

Received November 19, 2019, accepted December 3, 2019, date of publication December 19, 2019, date of current version January 3, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2960775

Data Augmentation for Deep Learning-Based Radio Modulation Classification

LIANG HUANG¹, (Member, IEEE), WEIJIAN PAN¹, YOU ZHANG¹,
LIPING QIAN^{1,5}, (Senior Member, IEEE), NAN GAO²,
AND YUAN WU^{3,4}, (Senior Member, IEEE)

¹College of Information Engineering, Zhejiang University of Technology, Hangzhou 310058, China

²College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310058, China

³State Key Laboratory of Internet of Things for Smart City, University of Macau, Macau 999078, China

⁴Department of Computer and Information Science, University of Macau, Macau 999078, China

⁵National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China

Corresponding author: Nan Gao (gaonan@zjut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Project 61572440 and Project 61702456, in part by the Zhejiang Provincial Natural Science Foundation of China under Project LR16F010003 and Project LY19F020033, in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University under Project 2019D11, and in part by the Research Grant of University of Macau under Project SRG2019-00168-IOTSC.

ABSTRACT Deep learning has recently been applied to automatically classify the modulation categories of received radio signals without manual experience. However, training deep learning models requires massive volume of data. An insufficient training data will cause serious overfitting problem and degrade the classification accuracy. To cope with small dataset, data augmentation has been widely used in image processing to expand the dataset and improve the robustness of deep learning models. However, in wireless communication areas, the effect of different data augmentation methods on radio modulation classification has not been studied yet. In this paper, we evaluate different data augmentation methods via a state-of-the-art deep learning-based modulation classifier. Based on the characteristics of modulated signals, three augmentation methods are considered, i.e., rotation, flip, and Gaussian noise, which can be applied in both training phase and inference phase of the deep learning-based classifier. Numerical results show that all three augmentation methods can improve the classification accuracy. Among which, the rotation augmentation method outperforms the flip method, both of which achieve higher classification accuracy than the Gaussian noise method. Given only 12.5% of training dataset, a joint rotation and flip augmentation policy can achieve even higher classification accuracy than the baseline with initial 100% training dataset without augmentation. Furthermore, with data augmentation, radio modulation categories can be successfully classified using shorter radio samples, leading to a simplified deep learning model and a shorter classification response time.

INDEX TERMS Data augmentation, deep learning, modulation classification, wireless communication.

I. INTRODUCTION

Benefiting from the improvement of computing power and big data, deep learning has achieved unprecedented development in many applications, i.e., speech and audio processing [1], natural language processing [2], object detection [3], and so on. In recent years, it also achieves dramatic development in the field of wireless communications, e.g., modulation classification [4], symbol detection [5], end-to-end communication [6], and mobile edge computing [7].

The associate editor coordinating the review of this manuscript and approving it for publication was Guan Gui¹.

Deep learning-based modulation classification automatically and efficiently classify received signals without prior knowledge. Modulation classification is a fundamental step for many applications in wireless communication systems, such as spectrum management in cognitive communication systems [8] and unauthorized signal detection in secure communications [9], [10]. Traditional modulation classification method either requires high computational complexity or greatly depends on manual operations [9]. Recently, deep learning is successfully introduced to classify signals [11]–[16], which feeds raw signal data or its

transforms into a deep neural network and instantly obtains the modulation category at the network output. It achieves higher classification accuracy than traditional methods for automatic modulation classification based on expert features such as higher order cumulants based features [17], while requiring a little extra computational overhead computation time.

Although deep learning-based approaches can greatly improve the performance of the modulation classifier, it requires a large volume of training radio samples. However, in practice, collecting a large amount of high quality and reliable training radio samples sometimes is costly and difficult. Data augmentation has been widely used to deal with lack of training data by artificially expanding the training dataset with label preserving transformation. Different data augmentation methods have been proposed in the literature, i.e., random cropping, rotation and mirroring in image classification [18], [19] and pitch shifting, time stretching and random frequency filtering in speech recognition [20]. For deep learning-based radio modulation classification, data augmentation can improve its invariant, especially for small radio signal dataset.

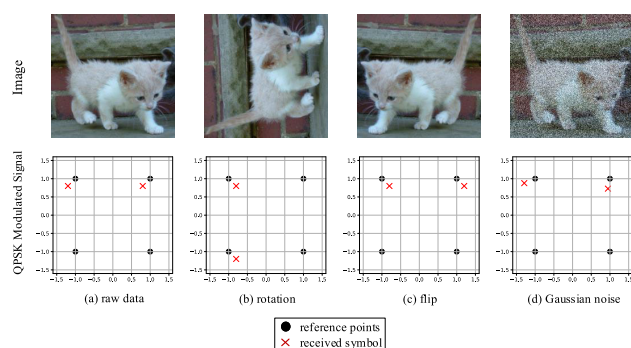


FIGURE 1. Different data augmentation methods for both image and modulated signal: (a) raw data, (b) rotation, (c) flip, (d) Gaussian noise.

Augmenting modulated radio signal is similar to augment images as shown in Fig. 1. Specifically, we consider three basic augmentation methods, i.e. rotation, flip, and Gaussian noise, for both an image and a quadrature phase-shift keying (QPSK) modulated radio signal sample illustrated in constellation diagram. For the image, after rotation or flip augmentation, the same cat is displayed but from different viewpoints. In the constellation diagram of the QPSK modulated radio signal, the black circles indicate four ideal reference points, and the red crosses are the received symbols which are shifted due to the imperfection of transmitter/receiver hardware and wireless channel [21]. In Fig. 1, we consider two received symbols with positive phase shift $(1, 1)$ and $(-1, 1)$, which are counter-clockwise shifted from their reference points. In wireless communication, each received symbol will be demodulated and mapped to one of the reference points based on the transmitted content. After rotation augmentation, two new symbols $(-1, 1)$ and

$(-1, -1)$ are generated as shown Fig. 1(b), which are also positively phase shifted. Therefore, for the radio modulation classification task considered in this paper, rotating the modulated radio signal is similar to rotating an image, without losing features for classification. However, flipping the radio signal generates two new QPSK modulated symbols whose phases are negatively shifted in the clockwise direction, as shown in Fig. 1(c). Although both rotation and flip augmentation methods achieve similar accuracy improvements for image classification [22], [23], it is an open question about which one is preferred for radio modulation classification. After the Gaussian noise augmentation, the image is full of 'snow' and the received radio symbols are deviated as shown in Fig. 1(d). Can all these three augmentation methods improve the classification accuracy for deep learning-based radio modulation classification? To the best of our knowledge, the effect of different data augmentation methods on radio modulation classification has not been evaluated yet.

In this paper, we study data augmentation methods for deep learning-based radio modulation classification. Specifically, a state-of-the-art deep learning-based modulation classifier, is used to automatically classify the modulation category of each radio signal sample. Based on the characteristics of the modulated signal, we study three augmentation methods, i.e., rotation, flip, and Gaussian noise. After extensive numerical evaluations on an open radio signal dataset, we obtain the following contributions:

(1) We propose algorithms to augment radio signals at both training phase and inference phase of the deep learning algorithm, which achieves around 2.5% improvement on the baseline in terms of classification accuracy.

(2) We discover that the rotation augmentation method outperforms the flip method, both of which achieve higher classification accuracy than the Gaussian noise method.

(3) We propose a joint augmentation policy with both rotation and flip methods for insufficient training dataset. Given only 12.5% of training dataset, the joint augmentation method expands the dataset to be a size of 75% of the initial dataset and achieves an even higher classification accuracy than the baseline with 100% training dataset without augmentation.

(4) With data augmentation, we successfully classify radio samples by using only one half of the sampling points. Therefore, the deep learning model can be simplified with a significantly reduced inference complexity. Furthermore, in the future field deployment, the modulation category can be successfully classified upon receiving only half number of radio sampling points, which greatly reduces the classification response time.

The remainder of this paper is organized as follows. Section II presents related work. Section III provides an overview of the studied radio signal dataset and the deep learning-based modulation classifier. We introduce three data augmentation methods in Section IV and propose an algorithm to augment signals at both deep learning phases in Section V. In Section VI, we present the simulation setup and

the final experimental results. We finally conclude this paper in Section VII.

II. RELATED WORK

A. DEEP LEARNING IN RADIO MODULATION CLASSIFICATION

Deep learning has been applied to automatically classify radio modulation categories in recent literature. By converting radio signals into images, two convolutional neural network (CNN)-based deep learning models, GoogleNet [24] and AlexNet [18], originally developed for image classification, are used for modulation classification [13], [14]. The modulation classification accuracy is further improved by a modified deep residual network (ResNet) [12], which is fed with modulated in-phase (I) and quadrature phase (Q) signals. Considering channel interference, the CNN structure also achieves a considerable classification accuracy [11]. In addition to the CNN-based models, the Long Short-Term Memory (LSTM) architecture with time-dependent amplitude and phase information can achieve the state-of-the-art classification accuracy [16]. To reduce the training time of deep learning models, different subsampling techniques are investigated in [15] to reduce the dimensions of input signals.

B. DATA AUGMENTATION IN DEEP LEARNING

Data augmentation is widely used in deep learning algorithms to increase the diversity of training dataset, prevent model overfitting, and improve the robustness of the model. For image classification tasks, generic data augmentation methods include flip, rotation, cropping, color jittering, edge enhancement, and Fancy PCA [22]. Other complex data augmentation methods synthesize a new image from two training images [25] or from Generative Adversarial Nets (GAN) [26]. Although there are many augmentation methods for images, AutoAugment [27] is proposed to automatically search for augmentation policies based on the dataset. In addition to images, augmentation methods such as synonym replacement, random insertion, random swap, and random deletion are used for text classification [28], where the same accuracy as normal in all training data is achieved when only half of the training data is available. For speech recognition tasks, training audio is augmented by changing the audio speed [29], warping features, masking blocks of frequency channels, and masking blocks of time steps [30].

There are few related works on data augmentation for radio modulation classification in the literature. The most related work is a GAN based data augmentation method proposed in [14]. The authors first converted the signal samples into Contour Stellar Images which were further used to train the GAN network so as to generate new signal training samples. With GAN-based augmentation, the modulation classification accuracy is improved by no more than 6%. However, training GAN network still requires sufficient signal samples to guarantee the convergence. Moreover, as reported in [14], the classification accuracy based on augmented dataset is

lower than on real dataset with the same amount of signal samples. Therefore, an efficient augmentation method for insufficient radio signal dataset is still absent.

III. PRELIMINARIES

In this section, we introduce the radio signal dataset and the architecture of the state-of-the-art LSTM model [32], which will be used to evaluate different data augmentation methods presented in Sec. IV.

A. RADIO SIGNAL DATASET

We evaluate the radio signal modulation classification based on an open radio signal dataset, RadioML2016.10a [31]. The radio signals in the dataset consider sample rate offset, center frequency offset, multi-path fading and additive white Gaussian noise. Specifically, there are 220,000 modulated radio signal segments belonging to 11 different modulation categories, i.e., binary phase-shift keying (BPSK), QPSK, eight phase-shift keying (8PSK), continuous phase frequency-shift keying (CPFSK), Gauss frequency-shift keying (GFSK), pulse-amplitude modulation four (PAM4), quadrature amplitude modulation 16 (QAM16), quadrature amplitude modulation 64 (QAM64), double-sideband AM (AM-DSB), single-sideband AM (AM-SSB) and wide-band FM (WB-FM). Each radio signal sample is composed of 128 consecutive modulated in-phase (I) signal and quadrature phase (Q) signal. The labels of each signal sample include its value of signal-to-noise ratio (SNR) and its corresponding modulation category. There are total 20 different SNRs ranging from -20dB to 18dB with a step size of 2dB . In the dataset, these 220,000 signal samples are uniformly distributed among 11 modulation categories and 20 SNRs. In other words, there are 1,000 signal samples for each modulation category at each SNR. In Fig. 2, we plot examples of 11 modulation categories in forms of constellation diagrams under different SNRs. In the following subsection, we introduce a deep learning algorithm which automatically predicts the radio's modulation category based on its raw I/Q signals.

B. LSTM NETWORK ARCHITECTURE

LSTM is a special category of Recurrent neural network (RNN), which is widely used to process time series data. Benefited from a specific LSTM memory cell mechanism, LSTM effectively solves the exploding and vanishing gradient problem of traditional RNN during training process and learns long-term dependencies in sequential data. The LSTM memory cell mainly consists of a forget gate, an input gate and an update gate [33], which implement selective retention and discard of input information.

The LSTM network takes each data sample with consecutive modulated in-phase (I) and quadrature phase (Q) signals as input and maps them to a specific modulation category, whose architecture is shown in Fig. 3. Specifically, the modulated I/Q signals are first converted into amplitudes

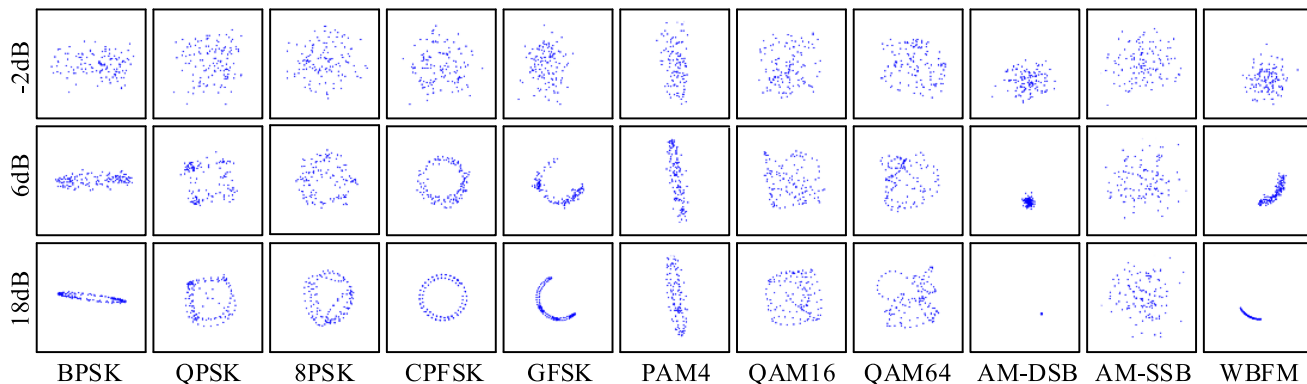


FIGURE 2. Constellation diagrams of 11 modulated signals [31] under different SNRs.

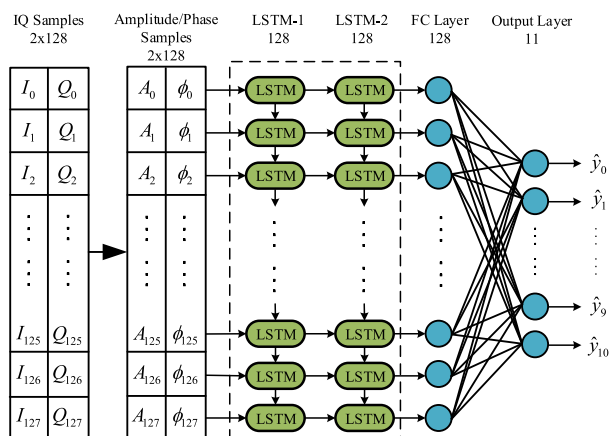


FIGURE 3. The architecture of LSTM network.

and phases [16], as:

$$\begin{cases} A = \sqrt{I^2 + Q^2} \\ \phi = \arctan(Q/I), \end{cases}$$

where A and ϕ represent the amplitude and phase of the modulated signal, respectively. The obtained signals are then fed into a two-layer LSTM network to extract characteristic features, where each layer has 128 LSTM cells. Finally, a fully connected layer with Softmax function is used to map the radio signal sample to one of these 11 modulation categories. Adam optimizer [34] with dynamic learning rate is used to minimize the cross-entropy loss as follows:

$$\ell = - \sum_{k=1}^K y_k \log(\hat{y}_k),$$

where K is the number of classes, y_k represents the ground truth label, and \hat{y}_k denotes the probability that the input sample will be predicted as k -th class.

IV. DATA AUGMENTATION METHODS

Data augmentation is a method widely used in deep learning because it improves the generalization ability of the model

and alleviates overfitting. In this section, we describe in detail three data augmentation methods for modulation signal recognition, including rotation, flip, and Gaussian noise. The dataset is expanded by a scale factor N .

A. ROTATION

By rotating a modulated radio signal (I, Q) around its origin, we obtain augmented signal sample (I', Q') as follows:

$$\begin{bmatrix} I' \\ Q' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} I \\ Q \end{bmatrix},$$

where θ is the angle of rotation. In this paper, the radio signal is rotated in the counter-clockwise direction by $0, \pi/2, \pi,$ and $3\pi/2$. In Fig. 4(a), we plot the constellation diagram of

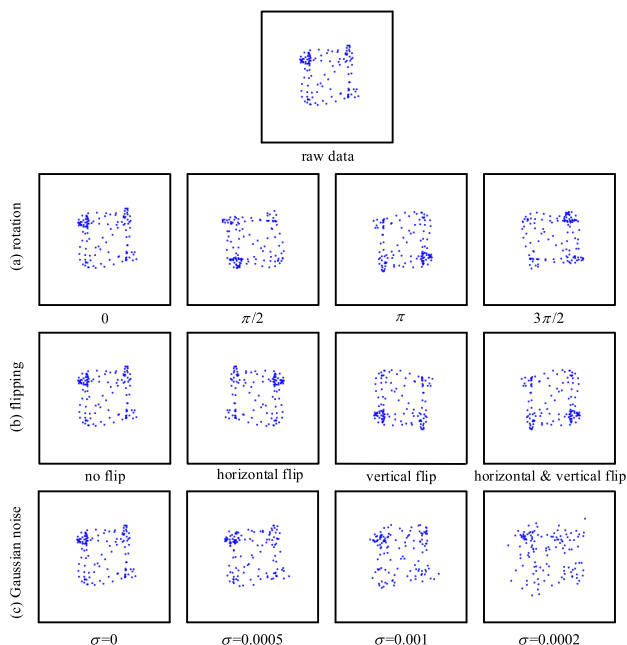


FIGURE 4. Constellation diagram of an QPSK radio signal sample with different data augmentation methods.

an QPSK sample where one set of raw data is augmented into four radio signal samples.

B. FLIP

For a given modulated radio signal (I, Q) , we define the horizontal flip by switching the I value to its opposite, as:

$$\begin{bmatrix} I' \\ Q' \end{bmatrix} = \begin{bmatrix} -I \\ Q \end{bmatrix},$$

and define the vertical flip by switching the Q value to its opposite, as:

$$\begin{bmatrix} I' \\ Q' \end{bmatrix} = \begin{bmatrix} I \\ -Q \end{bmatrix},$$

to augment the radio signals. We can perform horizontal flip, vertical flip, or both flips at the same time such that the signal dataset is expanded by a scale factor $N = 4$, as shown in Fig. 4(b).

C. GAUSSIAN NOISE

By adding a Gaussian noise $\mathcal{N}(0, \sigma^2)$ to the modulated radio signal (I, Q) , we obtain the augmented signal sample (I', Q') as:

$$\begin{bmatrix} I' \\ Q' \end{bmatrix} = \begin{bmatrix} I \\ Q \end{bmatrix} + \mathcal{N}(0, \sigma^2),$$

where σ^2 is the variance of noise. In Fig. 4(c), we show the augmented signal samples by adding Gaussian noise with different standard deviations $\sigma = 0, \sigma = 0.0005, \sigma = 0.001$ and $\sigma = 0.002$. For each data augmentation method, the original radio signal dataset is expanded by a default scale factor $N = 4$, as illustrated in Fig. 4. Note that the Gaussian noise data augmentation is supposed to significantly expand the dataset by choosing enough different values of σ . However, in the next section, we show that the Gaussian noise data augmentation is not preferred for radio data augmentation.

V. DATA AUGMENTATION TIME

The execution of a deep learning algorithm includes training phase and inference phase. Data augmentation can be performed in both phases, resulting in three possible combinations of augmentations, i.e., test-time augmentation, train-time augmentation, and train-test-time augmentation.

A. TRAIN-TIME AUGMENTATION

Train-time augmentation performs data augmentation during the training stage of the model. That is the training dataset is augmented and expanded by a scale factor N while the test dataset remains the same. Taking the rotation data augmentation as an example, the training dataset is expanded from 110,000 radio signal samples to 440,000 samples after train-time augmentation. In general, a larger size of training dataset leads to a higher modulation classification accuracy.

B. TEST-TIME AUGMENTATION

Test-time augmentation fuses features of all augmented radio signal samples in inference phase. In the inference phase, one radio signal sample (I, Q) in the test dataset is augmented into N samples $\{(I', Q')^n | n \in N\}$. Then each augmented sample $(I', Q')^n$ is fed into the LSTM network, and we obtain a vector of corresponding predicted probabilities \hat{y}_k^n . The predicted modulation category is decided through summing the predicted probabilities \hat{y}_k^n over all N augmented samples and choosing the one with maximum conference [35], as:

$$\arg \max_{1 \leq k \leq K} \sum_{n=1}^N \hat{y}_k^n.$$

C. TRAIN-TEST-TIME AUGMENTATION

Train-test-time augmentation conducts both train-time augmentation and test-time augmentation, where both training and test datasets are augmented and expanded by a factor N .

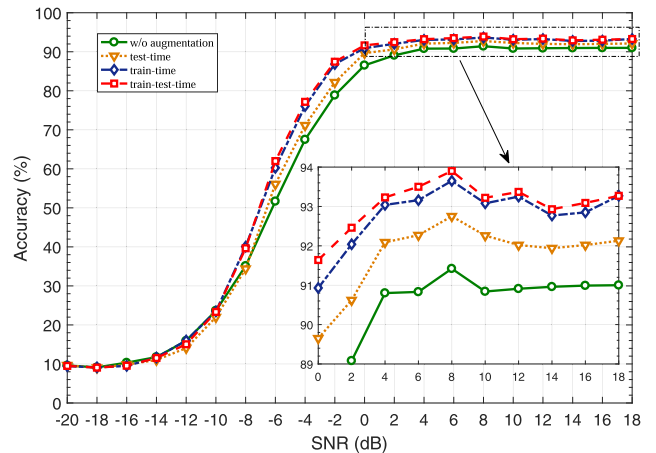


FIGURE 5. Classification accuracy under different augmentation times.

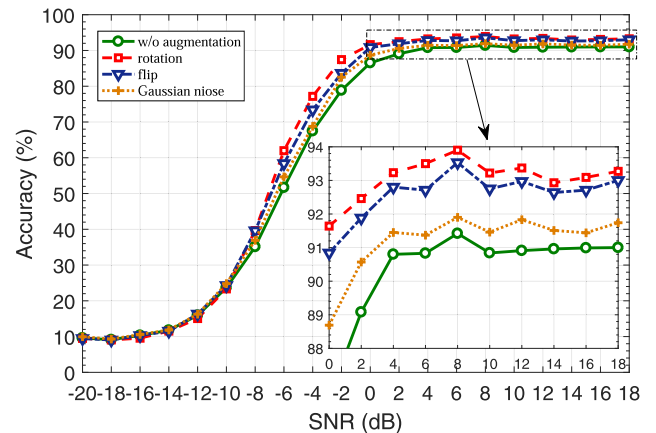


FIGURE 6. Classification accuracy under different augmentation methods.

In Fig. 5, we numerically study the performance of the data augmentation at different phases, where the rotation augmentation with a scale factor $N = 4$ is considered.

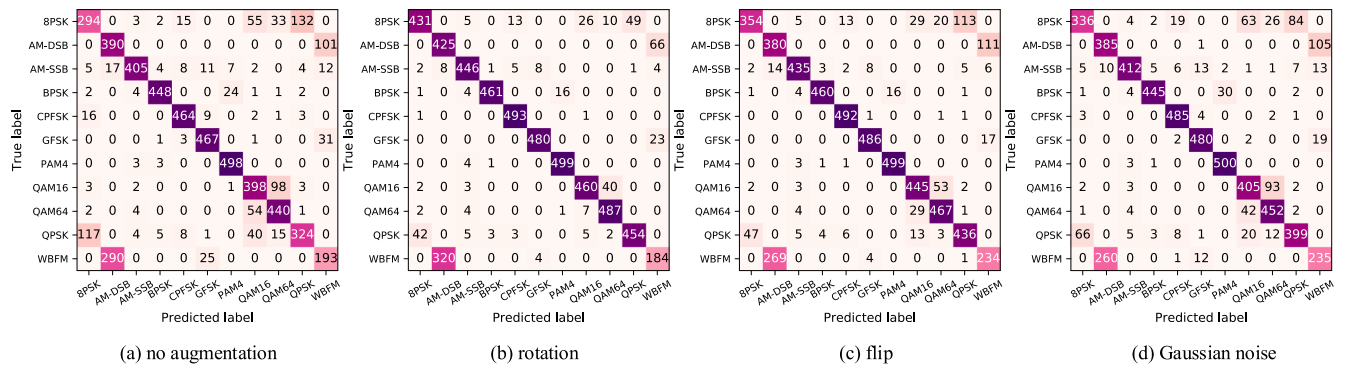


FIGURE 7. Confusion matrices under different data augmentation methods with 100% training dataset when SNR is -2dB .

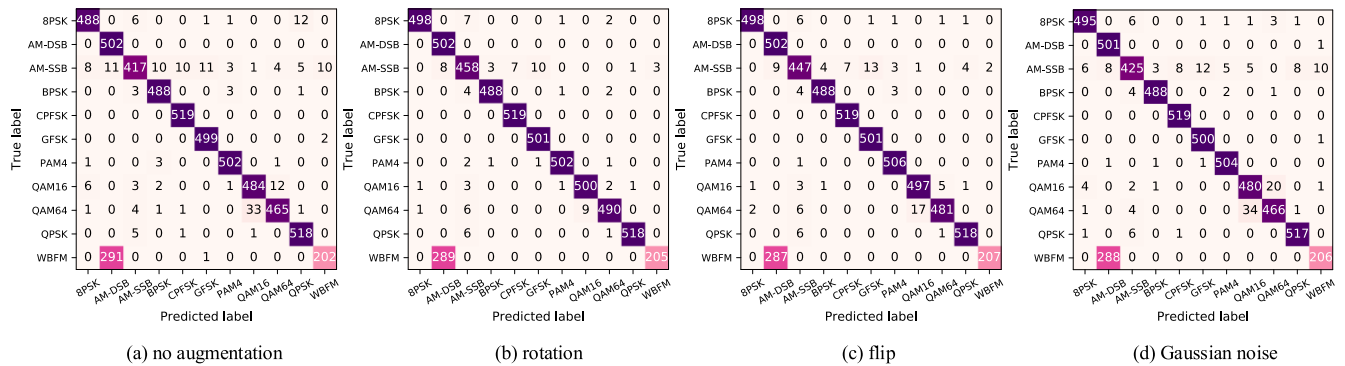


FIGURE 8. Confusion matrices under different data augmentation methods with 100% training dataset when SNR is 18dB .

Comparing with the baseline without augmentation, augmentations at different phases all improve the classification accuracy when the SNR is greater than -10dB . The train-time augmentation achieves better performance than test-time augmentation, and the train-test-time augmentation generates the highest accuracy. Specifically, comparing with the baseline, the train-test-time augmentation improves the modulation classification accuracy by 8.87% when SNR is -6dB and by about 2.2% when SNR is greater than 4dB . In the following numerical studies, we use the train-test-time augmentation by default.

VI. AUGMENTATION PERFORMANCE

In this section, we numerically study the performance of different radio data augmentation methods in terms of modulation classification accuracy. The open dataset, RadioML2016.10a, is divided equally into a training dataset and a test dataset, each containing 110,000 radio signal samples. In order to avoid overfitting, we set dropout rate to be 0.5 at both two LSTM layers. The number of training epoch is 80 and the mini-batch size is 128. The value of the learning rate is initially set as 0.001 and is halved when the training accuracy is not improved during three consecutive epochs. The model is implemented based on PyTorch [36].

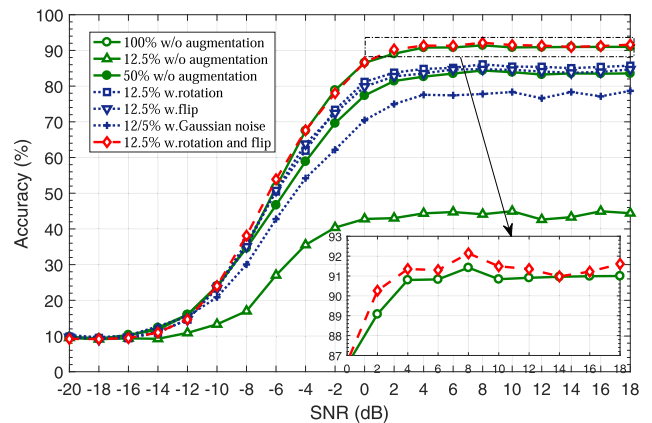


FIGURE 9. Classification accuracy versus different data augmentation methods under different SNRs with 12.5% training dataset.

A. AUGMENTATIONS ON FULL DATASET

In Fig. 6, we study the modulation classification accuracies of the LSTM model after deploying all three data augmentation methods presented in Sec. VI. Comparing with the baseline without augmentation, all augmentation methods improve the classification accuracy when the SNR is greater than -10dB , especially for the rotation data augmentation and flip data augmentation. In particular, the rotation data augmentation

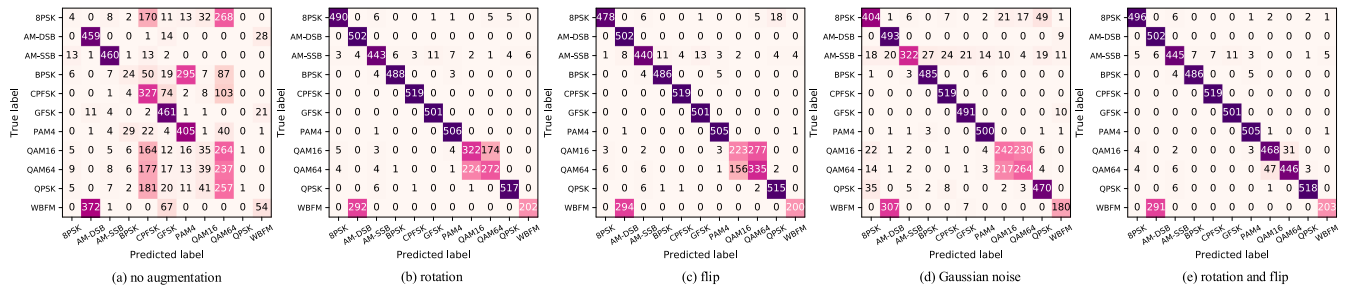


FIGURE 10. Confusion matrices under different data augmentation methods with 12.5% training dataset when SNR is 18dB.

method achieves the greatest improvement by 8% when SNR is between -6dB and -2dB and by about 2% at higher SNR ($\geq 4\text{dB}$). Meanwhile, the Gaussian noise data augmentation performs better at lower SNR when it is between -16dB and -10dB . Intuitively, adding Gaussian noise reduces the SNR of the original data sample which in turn generates more signal samples with low SNR. However, the improvement is trivial since the resulting classification accuracy is too small, less than 2% when SNR is smaller than 10 dB. Therefore, rotation data augmentation and flip data augmentation are more preferred for radio signals in modulation classification.

To further evaluate the improvements of different augmentation methods on classification accuracy, we present the corresponding confusion matrices these at low SNR (-2dB) and high SNR (18dB) in Fig. 7 and Fig. 8, respectively. Most values at diagonal entries of these matrices are increased after augmentation, which means the modulation classification accuracy is improved. Specifically, the proposed augmentation methods successfully reduce the confusion between QAM16 and QAM64 and solve the short-time observation problem presented in [37]. At low SNR, the LSTM model is difficult to classify 8PSK and QPSK, whose classification accuracy is greatly improved after rotation augmentation as shown in Fig. 7. At high SNR, the accuracy of the LSTM model is mainly limited by the confusion between AM-DSB and WBFM, which dues to frequent radio samples without information in the dataset [37]. In general, rotation and flip achieve better classification accuracy than Gaussian noise for all modulation categories.

B. AUGMENTATIONS ON PARTIAL DATASET

In Fig. 9, we further study the performance of different data augmentation methods with insufficient training dataset. To form new training sub-dataset, we randomly sample partial radio signal samples from the initial 110,000 radio signal training samples, i.e., 12.5% of the initial training dataset. Then, the LSTM network is trained by feeding the obtained training sub-dataset and is tested with the initial 110,000 radio signal testing samples. Note that 12.5% of the training dataset is insufficient to train the LSTM network, resulting a low modulation classification accuracy around 45% under high SNR, as shown in Fig. 9. After deploying different radio

data augmentation methods, the classification accuracy is improved. As expected, both the rotation augmentation and the flip augmentation outperform the Gaussian noise data augmentation. Interestingly, while the training sub-dataset is expanded by a scale factor $N = 4$ after augmentation, in the same size of 50% of the initial dataset, the rotation/flip augmentation achieves a higher classification accuracy, around 0.04%-4.03%, than the baseline by training the LSTM with 50% of the initial training dataset without augmentation.

We further consider a joint augmentation policy with both rotation and flip methods, which expands the dataset by a scale factor $N = 6$ (with 2 redundant augmented radio signal samples) as shown in Fig. 4(a-b). After this joint augmentation, the size of the training dataset is expanded from 12.5% to be 75% of the initial training dataset. Interestingly, we obtain similar classification accuracies at different SNRs as the baseline with 100% training dataset without augmentation, as plotted in Fig. 9. Note that such a classification accuracy is achieved by using 25% less training data.

To further evaluate the advantages of joint rotation and flip augmentation, we present confusion matrices in different augmentation methods with 12.5% training dataset at 18dB in Fig. 10. When training dataset is insufficient, it is difficult to classify BPSK, WBFM, QAM16 and QAM64, whose classification accuracies are significantly improved after joint augmentation. Specifically, in reducing the confusion between QAM16 and QAM64, the joint augmentation performs better than both the rotation augmentation and the flip augmentation.

We have also evaluated another joint augmentation with all three augmentation methods. However, adding Gaussian noise method to the joint rotation and flip augmentation slightly reduces the classification accuracy. Therefore, we conclude that both rotation and flip methods are preferred for radio data augmentation and they can be jointly applied to further improve the augmentation performance.

C. AUGMENTATIONS ON SHORT SAMPLE

We further evaluate data augmentation methods for modulated radio signals with fewer sampling points. We halve each original 128-point radio signal sample into two new samples and obtained a new dataset consisting of 440,000 entries

of 64-point radio signal samples. Similar to previous evaluations, we randomly choose half of them to the LSTM network, which is further tested with the remaining half dataset. With a shorter radio signal sample, the number of LSTM cells in each LSTM layer in Fig. 3 is reduced from 128 to 64, resulting a simpler inference model. Specifically, the number of parameters of the LSTM network is reduced from 201.1K to 54.1K and the inference complexity in FLOPs (floating-point operations) is reduced from 2.8K to 1.4K.

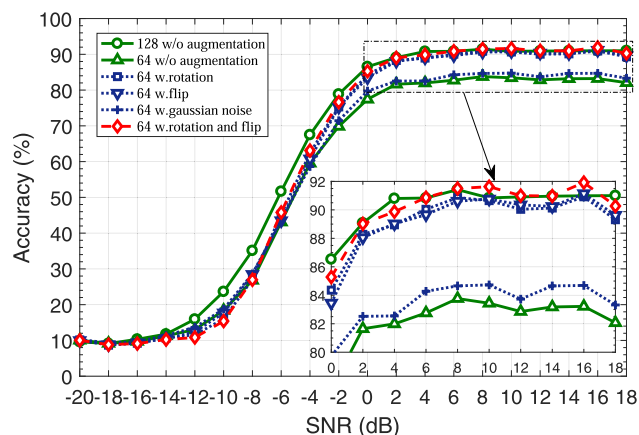


FIGURE 11. Classification accuracy versus different data augmentation methods under different SNRs with 64-length signal.

In Fig. 11, we evaluate modulation classifications with 64-point radio samples. Without augmentation, 64-point modulated radio samples always lead to lower classification accuracy than the baseline with 128-point, around an 8% reduction when SNR is greater than 0 dB. The classification accuracy is improved after deploying either rotation or flip augmentation. Especially, the joint rotation and flip augmentation can achieve 1% higher classification accuracy than the baseline under high SNR. Therefore, with data augmentation, the radio signal modulations can be successfully classified upon receiving only half number of sampling points, which significantly reduces the classification response time.

VII. CONCLUSION

In this paper, we studied radio data augmentation methods for deep learning-based modulation classification. Specifically, three typical augmentation methods, i.e., rotation, flip, and Gaussian noise, were studied based on a well-known LSTM model. We first studied radio data augmentations at training and inference phases and revealed that train-test-time augmentation achieves the highest accuracy. Then, we numerically evaluated all three augmentation methods based on the full and partial training dataset. All numerical results show that both the rotation and the flip methods achieve higher classification accuracy than the Gaussian noise method and the rotation method achieves the highest accuracy. Meanwhile, a joint augmentation policy with both rotation and flip methods can further improve the classification accuracy, especially with insufficient training samples. Given only 12.5% of

initial training dataset, the joint augmentation method expands the dataset to be a size of 75% of the initial dataset and obtains even higher than the baseline with 100% training datasets without augmentation. Furthermore, after deploying data augmentation, a radio sample can be classified based on only one half of the radio sampling points, resulting in a simplified deep learning model and a shorter classification response time.

REFERENCES

- [1] T. Afouras, J. S. Chung, A. Senior, O. Vinyals, and A. Zisserman, "Deep audio-visual speech recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, to be published, doi: 10.1109/TMC.2019.2928811.
- [2] M. M. Lopez and J. Kalita, "Deep learning applied to NLP," 2017, *arXiv:1703.03091*. [Online]. Available: <https://arxiv.org/abs/1703.03091>
- [3] X. Li, L. Zhao, L. Wei, M.-H. Yang, F. Wu, Y. Zhuang, H. Ling, and J. Wang, "DeepSaliency: Multi-task deep neural network model for salient object detection," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3919–3930, Aug. 2016.
- [4] A. Ali and F. Yangyu, "Automatic modulation classification using deep learning based on sparse autoencoders with nonnegativity constraints," *IEEE Signal Process. Lett.*, vol. 24, no. 11, pp. 1626–1630, Nov. 2017.
- [5] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.
- [6] A. Anderson, S. R. Young, T. P. Karnowski, and J. M. Vann, "Deepmod: An over-the-air trainable machine modem for resilient PHY layer communications," in *Proc. IEEE Mil. Commun. Conf.*, Oct. 2018, pp. 213–218.
- [7] L. Huang, S. Bi, and Y. J. Zhang, "Deep reinforcement learning for online computation offloading in wireless powered mobile-edge computing networks," *IEEE Trans. Mobile Comput.*, to be published, doi: 10.1109/TMC.2019.2928811.
- [8] A. Ali and W. Hamouda, "Advances on spectrum sensing for cognitive radio networks: Theory and applications," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1277–1304, 2nd Quart., 2016.
- [9] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: Classical approaches and new trends," *IET Commun.*, vol. 1, no. 2, pp. 137–156, Apr. 2007.
- [10] H. S. Jang and B. C. Jung, "Performance analysis of reactive symbol-level jamming techniques," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12432–12437, Dec. 2018.
- [11] P. Triantaris, E. Tsimbalo, W. H. Chin, and D. Gündüz, "Automatic modulation classification in the presence of interference," in *Proc. Eur. Conf. Netw. Commun.*, Jun. 2019, pp. 549–553.
- [12] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 168–179, Feb. 2018.
- [13] S. Peng, H. Jiang, H. Wang, H. Alwageed, Y. Zhou, M. M. Sebani, and Y.-D. Yao, "Modulation classification based on signal constellation diagrams and deep learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 3, pp. 718–727, Mar. 2019.
- [14] B. Tang, Y. Tu, Z. Zhang, and Y. Lin, "Digital signal modulation classification with data augmentation using generative adversarial nets in cognitive radio networks," *IEEE Access*, vol. 6, pp. 15713–15722, 2018.
- [15] S. Ramjee, S. Ju, D. Yang, X. Liu, A. E. Gamal, and Y. C. Eldar, "Fast deep learning for automatic modulation classification," 2019, *arXiv:1901.05850*. [Online]. Available: <https://arxiv.org/abs/1901.05850>
- [16] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Trans. Cogn. Commun. Netw.*, vol. 4, no. 3, pp. 433–445, Sep. 2018.
- [17] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *Proc. Int. Conf. Eng. Appl. Neural Netw.*, 2016, pp. 213–226.
- [18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2012, pp. 1097–1105.
- [19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.

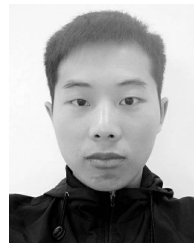
- [20] J. Salamon and J. P. Bello, "Deep convolutional neural networks and data augmentation for environmental sound classification," *IEEE Signal Process. Lett.*, vol. 24, no. 3, pp. 279–283, Mar. 2017.
- [21] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, "ORACLE: Optimized radio classification through convolutional neural networks," in *Proc. IEEE INFOCOM*, Apr./May 2019, pp. 370–378.
- [22] L. Taylor and G. Nitschke, "Improving deep learning with generic data augmentation," in *Proc. IEEE Symp. Ser. Comput. Intell.*, Nov. 2018, pp. 1542–1547.
- [23] J. Shijie, W. Ping, J. Peiyi, and H. Siping, "Research on data augmentation for image classification based on convolution neural networks," in *Proc. Chin. Automat. Congr.*, Oct. 2017, pp. 4165–4170.
- [24] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1–9.
- [25] H. Inoue, "Data augmentation by pairing samples for images classification," 2018, *arXiv:1801.02929*. [Online]. Available: <https://arxiv.org/abs/1801.02929>
- [26] A. Antoniou, A. Storkey, and H. Edwards, "Data augmentation generative adversarial networks," 2017, *arXiv:1711.04340*. [Online]. Available: <https://arxiv.org/abs/1711.04340>
- [27] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. V. Le, "Autoaugment: Learning augmentation policies from data," 2018, *arXiv:1805.09501*. [Online]. Available: <https://arxiv.org/abs/1805.09501>
- [28] J. W. Wei and K. Zou, "EDA: Easy data augmentation techniques for boosting performance on text classification tasks," 2019, *arXiv:1901.11196*. [Online]. Available: <https://arxiv.org/abs/1901.11196>
- [29] T. Ko, V. Peddinti, D. Povey, and S. Khudanpur, "Audio augmentation for speech recognition," in *Proc. Annu. Conf. Int. Speech Commun. Assoc.*, 2015.
- [30] D. S. Park, W. Chan, Y. Zhang, C.-C. Chiu, B. Zoph, E. D. Cubuk, and Q. V. Le, "SpecAugment: A simple data augmentation method for automatic speech recognition," 2019, *arXiv:1904.08779*. [Online]. Available: <https://arxiv.org/abs/1904.08779>
- [31] T. J. O'shea and N. West, "Radio machine learning dataset generation with gnu radio," in *Proc. GNU Radio Conf.*, 2016, vol. 1, no. 1.
- [32] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [33] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," in *Proc. ICANN*, vol. 2, Sep. 1999, pp. 850–855.
- [34] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [35] S. Zheng, P. Qi, S. Chen, and X. Yang, "Fusion methods for cnn-based automatic modulation classification," *IEEE Access*, vol. 7, pp. 66496–66504, 2019.
- [36] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017.
- [37] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.



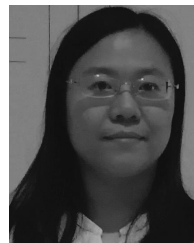
LIANG HUANG (M'16) received the B.Eng. degree in communications engineering from Zhejiang University, Hangzhou, China, in 2009, and the Ph.D. degree in information engineering from The Chinese University of Hong Kong, Hong Kong, in 2013. He is currently an Assistant Professor with the College of Information Engineering, Zhejiang University of Technology, China. His research interest includes the areas of queuing and scheduling in communication systems and networks.



WEIJIAN PAN received the B.E. degree from the College of Mechanical and Electrical Engineering, China Jiliang University, Hangzhou, China, in June 2018. He is currently pursuing the master's degree with the Information Engineering College, Zhejiang University of Technology. His research interests are deep learning and pattern recognition.



YOU ZHANG received the B.E. degree in electrical engineering and automation from the Luoyang Institute of Science and Technology, China, in June 2018. He is currently pursuing the master's degree with the Information Engineering College, Zhejiang University of Technology. His research interests are deep learning and signal communication.



LIPING QIAN received the Ph.D. degree in information engineering from The Chinese University of Hong Kong, in 2010. She is currently a Professor with the College of Information Engineering, Zhejiang University of Technology, China. She is also with the National Mobile Communications Research Laboratory, Southeast University, China. Her research interests include wireless communication and networking, resource management in wireless networks, the massive IoTs, mobile edge computing, emerging multiple access techniques, and machine learning oriented toward wireless communications. She was a co-recipient of the IEEE Marconi Prize Paper Award in Wireless Communications, in 2011.



NAN GAO received the master's degree in safety technology and engineering from Tsinghua University, in 2008, and the Ph.D. degree in computer science from the University of South Carolina, in 2014. She is currently an Assistant Professor with the College of Computer Science and Technology, Zhejiang University of Technology, and the Director of the Natural Language Processing Institute of Guoxin Big Data Research Institute. Her research interests include algorithm optimization, data mining, and artificial intelligence.



YUAN WU (S'08–M'10–SM'16) received the Ph.D. degree in electronic and computer engineering from The Hong Kong University of Science and Technology, Hong Kong, in 2010. He is currently an Associate Professor with the State Key Laboratory of Internet of Things for Smart City, University of Macau, Macao, China, and also with the Department of Computer and Information Science, University of Macau, Macao. Prior to that, he was a Full Professor with the College of Information Engineering, Zhejiang University of Technology, China. His research interests focus on resource management for wireless networks, green communications and computing, mobile edge computing, and smart grid. He was a recipient of the Best Paper Award of the IEEE International Conference on Communications, in 2016, and the Best Paper Award of the IEEE Technical Committee on Green Communications and Computing, in 2017.

• • •