The Journal of Supercomputing (2022) 78:12686–12709
https://doi.org/10.1007/s11227-022-04386-7

Attention-based Conv-LSTM and Bi-LSTM networks for large-scale traffic speed prediction

Xiaojian Hu[1,2,3,4] · Tong Liu[1] · Xiatong Hao[1] · Chenxi Lin[1]

Accepted: 17 February 2022 / Published online: 11 March 2022

Abstract
Timely and accurate traffic speed prediction has gained increasing importance for urban traffic management and helping one to make advisable travel decision. However, the existing approaches have difficulty extracting features of large-scale traffic data. This study proposed a hybrid deep learning method named AB-ConvLSTM for large-scale traffic speed prediction. The proposed model consists of a convolutional-long short-term memory (Conv-LSTM) module, an attention mechanism module, and two bidirectional LSTM (Bi-LSTM) modules. Conv-LSTM networks are used to extract the spatiotemporal features of traffic speed data. In addition, the attention mechanism module is introduced to enhance the performance of Conv-LSTM by automatically capturing the importance of different historical periods to the final prediction and assigning corresponding weights. What's more, two Bi-LSTM networks are designed to extract daily and weekly periodic features and capture variation tendency from forward and backward traffic data. Experimental results carried out on urban road networks show that the proposed model consistently outperforms the competing models.

Keywords Traffic speed prediction · Conv-LSTM · Bi-LSTM · Attention mechanism · Spatiotemporal · Periodic

1 Introduction

The increasing number of vehicles has led to a series of problems, such as traffic congestion, traffic accidents, energy dependence, and air pollution. To solve these problems, intelligent transportation system (ITS) is implemented in the city, which is a large-scale, real-time and efficient system to realize intelligent management of urban road networks based on advanced information technology and

✉ Tong Liu
lxtzmz@163.com

Extended author information available on the last page of the article communication technology. Large-scale traffic prediction plays a crucial role in ITS. It enables the system to achieve overall transportation management in a proactive way [1]. Accurate large-scale traffic prediction can help managers to carry out effective traffic management and allocate resources systematically. Besides, it can provide travellers with an awareness of overall traffic condition of the road networks, enabling them to make more appropriate travel choices.

Traffic prediction models can be roughly divided into parametric models and non-parametric models [2]. The parametric model is a mathematical model with a specific function expression form, and the model parameters can be calculated based on theoretical assumptions and empirical data. The classical parameter models include historical average method (HA), autoregressive integrated moving average (ARIMA) [3, 4], and Kalman filtering [5]. Non-parametric models are data-driven methods, including machine learning methods and deep learning methods [6]. These methods do not require any prior knowledge and can obtain eximious performance based on sufficient historical data. Common machine learning methods support vector machines (SVM) [7–10], Bayesian [11] and k-nearest neighbor (KNN) model [12–14] have been applied to traffic prediction. Traditional parametric models and machine learning methods are widely used in road segments prediction, but they have poor performance in large-scale traffic prediction [15]. Large-scale traffic prediction is a challenging task due to the uncertainty of the traffic condition and complex network topology. Specifically, the traffic demands and volume change from time to time and the traffic condition of road segments interact with one another. The modeling of nonlinear spatiotemporal dynamics in traffic will be a significant issue for traffic prediction.

In recent years, the deep learning methods have witnessed a great success in traffic prediction [16, 17]. They can handle multi-dimensional data and have substantial flexibility and extract traffic data's temporal and spatial characteristics [18]. Long short-term memory network (LSTM) is used to handle traffic flow with long time lags [19–21]. The convolutional neural network (CNN) is used to extract spatial correlations of a road network. Deep neural network (DNN) methods are used in traffic prediction to enhance the model's performance by deeper architectures [22–24]. Besides, many researchers construct a hybrid model to learn more features of traffic data, such as combining machine learning methods with neural networks [25, 26]. These deep learning methods achieve better performance than traditional models. However, most researchers build traffic prediction models considering nonlinear spatiotemporal features on a single road. Limited studies have addressed the modeling of nonlinear spatiotemporal and periodic dynamics in large-scale traffic data.

In this paper, we propose a hybrid deep learning method based on ConvLSTM, attention mechanism and Bi-LSTM, called AB-ConvLSTM, for large-scale traffic speed prediction. Three main features of traffic data in road networks are considered, including temporal feature, spatial feature, and periodicity feature. ConvLSTM combined with attention mechanism is used to extract temporal and spatial features. Bi-LSTM is used to extract periodicity feature. The main contributions of this paper are as follows.

- We propose a large-scale traffic speed prediction model, which exploits spatiotemporal features and periodic features of network traffic. These features can be well fused by the deep learning architecture of the proposed model.
- We use attention mechanism module to enhance the performance of ConvLSTM by capturing the importance that each historical sequence contributes to the final prediction.
- We use two Bi-LSTM networks as an auxiliary module to extract periodic features and variation tendencies of the traffic data.
- We use the urban road networks dataset of Nanjing to evaluate the performance of the proposed model. The results demonstrate that the proposed model consistently outperforms the competing models within an acceptable training time.
- We estimate the traffic condition of urban road network with high accuracy based on the predicted traffic speed and its corresponding relationship with traffic condition.

The remainder of this paper is organized as follows: Sect. 2 introduces the related work about deep learning models for traffic speed prediction. Section 3 presents the structure of the proposed model. Section 4 conduct experiments on urban road networks dataset and evaluate the performance of different models. Finally, Sect. 5 summarizes the work of the full text and discusses future research.

2 Related work

Recently, deep learning has drawn much academic attention. It has achieved abundant achievements in many fields, such as data mining, machine translation, natural language processing, recommendation and personalization technology. The latest deep learning algorithms have far surpassed traditional machine learning algorithms for data prediction and classification accuracy. The deep learning method can automatically filter data and extract high-dimensional features of the data by deep architectures. Many scholars have applied deep learning to extract spatiotemporal features of traffic and achieved good results [27].

Ma et al. [21] first proposed a LSTM network for traffic speed prediction using remote microwave sensor data. The result showed that LSTM is an effective method for short-term speed prediction without prior information of time lag. Khan et al. [28] used recurrent neural network (RNN), the gated recurrent unit (GRU) and LSTM units to predict hourly traffic volume and annual average daily traffic. Lv et al. used the deep learning architecture embedding stacked autoencoder (SAE) as the primary network structure block to predict traffic flow. The PEMS dataset is used to train the model, and different network structures are designed for different prediction steps [29]. LSTM has been proven to have powerful time feature extraction capability. However, it cannot extract spatial features and performs poorly in multiple road segments prediction. To better extract the spatial characteristics of the data, many scholars use CNN method to predict the traffic of multiple road segments [30, 31]. Ma et al. [15] tried CNN for large-scale traffic speed prediction, and this method learned traffic as images, which can achieve better performance in spatiotemporal traffic feature extraction.

Attention mechanism is used in neural networks to improve the performance of models such as LSTM and CNN [32]. Attention mechanism can automatically identify which part of the model input is more important, and then assign greater weight to this part. This method has become a hot issue in recent years and has been used in traffic prediction [33]. Wu et al. [34] proposed an attention-based LSTM (ATT-LSTM) model for predicting short-term traffic speed on urban roads. The results shown that attention mechanism can properly assign weights to distinguish the importance of traffic time sequences and improve the computational efficiency of the prediction model. Fei et al. [35] proposed an attention-based Bi-LSTM model for travel order quantity prediction, which proved that more hidden information can be learned by Bi-LSTM layers. Liu et al. [36] proposed an attention mechanism based model for traffic flow prediction, which extracts the spatiotemporal characteristics of traffic flow data through multi-head attention. Sun et al. [37] propose a multi-component attention (MCA) method for traffic flow prediction. Through the introduction of the attention mechanism, highly related historical information may be connected for multi-component flow data in the final prediction.

Some scholars combined CNN and LSTM to construct a better prediction model that can extract spatiotemporal features of data. Shi et al. [38] first proposed a convolutional LSTM (ConvLSTM) structure based on a fully connected LSTM network to cope with precipitation nowcasting problems. They added a Conv-LSTM structure to the fully connected LSTM network, the networks can obtain the timing relationship and use the convolutional layer to extract spatial features. Consequently, Conv-LSTM can extract temporal characteristics and spatial characteristics. Du et al. [39] designed a hybrid deep learning framework consisting of one-dimensional convolutional neural networks (ID CNN) and gated recurrent units (GRU). CNN is used to capture the local trend features and the GRU is used to capture the long temporal dependencies. Hou et al. [40] proposed a deep learning model based on LSTM and CNN to predict travel time in a road network. Xiao et al. proposed an attention-based Conv-LSTM network for time series prediction. The attention mechanism can effectively eliminate irrelevant information, select the relevant exogenous sequence, give it higher weight, and increase the past value of the target sequence to further eliminate irrelevant information [41]. The Conv-LSTM networks also achieved good performance in traffic speed prediction [38, 42].

The above-mentioned deep learning methods show that fully mining traffic data features can improve prediction accuracy. The expression and extraction of traffic data's features will still be a significant issue for traffic prediction. Besides, attention mechanism is an effective way to enhance the performance of neural networks. Inspired by these recent studies, a novel model integrating ConvLSTM, Bi-LSTM networks and attention mechanism is proposed in this study. The proposed model considers both spatiotemporal and periodic features and improves the performance for large-scale traffic speed prediction.

3 Methodology

In this section, we first introduce the problem formulation and then provide the detailed information about the AB- ConvLSTM model. Essentially, the proposed model consists of a Conv-LSTM module, an attention mechanism module, and two Bi-LSTM modules. We will describe these modules in considerable detail.

3.1 Problem formulation

In this study, the traffic speed we predicted refers to the space mean speed, which is defined as the average speed of all the vehicles occupying a given section of a highway over some specified time period. To achieve accurate prediction, it is necessary to mine the features of traffic data. This study uses a spatial–temporal traffic speed matrix to denote the traffic condition of the entire road network. On one hand, the time relevance of traffic data changes over time needs to be considered. Multiple historical periods are used to predict traffic in a specific period in the future. On the other hand, the traffic condition of the road segment itself is affected by the neighbouring road segments, which reflect the spatial correlation. Road segments are connected in the urban road network. The upstream vehicles will arrive downstream, and congestion in the downstream segments will also spread upstream. To extract the spatial relationship between road segments, they are ordered according to the traffic flow direction. Finally, a spatial–temporal matrix is constructed as Eq. (1). Before the data are input into the prediction model, it needs to be converted into the channel data format. Next, the data are put into Conv-LSTM networks for extracting spatiotemporal features. The whole process is shown in Fig. 1.

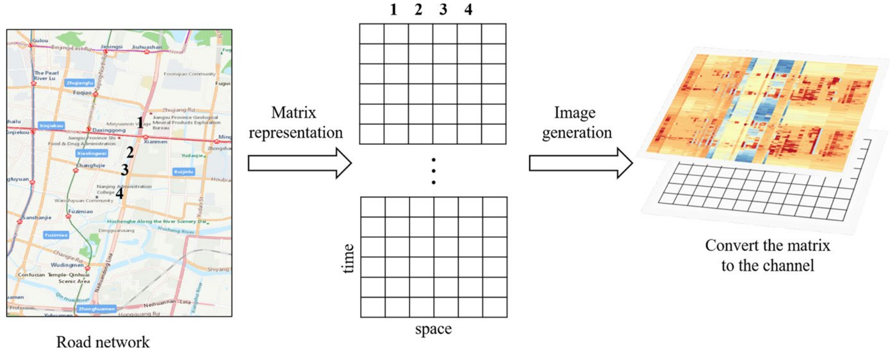


Fig. 1 Convert the matrix to the channel on road network

$$X_t^s = \begin{bmatrix} V_{t-m}^s \\ \vdots \\ V_{t-1}^s \\ V_t^s \end{bmatrix} = \begin{bmatrix} v_{t-m}^1 & v_{t-m}^2 & \cdots & v_{t-m}^n \\ \vdots & \vdots & \ddots & \vdots \\ v_{t-1}^1 & v_{t-1}^2 & \cdots & v_{t-1}^n \\ v_t^1 & v_t^2 & \cdots & v_t^n \end{bmatrix}, \tag{1}$$

where X_t^s denotes the state of the road networks, including $m+1$ time periods and n road segments. $V_t^s = [v_t^1, v_t^2, \dots, v_t^n]$ denotes traffic speed of all road segments in the road networks at time t.

In addition, traffic condition also presents periodic features. For example, the traffic speed on weekdays has the same variation tendency, and the traffic speed on the same day for 2 consecutive weeks has high similarity. As we can see from Fig. 2, adjacent road sections have similar traffic condition, and the traffic condition of current day has a high level of similarity to last day and last week.

This study takes both daily and weekly periodicity features of the traffic speed into consideration. The daily periodicity of the data can be extracted by considering the previous n time intervals and subsequent n time intervals of the same moment as time t in the last day. The weekly periodicity of the data can be extracted by considering the previous n time intervals and subsequent n time intervals of the same moment as time t in the last week. The daily and weekly data matrix can be represented as Eqs. (2) and (3). They will be put into the Bi-LSTM networks to extract periodic features.

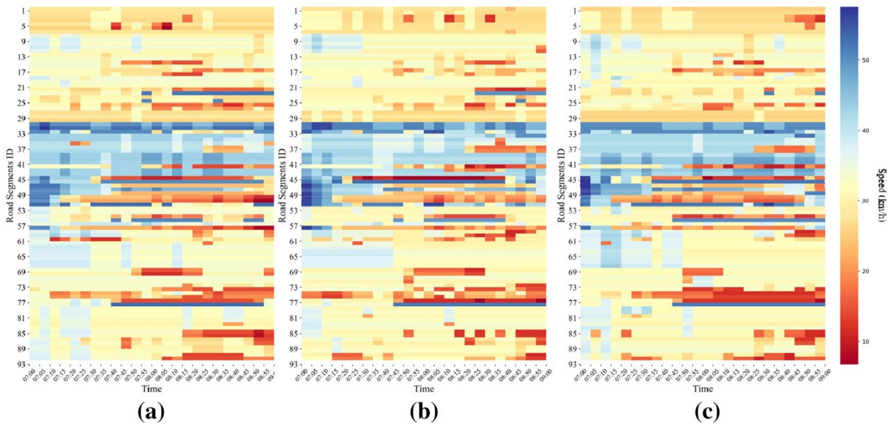


Fig. 2 Traffic condition of road networks, **a** denotes the traffic condition in last week, **b** denotes the traffic condition in last day, **c** denotes the traffic condition of current day

$$X_t^d = \begin{bmatrix} V^s_{t^d-m} \\ \vdots \\ V^s_{t^d-1} \\ V^s_{t^d} \\ \vdots \\ V^s_{t^d+m} \end{bmatrix} = \begin{bmatrix} v^1_{t^d-m} & v^2_{t^d-m} & \cdots & v^n_{t^d-m} \\ \vdots & \vdots & \iddots & \vdots \\ v^1_{t^d-1} & v^2_{t^d-1} & \cdots & v^n_{t^d-1} \\ v^1_{t^d} & v^2_{t^d} & \cdots & v^n_{t^d} \\ \vdots & \vdots & \iddots & \vdots \\ v^1_{t^d+m} & v^2_{t^d+m} & \cdots & v^n_{t^d+m} \end{bmatrix}, \tag{2}$$

$$X_t^w = \begin{bmatrix} V^s_{t^w-m} \\ \vdots \\ V^s_{t^w-1} \\ V^s_{t^w} \\ \vdots \\ V^w_{t^w+m} \end{bmatrix} = \begin{bmatrix} v^1_{t^w-m} & v^2_{t^w-m} & \cdots & v^n_{t^w-m} \\ \vdots & \vdots & \iddots & \vdots \\ v^1_{t^w-1} & v^2_{t^w-1} & \cdots & v^n_{t^w-1} \\ v^1_{t^w} & v^2_{t^w} & \cdots & v^n_{t^w} \\ \vdots & \vdots & \iddots & \vdots \\ v^1_{t^w+m} & v^2_{t^w+m} & \cdots & v^n_{t^w+m} \end{bmatrix}, \tag{3}$$

where t^d denotes the same moment as time t in the last day. t^w denotes the same moment as time t in the last week [18]. $V^s_{t^d} = [v^1_{t^d}, v^2_{t^d}, \ldots, v^n_{t^d}]$ denotes the traffic speed of the n road segments at the same moment as time t in the last day. $V^s_{t^w} = [v^1_{t^w}, v^2_{t^w}, \ldots, v^n_{t^w}]$ denotes the traffic speed of the n road segments at the same moment as time t in the last week.

The problem of traffic speed prediction can be formulated as follows. Based on the current traffic data and historical traffic data X^s_t, X^d_t, and X^w_t, the traffic of the road networks in the future can be predicted. The future traffic can be expressed as $V^p_{t+\Delta} = [v^1_{t+\Delta}, v^2_{t+\Delta}, \ldots, v^n_{t+\Delta}]$, the predicting step Δ can be 5 min, 15 min, 30 min, and 60 min. An example of constructing the input time series is illustrated in Fig. 3.

3.2 Overview of the proposed model

In this study, a deep learning framework called AB-ConvLSTM is proposed for traffic speed prediction. As Fig. 4 shows, the deep learning framework consist of a Conv-LSTM network, an attention mechanism module, and two Bi-LSTM networks. The Conv-LSTM is used to extract the spatiotemporal features of the traffic data. In addition, the attention mechanism is introduced in the model to enhance the performance of the networks. Attention-based Conv-LSTM networks can automatically extract the importance of different time series in the time dimension and assigned

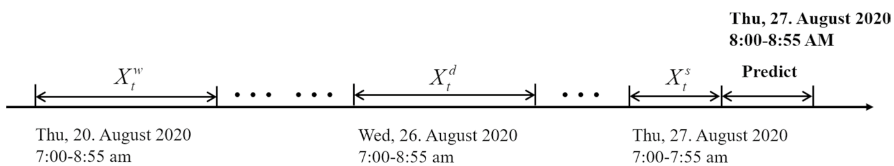


Fig. 3 An example of constructing the input time series (the size of predicting window is 1 h)

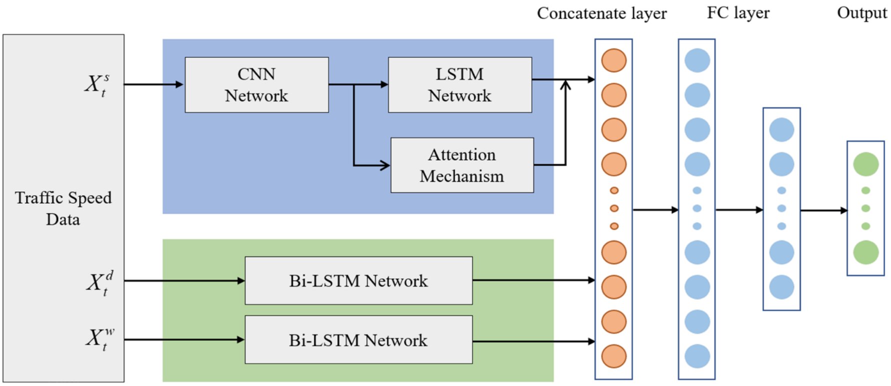


Fig. 4 Structure of the AB-ConvLSTM networks

corresponding weights. The Bi-LSTM networks are used to extract the periodic features of the traffic data. Then the outputs of Conv-LSTM networks and Bi-LSTM networks are connected by concatenate layer. Finally, the concatenate layer is connected to two fully connected layers (FC layer) to output prediction results.

3.3 Conv-LSTM

The structure of the Conv-LSTM networks is shown in Fig. 5. The input data are the traffic speed matrix, which can be expressed as $X_t^s = (V_{t-m}^s, V_{t-m+1}^s, \ldots, V_t^s)$ shown in Eq. (1). Firstly, the data matrix needs to be converted into "images" with channels.

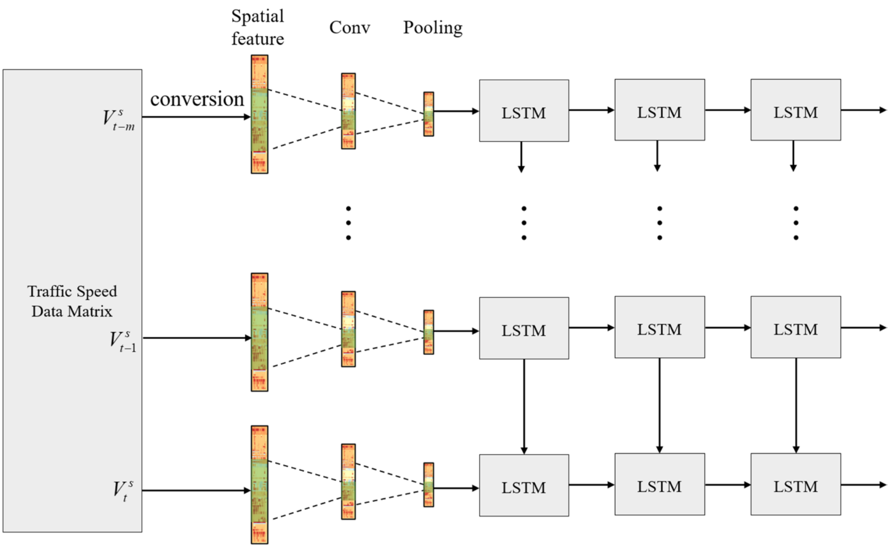


Fig. 5 Structure of Conv-LSTM networks

Then, the data are put into the convolutional layer and pooling layer. One-dimensional convolution filter is used to process each time step data V_t^s. Next, the maximum pooling is used to reduce the scale of the networks, which can make the model have higher distortion tolerance. The convolutional and pooling layers processing can well extract the spatial features. The output of Conv-LSTM can be expressed as:

$$O_t^s = \mathrm{pool}(\sigma(W_s * X_t^s + b_s)), \tag{4}$$

where W_s is the weights of filters, X_t^s is the input of the model, b_s is bias, σ is the activation function, pool is the pooling procedure, and O_t^s is the output of the model.

After the convolutional and pooling layers processing, suppose the final output is $O_t^s = [O_{t-m}^s, \dots, O_{t-1}^s, O_t^s]$. Each element in the vector contains the spatial correlation between road segments, and it will be put into LSTM networks. Traditional multiple layers of LSTM can be connected in time to form a more complex structure. These models have been applied to solve many practical sequence modelling problems. Such a model is very good at capturing temporal features but not good at capturing spatial features. The input of LSTM is 1D data, which cannot reflect the state information of the space. Therefore, LSTM does not perform well in processing spatiotemporal sequence data. LSTM uses the full connection in input-to-state and state-to-state processes, which means that the full information is directly multiplied to equal a value. It is impossible to capture spatiotemporal information. Compared with the original LSTM, Conv-LSTM replaces matrix multiplication ($\circ$) with convolution operations ($*$) in the input-to-state and state-to-state processes. There is a convolution kernel that considers the spatial features of the traffic data. The structure of the LSTM networks is shown in Fig. 6. The procedure of the LSTM layers can be explained as follows:

$$i_t = \sigma(W_{gi} * G_t^s + W_{hi} * H_{t-1}^s + W_{ci} \circ C_{t-1} + b_i), \tag{5}$$

$$f_t = \sigma(W_{gf} * G_t^s + W_{hf} * H_{t-1}^s + W_{cf} \circ C_{t-1} + b_f), \tag{6}$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tanh(W_{gc} * G_t^s + W_{hc} * H_{t-1}^s + b_c), \tag{7}$$

$$o_t = \sigma(W_{go} * G_t^s + W_{ho} * H_{t-1}^s + W_{co} \circ C_t + b_o), \tag{8}$$

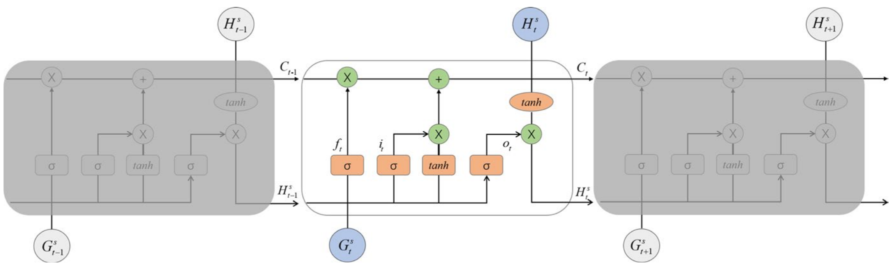


Fig. 6 The LSTM neural network structure

$$H_t^s = o_t \circ \tanh(C_t), \tag{9}$$

where i_t is the input gate, f_t is the forget gate, C_t is the update cell, o_t is the output gate, σ is the activation function, G_t^s is the input of the LSTM layer at time step t, W_{gi}, W_{gf}, W_{go}, W_{hi}, W_{hi}, W_{hi}, W_{ci}, W_{cf}, W_{co} are the weights of input gate, forget gate, and output gate, b_i, b_f, b_c, b_o are bias, and H_t^s is the output of the LSTM.

3.4 Attention mechanism

In this study, m time steps data matrix is designed to predict $t+\Delta$ time period traffic. However, the data in the different time period have various influences on the prediction time period. In general, the closer the time period to the prediction time period ($t+\Delta$ time period), the more significant impact of historical data on the target time period, and the higher the weight of the data should be assigned to [43]. To automatically capture the impact of different time periods on the prediction time period, this study has introduced an attention mechanism in Conv-LSTM. Figure 7 shows the structure of the model.

In the attention mechanism, each time period gets a score, denoted as $e = (e_1, e_2, \ldots, e_{m+1})^T$. The score is directly determined by the correlation between G_t^s and H_t^s. The higher the score, the more critical the impact of the current time series on the final prediction result. Bahdanau's additive style [44] is used to compute the score, which is shown in Eq. (10). The score is processed by weighted average as the weight of different time steps. The attention weights α_k can be obtained as Eq. (11). After getting the attention weight of each time step, the new output H_t^o can be calculated as Eq. (12).

$$e_t = v_s^T \tanh(W_s G_t^s + U_s H_t^s), \tag{10}$$

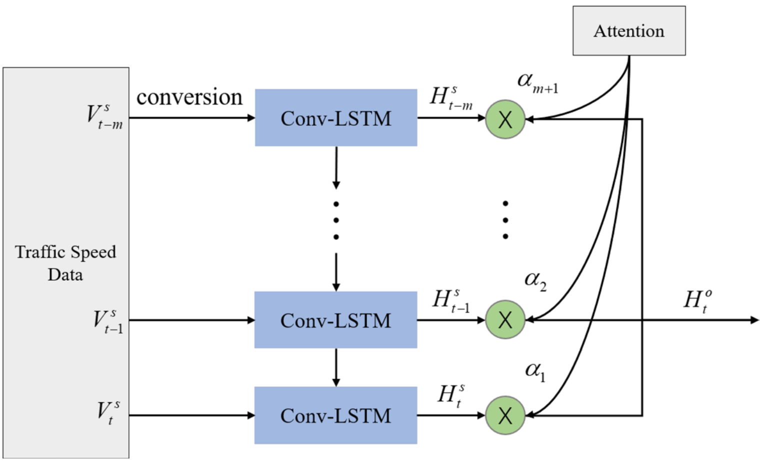


Fig. 7 Conv-LSTM networks with attention mechanism

$$\alpha_k = \frac{\exp(e_k)}{\sum_{k=1}^{m+1} \exp(e_k)}, \tag{11}$$

$$H_t^o = \sum_{k=1}^{m+1} \alpha_k H_{t-m+1}^s, \tag{12}$$

where e_t are the scores of different time period, v_s, W_s, U_s are learnable parameters, G_t^s is the input of the LSTM layer at time step t, H_t^s is hidden output of the Conv-LSTM, α_k is the attention weights, and H_t^o is the output of attention-based Conv-LSTM.

3.5 Bi-directional LSTM

The Conv-LSTM networks can well extract the spatiotemporal features of the traffic data. In addition, we consider the periodic features of the road networks traffic. In urban road networks, traffic data have prominent periodic features due to commuting activities. In general, the traffic condition at the current moment has a high level of similarity with the previous days' traffic condition, and each week's traffic condition also has a periodic relationship with the traffic condition of the previous week. In addition, the traffic speed in a given time period depends on the traffic speed of its previous time, but also in turn affects its speed of its upcoming time. This study uses the Bi-LSTM network to capture these features.

The structure of Bi-LSTM networks consists of two unidirectional LSTMs stacked up and down, one LSTM is forward propagation and another is backward propagation. The inputs of Bi-LSTM networks are X_t^d and X_t^w shown in Eqs. (2) and (3). The input is historical data (without current time data), including time series before and after the same moment as current time t in the last day and the last week, which are used as a supplement to the current moment data. In this way, more features can be extracted from both directions, which improves the prediction performance.

The structure of Bi-LSTM networks is shown in Fig. 8, O_f^d and O_b^d are the outputs of forward LSTM and backward LSTM respectively, which contains daily periodic information. O_f^w and O_b^w are the outputs of forward LSTM and backward LSTM respectively, which contains weekly periodic information.

3.6 Feature concatenation and prediction

After the processing by attention-based Conv-LSTM networks and two Bi-LSTM networks, we can get spatiotemporal features H_t^o, daily periodic features O_f^d and O_b^d, and weekly periodic features O_f^w and O_b^w. Next, they are merged in concatenate layer. Finally, we use two fully connected layers to output prediction results $V_{t+\Delta}^p$.

$$V_{t+\Delta}^p = W_{fc} H_t^m + b_{fc}, \tag{13}$$

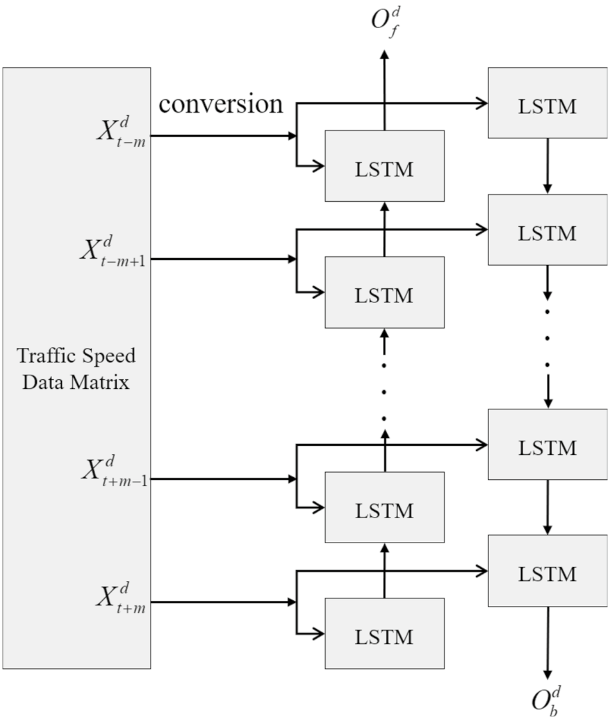


Fig. 8 The structure of Bi-LSTM networks

where H_t^m is the concatenated features, W_{fc} is the weights of fully connected layers, and b_{fc} is the bias of fully connected layers.

In the training phase, mean squared error (MSE) is used as the loss function for the proposed model. MSE can be expressed as Eq. (14). Then, this study used Adam algorithm to optimize the deep learning framework, which has an adaptive learning rate.

$$\text{MSE} = (1/n)\sum_{i=1}^{n}(v_i - v_i^*)^2. \tag{14}$$

4 Case study

4.1 Data preparation

In this study, the real-world traffic data at Nanjing are used to train the model. The data are obtained from Amap Open Platform [45], including the traffic speed, traffic condition, latitude, and longitude of all road segments in the road networks. Traffic speed here is the average speed of all vehicles occupying a segment over 5 min.

Table 1 shows the explanation of data. The data table contains nine parameters and are aggregated every 5 min from August 10, 2020 to September 4, 2020.

The missing values are filled by the linear interpolation and the MinMaxScaler method is used to normalize the data, which shown in Eq. (15). 60% of data is used for training the model, 30% for testing the model and 10% for validation. Figure 9 shows the scope of predicted road networks, which consists of 93 road segments.

$$X^* = \frac{X - X_{\min}}{X_{\max} - X_{\min}}, \tag{15}$$

where X^* denotes the traffic speed normalized to [0,1], $X_{\min}$ denotes the minimum value, and $X_{\max}$ denotes the maximum value.

4.2 Training parameters

The details of the AB-ConvLSTM are listed in Table 2. There are three input layers, three Conv-LSTM layers, one Attention layer, two Bi-LSTM networks including four Bi-LSTM layers, one concatenate layer, two FC layers, and four output layers. Input_1 has five dimensions (64, 8, 1, 93, 1), where the first number 64 is the batch size, the second number denotes the input time span (40 min), the third number represents one time series, the fourth number represents the total number of road segments in road network, and the fifth number denotes that the input data have one channel. In Conv-LSTM layer, each convolutional layer has 16 filters and the kernel size is (1,3). There are two Bi-LSTM layers in one Bi-LSTM networks and the hidden size is 64 and 16, respectively. The output of Attention-based ConvLSTM and two Bi-LSTM networks are merged in concatenate layer. Then, two fully connected layers are used to output the results with the dimension (64, 93).

Table 1 Explanation of original data

Parameters	Data type	Example	Remarks
Date	String	15 September 2020	DD/MM/YYYY
Time	String	08:30:00	UTC + 8
Road name	String	Yunqi Road	Roads in Nanjing, China
Road segment ID	Int	4120	A road consists of multiple road segments
Longitude	Float	112.8927	East longitude
Latitude	Float	28.1409	North latitude
Angle	Int	45	Vehicle driving direction
Speed	Float	36.5	Traffic speed of road segment in 5 min (km/h)
Traffic condition	Int	2	1: Smooth 2: Basically smooth 3: Lightly congested 4: Moderately congested 5: Severely congested

Fig. 9 Locations of the road segments in our experiment

The rectified linear activationunit (ReLU) is adopted as the activation function. Adam algorithm is used to optimize the deep learning framework with an initial learning rate of lr = 0.001. Adam is a very popular algorithm in the field of deep learning. Kingma et al. demonstrate that Adam can efficiently solve practical deep learning problems using large models and datasets. [46]. The experiments are conducted on a workstation with an Intel (R) Core (TM) i7-10875H CPU and one NVIDIA GeForce RTX 2060 Graphics Card.

4.3 Indicators of performance

General accuracy indicators are the mean absolute error (MAE), the mean absolute percentage error (MAPE), and the root mean square error (RMSE). This study uses MAE, MAPE, and RMSE to evaluate the performance of the prediction model. Three indicators are shown in the following Equations.

$$\mathrm{MAE} = (1/n)\sum_{i=1}^{n}\left|v_i - v_i^*\right|, \tag{16}$$

$$\mathrm{MAPE} = (1/n)\sum_{i=1}^{n}\left|(v_i - v_i^*)/v_i\right| \times 100\%, \tag{17}$$

Table 2 Hyperparameters of AB-ConvLSTM

Layer	Structure	Parameters	Dimensions
Input Layer_1	Input_1		(64, 8, 1, 93, 1)
Conv-LSTM layer	Conv-LSTM2D	Filters = 16, kernel_size = (1,3)	(64, 8, 1, 93, 16)
	Dropout	0.2	(64, 8, 1, 93, 16)
	Batch normalization		(64, 8, 1, 93, 16)
Conv-LSTM layer	Conv-LSTM2D	Filters = 16, kernel_size = (1,3)	(64, 8, 1, 93, 16)
	Dropout	0.2	(64, 8, 1, 93, 16)
	Batch normalization		(64, 8, 1, 93, 16)
Conv-LSTM layer	Conv-LSTM2D	Filters = 16, kernel_size = (1,3)	(64, 8, 1, 93, 16)
	Dropout	0.2	(64, 8, 1, 93, 16)
	Batch normalization		(64, 1, 93, 16)
Attention layer	Self-attention		(64, 1, 93, 16)
Output Layer_1	Conv2D	Filters = 1, kernel_size = (1,3)	(64, 1, 93, 1)
	Batch normalization		(64, 1, 93, 1)
	Flatten		(64, 93)
Input Layer_2	Input_2		(64, 16, 93)
Bi-LSTM layer	Bidirectional LSTM	Units = 64	(64, 16, 128)
Bi-LSTM layer	Bidirectional LSTM	Units = 16	(64, 16, 32)
Output Layer_2	Flatten		(64, 512)
Input Layer_3	Input_3		(64, 16, 93)
Bi-LSTM layer	Bidirectional LSTM	Units = 64	(64, 16, 128)
Bi-LSTM layer	Bidirectional LSTM	Units = 16	(64, 16, 32)
Output Layer_3	Flatten		(64, 512)
Concatenate layer	Concatenate (output Layer_1, 2, 3)		(64, 1117)
FC Layer	Dense	Units = 512	(64, 512)
FC Layer	Dense	Units = 256	(64, 256)
Output layer	Dense	Units = 93	(64, 93)

$$\mathrm{RMSE} = \sqrt{(1/n)\sum_{i=1}^{n}(v_i - v_i^*)^2}, \tag{18}$$

where v_i is the true value of the traffic speed, and v_i^* is the predicted value of the traffic speed.

4.4 Cross validation

Cross-validation is used to evaluate the predictive performance of the model, especially the performance of the trained model on new data, which can reduce overfitting to a certain extent. What's more, it can help the model obtain as much useful information as possible from limited data. As the complexity of the deep learning model increases, the model is prone to overfitting. To enhance the generalization capability of the model, we use cross-validation to form hyperparameters and to some extent reduce the occurrence of overfitting.

In this study, we use TimeSeriesSplit method to train the model, which is a Time Series cross-validator. This cross-validation object is a variation of KFold. In the k-th split, it returns first k folds as train set and the $(k+1)$th fold as test set. It can provide train/test indices to split time series data samples that are observed at fixed time intervals, in train/test sets. In each split, test indices must be higher than before, and thus shuffling in cross validator is inappropriate. Figure 10 shows the partition of train set and test set of TimeSeriesSplit. The dataset is divided into ten splits, we conducted five different experiments and averaged the results to evaluate the model.

4.5 Performance evaluation

We compare our model with the following models including the traditional time-series prediction methods, machine learning baselines and deep learning baselines.

- ARIMA [3]. Auto-regressive integrated moving average model is a well-known time series analysis and forecasting method.
- RF [47]. Random forest is a joint prediction model composed of multiple decision trees, which can be used for classification and prediction.
- SVR [7], Support vector regression is an important branch of support vector machine (SVM), SVM is used for classification, SVR is used for regression.
- LSTM [21]. Long short-term memory network is a special type of RNN model, which is better at handling time series with long time lags than RNN.
- SAE [29]. Stacked autoencoder is a kind of deep learning architecture and uses the layerwise greedy algorithm to train the model.
- CNN [15]. Convolutional neural network contains convolutional layers and pooling layers, which can extract more spatial features of traffic data.
- AT-BLSTM [35]. An attention-based Bi-LSTM network (AT-BLSTM) model, which consists of Bi-LSTM layer and attention layer.

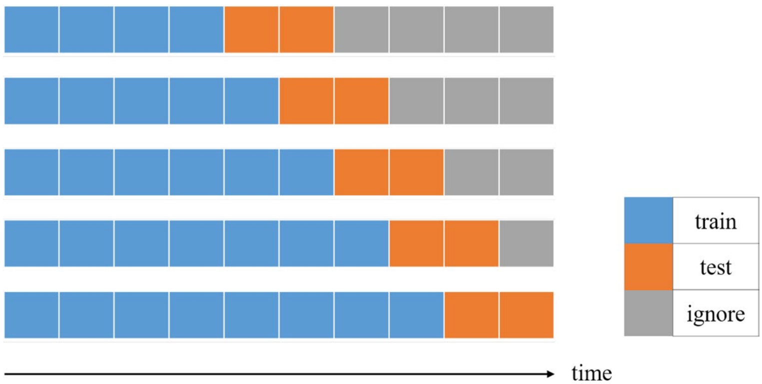


Fig. 10 Partition of train set and test set of TimeSeriesSplit

- ConvLSTM [40]. A deep learning model with 1D convolution units and two LSTM layers.

4.5.1 Performance of the proposed model

We first evaluate the ability of Bi-LSTM and attention mechanism. As aforementioned, attention mechanism can automatically capture the impact of different time periods on the prediction time period and assign corresponding weights. Bi-LSTM networks can extract daily and weekly periodic features. Figure 11 shows the MAE of ConvLSTM and improved ConvLSTM models under different predicting time windows. It can be obviously observed that both Bi-LSTM and attention mechanism help improve the performance of ConvLSTM. Bi-LSTM is more effective than attention mechanism, which means that periodic features play an important role in traffic speed prediction.

As shown in Fig. 12, the training loss of AB-ConvLSTM decreases faster than ConvLSTM, which indicates that AB-ConvLSTM have a high convergence speed. AB-ConvLSTM cost 12.3 min for 50 epochs' training and ConvLSTM cost 9.3 min. Although the introduction of Bi-LSTM and attention mechanism increases the model's complexity, the computational expense does not increase much, and the error of the model is smaller. In conclusion, AB-ConvLSTM provides a significant performance improvement and can realize road networks traffic speed prediction with high efficiency.

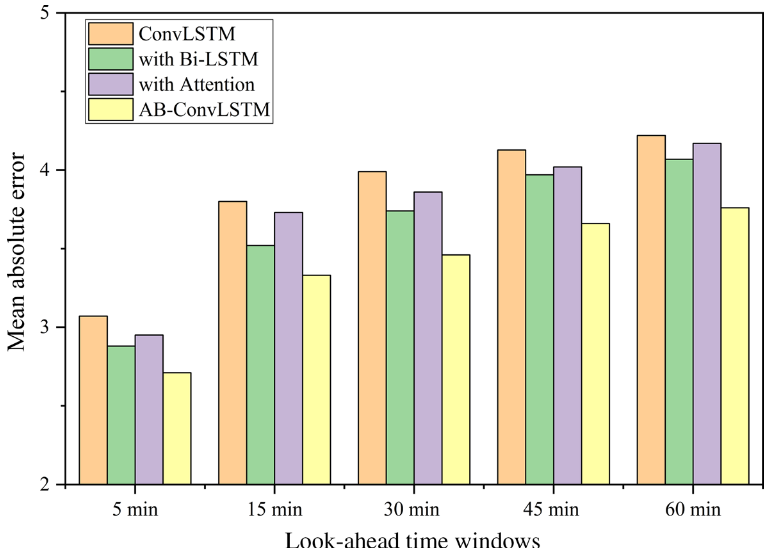
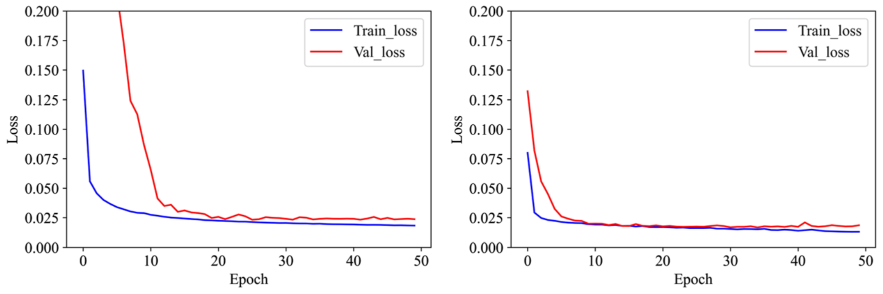


Fig. 11 MAE of ConvLSTM and improved ConvLSTM models



(a) ConvLSTM (b) AB-ConvLSTM

Fig. 12 Training loss of ConvLSTM and AB-ConvLSTM models

4.5.2 Performance comparison with different methods

In addition, our models are compared with eight baselines, and the predicting results are shown in Table 3. It can be seen clearly that ConvLSTM and the improved ConvLSTM have better performance than other models. As the bold values in the table show, AB-ConvLSTM achieves the lowest MAE, the lowest RMSE, and the lowest MAPE among all the methods. The performance of ARIMA, RF and SVM are worst because they can only extract features in the time dimension. CNN, LSTM and ST-LSTM perform better than SAE, but they cannot extract spatiotemporal features. The performance of Conv-LSTM is better than LSTM and CNN, which indicates that ConvLSTM can fully extract the temporal and spatial features of road network data.

Table 3 Prediction performances of different models with urban road networks data

Model	15 min			30 min			60 min		
	MAE	MAPE/%	RMSE	MAE	MAPE/%	RMSE	MAE	MAPE/%	RMSE
ARIMA	8.40	22.99	10.45	8.72	23.81	10.69	9.47	25.04	11.06
RF	7.34	21.95	9.41	7.69	22.71	9.69	8.41	24.02	10.03
SVR	6.30	20.94	8.36	6.68	21.67	8.64	7.38	23.02	8.96
LSTM	3.99	18.10	6.09	4.13	18.16	6.06	4.61	20.87	6.97
SAE	5.13	19.91	7.23	5.55	20.56	7.57	6.24	21.91	7.93
CNN	4.45	18.89	6.56	4.65	19.82	6.83	5.22	21.3	7.58
AT-BLSTM	4.02	18.00	5.98	4.03	18.05	5.96	4.50	20.77	6.87
ConvLSTM	4.17	18.36	6.32	4.31	18.51	6.36	4.81	21.64	7.31
Our model									
With Bi-LSTM	3.52	16.18	5.49	3.74	16.87	5.66	4.07	17.55	6.15
With attnetion	3.73	16.73	6.04	3.86	17.23	5.93	4.17	17.87	6.45
AB-ConvLSTM	**3.43**	**15.34**	**5.36**	**3.56**	**16.36**	**5.61**	**3.86**	**17.12**	**6.03**

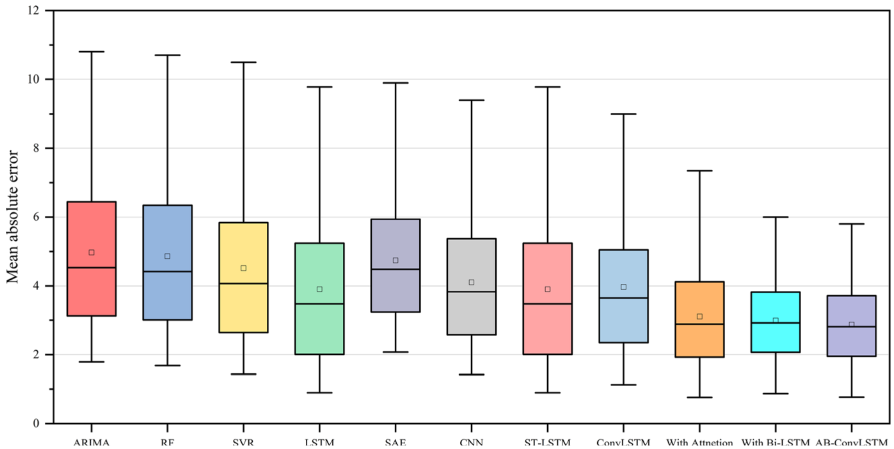


Fig. 13 Comparison of forecast errors for each method

Figure 13 shows the MAE of each method under 5-min prediction. Our model has a more concentrated error distribution and displays a more stable prediction ability than other methods. In conclusion, the proposed AB-ConvLSTM model can fully extract the road network data's spatiotemporal features and periodic features and consistently outperforms the competing models.

4.5.3 Run time analysis

Table 4 shows the specific training time and testing time of different models. The results show that the traditional machine learning methods including ARIMA, RF, and SVR have faster training time due to their simple calculation process and fewer

Table 4 Runtime of different models

Model	Computation time	
	Training time (s)	Testing time (s)
ARIMA	19.94	3.67
RF	46.13	11.56
SVR	23.15	9.27
LSTM	217.39	0.76
SAE	288.12	0.82
CNN	339.63	0.73
AT-BLSTM	301.29	0.88
ConvLSTM	418.84	1.53
Our model		
With Bi-LSTM	443.54	1.87
With Attnetion	484.34	2.54
AB-ConvLSTM	557.68	2.91

parameters. These traditional machine learning methods can only establish models for one road segment, and cannot produce prediction results of the entire road network at once. They are not suitable for large-scale traffic prediction. It is worth noting that deep learning methods have faster testing time than traditional machine learning methods. ConvLSTM and its variants increase computational complexity, resulting in longer training time. However, AB-ConvLSTM outperforms other baselines in traffic speed prediction within an acceptable training time.

Several valuable findings from the above results and discussion can be summarized as follows:

- The introduction of attention mechanism and Bi-LSTM is beneficial to improving the model's performance. The attention mechanism module can automatically capture each historical sequence's importance to the final prediction. Bi-LSTM networks can effectively extract periodic features of the data.
- The performance of the proposed model consistently outperforms the competing models. The AB-ConvLSTM model can fully extract the spatiotemporal features and periodic features of the road network data.
- AB-ConvLSTM has a high convergence speed and can realize road networks traffic speed prediction with high efficiency within an acceptable training time.

4.6 Estimation of traffic condition

According to the relationship between the Traffic speed and the traffic condition in Table 5, the traffic condition of the road segment can be derived from the predicted Traffic speed by the AB-ConvLSTM model. The traffic condition of the road has five values so that the speed is divided into five sections. The prediction results show that the accuracy of the traffic condition can reach almost 100%. Figure 14 shows the predicted traffic condition of different road segments in the urban road networks at 8:00 am. The results are the same as the real traffic condition.

The above analysis shows that the traffic condition of the urban road network can be predicted by AB-ConvLSTM with high accuracy. Large-scale traffic prediction can provide reliable traffic condition information for travellers and traffic managers. Travellers can adjust their travel plans in time to avoid entering traffic jams. It is also conducive for traffic managers to conduct overall management of congested

Table 5 Classification of different road conditions

Road level	Traffic condition				
	Smooth (1)	Basically smooth (2)	Lightly congested (3)	Moderately congested (4)	Severely congested (5)
Expressway	$S>55$	$40<S\leq 55$	$30<S\leq 40$	$20<S\leq 30$	$S\leq 20$
Trunk road	$S>40$	$30<S\leq 40$	$20<S\leq 30$	$15<S\leq 20$	$S\leq 15$
Branch road	$S>30$	$20<S\leq 30$	$15<S\leq 20$	$10<S\leq 15$	$S\leq 10$

Fig. 14 Traffic condition of different road segments in the urban road network (five numbers denote different traffic condition from smooth to severely congested)

road segments and surrounding-related road segments, which can reduce congestion and waste of resources. In conclusion, AB-ConvLSTM is promising in large-scale traffic prediction.

5 Conclusions

A considerable number of researches have been carried out on traffic speed prediction. However, limited studies have addressed spatiotemporal and periodic features of large-scale traffic data. This study aims to fully extract the traffic data's features and achieve the goal of large-scale traffic speed prediction with high accuracy.

In this paper, a hybrid deep learning method based on ConvLSTM, attention mechanism and Bi-LSTM, called AB-ConvLSTM, is proposed for large-scale traffic speed prediction. ConvLSTM networks can fully extract the spatiotemporal features of the traffic speed data. Besides, attention mechanism can enhance the performance of ConvLSTM by automatically capture the impact of different time periods on the final prediction. What's more, Bi-LSTM networks can effectively extract periodic features of the data and capture variation tendencies from forward and backward traffic data. The extensive experimental results show that the proposed model can realize large-scale traffic speed prediction with high efficiency and consistently outperforms the competing models. In addition, we have estimated the traffic condition of urban road network with high accuracy based on the predicted traffic speed and its corresponding relationship with traffic condition.

The periodicity of traffic speed is affected by holidays and weather. In these particular scenarios, the accuracy of the proposed model cannot be guaranteed. In future studies, we will focus on the collection and processing of data from special scenarios, and feed them into the model to fully extract periodic features. Besides, we will expand the road network size and carry out citywide traffic speed prediction.

Acknowledgements This work was supported by Jiangsu Provincial Key Research and Development Program (No. BE20187544).

References

1. Mackenzie J, Roddick JF, Zito R (2019) An evaluation of HTM and LSTM for short-term arterial traffic flow prediction. IEEE Trans Intell Transp Syst 20(5):1847–1857. https://doi.org/10.1109/tits.2018.2843349
2. Vlahogianni EI, Karlaftis MG, Golias JC (2014) Short-term traffic forecasting: where we are and where we're going. Transp Res Pt C Emerg Technol 43:3–19. https://doi.org/10.1016/j.trc.2014.01.005
3. Williams BM, Hoel LA (2003) Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results. J Transp Eng A Syst 129(6):664–672. https://doi.org/10.1061/(Asce)0733-947x(2003)129:6(664)
4. Fusco G, Colombaroni C, Isaenko N (2016) Short-term speed predictions exploiting big data on large urban road networks. Transp Res Pt C Emerg Technol 73:183–201. https://doi.org/10.1016/j.trc.2016.10.019
5. Ojeda LL, Kibangou AY, Wit CD (2013) Adaptive Kalman filtering for multi-step ahead traffic flow prediction. In: American Control Conference (ACC), pp 4724–4729
6. Hu X, Liu T, Li Y, Cui Z (2021) An improved Euclidean distance weighted K-nearest neighbor model for traffic state forecasting. In: Transportation research board 100th annual meeting. Transportation Research Board, Washington DC, p 18
7. Tang J, Chen X, Hu Z, Zong F, Han C, Li L (2019) Traffic flow prediction based on combination of support vector machine and data denoising schemes. Phys A. https://doi.org/10.1016/j.physa.2019.03.007
8. Sun J, Sun J, Chen P (2014) Use of support vector machine models for real-time prediction of crash risk on urban expressways. Transport Res Rec 2432(1):91–98. https://doi.org/10.3141/2432-11
9. Wang J, Shi Q (2013) Short-term traffic speed forecasting hybrid model based on Chaos-wavelet analysis-support vector machine theory. Transp Res Pt C Emerg Technol 27:219–232. https://doi.org/10.1016/j.trc.2012.08.004
10. Ren Q (2007) Research on time-space road network traffic congestion prediction and guide decision method. Southwest Jiaotong University
11. Castillo E, Jiménez P, Menéndez JM, Nogal M (2012) A Bayesian method for estimating traffic flows based on plate scanning. Transportation 40(1):173–201. https://doi.org/10.1007/s11116-012-9443-4
12. Yu B, Song X, Guan F, Yang Z, Yao B (2016) k-Nearest neighbor model for multiple-time-step prediction of short-term traffic condition. J Transp Eng A Syst 142(6):04016018. https://doi.org/10.1061/(asce)te.1943-5436.0000816
13. Cai P, Wang Y, Lu G, Chen P, Ding C, Sun J (2016) A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting. Transp Res Pt C Emerg Technol 62:21–34. https://doi.org/10.1016/j.trc.2015.11.002
14. Zheng Z, Su D (2014) Short-term traffic volume forecasting: a k-nearest neighbor approach enhanced by constrained linearly sewing principle component algorithm. Transp Res Pt C Emerg Technol 43:143–157. https://doi.org/10.1016/j.trc.2014.02.009
15. Ma X, Zhuang D, He Z, Ma J, Wang Y (2017) Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction. Sensors 17(4):818. https://doi.org/10.3390/s17040818

16. Zheng F, Zuylen HV (2013) Urban link travel time estimation based on sparse probe vehicle data. Transp Res Pt C Emerg Technol 31:145–157. https://doi.org/10.1016/j.trc.2012.04.007
17. Ye Q, Szeto WY, Wong SC (2012) Short-term traffic speed forecasting based on data recorded at irregular intervals. IEEE Trans Intell Transp Syst 13(4):1727–1737. https://doi.org/10.1109/Tits.2012.2203122
18. Zheng H, Lin F, Feng X, Chen Y (2020) A hybrid deep learning model with attention-based Conv-LSTM networks for short-term traffic flow prediction. IEEE Trans Intell Transp Syst 22(11):6910–6920. https://doi.org/10.1109/TITS.2020.2997352
19. Zhao L, Wang Q, Jin B, Ye C (2020) Short-term traffic flow intensity prediction based on CHS-LSTM. Arab J Sci Eng 45(12):10845–10857
20. Zhao Z, Chen W, Wu X, Chen PC, Liu J (2017) LSTM network: a deep learning approach for short-term traffic forecast. IET Intell Transp Syst 11(2):68–75. https://doi.org/10.1049/iet-its.2016.0208
21. Ma X, Tao Z, Wang Y, Yu H, Wang Y (2015) Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transp Res Pt C Emerg Technol 54:187–197. https://doi.org/10.1016/j.trc.2015.03.014
22. Yang HF, Dillon TS, Chen YP (2017) Optimized structure of the traffic flow forecasting model with a deep learning approach. IEEE Trans Neural Netw Learn Syst 28(10):2371–2381. https://doi.org/10.1109/TNNLS.2016.2574840
23. Duan Y, Lv Y, Wang FY (2016) Performance evaluation of the deep learning approach for traffic flow prediction at different times. In: IEEE International Conference on Service Operations & Logistics, pp 223–227
24. Huang W, Song G, Hong H, Xie K (2014) Deep architecture for traffic flow prediction: deep belief networks with multitask learning. IEEE trans Intell Transp Syst 15(5):2191–2201. https://doi.org/10.1109/TITS.2014.2311123
25. Bustillos BI, Chiu Y-C (2011) Real-time freeway-experienced travel time prediction using N-curve and k nearest neighbor methods. Transport Res Rec J Transport Res Board 2243(1):127–137. https://doi.org/10.3141/2243-15
26. Zhang N, Zhang Y, Lu H (2011) Seasonal autoregressive integrated moving average and support vector machine models. Transport Res Rec J Transport Res Board 2215(1):85–92. https://doi.org/10.3141/2215-09
27. Yin X, Wu G, Wei J, Shen Y, Qi H, Yin B (2021) Deep learning on traffic prediction: methods, analysis and future directions. IEEE Trans Intell Transp Syst. https://doi.org/10.1109/tits.2021.3054840
28. Khan Z, Khan SM, Dey K, Chowdhury M (2019) Development and evaluation of recurrent neural network-based models for hourly traffic volume and annual average daily traffic prediction. Transport Res Rec 2673(7):489–503. https://doi.org/10.1177/0361198119849059
29. Lv YS, Duan YJ, Kang WW, Li ZX, Wang FY (2015) Traffic flow prediction with big data: a deep learning approach. IEEE Trans Intell Transp Syst 16(2):865–873. https://doi.org/10.1109/Tits.2014.2345663
30. Ke R, Li W, Cui Z, Wang Y (2020) Two-stream multi-channel convolutional neural network for multi-lane traffic speed prediction considering traffic volume impact. Transport Res Rec 2674(4):459–470. https://doi.org/10.1177/0361198120911052
31. Hosseini MK, Talebpour A (2019) Traffic prediction using time-space diagram: a convolutional neural network approach. Transport Res Rec 2673(7):425–435. https://doi.org/10.1177/0361198119841291
32. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is all you need. In: 31st Annual Conference on Neural Information Processing Systems (NIPS), Long Beach
33. Ji Z, Xiong KL, Pang YW, Li XL (2020) Video summarization with attention-based encoder-decoder networks. IEEE Trans Circuits Syst Video Technol 30(6):1709–1717. https://doi.org/10.1109/tcsvt.2019.2904996
34. Wu P, Huang Z, Pian Y, Xu L, Li J, Chen K (2020) A combined deep learning method with attention-based LSTM model for short-term traffic speed forecasting. J Adv Transp. https://doi.org/10.1155/2020/8863724
35. Yang F, Zhang H, Tao S (2021) Travel order quantity prediction via attention-based bidirectional LSTM networks. J Supercomput. https://doi.org/10.1007/s11227-021-04032-8

36. Liu Q, Liu T, Cai Y, Xiong X, Hu Z (2021) Explanatory prediction of traffic congestion propagation mode: a self-attention based approach. Phys A 573:125940. https://doi.org/10.1016/j.physa.2021.125940
37. Sun B, Sun T, Zhang Y, Jiao P (2020) Urban traffic flow online prediction based on multi-component attention mechanism. IET Intell Transport Syst 14(10):1249–1258. https://doi.org/10.1049/iet-its.2020.0004
38. Shi XJ, Chen ZR, Wang H, Yeung DY, Wong WK, Woo WC (2015) Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In: Advances in neural information processing systems, vol 28, La Jolla
39. Du SD, Li TR, Gong X, Horng SJ (2020) A hybrid method for traffic flow forecasting using multimodal deep learning. Int J Comput Intell Syst 13(1):85–97. https://doi.org/10.2991/ijcis.d.200120.001
40. Hou Y, Edara P (2018) Network scale travel time prediction using deep learning. Transport Res Rec 2672(45):115–123. https://doi.org/10.1177/0361198118776139
41. Xiao Y, Yin H, Zhang Y, Qi H, Zhang Y, Liu Z (2021) A dual-stage attention-based Conv-LSTM network for spatio-temporal correlation and multivariate time series prediction. Int J Intell Syst 36(5):2036–2057. https://doi.org/10.1002/int.22370
42. Shabarek A, Chien S, Hadri S (2020) Deep learning framework for freeway speed prediction in adverse weather. Transport Res Rec 2674(10):28–41. https://doi.org/10.1177/0361198120947421
43. Bin Y, Shanhua W, Minghua W, Zhihong Z (2012) K-nearest neighbor short-term traffic flow prediction model. J Transp Eng 12(02):105–111
44. Bahdanau D, Cho K, Bengio Y (2014) Neural machine translation by jointly learning to align and translate. Computer Science
45. Amap (2021) Obtain traffic situation data. https://lbs.amap.com/api/webservice/guide/api/trafficstatus/. Accessed 8 July 2021
46. Kingma D, Ba J (2014) Adam: a method for stochastic optimization
47. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP (2003) Random forest: a classification and regression tool for compound classification and QSAR modeling. J Chem Inf Comput Sci 43(6):1947–1958. https://doi.org/10.1021/ci034160g

Authors and Affiliations

Xiaojian Hu[1,2,3,4] · Tong Liu[1] · Xiatong Hao[1] · Chenxi Lin[1]

Xiaojian Hu
huxiaojian@seu.edu.cn

Xiatong Hao
xthao96@foxmail.com

Chenxi Lin
918532967@qq.com

[1] Jiangsu Key Laboratory of Urban ITS, Southeast University, Southeast University Road #2, Nanjing 211189, China

[2] Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University, Nanjing, China

[3] National Demonstration Center for Experimental Road and Traffic Engineering Education (Southeast University), Nanjing 211189, China

[4] School of Transportation, Southeast University, Southeast University Road #2, Nanjing 211189, China