

## Research Article

# Physical Model versus Artificial Neural Network (ANN) Model: A Comparative Study on Modeling Car-Following Behavior at Signalized Intersections

Lan Yang <sup>1</sup>, Shan Fang <sup>1</sup>, Guoyuan Wu <sup>2</sup>, He Sheng,<sup>1</sup> Zhigang Xu <sup>1</sup>, Mengxiao Zhang,<sup>1</sup> and Xiangmo Zhao<sup>1</sup>

<sup>1</sup>School of Information Engineering, Chang'an University, Xi'an, Shaanxi, China

<sup>2</sup>Center for Environmental Research and Technology, University of California at Riverside, Riverside, CA, USA

Correspondence should be addressed to Shan Fang; [fang6100146@gmail.com](mailto:fang6100146@gmail.com)

Received 24 November 2021; Revised 1 March 2022; Accepted 3 March 2022; Published 25 April 2022

Academic Editor: Xianyu Wu

Copyright © 2022 Lan Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Many studies have simulated traffic behavior at signalized intersections using various Car-Following (CF) models. However, the performance of which CF Model is superior at signalized intersections has not been thoroughly analyzed and evaluated. In this study, two novel Artificial Neural Network (ANN) CF models, the Convolutional Neural Network—Long Short-term Memory (CNN-LSTM) and the Convolution-LSTM (Conv-LSTM)—are first applied to predict CF behaviors at signalized intersections. Both models can extract spatial and temporal information to address the long-term dependency problem more effectively. Based on the filtered NGSIM dataset, we conduct a comparative empirical study of three conventional CF models and five ANN CF models. The dataset is divided into two categories based on the characteristics of CF behavior at signalized intersections: continuous and discontinuous. The experiments demonstrated that ANN CF models outperformed conventional CF models when the output was the velocity in two categories of traffic flow but only failed to do so when the output was acceleration in discontinuous traffic flow. The proposed models were capable of accurately predicting acceleration, but the traffic fluctuations also existed as time passed. Additionally, it was discovered that while the ANN CF model is preferable for traffic flow simulation, the conventional CF model still cannot be ignored for discontinuous traffic flow simulation, particularly when acceleration is required.

## 1. Introduction

With the development of Intelligent Transportation System (ITS), the multitransport models which include human-driven vehicles, Connected and Automated vehicles (CAVs) have become the usual pattern of urban traffic flow research. The response mechanism of CAVs is more rapid and accurately compared with human-driven vehicles, so how to guide the CAVs in different categories of traffic flow to steadily follow the preceding vehicle is significant [1]. In comparison to the Car-Following (CF) behavior on the highway, the CF behavior at signalized intersections is more unstable. Drivers are frequently caught in a bind and unable to make quick decisions due to traffic signal light control.

Thus, the CF model at signalized intersections is a thriving field of research. Numerous studies have used various conventional CF models to simulate traffic flow at signalized intersections and to analyze the time-varying characteristics of single-lane traffic dynamics.

Several studies [2–6] simulated the trajectory of a human-driven vehicle at intersections using the Intelligent Driver Model (IDM). Han et al. and Yao et al. [7, 8] used the Gipps CF model to simulate human-driven vehicles crossing the intersection. Zhao et al. [9] enhanced the full velocity difference (FVD) model to better describe the CF behavior of an approach lane with a special width at an intersection. Zhang et al. [10] studied the CF behavior at intersections quantitatively and proposed a new CF model based on the FVD.

Tang et al. [11] proposed a method for speed guidance at intersections that used the FVD model as an input. Zhao et al. [12] and Yu and Shi [13] combined a new CF model of traffic flow at intersections with a basic FVD model based on multiple datasets to propose a new CF model. Liu et al. [14] improved the FVD model's stability and factored in the effect of short-term driving memory. Arvin et al. [15] used the Wiedemann CF model to evaluate the safety of connected and automated vehicles at intersections. Liu et al. [16] proposed an eco-speed guidance method for the mixed traffic flow of electric and conventional human-driven vehicles at an intersection, simulating CF behavior using the Wiedemann model. On the other hand, these conventional CF models are incapable of accurately simulating sophisticated vehicle movement at signalized intersections. Brockfeld et al. [17] stated that the error of conventional CF models is between 15% and 25%, which is difficult to reduce further. Albeaik et al. [18] found that the velocity that is simulated by IDM is unrealistic for some specific initial data and weakened the generalization ability of the model. With the advancement of artificial intelligence and data collection technologies, data-driven CF models, particularly Artificial Neural Network (ANN) CF models, are gradually emerging as a new area of research. The advantage of ANN CF models is that they do not depend on theoretical assumptions and do not strictly adhere to any mathematical derivation. Instead, they use a nonparametric approach to extract the inherent information in traffic data and build highly accurate CF models.

While several studies have explored the application of CF models at signalized intersections, few have examined the robustness of various CF models. The conventional CF models and the ANN CF models will be evaluated in this study using the NGSIM ground-truth dataset. Two improved ANN CF models are proposed to address the limitations of existing CF models when confronted with sophisticated traffic flows. Finally, we conduct simulations to determine which CF models perform the best under complicated traffic scenarios at signalized intersections. The following sections highlight the paper's main contributions:

- (1) The CNN-LSTM CF model is proposed, which can use Convolutional Neural Networks (CNN) to extract key features of each time step's CF behavior and the Long Short-Term Memory (LSTM) model to predict the CF behavior of future time steps.
- (2) The Conv-LSTM model is first applied to the simulation of CF behavior, as it can extract spatial and temporal information from multiple time steps in a single LSTM cell, thereby resolving the long-term dependency problem more effectively.
- (3) Three conventional CF models, five ANN CF models, and two machine learning CF models are compared at signalized intersections for two typical traffic flow categories. The acceleration, velocity, trajectory, and hysteresis loops, as well as the RMSE (Root Mean Square Error), are used to quantify the performance of all CF models in predicting the following vehicle's movement state.

The remainder of this study is organized as follows. Section 2 reviews the existing CF models and their application in traffic modeling. Section 3 preprocesses the training and cross-validation dataset. Section 4 analyzes three representative ANN CF models and proposes two novel ANN CF models. Section 5 analyzes the conventional CF models and calibrates their parameters using ground truth data. Section 6 explores the performance of different types of CF models in continuous/discontinuous traffic flow where the acceleration and the velocity are used as single model output, respectively.

## 2. Literature Review

*2.1. Conventional CF Models.* Pipes presented the first CF models in 1953 [19], and various CF models have been proposed since then. The conventional CF model can be classified into four broad categories: stimulus-response, safety distance, psycho-physical, and optimal velocity (OV). The stimulus-response model focuses on the reaction of subject vehicles, and notable stimulus-response models include the Gazis-Herman-Rothery (GHR) model [20], the Newell model [21], and the IDM model [22]. Wu et al. [23] used the IDM model to validate the proposed variable speed limit control strategy. Cui et al. [24] used the IDM to investigate the fuel consumption and emission characteristics of adaptive cruise control/cooperative adaptive cruise control at a signalized intersection. Gipps [25] proposed the safety distance model regarded the gap between continuous vehicles as the most significant parameter. Yao and Li [5] simulated the trajectories of human-driven vehicles using the Gipps model. Additionally, the psycho-physical model is a decision-making model that employs various thresholds to denote the various driving stages of CF behavior. For instance, the Wiedemann model [26] presupposes that the driver employs various driving strategies depending on the traffic environment. Wang et al. [27] determined expected vehicle behaviors using the Wiedemann model. Chauhan et al. [28] calibrated the Wiedemann model's parameters using an Indian naturalistic dataset. Bando et al. [29] developed the optimal velocity model (OVM), and Jiang et al. [30] incorporated the effect of the velocity difference into the OVM model and proposed Full Velocity Model (FVD). Zhao et al. [31] used an OVM to model human vehicle behavior. Qin et al. [32] used OVM to analyze the stability of connected and automated vehicles. Although these conventional CF models are useful for studying the dynamics of micro traffic flows, they have some limitations. For instance, the robustness of the conventional CF model is insufficient because the model's parameters do not account for all traffic-influencing factors.

*2.2. Data-Driven CF Models.* Data-driven methods include fuzzy logic, case paper, support vector regression, and ANN. The fuzzy logic method applies fuzzy sets and fuzzy rules to a problem to describe it, and it can be used to perform comprehensive analyses of unknown models and systems

[33]. The addition of fuzzy logic to the CF model improves the model's ability to describe the driver's CF behavior [34]. The case paper method models the CF behavior using locally weighted regression or the k-nearest neighbor method. Toledo et al. [35] constructed a CF model using a locally weighted regression method and fitted the position of each vehicle sequentially using the weighted least squares method. Machine Learning (ML) is one of the most significant branches of data driven research. He et al. [36] proposed a data-driven CF model based on the K-Nearest Neighbor (KNN) algorithm, eliminating the need for complicated mathematical formulations or parameter correction. Furthermore, Yu et al. [37] proposed the novel CF model based Fixed Radius Nearest Neighbor (FRNN) with a consideration to balance accuracy and compute efficiency. The simulation experiments proved that the FRNN perform better than the KNN CF model. Support vector regression is a technique based on the support vector machine (SVM); it augments the SVM with a regression algorithm. Lu et al. [38] developed a CF model based on the support vector regression CF (SVRCF) model and calibrated the SVRCF model's parameters. The experiments showed that the performance of the SVRCF model was further improved compared with that of the conventional CF models.

Whereas fuzzy logic, case papers, and SVM models can more precisely simulate CF behavior, they performed worse than neural network models when multisource data fusion is performed. A neural network-based CF model includes a simple network, such as a multilayer perceptron (MLP) or radial basis function (RBF) network. Other models are built using deep learning techniques such as recurrent neural networks (RNNs) and long short-term memory (LSTM). A few researchers have conducted prospective studies on these models. For example, Jia et al. [39] proposed a CF model based on a BP neural network. Khodayari et al. [40] developed a neural network CF model that considered the drivers' behaviors. Following that, Zheng et al. [41] developed a CF model based on neural networks characterized by an instantaneous driver-vehicle response delay. Colombaroni and Fusco [42] proposed a CF model based on a feedforward neural network that balanced efficiency and complexity via a particle swarm optimization algorithm. Zhu et al. [43] developed a deep deterministic policy gradient CF model (DDPGvRT) based on deep reinforcement learning by training the designed neural network on Shanghai's driving data. DDPGvRT is more consistent with actual driving than IDM or the other CF models. To overcome the problem of selfish objectives of human-driven vehicles, Peng et al. [44] redesigned a novel altruistic reward function of the deep reinforcement learning to improve the traffic efficiency in the unsignalized intersection. Papathanasopoulou and Antoniou [45] proposed a locally optimal neural network CF model with a lower RMSE than other CF models, such as the Gipps model. Wang et al. [46] established an RNN CF model characterized by the previous time segment's input of vehicle information; the author then simulated the driver's driving behavior and improved the model's accuracy. Zhou et al. [47]

predicted traffic flow oscillations and distinguished driving behaviors using a neural network model. Ma and Qu [48] proposed a new CF model based on sequence to sequence, and the proposed model outperformed conventional models in the CF pair and platoon experiments. Hui et al. [49] developed a novel mixed deep encoder-decoder neural network to predict the vehicle trajectory and the optimal network structure was selected by different trajectory samples training. These models demonstrated the neural network's nonlinear regression capability and its ability to simulate CF behavior.

According to the discussion above, only a few CF models based on neural networks have been applied to traffic flow research. The CF model's performance at signalized intersections has not been thoroughly investigated. Existing ANN CF models are primarily concerned with highways and pay little attention to urban roads. To address these gaps, this study compares the performance of ANN CF models and conventional CF models at signalized intersections.

### 3. Data Preparation

The precision of datasets is important for CF modeling [50, 51]. Two datasets are used in this study to train and calibrate CF models of signalized intersections: the Lankershim dataset and the Peach Tree dataset, both of which are provided by the Federal Highway Administration's NGSIM program [52]. It extracted detailed information from the video data, including the vehicle's length, type, position, lane identification, velocity, and acceleration. The training and testing datasets were gathered from a section of Lankershim Boulevard in Los Angeles, California, USA, as illustrated in Figure 1(a). Five video cameras have been installed on a 36-story building adjacent to US Highway 101 and Lankershim Boulevard. On June 16, 2005, between 8:45 and 9:00 am, the vehicles' trajectories were recorded at a frequency of 10 Hz. The monitored area is approximately 1600 ft long and included four signalized intersections. As illustrated in Figure 1(b), this study's validation dataset is the Peachtree Street dataset that comprised of data collected on urban roads.

As illustrated red circle in Figure 2, these original datasets contain noise and redundant information, preventing them from directly applying to CF behavior modeling. However, both conventional CF models and ANN CF models require high-quality data. As a result, we filter the datasets using the four conditions listed in the metadata documentation.

- (i) Vehicle\_type = 2: It narrows the focus only on the CF behavior of the automobile while excluding motorcycles and trucks
- (ii) Following  $\neq 0$ : It filters out other vehicle behaviors, such as lane changing and overtaking
- (iii)  $5 \text{ ft} < \text{space-headway} < 130 \text{ ft}$ : It eliminates the crash and free-flow traffic data
- (iv) CF duration  $> 15 \text{ s}$ : It ensures that the CF duration is sufficient for the study

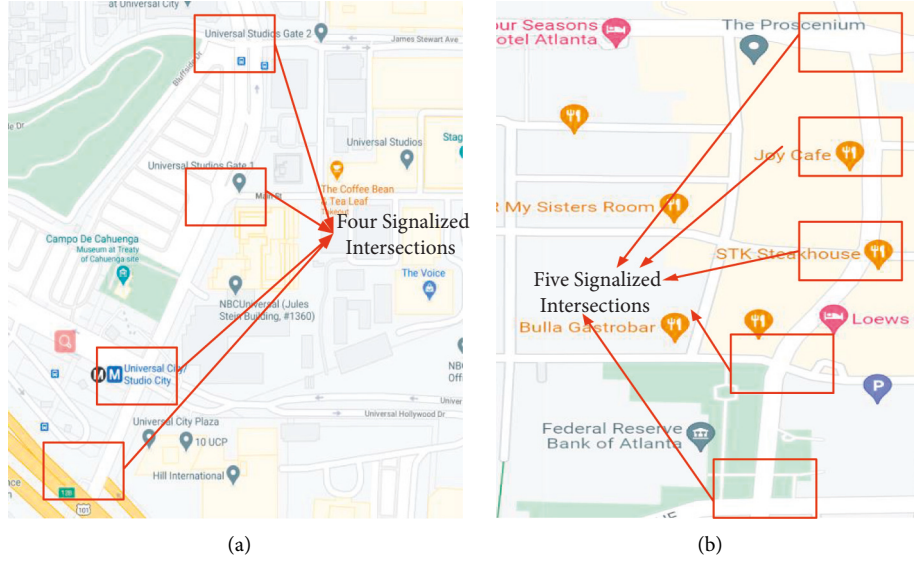


FIGURE 1: Schematic of the data collection area. (a) Lankershim Boulevard for data training. (b) Peachtree Street for data validation.

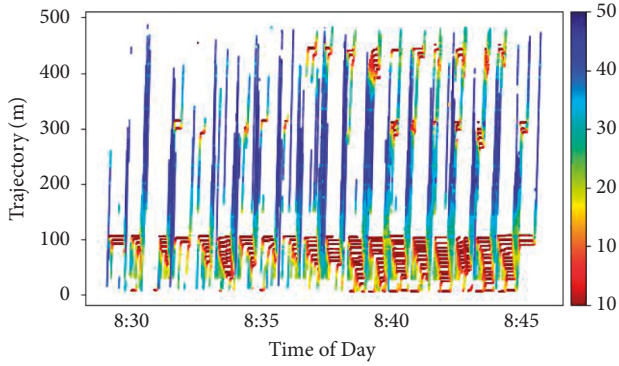


FIGURE 2: Trajectory of Lankershim Street, lane 2.

After data filtering, the filtered Lankershim dataset contained 377,000 data sample points: 90% (339,300) of the data were used in the training phase, whereas 10% (37,700) were used in the test phase. Cross-validation is performed on the filtered Peachtree dataset, which contains 83,000 data sample points.

Due to noise and abnormal data points, the datasets that satisfied the preceding four conditions are still unavailable. Montanino and Punzo [53] proposed a multistep filtering procedure to eliminate outliers from the dataset in order to address this issue. Coifman and Li [54] reextracted data from raw NGSIM video data via manual extraction. It indicated that the dataset contained errors, necessitating data filtering. The NGSIM dataset is filtered in this study using the moving average algorithm, and the filtered results are shown in Figure 3. In comparison to raw data, filtered velocity and acceleration data are smoother without sharp fluctuations. Additionally, the Pearson correlation coefficient ( $R$ ) is used to evaluate the consistency between input data and filter data. It can be seen from Table 1 that the filtered data have the high consistency with input data.

## 4. ANN CF Models

**4.1. Model Inputs and Outputs.** Several variables in the CF behavior influence the following vehicle's movement, especially the velocity, space headway, and velocity difference [36–43]. These variables are also used as inputs in this paper's ANN CF models. There are two options for determining the output of CF models. For example, some models' output is the acceleration [40, 48], while others' output is the velocity [41, 42, 45]. To examine the effect of varying the model's output, both acceleration and velocity are used as the output of ANN CF models, respectively. It can be formulated as follows:

$$\begin{aligned} \text{input}(t) &= (v_n(t), \Delta s_n(t), \Delta v_n(t)), \\ \text{output1}(t + \tau) &= v_n(t + \tau), \\ \text{output2}(t + \tau) &= a_n(t + \tau), \end{aligned} \quad (1)$$

where  $\tau$  is the simulation time step;  $v_n(t)$ ,  $\Delta s_n(t)$ ,  $\Delta v_n(t)$ , and  $a_n(t)$  represent the velocity, space headway, velocity difference, and acceleration of the following vehicle  $n$  at time  $t$ , respectively. Additionally, the ANN CF models that used acceleration as output are named ANN CF-A, and the ANN CF models that used velocity as output are named by ANN CF-V.

### 4.2. Architecture of ANN Models

**4.2.1. BP.** The BP neural network is a popular ANN algorithm used to solve classification and regression problems. Figure 4 illustrates the structure of the BP neural network used in this study.

As illustrated in Figure 4, the network contains several layers. The first layer is called the input layer, and it contains three input variables. The second and third layers are hidden

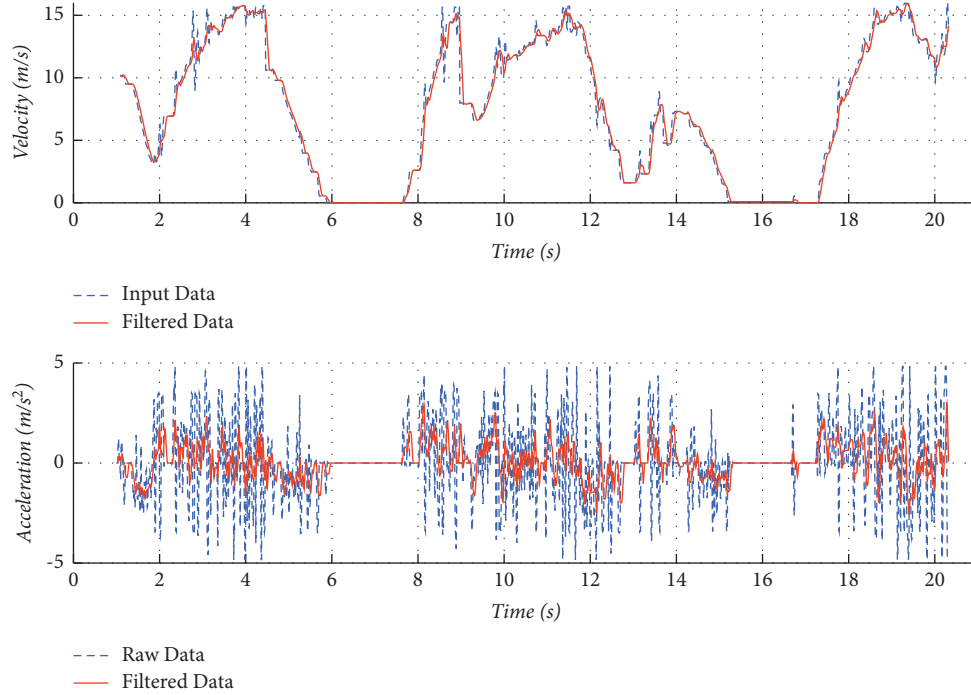


FIGURE 3: Comparison between the raw and filtered data.

TABLE 1: Results of data filtering.

| Values | Acceleration | Velocity | Space velocity | Space headway |
|--------|--------------|----------|----------------|---------------|
| $R$    | 0.996        | 0.998    | 0.991          | 0.995         |

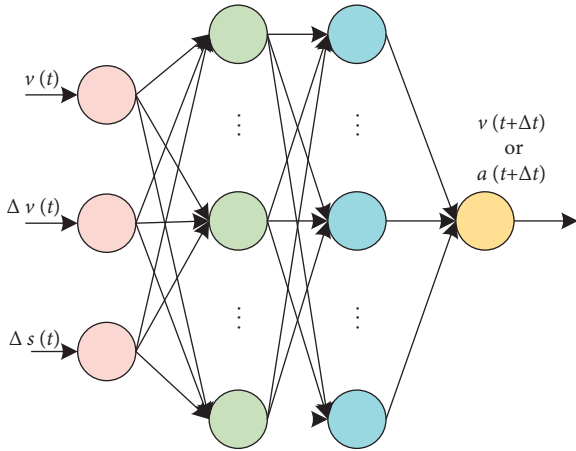


FIGURE 4: The structure of the BP neural network.

layers, whereas the fourth layer is an output layer that contains a single output variable. The selection of the number of hidden layers and hidden units is a difficult task, and in Section 4.3, we will provide the optimal selection strategy for each ANN CF model.

**4.2.2. RNN.** As shown in Figure 5,  $I$ ,  $S$ , and  $O$  represent the input, hidden, and output units, respectively.  $u$  is the weight matrix between the input and hidden layers;  $v$  is the weight

matrix between the hidden and output layers. The output layer is the fully connected layer to calculate the final value.  $W$  is the last value of the hidden layer as the weight of the input, and  $n$  refers to the number of data labels. Compared with feedforward neural networks, RNN saves the state of the training data at the final stage, thereby improving the network accuracy.

**4.2.3. LSTM.** The gradient vanishing or explosion problem may occur in a conventional feedforward neural network, limiting the applicability of such networks. Hochreiter and Schmidhuber [55] proposed the LSTM neural network to address these issues. In contrast to a standard RNN network unit, which contains only one hidden layer, a typical LSTM network unit (as illustrated in Figure 6) contains an input gate, an output gate, a forget gate, and a cell.  $h$  is the filtered output,  $c$  is the cell state, and  $I$  is the input vector.

**4.2.4. CNN-LSTM.** Convolutional neural networks are a subclass of deep learning networks that have achieved remarkable results in various research areas, especially computer vision, image classification, and speech recognition. A conventional CNN model consists of three layers: convolution, pooling, and fully connected. The key equation for the discrete convolutional operation in one dimension is as follows:

$$s(t) = (x * \omega)(t) = \sum_{a=-\infty}^{\infty} x(a)\omega(t-a), \quad (2)$$

where  $x$  is the input array,  $\omega$  is the one-dimension filter,  $a$  is the measurement step, and  $t$  is the current time. However,

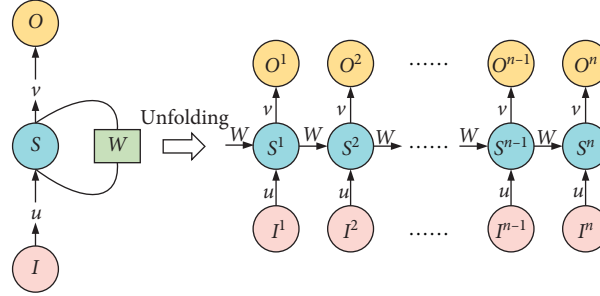


FIGURE 5: Structure of a recurrent neural network model.

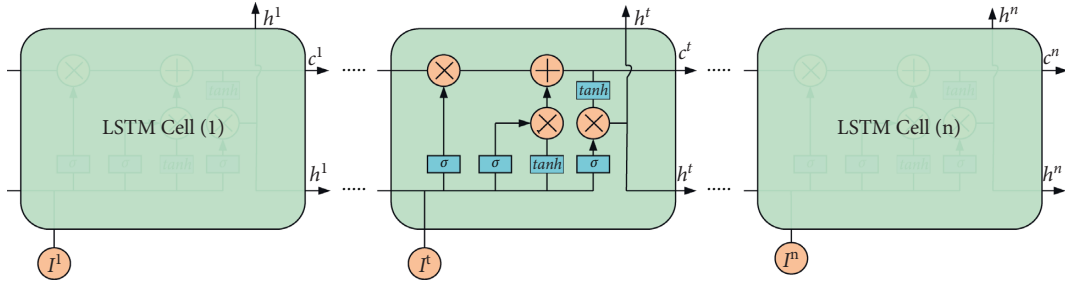


FIGURE 6: Structure of a long short-term memory model.

the conventional CNN model cannot process the time series problem well, such as subtitle generation from video.

To address this issue, a combination model called CNN-LSTM has been proposed that combines the advantages of CNN and LSTM, and the architecture of models is shown in Figure 7. Additionally, CNN-LSTM is used for the first time in the CF model study, which can use CNN to extract key features of each time step's CF behavior and predict the CF behavior of future time steps using an LSTM model. The CNN-LSTM architecture used in this study has several layers: an input layer, some convolutional layers, some LSTM layers, some fully connected layers, and an output layer.

**4.2.5. Convolution-LSTM.** While the convolution-LSTM can extract critical temporal information from the time series input array, it does not perform well in capturing spatial information. Shi et al. proposed the Conv-LSTM model and successfully applied it to convective precipitation prediction [56]. It is first applied to the simulation of CF behavior, which can extract spatial and temporal information from multiple time steps at a single Conv-LSTM cell, allowing for a more accurate solution to the long-term dependency problem. The primary formulation of Conv-LSTM is as follows:

$$\begin{aligned}
 i_t &= \sigma(W_{xi} * x_t + W_{hi} * h_{t-1} + W_{ci} \bullet c_{t-1} + b_i), \\
 f_t &= \sigma(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} \bullet c_{t-1} + b_f), \\
 c_t &= f_t \bullet c_{t-1} + i_t \bullet \tanh(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c), \\
 o_t &= (W_{xo} * x_t + W_{ho} * h_{t-1} + W_{co} \bullet c_t + b_o), \\
 h_t &= o_t \bullet \tanh(c_t),
 \end{aligned} \quad (3)$$

where  $i_t$ ,  $f_t$ , and  $o_t$  represents the input gate, forget gate, and output gate, respectively.  $c_t$  is the output of a single cell.  $W_{xi}$ ,  $W_{hi}$ ,  $W_{ci}$ ,  $W_{xf}$ ,  $W_{hf}$ ,  $W_{cf}$ ,  $W_{xc}$ ,  $W_{hc}$ ,  $W_{xo}$ ,  $W_{ho}$ , and  $W_{co}$  are weighted matrices, and  $b_i$ ,  $b_f$ ,  $b_c$ , and  $b_o$  refer to the bias of different gates.  $\sigma$  and  $\tanh$  are sigmoid activation function and hyperbolic tangent activation function. Additionally, asterisk "\*" is the convolutional operator, and " $\bullet$ " is the Hadamard product. The major difference between CNN-LSTM and Conv-LSTM is that in CNN-LSTM, the convolutional operation is performed as a single layer. In contrast, in Conv-LSTM, the convolutional operation is performed in each Conv-LSTM cell. Additionally, gate layers can be used to visualize the cell state, as indicated by the red line in Figure 8. The proposed model is composed of five layers: an input layer, a Conv-LSTM layer, one fully connected layer that are wrapped in time distribution layers, and an output layer.

**4.3. Models Training and Cross-Validation for ANN CF Models.** The Adam optimizer [57] (learning rate = 0.001), a stochastic gradient descent extension, was used to train the ANN CF model. The hyperparameter tuning is the tricky work, the different values of epoch, batch size, hidden layers, and cells are set to find the optimal model structure. The training and validations indicated that the accuracy of ANN car-following models cannot still be improved by increasing the hidden layers and the impact of different variables are independent. Therefore, a parameters sensitivity simulation experiment which included four mentioned variables was conducted, and the optimal parameters settings each model are shown in Table 2.

For the overfitting problem, the dropout layer is used on each fully connected layer preceding the output layer at a rate of 0.5 [58]. The validation dataset that was not used for

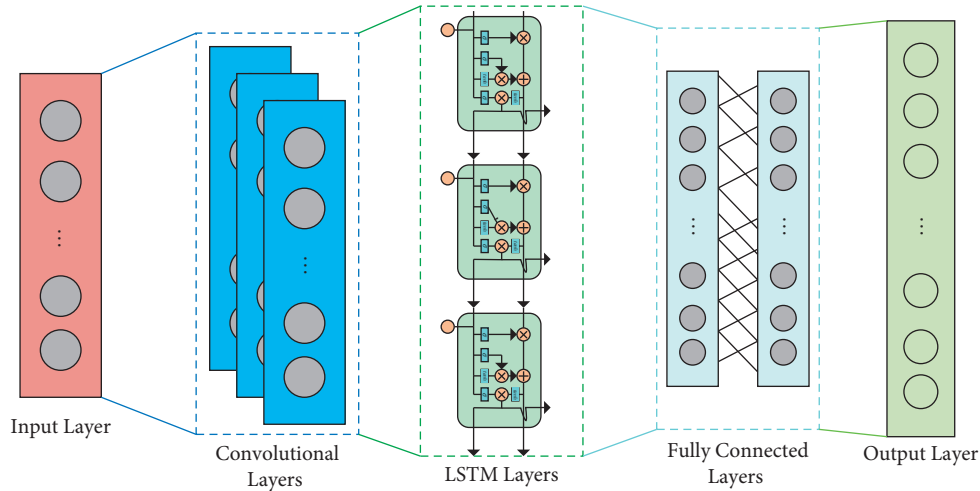


FIGURE 7: Architecture of CNN-LSTM CF model.

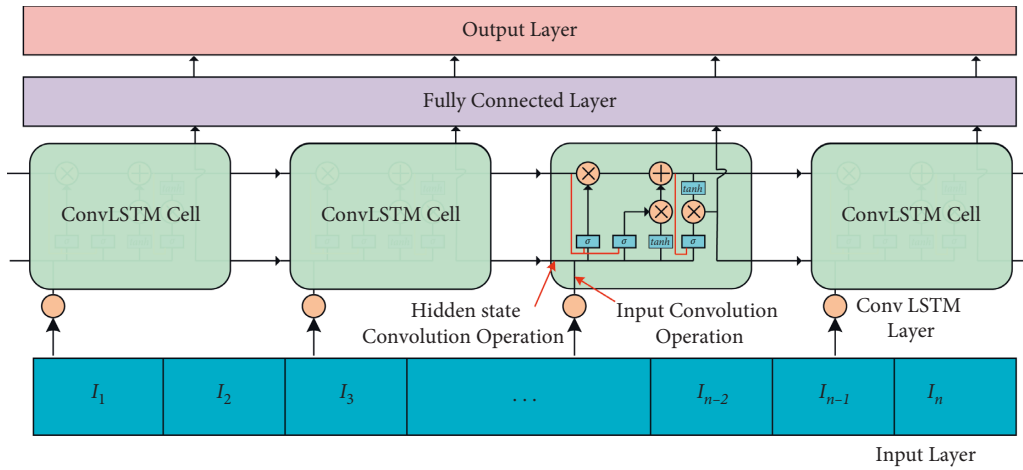


FIGURE 8: Architecture of ConvLSTM CF model.

TABLE 2: The hyperparameters of ANN CF models.

| Hyperparameters  | BP         | RNN          | LSTM         | CNN-LSTM     | Conv-LSTM    |
|------------------|------------|--------------|--------------|--------------|--------------|
| Number of layers | 4          | 4            | 4            | 4            | 4            |
| Number of cells  | 3, 9, 9, 1 | 3, 32, 10, 1 | 3, 32, 10, 1 | 3, 32, 10, 1 | 3, 32, 10, 1 |
| Batch size       | 128        | 128          | 128          | 128          | 128          |

training is used to solve the cross-validation problem. The cost of training and validation is depicted in Figure 9 using Mead-Square-Error.

For acceleration prediction, as illustrated in Figures 9(a) and 9(b), two proposed ANN CF models outperformed other models not only in the training loss comparison but also in the cross-validation loss. Due to the convolution operation, two novel ANN CF models were able to extract more features of vehicle movement at each time step compared to other CF models. Although BP CF model converged quickly, they had the highest training and cross-validation losses due to the BP model’s poor predictive ability and feature extraction limitations.

For velocity prediction, similar results are obtained for all models’ training and cross-validation losses. However, there are some notable findings. First, RNNs have the same training and cross-validation losses as LSTM CF models and perform worse than two novel ANN CF models. Second, both models’ cross-validation is more fluctuated than the training loss. The possible reason may be that it is ideally included all of the training data when the backpropagation algorithm is applied, but it is impossible due to the complexity of gradient calculation for all training data. Therefore, the smaller batch size is selected to accelerate the training process with the trade-off performance loss.

## 5. Conventional CF Models

In this study, three representative physical-based conventional CF models, IDM, Gipps, and FVD are chosen for comparison in the simulation experiment with the ANN CF models.

### 5.1. Physical-Based CF Models

**5.1.1. IDM.** IDM [22] is the most widely used CF model. In comparison to other models, the IDM model is more accurate and capable of simulating a variety of traffic flow scenarios, as defined in

$$\dot{v} = a \left[ 1 - \left( \frac{v}{v_0} \right)^\delta - \left( \frac{s^*(v, \Delta v)}{s} \right)^2 \right], \quad (4)$$

$$s^*(v, \Delta v) = s_0 + \max \left( 0, vT + \frac{v\Delta v}{2\sqrt{ab}} \right),$$

where  $v$ ,  $a$ ,  $b$ , and  $s$  are the velocity, acceleration, deceleration, and trajectory of the following vehicle, respectively;  $v_0$  is the desired velocity;  $\Delta v$  is the velocity difference;  $s_0$  is the minimum distance,  $T$  is the time gap, and  $\delta$  is the acceleration exponent.

**5.1.2. Gipps Model.** The Gipps model [25] is a typical safety distance CF model capable of simulating two distinct driver behaviors: CF and free flow. The following equation defines the model:

$$v_n(t + \tau) = \min \left( v_n(t) + 2.5a_{\max} * t * \left( 1 - \frac{v_n(t)}{V} \right) * \sqrt{0.025 + \frac{v_n(t)}{V}}, b_{\max} * t \right. \\ \left. + \sqrt{b_{\max}^2 * t^2 - b_{\max} * \frac{[2(\Delta x_n(t) - s_{n-1}) - v_n(t)t + v_{n-1}^2(t)]}{b_0}} \right), \quad (5)$$

where  $v_n$  and  $v_{n-1}$  are velocities of the following vehicle and preceding vehicle, respectively;  $a_{\max}$  and  $b_{\max}$  are the desired acceleration and deceleration rates of the following vehicle, respectively;  $b_0$  is the desired deceleration rate of the preceding vehicle;  $V$  is the desired velocity of the following vehicle;  $\Delta x_n(t)$  is the space headway;  $d$  is the minimum distance when the vehicle is stationary; and  $s_{n-1}$  is the length of the preceding vehicle.

**5.1.3. FVD.** The FVD [30] model is an extension of the CF model with optimal velocity. As illustrated in equation (6), it takes the effect of the velocity difference on the CF behavior into account:

$$a_n(t) = \alpha [V(\Delta x_n(t)) - v_n(t)] + \lambda \Delta v_n(t), \quad (6)$$

where  $\Delta x_n(t)$  is the space headway,  $v_n(t)$  is the velocity of the following vehicle,  $\Delta v_n(t)$  is the velocity difference,  $\alpha$  and  $\lambda$  are sensitivity coefficients, and  $v_n(\Delta x_n(t))$  is the optimal velocity function. The sensitivity coefficients and optimal velocity function can be defined as follows:

$$\lambda = \begin{cases} \lambda_0: \Delta x_n(t) \leq s_c, \\ 0: \Delta x_n(t) > s_c, \end{cases}$$

$$V(\Delta x_n(t)) = \frac{v_0}{2} \left[ \tanh \left( \frac{\Delta x_n(t) - l_{n-1}}{b} - \beta \right) - \tanh(-\beta) \right], \quad (7)$$

where  $\lambda_0$  is the constant sensitivity coefficient,  $v_0$  is the desired velocity of the following vehicle, and  $b$  and  $\beta$  are the parameters.

**5.2. Parameters Calibration.** Several parameters are included in conventional CF models, including the space headway and velocity difference. The values of the parameters vary according to the traffic scenario, and leaving the parameters uncalibrated makes evaluating the CF model's performance difficult. The simulated annealing algorithm [59] is used in this study to calibrate three different types of conventional CF models. Simulated annealing begins with a high initial temperature, steadily decreases the temperature parameters, and the global optimal solution of the given objective function is randomly found in the solution space. In other words, simulated annealing can approximate the global optimal solution from the local optimal solutions.

The CF model's calibration process utilizes the value bounds defined in [60] to calibrate the model's parameters using the NGSIM dataset. Table 3 summarizes the results of parameter calibration for three CF models.

As illustrated in Table 3, when compared to the original model parameters, the calibrated model parameters more accurately reflect the following vehicle's movement state. As can be seen, the deceleration is greater than the acceleration, indicating that vehicles typically decelerate rapidly at signalized intersections, resulting in frequent variations in the space headway. Additionally, the shorter minimum distance indicates that a signalized intersection has a high volume of traffic.



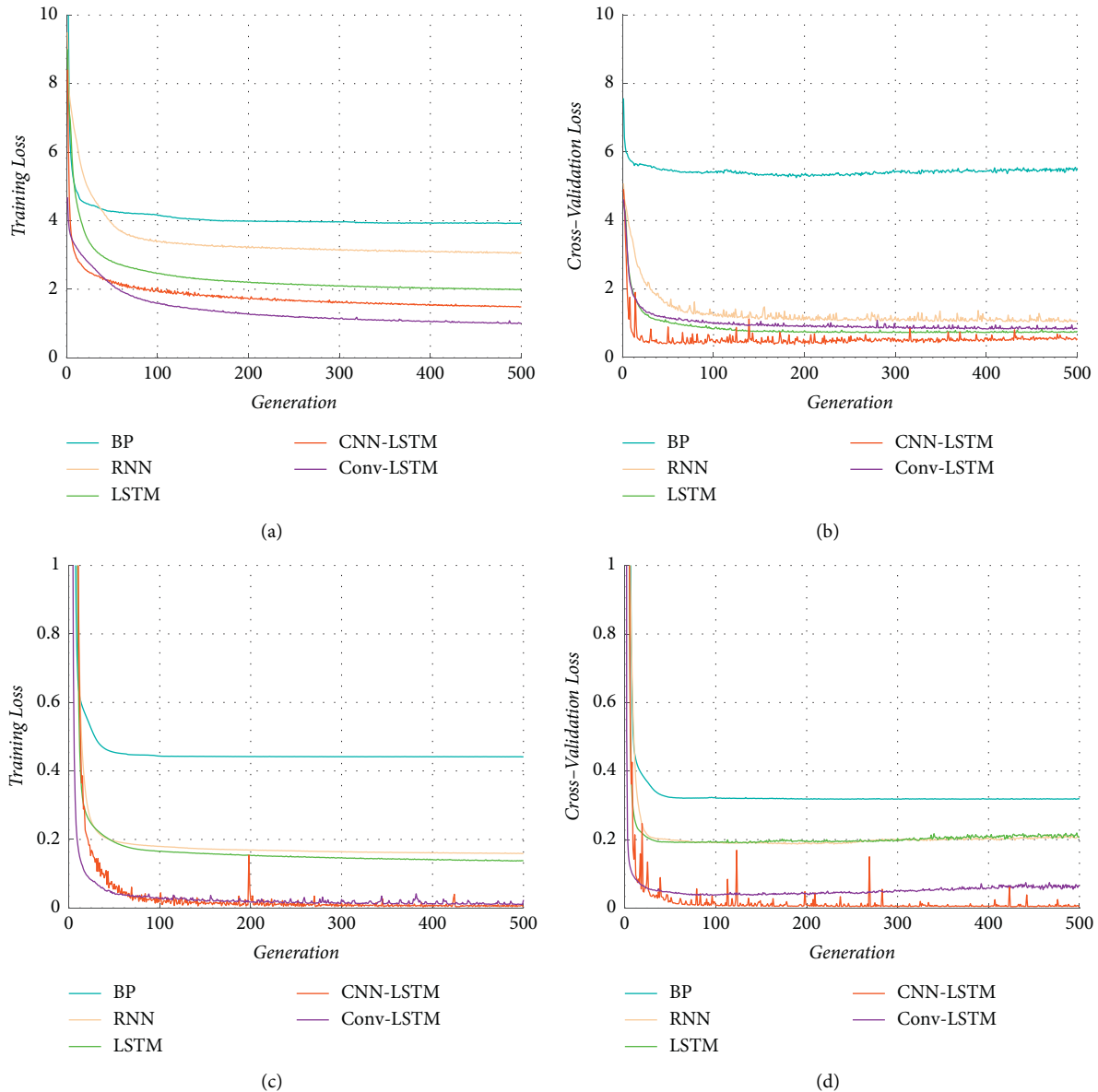


FIGURE 9: Training loss and validation of ANN CF models. (a) Training loss of ANN CF-A. (b) Cross-validation loss of ANN CF-A. (c) Training Los of ANN CF-V. (d) Cross-validation of ANN CF-V

## 6. Simulations and Analyses

In this section, the three representative CF models (conventional, ANN, and ML) are compared. The parameters settings of ML CF models are consistent with the literature [36, 37]. Additionally, all model parameters and simulation environment parameters used in the study are given to ensure the reproducibility of the experimental results [61].

**6.1. Continuous Traffic Flow.** The continuous traffic flow that occurs when the traffic light turns green or during the remaining red time allows the vehicle to pass through the signalized intersection. More precisely, the vehicle could proceed through the signalized intersection without

performing a stop-go or idling maneuver. Table 4 contains the parameters for the continuous traffic flow simulation experiment.

Figures 10–12 illustrate the CF models’ acceleration, velocity, trajectory, and hysteresis loops in continuous traffic flow.

For the acceleration comparison, all CF models could simulate the acceleration fluctuation of human-driven vehicles in Figure 10 of the acceleration comparison. However, some differences still existed in three types of models. It can be seen that the ANN CF-A models capture more details, and in particular, the proposed CNN-LSTM CF models outperformed other ANN CF models.

For the velocity comparison, as illustrated in Figure 11(a), the Gipps model is more unstable than the

TABLE 3: Model parameters calibration results.

| Model | Parameters                     | Description                           | Bounds     | Calibrated |
|-------|--------------------------------|---------------------------------------|------------|------------|
| Gipps | $a_{\max}$ (m/s <sup>2</sup> ) | Max acceleration of following vehicle | [0.1, 5]   | 2.14       |
|       | $b_{\max}$ (m/s <sup>2</sup> ) | Max deceleration of following vehicle | [0.1, 5]   | 2.70       |
|       | $b_0$ (m/s <sup>2</sup> )      | Max deceleration of preceding vehicle | [0.1, 5]   | 1.44       |
|       | $V$ (km/h)                     | Desired velocity of following vehicle | [1, 150]   | 59.79      |
|       | $t$ (s)                        | Reaction time                         | [0.3, 3]   | 2.03       |
| FVD   | $\alpha$                       | Constant sensitivity                  | [0.05, 20] | 0.06       |
|       | $\lambda_0$                    | Sensitivity to relative velocity      | [0, 3]     | 0.80       |
|       | $v_0$ (km/h)                   | Desired of following vehicle          | [1, 252]   | 40.84      |
|       | $b$ (m)                        | Interaction length                    | [0.1, 100] | 6.47       |
|       | $\beta$                        | Form factor                           | [0.1, 10]  | 0.75       |
|       | $s_c$ (m)                      | Max following distance                | [10 120]   | 48.24      |
|       | $a$ (m/s <sup>2</sup> )        | Acceleration of following vehicle     | [0.1, 5]   | 1.98       |
| IDM   | $v_0$ (km/h)                   | Desired velocity of following vehicle | [1, 252]   | 45.29      |
|       | $\delta$                       | Acceleration exponent                 | [1, 40]    | 1.34       |
|       | $b$ (m/s <sup>2</sup> )        | Deceleration of following vehicle     | [0.1, 5]   | 4.37       |
|       | $s_0$ (m)                      | Minimum distance                      | [0.1, 10]  | 0.31       |
|       | $T$ (s)                        | Time gap                              | [0.1, 5]   | 0.48       |

TABLE 4: Attributes setting of continuous traffic flow.

| Attributes                                | Value    |
|---|----------|
| Initial velocity of the preceding vehicle | 7.78 m/s |
| Initial velocity of the following vehicle | 6.97 m/s |
| Initial headway                           | 18.72 m  |

other two conventional CF models. The velocity of the following vehicle is simulated in Figure 11(b) by integrating the acceleration predicted by ANN CF-A (acceleration) models. The following vehicle's velocity was also simulated in Figure 11(c), but it is directly predicted by the ANN CF-V (velocity) models. As can be seen, the ANN CF-V simulates a velocity closer to reality than the ANN CF-A, and the accumulative integration error negatively affects the ANN CF-A.

For the trajectory comparison, as shown in Figure 12(a), the FVD CF model is highly consistent with the real trajectory during the first twenty seconds, but the error gradually increases during the subsequent period. The trajectory obtained by double integrating acceleration and single integrating velocity is shown in Figures 12(b) and 12(c). Compared to Figure 12(b), the trajectory error obtained by ANN CF-V is less than that obtained by ANN CF-A. This is consistent with the findings of the literature [47]. Additionally, as indicated by the red circle in Figures 12(b) and 12(c), the KNN, FRNN, CNN-LSTM, and Conv-LSTM perform better than other CF models, and the trajectory errors of the ANN CF-A and ANN CF-V remain low after twenty seconds of CF behavior, validating the proposed models' robustness.

Each model's Root Mean Square Error (RMSE) is calculated to provide a more intuitive assessment of its performance. The continuous traffic flow experiment contains 250 data points. Figure 13 depicted the RMSE of conventional, ML, and ANN CF-A models. As can be seen, the IDM and BP had a higher RMSE for acceleration, velocity, and trajectory, indicating that they were unable to accurately

simulate the vehicle movement of the following vehicle. The FVD model outperformed the RNN and LSTM models. Additionally, the proposed CNN-LSTM and Conv-LSTM models performed better than the KNN CF models but worse than the FRNN models. Figure 14 depicted the RMSE of conventional, ML, and ANN CF-V models. Compared with ANN CF-A models, the RMSE of ANN CF-V models is further reduced. Especially, the Conv-LSTM and FRNN had the lowest RMSE.

As illustrated in Figures 13 and 14, the above three labels indicate the difference between CF models' simulated vehicle movement feature and the real vehicle movement feature. However, there are distinct stages of CF behavior that must be explored further. As a result, the hysteresis loop is used to evaluate CF models. As illustrated in Figure 15, the following vehicle's space-headway range is 10–60 m, and its velocity range is 6–18 m/s. It indicates that the following vehicle will proceed through the signalized intersection without encountering any traffic congestion. As illustrated in Figure 15(a), the hysteresis loop of IDM and Gipps are more "aggressive" than the real data and the FVD model. Additionally, although the FVD is closer to the real data, it has a limited capacity for detecting subtle changes in the real data. The hysteresis loops of ML, ANN CF-A, and ANN CF-V are depicted in Figures 15(b) and 15(c). Clearly, the ANN CF-V is more consistent with real data than the ANN CF-A, particularly the BP and RNN, which perform poorly and are more "conservative" than real data. Additionally, it can be seen that the KNN CF models performed more "aggressive" like Gipps and IDM.

**6.2. Discontinuous Traffic Flow.** When the traffic light is red or the green time is insufficient to allow the vehicle to pass through the signalized intersection, discontinuous traffic flow frequently occurs. This means that the vehicle must adopt a stop-go behavior in order to pass through a signalized intersection. The parameters of the discontinuous traffic flow simulation experiment are presented in Table 5.

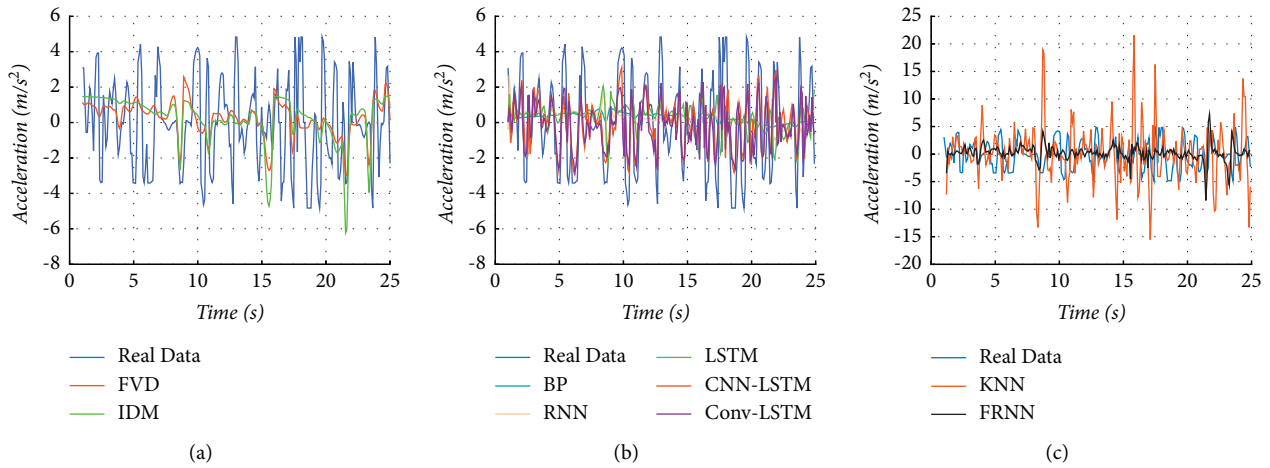


FIGURE 10: Acceleration comparison of continuous traffic flow. (a) Conventional models. (b) ANN CF-A. (c) ML CF.

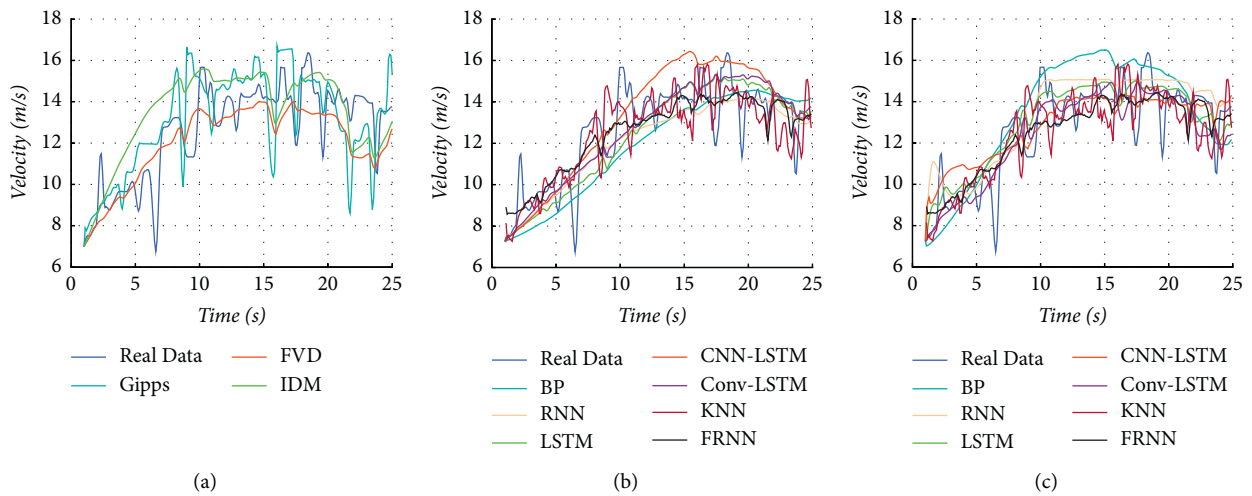


FIGURE 11: Velocity comparison of continuous traffic flow. (a) Conventional model. (b) ANN CF-A. (c) ANN CF-V

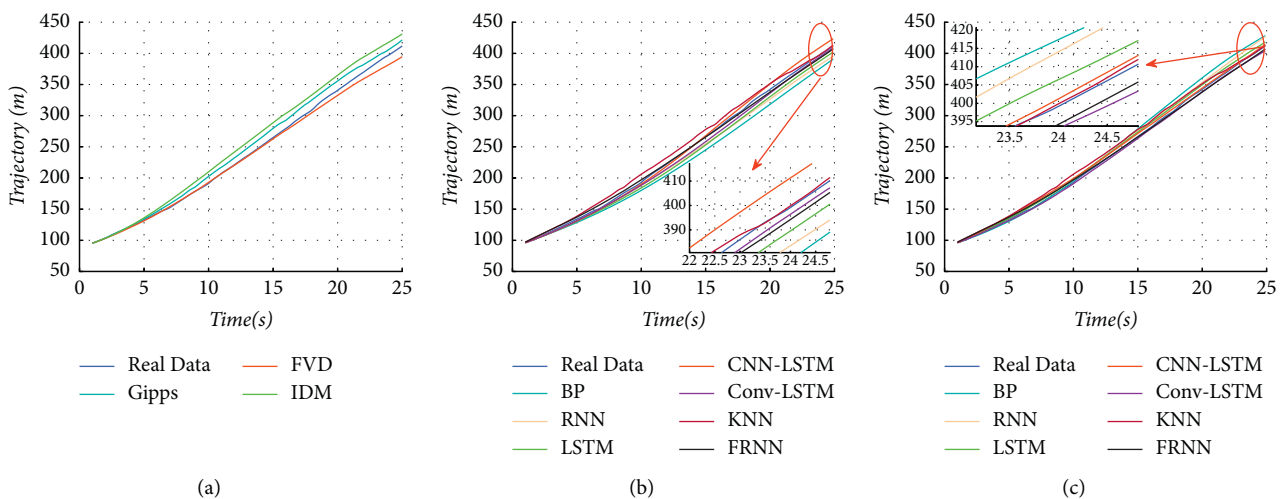


FIGURE 12: Trajectory comparison of continuous traffic flow. (a) Conventional model. (b) ANN CF-A. (c) ANN CF-V

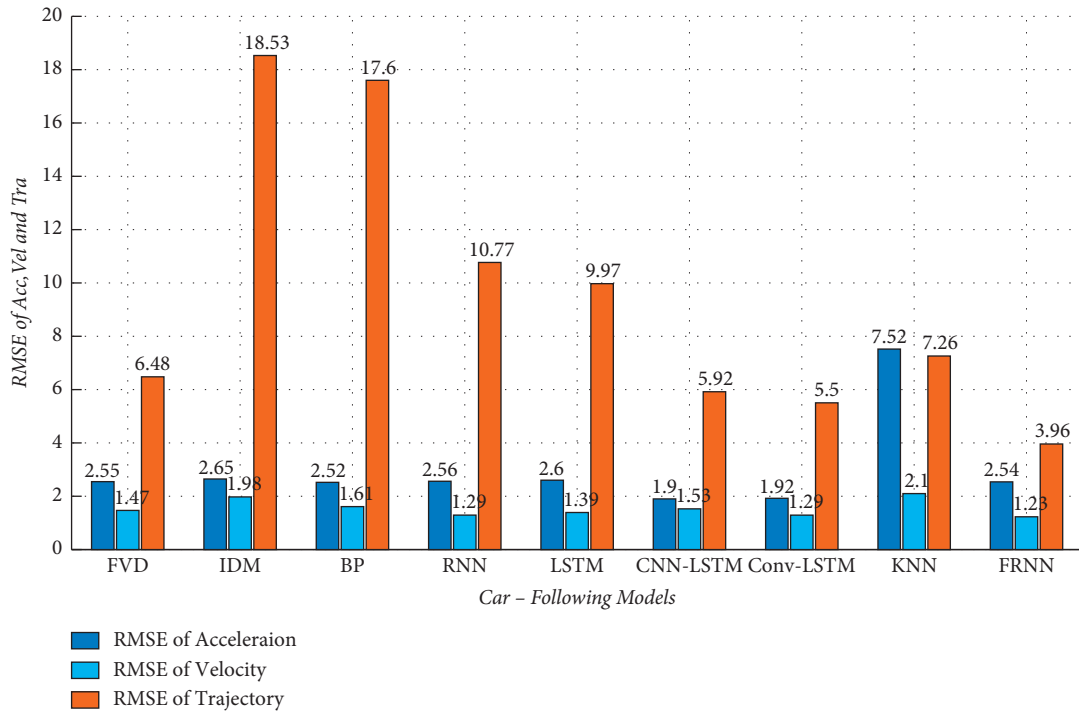


FIGURE 13: RMSE of conventional models and data-driven models in the continuous traffic flow (ANN CF-A).

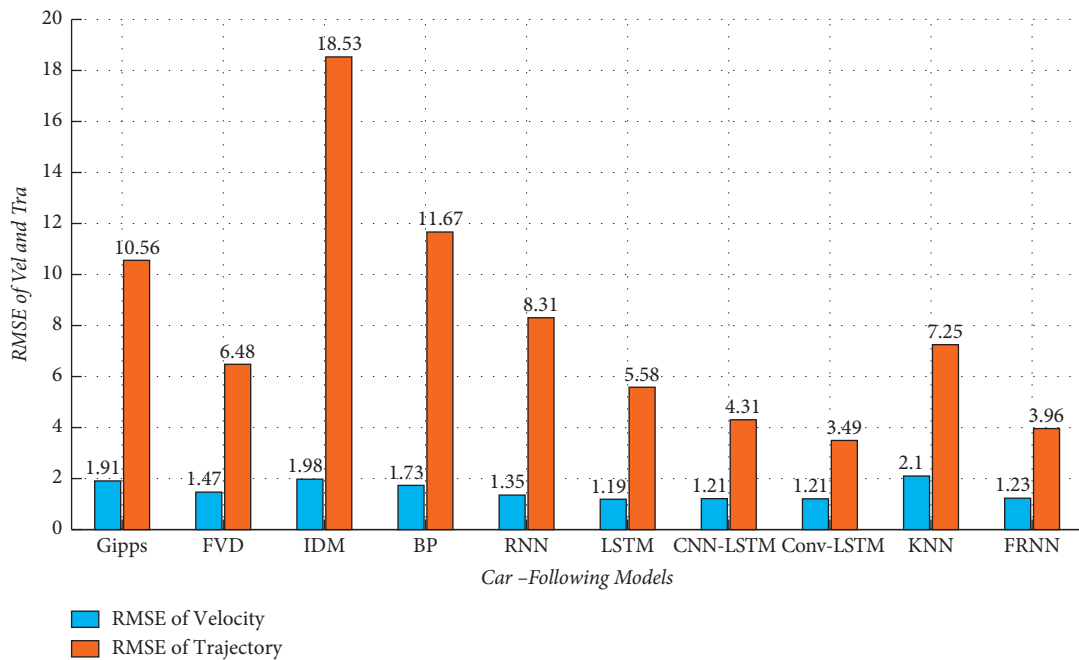


FIGURE 14: RMSE of conventional models and data-driven models in the continuous traffic flow. (ANN CF-V).

For the acceleration comparison, the acceleration which was simulated by ANN CF-A models were more consistent with the real acceleration of following vehicle, as shown in Figure 16. The conventional CF models cannot simulate many minor changes of acceleration. Additionally, even though the ML CF models have the ability to simulate the acceleration, the KNN CF model does not perform well, and

some acceleration values that are generated by KNN CF model is unrealistic.

For the velocity comparison, Figure 17 exhibited the velocity comparison results of conventional CF, ML CF, and ANN CF models. The CNN-LSTM CF and the Conv-LSTM CF models of the ANN CF-A models could simulate velocity simulations more realistically than

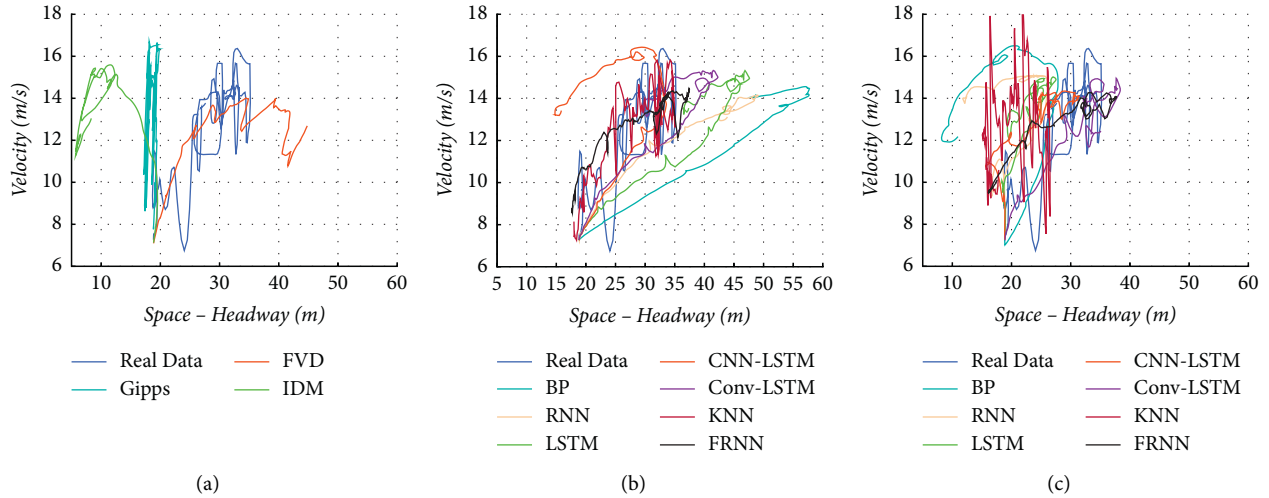


FIGURE 15: Hysteresis loops comparison of continuous traffic flow. (a) Conventional model; (b) ANN CF (based on the acceleration); (c) ANN CF (based on the velocity).

TABLE 5: Attribute settings of discontinuous traffic flow.

| Attributes                                | Value     |
|---|-----------|
| Initial velocity of the preceding vehicle | 7.77 m/s  |
| Initial velocity of the following vehicle | 13.50 m/s |
| Initial headway                           | 54.60 m   |

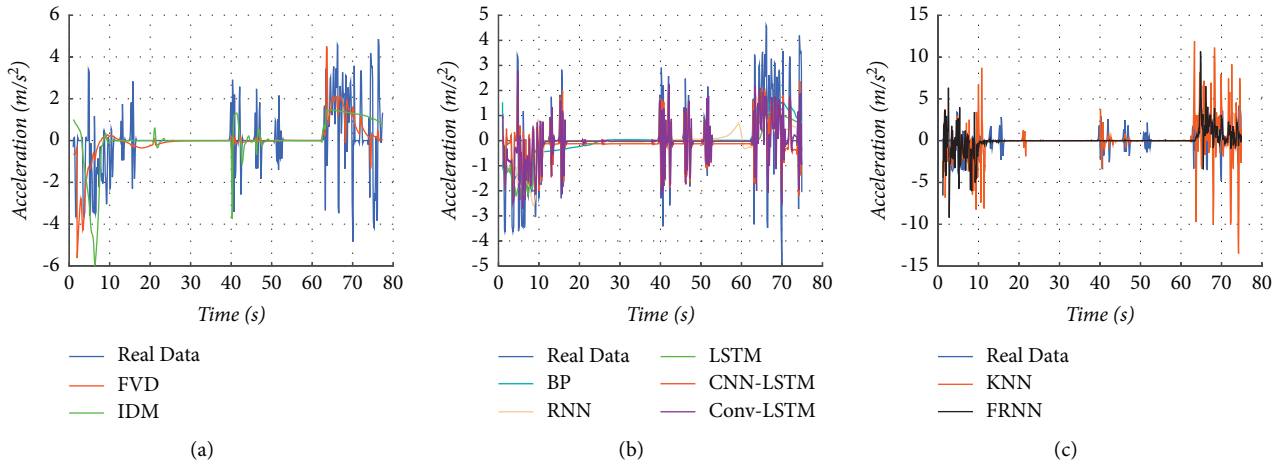


FIGURE 16: Acceleration comparison of discontinuous traffic flow. (a) Conventional models; (b) ANN CF- (A); and (c) ML CF.

conventional CF models and BP, RNN, and LSTM CF models, but when the traffic light is green, only KNN and FRNN CF models could simulate the velocity of following vehicle. Additionally, they both performed worse than the ANN CF-V as shown in Figure 17(c). At approximately 40th seconds, the velocity simulated by IDM is below zero, while the velocity simulated by FVD is abnormal, as indicated by the red circle in Figure 17(a). Furthermore, as illustrated in Figure 17(b), the velocity simulated by BP and RNN is below zero. They are not consistent with reality.

For the trajectory comparison, the unusual velocity fluctuations directly resulted in the same problems that occur during the trajectory simulation. The trajectory simulated by FVD is inconsistent with the real trajectory of the following vehicle between the 10th and 20th seconds, and the trajectory simulated by BP and RNN is also inconsistent with the real trajectory of the following vehicle when it came to a stop, as indicated by the red circle in Figures 18(a) and 18(b). The simulation results in Figures 16–18 indicate that when the loss of ANN CF-A models is not sufficiently low, they may have a limited ability to simulate long-term

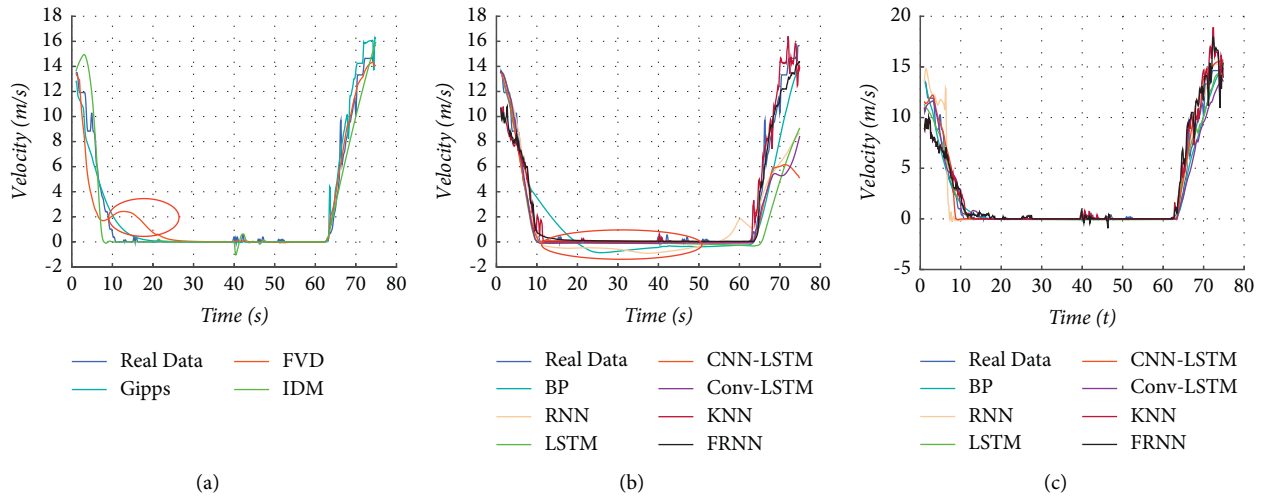


FIGURE 17: Velocity comparison of discontinuous traffic flow. (a) Conventional models; (b) ANN CF-A; and (c) ANN CF-V

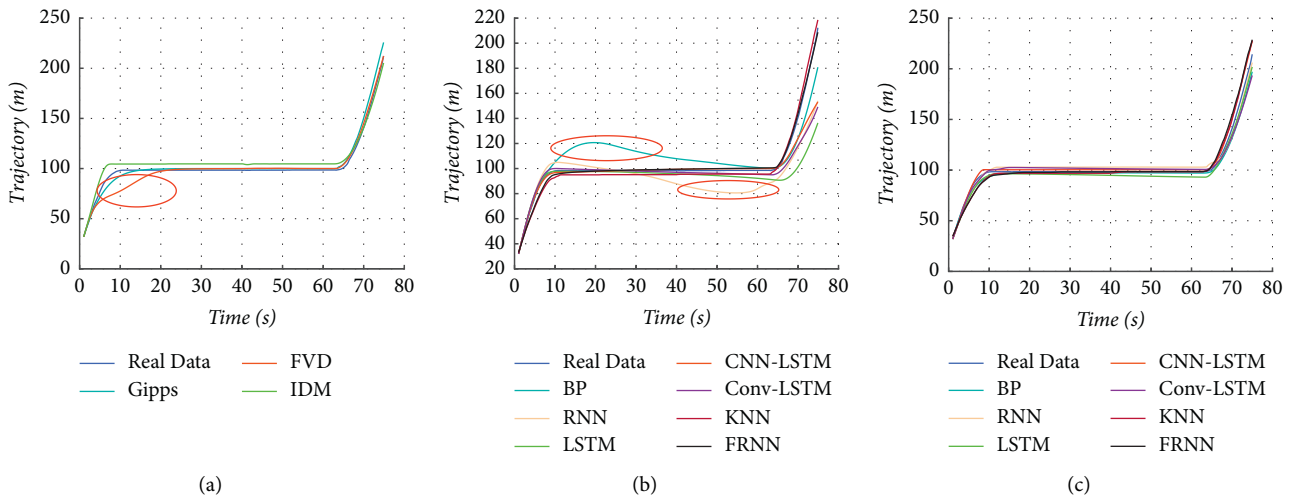


FIGURE 18: Trajectory comparison of discontinuous traffic flow. (a) Conventional model; (b) ANN CF-A; and (c) ANN CF-V

stop-go behavior in discontinuous traffic flows. In comparison, although conventional CF models could not accurately simulate vehicle movement, they were more stable than ANN CF-A models, but ANN CF-V models especially CNN-LSTM and Conv-LSTM still performed best, as shown in Figure 18(c).

Similar to the simulation experiment for continuous traffic flow, the RMSE of acceleration, velocity, and trajectory for ANN CF-A and ANN CF-V are shown in Figures 19 and 20. The RMSE of ANN CF-A and ML are greater than that of conventional CF models and produces the same results as stated earlier. We are impressed to see that all ANN CF-V models had a lower RMSE of trajectory, and the CNN-LSTM and Conv-LSTM models had the

lowest errors. Besides, it can be seen that KNN-FRNN simulate trajectory better than the velocity, and the FRNN-CF model achieved similar results to the ANN CF-V models.

As illustrated in Figure 21, the ANN CF-V model outperforms three different types of CF models. Gipps and IDM simulate a narrower hysteresis loop for conventional CF models than the FVD model. The reason for this is that the FVD's abnormal simulated velocity had a negative effect on its performance. For ANN-A CF models, the velocity of BP is negative, indicating that it violates basic rules of CF rules. For ANN-V CF models, it can be seen that CNN-LSTM and Conv-LSTM CF models are more consistent with the real data than KNN and ML CF models.

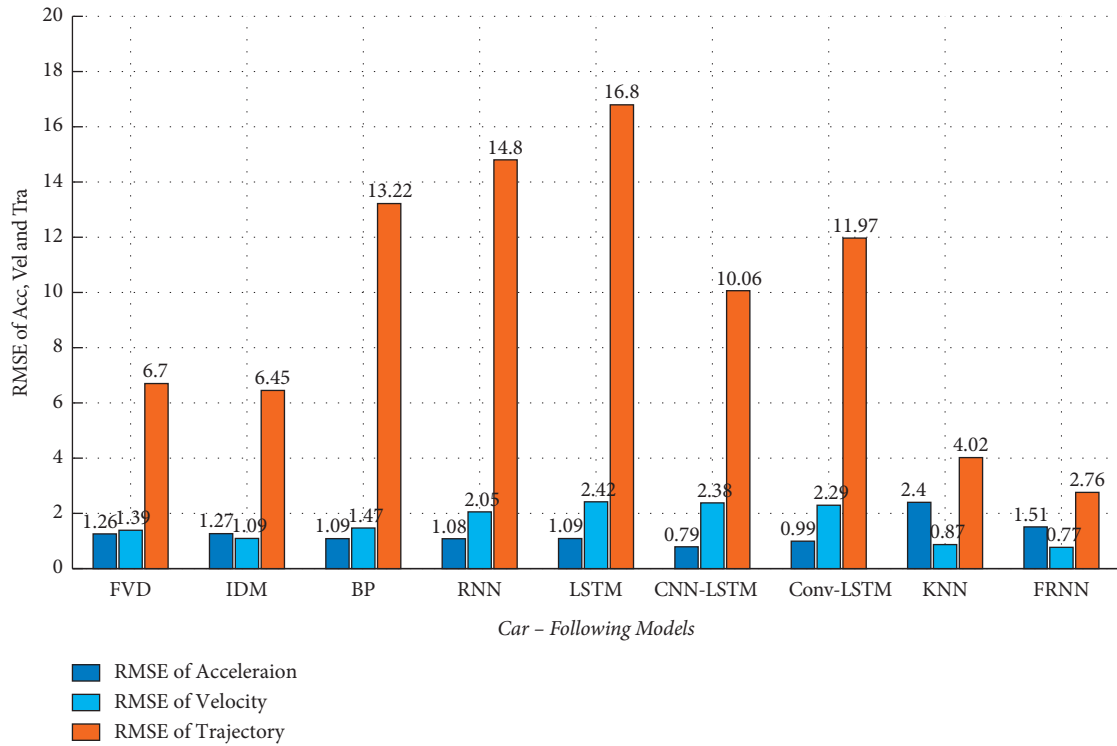


FIGURE 19: RMSE of conventional models and data-driven models in the discontinuous traffic flow (ANN CF-A).

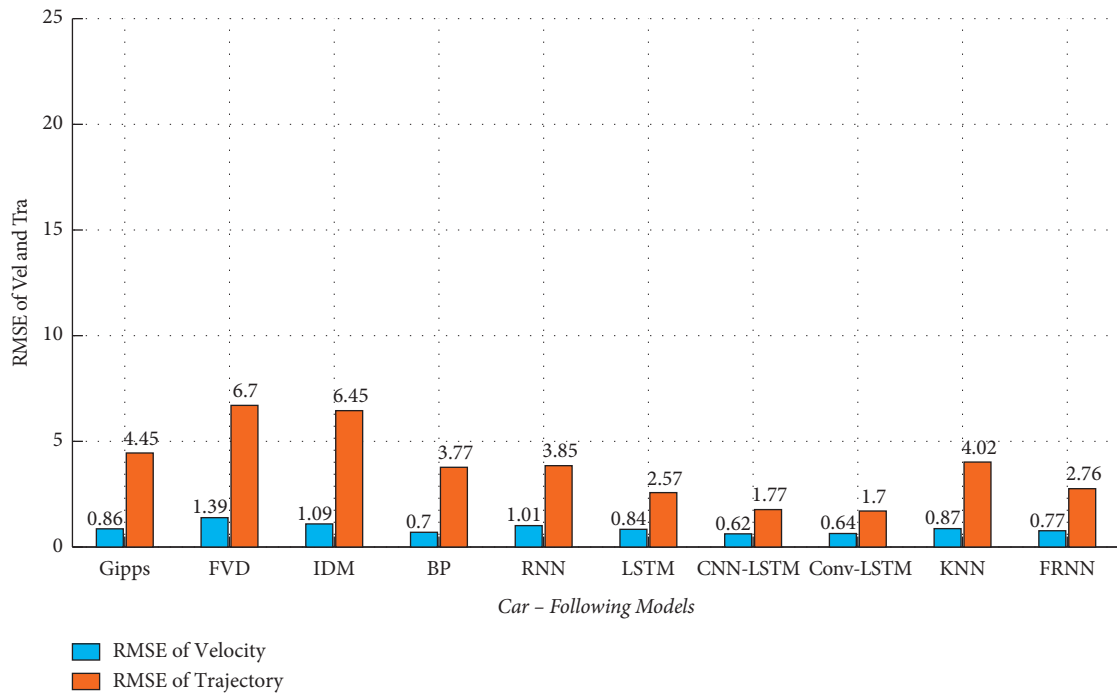


FIGURE 20: RMSE of conventional models and data-driven models in the discontinuous traffic flow (ANN CF-V).

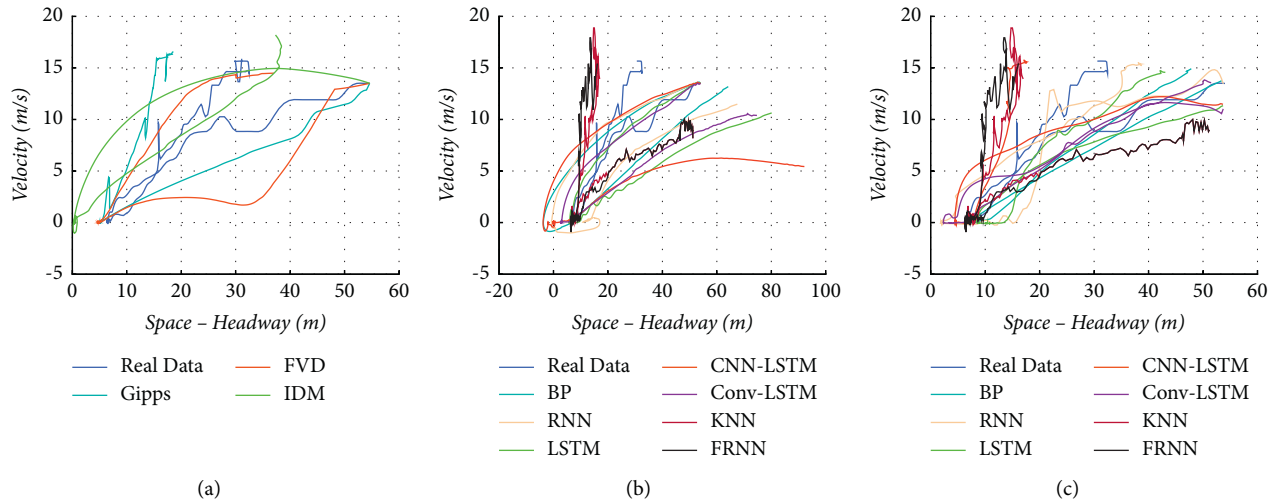


FIGURE 21: Hysteresis loops comparison of discontinuous traffic flow. (a) Conventional model; (b) ANN CF-A; and (c) ANN CF-V.

## 7. Conclusion

In this study, we compare conventional and ANN CF models to analyze and evaluate their performance to describe the vehicle behavior at the signalized intersection. The filtered NGSIM data are used to train and validate CF models. We propose the CNN-LSTM and Conv-LSTM CF models. The simulation is based on two typical categories of traffic flow at a signalized intersection. The acceleration, velocity, trajectory, and hysteresis loops, as well as the RMSE, are used to quantify the performance of all CF models in predicting the movement state of the following vehicle. The following summarizes the findings:

- (1) While conventional CF models perform poorly in continuous traffic flow simulation-based acceleration/velocity prediction and discontinuous traffic flow simulation-based velocity prediction, they perform better in discontinuous traffic flow-based acceleration prediction than ANN CF-A models. It is demonstrated that conventional CF models have an advantage in the discontinuous traffic flow simulation, particularly for forecasting acceleration of vehicle with long-term stop-go behavior. Because the parameters of conventional CF models have explicit physical meanings, they can effectively follow the human driver rules integrated into the model. Thus, when the accuracy of ANN CF-A models is insufficient for discontinuous traffic flow, conventional CF models are recommended.
- (2) The proposed two novel ANN CF models outperformed other CF models except the discontinuous traffic flow simulation-based acceleration prediction. The CNN-LSTM and Conv-LSTM models fully exploit the advantages of convolution and LSTM. The ANN CF models are recommended for velocity prediction at signalized intersections using traffic flow simulations and short-term

acceleration prediction using traffic flow simulations. Additionally, the accuracy of the ANN CF models must be improved further if they are to be used in more complicated traffic flow scenarios at signalized intersections.

The study's limitation is that it considers only locally following behavior (i.e., between two consecutive vehicles) and that different datasets may affect the performance of the ANN CF models. Vehicle platooning will be considered in future work, and extended experiments with diverse datasets will be required to gain additional insight into the proposed ANN-based CF models.

## Data Availability

Two datasets are used in this study to train and calibrate car-following models of signalized intersections: the Lankershim dataset and the Peach Tree dataset, both of which are provided by the Federal Highway Administration's NGSIM program. The NGSIM datasets include detailed vehicle and road information for micro traffic flow research.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research has been supported by the National Key Research and Development Program of China (No. 2021YFB2501200), the National Natural Science Foundation of China (No. 61703053), the Shaanxi Province Key Research and Development Program (2022GY-300), the Shaanxi Province Postdoctoral Science Foundation (No. 2018BSHYDZZ64), and Chang'an University Graduate Student Innovation Ability Cultivation Project (No. 300203211242).



## References

- [1] H. Yu, R. Jiang, Z. He et al., "Automated vehicle-involved traffic flow studies: A survey of assumptions, models, specifications, and perspectives," *Transportation Research Part C: Emerging Technologies*, vol. 127, Article ID 103101, 2021.
- [2] H. Jiang, J. Hu, S. An, M. Wang, and B. B. Park, "Eco approaching at an isolated signalized intersection under partially connected and automated vehicles environment," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 290–307, 2017.
- [3] M. Li, X. Wu, X. He, G. Yu, and Y. Wang, "An eco-driving system for electric vehicles with signal control under V2X environment," *Transportation Research Part C: Emerging Technologies*, vol. 93, pp. 335–350, 2018.
- [4] S. Fang, L. Yang, T. Wang, and S. Jing, "Trajectory planning method for mixed vehicles considering traffic stability and fuel consumption at the signalized intersection," *Journal of Advanced Transportation*, vol. 2020, 2020.
- [5] H. Yao and X. Li, "Decentralized control of connected automated vehicle trajectories in mixed traffic at an isolated signalized intersection," *Transportation Research Part C: Emerging Technologies*, vol. 121, Article ID 102846, 2020.
- [6] S. Soleimaniamiri, A. Ghiasi, X. Li, and Z. Huang, "An analytical optimization approach to the joint trajectory and signal optimization problem for connected automated vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 120, Article ID 102759, 2020.
- [7] X. Han, R. Ma, and H. M. Zhang, "Energy-aware trajectory optimization of CAV platoons through a signalized intersection," *Transportation Research Part C: Emerging Technologies*, vol. 118, Article ID 102652, 2020.
- [8] H. Yao, J. Cui, X. Li, Y. Wang, and S. An, "A trajectory smoothing method at signalized intersection based on individualized variable speed limits with location optimization," *Transportation Research Part D: Transport and Environment*, vol. 62, pp. 456–473, 2018.
- [9] J. Zhao, K. K. Kigen, and M. Wang, "Modelling the operation of vehicles at signalised intersections with special width approach lane based on field data," *IET Intelligent Transport Systems*, vol. 14, no. 12, pp. 1565–1572, 2020.
- [10] J. Zhang, T. Q. Tang, and T. Wang, "Some features of car-following behaviour in the vicinity of signalised intersection and how to model them," *IET Intelligent Transport Systems*, vol. 13, no. 11, pp. 1686–1693, 2019.
- [11] T.-Q. Tang, Z.-Y. Yi, J. Zhang, T. Wang, and J.-Q. Leng, "A speed guidance strategy for multiple signalized intersections based on car-following model," *Physica A: Statistical Mechanics and Its Applications*, vol. 496, pp. 399–409, 2018.
- [12] H. Zhao, R. He, and C. Ma, "An extended car-following model at signalized intersections," *Journal of Advanced Transportation*, vol. 2018, 2018.
- [13] S. Yu and Z. Shi, "An extended car-following model at signalized intersections," *Physica A: Statistical Mechanics and Its Applications*, vol. 407, pp. 152–159, 2014.
- [14] D.-W. Liu, Z.-K. Shi, and W.-H. Ai, "Enhanced stability of car-following model upon incorporation of short-term driving memory," *Communications in Nonlinear Science and Numerical Simulation*, vol. 47, pp. 139–150, 2017.
- [15] R. Arvin, A. J. Khattak, M. Kamrani, and J. Rio-Torres, "Safety evaluation of connected and automated vehicles in mixed traffic with conventional vehicles at intersections," *Journal of Intelligent Transportation Systems*, vol. 25, no. 2, pp. 170–187, 2020.
- [16] K. Liu, D. Liu, C. Li, and T. Yamamoto, "Eco-speed guidance for the mixed traffic of electric vehicles and internal combustion engine vehicles at an isolated signalized intersection," *Sustainability*, vol. 11, no. 20, p. 5636, 2019.
- [17] E. Brockfeld, R. D. Kühne, and P. Wagner, "Calibration and validation of microscopic traffic flow models," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1876, no. 1, pp. 62–70, 2004.
- [18] S. Albeaik, A. Bayen, M. T. Chiri et al., "Limitations and improvements of the intelligent driver model (IDM)," 2021, <https://arxiv.org/abs/2104.02583>.
- [19] L. A. Pipes, "An operational analysis of traffic dynamics," *Journal of Applied Physics*, vol. 24, no. 3, pp. 274–281, 1953.
- [20] D. C. Gazis, R. Herman, and R. W. Rothery, "Nonlinear follow-the-leader models of traffic flow," *Operations Research*, vol. 9, no. 4, pp. 545–567, 1961.
- [21] G. F. Newell, "Nonlinear effects in the dynamics of car following," *Operations Research*, vol. 9, no. 2, pp. 209–229, 1961.
- [22] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Physical Review*, vol. 62, no. 2, pp. 1805–1824, 2000.
- [23] Y. Wu, M. Abdel-Aty, L. Wang, and M. S. Rahman, "Combined connected vehicles and variable speed limit strategies to reduce rear-end crash risk under fog conditions," *Journal of Intelligent Transportation Systems*, vol. 24, no. 5, pp. 494–513, 2020.
- [24] H. Cui, J. Xing, X. Li, and M. Zhu, "Effects of adaptive and cooperative adaptive cruise control on the fuel consumption and emissions at the signalized intersection," *Modern Physics Letters B*, vol. 32, no. 32, Article ID 1850396, 2018.
- [25] P. G. Gipps, "A behavioural car-following model for computer simulation," *Transportation Research Part B: Methodological*, vol. 15, no. 2, pp. 105–111, 1981.
- [26] R. J. T. E. Wiedemann, *Simulation des Strassenverkehrsflusses*, Schriftenreihe des Instituts für Verkehrswesen, Heft 8, Universität Karlsruhe, Karlsruhe, Germany, 1974.
- [27] Y. Wang, L. Wei, and P. Chen, "Trajectory reconstruction for freeway traffic mixed with human-driven vehicles and connected and automated vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 111, pp. 135–155, 2020.
- [28] B. P. Chauhan, G. J. Joshi, and P. Parida, "Car following model for urban signalised intersection to estimate speed based vehicle exhaust emissions," *Urban Climate*, vol. 29, Article ID 100480, 2019.
- [29] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, "Dynamical model of traffic congestion and numerical simulation," *Physical Review*, vol. 51, no. 2, pp. 1035–1042, 1995.
- [30] R. Jiang, Q. Wu, and Z. Zhu, "Full velocity difference model for a Car-following theory," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 64, no. 1, Article ID 017101, 2001.
- [31] W. Zhao, D. Ngoduy, S. Shepherd, R. Liu, and M. Papageorgiou, "A platoon based cooperative eco-driving model for mixed automated and human-driven vehicles at a signalised intersection," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 802–821, 2018.
- [32] Y. Qin, H. Wang, and B. Ran, "Stability analysis of connected and automated vehicles to reduce fuel consumption and emissions," *Journal of Transportation Engineering, Part A: Systems*, vol. 144, no. 11, Article ID 04018068, 2018.
- [33] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Transactions*

- on systems, *Man, and Cybernetics*, vol. SMC-3, no. 1, pp. 28–44, 1973.
- [34] H. Hao, W. Ma, and H. Xu, “A fuzzy logic-based multi-agent car-following model,” *Transportation Research Part C: Emerging Technologies*, vol. 69, pp. 477–496, 2016.
- [35] T. Toledo, H. N. Koutsopoulos, and K. I. Ahmed, “Estimation of vehicle trajectories with locally weighted regression,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1999, no. 1, pp. 161–169, 2007.
- [36] Z. He, L. Zheng, and W. Guan, “A simple nonparametric car-following model driven by field data,” *Transportation Research Part B: Methodological*, vol. 80, pp. 185–201, 2015.
- [37] Y. Yu, Z. He, and X. Qu, “On the impact of prior experiences in car-following models: model development, computational efficiency, comparative analyses, and extensive applications,” *IEEE Transactions on Cybernetics*, 2021.
- [38] B. Lu, S. Ni, and S. S. Washburn, “A support vector regression approach for investigating multianticipative driving behavior,” *Mathematical Problems in Engineering*, vol. 2015, Article ID 701926, 2015.
- [39] J. Hongfei, J. Zhicai, and N. Anning, “Develop a car-following model using data collected by “five-wheel system”” vol. 1, pp. 346–351, in *Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems*, vol. 1, pp. 346–351, IEEE, Shanghai, China, October 2003.
- [40] A. Khodayari, A. Ghaffari, R. Kazemi, and R. Braunstingl, “A modified car-following model based on a neural network model of the human driver effects,” *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 42, no. 6, pp. 1440–1449, 2012.
- [41] J. Zheng, K. Suzuki, and M. Fujita, “Car-following behavior with instantaneous driver-vehicle reaction delay: A neural-network-based methodology,” *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 339–351, 2013.
- [42] C. Colombaroni and G. Fusco, “Artificial neural network models for car following: experimental analysis and calibration issues,” *Journal of Intelligent Transportation Systems*, vol. 18, no. 1, pp. 5–16, 2014.
- [43] M. Zhu, X. Wang, and Y. Wang, “Human-like autonomous car-following model with deep reinforcement learning,” *Transportation Research Part C: Emerging Technologies*, vol. 97, pp. 348–368, 2018.
- [44] B. Peng, M. F. Keskin, B. Kulcsár, and H. Wymeersch, “Connected autonomous vehicles for improving mixed traffic efficiency in unsignalized intersections with deep reinforcement learning,” *Communications in Transportation Research*, vol. 1, Article ID 100017, 2021.
- [45] V. Papathanasopoulou and C. Antoniou, “Towards data-driven car-following models,” *Transportation Research Part C: Emerging Technologies*, vol. 55, pp. 496–509, 2015.
- [46] X. Wang, R. Jiang, L. Li, Y. Lin, X. Zheng, and F. Y. Wang, “Capturing car-following behaviors by deep learning,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 3, pp. 910–920, 2017.
- [47] M. Zhou, X. Qu, and X. Li, “A recurrent neural network based microscopic car following model to predict traffic oscillation,” *Transportation Research Part C: Emerging Technologies*, vol. 84, pp. 245–264, 2017.
- [48] L. Ma and S. Qu, “A sequence to sequence learning based car-following model for multi-step predictions considering reaction delay,” *Transportation Research Part C: Emerging Technologies*, vol. 120, Article ID 102785, 2020.
- [49] F. Hui, C. Wei, W. ShangGuan, R. Ando, and S. Fang, “Deep encoder-decoder-NN: A deep learning-based autonomous vehicle trajectory prediction and correction model,” *Physica A: Statistical Mechanics and Its Applications*, Article ID 126869, 2022.
- [50] Z. Yuan, K. He, and Y. Yang, “A roadway safety sustainable approach: modeling for real-time traffic crash with limited data and its reliability verification,” *Journal of Advanced Transportation*, vol. 2022, 2022.
- [51] Y. Yang, K. He, Y. P. Wang, Z. Z. Yuan, Y. H. Yin, and M. Z. Guo, “Identification of dynamic traffic crash risk for cross-area freeways based on statistical and machine learning methods,” *Physica A: Statistical Mechanics and Its Applications*, vol. 595, Article ID 127083, 2022.
- [52] V. Alexiadis, J. Colyar, J. Halkias, R. Hranac, and G. McHale, “The next generation simulation program. Institute of Transportation Engineers,” *ITEA Journal*, vol. 74, no. 8, p. 22, 2004.
- [53] M. Montanino and V. Punzo, “Making NGSIM data useable for studies on traffic flow theory,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2390, no. 1, pp. 99–111, 2013.
- [54] B. Coifman and L. Li, “A critical evaluation of the Next Generation Simulation (NGSIM) vehicle trajectory dataset,” *Transportation Research Part B: Methodological*, vol. 105, pp. 362–377, 2017.
- [55] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [56] S. H. I. Xingjian, 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 *Advances in Neural Information Processing Systems*, pp. 802–810, 2015.
- [57] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, 2014, arXiv:1412.6980.
- [58] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting,” *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [59] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
- [60] M. Zhu, X. Wang, A. Tarko, and S. E. Fang, “Modeling car-following behavior on urban expressways in Shanghai: A naturalistic driving study,” *Transportation Research Part C: Emerging Technologies*, vol. 93, pp. 425–445, 2018.
- [61] Z. Zheng, “Reasons, challenges, and some tools for doing reproducible transportation research,” *Communications in Transportation Research*, vol. 1, Article ID 100004, 2021.