Received July 6, 2019, accepted July 28, 2019, date of publication August 6, 2019, date of current version August 21, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2933448

Deep Learning Aided Method for Automatic Modulation Recognition

CHENG YANG[1], ZHIMIN HE[2], (Student Member, IEEE), YANG PENG[2], YU WANG[2], (Student Member, IEEE), AND JIE YANG[2], (Member, IEEE)

[1]Changzhou College of Information Technology, Changzhou 213164, China
[2]College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Corresponding authors: Zhimin He (b15010715@njupt.edu.cn) and Jie Yang (jyang@njupt.edu.cn)

ABSTRACT Automatic modulation recognition (AMR) is considered one of most important techniques in the non-cooperative wireless communication systems. Traditional algorithms, e.g., support vector machine (SVM) based on high order cumulants (HOC), are hard to achieve the reliable performance. In this paper, we propose an effective AMR algorithm based on deep learning (DL) with capabilities of automatically extracting representative and effective features. Our proposed method resorts to in-phase and quadrature (IQ) samples which are IQ components of received baseband signal, respectively. We adopt convolutional neural networks (CNN) and recurrent neural networks (RNN) to classify six types of signal modulations over additive white Gaussian noise (AWGN) channel and Rayleigh fading channel, respectively. Simulation results show that DL-AMR is much better than traditional algorithms under two fading channels.

INDEX TERMS Automatic modulation recognition (AMR), deep learning (DL), convolutional neural networks (CNN), recurrent neural networks (RNN).

I. INTRODUCTION

Automatic modulation recognition (AMR) refers to identify the types of modulation signal efficiently and accurately under the condition of insufficient or lack of prior information [1]–[4]. When analyzing the received unknown signal, it is necessary to recognize the modulation type of target signal in various applications [5]–[10]. Hence, AMR technique is extremely important in the complicate wireless communications environments. The research methods on the AMR are mainly divided into two categories: One is based on maximum likelihood (ML) theory [11] and the other is feature-based (FB) method.

The main idea of the ML method is Bayesian theory, which needs to acquire detailed channel information as much as possible to design likelihood function. It compares the statistical characteristics of the received signals with the best decision thresholds obtained by the calculation, and then the corresponding conclusions are reached. This method is to make approximate judgment rather than optimal judgment. Its essence is the problem formulation of multiple hypothesis testing. Whereas FB method requires no priori information, it can extract features from the received signals and designs a proper classifier for modulation recognition.

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

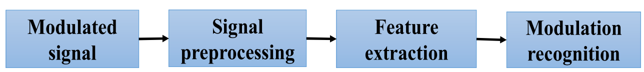

FIGURE 1. The process of FB method.

Although the ML method is relatively simple and convenient, it demands sufficient and accurate prior information. Furthermore, the whole calculation process is rather complicated. Consequently, the current AMR mainly adopts the FB method. The FB method includes three steps of signal preprocessing, feature extraction, and modulation recognition using the extracted features. The structure of FB method is shown in Fig. 1.

The main purpose of signal preprocessing is signal denoising, which is helpful for feature extraction in the next step. Feature extraction and design of classifier have great impact on the performance of FB method. In fact, there are many kinds of features for AMR, including instantaneous features [1], wavelet transform [12], cyclic spectrum [13], high order cumulants (HOC) [14], and so on. Besides, there are also many types of AMR classifiers, such as decision tree [14], support vector machine (SVM) [15].

In recent, many researchers have focused on deep learning (DL) applications in physical layer [16]–[20], especially in AMR, due to its outstanding performances in various tasks. O'Shea *et. al* proposed an adoption of convolutional neural networks (CNN) to directly extract the features

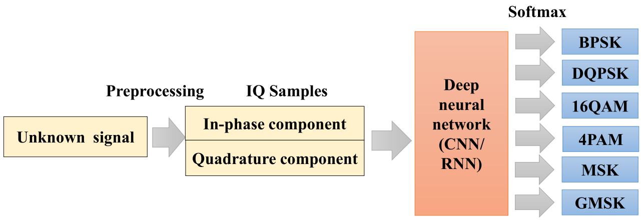

FIGURE 2. The structure of system model.

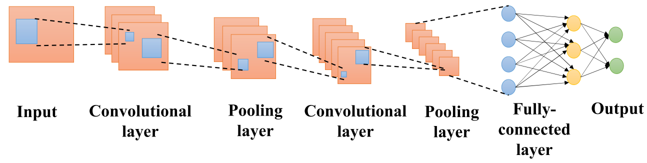

FIGURE 3. The structure of CNN.

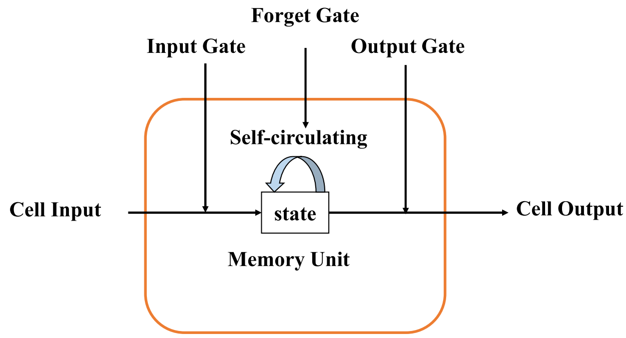

FIGURE 4. The structure of LSTM memory unit.

from a modulated signal without designing manmade features [21]. Lin *et al*. applied several well-known DL network models to recognize modulation signal, and made significant progress in improving modulation signal classification accuracy [22].

In this paper, we propose a DL-based FB method to classify different types of modulation signals. We use the data generated by Matlab simulation as the input for training in CNN and recurrent neural networks (RNN) which is also adopted because of its excellent performance in the serialization problem. In addition, a traditional method with HOC features and SVM classifier is adopted as the benchmark. Our main contributions are summarized as follows.

First, IQ samples are fed into neural network directly, and no additional feature extraction is required. Second, we design a novel and effective DL-based method, which can automatically extract features from original data and identify the type of modulation signal. At last, simulation results show that the performance of DL algorithm is much better than traditional algorithm. Particularly, the classification accuracy of CNN is nearly 100% at 0 dB in AWGN channel.

This paper is organized as follows. Section II shows system model and DL networks, respectively. Section III proposes DL-based AMR method using CNN and RNN, respectively. In Section IV, we conduct the experiment to evaluate the proposed methods. Section V gives a conclusion.

II. SYSTEM MODEL AND DL NETWORKS

A. SYSTEM MODEL

In this work, we decide to classify six modulation signals that are often used in digital communications. They are binary frequency shift keying (BFSK), differential quadrature phase shift keying (DQPSK), 16 quadrature amplitude modulation (16QAM), quaternary pulse amplitude modulation (4PAM), minimum shift keying (MSK), Gaussian minimum shift keying (GMSK) respectively. And the system model is shown in Fig. 2.

Firstly, the preprocessed unknown signal is converted into IQ samples, and then use them as input data to train the deep neural network. After neural network training, the correct classification probability of the corresponding modulation method can be obtained.

B. CNN

CNN, which consists input layer, convolutional layer, pooling layer, fully-connected layer, output layer and activation function, has shown a wide range of applications for its excellent performance. Its basic structure is shown in Fig. 3.

- The convolutional layer is used for feature extraction, and it obtains local information by acting on a local region through a convolution kernel.
- The pooling layer is designed to compress the input feature map, but not necessary. It makes the feature map smaller and simplifies the computational complexity of the network, furthermore, it makes the main features extraction easier.
- Fully-connected layer combines all the features of the input data and sends the output value to the classifier.
- The activation function adds nonlinear factors to the neural network model, so the neural networks can solve nonlinear problems.

C. RNN

RNN is an important branch of deep learning, which deals with serialization problems. Unlike feedforward neural networks (e.g., artificial neural networks (ANN), CNN), RNN pays more attention to feedback, which processes the sequence data by using their internal state (memory). In this paper, as a variant of RNN, long-short term memory (LSTM) neural network is adopted to classify the modulation methods. Its basic unit structure is shown in Fig. 4.

- The input gate determines new input information into the memory unit.
- The forget gate is used to control whether the memory unit should remember or forget the previous state.
- The output gate decides which information in memory unit is allowed to be output.

III. OUR PROPOSED AMR ALGORITHM

A. DATASET

When the unknown signals arrive at the receiver, it is preprocessed to obtain the complex-valued baseband signal.

TABLE 1. The network layers structure of CNN.

Layer	Activation function	Output Dimensions
Input	/	2×256×1
Convolution (128 fiters,1×8)	PReLU	2×249×128
Dropout (0.6)	/	/
Convolution (64 filters, 1×4)	PReLU	2×246×64
Dropout (0.6)	/	/
Flatten	/	31488
Dense	PReLU	256
Dropout (0.6)	/	/
Dense	PReLU	128
Dropout (0.6)	/	/
Dense	Softmax	6

Then the baseband signal is simultaneously sampled by the IQ components to obtain the IQ samples which are required by the training network. Finally, they are combined into a $2 \times N$ matrix, where N is the number of sampling points. The value of N is 256. In this work, we only consider AWGN channel and Rayleigh channel, respectively. The IQ samples of each modulation type is 2000. So there are 12000 training data in total as the input of the neural network, which are divided into training sets and test sets in the ratio of 7:3.

B. CNN FOR AMR

The network layers structure of CNN is shown in Table 1. It has two convolution layers and three fully-connected layers. The first convolutional layer has 128 convolution kernels with size of 1×8, and 1×4 convolution kernels is employed in the second convolutional layer. After two convolution layers, flatten layer and three full-connected layers are applied to integrate and normalize highly abstract features, which is beneficial for classification of the classifier.

There is a dropout layer behind each convolution layer and fully-connected layer, except the last layer. The purpose of that is to prevent overfitting and improve the generalization ability of the network. The parameter of dropout is set as 0.6 when the best classification performance is achieved in our experiment.

Parametric rectified linear unit (PReLU) is used as activation function in almost the entire network structure, in addition to the last layer. PReLU is an improvement of rectified linear unit (ReLU), which can learn parameters from data adaptively. And it has distinct advantages of fast convergence speed and low error rate. The activation function of the last layer is Softmax, which is mostly applied for multi-classification problems. It takes the maximum output probability of all categories as the final result.

C. RNN FOR AMR

The specific network layer structure of LSTM is shown in Table 2, which makes up of three LSTM layers and three fully connected layers. The size of LSTM layer are set as 256, 128 and 64, respectively. Similar to the CNN, the activation function of the whole network is ReLU except the last layer. ReLU converges faster in the stochastic gradient descent (SGD) algorithm and can alleviate the gradient diffusion to some extent. Likewise, softmax is adopted as the activation function to produce the probability score about six modulation categories in the last layer. And dropout is used to reduce the risk of overfitting. The loss function applied in the network training process is the cross-entropy.

D. SVM FOR AMR

A traditional algorithm based on HOC features is also applied as a comparison. The HOC features are extracted from the sample data used in the DL algorithm. Assuming $\{X(t)\}$ is a stationary stochastic process, and its p order mixed moments is defined as

$$M_{pq} = E\left[X(t)^{p-q} X^*(t)^q\right] \tag{1}$$

where $X^*(t)$ is the conjugate of $X(t)$, p denotes the value of order. For a complex stochastic process with zero mean, the cumulants of various orders [23] are defined as below.

- Second-order cumulants:

$$C_{20} = Cum(X, X) = M_{20} \tag{2}$$
$$C_{21} = Cum\left(X, X^*\right) = M_{21} \tag{3}$$

- Fourth-order cumulants:

$$C_{40} = Cum(X, X, X, X) = M_{40} - 3M_{20}^2 \tag{4}$$
$$C_{41} = Cum\left(X, X, X, X^*\right) = M_{41} - 3M_{20}M_{21} \tag{5}$$
$$C_{42} = Cum\left(X, X, X^*, X^*\right) = M_{42} - |M_{20}|^2 - 2M_{21}^2 \tag{6}$$

- Sixth-order cumulants:

$$\begin{aligned} C_{60} &= Cum(X, X, X, X, X, X) \\ &= M_{60} - 15M_{40}M_{20} + 30M_{20}^3 \end{aligned} \tag{7}$$
$$\begin{aligned} C_{63} &= Cum\left(X, X, X, X^*, X^*, X^*\right) \\ &= M_{63} - 6M_{41}M_{20} - 9M_{42}M_{21} \\ &\quad + 18M_{20}^2M_{21} + 12M_{21}^3 \end{aligned} \tag{8}$$

Note that C_{pq} represents p-th order cumulants with q conjugate variables. In this paper, we eventually used the following

TABLE 2. The network layers structure of RNN.

Layer	Activation function	Output Dimensions
Input	/	2×256
LSTM	ReLU	2×256
Dropout (0.35)	/	/
LSTM		2×128
Dropout (0.35)	/	/
LSTM	ReLU	2×64
Dropout (0.35)		/
Dense	ReLU	256
Dropout (0.35)	/	/
Dense	ReLU	128
Dropout (0.35)	/	/
Dense	ReLU	64
Dropout (0.35)	/	/
Dense	Softmax	6

seven HOC features for classification,

$$\begin{aligned} f_1 &= |C_{40}| \,/\, |C_{42}| \,, \\ f_2 &= |C_{41}| \,/\, |C_{42}| \,, \\ f_3 &= |C_{42}| \,/\, |C_{21}|^2 \,, \\ f_4 &= |C_{60}| \,/\, |C_{21}|^3 \,, \\ f_5 &= |C_{63}| \,/\, |C_{21}|^3 \,, \\ f_6 &= |C_{60}|^2 \,/\, |C_{42}|^3 \,, \\ f_7 &= |C_{63}|^2 \,/\, |C_{42}|^3 \end{aligned} \tag{9}$$

After obtained the HOC features set, it can be used as the input of SVM for classification. SVM is a classifier with excellent performance in traditional algorithms, which distinguishes different categories by finding a hyperplane in the feature space. For nonlinear classification problems, SVM can map input data into high-dimensional feature space by using what is called the kernel trick. The kernel trick applied in this paper is Gaussian kernel.

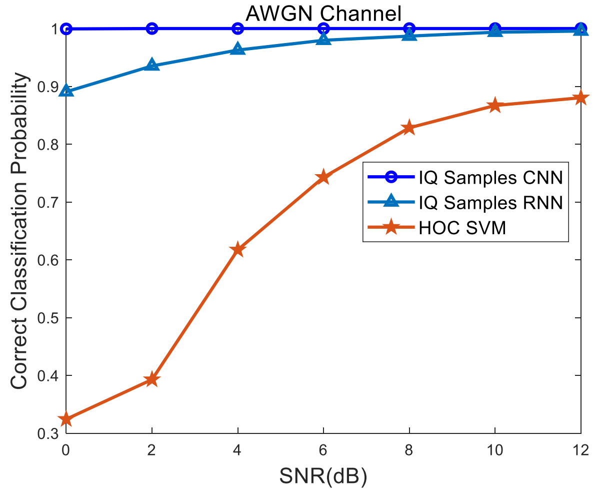

FIGURE 5. The performance of deep learning algorithm and traditional algorithm under AWGN channel.

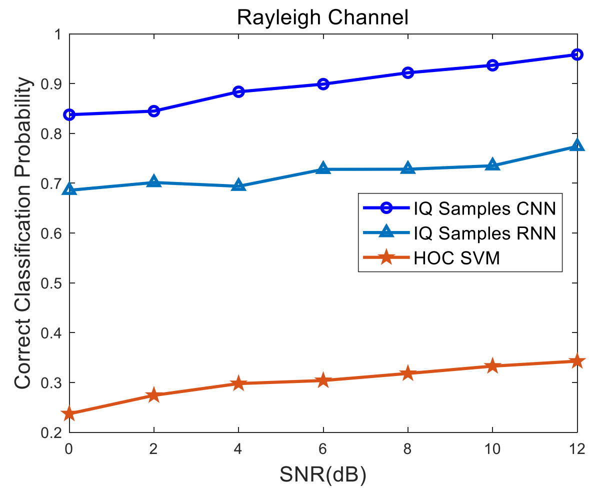

FIGURE 6. The performance of deep learning algorithm and traditional algorithm under Rayleigh fading channel.

E. IMPLEMENTTATION PLATFORM

The platform used for DL in this experiment is Keras library with tensorflow as the backend. Matlab 2018a is applied to generate and preprocess the data used in this work. NVIDIA GTX1080Ti is also employed to accelerate the computation speed.

IV. EXPERIMENT RESULTS

A. PERFORMANCE COMPARISIONS

In this paper, we train the CNN and RNN with IQ samples. Meanwhile, experiment based on traditional algorithm is conducted to compare with the DL algorithm. The final experiment results are as follows.

The DL algorithm does not need complex feature extraction process and saves a lot of labor cost compared with the traditional algorithm. Moreover, the DL algorithm is more robust and can maintain excellent performance even in the

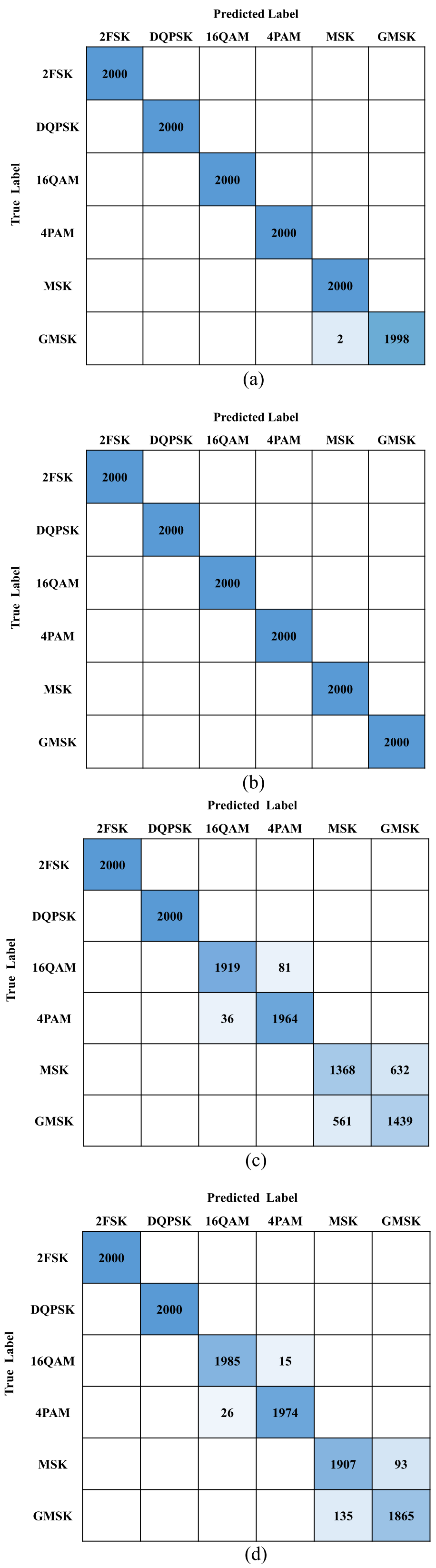

FIGURE 7. Confusion matrices of CNN in different channel: (a) SNR = 0 dB under AWGN channel, (b) SNR = 12 dB under AWGN channel, (c) SNR = 0 dB under Rayleigh channel, (d) SNR = 12 dB under Rayleigh channel.

environment with more serious signal attenuation, such as Rayleigh channel.

Figs. 5∼6 show that the proposed algorithm is increasing as the higher SNR, and vice versa. In addition, the proposed DL algorithms (e.g., CNN) are much better than traditional algorithm under either AWGN or Rayleigh fading channel. Obviously, compared with the other two algorithms, CNN is more robust and superior under AWGN channel and Rayleigh channel.

B. SUPERIOR PERFORMANCES OF CNN

Because of its huge advantages in independence from prior knowledge and feature extraction, CNN has outstanding performance. Furthermore, CNN can reduce the computations by sharing the parameters of convolution kernels, and it has no pressure to deal with the huge data of CNN network. By tuning the parameters of CNN network reasonably, we can obtain good performance in both AWGN channel and Rayleigh channel.

As we can see from the Fig. 5, the correct classification probability of CNN is always close to 100% in AWGN channel. Even in Rayleigh channel, the minimum classification accuracy still approach to 84%, whereas the maximum is near to 96%. The details are shown in Fig. 7 with the form of confusion matrices.

V. CONCLUSION

In this paper, a novel AMR algorithm based on DL is proposed for identify different categories of modulated signals rapidly and precisely. The proposed DL based AMR algorithm is much better than traditional method over AWGN channel and Rayleigh channel, respectively. In addition, DL based method avoids feature extraction for saving a lot of computational time.

REFERENCES

[1] A. K. Nandi and E. E. Azzouz, "Algorithms for automatic modulation recognition of communication signals," *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 431–436, Apr. 1998.

[2] J. L. Xu, W. Su, and M. Zhou, "Software-defined radio equipped with rapid modulation recognition," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1659–1667, May 2010.

[3] M. Lichtman, W. C. Headley, and J. H. Reed, "Automatic modulation classification under IQ imbalance using supervised learning," in *Proc. IEEE Military Commun. Conf. (MILCOM)*, Nov. 2013, pp. 1622–1627.

[4] F. Wen, Z. Zhang, and X. Zhang, "CRBs for direction-of-departure and direction-of-arrival estimation in collocated MIMO radar in the presence of unknown spatially coloured noise," *IET Radar, Sonar Navig.*, vol. 13, no. 4, pp. 530–537, Apr. 2019.

[5] X. Fu, R. Cao, and F. Wen, "A de-noising 2-D-DOA estimation method for uniform rectangle array," *IEEE Commun. Lett.*, vol. 22, no. 9, pp. 1854–1857, Sep. 2018.

[6] F. Wen, "Computationally efficient DOA estimation algorithm for MIMO radar with imperfect waveforms," *IEEE Commun. Lett.*, vol. 23, no. 6, pp. 1037–1040, Jun. 2019.

[7] F. Wen, Z. Zhang, K. Wang, G. Sheng, and G. Zhang, "Angle estimation and mutual coupling self-calibration for ULA-based bistatic MIMO radar," *Signal Process.*, vol. 144, no. 3, pp. 61–67, Mar. 2018.

[8] F. Wen, X. Xiong, J. Su, and Z. Zhang, "Angle estimation for bistatic MIMO radar in the presence of spatial colored noise," *Signal Process.*, vol. 134, no. 5, pp. 261–267, May 2017.

[9] F. Wen, X. Xiong, and Z. Zhang, "Angle and mutual coupling estimation in bistatic MIMO radar based on PARAFAC decomposition," *Digit. Signal Process.*, vol. 65, no. 3, pp. 1–10, Jun. 2017.
[10] G. Gui, H. Sari, and E. Biglieri, "A new definition of fairness for non-orthogonal multiple access," *IEEE Commun. Lett.*, vol. 23, no. 7, pp. 1267–1271, Jul. 2019.
[11] W. Wei and J. M. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *IEEE Trans. Commun.*, vol. 48, no. 2, pp. 189–193, Feb. 2000.
[12] K. C. Ho, W. Prokopiw, and Y. T. Chan, "Modulation identification of digital signals by the wavelet transform," *IEE Proc.-Radar, Sonar Navigat.*, vol. 147, no. 4, pp. 169–176, Aug. 2000.
[13] X. Yan, G. Feng, H. C. Wu, W. Xiang, and Q. Wang, "Innovative robust modulation classification using graph-based cyclic-spectrum analysis," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 16–19, Jan. 2017.
[14] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 416–429, Mar. 2000.
[15] C.-S. Park, J.-H. Choi, S.-P. Nah, W. Jang, and D.-Y. Kim, "Automatic modulation recognition of digital signals using wavelet features and SVM," in *Proc. Int. Conf. Adv. Commun. Technol. (ICACT)*, Feb. 2008, pp. 387–390.
[16] G. Gui, H. Huang, Y. Song, and H. Sari, "Deep learning for an effective nonorthogonal multiple access scheme," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8440–8450, Sep. 2018.
[17] M. Liu, J. Yang, T. Song, J. Hu, and G. Gui, "Deep learning-inspired message passing algorithm for efficient resource allocation in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 641–653, Jan. 2018.
[18] M. Liu, T. Song, G. Gui, J. Hu, and H. Sari, "Deep cognitive perspective: Resource allocation for NOMA-based heterogeneous IoT with imperfect SIC," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2885–2894, Apr. 2019.
[19] C.-K. Wen, W.-T. Shih, and S. Jin, "Deep learning for massive MIMO CSI feedback," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 748–751, Oct. 2018.
[20] H. Huang, Y. Song, J. Yang, G. Gui, and F. Adachi, "Deep-learning-based millimeter-wave massive MIMO for hybrid precoding," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3027–3032, Mar. 2019.
[21] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cognit. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.
[22] Y. Lin, Y. Tu, Z. Dou, and Z. Wu, "The application of deep learning in communication signal modulation recognition," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Oct. 2017, pp. 1–5.
[23] Z. Yin, B. Ma, S. Zhou, Z. Yang, and Z. Wu, "Robust automatic modulation classification under varying noise conditions," *IEEE Access*, vol. 5, pp. 19733–19741, 2017.

CHENG YANG received the master's degree in computer science from Jiangsu University, in 2017. He is currently a Faculty Member with the Changzhou College of Information Technology and also a Visiting Scholar with the FocusLab, Nanjing University of Posts and Telecommunications (NJUPT), Nanjing, China. His research interest includes deep learning for physical layer wireless communications.

ZHIMIN HE (S'19) received the B.S. degree in communication engineering from the Nanjing University of Posts and Telecommunications (NJUPT), Nanjing, China, in 2019, where he is currently pursuing the master's degree in signal processing for wireless communications with the College of Telecommunications and Information Engineering. His research interest includes deep learning and its applications in wireless communications.

YANG PENG received the B.S. degree in communication engineering from Nanchang Hangkong University, Nanchang, China, in 2019. He is currently pursuing the master's degree in signal processing for wireless communications with the College of Telecommunications and Information Engineering, NJUPT. His research interest includes deep learning and its applications in wireless communications.

YU WANG (S'18) received the B.S. degree in communication engineering from the Nanjing University of Posts and Telecommunications (NUPT), Nanjing, China, in 2018, where he is currently pursuing the Ph.D. degree. His research interests include deep learning, optimization, and its application in wireless communications.

JIE YANG (M'18) received the B.Sc., M.Sc., and Ph.D. degrees in communication engineering from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2003, 2006, and 2018, respectively, where she is currently an Assistant Professor.