Complex-Valued Networks for Automatic Modulation Classification

Ya Tu, *Student Member, IEEE*, Yun Lin, *Member, IEEE*, Changbo Hou, and Shiwen Mao, *Fellow, IEEE*

***Abstract*—Deep learning (DL) has been recognized as an effective solution for automatic modulation classification (AMC). However, most recent DL based AMC works are based on real-valued operations and representations. In this correspondence, we aim to demonstrate the high potential of complex-valued networks for AMC. We present the design of several key building blocks for complex-valued networks, such as complex convolution, complex batch-normalization, complex weight initialization, and complex dense strategies. We then provide a comparison study of three different neural network models and their complex-valued counterparts using the RadioML 2016.10 A dataset. Our results validate the superior performance in AMC achieved by the complex-valued networks.**

***Index Terms*—Automatic modulation classification, deep learning, complex-valued networks.**

I. Introduction

Automatic modulation classification (AMC) is a classical problem in modern wireless communications. AMC is to understand and label the radio spectrum under non-cooperative communication scenarios, which will greatly facilitate radio fault detection, spectrum interference monitoring, and dynamic spectrum access [1]–[3].Vehicular networks aim to provide sustainable and reliable communications for smart vehicles. The vehicle-to-vehicle (V2V) communication in vehicular networks was standardized in the form of an amendment to the IEEE 802.11 standard. AMC is one the most important parts for link adaption (LA) [4], which is widely applied into vehicle-to-vehicle networks [5] or vehicle-to-grid networks [6]. At present, the high-data rates and the explosive increase of wireless communications data make it challenging to achieve precise classification by traditional algorithms. To this end, deep learning (DL) provides a highly promising solution for handling the raw, high dimensional wireless data, rather than extracting hand-crafted and specialized features.

Manuscript received January 5, 2020; revised April 12, 2020; accepted June 16, 2020. Date of publication June 26, 2020; date of current version October 13, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61771154, in part by the Fundamental Research Funds for the Central Universities under Grant 3072020CF0813, and in part by the Ph.D. Student Research and Innovation Fund of the Fundamental Research Funds for the Central Universities under Grant 3072020GIP0813. The review of this article was coordinated by Prof. H. Zhu. *(Corresponding author: Yun Lin.)*

Ya Tu, Yun Lin, and Changbo Hou are with the College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China (e-mail: tuya@hrbeu.edu.cn; linyun@hrbeu.edu.cn; houchangbo@hrbeu.edu.cn).

Shiwen Mao is with the Department of Electricla and Computer Engineering, Auburn University, Auburn, AL 36849 USA (e-mail: smao@ieee.org).

Digital Object Identifier 10.1109/TVT.2020.3005707

Recent research advances have made significant progress in designing and applying DL framework for wireless communications [7]–[19]. For example, O'Shea *et al.* in [20] proposed the first modulation classification dataset, which is somewhat comparable to the canonical MNIST dataset (Modified National Institute of Standards and Technology dataset). In order to make full use of DL's capacity as in the computer vision field, Peng *et al.* in [21] exploited colored constellation diagrams to represent digital signals and AlexNet to precisely recognize them. Hu *et al.* in [22] proposed a multiple recurrent neural network for AMC. Their experiment result showed that the proposed RNN-based classifier was robust to uncertain noise conditions. To deal with the varying input dimensions, Meng *et al.* in [23] proposed a novel two-step training procedure and used transfer learning to improve the efficiency of retraining. Then, Wang *et al.* in [24] combined two convolutional neural networks (CNNs) trained on different datasets to achieve higher AMC accuracy. Recently, Tu *et al.* in [25] also utilized semi-supervised learning generative adversarial networks to conduct semi-supervised learning on AMC, where dataset is incompletely labeled. Despite the relatively rich literature on application of DL in communications, there are only few examples considering complex-valued representation of signal attributes.

It is well-know that the amount of information present in the phase of a signal is sufficient to recover the majority of the information encoded in its magnitude [26]. Most communication models favor a complex-valued signal to represent its magnitude and phase information. However, complex-valued deep neural networks for wireless communications have been largely ignored since complex valued operations are not yet mature enough for practical use [1]. Therefore, researchers often use an ordered pair of real-valued numbers to represent I/Q data and design various architectures for this data format [1], [20], [23], [27].

In this correspondence, our motivation to exploit complex-valued networks for AMC is twofold. First, the real and imaginary parts of the signal are dependent on each other under any change in phase due to the shift effect. A real-valued model will deem the real and imaginary parts of the signal as independent, while a complex-valued model will consider the correlation between the signal's real and imaginary parts. Second, as pointed in [28], magnitude and phase are important to the learning objective, so it make sense to employ a complex-valued model. Representing a univariate complex parameter rather than two bivariate real parameters, we have a solution space with fewer degrees of freedom, which is easily interpreted in terms of magnitude and phase. In order to fully harvest the high potential of complex representations, our contributions in this correspondence are summarized as follows:

1) present a general formulation of the building blocks of complex-valued deep neural networks.
2) conduct a comparison study of complex-valued and real-valued models using the RadioML 2016.10 A dataset.
3) demonstrate the superior performance of complex-valued models through our experimental study and discuss future works in this area.

The remainder of this paper is organized as follows. Section II presents the design of complex neural network building blocks. In Section III, we design several different network architectures and our experimental comparison of complex-valued and real-valued networks for AMC. Finally, conclusions and future work are presented in Section IV.

II. Complex Neural Network Building Blocks

In this section, we present the implementation of the complex-valued building blocks of the complex neural network, including the complex convolution block, complex batch normalization, complex dense block and complex weight initialization.

A. Complex Convolutional Layer

To simulate complex arithmetic using real-valued arithmetic internally, we set the first half feature maps to represent the real components and the second half feature maps to represent the imaginary parts. We define a complex filter weight matrix $\mathbf{W} = \mathbf{A} + i\mathbf{B}$ and a complex vector $\vec{s} = x + iy$ to represent the I/Q signal. Convolving the complex vector by the filter we obtain:

$$\mathbf{W} * \vec{s} = (\mathbf{A} * x - \mathbf{B} * y) + i(\mathbf{B} * x + \mathbf{A} * y). \quad (1)$$

This process is illustrated in Fig. 1.

We then obtain the real part and the imaginary part of the convolution result as:

$$Real(\mathbf{W} * \vec{s}) = (\mathbf{A} * x - \mathbf{B} * y) \quad (2)$$

$$Imag(\mathbf{W} * \vec{s}) = (\mathbf{A} * y + \mathbf{B} * x). \quad (3)$$

B. Complex Batch Normalization

Batch Normalization is an essential technique to optimize the DL model [29]. The real-valued batch normalization layer can be formulated as:

$$BN(\vec{x}) = \gamma \cdot \frac{\vec{x} - \mathbb{E}(\vec{x})}{\sqrt{\mathbf{V} + \epsilon}} + \beta, \quad (4)$$

where $\mathbb{E}(\vec{x})$ and $\mathbf{V}$ are the mean and variance of the output of mini-batch data, respectively; γ and β are trainable parameters, and ϵ is a minimum value to prevent denominators from being zero. However, batch normalization applies only to real values in prior work. For complex-valued neural network, complex batch normalization [28] will firstly obtain 0-centered data as in (5).

$$\vec{x}_0 = \vec{x} - \mathbb{E}\left[\vec{x}\right]. \quad (5)$$

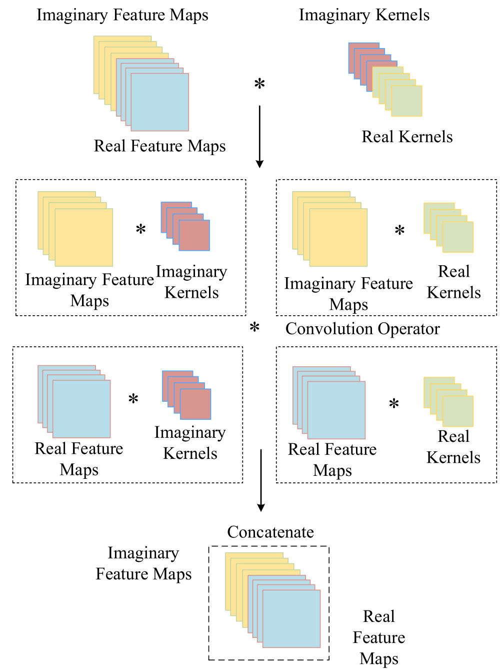

Fig. 1. Complex-valued convolution operation.

We then multiply (5) by the inverse square root of the 2×2 covariance matrix to obtain the normalized complex vector.

$$\tilde{\vec{x}} = (\mathbf{V})^{-\frac{1}{2}}(\vec{x} - \mathbb{E}[\vec{x}]), \quad (6)$$

where the covariance matrix $\mathbf{V}$ is:

$$\mathbf{V} = \begin{pmatrix} \mathrm{Cov}(Real\{\vec{x}\}, Real\{\vec{x}\}) & \mathrm{Cov}(Real\{\vec{x}\}, Imag\{\vec{x}\}) \\ \mathrm{Cov}(Imag\{\vec{x}\}, Real\{\vec{x}\}) & \mathrm{Cov}(Imag\{\vec{x}\}, Imag\{\vec{x}\}) \end{pmatrix}. \quad (7)$$

As batch normalization for real values, we also set the learnable shift parameter $\vec{\beta}$ and scaling parameter $\boldsymbol{\gamma}$ for complex batch normalization. Scaling parameter $\boldsymbol{\gamma}$ is given by:

$$\boldsymbol{\gamma} = \begin{pmatrix} \gamma_{rr} & \gamma_{ri} \\ \gamma_{ri} & \gamma_{ii} \end{pmatrix}. \quad (8)$$

The real part and imaginary part of the shift parameter β and γ_{ri} in the scaling parameter $\boldsymbol{\gamma}$ will be initialized to zero, while γ_{rr} and γ_{ii} in the scaling parameter $\boldsymbol{\gamma}$ will be initialized to $\frac{1}{\sqrt{2}}$. The initialization will satisfy a modulus of 1 for the variance of the normalized value [28].

C. Complex Dense Layer

The dense layer often serves as the classifier in DNN. To make full use of the complex-valued statistical information, we also present a mechanism, named Complex Dense Layer, to accept complex-valued feature while computing complex-valued classification results. We define a complex dense vector weight $\mathbf{W} = \mathbf{A} + i\mathbf{B}$ and a complex vector $\vec{s} = x + iy$ to represent complex-valued input. Similar to the complex convolutional

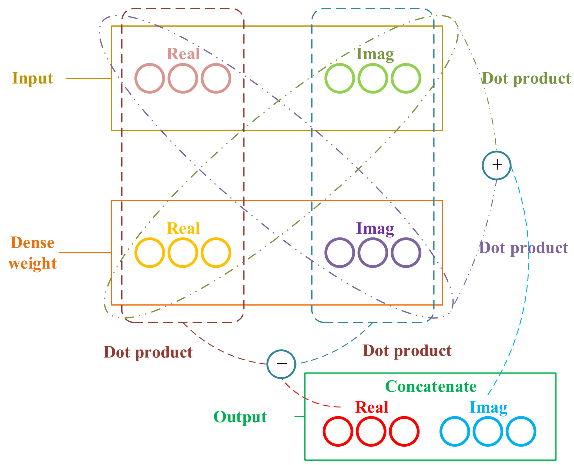

Fig. 2. The complex-valued dense layer.

operation, we have:

$$\mathbf{W} \cdot \vec{s} = (\mathbf{A} \cdot x - \mathbf{B} \cdot y) + i(\mathbf{B} \cdot x + \mathbf{A} \cdot y), \tag{9}$$

where $(\cdot)$ denotes the dot product. The overall process is depicted in Fig. 2.

D. Complex Weight Initialization

It is well known that proper initialization for neural network is vital in reducing the risks of vanishing or exploding gradients. To accomplish this, we follow the same steps as in [28], where a complex weight can be represented as:

$$\mathbf{W} = Real\{\mathbf{W}\} + iImag\{\mathbf{W}\}. \tag{10}$$

In the case when $\mathbf{W}$ is symmetrically distributed around 0, the variance of $\mathbf{W}$ can be estimated from a Rayleigh distribution's single parameter σ [28]. According to [28], we initialize the complex weight magnitude $\mathbf{W}$ from a Rayleigh distribution, with an expectation parameter 0 and variance parameter $2\sigma^2$. The parameter σ will be set differently according to different neural network architecture.

III. Experimental Validation

In this section, we present a comparison study between the real-valued model and complex-valued model. We will first present the dataset used in our experiments. We will then introduce our proposed deep signal network (DSN) model, the model from [20], [23], and their complex-valued counterparts. The real-valued model has double-input-terminal and will take I/Q waveform data as its input. Finally we will present our experimental results.

A. Dataset

In this paper, we choose the RadioML 2016.10A [20] dataset for our experiments. RadioML 2016.10 A is a synthetic dataset generated with GNU Radio. It consists of 11 modulations (8 digital and 3 analog) at various signal-to-noise ratios (SNR). In every category, there are 1,000 samples of modulated signals. To model the real communication scenario, this dataset includes a number of desired effects such as random processes for the center frequency offset, sample rate offset, additive white Gaussian noise, multi-path, and fading, respectively. In our experiments, we choose 8,800 I/Q data samples as the training dataset (denoted by "trainset") and 2,200 I/Q data samples as the test dataset (denoted by "testset") per SNR level. It is ensured that the selection of examples is random and that the classes are balanced. Each of the 11 classes of modulated signal will have 800 randomly selected samples in the trainset and 200 remaining samples in the testset. The SNR ranges from −20 dB to 18 dB.

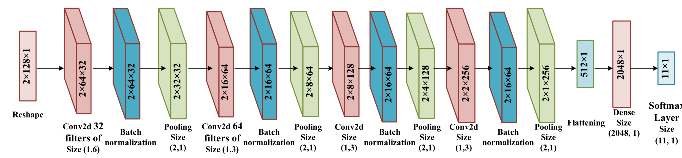

Fig. 3. The proposed DSN architecture.

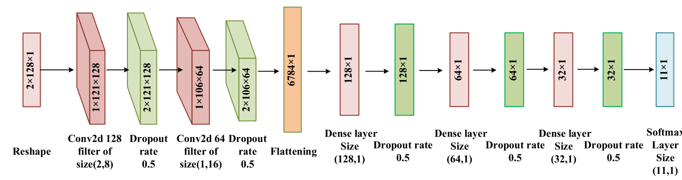

Fig. 4. The DrCNN architecture [23].

B. Network Architecture

1) Deep Signal Network: We propose a deep signal network (DSN) for AMC. DSN has four convolution units, each comprised of one convolutional layer, followed by one batch normalization layer and one average pooling layer. The next classification layer consisted of two fully-connected layers, with 2,048 and 11 neurons, respectively. The rectified linear unit (ReLU) is selected as the activation function for all the layers except the last fully connected layer, where Softmax is used to obtain the probability distribution matrix of the last layer. The DSN model and each layer's output shape are depicted in Fig. 3.

The complex-valued Deep Signal Network has the same architecture as DSN, but the covolutional layer and the dense layer are replaced by the complex convolutional layer and complex dense layer, respectively.

2) DrCNN: DrCNN [23] has six layers (two convolution layers and four fully connected layers), as shown in Fig. 4. The parametric rectified linear unit (PReLU) is selected as the activation function for all the layers except the last fully connected layer, where Softmax is used to obtain the probability distribution matrix of the last layer. Different from the CNN model proposed in [30], where the maximum pooling operation is applied, a dropout layer (rather than a pooling operation) follows every convolutional layer instead. Dropout replaces the pooling operation in this model because the pooling operation involves down-sampling, which can result in loss of signal characteristics, while dropout will not neglect important features of the signal. In addition, pooling is applied for dimensionality reduction to accelerate computation and avoid overfitting when the network size is large. Dropout also helps to avoid overfitting.

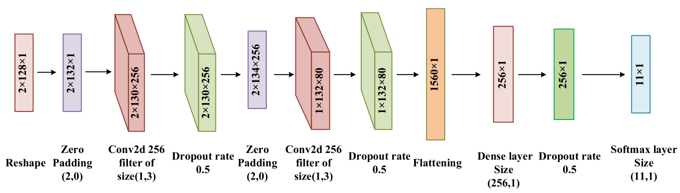

Fig. 5. The VTCNN2 architecture [20].

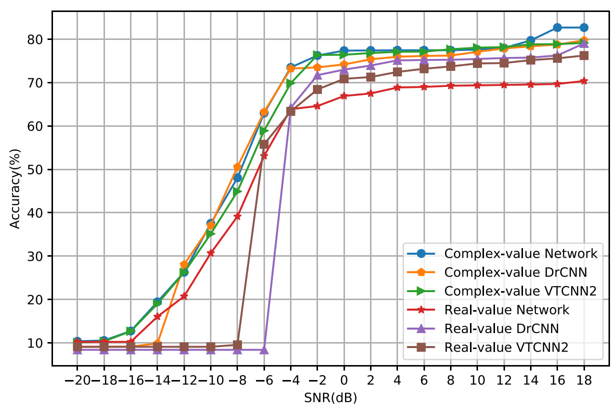

Fig. 6. Classification result comparison between complex-valued model and real-valued model.

The complex-valued DrCNN has the same architecture as DrCNN, but the covolutional layer and the dense layer are replaced by the complex convolutional layer and the complex dense layer, respectively. The architecture and each layer's output shape of DrCNN are illustrated in Fig. 4.

3) VTCNN2: VTCNN2 [20] is a 4-layer network utilizing two convolutional layers and two dense fully connected layers. All the layers use the ReLU activation function except that the output layer uses the Softmax activation. Ref. [20] use this network depth as it is roughly equivalent to networks that work well on similar simple datasets in the vision domain such as the MNIST dataset. The complex-valued VTCNN2 has the same architecture, but the covolutional layer and the dense layer are replaced by the complex convolutional layer and the complex dense layer, respectively. The architecture and each layer's output shape of VTCNN2 are described in Fig. 5.

In our experiments, all the networks were optimized with Adam and the learning rate is set to 10^{-4}. Epochs means the entire dataset is passed forward and backward through the neural network only once. The early stop patience is set to 20 epochs, which means the training process will be terminated after 20 epochs. The implementation of the models is based on Keras with Tensorflow as the backend. To have a fair comparison, we only set half real-valued neural network parameters amount for complex-valued neural network parameters amount to make both representation capacity equivalent [28].

C. Experiment Results

After training the models with the trainset and test the trained models with the testset, we obtain the results as presented in Fig. 6. We can make the following observations.

1) Each complex-valued model performance is better than that of the corresponding real-valued models.
2) At relative higher SNRs (0$\sim$18 dB), DSN performs around 10% better than RSN, complex-valued DrCNN2 performs about 1$\sim$3% better than real-valued DrCNN2, and complex-valued VTCNN2 performs around 3$\sim$6% better than real-valued VTCNN2.
3 At relative lower SNR($-20\sim0$ dB), we can see that DSN, complex-valued DrCNN2, and complex-valued VTCNN2 start to extract useful features from the I/Q waveforms at -18, -16, and -14 dB, respectively, while their real-valued counterparts start to learn useful features at -16, -8, -6 dB, respectively. This indicates that complex-valued operations can help the neural network extract correct statistical information than real-valued operations.
4) Our proposed model, i.e., the complex-valued DSN, outperforms both DrCNN2 and VTCNN2, as well as its complex-valued counterpart. Therefore, we believe the complex-valued DSN will be a good choice for AMC.
5) As has been mention in [28], a complex multiplication is 4 times more expensive than its real counterpart if they have the same parameters. We set the half real-valued neural network parameter amount to the complex-valued neural network. Therefore, the complex-valued neural network in this paper is about twice more expensive than its real counterpart

We conjecture that the reason behind this phenomenon is that complex-valued networks can effectively detect and extract the correlation between real and imaginary parts of the signal. Moreover, a complex-valued model provides a more constrained system than that based on real numbers.

IV. Conclusion

Complex neural networks have existed for some time, in the research community. In this paper, we proposed to apply complex neural networks for AMC. We presented the key building blocks in complex valued neural networks, such as the complex convolutional layer, complex batch normalization, complex weight initialization, and complex dense layer. Through experiments, we demonstrated that the complex-valued network outperforms its real-valued counterpart in AMC. To the best of our knowledge, this is the first work on applying complex neural networks to solving the AMC problem. We believe that complex neural training have been proposed for a while, but received few attend inference of deep models is an emerging and exciting field of machine learning for signal processing, e.g., MIMO systems, radio signal analysis, etc.

However, we also found that complex-valued model has a higher computational complexity, since it computes the correlation information between input real and imaginary parts and layer weight's real and imaginary parts. So it is a challenge to deploy a complex-valued neural network at edge devices with limited computational capabilities and storage. Fortunately, the neural network compression technique provides a solution to alleviate the computation and storage burden without sacrificing much accuracy. We will explore effective neural network compression techniques for complex-valued networks in our future.

REFERENCES

[1] 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.

[2] 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. 12, no. 1, pp. 168–179, Jan. 2019.

[3] Y. Kawamoto, H. Takagi, H. Nishiyama, and N. Kato, "Efficient resource allocation utilizing Q-learning in multiple UA communications," *IEEE Trans. Netw. Sci. Eng.*, vol. 6, no. 3, pp. 293–302, Jan. 2019.

[4] G. Gui, M. Liu, F. Tang, N. Kato, F. Adachi, "6G: Opening new horizons for integration of comfort, security and intelligence," *IEEE Wireless Commun. Mag.*, to be published, doi: 10.1109/MWC.001.1900516.

[5] R. C. Daniels and R. W. Heath, "Link adaptation with position/motion information in vehicle-to-vehicle networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 505–509, Feb. 2012.

[6] E. Mohamed, T. Wang, and S. Zhang, "Deep learning for joint adaptations of transmission rate and payload length in vehicular networks," *Sensors*, vol. 19, no. 5 pp. 1113–1118.

[7] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surv. Tut.*, vol. 21, no. 3, pp. 2224–2287, Mar. 2019.

[8] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key technologies and open issues," *IEEE Commun. Surv. Tut.*, vol. 12, no. 4, pp. 3072–3108, Oct.–Dec. 2019.

[9] Z. Md. Fadlullah, F. Tang, B. Mao, J. Liu, and N. Kato, "On intelligent traffic control for large-scale heterogeneous networks: A value matrix-based deep learning approach," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2479–2482, Dec. 2018.

[10] F. Tang, Z. M. Fadlullah, B. Mao, and N. Kato, "An intelligent traffic load prediction-based adaptive channel assignment algorithm in SDN-IoT: A deep learning approach," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 5141–5154, Dec. 2018.

[11] Z. M. Fadlullah, B. Mao, F. Tang, and N. Kato, "Value iteration architecture based deep learning for intelligent routing exploiting heterogeneous computing platforms," *IEEE Trans. Comput.*, vol. 68, no. 6, pp. 939–950, Jan. 2019.

[12] G. Gui, F. Liu, J. Sun, J. Yang, Z. Zhou, and D. Zhao, "Flight delay prediction based on aviation big data and machine learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 140–150, Jan. 2020.

[13] H. Huang, Y. Peng, J. Yang, W. Xia, and G. Gui, "Fast beamforming design via deep learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 1065–1069, Jan. 2020.

[14] L. Zhao, J. Wang, J. Liu, and N. Kato, "Routing for crowd management in smart cities: A deep reinforcement learning perspective," *IEEE Commun. Mag.*, vol. 57, no. 4, pp. 88–93, Apr. 2019.

[15] N. Kato *et al.*, "The deep learning vision for heterogeneous network traffic control: Proposal, challenges, and future perspective," *IEEE Wireless Commun.*, vol. 24, no. 3, pp. 146–153, Mar. 2017.

[16] W. Wang, J. Yang, M. Liu, and G. Gui, "LightAMC: Lightweight automatic modulation classification using deep learning and compressive sensing," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3491–3495, Mar. 2020.

[17] N. Kato, B. Mao, F. Tang, Y. Kawamoto, and J. Liu, "Ten challenges in advancing machine learning technologies towards 6G," *IEEE Wireless Commun. Mag.*, vol. 27, no. 3, pp. 96–103, Jun. 2020.

[18] J. Sun, W. Shi, Z. Han, J. Yang, and G. Gui, "Behavioral modeling and linearization of wideband RF power amplifiers using BiLSTM networks for 5G wireless systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 10348–10356, Nov. 2019.

[19] L. Huang, W. Pan, Y. Zhang, L Qian, N. Gao, and Y. Wu, "Data augmentation for deep learning-based radio modulation classification," *IEEE Access*, vol. 8, no. 1, pp. 1498–1506, 2019

[20] T. J. O'Shea, and N. West, "Radio machine learning dataset generation with GNU radio," in *Proc. GNU Radio Conf.*, Boulder, CO, USA, Sep. 2016, pp. 1–6.

[21] S. Peng *et al.*, "Modulation classification based on signal constellation diagrams and deep learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 99, pp. 1–10, Mar. 2019.

[22] S. Hu, Y. Pei, P. Liang, and Y. Liang, "Deep neural network for robust modulation classification under uncertain noise conditions," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 564–577, Jan. 2020.

[23] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4074–4077, Apr. 2019.

[24] F. Meng, P. Chen, L. Wu, and X. Wang, "Automatic modulation classification: A deep learning enabled approach," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10760–10772, Nov. 2018.

[25] Y. Tu, Y. Lin, J. Wang, and J. U. Kim, "Semi-supervised learning with generative adversarial networks on digital signal modulation classification," *Comput. Mater. Continua*, vol. 55, no. 2, pp. 243–254, May 2018.

[26] G. Shi, M. M. Shanechi, and P. Aarabi, "On the importance of phase in human speech recognition," *IEEE Trans. Audio Speech Lang. Process.*, vol. 14, no. 5, pp. 1867–1874, Sep. 2006.

[27] Y. Zhou, Z. Md. Fadlullah, B. Mao, and N. Kato, "A deep-learning-based radio resource assignment technique for 5G ultra dense networks," *IEEE Netw.*, vol. 32, no. 6, pp. 28–34, Nov. 2018.

[28] C. Trabelsi *et al.*, "Deep complex networks," in *Proc. Int. Conf. Learn. Representations*, Vancouver, BC, Canada, 2018, pp. 1–19.

[29] S. Loffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.

[30] 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.

Ya Tu (Student Member, IEEE) received the B.S. degree from the College of Computer Science and Technology, Taiyuan University of Technology, Taiyuan, China, in 2016. He is working toward the Ph.D. degree in the College of Information and Communication Engineering, Harbin Engineering University, Harbin, China. His current research interests include signal processing, machine learning, and data analysis.

Yun Lin (Member, IEEE) received the B.S. degree from Dalian Maritime University in 2003, the M.S. degree from the Harbin Institute of Technology in 2005, and the Ph.D. degree from Harbin Engineering University in 2010. He was a Visiting Scholar at Wright State University, USA, from 2014 to 2015. He is currently an Associate Professor at the College of Information and Communication Engineering, Harbin Engineering University, Harbin China. His research interests include machine learning, information fusion, cognitive and software defined radio.

Changbo Hou received the B.S. and M.S. degrees from the College of Information and Communication Engineering, Harbin Engineering University, Heilongjiang, China, in 2008 and 2011, respectively. He is currently a Lecturer with the College of Information and Communication Engineering, Harbin Engineering University, where he is also a Doctor in the Key Laboratory of In-ber Integrated Optics, Ministry Education of China. His research interests include wideband signal processing, optical sensors, image processing and deep learning.

Shiwen Mao (Fellow, IEEE) received the Ph.D. degree in electrical and computer engineering from Polytechnic University, Brooklyn, N.Y. in 2004. He is the Samuel Ginn Professor and Director of Wireless Engineering Research and Education Center at Auburn University, Auburn, AL. His research interests include wireless networks, multimedia communications, and smart grid. He is a recipient of the IEEE ComSoc TC-CSR Distinguished Technical Achievement Award in 2019, the NSF CAREER Award in 2010, and the 2004 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems.