at the previous moment, x_t is the input of the current moment, h_t is the output of the hidden layer at the current time, w_t is the loss function of the current moment, b is the bias, f is the activation function, as can be seen from Fig. 1:

Hidden layer output from the previous moment:

$$h_{t-1} = f(\mathrm{w}_{t-1}\mathrm{x}_{t-1} + \mathrm{b}) \tag{1}$$

Hidden layer output at the current moment:

$$h_t = f(\mathrm{w}_t\mathrm{x}_t + \mathrm{w}_{t-1}\mathrm{h}_{t-1} + \mathrm{b}) \tag{2}$$

Hidden layer output at t+1:

$$h_{t+1} = f(\mathrm{h}_{t+1}\mathrm{x}_{t+1} + \mathrm{h}_t\mathrm{x}_t + \mathrm{b}) \tag{3}$$

When the input variable is x_t, the hidden layer output is not only related to the current input x_t, but also related to the output $h_{t-1}, h_{t-2}, \cdots, h_{t-n}$ of the hidden layer that has been experienced before, which allows the RNN to use the information of the previous and subsequent time points to find the output of the corresponding input sequence. Thus having time memory.

A. LSTM(Long Short-Term Memory)

In order to overcome the problems of gradient explosion, gradient disappearance and long-term dependence in traditional RNN, Hochreiter and Schmidhuber proposed the long-short-time memory model LSTM in 1997, and Fig. 2 shows the typical structure of LSTM.

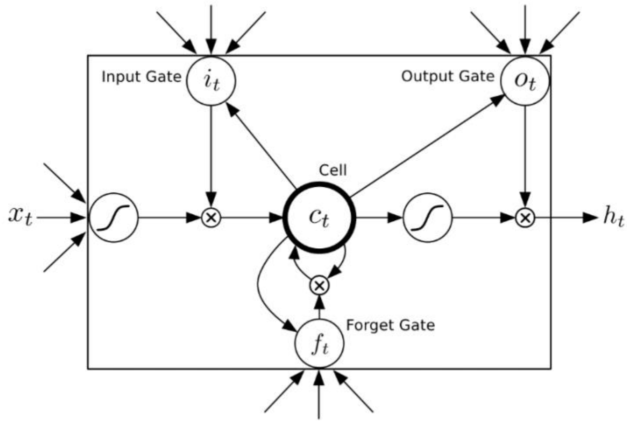

Fig. 2 LSTM structure.

Suppose the input time series is $X = (x_1, x_2, \cdots, x_n)$, hidden state memory unit is $H = (h_1, h_2, \cdots, h_n)$, output time series is $Y = (y_1, y_2, \cdots, y_n)$, the calculation formula of LSTM is as follows :

$$h_t = H(W_{hx}x_t + W_{hh}h_{t-1} + b_n) \tag{4}$$

$$p_t = W_{hy}y_{t-1} + b_y \tag{5}$$

Where W represents the weight matrix, b represents the offset vector, y represents the actual output, and p represents the predicted traffic flow. The calculation formula for the hidden state memory unit is as follows:

$$i_t = \sigma(W_{ix}x_t + W_{hh}h_{t-1} + W_{ic}c_{t-1} + b_i) \tag{6}$$

$$f_t = \sigma(W_{fx}x_t + W_{hh}h_{t-1} + W_{fc}c_{t-1} + b_f) \tag{7}$$

$$c_t = f_t * c_{t-1} + i_t * g(W_{cx}x_t + W_{hh}h_{t-1} + W_{cc}c_{t-1} + b_c) \tag{8}$$

$$o_t = \sigma(W_{ox}x_t + W_{hh}h_{t-1} + W_{oc}c_{t-1} + b_o) \tag{9}$$

$$h_t = o_t * h(c_t) \tag{10}$$

Here σ is the standard sigmoid function, defined as (11), * represents the scalar product of two matrices or vectors, and g and h are the two ranges of sigmoid function from [-2, 2] and [-1, 1], respectively.

$$\sigma(x) = \frac{1}{1+e^{x}} \tag{11}$$

For the objective function we use the following square loss function:

$$e = \sum_{t=1}^{n} (y_t - p_t)^2 \tag{12}$$

B. GRU(Gated Recurrent Unit)

LSTM has a good performance in processing sequences, but its structure is more complicated. In order to simplify the structure and obtain better performance, many variants have appeared. GRU is a model that changes the internal structure based on LSTM.

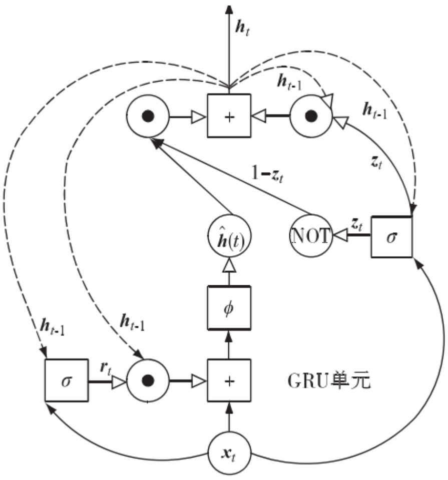

Fig. 3 GRU structure.

1. Enter the reset gate together with the output h_{t-1} of the previous hidden layer. The two are weighted and added together to form the output r_t of the reset gate through the activation function σ . This process discards the useless information. Reset the gate output as shown in (13):

$$r_t = \sigma(\mathrm{W}_{rx}\mathrm{x}_t + \mathrm{W}_{rh}\mathrm{h}_{t-1}) \tag{13}$$

2. The hidden layer output h_{t-1} of the previous node and the reset gate output r_t are Hadamard products, and then weighted and added with the current input x_t to form an output h_t .

$$\hat{h}_t = \varphi(\mathrm{W}_{hx}\mathrm{x}_t + \mathrm{r}_t \odot \mathrm{W}_{hh}\mathrm{h}_{t-1}) \tag{14}$$

3. The update gate weight z_t is obtained by updating the gate, and is used to determine how much weight in $\hat{h}_{t-1}$ can enter the output of the current data, and then weight the output of the reset gate to obtain the final output.

$$z_t = \sigma(\mathrm{W}_{zx}\mathrm{x}_t + \mathrm{W}_{zh}\mathrm{h}_{t-1}) \tag{15}$$

$$h_t = (1-\mathrm{z}_t) \odot \hat{\mathrm{h}} + \mathrm{z}_t \odot h_{t-1} \tag{16}$$

Where x_t is the input to the current neural network and h_{t-1} is the activation value of the previous hidden node output. σ denotes the sigmoid function, φ denotes the tanh function, $\odot$ denotes the Hadamard product, and all W are weight parameter matrices to be trained by the model.

C. BRNN（Bi-directional RNN）

BRNN (Bi-directional RNN) is an extended form of unidirectional RNN. Ordinary RNNs only focus on the previous content, while BRNN focuses on the context before and after, and can use more information to make predictions.

Structurally, the BRNN consists of two RNNs that are opposite in direction, and the two RNNs are connected to the same output layer. The BRNN not only accepts the hidden layer output of the previous moment as an input, but also accepts the hidden layer output of the next moment as an input, which achieves the above-mentioned simultaneous attention to the context. The specific structure is as follows:

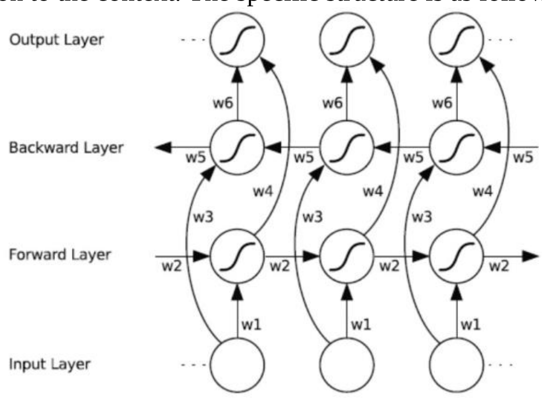

Fig. 4 BRNN structure.

IV. SHORT-TERM TRAFFIC FLOW FORECAST

The data in this paper is the GPS data of K5, 32, and 73 buses obtained from the Hohhot Bus Corporation of Inner Mongolia Autonomous Region. It includes the running hours from 6 o'clock to 20 o'clock in the six months from May 2017 to October 2017. Data and data are automatically collected by the GPS acquisition device on the bus. GPS data is saved in TXT text format, and Fig. 5 is part of the data.

[{"lng":"111.68443877362611","odometer":0,"busSpeed":"0.00","relativeLocati
"lat":"40.867908653259384","devUuid":"65b531cbb6cb4c76bb7e","sationUuid":""
"gatherTime":1495489853000,"allAlarms":"7","busUuid":"2f12cdc126854e86a5b8"
","drvIcCard":"3879948476","lineType":"0","isOffset":"1","driverName":"云志
","distanceToPrePosition":0,"lineUuid":"c2e0313000fba04c940a4be0f9edb9c521c
71cd5d740c18101","cursorOverGround":"486","posIsInStation":"","sationName":
us":"0","driverUuid":"2b4bcd560d1b404aa909cdb01014ab21851e","posUuid":"5934
877362611","odometer":0,"busSpeed":"0.00","relativeLocation":0,"realTimeSta
3259384","devUuid":"65b531cbb6cb4c76bb7e","sationUuid":""},{"lng":"111.6844
89932000,"allAlarms":"7","busUuid":"2f12cdc126854e86a5b8","lat":"40.8679086
9948476","lineType":"0","isOffset":"1","driverName":"云志军","gatherTime":1
sition":0,"lineUuid":"c2e0313000fba04c940a4be0f9edb9c521c8","drvIcCard":"38
cursorOverGround":"486","posIsInStation":"","sationName":"","distanceToPreP
d":"2b4bcd560d1b404aa909cdb01014ab21851e","posUuid":"9f811b52b1844cbca6ce",
er":0,"busSpeed":"0.00","relativeLocation":0,"realTimeStatus":"0","driverUu
:"65b531cbb6cb4c76bb7e","sationUuid":""},{"lng":"111.68443877362611","odome
s":"7","busUuid":"2f12cdc126854e86a5b8","lat":"40.867908653259384","devUuid
":"0","isOffset":"1","driverName":"云志军","gatherTime":1495490022000,"allA
d":"c2e0313000fba04c940a4be0f9edb9c521c8","drvIcCard":"3879948476","lineTyp
"486","posIsInStation":"","sationName":"","distanceToPrePosition":0,"lineUu
4aa909cdb01014ab21851e","posUuid":"a2c9bef3f1da44c1a9cf","cursorOverGround"

Fig.5 TXT text data

Contains 22 parameter information, but in terms of traffic state prediction and recognition, not every parameter has value. After analyzing the meaning of each parameter, the meaning of each parameter is as follows:

TABLE II
THE PARAMETER ANALYSIS OF BUS GPS DATA

Parameter	Comment	Parameter	Comment
Lng	GPS longitude	devUuid	Device ID
Busspeed	Bus speed	odometer	Odometer
realTimeStatus	Real-time location	relativeLocation	Relative position
posUuid	GPS device number	driverUuid	Driver number
posIsInStation	In and out station identification	cursorOverGround	Bus azimuth
distanceToPre Position	Distance from next stop	sationName	Bus station name
drvIcCard	Driver IC card number	lineUuid	Line number
isOffset	Whether to leave	lineType	Line type
gatherTime	GPS data acquisition time	driverName	Driver name
busUuid	Vehicle number	allAlarms	Number of warnings
lat	GPS latitude	sationUuid	Bus stop number

Considering the research needs of the subject, the five parameters of longitude, latitude, speed, site and time are retained, and the remaining parameters are deleted. Part of the data extraction results are shown in Fig. 6:

Longitude	Latitude	Speed	Time	Site
111.6711778	40.8375302	0	2017-9-1 14:23	八一幼儿园
111.6711778	40.8375302	0	2017-9-1 14:23	八一幼儿园
111.6711778	40.8375302	0	2017-9-1 14:23	八一幼儿园
111.6711778	40.8375302	0	2017-9-1 14:23	八一幼儿园
111.6711778	40.8375302	0	2017-9-1 14:23	八一幼儿园
111.6713032	40.83755053	0	2017-9-1 14:23	八一幼儿园
111.6715403	40.83757433	0	2017-9-1 14:23	八一幼儿园
111.6718395	40.83761671	20	2017-9-1 14:23	八一幼儿园
111.6722307	40.83767256	20	2017-9-1 14:23	八一幼儿园
111.6727219	40.83773838	23	2017-9-1 14:23	八一幼儿园
111.6731731	40.83779598	23	2017-9-1 14:24	八一幼儿园
111.6733921	40.83781812	22	2017-9-1 14:24	土默特中学
111.6736711	40.83784683	21	2017-9-1 14:24	土默特中学
111.6739985	40.83787922	20	2017-9-1 14:24	土默特中学
111.6743679	40.83791821	18	2017-9-1 14:24	土默特中学
111.6746119	40.83794392	5	2017-9-1 14:24	土默特中学
111.6746187	40.83794203	0	2017-9-1 14:24	土默特中学
111.6746187	40.83794393	0	2017-9-1 14:24	土默特中学

Fig.6 The Extraction of GPS Data

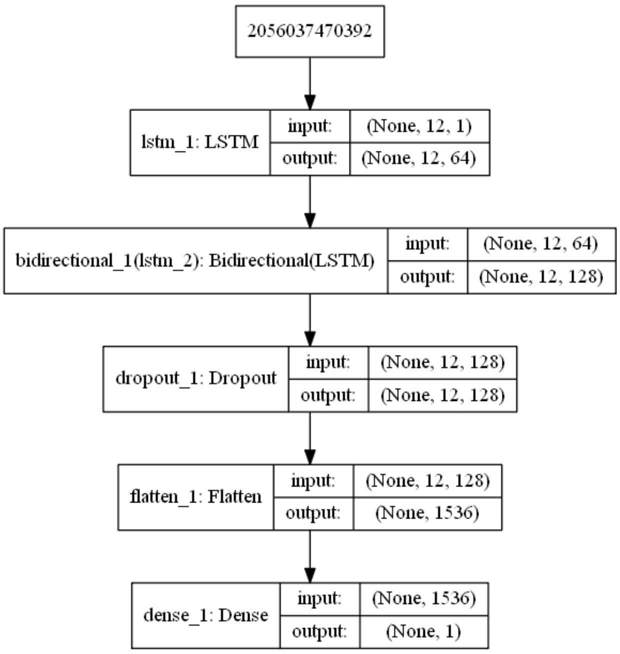

Fig. 7 Network structure and number of nodes.

The LSTM, GRU, and BRNN in the recurrent neural network are used as prediction algorithms and compared. The selected network structure and number of nodes are shown in Fig. 7.

V. CONCLUSION

The predictions are made by LSTM, GRU, and BRNN. The prediction results from 8:00-8:20 in the morning peak hours of a certain day are compared with the real values as shown in Fig. 8. Fig. 8 shows that the prediction results are basically consistent with the actual data, and the prediction results are informative.

TABLE III
COMPARISON OF THREE METHODS

Average of five training sessions	LSTM	GRU	BRNN
Precision	0.7523	0.8117	0.7861
Loss	0.0072	0.0164	0.0128

After five trainings to obtain the average accuracy, the LSTM algorithm has an accuracy of 75.23%, the BRNN algorithm has a precision of 78.61%, and the GRU algorithm has an accuracy of 81.17%. The results show that the recurrent neural network is a feasible traffic flow prediction model.

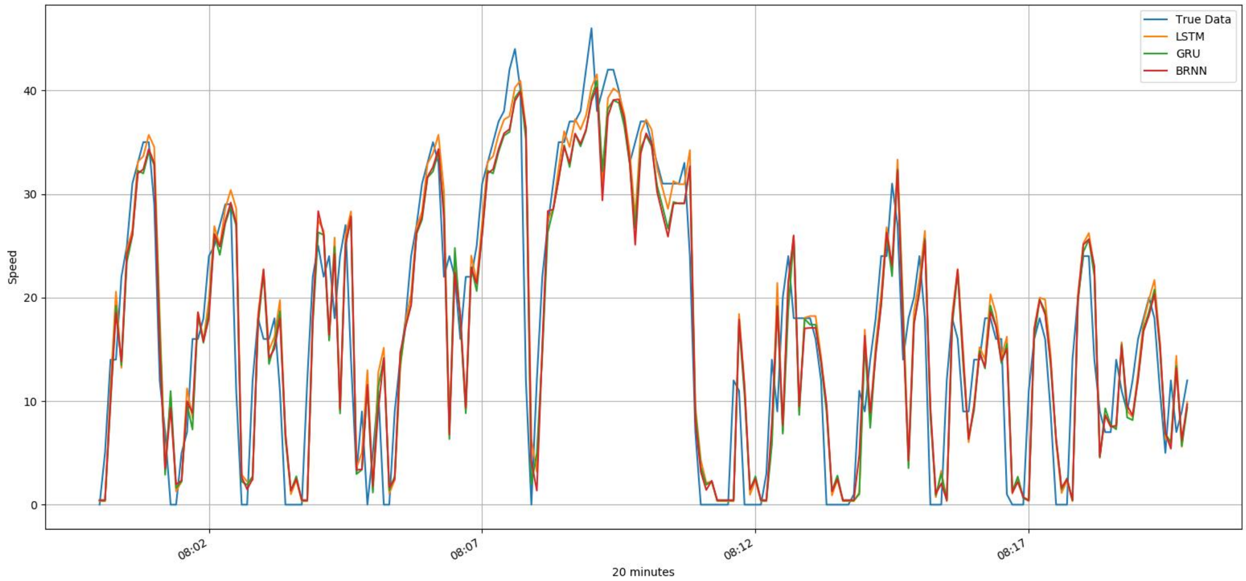

Fig.8 Comparison of predicted and true values

ACKNOWLEDGMENT

This work is supported by the School Key Fund Project under Grant No. ZD201226, Inner Mongolia Autonomous Region Graduate Research and Innovation Project under Grant No. S20181111952Z.

REFERENCES

[1] Intelligent Traffic System [EB/OL]. (2015-07-07)[2019-05-20]. https://baike.baidu.com/item/%E6%99%BA%E8%83%BD%E4%BA%A4%E9%80%9A%E7%B3%BB%E7%BB%9F/1188149.
[2] Su Z. Research of Intelligent Traffic Development in ShenZhen[D].Dalian Maritime University, 2012.
[3] Fu R , Zhang Z , Li L . Using LSTM and GRU neural network methods for traffic flow prediction[C]// 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC). IEEE, 2016.
[4] The Status Quo of China's Intelligent Transportation Development in 2018[EB/OL]. (2018-07-31)[2019-05-23]. http://www.gdupi.com/Common/news_detail/article_id/3384.html
[5] Wang X, Xu L. Short-term Traffic Flow Prediction Based on Deep Learning[J]. Journal of Transportation Systems Engineering and Information Technology, 2018
[6] Qian W, Che K, Li B. Short-term Traffic Flow Prediction Based on Combined Models[J]. Control Engineering of China,2019.
[7] LV Y ,Duan Y Kang W.Traffic flow prediction with big data:A deep learning approach.[J]IEEE Transactions on Intelligent Transportation Systems.2015,16(2): 865-873.
[8] Kuremeo "f. Kimura S. Kobavashi K. "Time series forecasting using a deep belief network with restricted Boltzman machine,[J]Neuro computing.2014,137:47-56.