IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 15, 2021, accepted November 5, 2021, date of publication November 11, 2021,
date of current version December 2, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3127584

Hybrid Deep Spatio-Temporal Models for Traffic Flow Prediction on Holidays and Under Adverse Weather

WENSONG ZHANG[1], RONGHAN YAO[1], XIAOJING DU[1], AND JINSONG YE[2]
[1]School of Transportation and Logistics, Dalian University of Technology, Dalian 116024, China
[2]Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, China Academy of Transportation Sciences, Beijing 100029, China

Corresponding authors: Ronghan Yao (cyanyrh@dlut.edu.cn) and Jinsong Ye (yejs@motcats.ac.cn)

This work was supported in part by the Open Foundation of Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport under Grant 2020B1203, in part by the Fundamental Research Funds for the Central Universities of Ministry of Education of China under Grant DUT20JC40, and in part by the National Natural Science Foundation of China under Grant 52172314.

ABSTRACT Three hybrid deep spatio-temporal models are proposed to accurately predict traffic flow under normal conditions, on holidays and under adverse weather. Each of the proposed models consists of the global and target parts, and fuses the weather and traffic flow data obtained from the target and upstream sections. The convolutional neural network (CNN), and the gated recurrent unit (GRU) and convolutional long short-term memory (ConvLSTM) neural networks are selected to analyze the spatio-temporal characteristics of traffic flow data. Then, the three proposed models are verified using three actual cases, including traffic flow prediction under normal conditions, on holidays and under adverse weather. Moreover, the characteristics of traffic flow data on the Independence Day and Thanksgiving Day are discussed, as do the patterns of traffic flow data under heavy rain and strong wind. The experimental results show that: the three new models usually perform better than the existing models under all the situations; different holidays and different types of adverse weather have various impacts on the characteristics of traffic volume and speed data.

INDEX TERMS Traffic flow prediction, holidays, adverse weather, hybrid deep spatio-temporal model, convolutional long short-term memory (ConvLSTM).

I. INTRODUCTION

With the development of artificial intelligence, more and more advanced technologies have been applied in the intelligent transportation systems (ITS) [1]. Thus, ITS plays more and more important role in traffic control and management. A lot of ITS basic data are usually obtained from traffic flow prediction so that a large number of researchers have paid attention to traffic flow prediction [2]–[5].

To get more accurate traffic flow information, the patterns of traffic flow under all the situations, including normal conditions, holidays and adverse weather, should be analyzed in detail. The characteristics of traffic flow on holidays or under adverse weather are different from those under normal conditions. On important holidays, some governments often hold celebrations, and some highways are free of charge [6]. The corresponding measures will affect people's travel behaviors so that the characteristics of traffic flow will change. In the meantime, adverse weather happens more frequently due to global warming [7]. Under adverse weather, the road environment will become worse because of rainfall, wind, storm, etc. so that the patterns of traffic flow have to vary. According to the statistical analysis of highway traffic accidents in Zhejiang Province of China, about 85.00% of the accidents happened under adverse weather such as snowy days, rainy days and foggy days [8]. In summary, holidays and adverse weather have significant impacts on traffic flow. Therefore, it is necessary to propose a more reliable model for traffic flow prediction with the concern of holidays and adverse weather.

To find a more reliable model under various conditions, three hybrid deep spatio-temporal models (i.e., the CL-CN-G, CL-CNG and G-CN-CL models) for traffic flow prediction considering holidays and adverse weather are proposed by

The associate editor coordinating the review of this manuscript and approving it for publication was Emanuele Crisostomi.

combining the convolutional neural network (CNN), and the gated recurrent unit (GRU) and convolutional long short-term memory (ConvLSTM) models [9]. The CNN model is good at capturing the spatial rules of data, the GRU model can better analyze the temporal characteristics of data, and the spatio-temporal characteristics of traffic flow data can be captured by the ConvLSTM model. Thus, these three models are selected to combine. To better capture the rules of traffic flow data under various conditions, each of the three proposed models is composed of the global and target parts. The global part is used to analyze the characteristics of the traffic flow and weather data obtained from the target section during all the periods. Not only the traffic flow data, but also the weather data obtained from the target section during all the periods (which may include adverse weather and holidays) are selected as the input data of the global part. Then, more rules about adverse weather and holidays can be found. The target part is utilized to analyze the traffic flow data originated from the target and upstream sections during the target period. Then, the predicted values of the three proposed models are gotten by combining the outputs of the global and target parts. Three cases are designed to verify the three new models, and the most reliable model is recommended by comparing the performance of the selected models. In addition, the characteristics of traffic flow on holidays and those under adverse weather are analyzed.

The remainder of this paper is organized as below. In the ''Literature Review'' section, the existing models and methods for traffic flow prediction under normal conditions, on holidays and under adverse weather are summarized. The ''Methodology'' section describes the basic principles of the CNN, GRU, ConvLSTM models and the three proposed models. In the ''Experiment Setup and Case Studies'' section, three cases are designed to verify the three proposed models, and the characteristics of traffic flow on holidays and those under adverse weather are investigated in detail. Conclusions and discussions are presented in the ''Conclusions'' section.

II. LITERATURE REVIEW

Overall, the researches about traffic flow prediction can be divided into two categories, including traffic flow prediction under normal conditions and traffic flow prediction under abnormal conditions. The frequency of holidays and adverse weather is the highest under all the abnormal conditions. Meanwhile, holidays and adverse weather all have huge impacts on the patterns of traffic flow. Thus, the researches about traffic flow prediction under normal conditions, traffic flow prediction on holidays and traffic flow prediction under adverse weather are summarized below.

A. TRAFFIC FLOW PREDICTION UNDER NORMAL CONDITIONS

In the past decades, many models and methods have been applied for traffic flow prediction under normal conditions, including the regression method [2], the Markov model [10], the hidden Markov model [11], the autoregressive integrated moving average (ARIMA) model [12], [13], the K nearest neighbor (KNN) model [14], [15], the support vector machine (SVR) model [16], [17], the extreme learning machine method [18], the random forests method [19], the shallow neural network [20], [21], the deep belief network [22], [23], the CNN model [24], [25] and the recurrent neural network [26]–[28]. To consider the characteristics of traffic flow in a large-scale network, the graph neural network was utilized for traffic flow prediction [29]–[32]. The above studies only considered the characteristics of traffic flow under normal conditions. However, the patterns of traffic flow may be impacted by holidays and adverse weather.

To take into account the impacts of holidays and weather on traffic flow operations, some scholars proposed new models and methods with the concern of holiday and weather data. Based on the KNN model, Zhao *et al.* [33] and Qiao *et al.* [34] proposed the traffic volume and travel time prediction methods, and the features of day, precipitation intensity and wind speed were used to analyze the characteristics of traffic flow. In order to consider the influences of heavy snow on traffic flow, Tanimura *et al.* [35] proposed a traffic speed prediction model according to the multiple regression function. With the development of deep learning, some researchers put forward the traffic flow prediction methods based on the stacked autoencoder [36], the long short-term memory (LSTM) neural network [37], the GRU neural network [38], and the hybrid deep learning approaches [39], [40]. The holidays, weather and traffic flow data were selected to build the input matrices in the above models and methods, but these models and methods were not verified on holidays and under adverse weather.

B. TRAFFIC FLOW PREDICTION ON HOLIDAYS

The patterns of traffic flow on holidays are obviously different from those on weekdays because of the celebrations. Thus, Bai [41] proposed a passenger flow prediction method on holidays based on the ARIMA model, and the dummy variables and similar daily sample data were also considered. Similarly, Xie *et al.* [42] put forward an improved genetic algorithm optimized back propagation neural network for passenger flow prediction. In China, the freeway free policy is usually implemented on holidays so that the characteristics of traffic flow will change. To focus on such characteristics, some researchers proposed a hybrid method based on the discrete Fourier transform and the SVR model [6]. For highway traffic flow on holidays, Ji and Ge [43] developed a prediction method by combining the SVR model and the LSTM neural network, and the proposed method showed good performance under different conditions. The experimental results in the above studies reveal that it is necessary to consider the impacts of holidays on traffic flow.

C. TRAFFIC FLOW PREDICTION UNDER ADVERSE WEATHER

Drivers' travel behaviors will change under adverse weather, especially under heavy snow, heavy rain, fog, strong wind

and sandstorm [44]. To consider such impacts, Sun *et al.* [45] and Xu *et al.* [46] proposed the traffic flow prediction methods under adverse weather based on the fuzzy neural network and the random forest method. For urban rail transit ridership under rainfall, Xue *et al.* [47] put forward the hybrid prediction methods based on the seasonal autoregressive integrated moving average model, the SVR model and the multiple linear regression model. Some researchers presented a Markov-based time series model with the concern of rain and snow weather, and the results revealed that the formulated model was good [48]. Because deep learning methods can better capture the rules of traffic flow data, Polson and Sokolov [49] and Shabarek *et al.* [50] both proposed the deep learning models for traffic flow prediction under adverse weather. The experimental outcomes showed that the deep learning models can get more accurate prediction values of traffic flow data under adverse weather.

The models and methods proposed in the aforementioned first aspect can accurately predict the traffic flow data under normal conditions. In the aforementioned latter two aspects, some researchers make efforts to investigate traffic flow prediction on holidays and under adverse weather. However, there are still two main problems: (1) When predicting traffic flow, the above studies usually only considered holidays or one type of adverse weather. However, the spatio-temporal characteristics of traffic flow on different holidays and under different types of adverse weather are various. (2) With the development of deep learning, more advanced models and methods have been proposed, such as the GRU, ConvLSTM and graph neural networks. It is necessary to use these new models to put forward a more reasonable prediction method for traffic flow to cover different situations. Therefore, it needs to be further discussed how to build a more general prediction model.

To solve the above problems, three hybrid deep spatio-temporal models for traffic flow prediction under normal conditions, on holidays and under adverse weather are constructed. The three proposed models (i.e., the CL-CN-G, CL-CNG and G-CN-CL models) are built by combining the CNN, GRU and ConvLSTM models, and each of them consists of the global and target parts. To consider the relationships among traffic flow data, holidays data and adverse weather data, the traffic flow and weather data obtained from the target section during all the periods are used to build the input matrices in the global part, and the traffic flow data obtained from the target and upstream sections during the target period are selected to build the input matrices in the target part. The construction of the input matrices will be explained in detail in the next section.

III. METHODOLOGY

Traffic flow prediction means to estimate the traffic flow data in the future. The traffic flow data for section m is denoted by $X_m = (x_{1,m}, x_{2,m}, x_{3,m}, \cdots, x_{t,m}, \cdots, x_{T,m})$, where $x_{t,m}$ is the traffic flow data at time interval t for section m and T is the number of time intervals. In this section, the CNN, GRU and ConvLSTM models are described. Then, the structures of the CL-CN-G, CL-CNG and G-CN-CL models are explained.

A. CNN MODEL

The CNN model is composed of the input, convolutional, pooling and output layers. The input matrix $\tilde{\mathbf{X}}_m = [\tilde{x}_{h+1,m}; \tilde{x}_{h+2,m}; \tilde{x}_{h+3,m}; \cdots; \tilde{x}_{t,m}; \cdots; \tilde{x}_{T,m}]$ of the input layer is built based on X_m, where $\tilde{x}_{t,m} = \left(x_{t-h,m}, x_{t-h+1,m}, x_{t-h+2,m}, \cdots, x_{t-1,m}\right)$ is the input vector at time interval t for section m, and h is the number of time lags of traffic flow data. Then, the input matrix $\tilde{\mathbf{X}}_m$ is processed by the convolutional layer. The output of the convolutional layer can be given by [51]

$$C_{n,m}^{c_n} = \sigma_m^{c_n}(W_{n,m}^{c_n} \otimes \tilde{x}_{n,t,m} + b_{n,m}^{c_n}) \tag{1}$$

where $C_{n,m}^{c_n}$ is the output of the n^{th} layer when using the c_n^{th} convolutional filter for section m, $c_n \in \{1, 2, 3, \cdots, C_n\}$, C_n is the number of convolutional filters in the n^{th} layer, $n \in \{1, 2, 3, \cdots, N\}$, N is the depth of the CNN model; $\tilde{x}_{n,t,m}$ is the input vector of the n^{th} layer at time interval t for section m, $\tilde{x}_{1,t,m} = \tilde{x}_{t,m}$; $\otimes$ is the process of convolution; $\sigma_m^{c_n}$, $W_{n,m}^{c_n}$ and $b_{n,m}^{c_n}$ are the activation function, weight and bias vector of the n^{th} layer when using the c_n^{th} convolutional filter for section m, respectively.

Next, the pooling layer is used to extract the features from the outputs of the convolutional layer. The output of the pooling layer, denoted by $P_{n,m}^{c_n}$, is

$$P_{n,m}^{c_n} = \text{pool}(C_{n,m}^{c_n}) \tag{2}$$

where pool is the process of pooling.

Then, the outputs of the pooling layer need to be flattened, namely

$$P_{N-1,m} = \text{flatten}([P_{N-1,m}^1; P_{N-1,m}^2; P_{N-1,m}^3; \cdots; P_{N-1,m}^{c_{N-1}}; \cdots; P_{N-1,m}^{C_{N-1}}]) \tag{3}$$

where $P_{N-1,m}$ is the output of the flattening of the $N-1^{\text{th}}$ layer for section m; flatten is the process of flattening; $P_{N-1,m}^{c_{N-1}}$ is the output of the pooling for the $N-1^{\text{th}}$ layer when using the c_{N-1}^{th} convolutional filter for section m, $c_{N-1} \in \{1, 2, 3, \cdots, C_{N-1}\}$, C_{N-1} is the number of the convolutional filters in the $N-1^{\text{th}}$ layer.

Finally, the predicted value of traffic flow data can be given by the output layer, i.e.

$$\hat{y}_{t,m}^{\text{CN}} = \sigma_m^{\text{CN}}(W_m^{\text{CN}} \cdot P_{N-1,m} + b_m^{\text{CN}}) \tag{4}$$

where $\hat{y}_{t,m}^{\text{CN}}$ is the predicted value obtained from the CNN model at time interval t for section m; σ_m^{CN}, W_m^{CN} and b_m^{CN} are the activation function, weight and bias vector of the output layer in the CNN model for section m, respectively.

B. GRU MODEL

The GRU model contains the input, hidden and output layers. The hidden layer consists of the update and reset gates.

The outputs of the update and reset gates are computed as [52]

$$z_{t,m} = \sigma_m^z(W_m^z \cdot [h_{t-1,m}, \tilde{x}_{t,m}] + b_m^z) \quad (5)$$
$$r_{t,m} = \sigma_m^r(W_m^r \cdot [h_{t-1,m}, \tilde{x}_{t,m}] + b_m^r) \quad (6)$$

where $z_{t,m}$ is the output of the update gate at time interval t for section m; $h_{t-1,m}$ is the memory information of the hidden layer at time interval $t-1$ for section m; $r_{t,m}$ is the output of the reset gate at time interval t for section m; σ_m^z, W_m^z and b_m^z are the activation function, weight and bias vector of the update gate for section m, respectively; σ_m^r, W_m^r and b_m^r are the activation function, weight and bias vector of the reset gate for section m, respectively.

The memory information at time interval t for section m, denoted by $h_{t,m}$, is

$$h_{t,m} = (1 - z_{t,m}) \odot h_{t-1,m} + z_{t,m} \odot h'_{t,m}$$
$$h'_{t,m} = \tanh(W_m^{h'} \cdot [h_{t-1,m} \odot r_{t,m}, \tilde{x}_{t,m}] + b_m^{h'}) \quad (7)$$

where $h'_{t,m}$ is a variable when calculating $h_{t,m}$; $\odot$ is the Hadamard product; tanh is the hyperbolic tangent function; $W_m^{h'}$ and $b_m^{h'}$ are the weight and bias vector when calculating $h'_{t,m}$, respectively.

Next, the predicted value of traffic flow data can be obtained based on the memory information, i.e.

$$\hat{y}_{t,m}^{\mathrm{GR}} = \sigma_m^{\mathrm{GR}}(W_m^{\mathrm{GR}} \cdot h_{t,m} + b_m^{\mathrm{GR}}) \quad (8)$$

where $\hat{y}_{t,m}^{\mathrm{GR}}$ is the predicted value obtained from the GRU model at time interval t for section m; σ_m^{GR}, W_m^{GR} and b_m^{GR} are the activation function, weight and bias vector of the output layer in the GRU model for section m, respectively.

C. CONVLSTM MODEL

The ConvLSTM model consists of the input, hidden and output layers. The hidden layer is composed of the input, forget and output gates. The outputs of these three gates and the memory information can be calculated by the following group of formulas [9]:

$$\begin{aligned} i_{t,m} &= \sigma_m^i(W_m^i \otimes [\bar{h}_{t-1,m}, \tilde{x}_{t,m}] + b_m^i) \\ f_{t,m} &= \sigma_m^f(W_m^f \otimes [\bar{h}_{t-1,m}, \tilde{x}_{t,m}] + b_m^f) \\ o_{t,m} &= \sigma_m^o(W_m^o \otimes [\bar{h}_{t-1,m}, \tilde{x}_{t,m}] + b_m^o) \\ c_{t,m} &= f_{t,m} \odot c_{t-1,m} + i_{t,m} \odot \tanh(W_m^c \\ &\quad \otimes [\bar{h}_{t-1,m}, \tilde{x}_{t,m}] + b_m^c) \\ \bar{h}_{t,m} &= o_{t,m} \odot \tanh(c_{t,m}) \end{aligned} \quad (9)$$

where $i_{t,m}$, $f_{t,m}$ and $o_{t,m}$ are the outputs of the input, forget and output gates at time interval t for section m, respectively; $c_{t,m}$ and $c_{t-1,m}$ are the states of the cell in the hidden layer at time intervals t and $t-1$ for section m, respectively; $\bar{h}_{t,m}$ and $\bar{h}_{t-1,m}$ are the memory information of the hidden layer in the ConvLSTM model at time intervals t and $t-1$ for section m, respectively; σ_m^i, W_m^i and b_m^i are the activation function, weight and bias vector when calculating $i_{t,m}$, respectively; σ_m^f, W_m^f and b_m^f are the activation function, weight and bias vector when calculating $f_{t,m}$, respectively; σ_m^o, W_m^o and b_m^o are the activation function, weight and bias vector when calculating $o_{t,m}$, respectively; W_m^c and b_m^c are the weight and bias vector when calculating $c_{t,m}$, respectively.

Then, the predicted value obtained from the ConvLSTM model, denoted by $\hat{y}_{t,m}^{\mathrm{CL}}$, can be expressed as

$$\hat{y}_{t,m}^{\mathrm{CL}} = \sigma_m^{\mathrm{CL}}(W_m^{\mathrm{CL}} \cdot \bar{h}_{t,m} + b_m^{\mathrm{CL}}) \quad (10)$$

where σ_m^{CL}, W_m^{CL} and b_m^{CL} are the activation function, weight and bias vector of the output layer in the ConvLSTM model for section m, respectively.

D. HYBRID MODELS

Three hybrid models are proposed by combining the CNN, GRU and ConvLSTM models, namely, the CL-CN-G model, the CL-CNG model, and the G-CN-CL model. Every proposed model consists of the global and target parts. In the CL-CN-G model, the global part is built using the ConvLSTM model, and the target part is built by connecting the CNN and GRU models in parallel. In the CL-CNG model, the global part is also built using the ConvLSTM model, whereas the target part is built by connecting the CNN and GRU models in series. In the G-CN-CL model, the global part is constructed using the GRU model, and the target part is constructed using the CNN and ConvLSTM models in parallel. The global part is used to analyze the rules of the weather and traffic flow data originated from the target section during all the periods. The target part aims to analyze the characteristics of the traffic flow data obtained from the target and upstream sections during the target period.

The general frameworks of the three new models are illustrated in Fig. 1. In this figure, L^{G} is the depth of the selected model in the global part, L^{TC}, L^{TG} and L^{TL} are the depth of the CNN, GRU and ConvLSTM models in the target part, respectively; $\hat{y}_m^{\mathrm{G1}}$, $\hat{y}_m^{\mathrm{G2}}$ and $\hat{y}_m^{\mathrm{G3}}$ are the vectors denoting the predicted values obtained from the global parts in the CL-CN-G, CL-CNG and G-CN-CL models, respectively; $\hat{y}_m^{\mathrm{T1}}$, $\hat{y}_m^{\mathrm{T2}}$ and $\hat{y}_m^{\mathrm{T3}}$ are the vectors denoting the predicted values obtained from the target parts in the CL-CN-G, CL-CNG and G-CN-CL models, respectively; $\hat{y}_m^1$, $\hat{y}_m^2$ and $\hat{y}_m^3$ are the vectors denoting the predicted values of the traffic flow data obtained by the CL-CN-G, CL-CNG and G-CN-CL models, respectively; Dense means the Dense layer, and Fusion indicates the outputs are combined by a fully connected layer.

When using the three proposed models, the input matrices need to be constructed firstly. Considering vehicle detectors, the traffic flow data for the target section X_m and those for the upstream section X_{m-1} can be obtained. The weather data $W_m = (w_{1,m}, w_{2,m}, w_{3,m}, \cdots, w_{t,m}, \cdots, w_{T,m})$ (e.g., temperature, wind speed, precipitation, etc.) can be collected by weather stations, where $w_{t,m}$ is the weather data at time interval t for section m. The input matrix of the global part $\bar{\mathbf{X}}_m = [\bar{x}_{\tilde{h}+1,m}; \bar{x}_{\tilde{h}+2,m}; \bar{x}_{\tilde{h}+3,m}; \cdots; \bar{x}_{t,m}; \cdots; \bar{x}_{T,m}]$ is constructed by X_m and W_m, where $\bar{x}_{t,m} = (w_{t-\tilde{h},m}, \cdots, w_{t-1,m}, x_{t-\tilde{h},m}, \cdots, x_{t-1,m})$ is the input vector of the global part at time interval t for section m, $\tilde{h}$ and

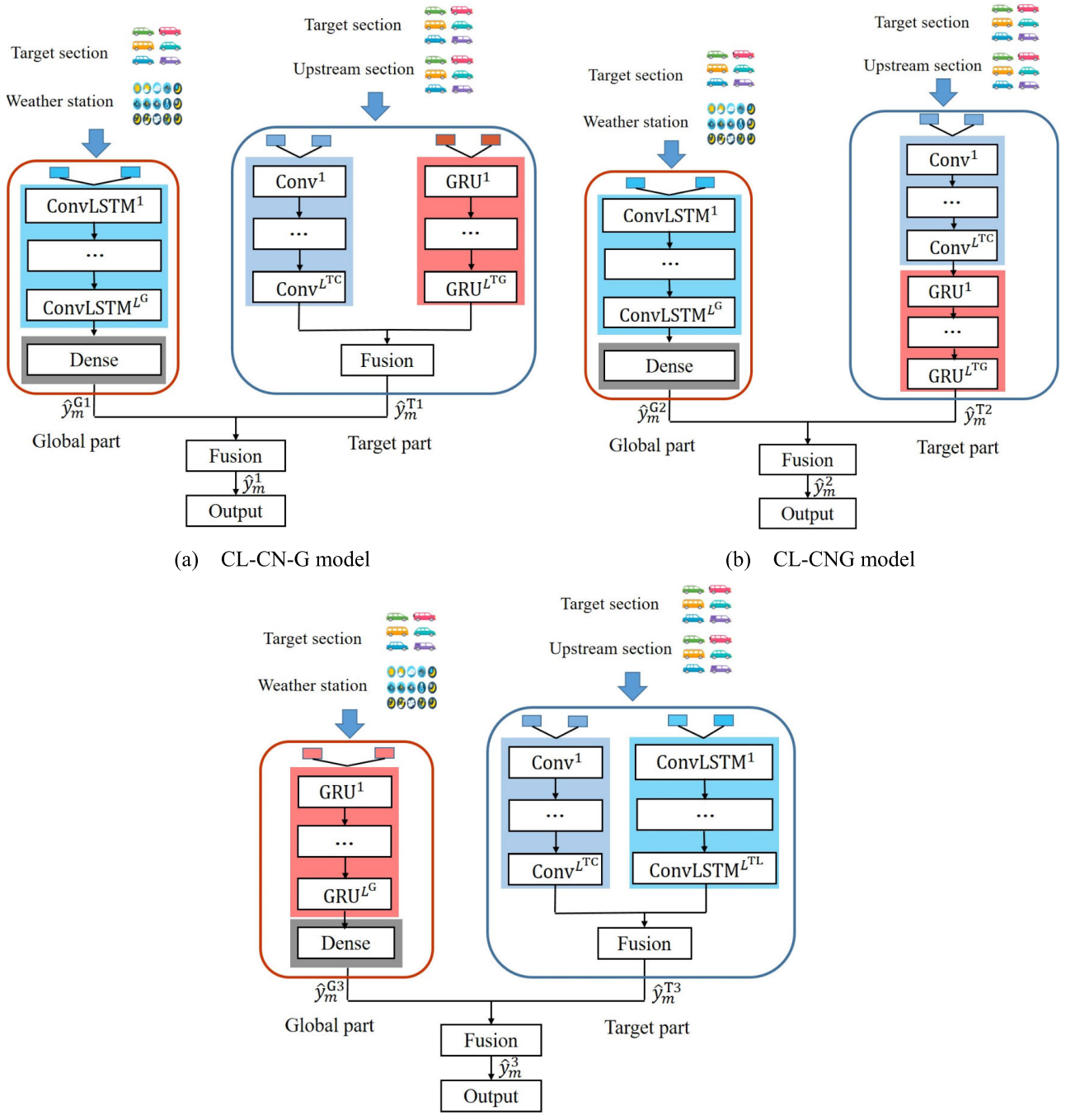


FIGURE 1. General frameworks of the proposed models.

$\bar{h}$ are individually the number of time lags of weather data and that of traffic flow data for the global part, $\tilde{h} \leq \bar{h}$. Then, the traffic flow data for section m during period i, denoted by $X_m^i = (x_{1,m}^i, x_{2,m}^i, x_{3,m}^i, \cdots, x_{t^i,m}^i, \cdots, x_{T^i,m}^i)$, can be gotten, where $x_{t^i,m}^i$ is the traffic flow data at time interval t^i during period i for section m, $i \in \{0, 1\}$; if $i = 0$, $X_m^i = X_m^0$ is the traffic flow data on weekends for section m; if $i = 1$, $X_m^i = X_m^1$ is the traffic flow data on weekdays for section m, $T^0 + T^1 = T$. For example, X_m indicates the traffic flow data obtained from section m during a week, T is 2016 if the time interval to aggregate traffic flow data is 5 minutes. Then, X_m^0 means the traffic flow data obtained from section m on Saturday and Sunday, and T^0 is 576; X_m^1 indicates the traffic flow data originated from section m on five weekdays, and T^1 is 1440. Similarly, series X_{m-1}^0 and X_{m-1}^1 can be gotten based on X_{m-1}. According to

X_m^i and X_{m-1}^i, the input matrix in the target part $\bar{\mathbf{X}}_m^i = [\bar{x}_{\bar{h}^i+1,m}^i; \bar{x}_{\bar{h}^i+2,m}^i; \bar{x}_{\bar{h}^i+3,m}^i; \cdots; \bar{x}_{t^i,m}^i; \cdots; \bar{x}_{T^i,m}^i]$ can be constructed, where $\bar{x}_{t^i,m}^i = (x_{t^i-\bar{h}^i,m-1}^i, \cdots, x_{t^i-1,m-1}^i, x_{t^i-\bar{h}^i,m}^i, \cdots, x_{t^i-1,m}^i)$ is the input vector at time interval t^i during period i for section m, $\bar{h}^i$ is the number of time lags of traffic flow data during period i for the target part.

Then, the predicted values obtained from the global part in the three new models, denoted by $\hat{y}_{t,m}^{G1}$, $\hat{y}_{t,m}^{G2}$ and $\hat{y}_{t,m}^{G3}$, can be calculated as

$$\hat{y}_{t,m}^{G1} = \sigma_m^{G1}(W_m^{G1} \cdot \bar{h}_{t,m}^{G1} + b_m^{G1}) \quad (11)$$
$$\hat{y}_{t,m}^{G2} = \sigma_m^{G2}(W_m^{G2} \cdot \bar{h}_{t,m}^{G2} + b_m^{G2}) \quad (12)$$
$$\hat{y}_{t,m}^{G3} = \sigma_m^{G3}(W_m^{G3} \cdot h_{t,m}^{G3} + b_m^{G3}) \quad (13)$$

where $\bar{h}_{t,m}^{G1}$, $\bar{h}_{t,m}^{G2}$ and $h_{t,m}^{G3}$ are the memory information in the hidden layer of the global part in the CL-CN-G, CL-CNG and G-CN-CL models at time interval t for section m, respectively; σ_m^{G1}, W_m^{G1} and b_m^{G1} are the activation function, weight and bias vector when calculating $\hat{y}_{t,m}^{G1}$, respectively; σ_m^{G2}, W_m^{G2} and b_m^{G2} are the activation function, weight and bias vector when calculating $\hat{y}_{t,m}^{G2}$, respectively; σ_m^{G3}, W_m^{G3} and b_m^{G3} are the activation function, weight and bias vector when calculating $\hat{y}_{t,m}^{G3}$, respectively.

The vectors denoting the predicted values obtained from the global parts in the CL-CN-G, CL-CNG and G-CN-CL models are denoted by $\hat{y}_m^{G1}$, $\hat{y}_m^{G2}$ and $\hat{y}_m^{G3}$, respectively.

The predicted values obtained from the target parts in the three proposed models, denoted by $\hat{y}_{t^i,m}^{T1}$, $\hat{y}_{t^i,m}^{T2}$ and $\hat{y}_{t^i,m}^{T3}$, are

$$\hat{y}_{t^i,m}^{T1} = \sigma_m^{T1}(W_m^{T1} \cdot [P_{N-1,m}^{T1^i}, h_{t^i,m}^{T1}] + b_m^{T1}) \quad (14)$$
$$\hat{y}_{t^i,m}^{T2} = \sigma_m^{T2}(W_m^{T2} \cdot P_{N-1,m}^{T2^i} + b_m^{T2}) \quad (15)$$
$$\hat{y}_{t^i,m}^{T3} = \sigma_m^{T3}(W_m^{T3} \cdot [P_{N-1,m}^{T3^i}, \bar{h}_{t^i,m}^{T3}] + b_m^{T3}) \quad (16)$$

where $P_{N-1,m}^{T1^i}$, $P_{N-1,m}^{T2^i}$ and $P_{N-1,m}^{T3^i}$ are the outputs of the $N-1^{\text{th}}$ layer of the target parts in the CL-CN-G, CL-CNG and G-CN-CL models during period i for section m, respectively; $h_{t^i,m}^{T1}$ and $\bar{h}_{t^i,m}^{T3}$ are the memory information of the hidden layers of the target parts in the CL-CN-G and G-CN-CL models at time interval t^i for section m, respectively; σ_m^{T1}, W_m^{T1} and b_m^{T1} are the activation function, weight and bias vector of the target part in the CL-CN-G model, respectively; σ_m^{T2}, W_m^{T2} and b_m^{T2} are the activation function, weight and bias vector of the target part in the CL-CNG model, respectively; σ_m^{T3}, W_m^{T3} and b_m^{T3} are the activation function, weight and bias vector of the target part in the G-CN-CL model, respectively.

The vectors denoting the predicted values obtained from the target parts in the CL-CN-G, CL-CNG and G-CN-CL models are denoted by $\hat{y}_m^{T1}$, $\hat{y}_m^{T2}$ and $\hat{y}_m^{T3}$, respectively.

Finally, the vectors denoting the predicted values of the traffic flow data acquired by the CL-CN-G, CL-CNG and G-CN-CL models can be calculated by combining the outputs of the global and target parts, i.e.

$$\hat{y}_m^1 = \sigma_m^1(W_m^1 \cdot [\hat{y}_m^{G1}, \hat{y}_m^{T1}] + b_m^1) \quad (17)$$
$$\hat{y}_m^2 = \sigma_m^2(W_m^2 \cdot [\hat{y}_m^{G2}, \hat{y}_m^{T2}] + b_m^2) \quad (18)$$
$$\hat{y}_m^3 = \sigma_m^3(W_m^3 \cdot [\hat{y}_m^{G3}, \hat{y}_m^{T3}] + b_m^3) \quad (19)$$

where σ_m^1, W_m^1 and b_m^1 are the activation function, weight and bias vector of the output layer in the CL-CN-G model, respectively; σ_m^2, W_m^2 and b_m^2 are the activation function, weight and bias vector of the output layer in the CL-CNG model, respectively; σ_m^3, W_m^3 and b_m^3 are the activation function, weight and bias vector of the output layer in the G-CN-CL model, respectively.

The parameters of each layer in the three new models are trained by minimizing the loss function L, and the loss function can be given by

$$L = \sum\nolimits_m [(x_m - \hat{y}_m)^2] \quad (20)$$

where x_m is the vector denoting the observed values of the traffic flow data for section m; $\hat{y}_m$ is the vector denoting the predicted values of the traffic flow data obtained from the selected model for section m.

IV. EXPERIMENT SETUP AND CASE STUDIES

To testify the effectiveness of the three proposed models, three situations including normal conditions, holidays and adverse weather, are considered in this section. Case one displays traffic flow prediction during a week under normal conditions. Case two shows traffic flow prediction on holidays, and case three presents traffic flow prediction under adverse weather. The raw traffic flow data were downloaded from the Caltrans Performance Measurements Systems. The weather data were obtained from the website of "https://mesowest.utah.edu/". Furthermore, the characteristics of traffic flow data on holidays and under adverse weather are discussed. The KNN, CNN, LSTM, GRU, ConvLSTM models and the CNNs-LSTM model [53], [54] are selected as the comparative models, and the mean absolute percentage error (MAPE), the mean absolute error (MAE) and the rooted mean square error (RMSE) are adopted as the performance indices. The formulas of MAPE, MAE and RMSE are as below:

$$\text{MAPE} = \frac{1}{\tilde{T}} \sum_{t=1}^{\tilde{T}} \left| \frac{\hat{y}_{t,m} - x_{t,m}}{x_{t,m}} \right| \quad (21)$$

$$\text{MAE} = \frac{1}{\tilde{T}} \sum_{t=1}^{\tilde{T}} \left| \hat{y}_{t,m} - x_{t,m} \right| \quad (22)$$

$$\text{RMSE} = \sqrt{\frac{1}{\tilde{T}} \sum_{t=1}^{\tilde{T}} (\hat{y}_{t,m} - x_{t,m})^2} \quad (23)$$

where $\hat{y}_{t,m}$ is the predicted value of traffic volume at time interval t for section m; $\tilde{T}$ is the number of the time intervals selected for traffic flow prediction.

TABLE 1. Average performance indices obtained from each model for traffic volume prediction in case one.

Performance indices	Model								
	KNN	CNN	LSTM	GRU	ConvLSTM	CNNs-LSTM	CL-CN-G	CL-CNG	G-CN-CL
MAPE (%)	11.14	9.40	9.05	9.12	8.98	9.38	**8.89**	9.09	**8.85**
MAE (veh/5mins)	30.72	24.48	24.29	23.97	24.08	24.46	**23.43**	**23.44**	**23.30**
RMSE (veh/5mins)	41.15	32.61	32.54	32.08	32.22	32.55	**31.36**	**31.11**	**31.10**

Note: Bold numbers indicate that the values of the proposed models are less than those of all the existing models.

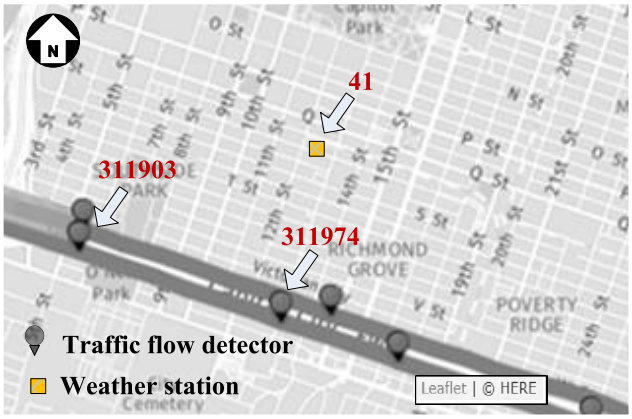


FIGURE 2. Locations of the selected detectors and weather station.

A. TRAFFIC FLOW PREDICTION UNDER NORMAL CONDITIONS

To test the proposed models, the three models are firstly applied under normal conditions. In this case, the sections for detectors 311974 and 311903 in the US-50 corridor in Sacramento of United States are selected as the study sites. Detector 311903 is located at the upstream section for detector 311974. The locations of the selected detectors and weather station are depicted in Fig. 2. The traffic flow data obtained by the two selected detectors from September 14 to November 8 in 2020 are selected, and the data on September 17, 24, 25 and 26 are removed because a lot of data are missing on these days. The time interval by which traffic flow data are aggregated is 5 minutes. Thus, one day includes 288 time intervals. To predict the traffic volume and the traffic speed during the last week (i.e., November 2 to 8), the traffic flow data on the former 45 days are used to train the prediction models. There are no special conditions in the last week. By testing the model, the number of time lags for the traffic flow data for the target section and that for the upstream section in the target part are both set to 6, and that for the target section in the global part is also set to 6 [55].

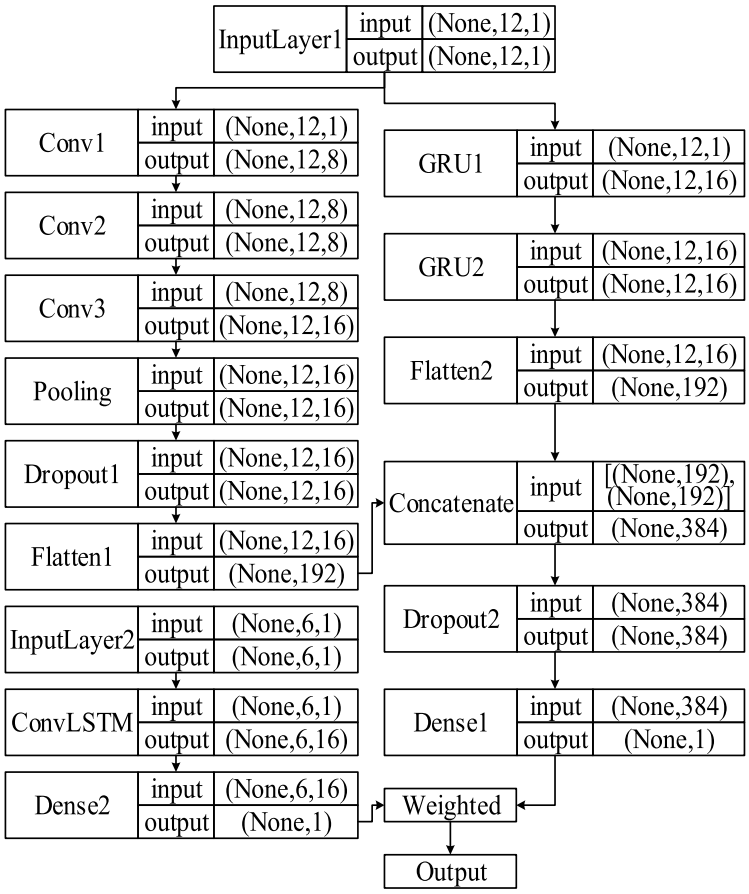
Fig. 3 displays the specific structures of the proposed models in case one, including the number of layers and the number of nodes in each layer. As shown in Fig. 3(a), the global part of the CL-CN-G model contains one ConvLSTM layer and one Dense layer. The target part of the CL-CN-G model contains two paths: (1) path 1 has three Conv layers, one Pooling layer, one Dropout layer and one Flatten layer; path 2 includes two GRU layers and one Flatten layer. Then, the outputs of the two paths are concatenated by a Concatenate layer. As a result, the predicted values can be obtained by combining the outputs of the global and target parts.

Fig. 3(b) shows the specific structure of the CL-CNG model in this case. In this figure, the global part of the CL-CNG model contains one ConvLSTM layer and one Dense layer; the target part of the CL-CNG model has two Conv layers, one Pooling layer, two GRU layers, one Flatten layer and one Dense layer. Then, the predicted values of the CL-CNG model can be calculated by the outputs of the two parts.

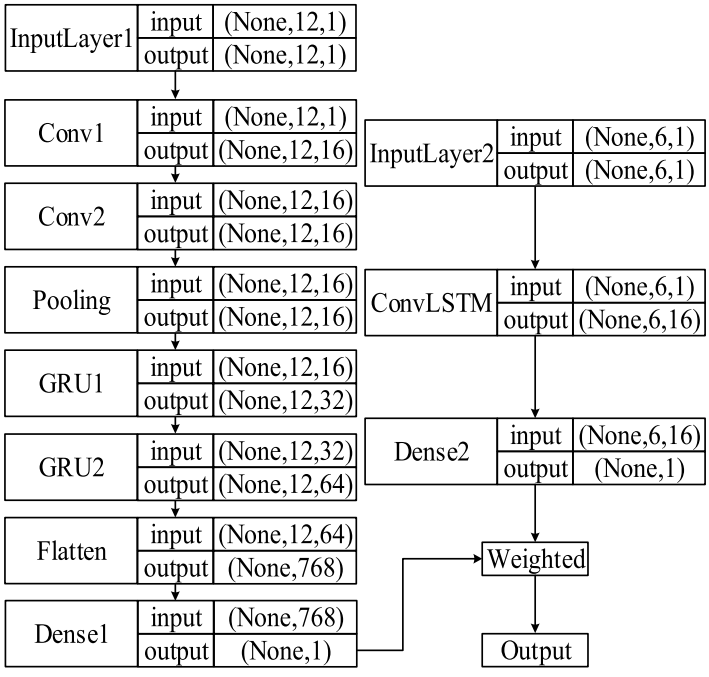
The specific structure of the G-CN-CL model in this case is shown in Fig. 3(c). It can be seen that the global part of the G-CN-CL model contains two GRU layers and one Dense layer. The target part of the G-CN-CL model is composed of two paths: (1) path 1 has two Conv layers, one Pooling layer, one Dropout layer and one Flatten layer; (2) path 2 consists of two ConvLSTM layers and one Flatten layer. Then, the outputs of the two paths are concatenated by a Concatenate layer. Next, the predicted values can be obtained by combining the outputs of the global and target parts.

By testing the model, the dropout rates of the Dropout1 and Dropout2 layers in the CL-CN-G and G-CN-CL models are both set to 0.5 and 0.05, respectively. The Adam algorithm is used to train the model. The activation function in each layer adopts the ReLU function, the number of training is set to 100, and the batch size is set to 64. When predicting traffic volume and speed, the models are trained using traffic volume data and traffic speed data, respectively. All the experiments are compiled on a Windows server (CPU: Intel(R) Core(TM) i7-7700HQ @ 2.80 GHz, GPU: NVIDIA GeForce GTX 1050Ti).

Tables 1 and 2 list the average performance indices (i.e., MAPE, MAE and RMSE) obtained from each model for traffic volume prediction and those for traffic speed prediction, respectively. The bold digitals indicate the values calculated by the three proposed models are smaller than those obtained from the existing models. Then, the selected models are compared. It can be revealed that: (1) the performance indices obtained from the CL-CN-G, CL-CNG and G-CN-CL models are usually smaller than those obtained from the existing models, which means the three proposed

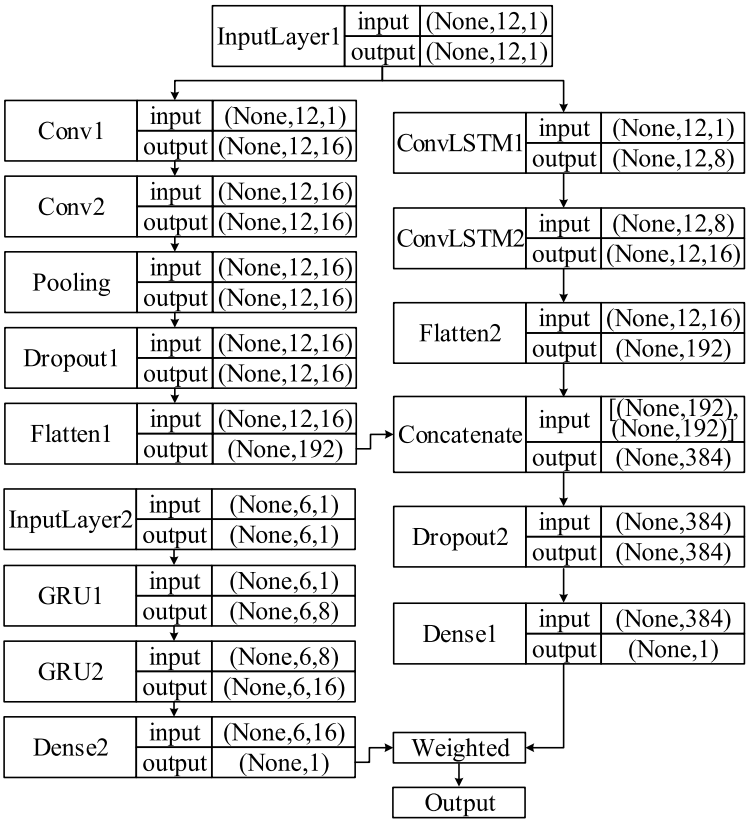


(a) Specific structure of the CL-CN-G model



(b) Specific structure of the CL-CNG model

FIGURE 3. Specific structures of the proposed models in case one.



(c) Specific structure of the G-CN-CL model

FIGURE 3. *(Continued.)* **Specific structures of the proposed models in case one.**

TABLE 2. **Average performance indices obtained from each model for traffic speed prediction in case one.**

Performance indices	Model								
	KNN	CNN	LSTM	GRU	ConvLSTM	CNNs-LSTM	CL-CN-G	CL-CNG	G-CN-CL
MAPE (%)	4.00	2.46	2.56	2.51	2.48	2.60	**2.33**	**2.40**	**2.36**
MAE (mph)	2.26	1.43	1.50	1.46	1.46	1.46	**1.35**	**1.40**	**1.37**
RMSE (mph)	3.66	2.24	2.30	2.23	2.23	2.31	**2.11**	**2.15**	**2.14**

Note: Bold numbers indicate that the values of the proposed models are less than those of all the existing models.

models perform better than the existing models; (2) among the three proposed models, the performance indices obtained from the CL-CNG model are usually larger than those obtained from the CL-CN-G and G-CN-CL models, which demonstrates the CL-CNG model performs worse than the CL-CN-G and G-CN-CL models; (3) the performance indices obtained from the deep learning models are always smaller than those calculated by the KNN model, which shows the deep learning models perform better than the KNN model

The above outcomes indicate that: (1) the CL-CN-G, CL-CNG and G-CN-CL models perform better than the existing models, which denotes the three proposed models are effective and efficient for traffic flow prediction under normal conditions; (2) the deep learning models perform better than the traditional machine learning model mentioned in this study; (3) the CL-CN-G and G-CN-CL models show more stable performance than the CL-CNG model.

B. TRAFFIC FLOW PREDICTION ON HOLIDAYS

In this case, detectors 311974 and 311903 shown in Fig. 2 are still selected. To verify the performance of the three proposed models on holidays, the Independence Day in 2019 (July 4, 2019) and the Thanksgiving Day in 2020 (November 26, 2020) are selected. The traffic flow data collected from detectors 311974 and 311903 from May 13 to July 7 in 2019 and those from October 5 to November 29 in 2020 were downloaded. The weather data from May 13 to July 7 in 2019 and those from October 5 to

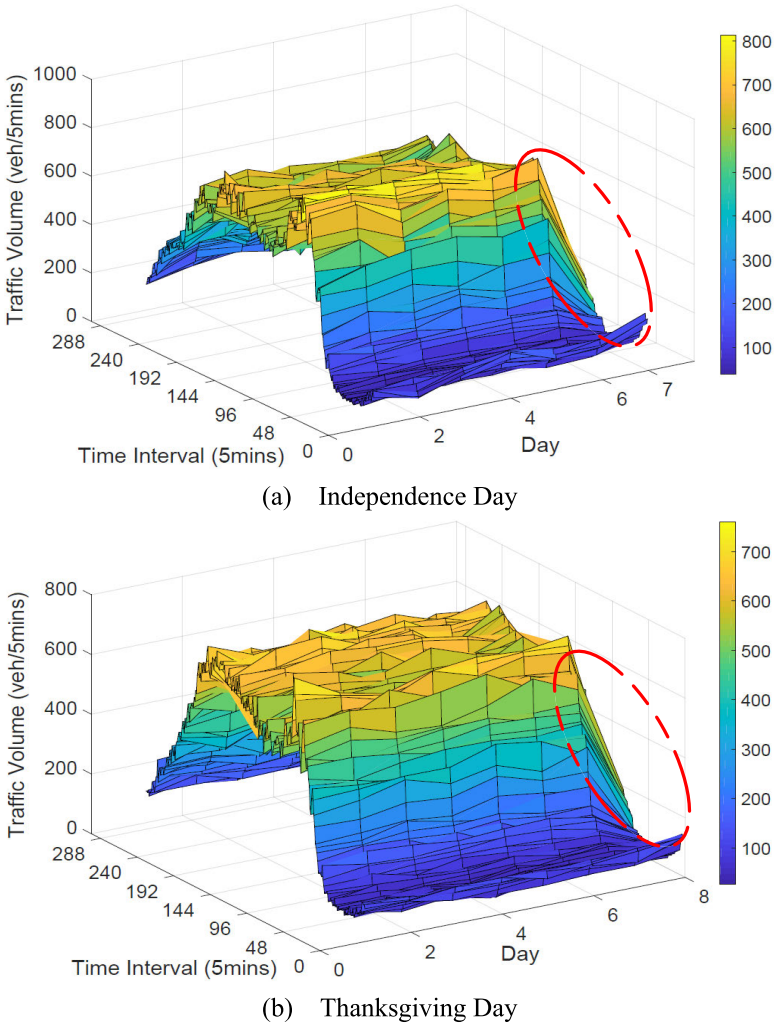


FIGURE 4. Variations of traffic volumes on thursdays.

November 29 in 2020 at weather station 41 shown in Fig. 2 were downloaded, too. To predict the traffic volume and speed on the Independence Day and Thanksgiving Day, the traffic flow data from May 13 to July 3 in 2019 and those from October 5 to November 25 in 2020 are selected as the historical data. The data on May 21, 22, 23 and June 14 in 2019 are removed because a lot of data are missing on these days. The traffic flow patterns on holidays are also discussed.

Figs. 4(a) and 4(b) show the variations of traffic volumes on Thursdays, including the Independence Day and Thanksgiving Day. One day includes 288 time intervals, and 0 in the figures represents the start time. The seven days in Fig. 4(a) are May 16, May 30, June 6, June 13, June 20, June 27 and July 4 (the Independence Day) in 2019, and the eight days in Fig. 4(b) are October 8, October 15, October 22, October 29, November 5, November 12, November 19 and November 26 (the Thanksgiving Day) in 2020. It can be seen that the traffic volumes on the Independence Day and Thanksgiving Day (i.e., the traffic volumes in the red circles) are obviously smaller than those on other Thursdays, which indicates that the traffic flow patterns on holidays are different from those on workdays.

Figs. 5(a) and 5(b) display the variations of traffic speeds on Thursdays, including the Independence Day and Thanksgiving Day. In Fig. 5(a), 5.16 means May 16 in 2019, and 10.08 in Fig. 5(b) means October 8 in 2020. It can be seen that: the traffic speeds in daytime are usually lower than those at night; however, the traffic speeds in daytime on the Independence Day and Thanksgiving Day are close to those at night, which again reveals that the characteristics of traffic flow will change on holidays. Thus, it is necessary to consider the impacts of holidays on traffic flow.

In this case, the structures of the CL-CN-G, CL-CNG and G-CN-CL models are the same as those in case one. The parameters in each layer and the experimental environment are also the same as those in case one. To analyze the performance of each model, Table 3 lists the average performance indices obtained from all the selected models for traffic

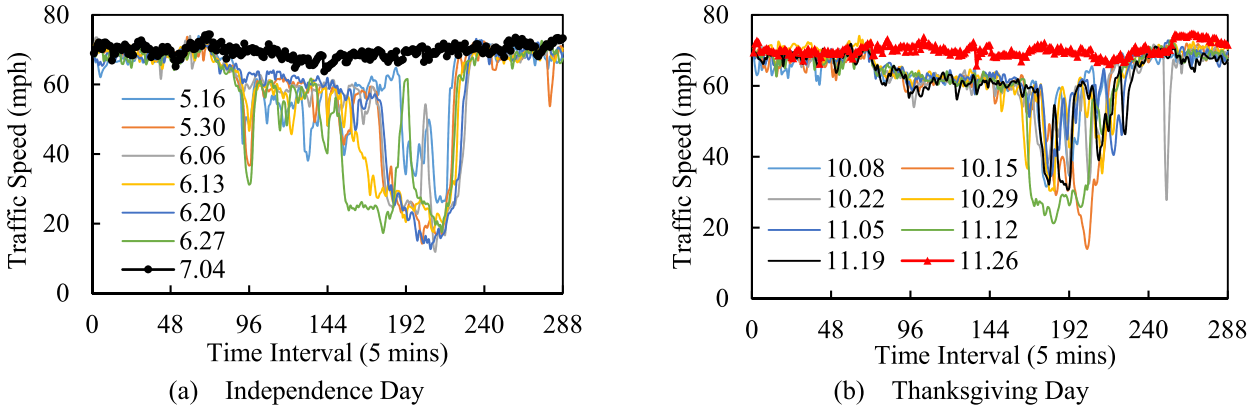


FIGURE 5. Variations of traffic speeds on thursdays.

TABLE 3. Average performance indices obtained from each model for traffic volume prediction on holidays.

Performance indices	Model								
	KNN	CNN	LSTM	GRU	ConvLSTM	CNNs-LSTM	CL-CN-G	CL-CNG	G-CN-CL
MAPE (%)	11.96	9.94	9.06	9.36	9.25	9.65	9.20	9.15	9.08
MAE (veh/5mins)	27.40	20.87	21.11	21.65	20.80	21.14	**20.34**	**20.31**	**20.35**
RMSE (veh/5mins)	36.09	27.23	27.74	28.30	27.26	27.71	**26.60**	**26.53**	**26.51**

Note: Bold numbers indicate that the values of the proposed models are less than those of all the existing models.

TABLE 4. Average performance indices obtained from each model for traffic speed prediction on holidays.

Performance indices	Model								
	KNN	CNN	LSTM	GRU	ConvLSTM	CNNs-LSTM	CL-CN-G	CL-CNG	G-CN-CL
MAPE (%)	1.67	1.56	1.57	1.59	1.57	1.63	**1.33**	**1.50**	**1.24**
MAE (mph)	1.16	1.09	1.09	1.11	1.09	1.13	**0.93**	**1.04**	**0.86**
RMSE (mph)	1.50	1.41	1.43	1.42	1.42	1.47	**1.17**	**1.35**	**1.12**

Note: Bold numbers indicate that the values of the proposed models are less than those of all the existing models.

volume prediction on the Independence Day and Thanksgiving Day. The bold digitals still indicate that the values calculated by the three proposed models are smaller than those obtained from the existing models. It can be found that: (1) the CL-CN-G, CL-CNG and G-CN-CL models outperform the best existing model with the improvement of 2.16% and 2.31% by MAE and RMSE at least; (2) the values of MAPE obtained from the three proposed models are smaller than those calculated by the existing models except the LSTM model. These analyses reveal that the three proposed models are effective.

Table 4 shows the average performance indices obtained from each model for traffic speed prediction on the Independence Day and Thanksgiving Day. It can be seen that: (1) the CL-CN-G model outperforms the best existing model with the improvement of 14.74%, 14.68% and 17.02% by MAPE, MAE and RMSE; (2) the CL-CNG model outperforms the best existing model with the improvement of 3.84%, 4.59% and 4.26% by MAPE, MAE and RMSE; (3) the G-CN-CL model outperforms the best existing model with the improvement of 20.51%, 21.10% and 20.57% by MAPE, MAE and RMSE. The results show that the three proposed models perform better than the existing models when predicting traffic speed, and the performance of the G-CN-CL model is the best.

In summary, the above discussions reveal that: (1) the CL-CN-G, CL-CNG and G-CN-CL models perform better than the existing models when predicting traffic flow on holidays; (2) the existing deep learning models have similar performance on holidays, and the KNN model performs the worst.

C. TRAFFIC FLOW PREDICTION UNDER ADVERSE WEATHER

Since drivers' travel behaviors will be affected by adverse weather, the characteristics of traffic flow under adverse

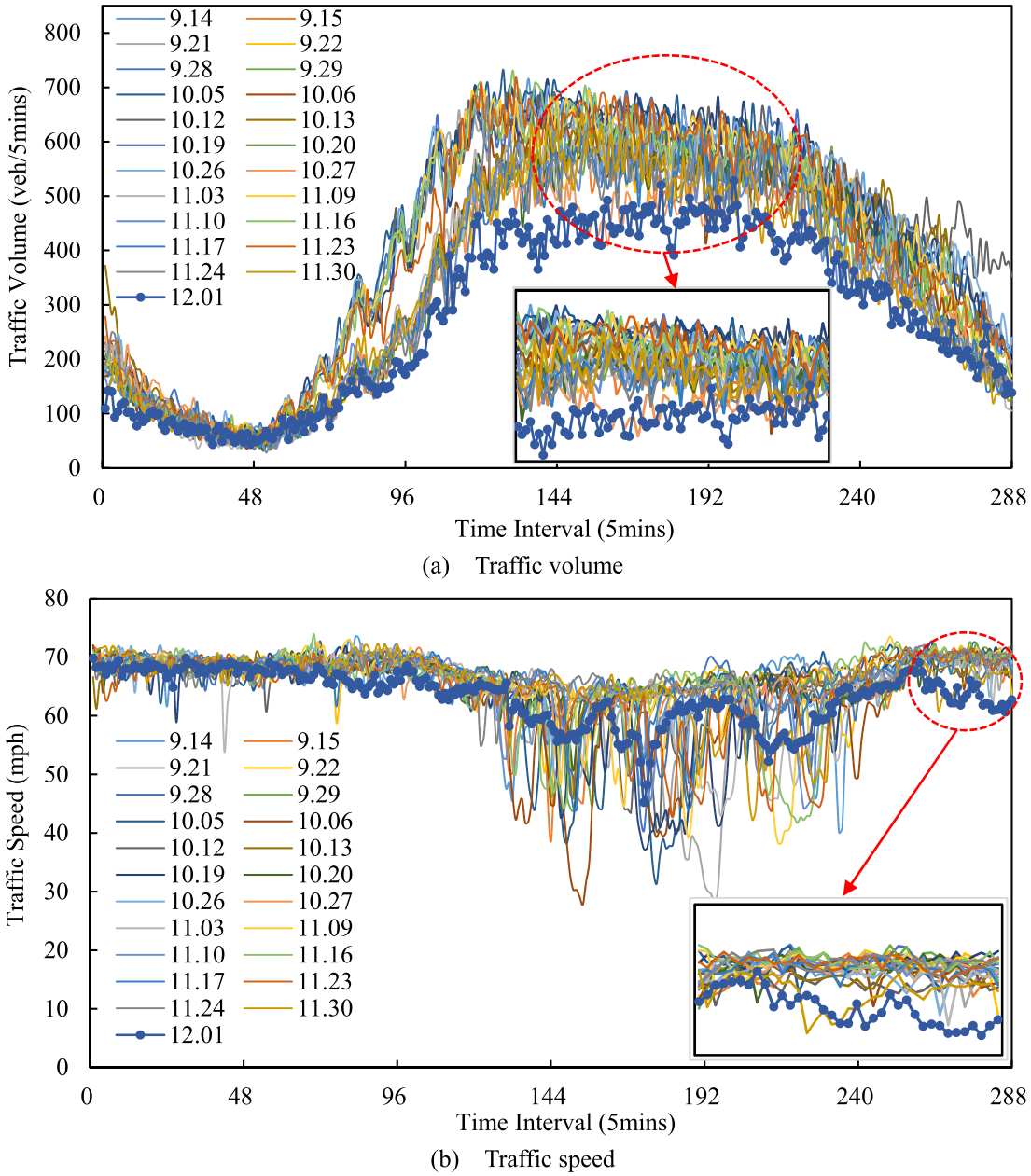


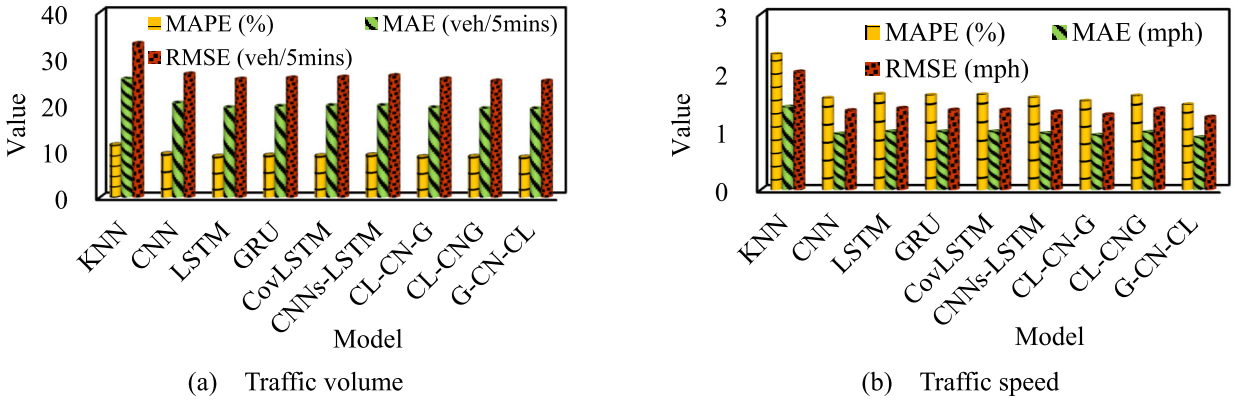
FIGURE 6. Variations of traffic flow data on weekends in case three.

weather will be different from those under normal conditions. In this case, detectors 311974 and 311903 shown in Fig. 2 are still selected, and the traffic flow data under heavy rain and strong wind were downloaded to test the proposed models. The weather, traffic volume and traffic speed data from August 19 to December 1 in 2019 are selected. The level of wind during 10:00 am to 13:00 pm on October 9 was 6 [56]. The rainfall on December 1 was 30.48 mm and bigger than 25 mm so that the weather type on December 1 was heavy rain [57]. The CL-CN-G, CL-CNG and G-CN-CL models are utilized to predict the traffic flow data on the two days again.

1) HEAVY RAIN

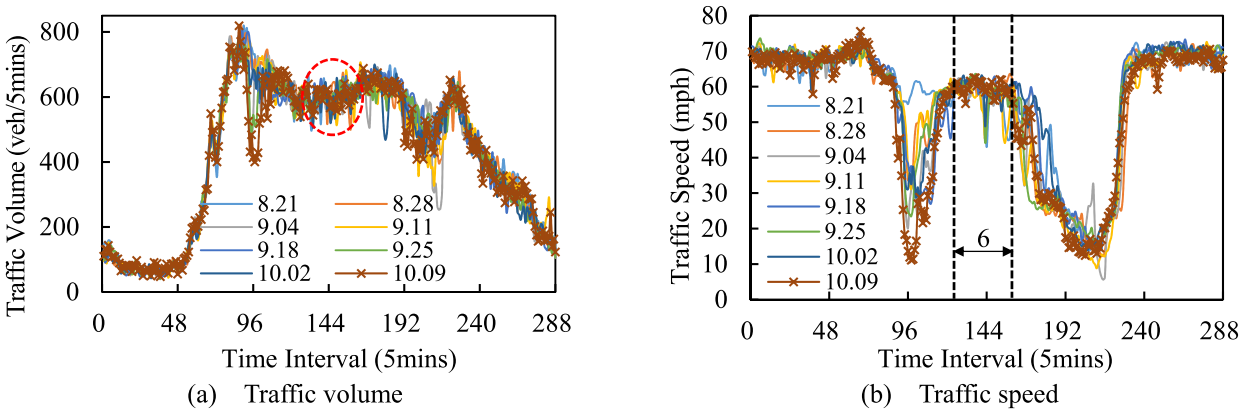
To predict the traffic volume and speed under heavy rain (the traffic flow data on December 1 in 2019), the traffic flow data from September 9 to November 31 are selected as the historical data. The traffic flow data on September 19, 20 and November 1, 2, 12, 13, 22 are removed because many data are missing on these days.

Fig. 6 shows the variations of traffic volumes and speeds on weekends from September 9 to December 1. From Fig. 6(a), it can be seen that: (1) the traffic volumes on weekends usually show the similar trends and have one distinct peak; (2) the amount of traffic volume under heavy rain is



(a) Traffic volume (b) Traffic speed

FIGURE 7. Performance indices obtained from each model under heavy rain.



(a) Traffic volume (b) Traffic speed

FIGURE 8. Variations of traffic volumes and traffic speeds on Wednesdays.

obviously smaller than that under normal conditions, especially in daytime (i.e., the traffic volumes in the red circle). Fig. 6(b) shows the variations of traffic speeds on weekends. It can be found that: (1) the traffic speeds in daytime are usually lower than those at night; (2) the fluctuations of traffic speeds under heavy rain are more stable than those under normal conditions; (3) the traffic speeds from 21:00 pm to 24:00 pm on December 1 (i.e., the traffic speeds in the red cycle) are clearly lower than those on other days, the reason is that it was raining hard during that period. On the whole, heavy rain has significant impacts on the patterns of traffic speed and volume. Thus, it is necessary to consider the impacts of adverse weather on traffic flow prediction.

In this section, the parameters and structures of the CL-CN-G, CL-CNG and G-CN-CL models are the same as those in Fig. 3, the input matrices of the global part are built using the temperature and traffic flow data. The number of time lags for the temperature data is set to 1. Fig. 7 displays the performance indices obtained from each model under heavy rain. It can be seen that: (1) the performance indices obtained from the CL-CN-G, CL-CNG and G-CN-CL models are always smaller than those calculated by the existing models; (2) the performance indices calculated by the KNN model are always the largest, and the performance indices obtained from the existing deep learning models are close. These discussions reveal that the CL-CN-G, CL-CNG and G-CN-CL models perform better than the existing models, which means that the three proposed models have excellent performance under heavy rain.

2) STRONG WIND

In this section, the traffic flow, wind speed and temperature data from August 19 to October 8 are selected to build the input matrices of the three proposed models. The number of time lags for the wind speed data and that for the temperature data are both set to 1. The structures and parameters of the CL-CN-G, CL-CNG and G-CN-CL models are also the same as those under heavy rain. Then, the traffic volume and speed under strong wind (i.e., the traffic flow data on October 9 in 2019) can be predicted based on the input matrices and models. Fig. 8 shows the variations of traffic volumes and speeds on the eight Wednesdays. The level of wind from 10:00 am to 13:00 pm on October 9 was 6. From Fig. 8(a), it can be found that the patterns of traffic volumes on these days are similar; the fluctuations of traffic volumes under

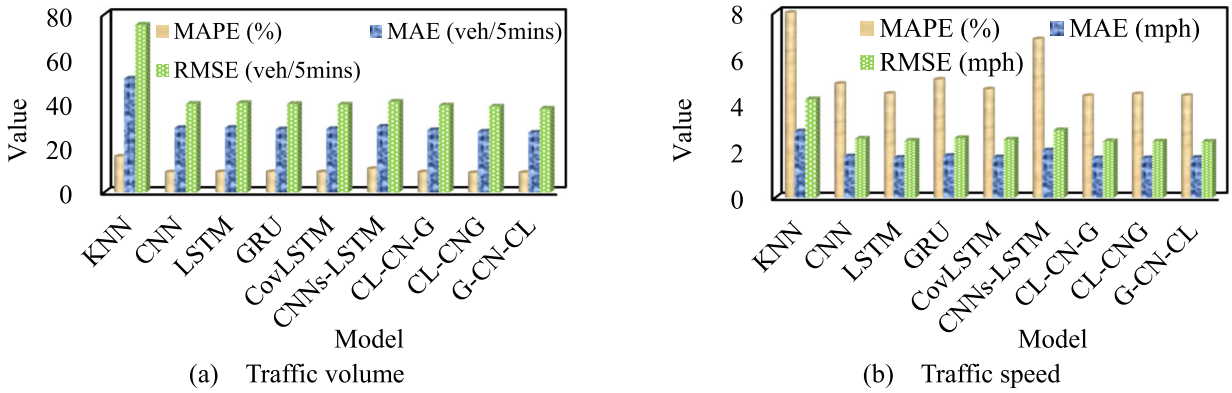


FIGURE 9. Performance indices obtained from each model under strong wind.

TABLE 5. Average performance indices obtained from each model for traffic volume prediction under adverse weather.

Performance indices	Model								
	KNN	CNN	LSTM	GRU	ConvLSTM	CNNs-LSTM	CL-CN-G	CL-CNG	G-CN-CL
MAPE (%)	13.73	9.23	8.99	9.12	9.03	9.86	**8.92**	**8.77**	**8.73**
MAE (veh/5mins)	38.32	24.66	24.21	23.98	24.14	24.68	**23.67**	**23.25**	**22.98**
RMSE (veh/5mins)	54.32	33.20	32.84	32.74	32.68	33.52	**32.32**	**31.84**	**31.34**

Note: Bold numbers indicate that the values of the proposed models are less than those of all the existing models.

TABLE 6. Average performance indices obtained from each model for traffic speed prediction under adverse weather.

Performance indices	Model								
	KNN	CNN	LSTM	GRU	ConvLSTM	CNNs-LSTM	CL-CN-G	CL-CNG	G-CN-CL
MAPE (%)	5.15	3.24	3.06	3.36	3.15	4.21	**2.95**	**3.04**	**2.93**
MAE (mph)	2.14	1.37	1.36	1.40	1.37	1.51	**1.32**	**1.35**	**1.31**
RMSE (mph)	3.13	1.95	1.93	1.97	1.94	2.12	**1.86**	**1.91**	**1.83**

Note: Bold numbers indicate that the values of the proposed models are less than those of all the existing models.

strong wind (i.e., the traffic volumes on October 9 in the red circle) are relatively smaller than those on other days. From Fig. 8(b), it can be seen that the variations of traffic speeds under strong wind are obviously more stable. Thus, strong wind also has impacts on the pattern of traffic flow.

Fig. 9 shows the performance indices obtained from each model under strong wind. It can be revealed that: (1) the performance indices obtained from the CL-CN-G, CL-CNG and G-CN-CL models are usually the smallest, which means that the three proposed models perform better than the existing models; (2) the performance indices calculated by the KNN model are always the largest, which indicates that the deep learning models perform better than the traditional machine learning model mentioned in this study; (3) the performance indices calculated by the existing deep learning models are usually close, and the performance indices obtained from the CNNs-LSTM model are sometimes obviously larger than those obtained from other existing deep learning models, which means that the performance of the CNNs-LSTM model is unstable under strong wind. These discussions reveal that the CL-CN-G, CL-CNG and G-CN-CL models show the excellent performance under strong wind.

To further analyze the performance of each model under adverse weather, Table 5 lists the average performance indices for traffic volume prediction under heavy rain and strong wind. From this figure, it can be found that: (1) the CL-CN-G model outperforms the best existing model with the improvement of 0.78%, 1.29% and 1.10% by MAPE, MAE and RMSE; (2) the CL-CNG model outperforms the best existing model with the improvement of 2.45%, 3.04% and 2.57% by MAPE, MAE and RMSE; (3) the G-CN-CL model outperforms the best existing model with the improvement of 2.89%, 4.17% and 4.10% by MAPE, MAE and RMSE. These analyses reveal that the three proposed models are effective for traffic volume prediction under adverse weather, and the G-CN-CL performs the best.

TABLE 7. Training time of each model.

Training time (s)		Model								
		KNN	CNN	LSTM	GRU	ConvLSTM	CNNs-LSTM	CL-CN-G	CL-CNG	G-CN-CL
Heavy rain	Volume	13.49	17.25	59.54	58.45	120.51	46.58	178.58	177.94	188.35
	Speed	13.30	20.13	68.15	65.93	130.43	50.36	206.70	172.74	201.60
Strong wind	Volume	20.80	30.05	106.74	101.78	213.75	83.17	150.70	142.04	330.86
	Speed	20.71	28.52	103.73	99.77	201.59	75.43	138.03	139.38	300.70

Table 6 shows the average performance indices obtained from each model for traffic speed prediction under heavy rain and strong wind. It can be seen that: (1) the CL-CN-G model outperforms the best existing model with the improvement of 3.59%, 2.94% and 3.63% by MAPE, MAE and RMSE; (2) the CL-CNG model outperforms the best existing model with the improvement of 0.65%, 0.74% and 1.04% by MAPE, MAE and RMSE; (3) the G-CN-CL model outperforms the best existing model with the improvement of 4.25%, 3.68% and 5.18% by MAPE, MAE and RMSE. These results also show that the three proposed models perform better than the existing models, and the G-CN-CL model still performs the best.

Table 7 lists the training time of each model under heavy rain and strong wind. It can be found that the training time of the KNN model is the least, and that of the ConvLSTM model and that of the proposed hybrid models are relatively longer. To sum up, the proposed hybrid models obtain the smallest performance indices, but the efficiency of the three models needs to be further improved.

V. CONCLUSION

Based on the CNN, GRU and ConvLSTM models, three hybrid deep spatio-temporal models (i.e., the CL-CN-G, CL-CNG and G-CN-CL models) are proposed with the concern of holidays and adverse weather. The three models consist of the global and target parts. Then, the input matrices of the global part are constructed by the traffic flow data and the weather data during all the periods, and the input matrices of the target part are established by the traffic flow data during the target period. To verify the accuracy of the three proposed models, three cases are designed, including traffic flow prediction under normal conditions, on holidays and under adverse weather. Additionally, the characteristics of traffic flow data under different situations are analyzed.

The experimental results show that: (1) the three proposed models perform better than the existing models mentioned in this study, and the G-CN-CL model performs the best; (2) in the three cases, the G-CN-CL model outperforms the best existing model with the improvement of –0.22% $\sim$ 2.89%, 2.16% $\sim$ 4.17% and 2.64% $\sim$ 4.10% by MAPE, MAE and RMSE for traffic volume prediction; (3) the G-CN-CL model outperforms the best existing model with the improvement of 4.07% $\sim$ 20.51%, 3.68% $\sim$ 21.10% and 4.04% $\sim$ 20.57% by MAPE, MAE and RMSE for traffic speed prediction; (4) on holidays, the traffic volume during peak hours will reduce, and the traffic speed in daytime will become larger; (5) the traffic volume in daytime on rainy day is obviously smaller than those on other days, and the traffic speed under heavy rain is lower than those on other days; under strong wind, the variations of traffic volume and speed are more stable. These results reveal that the accuracy and effectiveness of the CL-CN-G, CL-CNG and G-CN-CL models are good under all the situations. Because the three proposed models are all composed of the global and target parts, the spatio-temporal characteristics of the historical data under each situation can be further extracted, and the similarities of traffic flow between the target period and other periods can be captured. In addition, holidays, heavy rain and strong wind all have significant impacts on the characteristics of traffic flow data. In practice, not only the traffic flow data during target periods but also those during other periods should be used when predicting traffic flow on holidays or under adverse weather. The impacts of holidays and adverse weather should be fully considered, so that the prediction accuracy can be further improved. To get more accurate prediction results under all conditions, it is suggested to adopt the three proposed models, especially the G-CN-CL model.

The performance of the three proposed models needs to be further verified using traffic flow data under fog and other types of adverse weather. Also, more types of data (e.g. social media data) related to traffic flow should be concerned. To focus on traffic flow prediction in a large-scale network, the graph neural network will be introduced to improve the proposed models.

ACKNOWLEDGMENT

The data used to support the findings of this study are available from http://pems.dot.ca.gov and https://mesowest.utah.edu/.

REFERENCES

[1] J. Zeng, Y. Qian, F. Yin, L. Zhu, and D. Xu, "A multi-value cellular automata model for multi-lane traffic flow under Lagrange coordinate," *Comput. Math. Org. Theory*, early access, Oct. 2021, doi: 10.1007/s10588-021-09345-w.

[2] X. L. Zhang and H. P. Lu, "Non-parametric regression and application for short-term traffic flow forecasting," *J. Tsinghua Univ., Sci. Technol.*, vol. 49, no. 9, pp. 1471–1475, 2009.

[3] J. Wang and Q. Shi, "Short-term traffic speed forecasting hybrid model based on chaos-wavelet analysis-support vector machine theory," *Transp. Res. C, Emerg. Technol.*, vol. 27, pp. 219–232, Feb. 2013.

[4] W. Li, X. J. Ban, J. Zheng, H. X. Liu, C. Gong, and Y. Li, "Real-time movement-based traffic volume prediction at signalized intersections," *J. Transp. Eng. A, Syst.*, vol. 146, no. 8, Aug. 2020, Art. no. 04020081.
[5] H. Zhu, Y. Xie, W. He, C. Sun, K. Zhu, G. Zhou, and N. Ma, "A novel traffic flow forecasting method based on RNN-GCN and BRB," *J. Adv. Transp.*, vol. 2020, Oct. 2020, Art. no. 7586154.
[6] X. Luo, D. Li, and S. Zhang, "Traffic flow prediction during the holidays based on DFT and SVR," *J. Sensors*, vol. 2019, Jan. 2019, Art. no. 6461450.
[7] G. R. Walther, E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J.-M. Fromentin, O. Hoegh-Guldberg, and F. Bairlein, "Ecological responses to recent climate change," *Nature*, vol. 416, no. 6879, pp. 389–395, 2002.
[8] Y. Y. Pan, Z. Y. Zhu, P. Y. Shen, T. T. Gu, and W. W. Zhang, "Highway traffic accident features and analysis of effect and adverse weather conditions in Zhejiang province," *Highway*, vol. 58, no. 12, pp. 157–160, 2013.
[9] X. J. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Proc. 29th Annu. Conf. Neural Inf. Process. Syst.*, Montreal, QC, Canada, 2015, pp. 802–810.
[10] R. Yao, W. Zhang, and D. Zhang, "Period division-based Markov models for short-term traffic flow prediction," *IEEE Access*, vol. 8, pp. 178345–178359, 2020.
[11] Y. Qi and S. Ishak, "A hidden Markov model for short term prediction of traffic conditions on freeways," *Transp. Res. C, Emerg. Technol.*, vol. 43, pp. 95–111, Jun. 2014.
[12] C. Han, S. Song, and C. H. Wang, "A real-time short-term traffic flow adaptive forecasting method based on ARIMA model," *J. Syst. Simul.*, vol. 16, no. 7, pp. 1530–1532 and 1535, 2004.
[13] S. V. Kumar and L. Vanajakshi, "Short-term traffic flow prediction using seasonal ARIMA model with limited input data," *Eur. Transp. Res. Rev.*, vol. 7, no. 3, Sep. 2015, Art. no. 21.
[14] N. Huan, Q. Xie, H. X. Ye, and E. J. Ya, "Real-time forecasting of urban rail transit ridership at the station level based on improved KNN algorithm," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 18, no. 5, pp. 121–128, 2018.
[15] D. Xu, Y. Wang, P. Peng, S. Beilun, Z. Deng, and H. Guo, "Real-time road traffic state prediction based on kernel-KNN," *Transportmetrica A, Transp. Sci.*, vol. 16, no. 1, pp. 104–118, Dec. 2020.
[16] W.-C. Hong, "Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm," *Neurocomputing*, vol. 74, nos. 12–13, pp. 2096–2107, 2011.
[17] J. Tang, X. Chen, Z. Hu, F. Zong, C. Han, and L. Li, "Traffic flow prediction based on combination of support vector machine and data denoising schemes," *Phys. A, Stat. Mech. Appl.*, vol. 534, Nov. 2019, Art. no. 120642.
[18] W. Li, Y. Ji, and T. Wang, "Adaptive real-time prediction model for short-term traffic flow uncertainty," *J. Transp. Eng. A, Syst.*, vol. 146, no. 8, Aug. 2020, Art. no. 04020075.
[19] B. Yu, H. Wang, W. Shan, and B. Yao, "Prediction of bus travel time using random forests based on near neighbors," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 33, no. 4, pp. 333–350, 2018.
[20] W. S. Zhang, Z.-Q. Hao, J.-J. Zhu, T.-T. Du, and H.-M. Hao, "BP neural network model for short-time traffic flow forecasting based on transformed grey wolf optimizer algorithm," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 20, no. 2, pp. 196–203, 2020.
[21] Y. S. Qian, U.-W. Zeng, S.-F. Zhang, D.-J. Xu, and X.-T. Wei, "Short-term traffic prediction based on genetic algorithm improved neural network," *Tehnicki Vjesnik, Tech. Gazette*, vol. 27, no. 4, pp. 1270–1276, 2020.
[22] Y. Zhang and G. Huang, "Traffic flow prediction model based on deep belief network and genetic algorithm," *IET Intell. Transp. Syst.*, vol. 12, no. 6, pp. 533–541, 2018.
[23] C. Chen, H. Wang, F. Yuan, H. Jia, and B. Yao, "Bus travel time prediction based on deep belief network with back-propagation," *Neural Comput. Appl.*, vol. 32, no. 14, pp. 10435–10449, Jul. 2020.
[24] M. K. Hosseini and A. Talebpour, "Traffic prediction using time-space diagram: A convolutional neural network approach," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2673, no. 7, pp. 425–435, Jul. 2019.
[25] W. Zhang, Y. Yu, Y. Qi, F. Shu, and Y. Wang, "Short-term traffic flow prediction based on spatio-temporal analysis and CNN deep learning," *Transportmetrica A, Transp. Sci.*, vol. 15, no. 2, pp. 1688–1711, 2019.
[26] Y. Tian and L. Pan, "Predicting short-term traffic flow by long short-term memory recurrent neural network," in *Proc. IEEE Int. Conf. Smart City/SocialCom/SustainCom (SmartCity)*, Chengdu, China, Dec. 2015, pp. 153–158.
[27] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu, "LSTM network: A deep learning approach for short-term traffic forecast," *IET Intell. Transp. Syst.*, vol. 11, no. 2, pp. 68–75, 2017.
[28] X. Chen, H. Chen, Y. Yang, H. Wu, W. Zhang, J. Zhao, and Y. Xiong, "Traffic flow prediction by an ensemble framework with data denoising and deep learning model," *Phys. A, Stat. Mech. Appl.*, vol. 565, Mar. 2021, Art. no. 125574.
[29] Y. Zhang, T. Cheng, and Y. Ren, "A graph deep learning method for short-term traffic forecasting on large road networks," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 34, no. 10, pp. 877–896, Oct. 2019.
[30] Z. Y. Cui, K. Henrickson, R. Ke, and Y. Wang, "Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 11, pp. 4883–4894, Nov. 2020.
[31] J. Zhang, F. Chen, Y. Guo, and X. Li, "Multi-graph convolutional network for short-term passenger flow forecasting in urban rail transit," *IET Intell. Transp. Syst.*, vol. 14, no. 10, pp. 1210–1217, Oct. 2020.
[32] Z. Li, C. Lu, Y. Yi, and J. Gong, "A hierarchical framework for interactive behaviour prediction of heterogeneous traffic participants based on graph neural network," *IEEE Trans. Intell. Transp. Syst.*, early access, Jun. 29, 2021, doi: 10.1109/TITS.2021.3090851.
[33] Y. W. Zhao, Y. J. Chen, and W. Guan, "Prediction model of ETC short term traffic flow based on multidimensional time series," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 16, no. 4, pp. 191–198, 2016.
[34] W. Qiao, A. Haghani, C.-F. Shao, and J. Liu, "Freeway path travel time prediction based on heterogeneous traffic data through nonparametric model," *J. Intell. Transp. Syst.*, vol. 20, no. 5, pp. 438–448, Sep. 2016.
[35] R. Tanimura, A. Hiromori, T. Umedu, H. Yamaguchi, and T. Higashino, "Prediction of deceleration amount of vehicle speed in snowy urban roads using weather information and traffic data," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Gran Canaria, Spain, Sep. 2015, pp. 2268–2273.
[36] L. Liu and R.-C. Chen, "A novel passenger flow prediction model using deep learning methods," *Transp. Res. C, Emerg. Technol.*, vol. 84, pp. 74–91, Nov. 2017.
[37] Q. Chen, W. Wang, X. Huang, and H.-N. Liang, "Attention-based recurrent neural network for traffic flow prediction," *J. Internet Technol.*, vol. 21, no. 3, pp. 831–839, 2020.
[38] D. Zhang and M. R. Kabuka, "Combining weather condition data to predict traffic flow: A GRU-based deep learning approach," *IET Intell. Transp. Syst.*, vol. 12, no. 7, pp. 578–585, 2018.
[39] J. Ke, H. Zheng, H. Yang, and X. M. Chen, "Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach," *J. Transp. Res. C, Emerg. Technol.*, vol. 85, pp. 591–608, Dec. 2017.
[40] C. Zheng, X. Fan, C. Wen, L. Chen, C. Wang, and J. Li, "DeepSTD: Mining spatio-temporal disturbances of multiple context factors for citywide traffic flow prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 9, pp. 3744–3755, Sep. 2020.
[41] L. Bai, "Urban rail transit normal and abnormal short-term passenger flow forecasting method," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 17, no. 1, pp. 127–135, 2017.
[42] B. Xie, Y. Sun, X. Huang, L. Yu, and G. Xu, "Travel characteristics analysis and passenger flow prediction of intercity shuttles in the pearl river delta on holidays," *Sustainability*, vol. 12, no. 18, Sep. 2020, Art. no. 7249.
[43] X. F. Ji and Y. C. Ge, "Holiday highway traffic flow prediction method based on deep learning," *J. Syst. Simul.*, vol. 32, no. 6, pp. 1164–1171, 2020.
[44] X. H. Zhao, G. C. Ren, C. Chen, and J. Rong, "A review on driving behavior under adverse weather conditions," *J. Transp. Inf. Saf.*, vol. 35, no. 5, pp. 70–75 and 98, 2017.
[45] H. Y. Sun, I. S. Yang, L. B. Li, and B. Wu, "Fuzzy neural network system for urban expressway speed prediction on rainy day," *J. Tongji Univ., Natural Sci.*, vol. 44, no. 11, pp. 1695–1701, 2016.
[46] X. J. Xu, Y. L. Bai, L. Xu, Z. Z. Xu, and X. W. Zhao, "Traffic flow prediction based on random forest in severe weather conditions," *J. Shaanxi Normal Univ., Natural Sci. Ed.*, vol. 48, no. 2, pp. 25–31, 2020.
[47] F. Xue, E. Yao, N. Huan, B. Li, and S. Liu, "Prediction of urban rail transit ridership under rainfall weather conditions," *J. Transp. Eng. A, Syst.*, vol. 146, no. 7, Jul. 2020, Art. no. 04020061.

[48] T. Li, J. Ma, and C. Lee, "Markov-based time series modeling framework for traffic-network state prediction under various external conditions," *J. Transp. Eng. A, Syst.*, vol. 146, no. 6, Jun. 2020, Art. no. 04020042.
[49] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transp. Res. C, Emerg. Technol.*, vol. 79, pp. 1–17, Jun. 2017.
[50] A. Shabarek, S. Chien, and S. Hadri, "Deep learning framework for freeway speed prediction in adverse weather," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2674, no. 10, pp. 28–41, Oct. 2020.
[51] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang, "Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction," *Sensors*, vol. 17, no. 4, Apr. 2017, Art. no. 818.
[52] J. Zhao, Y. Gao, Y. Qu, H. Yin, Y. Liu, and H. Sun, "Travel time prediction: Based on gated recurrent unit method and data fusion," *IEEE Access*, vol. 6, pp. 70463–70472, 2018.
[53] D. Ma, B. Sheng, S. Jin, X. Ma, and P. Gao, "Short-term traffic flow forecasting by selecting appropriate predictions based on pattern matching," *IEEE Access*, vol. 6, pp. 75629–75638, 2018.
[54] Y. Qiao, Y. Wang, C. Ma, and J. Yang, "Short-term traffic flow prediction based on 1DCNN-LSTM neural network structure," *Mod. Phys. Lett. B*, vol. 35, no. 2, Jan. 2021, Art. no. 2150042.
[55] W. S. Zhang and R. H. Yao, "Traffic flow parameters estimation based on spatio-temporal characteristics and hybrid deep learning," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 21, no. 1, pp. 82–89, 2020.
[56] *Wind Scale*, Standard GB/T 28591-2012, 2012.
[57] *Grade of Precipitation*, Standard GB/T 28592-2012, 2012.

WENSONG ZHANG is currently pursuing the Ph.D. degree with the School of Transportation and Logistics, Dalian University of Technology, Dalian, China. His research interests include traffic flow prediction and its applications in urban traffic signal control.

RONGHAN YAO received the Ph.D. degree from the College of Transportation, Jilin University, Changchun, China, in 2007. She is currently an Associate Professor at the School of Transportation and Logistics, Dalian University of Technology, Dalian, China. Her research interests include intermittent traffic flow theory, urban traffic management and control, and traffic demand and behavior prediction.

XIAOJING DU is currently pursuing the Ph.D. degree with the School of Transportation and Logistics, Dalian University of Technology, Dalian, China. Her research interests include game theoretic analysis of driving behavior and its application in urban traffic planning and management.

JINSONG YE received the master's degree from the College of Information, Beijing Forestry University, Beijing, China, in 2005. He is currently a Senior Engineer at the Information Center of the China Academy of Transportation Sciences, Beijing, China. His research interests include traffic big data governance, mining and application, and statistical and analysis of expressway transportation volume.