IEEE Open Journal of
Intelligent Transportation
Systems

Received 11 January 2023; revised 4 April 2023; accepted 21 May 2023. Date of publication 23 May 2023; date of current version 1 June 2023.

Digital Object Identifier 10.1109/OJITS.2023.3279264

Mixed Traffic Flow State Detection: A Connected Vehicles-Assisted Roadside Radar and Video Data Fusion Scheme

RUI CHEN (Member, IEEE), JIAMENG NING (Graduate Student Member, IEEE), YU LEI (Graduate Student Member, IEEE), YILONG HUI (Member, IEEE), AND NAN CHENG (Member, IEEE)

State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China

CORRESPONDING AUTHOR: R. CHEN (e-mail: rchen@xidian.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2019YFB1600100, and in part by the Guangdong Natural Science Fund for Distinguished Young Scholar under Grant 2023B1515020079.

ABSTRACT An increasing number of connected vehicles (CVs) driving together with regular vehicles (RVs) on the road is an inevitable stage of future traffic development. As accurate traffic flow state detection is essential for ensuring safe and efficient traffic, the level of road intelligence is being enhanced by the mass deployment of roadside perception devices, which is capable of sensing the mixed traffic flow consisting of RVs and CVs. In this background, we propose a roadside radar and camera data fusion framework to improve the accuracy of traffic flow state detection, which utilizes relatively more accurate traffic parameters obtained from real-time communication between CVs and roadside unit (RSU) as calibration values for training the back propagation (BP) neural network. Then, with the perception data collected by roadside sensors, the BP neural network-based data fusion model is applied to all vehicles including RVs. Furthermore, considering the changes of road environments, a dynamic BP fusion method is proposed, which adopts dynamic training by updating samples conditionally, and are applied to fuse traffic flow, occupancy and RVs speed data. Simulation results demonstrate that for CVs data and all vehicles (including RVs) data, the proposed dynamic BP fusion method is more accurate than single sensor detection, entropy based Bayesian fusion method and traditional BP fusion without training by CVs. It can achieve smaller error, and the accuracies of vehicle speed, traffic flow, and occupancy are all above 95%.

INDEX TERMS Connected vehicles, V2X communication, radar, video, data fusion, state detection.

I. INTRODUCTION

A. BACKGROUND

With the rapid development of intelligent driving technology, the composition of traffic flow is changed by the emergence of connected vehicle (CV). It is foreseeable that the traffic flow will consist of regular vehicle (RV) and CV in the next few decades [1]. As a result, the characteristics of traffic flow will be changed accordingly [2]. Therefore, it is necessary to detect mixed traffic flow to realize the fine management of the road. At present, in order to detect mixed traffic flow, roadside sensors such as radar and camera are adopted to sense the road environment. Moreover, sensor fusion technology has been introduced into the problem of traffic state detection [3]. It is expected to obtain more reliable and precise traffic state by combining data from different sensors [4], [5], [6]. The advantage is that the fusion results have higher accuracy, reliability and robustness [7]. On the other hand, communication technology has the potential to provide more accessible and available information for data fusion in transportation systems [8]. For example, wireless communication has been adopted in data fusion to estimate vehicle travel time [9] and solve the problem of positioning in Vehicular ad-hoc

The review of this article was arranged by Associate Editor Stefania Santini.

networks(VANETs) [10], which motivates our research on data fusion with the aid of communications for mixed traffic flow state detection.

B. RELATED WORK

A variety of sensors are widely employed for traffic state detection, including camera in [11], [12], [13], LiDAR sensor in [14], mmWave radar in [15], loop detector in [16], satellite remote sensing in [17], float car and probe vehicle in [18], [19], etc. Thus, different traffic flow parameters can be detected such as vehicle trajectory, speed, road space occupancy, density, and volume. Furthermore, through data fusion technology, the data of these sensors is expected to be fully utilized to obtain more comprehensive and accurate traffic state information. For example, fusing camera and LiDAR sensor data to recognize the environment around the vehicle [20], fusing floating car data (FCD) and stationary detector data (SDD) to accurately detect traffic flow, density, vehicle speed [21], fusing video and GPS-equipped vehicle data to accurately detect travel time and density [22]. However, loop detector has low detection accuracy and is not suitable for installation and maintenance, LiDAR is vulnerable to weather. Considering that camera can provide visual data, millimeter wave radar can provide high-precision and robust detection data that is not affected by weather. The fusion of the two can complement each other's functions. Therefore, they are the most commonly adopted combination of sensing devices. The fusion of them has priority over other sensor fusion [23].

A number of fusion methods based on radar and camera sensors have been proposed [24], [25], [26], [27], [28], [29]. In [24], in order to improve the accuracy of target detection, the author proposed a Camera Radar Fusion-Net (CRF-NET) which is employed to fuse the data of camera and radar, pointing out a new direction for sensor data fusion work. In [25], YOLOv3 convolutional neural network is used to detect the vehicle. The camera calibration algorithm is utilized to combine the radar targets and video vehicle detection targets, simultaneously measure the speed of multiple vehicles. In [26], the main idea of sensor fusion is to extract image patches according to regions of interest (ROIs), generated by radar points in camera coordinates. Finally, the target data are fused based on Bayesian estimation. In [27], a model implemented using the Robot Operating System (ROS) environment is proposed to handle radar and camera sensor synchronization and fusion for accurate object detection. In [28], a spatio-temporal synchronization method for roadside radar and camera is proposed to fuse targets and match vehicle trajectories. In [29], a radar and camera fusion sensing method is proposed to improve target recognition and tracking, which reduces the missed detection rate of autonomous vehicle environmental awareness in severe weather. These demonstrate that the fusion of information from each sensor provides more accurate and faster information than using each sensor individually [30].

In addition, as an important technology of intelligent transportation system, the key features of vehicle-to-everything(V2X) focus on ultra-reliable and low latency communication in vehicular networks [31]. At present, taking long-term evolution (LTE)-V2X and new radio (NR)-V2X as the main technical line is the development trend of future intelligent transportation system. 5G NR V2X is designed to complement LTE V2X [32], which enables the exchange of sensor data between vehicles, RSUs, devices of pedestrians, and V2X application servers [33]. Among them, vehicle-to-infrastructure (V2I) is an important module of V2X [34]. Vehicles can achieve a global awareness of the network around them, whereas the RSUs can improve their information about traffic through the local information obtained from the vehicles [35]. In [36], a traffic environment collection method based on V2I is proposed. Location information from different sources is provided through roadside cameras and vehicle GPS. Then, sensor data is fused in the RSU and mapped onto the credibility map provided to vehicles, which effectively alleviates traffic congestion.

However, to the best of our knowledge, existing researches focus more on target-level fusion, such as radar and video correlation, target detection, and target classification. There are few specific researches on data-level fusion of traffic parameters from radar and camera. Meanwhile, existing researches usually use the data from a certain historical day or period as training samples, which can easily lead to low accuracy of the fusion results. In addition, existing researches focus on the on-board sensors, and few studies have considered taking advantage of mixed traffic flow characteristics to improve the accuracy of roadside sensor fusion.

Therefore, this paper proposes a data-level sensor fusion method using CVs communication data to improve the detection accuracy of traffic parameters. Moreover, the method utilizes updated samples to maintain the accuracy of fusion for a longer period. Its main applications include directly transmitting the fusion results back to CVs through V2I for assisted driving. Besides, it can be used in the edge computing unit in the mixed traffic flow. For example, the dynamic BP method is employed in MEC to obtain more accurate parameters. Then, the real-time state parameters are used in the demonstration of the traffic monitoring system to assist the traffic management department in effectively controlling the road.

C. CONTRIBUTIONS

The main contributions of this paper are summarized as follows:

- We propose a CV assisted data fusion framework for accurate mixed traffic flow detection, which utilizes relatively more accurate traffic parameters obtained from 5G V2I communication between CVs and RSU as calibration values for training the BP neural network based data fusion model.

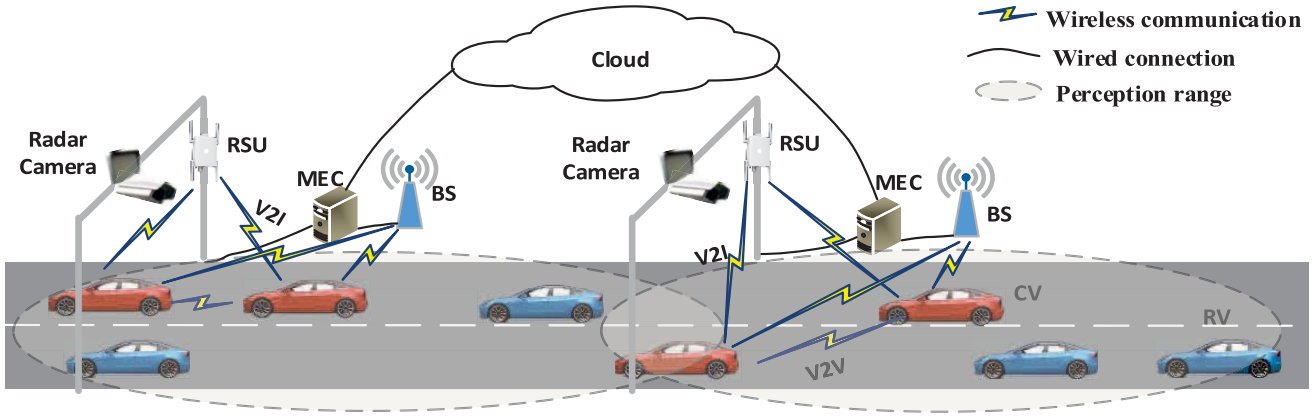


FIGURE 1. The scenario of mixed traffic flow based on roadside radar and video holographic perception.

- Considering the changes of road environments, we propose a dynamic BP fusion method that conditionally updates samples for training BP neural network, which are applied to fuse traffic flow, occupancy and RVs speed data. Simulation results show that the proposed data fusion method is more accurate than single sensor detection, the entropy based Bayesian fusion method and traditional BP fusion method.

The remainder of the paper is organized as follows. In Section II, we introduce the mixed traffic flow data fusion system framework upon which we raise questions of interest. In Section III, existing entropy based Bayesian and traditional BP neural network algorithm are reviewed. In Section IV, the fusion method of dynamic BP neural network is presented in detail. In Section V, numerical simulations and results of this study are presented, followed by conclusions in Section VI.

Notations: Bold lowercase letters denote vectors; bold uppercase letters denote matrices. $||\cdot||$ denotes the Frobenius norm, $\odot$ denotes the element-wise product and $(\cdot)^T$ denotes transpose. $(\cdot)'$ denotes the derivative. $\nabla(\cdot)$ is the partial derivative. $\mathbb{R}$ represents the set of real numbers.

II. MIXED TRAFFIC FLOW DATA FUSION SYSTEM

A. THE ROLE OF CV IN MIXED TRAFFIC FLOW

Vehicle-to-infrastructure (V2I) communication technology in mixed traffic flow can obtain more accurate and stable data in real-time [37]. Connected vehicles are able to share their state information (such as position, speed, acceleration, etc.) with infrastructure within a certain range through V2I communication. Due to the ability to share state information, compared with regular vehicles, connected vehicle has a shorter reaction time when it follows the car [2]. Therefore, it can contribute to the efficiency and safety of the transportation system. A number of studies have been devoted to the effect of CV on traffic flow. In [38], it is demonstrated that connected and autonomous vehicles can improve string stability, and its effects on throughput is also explored. In [39], analytical methods for the stability and the fundamental diagram models of mixed traffic flow are studied. Moreover, it is proved that intelligent and connected vehicles (ICVs) can improve the stability of mixed traffic flow under critical speed. It can be confirmed that CV plays an indispensable role in mixed traffic flow, which transmits perception information such as speed, position to the RSU. Due to the low delay of communication, the obtained communication data produce small errors, but its accuracy is still higher than that of roadside radar and camera sensors. Thus, we adopt the CVs communication data as the calibration value to train the neural network, input the CVs parameters corresponding to the radar and the camera, then obtain the trained fusion model, which is employed to fuse traffic flow, occupancy and RVs speed data. Previous studies in this area are very limited. We take advantage of the communication capabilities of CVs to detect all vehicles to improve the traffic efficiency of mixed traffic flow.

B. SYSTEM ARCHITECTURE

At present, a variety of advanced roadside sensors have been widely deployed to improve the perception of mixed traffic flow environment. In the system, roadside radar and camera sensors can detect the information of all the CVs and RVs on the road within their detection range. At the same time, CVs can transmit their real-time information to the roadside unit (RSU) through V2I communication, as shown in Fig. 1. The system is divided into data acquisition and data processing. The data acquisition part includes roadside radar, camera sensor and CVs. In contrast, the processing part includes RSU as communication terminals, mobile edge computing (MEC), base station (BS), as well as cloud platform. In mixed traffic flow, V2X technology can provide two communication modes. On the one hand, CVs can use the cellular mode to realize V2N communication with the BS through the UU interface for map downloading and other tasks. On the other hand, it can use direct mode to realize V2I communication with RSU through PC5 interface. RSU is responsible for receiving detection data of all

FIGURE 2. The demonstration of our developed radar and video fusion transportation monitoring system.

vehicles from different sensors and real-time communication data of CVs.

The data fusion process of mixed traffic flow can be summarized as follows. First, when CVs enter the coverage of an infrastructure, each CV can transmit its own vehicle information to the RSU through V2I communication, including ID, speed, location, etc. Simultaneously, RSU also receives the license plate, speed and other data of all vehicles including CV and RV within the detection range of the roadside perception sensors in the corresponding road section. Further, the roadside MEC can associate and fuse local sensing data from different sources. At last, different MEC transmit the fusion results to the cloud platform through optical fiber or UU interface to further splice the data of different road sections to obtain global road state fusion, so as to manage and control the road more effectively and improve operation efficiency. Simultaneously, the comprehensive traffic situation is displayed through the radar and video fusion transportation monitoring system, as shown in Fig. 2. The display interface shows detailed information of all vehicles on the road in the range of radar and camera, including real-time vehicle speed data, and traffic flow data, occupancy data and other statistics.

Based on the real-time and convenience of vehicle data acquisition, a fusion framework of roadside perception data assisted by connected vehicles is proposed. The fusion of perception data can improve the detection accuracy of real-time traffic flow state information to realize the identification and analysis of mixed traffic flow environment information.

III. REVIEW OF EXISTING PARAMETER FUSION METHODS

A. ENTROPY BASED BAYESIAN FUSION

A multi-sensor data fusion algorithm based on entropy weight method and Bayesian inference method is proposed [40]. The method has the following steps. First, $\mathbf{d}$ is a class variable representing different sensors, $\mathbf{c}_i = \{c_{i1}, c_{i2}, \ldots, c_{ij}\}$ is a feature variable consisting of observations, $i = 1, 2, \ldots, M$ and $j = 1, 2, \ldots, N$, where M is the number of sensors, and N is the number of traffic parameter samples. The probability density function (PDF) of $\mathbf{c}_i$ is expressed as

$$p(\mathbf{c}_i) = \frac{1}{(2\pi)^{M/2} \mid \boldsymbol{\Sigma}_i \mid^{1/2}} \exp\left\{-\left(\mathbf{c}_i - \boldsymbol{\mu}_i\right)^T \boldsymbol{\Sigma}_i{}^{-1} \left(\mathbf{c}_i - \boldsymbol{\mu}_i\right)\right\}, \tag{1}$$

where $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ are the maximum likelihood estimates of the mean and covariance respectively. The posterior probability $p(\mathbf{d}|\mathbf{c}_i)$ of each sensor is obtained according to the Bayes rule. Second, the joint entropy and mutual information of $\mathbf{c}_i$ and $\mathbf{d}$ are calculated using the above probabilities as

$$H(\mathbf{c}_i, \mathbf{d}) = -\sum_{j=1}^{N} \sum_{k=1}^{M} p\left(\mathbf{c}_j, \mathbf{d}_k\right) \log p\left(\mathbf{c}_j, \mathbf{d}_k\right), i = 1, 2, \ldots, M, \tag{2}$$

$$p(\mathbf{c}_i, \mathbf{d}_k) = p(\mathbf{d}_k|\mathbf{c}_i) p(\mathbf{c}_i), \tag{3}$$

$$I(\mathbf{c}_i, \mathbf{d}) = H(\mathbf{c}_i) + H(\mathbf{d}) - H(\mathbf{c}_i, \mathbf{d}). \tag{4}$$

Third, the weight of each sensor is given as

$$w_k = \frac{I_k}{I_1 + I_2 + \cdots + I_M}, k = 1, \ldots, M. \tag{5}$$

Based on (4), the fusion result can be expressed as

$$f_j = \sum_{j=1}^{M} w_j c_{ij}, \tag{6}$$

where c_{ij} is the jth observation of the ith sensor.

The limitation of this algorithm is that in reality, the acquisition of the prior probability and the conditional probability needs to be represented by the empirical value obtained by the sensor collecting the vehicle speed or traffic flow parameters in different weather and different time periods, which has a large deviation from the actual situation. Therefore, the data accuracy is low.

B. TRADITIONAL BP NEURAL NETWORK

BP neural network is a multi-layer feedforward neural network with neuron signal forward propagation and error back propagation. According to the statistics, 80%-90% of the neural network models have adopted BP network or its transformations [41]. In the current research, it can effectively process some complex nonlinear mapping data, which is an important method for processing traffic flow data. The actual speed, flow and other traffic data are used as the sequence of input vectors, and the gradient descent method is employed for training. During the training process, the output of each layer is obtained by sending the weighted sum of the input to an activation function $\sigma(\mathbf{x}) = \frac{1}{1+\mathbf{e}^{(-\mathbf{x})}}$, and the connection weights between layers are continuously adjusted. Finally, the network output value of the traffic parameter has the smallest error with the calibration value [42].

The traditional BP neural network algorithm has strong fault tolerance and anti-interference performance, and can meet the requirements of data processing of multi-sensor collection. However, it does not consider the adaptability of the algorithm when the traffic data changes.

IV. FUSION OF SENSORS DATA WITH THE PROPOSED FUSION FRAMEWORK

The Fig. 3 shows the specific process of mixed traffic flow data fusion architecture and data flow, mainly including data collection, dynamic neural network model construction and fusion application.

A. DATA COLLECTION

Existing researches usually select statistical traffic parameters such as flow [43], speed [44], occupancy [45], density [46] and travel time [47] to describe the traffic state within a certain period of time. Therefore, in the paper, the three most commonly used statistical parameters including flow, occupancy and speed data are adopted for simulation.

Currently, roadside sensors can collect all types of vehicle information including CVs and RVs within their respective detection ranges. HD cameras can monitor multiple lanes and have high environmental requirements. The embedded recognition software includes modules such as video acquisition, license plate detection, and data transmission, which can collect vehicle speed, traffic flow, and time occupancy. In addition, the GB/T 24726 standard specifies the parameters index of the traffic flow detector, requiring the visibility to be no less than 5500m, the sample size to be no less than 100, and the vehicle speed to be 0-120km/h. When the road illumination is not less than 5000lx, the accuracy of flow, vehicle speed and time occupancy is not less than 92%, 92% and 90% respectively; Otherwise, the accuracy of flow, vehicle speed and time occupancy is not less than 85%.

In the mixed traffic flow data fusion system, the radar and camera are installed on the crossbar on the roadside and are not blocked by vehicles with high height on the road, so as to detect the maximum range. We use the on-board GPS, IMU, etc. of CVs to obtain perception information such as vehicle speed and location. CVs transmit the speed and other sensing data corresponding to their own vehicle identifier (ID) to RSU through V2I communication. During the time interval T_{tol}, the road sections within the detection range of the sensor are observed, and the traffic flow data of CVs can be counted according to the license plate information uploaded by different CVs. CVs upload their own time of passing the section to the RSU for storage. At the same time, roadside radar and cameras separately measure the time that all types of vehicles pass through the section. The time occupancy data of CVs and all types of vehicles are obtained by formula

$$O = \frac{1}{T_{tol}} \sum_{v=1}^{V} T_v, \tag{7}$$

where O represents the time occupancy rate, T_{tol} represents the total observation time, T_v represents the time taken by the vth vehicle to pass through the section, and V represents the number of passing vehicles.

B. CONSTRUCTION OF DYNAMIC BP NEURAL NETWORK MODEL

The collected data is transmitted by RSU to MEC for network construction and data fusion. Due to only a few number of iterations being required for training, the roadside MEC is capable of performing real-time processing of the traffic data. Moreover, due to the fact that the updated samples are more consistent with the law of traffic flow state data in the current period, the fusion results are more real-time and have higher accuracy. In this paper, the parameters of the BP neural network model are determined into three categories, which are from radar, camera and V2I communication between CVs and RSU. The specific steps can be summarized as follows.

- First, dynamic BP uses N consecutive groups of CVs data collected from radar and camera within the observation time as input sequences, and N communication data of CVs within this time period as calibration value sequences. And aggregate them into a training group, which is divided into speed training group, flow training

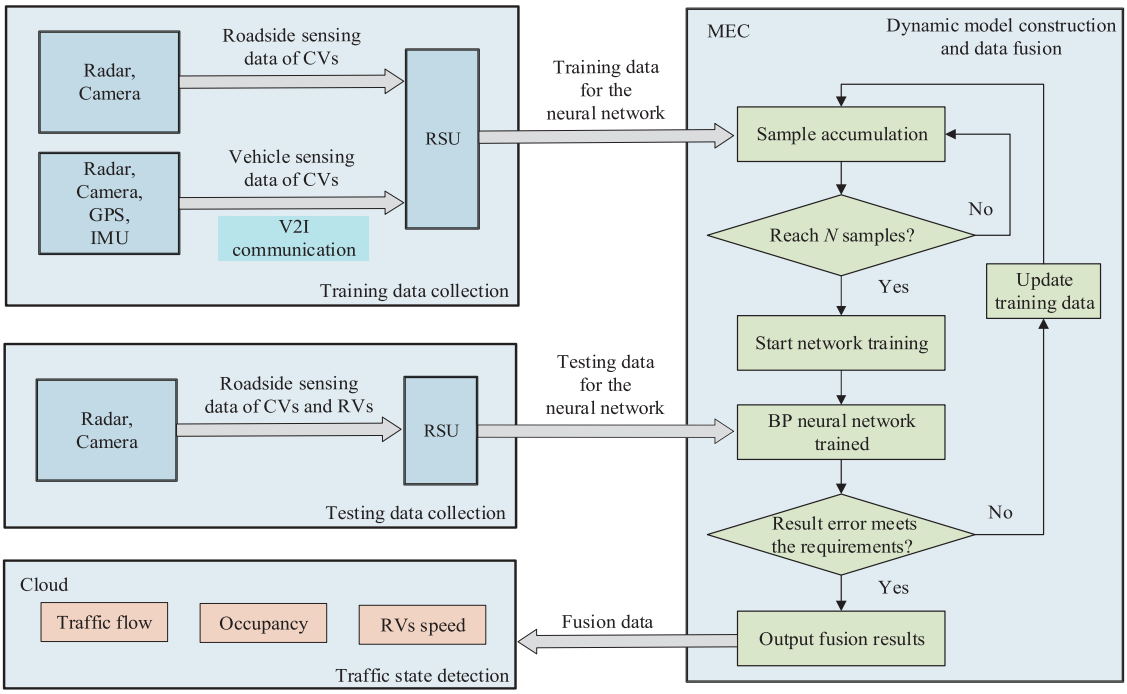


FIGURE 3. Mixed traffic flow data fusion architecture and data flow.

group and occupancy training group. Then immediately start training the fusion model.

- After training, a BP neural network fusion model is obtained. N_{tes} groups of radar and camera sensing data are directly input into the trained model for fusion, N_{tes} represents the number of test samples. The real-time fusion results are obtained and the accuracy of the results is calculated.
- If the accuracy meets the requirements, the results are directly output without having to train the fusion model again. Until the accuracy does not meet the requirements due to significant changes in road conditions or the accuracy of the camera deteriorates, the N groups of radar, camera and CVs data adjacent to the test data are used as updated samples to train the model in the MEC. Then, the N_{tes} groups and subsequent test data are fused using the retrained model to obtain real-time fusion results.
- Repeat the above dynamic update steps, and sequentially calculate the new result.

Among them, the working process of the neural network can be divided into two processes, one is the forward propagation process, and the other is the back propagation process [48].

In the first stage, the network receives the input of traffic data samples, the neurons are activated from the hidden layer to the output layer, and the output value of each node is calculated layer by layer until the summary calculation of each node is completed. In the second stage, if there is a difference between the output value and the calibration value, the error starts from the output layer and goes back through the hidden

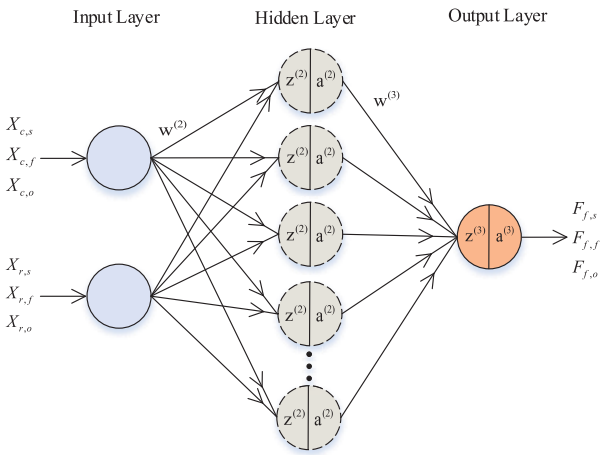


FIGURE 4. The structure of BP neural network.

layers to correct the weights between the neurons. Then go back to the input layer and enter the forward propagation again. Such repeated forward propagation and back propagation process is a process of continuously adjusting the weights between each neuron until the number of iterations is reached or the output layer obtains a value whose error with the calibration value meets the requirements. Fig. 4 shows the three-layer BP neural network structure designed in this paper. A neural network composed of multiple neurons includes L layers: input layer of two sensor measurements, $X_{c,s}, X_{c,f}, X_{c,o}$ are the speed, flow, occupancy data detected by the camera, $X_{r,s}, X_{r,f}, X_{r,o}$ are the speed, flow, occupancy data detected by the radar, $(L-2)$ hidden layer and the output layer with value corresponding to the fusion result $F_{c,s}, F_{c,f}, F_{c,o}$. Each layer can be interconnected, while the

neurons in the same layer are independent of each other, and the hidden layer can be set to a single layer. During the forward propagation of the neural network, sensor measurements are used as the input of the first layer, and the state and neuron output values of the lth layer are calculated as

$$\mathbf{z}^{(l)} = \mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}, \tag{8}$$

$$\mathbf{a}^{(l)} = \sigma\left(\mathbf{z}^{(l)}\right), \tag{9}$$

where $\mathbf{W}^{(l)} \in \mathbb{R}^{n_l \times n_{l-1}}$ is given as the weight matrix from layer $(l-1)$ to layer l; n_l is the number of neurons in layer l. $\mathbf{w}^{(l)}$ is an element in the weight matrix $\mathbf{W}^{(l)}$. $\mathbf{b}^{(l)}$ is given as the bias from layer $(l-1)$ to layer l; Parameters $\mathbf{W}^{(l)}$ and $\mathbf{b}^{(l)}$ are first initialized with extremely small random values. $\mathbf{z}^{(l)}$ represents the state of neurons in layer l; And $\mathbf{a}^{(l)}$ represents the output value of the neuron in layer l; $\sigma(\cdot)$ is an activation function.

The second step is the back propagation process. The cost function for training data can be defined as

$$E(\mathbf{W}, \mathbf{b}) = \frac{1}{2}\sum || \mathbf{y} - \mathbf{a} ||^2, \tag{10}$$

where $\mathbf{y}$ represents the expected output given by the training data, and $\mathbf{a}$ represents the actual output produced by the neural network for the input $\mathbf{x}$. In order to minimize E, the gradient descent method is used to update the weight. Let $\boldsymbol{\delta}$ represent the defined local gradient. Then, the local gradient of the output layer and hidden layer can be expressed separately as

$$\boldsymbol{\delta}^{(L)} = -\left(\mathbf{y} - \mathbf{a}^{(L)}\right) \odot \sigma'\left(\mathbf{z}^{(L)}\right), \tag{11}$$

$$\boldsymbol{\delta}^{(l)} = \left(\left(\mathbf{W}^{(l+1)}\right)^T \boldsymbol{\delta}^{(l+1)}\right) \odot \sigma'\left(\mathbf{z}^{(l)}\right). \tag{12}$$

Then, the partial derivatives of the cost function with respect to the parameters are given as

$$\nabla_{\mathbf{W}^{(l)}} E = \boldsymbol{\delta}^{(l)}\left(\mathbf{a}^{(l-1)}\right)^T, \tag{13}$$

$$\nabla_{\mathbf{b}^{(l)}} E = \boldsymbol{\delta}^{(l)}. \tag{14}$$

Finally, the updated weight and bias values can be calculated by

$$\mathbf{W}^{(l)} = \mathbf{W}^{(l)} - \eta \frac{\partial E(\mathbf{W}, \mathbf{b})}{\partial \mathbf{W}^{(l)}}, \tag{15}$$

$$\mathbf{b}^{(l)} = \mathbf{b}^{(l)} - \eta \frac{\partial E(\mathbf{W}, \mathbf{b})}{\partial \mathbf{b}^{(l)}}. \tag{16}$$

Repeat the above process continuously until the error meets the requirements and end.

The training process of dynamic BP neural network is summarized in Algorithm 1, in which $(\mathbf{x}, \mathbf{y}) = \{(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), (\mathbf{x}^{(2)}, \mathbf{y}^{(2)}), \ldots, (\mathbf{x}^{(N)}, \mathbf{y}^{(N)})\}$ is a set of N training patterns (input-output pair), $\mathbf{x}^{(j)}$ is the input vector in the M-dimensional pattern space, M is the number of sensors, $\mathbf{F}_j, j = 1, 2, \ldots, N$ is the corresponding fusion value, $\eta, \varepsilon, t_{max}, L$, are the learning rate of error back propagation, the minimum error of the training target, the number of iterations and the number of neural network layers respectively. The algorithm obtains a trained model by continuously updating the weights and bias parameters $\mathbf{w}^{(l)}, \mathbf{b}^{(l)}$ of each layer of the network. Then we input the testing data to be fused into the trained model and output the fusion result. Finally, the training data is updated according to the calculated result accuracy to repeat the above steps.

Algorithm 1 The Training Process of Dynamic BP Neural Network

Input: $(\mathbf{x}, \mathbf{y})$, η, ε, t_{max}, L, N
Output: $\mathbf{F}_j, j \in [1, N]$

procedure NEURAL-NETWORK($(\mathbf{x}, \mathbf{y})$, η, ε, t_{max}, L, N)
  while The number of samples reaches N **do**
    Initialize $\mathbf{W}^{(l)}$, $\mathbf{b}^{(l)}$
    for $l = 2{:}L$ **do**
      $\mathbf{z}^{(l)} = \mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}$
      $\mathbf{a}^{(l)} = \sigma(\mathbf{z}^{(l)})$
    end
    $\boldsymbol{\delta}^{(L)} \leftarrow$ use (11)
    for $l = (L-1){:}2$ **do**
      $\boldsymbol{\delta}^{(l)} \leftarrow$ use (12)
    end
    while $(\Delta\mathbf{b} > \varepsilon) or (t < t_{max})$ **do**
      $t = 0$
      for $l = 2{:}L$ **do**
        $t = t + 1$
        $\mathbf{W}^{(l)} \leftarrow$ use (15)
        $\mathbf{b}^{(l)} \leftarrow$ use (16)
      end
    end
    return $\mathbf{F}_j, j \in [1, N]$
  end
  $ACC \leftarrow$ use (19)
  if $ACC < 95\%$ *do*
    Update $(\mathbf{x}, \mathbf{y})$
    Back to step 2
  end
end procedure

C. COMPUTATIONAL COMPLEXITIES

Dynamic BP updates samples based on BP neural networks to train fusion models. The neural network includes L layers, with the number of neurons in each layer being n_{l_1}, n_{l_2}, n_L. Its time complexity is determined by forward propagation and backward propagation. During forward propagation, each training sample will perform $L-1$ matrix calculations. The number of neurons in the input and output layers are determined, so the computational complexity of each sample is $\mathcal{O}(n_{max}^2)$, where n_{max} is the maximum number of neurons. The computational complexity of the back propagation is the same as that of the forward propagation, and N samples are trained once, it can be obtained that the time complexity of training a neural network is $\mathcal{O}(N * n_{max}^2)$. When the accuracy

of the parameter fusion results meets the requirements, it is not necessary to train the neural network again, so its computational complexity is $\mathcal{O}(N * n_{max}^2)$. When the accuracy of the fusion results does not meet the requirements, it is necessary to update the samples and retrain the appropriate model, so the computational complexity of the dynamic BP algorithm is $\mathcal{O}(2N * n_{max}^2)$. The higher fusion accuracy is achieved as sacrifice algorithm complexity.

D. EVALUATION INDICATORS

This paper adopts formulas to calculate the fusion result accuracy, root mean square error (RMSE), and mean absolute error (MAE) to evaluate the model. RMSE is the root mean square error, which is used to measure the deviation between the calibration value and the fusion value. It is also a way to describe the accuracy of the fusion model. The smaller the value, the more accurate the fusion model is. It can be written as

$$RMSE = \sqrt{\frac{1}{N_{tes}} \sum_{t=1}^{N_{tes}} (R_t - F_t)^2}. \tag{17}$$

MAE is the mean absolute error, which prevents the errors from canceling each other to better reflect the actual situation of the error. It can be given as

$$MAE = \frac{\sum_{t=1}^{N_{tes}} |R_t - F_t|}{N_{tes}}. \tag{18}$$

According to the GB/T 20609 standard, ACC is defined as the average detection accuracy, which can be written as

$$ACC = \frac{1}{N_{tes}} \sum_{t=1}^{N_{tes}} \left(1 - \frac{|R_t - F_t|}{R_t} \times 100\%\right). \tag{19}$$

In the above formulas, R_t represents the calibration value of the connected vehicle in one direction, as the calibration value. F_t represents the fusion value of radar and camera sensors in one direction. N_{tes} represents the number of samples in the test group. In order to evaluate the fusion effect of the model, this paper compares the results of the dynamic BP fusion with the traditional BP fusion and entropy based Bayesian fusion method, and further highlights the adaptability of the data fusion model based on the dynamic BP neural network.

V. SIMULATION RESULTS AND DISCUSSION

A. SIMULATION SETTINGS AND RESULTS

This simulation adopts the gradient descent method as the training method of the network. Other simulation parameters are set in the Table 1. The datasets we use include speed, traffic flow, occupancy data measured by sensors as well as connected vehicles through V2I communication. The trained model is used to different state data fusion of roadside radar and camera sensors. Finally, we draw a comparison curve between the fusion result and the calibration value. RMSE, MAE and the accuracy are employed as evaluation indicators to evaluate the results.

TABLE 1. Network parameter settings.

Parameters	Value
The training function	The trainlm function
The hidden layer transfer function	The tansig function
The output layer transfer function	The purelin function
The learning rate η	0.01
The maximum number of training	10^3
The target error	10^{-6}
Training group of speed data	230
Testing group of speed data	70
Training group of flow or occupancy data	190
Testing group of flow or occupancy data	30

In the simulation, different numbers of CVs sample data are selected lots of times to train the network, and the fusion effects of the trained models are compared. The sample data is the characteristic value of the network input.

When training fusion models, there are certain requirements for setting the samples. Firstly, the updated training samples should correctly reflect the general laws of the current traffic environment. Therefore, the sensor and CV communication data adjacent to the test group are adopted as updated training samples. Secondly, the number of training samples remains consistent before and after the update. In this manuscript, the number of training samples is set to N. Finally, the number of training samples cannot be too small or too large, as too few training samples lead to increased inaccuracy of the model. An appropriate increase in samples improve the accuracy of training results, but excessive training samples lead to an increase in training volume and time costs. At the same time, there are significant changes in the sample rules, resulting in a decrease in the accuracy of the results. In this manuscript, according to the results shown in Table 2, the sample numbers of three different types of data are chosen to be 230, 190, and 190, respectively. The threshold values for the flow sample and the speed sample are 80. Since occupancy data is calculated based on traffic data, its sample number is set to be the same as the flow sample number.

Our method is compared with the entropy based Bayesian fusion method described above and with single-source data before fusion to show the advantages. Fig. 5 shows the comparison of the speed, flow, and occupancy of CVs before and after fusion. From Fig. 5(a), (b) and (c), it can be seen that CVs communication data is particularly close to the real value, proving its high accuracy. Moreover, V2I communication can meet the real-time requirements, so it can be adopted to train the fusion network model. In terms of overall trends, both the individual sensor data and the fusion result values agree with the target output. In contrast, the deviation of the camera perception data is the largest, and the effect is poor. The results obtained by dynamic BP neural network model (red solid curve) are generally the closest to

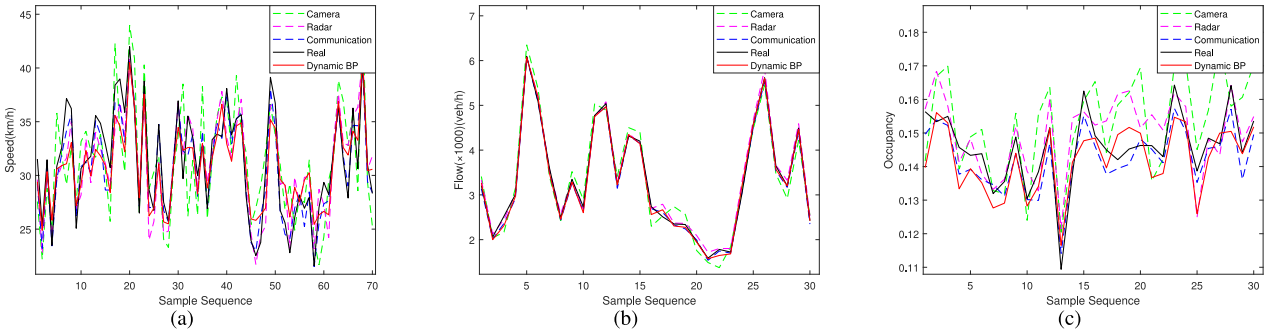


FIGURE 5. Comparison of different data of CVs before and after fusion: (a) CVs speed data comparison, (b) CVs flow data comparison, (c) CVs occupancy data comparison.

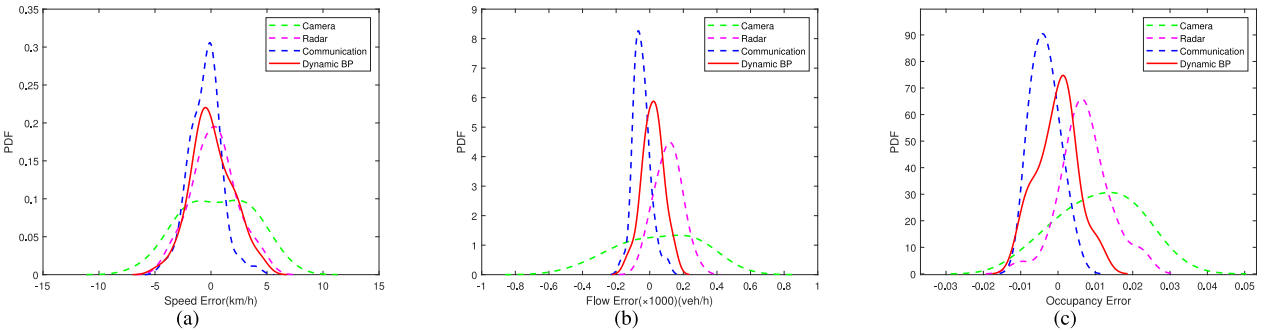


FIGURE 6. Error distribution of different data for CVs before and after fusion: (a) CVs speed data error, (b) CVs flow data error, (c) CVs occupancy data error.

TABLE 2. Fusion results using different training samples.

Type of Data	Number of training samples	Accuracy
Speed	290	96.19%
	230	96.22%
	170	95.93%
Flow	210	95.84%
	190	96.0%
	130	95.67%
Occupancy	210	95.81%
	190	95.88%
	130	95.87%

TABLE 3. Error evaluation index of different models.

Type of Data	Error	Bayes	BP	**Dynamic BP**
Speed	MAE	1.4806	1.3359	1.1332
	RMSE	1.971	1.8699	1.4502
Flow	MAE	0.1418	0.1572	0.1252
	RMSE	0.1738	0.1872	0.1491
Occupancy	MAE	0.0109	0.007	0.0058
	RMSE	0.0131	0.0089	0.0069

the target output (blue dotted curve). From Fig. 6(a), (b) and (c), it can be seen that the error distribution of different data of CVs before and after fusion. The errors of individual sensors and communication data tend to Gaussian distribution. The error generated after the fusion of the dynamic BP neural network model also tends to a Gaussian distribution, and its distribution is more concentrated than that of a single sensor, which proves that the fusion result obtained by using the fusion model has higher accuracy and satisfactory results.

Then, we utilize the RV data for testing. Specifically, it can be seen from Table 3 that for the speed, flow, occupancy data, the error generated by the dynamic BP model fusion is the smallest, the fusion value is numerically closer to the calibration value. The mean absolute error (MAE) of the fusion model based on dynamic BP neural network are 1.1332, 0.1252, and 0.0058 respectively, which are lower than the MAE produced by the other two fusion models. Moreover, the root mean square error (RMSE) generated by the dynamic BP model is also the lowest, which are 1.4502, 0.1491, and 0.0069. It shows better fusion performance. Fig. 7 shows the error distribution of different data of RVs before and after fusion. From Fig. 7(a), (b) and (c), it can be seen that the error value generated by using RVs data for fusion also tends to a Gaussian distribution, and the probability of small errors increases, which proves that the fusion model has good results for the data fusion of RVs.

As can be seen from Fig. 8, for the parameters of RVs, the accuracy of dynamic BP network fusion is significantly improved, the fluctuation of the accuracy curve is reduced and the value is closer to 1. When the camera accuracy is poor, the results are affected, but the model still shows good fusion performance. As shown in Table 4, the fusion results

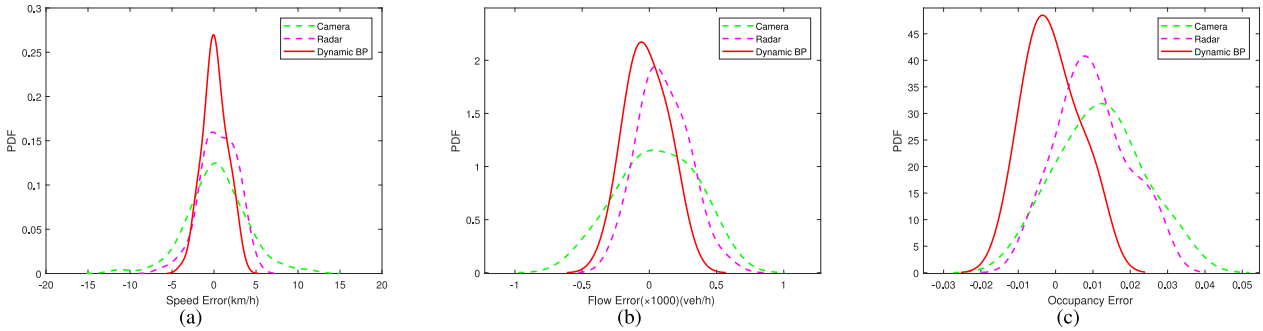


FIGURE 7. Error distribution of different data before and after fusion: (a) RVs speed data error, (b) traffic flow data error, (c) occupancy data error.

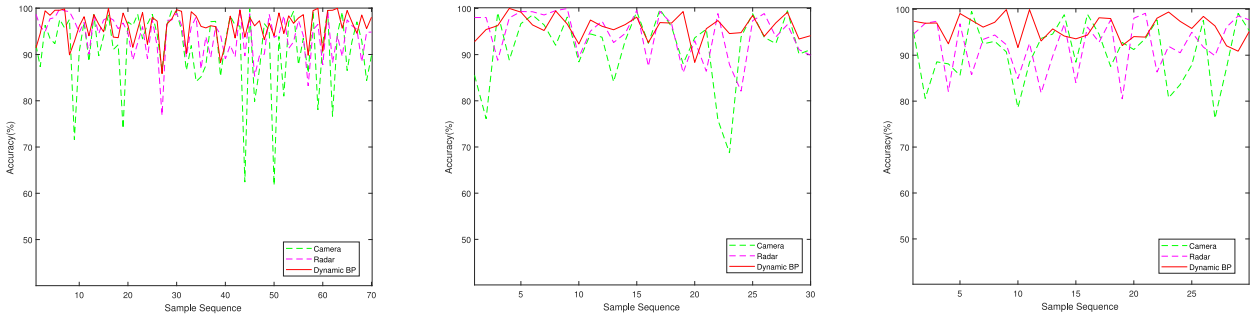


FIGURE 8. Accuracy of results using dynamic BP fusion models: (a) RVs speed data, (b) traffic flow data, (c) occupancy data.

TABLE 4. Average accuracy of single-source data and results by different fusion models.

Type of Data	Camera	Radar	Bayes	BP	**Dynamic BP**
Speed	91.34%	94.13%	95.17%	95.40%	96.22%
Flow	91.80%	94.44%	95.0%	94.91%	96.0%
Occupancy	90.67%	92.19%	92.36%	95.06%	95.88%

using the dynamic BP neural network model can be further improved on the basis of the traditional BP model, and the accuracy of the results is above 95%.

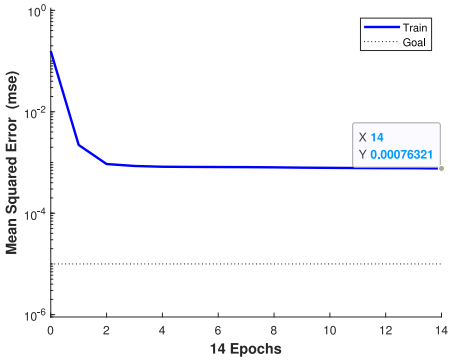
The improved module of the proposed dynamic BP method based on the traditional BP method is to calculate the accuracy of the first fusion results, and update the training samples for the second fusion. Therefore, it is necessary to compare traditional BP with dynamic BP in order to clearly understand the effectiveness of the proposed method.

B. MODEL CONVERGENCE

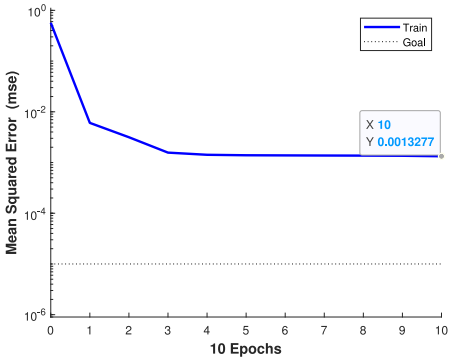
In the process of training the network model, the correction of weight values between each layer reduce the error. When the training is terminated, the error between the output value and the calibration value is less than the set error threshold, which means that the network model converges. In this manuscript, we have set an error threshold of 10^{-6}. Taking occupancy data as an example, we use the same 190 groups of samples to train the traditional BP and the dynamic BP model, and then use the same 30 groups of test samples to input the two trained models for fusion. When the accuracy of result does not meet the requirements, 190 groups of samples are updated to retrain the model, and the 30 groups of test samples are fused. Finally, the convergence of the algorithm is shown in Fig. 9, in which the output continuously decreases with the number of training times. The black dotted line indicates a preset error threshold. The traditional BP ends after 14 training epochs, and the dynamic BP ends after 10 training epochs. There is no significant difference in the final mean square error and training times between the two algorithm models, that is, the convergence speeds of the two methods are basically the same.

VI. CONCLUSION

To accurately detect mixed traffic flow states, we propose a roadside sensor data fusion framework, which takes into account the characteristics of mixed traffic flow and makes use of relatively more accurate data obtained from V2I real-time communication. Specifically, the V2I data between CVs and infrastructure is adopted as the calibration value of BP neural network to construct a model for fusing parameters including flow, occupancy and speed. Furthermore, we adopt a dynamic BP fusion method that conditionally updates samples for training BP neural network to adapt to the changes in the traffic environment and keep the accuracy for a long time. The simulation results show that for CVs data and all vehicles (including RVs) data, the fusion results of the dynamic BP model are more accurate than the fusion results of single sensor detection, entropy based Bayesian fusion



(a) traditional BP



(b) dynamic BP

FIGURE 9. The convergence of the algorithm.

and traditional BP fusion without CVs training. The dynamic BP method can achieve smaller errors, and the accuracies of vehicle speed, traffic flow, and occupancy are all above 95%.

REFERENCES

[1] Y. Wang, X. Yang, H. Liang, and Y. Liu, "A review of the self-adaptive traffic signal control system based on future traffic environment," *J. Adv. Transp.*, vol. 2018, Jun. 2018, Art. no. 1096123.

[2] D.-F. Xie, X.-M. Zhao, and Z. He, "Heterogeneous traffic mixing regular and connected vehicles: Modeling and stabilization," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 2060–2071, Jun. 2019.

[3] N.-E. El Faouzi, "Data fusion in road traffic engineering: An overview," in *Proc. Multisens. Multisource Inf. Fusion Archit. Algorithms Appl.*, Apr. 2004, pp. 360–371.

[4] F. A. Butt, J. N. Chattha, J. Ahmad, M. U. Zia, M. Rizwan, and I. H. Naqvi, "On the integration of enabling wireless technologies and sensor fusion for next-generation connected and autonomous vehicles," *IEEE Access*, vol. 10, pp. 14643–14668, 2022.

[5] J. Fayyad, M. A. Jaradat, D. Gruyer, and H. Najjaran, "Deep learning sensor fusion for autonomous vehicle perception and localization: A review," *Sensors*, vol. 20, no. 15, p. 4220, 2020.

[6] D. J. Yeong, G. Velasco-Hernandez, J. Barry, and J. Walsh, "Sensor and sensor fusion technology in autonomous vehicles: A review," *Sensors*, vol. 21, no. 6, p. 2140, 2021.

[7] J. Van Brummelen, M. O'Brien, D. Gruyer, and H. Najjaran, "Autonomous vehicle perception: The technology of today and tomorrow," *Transp. Res. C, Emerg. Technol.*, vol. 89, pp. 384–406, Apr. 2018.

[8] N.-E. El Faouzi, H. Leung, and A. Kurian, "Data fusion in intelligent transportation systems: Progress and challenges—A survey," *Inf. Fusion*, vol. 12, no. 1, pp. 4–10, 2011.

[9] R. D. Kühne, "Data fusion for dynamic route guidance systems," *IFAC Proc. Vol.*, vol. 30, no. 8, pp. 1319–1323, 1997.

[10] K. Golestan, F. Sattar, F. Karray, M. Kamel, and S. Seifzadeh, "Localization in vehicular ad hoc networks using data fusion and V2V communication," *Comput. Commun.*, vol. 71, pp. 61–72, Nov. 2015.

[11] X. Li, Y. She, D. Luo, and Z. Yu, "A traffic state detection tool for freeway video surveillance system," *Procedia Soc. Behav. Sci.*, vol. 96, pp. 2453–2461, Nov. 2013.

[12] J. Li, Z. Xu, L. Fu, X. Zhou, and H. Yu, "Domain adaptation from daytime to nighttime: A situation-sensitive vehicle detection and traffic flow parameter estimation framework," *Transp. Res. C, Emerg. Technol.*, vol. 124, Mar. 2021, Art. no. 102946.

[13] T. M. Deng and B. Li, "A detection method of traffic parameters based on EPI," *Procedia Eng.*, vol. 29, pp. 3054–3059, Feb. 2012.

[14] J. Zhao, H. Xu, H. Liu, J. Wu, Y. Zheng, and D. Wu, "Detection and tracking of pedestrians and vehicles using roadside LiDAR sensors," *Transp. Res. C, Emerg. Technol.*, vol. 100, pp. 68–87, Mar. 2019.

[15] Q. J. O. Tan and R. A. Romero, "Ground vehicle target signature identification with cognitive automotive radar using 24–25 and 76–77 GHz bands," *IET Radar Sonar Navig.*, vol. 12, no. 12, pp. 1448–1465, 2018.

[16] S. M. A. Rizvi, A. Ahmed, and Y. Shen, "Real-time incident detection and capacity estimation using loop detector data," *J. Adv. Transport.*, vol. 2020, Oct. 2020, Art. no. 8857502.

[17] S. A. Ahmadi, A. Ghorbanian, and A. Mohammadzadeh, "Moving vehicle detection, tracking and traffic parameter estimation from a satellite video: A perspective on a smarter city," *Int. J. Remote Sens.*, vol. 40, no. 22, pp. 8379–8394, 2019.

[18] B. S. Kerner et al., "Traffic state detection with floating car data in road networks," in *Proc. IEEE Int. Conf. Intell. Transp. Syst. (ITS)*, 2005, pp. 44–49.

[19] T. Seo, T. Kusakabe, and Y. Asakura, "Estimation of flow and density using probe vehicles with spacing measurement equipment," *Transp. Res. C, Emerg. Technol.*, vol. 53, pp. 134–150, Apr. 2015.

[20] G. H. Lee, J. D. Choi, J. H. Lee, and M. Y. Kim, "Object detection using vision and LiDAR sensor fusion for multi-channel V2X system," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, 2020, pp. 1–5.

[21] M. Treiber, A. Kesting, and R. E. Wilson, "Reconstructing the traffic state by fusion of heterogeneous data," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 26, no. 6, pp. 408–419, 2011.

[22] R. A. Anand, L. Vanajakshi, and S. C. Subramanian, "Traffic density estimation under heterogeneous traffic conditions using data fusion," in *Proc. IEEE Intell. Veh. Symp.*, 2011, pp. 31–36.

[23] X. Tang, Z. Zhang, and Y. Qin, "On-road object detection and tracking based on radar and vision fusion: A review," *IEEE Intell. Transp. Syst.*, vol. 14, no. 5, pp. 103–128, Sep./Oct. 2022.

[24] F. A. Jibrin, Z. Deng, and Y. Zhang, "An object detection and classification method using radar and camera data fusion," in *Proc. Int. Conf. Signal Inf. Data Process.*, 2019, pp. 1–6.

[25] J. Ren, Y. Wang, Y. Han, and R. Zhang, "Information fusion of digital camera and radar," in *Proc. IEEE MTT-S Int. Microw. Biomed. Conf.(IMBioC)*, vol. 1, 2019, pp. 1–4.

[26] Z. Yu, J. Bai, S. Chen, L. Huang, and X. Bi, "Camera-radar data fusion for target detection via Kalman filter and Bayesian estimation," SAE, Warrendale, PA, USA, SAE Technical Paper, 2018.

[27] R. Kumar and S. Jayashankar, "Radar and camera sensor fusion with ROS for autonomous driving," in *Proc. 5th Int. Conf. Image Inf. Process.*, 2019, pp. 568–573.

[28] Y. Du, B. Qin, C. Zhao, Y. Zhu, J. Cao, and Y. Ji, "A novel spatio-temporal synchronization method of roadside asynchronous MMW radar-camera for sensor fusion," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 22278–22289, Nov. 2022.

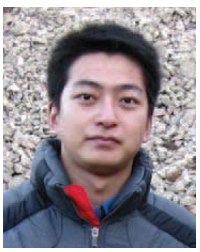
[29] Z. Liu et al., "Robust target recognition and tracking of self-driving cars with radar and camera information fusion under severe weather conditions," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 6640–6653, Jul. 2022.

[30] S. Park, J. A. Jang, and J. Kim, "Infrastructure based vehicle recognition system with multi sensor fusion," in *Proc. ICSPCS*, Dec. 2013, pp. 1–5.

[31] C. R. Storck and F. Duarte-Figueiredo, "A survey of 5G technology evolution, standards, and infrastructure associated with vehicle-to-everything communications by Internet of Vehicles," *IEEE Access*, vol. 8, pp. 117593–117614, 2020.

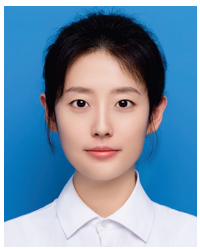
[32] S. Chen et al., "Vehicle-to-everything (V2X) services supported by LTE-based systems and 5G," *IEEE Commun. Standards Mag.*, vol. 1, no. 2, pp. 70–76, Jul. 2017.

[33] M. H. C. Garcia et al., "A tutorial on 5G NR V2X communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1972–2026, 3rd Quart., 2021.
[34] S. Chen, J. Hu, Y. Shi, L. Zhao, and W. Li, "A vision of C-V2X: Technologies, field testing, and challenges with Chinese development," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 3872–3881, May 2020.
[35] J. Barrachina et al., "V2X-d: A vehicular density estimation system that combines V2V and V2I communications," in *Proc. IEEE Wireless Days (WD) IFIP*, 2013, pp. 1–6.
[36] X. Zhao, K. Mu, F. Hui, and C. Prehofer, "A cooperative vehicle-infrastructure based urban driving environment perception method using a D-S theory-based credibility map," *Optik*, vol. 138, pp. 407–415, Jun. 2017.
[37] Y. Mo, P. Zhang, Z. Chen, and B. Ran, "A method of vehicle-infrastructure cooperative perception based vehicle state information fusion using improved Kalman filter," *Multimed. Tools. Appl.*, vol. 81, pp. 4603–4620, Feb. 2021.
[38] A. Talebpour and H. S. Mahmassani, "Influence of connected and autonomous vehicles on traffic flow stability and throughput," *Transp. Res. C, Emerg. Technol.*, vol. 71, pp. 143–163, Oct. 2016.
[39] X. Chang, H. Li, J. Rong, X. Zhao, and A. Li, "Analysis on traffic stability and capacity for mixed traffic flow with platoons of intelligent connected vehicles," *Phys. A*, vol. 557, Nov. 2020, Art. no. 124829.
[40] Z. Sun, M. Guo, W. Liu, J. Feng, and J. Hu, "Multisource traffic data fusion with entropy based method," in *Proc. Int. Conf. Artif. Intell. Comput. Intell.*, vol. 4, 2009, pp. 506–509.
[41] X. Pan, W. Zhou, Y. Lu, and N. Sun, "Prediction of network traffic of smart cities based on DE-BP neural network," *IEEE Access*, vol. 7, pp. 55807–55816, 2019.
[42] V. G. Gudise and G. K. Venayagamoorthy, "Comparison of particle swarm optimization and backpropagation as training algorithms for neural networks," in *Proc. IEEE Swarm Intell. Symp.*, 2003, pp. 110–117.
[43] D. Xu, C. Wei, P. Peng, Q. Xuan, and H. Guo, "GE-GAN: A novel deep learning framework for road traffic state estimation," *Transp. Res. C, Emerg. Technol.*, vol. 117, Aug. 2020, Art. no. 102635.
[44] E. F. Grumert and A. Tapani, "Traffic state estimation using connected vehicles and stationary detectors," *J. Adv. Transp.*, vol. 2018, pp. 1–14, Jan. 2018.
[45] E. Kidando, R. Moses, E. E. Ozguven, and T. Sando, "Evaluating traffic congestion using the traffic occupancy and speed distribution relationship: An application of Bayesian Dirichlet process mixtures of generalized linear model," *J. Transp. Technol.*, vol. 7, no. 3, p. 318, 2017.
[46] J. Xing, W. Wu, Q. Cheng, and R. Liu, "Traffic state estimation of urban road networks by multi-source data fusion: Review and new insights," *Physica A, Stat. Mech. Appl.*, vol. 592, Jun. 2022, Art. no. 127079.
[47] A. Bhaskar, T. Tsubota, L. M. Kieu, and E. Chung, "Urban traffic state estimation: Fusing point and zone based data," *Transp. Res. C, Emerg. Technol.*, vol. 48, pp. 120–142, Nov. 2014.
[48] A. K. Qin, V. L. Huang, and P. N. Suganthan, "Differential evolution algorithm with strategy adaptation for global numerical optimization," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 398–417, Apr. 2009.



RUI CHEN (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in communications and information systems from Xidian University, Xi'an, China, in 2005, 2007, and 2011, respectively. From 2014 to 2015, he was a Visiting Scholar with Columbia University, New York. He is currently an Associate Professor and a Ph.D. Supervisor with the School of Telecommunications Engineering, Xidian University. He is also the Director of INFOHAND Tech.-Xidian Guangzhou Institute Joint Intelligent Transportation Research Center. He has published over 80 papers in international journals and conferences and held 40 patents. His research interests include broadband wireless communication systems, array signal processing, and intelligent transportation systems. He serves as an Associate Editor for *International Journal of Electronics, Communications, and Measurement Engineering* (IGI Global), an Editor for *Information Countermeasure Technology*, and a guest editor for several journals. He is a member of the IMT-2030 Promotion Group.

JIAMENG NING (Graduate Student Member, IEEE) received the B.S. degree in electrical information engineering from the Chengdu University of Information Technology, Chengdu, China, in 2020. She is currently pursuing the M.S. degree in communications and information systems with Xidian University. Her research interests include intelligent transportation system and sensor fusion.



YU LEI (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering and automation from the Wuhan University of Science and Technology, Wuhan, China, in 2020. She is currently pursuing the M.S. degree in communications and information systems with Xidian University. Her research interests include intelligent transportation system and sensor fusion.

YILONG HUI (Member, IEEE) received the Ph.D. degree in control theory and control engineering from Shanghai University, Shanghai, China, in 2018. He is currently an Associate Professor with the State Key Laboratory of Integrated Services Networks and the School of Telecommunications Engineering, Xidian University, China. He has published over 50 scientific articles in leading journals and international conferences. His research interests include wireless communication, mobile-edge computing, vehicular networks, intelligent transportation systems, and autonomous driving. He was the recipient of the Best Paper Award of International Conference WiCon2016 and IEEE Cyber-SciTech2017.

NAN CHENG (Member, IEEE) received the B.E. and M.S. degrees from the Department of Electronics and Information Engineering, Tongji University, Shanghai, China, in 2009 and 2012, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo in 2016. He worked as a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, University of Toronto from 2017 to 2019. He is currently a Professor with the State Key Laboratory of Integrated Services Networks and the School of Telecommunications Engineering, Xidian University, Shaanxi, China. He has published over 70 journal papers in IEEE TRANSACTIONS and other top journals. His current research focuses on B5G/6G, space–air–ground-integrated network, big data in vehicular networks, and self-driving system. His research interests also include AI-driven future networks. He serves as an Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, and *Peer-to-Peer Networking and Applications*, and serves/served as a guest editor for several journals.