Joint Multislice and Cooperative Detection Aided Residual Network for Scenario Identification in Vehicle-to-Vehicle Communication Systems

Jiawei Huang, Yuxin Ji, Jie Yang, Miao Liu and Hikmet Sari

College of Telecommunications and Information Engineering
Nanjing University of Posts and Telecommunications
Nanjing 210003, China

2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall) | 978-1-6654-1368-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/VTC2021-FALL52928.2021.9625525

***Abstract*—Scenario identification plays a crucial role in enhancing the performance of vehicle-to-vehicle (V2V) communication systems. It enables smart vehicles to adjust driving speed in allowable range according to the surrounding circumstance automatically and avoid possible crashes. However, existing methods for scenario identification in vehicular networks have cumbersome processing of information sequence and huge energy consumption. This paper proposes a novel scenario identification method using joint multislice and cooperative detection aided residual network (Resnet), which can extract features from non-equalized signal at the receiver (Rx. signal) automatically and realize scenario recognition. Simulation results demonstrate that the proposed Resnet-based scenario identification method can achieve high classification accuracy with small model size.**

***Index Terms*—Scenario identification, joint multislice and cooperative detection, residual network (Resnet), vehicle-to-vehicle (V2V) communication systems.**

I. INTRODUCTION

Scenario identification is considered as one of the prospective techniques for wireless communication in beyond fifth-generation (B5G) and six-generation (6G) systems [1]–[6]. There are many existing applications of scenario identification in wireless communication such as adaptive modulation and coding (AMC) [7]–[9], radio frequency fingerprint identification (RFID) [10], spectrum sensing [11] and network intrusion detection [12], [13]. In recent years, intelligent transportation systems (ITS) have emerged as a new research hot topic. ITS are designed for offering different kinds of services on the basis of real-time circumstance and enabling users with safe and smart use in V2V systems. Some researches have been done in order to enhance the performance of ITS. For instance, L. Zhao *et al.* [14] proposed a temporal graph convolutional network (TGCN) to implement real-time traffic forecast and improve transport efficiency. L. Cui *et al.* [15] adopted edge learning in minimizing the uploading delay and realized real-time analysis of surveillance videos. M. Liu *et al.* [16] introduced deep-reinforcement learning into autonomous driving which performs well in realistic scenario with satisfactory running time and accuracy.

As mentioned above, scenario identification is a vital technique for the improvement of performance in V2V communication systems. Existing methods can be generally summarized into two categories. One approach is deploying sensors such as radars and cameras on the user end [15] or roadside [17], [18]. However, most of the measured data is stored and transmitted in the form of video which requires complex algorithm and devices with strong computing power to process [19]. In addition, the high latency and huge energy consumption involved during video uploading and analysis is also a problem to be solved [15]. Another approach is utilizing man-made features. In [20], the authors used shadow fading (SF), root mean square (RMS) delay spread, K factor and power delay spectrum (PDP) as features. Moreover, a back-propagation neural network (BPNN) was applied for establishing a mapping relationship between the four channel features and scenarios. However, it requires an extra equipment to extract channel features instantaneously with high computational complexity.

Existing methods for scenario identification have cumbersome processing of the received information. Hence, it is hard to be adopted into real scenarios. In this paper, we introduce a Resnet-based method to extract features automatically from Rx. signal which can be easily obtained by wireless receivers without deploying extra equipment. In addition, inspired by [21], we adopt multislice method to compress the model size and cooperative detection algorithm to compensate for the loss of features. The main contributions of this paper can be summarized as follows.

1) We propose a Resnet-based scenario identification method in V2V communication systems without using complex algorithm and extra equipment, which enables the construction of model easy and feasible.
2) A multislice method is utilized to reduce the input size and compress the model size by 83% at least, which enables the model to be deployed in parallel on devices with limited computing power.
3) A cooperative detection algorithm is applied to remedy the loss of features caused by multislice and improve identification accuracy by 5.78% at the minimum.

978-1-6654-1368-8/21/$31.00 ©2021 IEEE
Authorized licensed use limited to: Harbin Engineering Univ Library. Downloaded on November 26,2025 at 02:21:05 UTC from IEEE Xplore. Restrictions apply.

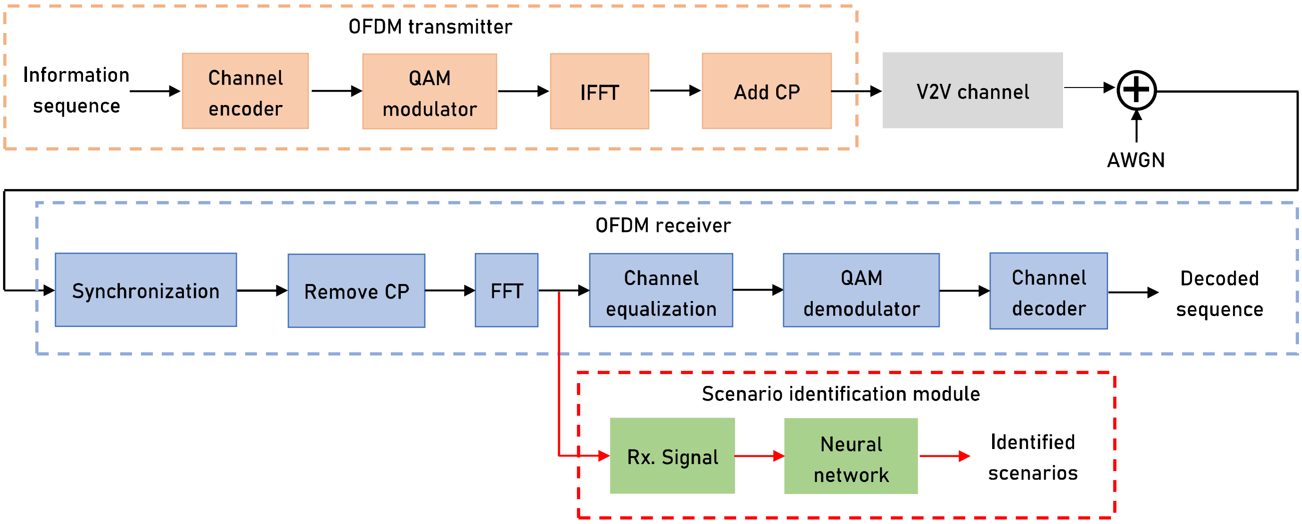

Fig. 1. The construction of V2V communication system implemented by scenario identification module.

II. System Model and Problem Formulation

The construction of IEEE 802.11p V2V communication system implemented by scenario identification module is illustrated in Fig. 1. In orthogonal frequency division multiplexing (OFDM) transmitter, the information sequence is first encoded by a channel encoder and modulated by quadrature amplitude modulation (QAM) to obtain symbols expressed as $X(n) \in \mathbb{C}$, where n denotes the number of data subcarrier index and $n \in [1, N_d]$, where N_d is the total number of data subcarriers and equals to 48. After modulation, the inverse fast Fourier transform (IFFT) is performed and cyclic prefix (CP) is added to get the OFDM symbols. Next, the OFDM symbols are passed over V2V channel which contains five scenarios named rural light-of-sight (R-LOS), urban approaching light-of-sight (UA-LOS), urban non-light-of-sight (U-NLOS), highway light-of-sight (H-LOS) and highway non-light-of-sight (H-NLOS). In OFDM receiver, after synchronization, CP is removed and the scenario identification module is inserted after fast Fourier transform (FFT). The non-equalized Rx signal $Y(n) \in \mathbb{C}$ used for scenario recognition in frequency domain is expressed as

$$Y(n) = H(n)X(n) + w(n), \tag{1}$$

where $H(n) \in \mathbb{C}$ is the channel gain of data subcarrier n and the additive white Gaussian noise (AWGN) $w(n) \in \mathbb{C}$ with zero mean value and σ^2 variance is given as $w(n) \sim \mathcal{CN}(0, \sigma^2)$. Eventually, after channel equalization and QAM demodulation, the symbols are passed over channel decoder to output the decoded sequence.

III. The Proposed Resnet-based Method and Cooperative Detection Algorithm

In this section, we first generate the specific dataset by considering multislice method. Secondly, Resnet-based scenario identification method is proposed and compared with other methods. Finally, cooperative detection algorithm is illustrated in detail.

A. Dataset Generation and Multislice Method

In this paper, the process of data generation and multislice is introduced as follows. Firstly, we utilize MATLAB simulation to generate the Rx. signal as training and testing dataset. Each signal at n-th data subcarrier has the length of L_s and the original size of sample is $N \times L$, where N and L denote the number and length of samples, respectively. Here, $L_s = 70$ and $L = L_s \times N_d = 3360$. Secondly, we decompose the complex-valued data as in-phase (I) and quadrature (Q) part given as

$$I = \{real[Y(1)], real[Y(2)], \cdots, real[Y(N_d)]\}, \tag{2}$$

$$Q = \{imag[Y(1)], imag[Y(2)], \cdots, imag[Y(N_d)]\}. \tag{3}$$

Then the two parts are spliced into a three-dimensional matrix Y with a shape of $N \times L \times 2$. Thirdly, the sample matrix is transformed into multi slices with the same size. The k-th sliced sample matrix is given by $Y_{slice}(k)$, $k \in [1, K]$ with the size of $N \times L_0 \times 2$. Hence, the sample matrix can be divided as

$$Y = \{Y_{slice}(1), Y_{slice}(2), \cdots, Y_{slice}(K)\}, \tag{4}$$

where L_0 and K represent the length and number of sliced sample, respectively.

B. Resnet-based Method

Resnet is a classical deep learning (DL)-based network. The basic unit of Resnet is defined as Res Block (a, (b, c)) shown in Fig. 2, where x is the identity mapping and $F(x)$ is the residual mapping. The original mapping is recast into $F(x) + x$. Batch Normalization (BN) layers are applied for training acceleration and convolutional layers are utilized for

Authorized licensed use limited to: Harbin Engineering Univ Library. Downloaded on November 26,2025 at 02:21:05 UTC from IEEE Xplore. Restrictions apply.

automatically extracting features with the kernel number of a and the kernel size of (b, c). In addition, rectified linear unit (ReLu) and Linear are adopted as activation function. Next, the Res Blocks are piled up to build the Res Stack (Res Block (a, (b, c)), d, e) shown in Fig. 3, where d and e represent the kernel number of the first convolutional layer and the number of Res Block. Finally, the structure of Resnet is proposed in Table. I. The dropout layer is adopted between layers to avoid over-fitting and softmax is used as activation function for the last dense layer. Here, cross entropy function is utilized as the loss function for multi-classification task.

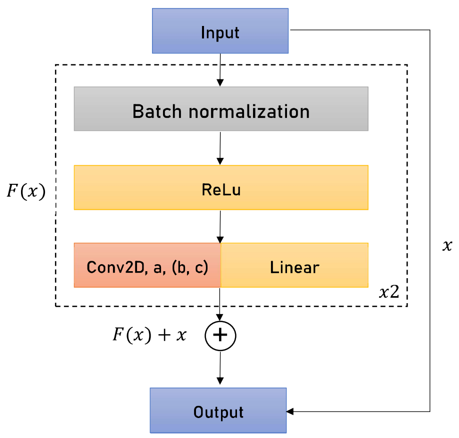

Fig. 2. The structure of Res Block.

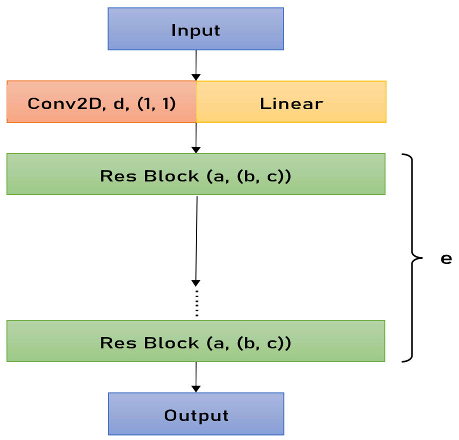

Fig. 3. The structure of Res Stack.

C. Benchmark Methods for Comparison

1) Conventional ML-based Methods: Four conventional ML-based methods are adopted for comparison including random forest (RF), Gaussian naive Bayes (GNB), k-nearest neighbor (KNN) and gradient boosting decision tree (GBDT).

2) DL-based Methods: Here, we consider two types of classical DL-based methods: convolutional neural network (CNN) and deep neural network (DNN) as shown in Table. II and III, respectively.

TABLE I
THE STRUCTURE OF RESNET.

Structure	**Parameters**	**Activation**
Input	$L_0\times 2$	–
Res Stack	Res Block (128, (16, 1), 128, 1)	ReLu & Linear
Dropout	0.3	–
Res Stack	Res Block (64, (8, 1), 64, 2)	ReLu & Linear
Dropout	0.3	–
Res Stack	Res Block (64, (8, 1), 64, 2)	ReLu & Linear
Dropout	0.3	–
Res Stack	Res Block (32, (4, 1), 32, 2)	ReLu & Linear
Dropout	0.3	–
Res Stack	Res Block (32, (4, 1), 32, 2)	ReLu & Linear
Dropout	0.3	–
Flatten	–	–
Dense	128	Linear
Dropout	0.5	–
Dense	64	Linear
Dropout	0.5	–
Dense	5	Softmax
Output	1×5	–

TABLE II
THE STRUCTURE OF CNN.

Structure	**Parameters**	**Activation**
Input	$L_0\times 2$	–
BN	–	–
Conv2D	128, (16, 1)	Linear
Conv2D	64, (8, 1)	Linear
Conv2D	32, (4, 1)	Linear
Flatten	–	–
Dense	128	Linear
Dropout	0.5	–
Dense	64	Linear
Dropout	0.5	–
Dense	5	Softmax
Output	1×5	–

D. Cooperative Detection Algorithm

In this paper, we propose a cooperative detection algorithm to compensate for the loss of features caused by multislice. The output of above methods is a 1×5 vector which can be expressed as

$$P_k = \left[P_k^1, P_k^2, P_k^3, P_k^4, P_k^5\right], \tag{5}$$

where P_k and P_k^j represent the probability vector of k-th slice and the probability vector of k-th slice corresponding to j-th scenario, respectively. Taking the slice length of 480 as example, the process of cooperative detection is shown in Fig. 4. Firstly, a sample is divided into seven slices and fed into model for training. Secondly, after training, each slice outputs a probability vector P_k and then put into cooperative detection module which is implemented based on the maximum possibility (MP) algorithm. The steps of MP algorithm are illustrated as follows. First, since each slice has the same length, the probability vector of k-th slice is added to get a whole probability vector

$$P_{sum} = \sum_{k=1}^{5} P_k = \left[P_{sum}^1, P_{sum}^2, P_{sum}^3, P_{sum}^4, P_{sum}^5\right]. \tag{6}$$

Authorized licensed use limited to: Harbin Engineering Univ Library. Downloaded on November 26,2025 at 02:21:05 UTC from IEEE Xplore. Restrictions apply.

TABLE III
THE STRUCTURE OF DNN.

Structure	Parameters	Activation
Input	$L_0\times 2$	–
BN	–	–
Flatten	–	–
Dense	128	Linear
Dropout	0.5	–
Dense	64	Linear
Dropout	0.5	–
Dense	5	Softmax
Output	1×5	–

TABLE IV
PARAMETER SETTINGS IN EXPERIMENT.

Parameter	Value
Learning rate	0.0001
ReduceLROnPlateau	factor: 0.6, patience: 15
Epochs	500
Batch size	256
Optimizer	Adam
Number of training data	20,000/SNR/scenario
Ratio of training and validation	7:3
Number of testing data	6,000/SNR/scenario
SNR	[15, 40] dB with 5 dB interval

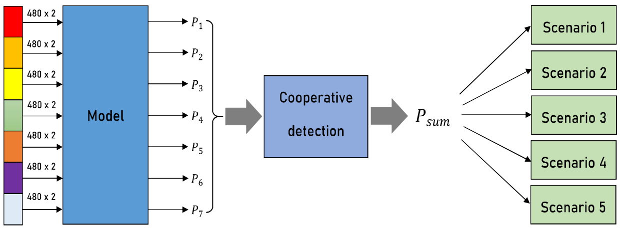

Fig. 4. The process of cooperative detection based on MP algorithm.

Next, since the largest $P_{sum}^j, j \in [1,5]$ represents the sample coming from the j-th scenario, we use the function $\arg\max\ (f(x))$ to output the identified scenario J as

$$J = \arg\max\ (P_{sum}),\ \ J \in [1,5]. \tag{7}$$

To verify the proposed algorithm, the raw model accuracy (RMA) and the cooperative detection aided model accuracy (CDAMA) are introduced to measure the accuracy of the model as

$$RMA = \frac{N_{slices}}{K}, \tag{8}$$

$$CDAMA = \frac{N_{samples}}{N}, \tag{9}$$

where N_{slices} and $N_{samples}$ are the correct prediction slices and correct prediction samples, respectively.

IV. EXPERIMENTAL RESULTS

A. Simulation Platform and Parameters

The simulation experiments are carried out on MATLAB 2020a, Keras 2.3.1 with Tensorflow 2.1 as backend and a GTX 1080Ti. The detailed parameter settings are given in Tab. IV. Noted that we use the ReduceLROnPlateau function built in Keras to realize the automatic adjustment of the learning rate. Thus, the model of higher accuracy for scenario identification can be obtained.

B. Resnet-based Method vs. Others

RMA is the base of cooperative detection and CDAMA cannot achieve at a high value with low RMA. In order to find out the best scenario identification method, we utilize testing set to calculate the RMA of each method under the same slice length of 480. As shown in Fig. 5, the RMA of DL-based methods is higher than ML-based ones. In addition, compared with CNN-based and DNN-based methods, Resnet-based method performs better especially at low SNRs. Simulation results demonstrate that Resnet-based method is the best of all seven methods.

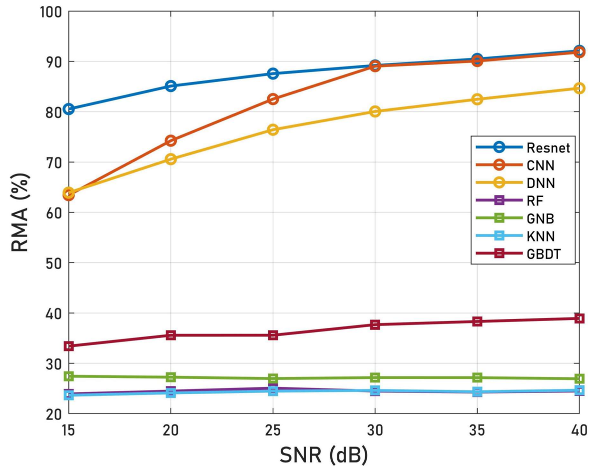

Fig. 5. The RMA of proposed methods under the slice length of 480.

C. Performance Comparison of Methods under Different Slice Lengths

1) Model Size: In this paper, the model size is measured by model parameters and floating-point operations per second (FLOPS). As shown in Tab. V, it is obvious that after multislice method is applied, the model size is smaller than before. Specially, under the slice length of 70, the model parameters and FLOPS are decreased by 95% and 98% compared with the Resnet-based method without multislice.

2) Model Accuracy: After the best Resnet-based method has been selected, its RMA and CDAMA are calculated under different slice lengths and SNRs shown in Tab. VI and Fig. 6, where the AveRMA and AveCDAMA represent the average value of RMA and CDAMA under the SNRs given in Tab. IV. It can be observed that the RMA and CDAMA decrease with the drop of slice length attributed to fewer features in each sample slice. In addition, our proposed cooperative detection

Authorized licensed use limited to: Harbin Engineering Univ Library. Downloaded on November 26,2025 at 02:21:05 UTC from IEEE Xplore. Restrictions apply.

TABLE V
THE MODEL SIZE OF DNN-BASED, CNN-BASED AND RESNET-BASED METHODS.

Model	Parameters	FLOPS
DNN	9.0×10^5	1.8×10^6
CNN	2.7×10^7	1.1×10^9
Resnet	2.8×10^7	1.1×10^{10}
Resnet ($L_0 = 480$)	4.7×10^6 (83%↓)	1.6×10^9 (85%↓)
Resnet ($L_0 = 240$)	2.8×10^6 (90%↓)	8.0×10^8 (93%↓)
Resnet ($L_0 = 70$)	1.4×10^6 (95%↓)	2.3×10^8 (98%↓)

method can improve identification accuracy especially under small slice length.

TABLE VI
THE AVERAGE MODEL ACCURACY OF RESNET-BASED METHOD UNDER DIFFERENT SLICE LENGTHS.

Model	AveRMA (%)	AveCDAMA (%)
Resnet ($L_0 = 480$)	87.49	93.27 (5.78%↑)
Resnet ($L_0 = 240$)	86.35	92.30 (5.95%↑)
Resnet ($L_0 = 70$)	61.77	78.66 (16.89%↑)

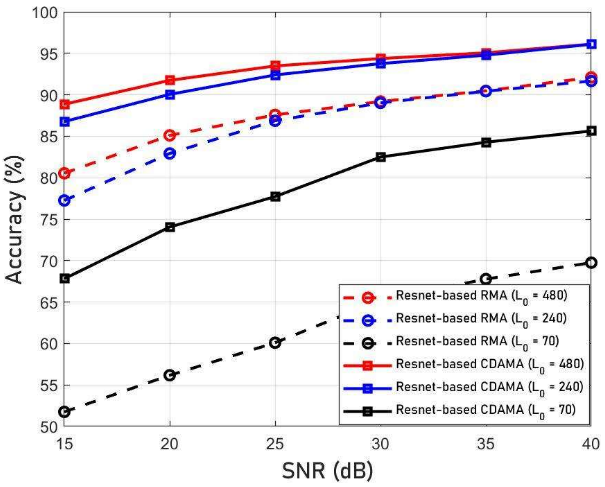

Fig. 6. The model accuracy of Resnet-based method under different slice lengths and SNRs.

V. CONCLUSION

In this paper, we propose a multislice method to lower the model size and a cooperative detection algorithm based on MP to enhance the identification accuracy. In addition, we adopt both of them in Resnet to realize a Resnet-based scenario identification method. Simulation results demonstrate that our method performs well in scenario identification with small model size. Under the slice length of 480, the FLOPS is decreased by 85% and the average identification accuracy can be achieved at 93.27%.

REFERENCES

[1] X. You, *et al.*, "Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts," *Science China Information Sciences*, vol. 64, no. 1, article number: 110301, Nov. 2021.
[2] S. Han, T. Xie, and C.-L. I, "Greener physical layer technologies for 6G mobile communications," *IEEE Communications Magazine*, vol. 59, no. 4, pp. 68–74, Apr. 2021.
[3] M. Matthaiou, *et al.*, "The road to 6G: Ten physical layer challenges for communications engineers," *IEEE Communications Magazine*, vol. 59, no. 1, pp. 64–69, Jan. 2021.
[4] G. Gui, M. Liu, F. Tang, N. Kato, and F. Adachi, "6G: Opening new horizons for integration of comfort, security, and intelligence," *IEEE Wireless Communications*, vol. 27, no. 5, pp. 126–132, Oct. 2020.
[5] J. Zhao, J. Liu, L. Yang, B. Ai and S. Ni, "Future 5G-oriented system for urban rail transit: Opportunities and challenges," *China Communications*, vol. 18, no. 2, pp. 1–12, Feb. 2021.
[6] J. Zhao, P. Dong, X. Ma, X. Sun and D. Zou, "Mobile-aware and relay-assisted partial offloading scheme based on parked vehicles in B5G vehicular networks," *Physical Communication*, vol. 42, pp. 101163, Jul. 2020.
[7] Y. Lin, Y. Tu, Z. Dou, L. Chen, and S. Mao, "Contour stella image and deep learning for signal recognition in the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 1, pp. 34–46, Mar. 2021.
[8] Y. Wang, G. Gui, T. Ohtsuki, and F. Adachi, "Multi-task learning for generalized automatic modulation classification under non-Gaussian noise with varying SNR conditions," *IEEE Transactions on Wireless Commuications*, early access, doi: 10.1109/TWC.2021.3052222
[9] Y. Lin, Y. Tu, and Z. Dou, "An improved neural network pruning technology for automatic modulation classification in edge device," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5703–5706, May 2020.
[10] D. Roy, T. Mukherjee, M. Chatterjee, and E. Pasiliao, "RF transmitter fingerprinting exploiting spatio-temporal properties in raw signal data," in *18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, Boca Raton, FL, USA, Dec. 16–19, 2019, pp. 89–96.
[11] S. Zheng, S. Chen, P. Qi, H. Zhou, and X. Yang, "Spectrum sensing based on deep learning classification for cognitive radios," *China Communications*, vol. 17, no. 2, pp. 138–148, Feb. 2020.
[12] R. Zhao, *et al.*, "An efficient intrusion detection method based on dynamic autoencoder," *IEEE Wireless Communication Letters*, early access, doi: 10.1109/LWC.2021.3077946.
[13] Z. He, *et al.*, "Edge device identification based on federated learning and network traffic feature engineering," *IEEE Trans. Cogn. Commun. Netw.*, early access.
[14] L. Zhao, *et al.*, "T-GCN: A temporal graph convolutional network for traffic prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3848–3858, Sept. 2020.
[15] L. Cui, D. Su, Y. Zhou, L. Zhang, Y. Wu, and S. Chen, "Edge learning for surveillance video uploading sharing in public transport systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2274–2285, Apr. 2021.
[16] M. Liu, F. Zhao, J. Niu, and Y. Liu, "Reinforcement driving: exploring trajectories and navigation for autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 808–820, Feb. 2021.
[17] Q. Wang, J. Zheng, H. Xu, B. Xu, and R. Chen, "Roadside magnetic sensor system for vehicle detection in urban environments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1365–1374, May. 2018.
[18] L. Wang, Z. Zhang, X. Di, and J. Tian, "A Roadside camera-radar sensing fusion system for intelligent transportation," in *17th European Radar Conference (EuRAD)*, Utrecht, Netherlands, Jan. 10–15, 2021, pp. 282–285.
[19] S. Wan, X. Xu, T. Wang, and Z. Gu, "An intelligent video analysis method for abnormal event detection in intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, early access, doi: 10.1109/TITS.2020.3017505.
[20] M. Yang, *et al.*, "Machine-learning-based scenario identification using channel characteristics in intelligent vehicular communications," *IEEE Transactions on Intelligent Transportation Systems*, early access, doi: 10.1109/TITS.2020.3001132.
[21] D. Jiao, Y. Peng, Y. Wang, and J. Yang, "Joint multislice and cooperative detection aided RFID method based on deep learning," *Physical Communication*, vol. 42, no. 101153, pp. 1–9, Oct. 2020.

Authorized licensed use limited to: Harbin Engineering Univ Library. Downloaded on November 26,2025 at 02:21:05 UTC from IEEE Xplore. Restrictions apply.