IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

IEEE RELIABILITY SOCIETY SECTION

Received April 5, 2022, accepted April 21, 2022, date of publication May 2, 2022, date of current version May 11, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3171856

The Performance Analysis of Complex-Valued Neural Network in Radio Signal Recognition

JIE XU[1,2], CHENGYU WU[3], SHUANGSHUANG YING[4], AND HUI LI[2]
[1]School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China
[2]No. 36 Research Institute of CETC, Jiaxing 314033, China
[3]School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China
[4]Taizhou Vocational & Technical College, Taizhou 318000, China

Corresponding author: Chengyu Wu (jerry916@zstu.edu.cn)

This work was supported by the Fundamental Research Funds of Zhejiang Sci-Tech University under Grant 2021Q029.

ABSTRACT Many techniques have been developed for wireless signal recognition in many fifth generation (5G) enabled derivatives. Many harsh constraints, such as the large amount of model parameters and complex signal characteristics, drives intelligent recognition method in real-world settings. In this paper, we propose a generalizable, practical method for raw IQ signal recognition. Specifically, deep complex-valued convolutional neural network models, including a Complex-valued Visual Geometry Group (VGG) (CvVGG) model and a Complex-valued ResNet (CvRN) model, are proposed for handling raw signal IQ data. We examine the merit of complex-valued neural networks (CvNN) and validate their performance with experiments using two public datasets. With an SNR of 10dB, the proposed algorithm achieves a recognition accuracy of 96% on the RadioML2018.10a dataset. When performing drone recognition, CvNN can achieve a recognition accuracy of 99%. Our experimental results verify that deep complex-valued neural network models can achieve considerably improved accuracy with lower computation complexity and fewer model parameters than their real-valued counterparts.

INDEX TERMS Signal recognition, modulation classification, convolutional neural network (CNN), complex-valued neural network (CvNN), fifth generation (5G).

I. INTRODUCTION

Over the years, various intelligent terminals and wearable sensing devices brings us more serves mode and connecting approaches, such as autonomous vehicles [1], privacy-preserving [2], [3], and blockchain [4], [5]. Identifica- tion of received signals has become a hot research topic in fifth generation (5G) communication, which is usually a challenging task executed at the receiver side before demodulation [6]. Generally, feature based (FB) methods are prominent in practical implementations because of the low complexity involved [7]–[12]. Usually, domain expertise is critical for such solutions. Recently, deep learning (DL) attracts great interest in the signal recognition field, which employs a hierarchical feature extraction approach. In other words, DL can reduce the pre-processing effort and efficiently distill useful information. Moreover, Graphics Processing Unit (GPU)-based parallel computing allows fast inference with DL. In addition, DL has the unique advantage of handling high dimensional feature spaces encountered in signal recognition problems, while the availability of public datasets facilities the research and wide deployment of DL [13]–[15]. Several recent works study the application of DL to wireless signal recognition. Specifically, DL models have been proposed to handle raw I/Q data of received signal [7] and transfer computer vision DL models to classify the statistical images constructed with the received signal [16], [17]. Y. Tu *et al.* [18] proposed large-scale real-world radio signal based on automatic Dependent Surveillance-Broadcast (ADS-B) and utilized several DL methods to give some results on common channel influence in communication system. For DL-aided approaches with raw I/Q data of signals as input, T. O'shea *et al.* [19] first introduce an open-access AMC dataset termed RadioML2016.10a, which comprises of eight digital-modulated signals and three analog-modulated signals. In [20], O'Shea *et al.* introduce several novel deep learning applications in the physical layer. They demonstrate a proof-of-concept study, where a convolutional neural network (CNN) is utilized for modulation classification and achieves satisfactory accuracy. Later in [7], the authors provide a more extensive dataset of additional radio signal types, a more realistic simulation of the wireless propagation environment, and new methods for signal classification which greatly outperform those they

The associate editor coordinating the review of this manuscript and approving it for publication was Zhaojun Steven Li.

initially introduced. M. Wang *et al.* [21] illustrated how to improve work efficiency while saving costs in the future DL-based scenarios. Y. Tu *et al.* [22] proposed activation-maximization based neural network pruning method, it will slim the neural network and make it much easier to be deployed in edge device. Some researchers [23], [24] also considered the security problems existed in DL-based communication system, and they used adversarial example to check the best way to deceive DL-based signal recognition. What is more, to overcome the constraint number of signal label, Y. Dong *et al.* [25] proposed SR2CNN to conduct zero-shot learning for signal type classification.

With regard to DL-aided statistical signal image recognition, deep convolutional neural networks (DCNN) is first applied to process the images constructed with channel state information phase difference data for indoor fingerprinting in 2017 [16], [17]. S. Peng *et al.* [26] exploit colored constellation diagrams to represent digital signals and utilize AlexNet to precisely recognize them. Y. Lin *et al.* [27] proposed contour stellar image (CSI) to bridge the gap between signal waveforms to DL data formats. CSI will extract the signal I/Q amplitude distribution and map it into different color. This will boost signal feature extraction performance and facility DL framework application in AMC, data augmentation, and transfer learning. S. Zhang *et al.* [28] utilized binary neural networks to reduce compute complexity, the experiment result showed the neural network only needs 5% and 50% run time to achieve the same accuracy in CSI. Wang *et al.* in [29] combine two CNNs trained with different datasets to achieve high AMC accuracy. In a recent work [30], the authors utilize generative adversarial networks (GAN), a semi-supervised learning approach, to conduct semi-supervised learning for AMC with incompletely labeled datasets. Wang *et al.* in [31] propose a fundamental privacy-preserving framework with differential privacy. They survey the adaptations and variants of differential privacy in emerging applications as well as the challenges to differential privacy. Zheng *et al.* in [32] propose adaptive hybrid communication protocols, including a novel position-prediction-based directional MAC protocol (PPMAC) and a self-learning routing protocol based on reinforcement learning (RLSRP). Tang *et al.* in [33] propose a smart approach to programmatic data augmentation by using the auxiliary classifier generative adversarial networks (ACGAN) to gain a better classification accuracy of communication signal modulation. Wang *et al.* in [34] also propose a novel AMC method. They firstly utilize Rényi entropy and singular entropy to obtain the modulation feature; present a novel basic probability assignment function (BPAF) based on the normal test theory; and utilize the Dempster-Shafer (DS) evidence theory to develop a classifier. Despite the relatively rich literature on application of DL in communications, there are only few works considering complex-valued representation of signal attributes in model training and inference.

Deep learning in the complex domain is challenging. Hirose *et al.* demonstrate that complex-valued networks are better than real-valued networks for denoising incoherent, or noise-corrupted, waveforms [35]–[37]. Nuaimi *et al.* propose an improvement to the MP3 codec using complex networks [38]. Zhang *et al.* in [39] propose a complex-valued CNN (CvCNN) specifically for synthetic aperture radar (SAR) image interpretation. It utilizes both amplitude and phase information of complex SAR imagery. However, the tools and algorithms for handling complex-valued signals are still lacking, or, are simply too scattered in the literature.

Complex-valued signals are encountered in a wide variety of applications, such as wireless communications, sensor array signal processing, as well as biomedical sciences and physics. Consequently, there is compelling need in science and engineering for a statistical and mathematical theory for processing complex-valued random signals. For example, most practical modulation schemes in communications are complex-valued. Many applications, such as radar and magnetic resonance imaging (MRI), generate data that are inherently complex-valued. In some scenes, two-dimensional real-valued data matrix is another way to present a complex vector and then conducting analysis in the complex domain (instead of the real domain). The complex-valued representation is also compact and simpler in terms of notation and for algebraic manipulations, and is convenient for computation operations. It is evident that the need for the expertise and theory in the processing, statistical modelling, and estimation of complex-valued multivariate signals and phenomena is rapidly increasing.

It is well known that CNNs actually learn discriminating features in computer vision applications. CNNs are expected to detect small-scale feature at shallow layers while complex features in deeper layers. In the wireless communications domain, however, CNNs are not trained to identify images but I/Q samples. Nowadays, the prevalent CNN framework [7] considers the particular transition signature in wireless signals.

Fig. 1 shows examples of transitions in the I/Q complex plane corresponding to QPSK, BPSK, and OQPSK. It can be seen that different modulation waveforms present different transition patterns in the I/Q plane. For example, the transitions between (1, 0) and (−1, 0) are unique to BPSK but do not appear in QPSK, which has a substantially different constellation. The transition patterns can constitute a unique signature of the modulated signal, which can eventually be learned by the CNN's filters.

Although the above CNN models are proved to be effective and useful [7], we still believe that they do not fully consider the inherent nature of the physical layer. Our motivation are twofold. First, the real and imaginary parts of wireless signals are statistically dependent on each other. For example, consider the circular rotation of a time-domain signal corresponds to a linear phase shift in the frequency domain. The real and imaginary parts of a complex number are dependent on each other under any change in phase. Unfortunately, the real-valued network model usually ignores the statistical correlation between the real and imaginary

parts. Second, a complex-valued model provides a more constrained system than real-numbered models. If we know in advance that both phase and amplitude are important to the learning objective, then it is sensible to employ a complex-valued model.

In this paper, we propose complex-valued neural network (CvNN) models for communication signal recognition, including a complex-valued Visual Geometry Group (VGG) model, termed CvVGG, and a deep complex-valued convolutional neural network model, termed Complex-valued ResNet (CvRN). We present their design, study their merit, and validate their performance with two public datasets.

The main contributions made in this paper are as follows.

- We examine how to incorporate CvNNs into smart communication systems. Furthermore, we derive useful insights and highlight the inherent merit of CvNN in handling raw I/Q signal waveforms.
- We propose new deep neural network architectures, i.e., the complex-valued VGG (CvVGG) and the complex-valued ResNet (CvRN) models for signal recognition problems. We present their designs and configurations, and explore their suitability for signal modulation classification and device fingerprint identification problems.
- We provide a thorough experimental study of the proposed CvNN models with respect to their classification accuracy, learning speed, computation complexity, and model parameters using two public datasets. We demonstrate the effectiveness of the proposed CvNN models under challenging and realistic scenarios, using raw, real-world I/Q waveform data.

The remainder of this paper is organized as follows. Section 2 presents the merit of CvNN for signal recognition. In Section 3, the building blocks and the architectures of the two proposed CvNN models are introduced. Subsequently, the performance of the proposed CvRN and CvVGG models is investigated using two public datasets in Section 4. Section 5 concludes this paper.

II. MERIT OF CVNN FOR SIGNAL RECOGNITION

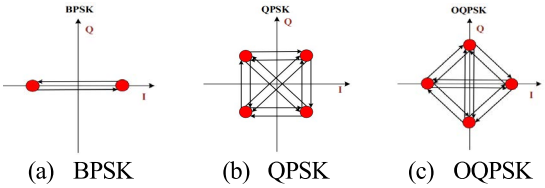
A. COMPLEXED-VALUED CONVOLUTIONAL KERNEL REGULARATION

A real convolution output can be interpreted as a heat map of similarity to the convolution kernel. That is, every output value of a convolutional layer is a dot product between a kernel and an input patch. Indeed, a dot product between a real patch and a kernel with L1-norm is maximized when they are equal to a scalar, given by:

$$\arg\max \{\mathbf{W}_r \cdot \mathbf{X}\} \tag{1}$$

where $\mathbf{W}_r$ denotes the convolutional kernel weight matrix and $\mathbf{X}$ denotes the input data. When it comes to complex-valued convolutional operation, we need to maximize the magnitude of the dot product between the complex-valued convolutional kernel weight matrix $\mathbf{W}_c$ and the input data $\mathbf{X}$, given by:

$$\arg\max |\mathbf{W}_c \cdot \mathbf{X}| \tag{2}$$



(a) BPSK (b) QPSK (c) OQPSK

FIGURE 1. Transitions in the I/Q complex plane corresponding to (a) BPSK, (b) QPSK, and (c) OQPSK, where the arrows indicate the symbol transition traces.

considering the representation of complex-valued amplitude and phase, we can rewrite (2) as:

$$\arg\max \sum_{mn} \dot{A}_{mn}\ddot{A}_{mn}e^{j(\alpha_{mn}+\beta_{mn})} \tag{3}$$

where $\mathrm{W}_{c,mn} = \ddot{\mathrm{A}}_{mn}e^{j\alpha_{mn}}$ and $\mathrm{X}_{mn} = \ddot{\mathrm{A}}_{mn}e^{j\beta_{mn}}$.

In (3), Z_{mn} is multiplied by $\dot{\mathrm{A}}_{mn}$ and rotated by β_{mn}. The sum of the multiplied and rotated vectors achieves its maximal magnitude if they all have the same phase and their magnitudes accumulate. Otherwise, the summed terms may cancel each other, resulting in a smaller magnitude. In other words, we claim that the complex-valued convolutional layer indeed operates as a regularization method.

B. SIGNAL COHERENCE

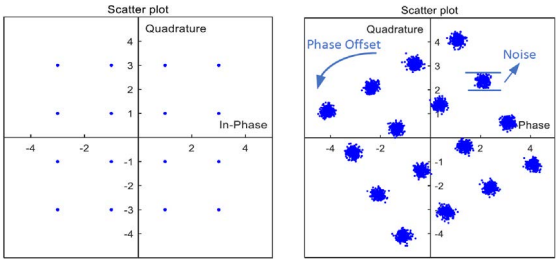
When we deal with I/Q raw waves, the real and imaginary axes are essentially less meaningful than amplitude and phase (or phase difference). This is because the real and imaginary axes are determined only relatively to an arbitrarily determined phase reference. The receiver determines the real and imaginary parts, which never exist beforehand [44]. Instead, the difference of two phase values is meaningful itself, which corresponds to the 1: time course and/or position difference. In this sense, the phase difference represents certain useful information. The amplitude, orthogonal to phase, is also meaningful since it signifies the energy or power of the waveform.

As an example, consider the 16-quadrature-amplitude modulation (16-QAM). Fig.2(a) shows an ideal signal constellation in the complex plane. When a receiver detects the signal, the constellation is affected by random noise, phase rotation/Doppler effect, and possible harmonic waves just like that shown in Fig. 2(b).

The outcome of the CvNNs operation contains values that carry features of both I/O parts, as a result of the complex operation. Let the complex kernel be A + jB and the complex input signal be X + jY. We can store the outcome as:

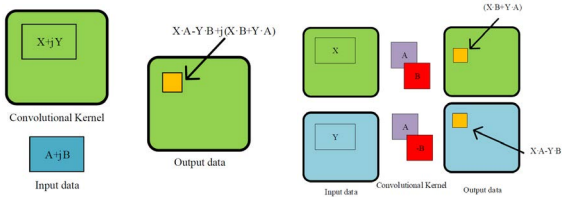
- Separated channel: $\mathbf{X} \cdot \mathbf{A} - \mathbf{Y} \cdot \mathbf{B}$
- Mixed channel: $\mathbf{X} \cdot \mathbf{B} + \mathbf{Y} \cdot \mathbf{A}$

Mathematically, the outcome of this operation is still real valued in one channel and imaginary in the other channel. This way, the mixed channel will allow CvNN to learn signal coherence information. Theoretically, CvNN outperforms real-valued Neural Network (RVNN) in high signal coherence regions (e.g., the high SNR region).



(a) The ideal constellation (b) The ideal constellation

FIGURE 2. Signal constellation in the complex plane of 16-QAM digital communications: (a) The idea constellation; (b) The received constellation that is affected by noise, phase rotation, and saturation.



(a) Complex-valued convolution (b) Equivalent real-valued convolution

FIGURE 3. Equivalent complex and real convolution layers: (a) Complex-valued convolution, where the output pixel is given by the dot product of the input patch and the kernel; (b) Equivalent real-valued convolution with two-channel inputs, kernels, and outputs.

III. DEEP COMPLEX-VALUED NEURAL NETWORK ARCHITECTURE

In this section, we will first discussion the complex-valued building blocks of CvNNs, including complex-valued convolutional layer, complex-valued batch normalization layer complex-valued activation, and complex-valued dense layer. We will then present the architectures of the proposed complex-valued VGG (CvVGG) and complex-valued ResNet (CvRN) models.

A. COMPLEXED-VALUED CONVOLUTIONAL LAYER

Based on the definition of complex number as an ordered pair of real numbers, we represent a complex number z as:

$$z \equiv (m, n) \tag{4}$$

where m and n are real numbers representing the real and imaginary parts of z. The addition and multiplication operations of two complex numbers can be respectively defined as:

$$(m_1, n_1) + (m_2, n_2) \equiv (m_1 + m_2, n_1 + n_2) \tag{5}$$
$$(m_1, n_1) \cdot (m_2, n_2) \equiv (m_1 m_2 - n_1 n_2, m_1 n_2 + n_1 m_2) \tag{6}$$

In addition, for 2D real values, the addition and multiplication operations can be respectively represented as:

$$(m_1, n_1) + (m_2, n_2) \equiv (m_1 + m_2, n_1 + n_2) \tag{7}$$
$$(m_1, n_1) \cdot (m_2, n_2) \equiv (m_1 m_2, n_1 n_2) \tag{8}$$

Nowadays, the fact that a complex number is defined by two real numbers may lead present-day neural-network researchers to consider a complex neural network equivalent, other than just a double-dimension, real-valued network. However, as can be seen from (5) to (8), although the addition processes are identical, the multiplication of two complex numbers is unique, involving both angle rotation and amplitude amplification. This feature is the result of the mixture of the real and imaginary components, making it more challenging to design a CvNN.

In the complex generalization, both the kernel and input are complex-valued. The only difference stems from the multiplication of complex numbers. When convoluting a complex matrix with a complex kernel $\mathbf{A} + \mathrm{j}\mathbf{B}$, the output corresponding to the input patch $\mathbf{X} + \mathrm{j}\mathbf{Y}$ is given by:

$$(\mathbf{X} \cdot \mathbf{A} - \mathbf{Y} \cdot \mathbf{B}) + j(\mathbf{X} \cdot \mathbf{B} + \mathbf{Y} \cdot \mathbf{A}) \tag{9}$$

To implement the same functionality with a real valued convolution, the input and output should be equivalent. Each complex matrix is represented by two real matrices, stacked together in a three dimensional array. Again use array $[\mathbf{A}, \mathbf{B}]$ to represent the convolutional kernel and let the input data be $[\mathbf{X}, \mathbf{Y}]$. For traditional two real-valued channel kernel, the dot product between the kernel and input data is:

$$\mathbf{X} \cdot \mathbf{A} + \mathbf{Y} \cdot \mathbf{B} \tag{10}$$

which is not the desired complex-valued output (which should be (9)). To obtain the desired outcome, we convolute with multiple kernels through multiple channels. That is, we use an equivalent real convolution scheme that has two kernels in the forms of $[\mathbf{A}, -\mathbf{B}]$ and $[\mathbf{B}, \mathbf{A}]$. Such a two-kernel approach, as illustrated in Fig. 3, can produce the desired output given in (9).

In summary, a convolution layer in a complex-valued network can be implemented in a restricted form of a real valued convolution laver with twice as many kernels.

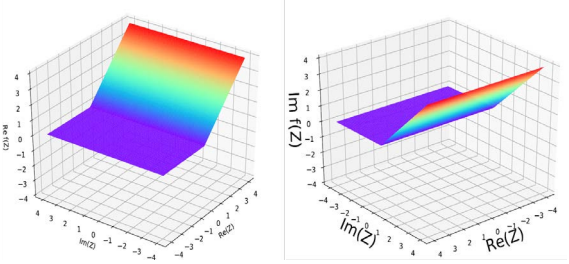
B. COMPLEXED-VALUED BATCH NORMALIZATION LAYER

Data normalization is commonly used for efficient training in most types of deep networks. However, when the underlying problem lies in the complex domain, the data requires special handling as outlined in [35]. Briefly, one cannot perform two-way in dependent normalization in the real and imaginary parts of a complex number, as there is also information precisely in the relation of the real and imaginary axes. To properly handle the normalization process, we treat the complex numbers as 2D vectors and the process as 2D whitening. We scale the 0 centered data x by the inverse square root of its covariance matrix V, as:

$$\tilde{\mathbf{X}} = (\mathbf{V})^{\frac{1}{2}}(\mathbf{x} - \mathbb{E}[\mathbf{x}]) \tag{11}$$

where V is a 2×2 covariance matrix given by:

$$\mathbf{V} = \begin{pmatrix} V_{rr} & V_{ri} \\ V_{ir} & V_{ii} \end{pmatrix} = \begin{pmatrix} \mathrm{cov}(\mathrm{Re}\{\mathbf{x}\}, \mathrm{Re}\{\mathbf{x}\}) & \mathrm{cov}(\mathrm{Re}\{\mathbf{x}\}, Im\{\mathbf{x}\}) \\ \mathrm{cov}(Im\{\mathbf{x}\}, \mathrm{Re}\{\mathbf{x}\}) & \mathrm{cov}(Im\{\mathbf{x}\}, Im\{\mathbf{x}\}) \end{pmatrix} \tag{12}$$



(a) Real part (b) Imaginary part

FIGURE 4. Surface plots for (a) the real and (b) imaginary parts of $\mathbb{C}$Re*lu*.

As in batch normalization for real values, we also set the learnable shift parameter p and the scaling parameter y for complex batch normalization. The scaling parameter is given by:

$$\gamma = \begin{pmatrix} \gamma_{rr} & \gamma_{ri} \\ \gamma_{ri} & \gamma_{ii} \end{pmatrix} \tag{13}$$

The real part and imaginary part of the shift parameter $\vec{\beta}$ and γ_{ri}; in the scaling parameter γ will be initialized to zero, while γ_{rr} and γ_{ii} in the scaling parameter γ will be initialized to $\frac{1}{\sqrt{2}}$. The initialization will satisfy a modulus of 1 for the variance of the normalized value [45]. Finally, we obtain the complex-valued batch normalization laver as:

$$BN(\tilde{\mathrm{x}}) = \gamma \cdot \tilde{\mathrm{x}} + \vec{\beta} \tag{14}$$

C. COMPLEXED-VALUED ACTIVATION

The rectified linear unit (ReLU) has become very popular for DNNs, since it avoids vanishing gradients usually associated with the sigmoidal activation. In this paper, we propose Complex ReLU (or, $\mathbb{C}$ ReLU), the complex-valued activation that applies separate ReLUs on both the real and imaginary parts of a neuron, which is given by:

$$\mathbb{C}Relu = Relu(\mathrm{Re}(z)) + Relu(jIm(z)) \tag{15}$$

The surface plots for the real and imaginary parts of CRelu are presented in Fig. 4.

D. COMPLEXED-VALUED DENSE LAYER

The dense layer often serves as a classifier in BNN. To make full use of the complex-valued statistical information, we also present a mechanism, named Complex Dense Layer, to learn complex-valued features while computing complex-valued classification results. Denote a complex-valued dense vector weight as $\vec{w} = \vec{a} + j\vec{b}$ and a complex-valued input as $\vec{s} = \vec{x} + j\vec{y}$.

Similar to the complex-valued convolutional operation presented in Section III part A, we have:

$$\vec{w} \cdot \vec{s} = (\vec{a} \cdot \vec{x} - \vec{b} \cdot \vec{y}) + i(\vec{b} \cdot \vec{x} + \vec{a} \cdot \vec{y}) \tag{16}$$

This process is illustrated in Fig.5.

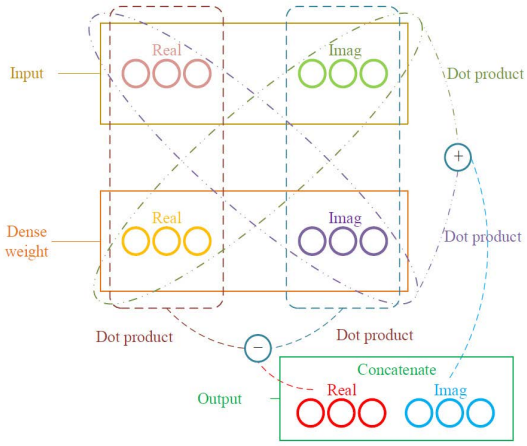


FIGURE 5. Complex-valued dense layer computation operation.

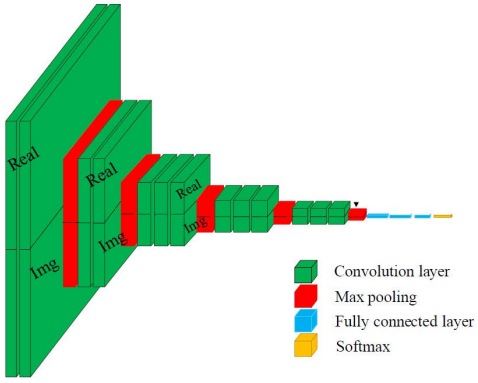


FIGURE 6. The proposed CvVGG architecture.

E. COMPLEXED-VALUED VGG ARCHITECTURE

The VGG model [40] is based on AlexNet [41] and has several unique features. Instead of using large receptive fields like AlexNet (11 × 11 with a stride of 4), VGG uses very small receptive fields (3 × 3 with a stride of 1). Because there are now three ReLU units instead of just one, the decision function is more discriminative. There are also fewer parameters (27 times the number of channels, while AlexNet has 49 times the number of channels). VGG incorporates 1x convolutional lavers to make the decision function more non-linear without changing the receptive fields. The small-size convolution filters allow VGG to have a large number of weight layers; and in most cases, more layers lead to improved performance. This is not an uncommon feature, though. GoogleNet from Google Research, another model that uses deep CNNs and small convolution filters, also performed well in the 2014 ImageNet competition.

The main challenge is how to make VGG accept complex-valued signal data format.

In this paper, we refer to [7] and design a VGG architecture for modulated signal I/O raw waveforms and replace the 2D Convolutional laver with the 1D Convolutional laver.

TABLE 1. The CvVGG network layout.

Layer	Output dimension
Input	1024×2
Complex-valued Convolution	1024×64
MaxPooling	512×64
Complex-valued Convolution	512×64
MaxPooling	256×64
Complex-valued Convolution	256×64
MaxPooling	128×64
Complex-valued Convolution	128×64
MaxPooling	64×64
Complex-valued Convolution	64×64
MaxPooling	32×64
Complex-valued Convolution	32×64
MaxPooling	16×64
Complex-valued Convolution	16×64
MaxPooling	8×64
FC/CRelu	128
FC/CRelu	128
FC/SoftMax	24

We do not perform any expert feature extraction or other pre-processing on the raw radio signal.

Instead, we allow the network to learn raw time series features directly from the high dimension data. The architecture of the proposed CvVGG is given in Fig. 6. The network layout parameters are given in Table 1.

F. COMPLEXED-VALUED VGG ARCHITECTURE

Deeper neural networks are usually more difficult to train. Deep Residual Network (ResNet) [42] is arguably the most groundbreaking work in the computer vision/DD community in the last few years. ResNet makes it possible to train up to hundreds or even thousands of layers while achieving a compelling performance. Taking advantage of its powerful representational ability, the performance of many applications other than image classification have been boosted, such as object detection, face recognition, and Wi-Fi fingerprinting [43].

The main challenges for creating a complex-valued ResNet model are twofold. First, how to make CvRN RN accept complex-valued signal data. Second, how to design the residual stack which can extract complex-valued signal features while guaranteeing projection shortcut match the dimension.

In this paper, we design a CvRN architecture for handling I/Q raw waveforms of complex-valued wireless signal. We replace the 2D Convolutional layer in the residual block with the 1D Convolutional layer, and replace the real-valued convolutional laver and the BN layer with their complex-valued counterparts, respectively.

We also utilize the 1×1 convolutional layer to match the dimension. A sketch of the proposed CvRN architecture is given in Fig. 7. The detailed CvRN network layout is given in Table 2 and the residual stack architecture is given in Fig. 8.

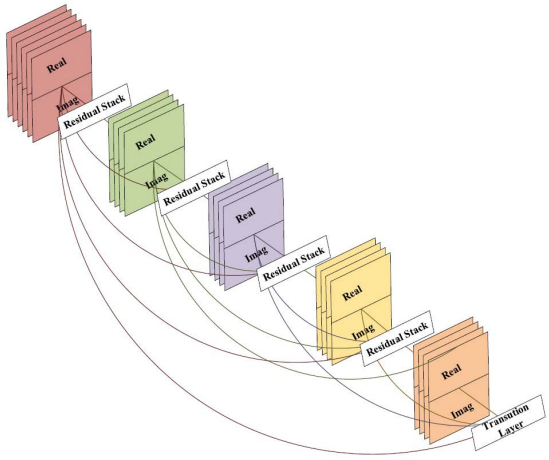


FIGURE 7. The proposed CvRN architecture.

TABLE 2. The CvRN network layout.

Layer	Output dimension
Input	1024×2
Residual Stack	512×32
Residual Stack	256×32
Residual Stack	128×32
Residual Stack	64×32
Residual Stack	32×32
Residual Stack	32×16
FC/CRelu	128
FC/CRelu	128
FC/SoftMax	24

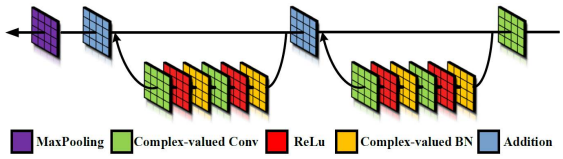


FIGURE 8. The residual stack architecture in CvRN.

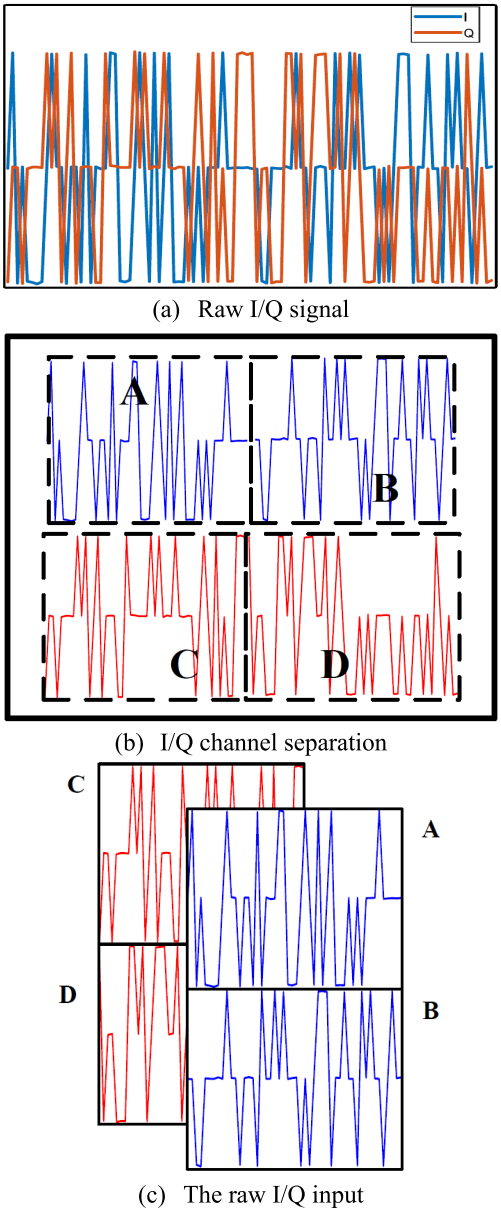
IV. EXPERIMENTAL

In this section, we will investigate the performance of the proposed CvVGG and CvRN models for wireless signal recognition, by comparing their classification accuracy, learning speed, computational complexity, and amount of parameters with their real-valued counterparts. We choose two public datasets in our study, including the RadioML2018.10a public dataset [19] and the real-world, over-the-air drone radio fingerprint dataset [46]. The experiment results validate that CvNN is more suitable for raw I/Q waveform recognition than their real-valued counterparts.

A. DATASETS

1) RADIOML2018.10A

The RadioML2018.10a dataset contains 24 types of modulations, including several high-order modulations (QAM256 and APSK256) [19]. Generally speaking, they are often used



(a) Raw I/Q signal

(b) I/Q channel separation

(c) The raw I/Q input

FIGURE 9. Relatively point destiny colorbar.

in low fading channels and high SNR environments, such as impulsive satellite links (e.g., DVB-S2X) [47]. The dataset was generated with many transmission impairments, such as carrier frequency offset (CFO), symbol rate offset (SRO), delay spread, and thermal no This dataset only takes into account the observation in relatively short time windows. The number of samples is 1,024. When the decision-making process does not have enough time to wait for more data to improve certainty, such short time classification would be painful but inevitable. This is particularly common in objective real-world systems, such as those in the environment where short bursts of signals occur or where observations are processed over time. One would not expect a classification rate close to 100% with this dataset when the signal-to-noise ratio (SNR) is low (i.e., Es/No is from −20dB to +30dB),

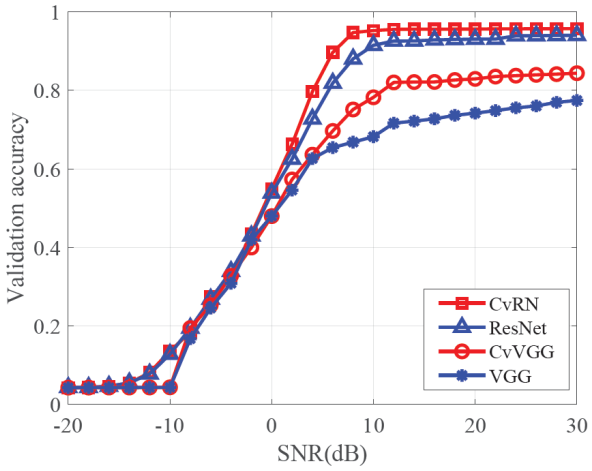


FIGURE 10. Comparison of the validation accuracy of the four models under various SNR level.

which makes it a good benchmark for us to study the proposed CvNNs. In this experiment, 70% of the ww dataset is used for training and the remaining 30% is used for validation and testing.

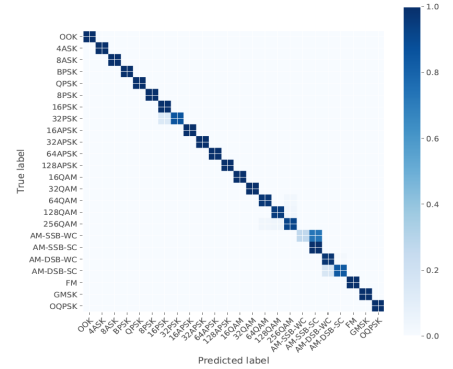
2) REAL-WORLD DRONE FINGERPRINT DATASET

Based on the system model, a drone dataset is composed of the complex-valued (radio fingerprint) RF data collected from real drones [46]. These real drone signals are provided by a large drone database. This database has collected many valuable real drone RF data through the following three modules: drones under analysis module, fight control module, and RF sensing module. Furthermore, we used three types of drone signals and noise from this database. More important, three types of drone activities were carried out by three different brands of drones, where different brands have different prices, protocols, and technologies. After IQ sampling on the RF data of each drone activity, we can obtain the drone signal dataset.

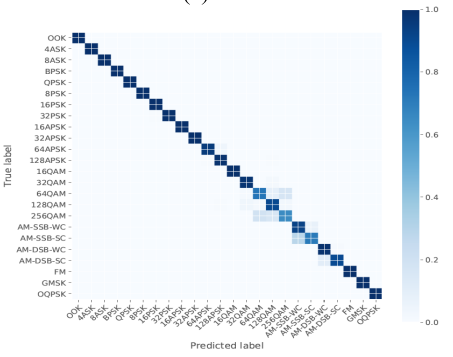
A total of 4,400 drone signal samples are used in our experiment, with an average of 1,100 signal samples per drone activity. These samples are randomly split with a proportion of 7:3 for the training and testing of the CvNNs. The length of each sample is 1,024. In order to better extract the features in the dataset, we split each sample into in-phase and quadrature components, in the form of a real-valued matrix of dimension $1,024 \times 2$. This dataset will be used as input to the CvNN drone identification system for identification of different drone signals. Fig. 9 shows how to construct an input tensor from I/Q waveforms.

B. HARDWARE AND MODEL CONFIGURATIONS

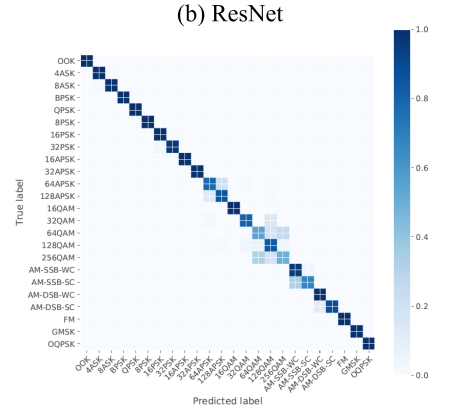
Experiments are performed on a SUGON sever equipped with two Intel Xeon Bronze 3104 CPUs (6 cores, each at 1.7GHz) with 256GB of DDR4 RAM and two Nvidia TitanXP GPUs, each with 12GB of VRAM. Each model we tested occupies a maximum of 8GB space on the GPU. The software we use



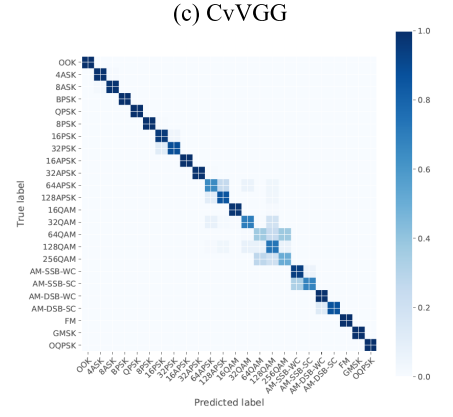
(a) CvRN



(b) ResNet



(c) CvVGG



(d) VGG

FIGURE 11. 24-modulation confusion matrices obtained with the RadioML2018.10a test set for the four schemes when the SNR is 10dB.

is Keras 2.2.4 with TensorFlow 1.12.0 and Python 3.7.5 as y backend.

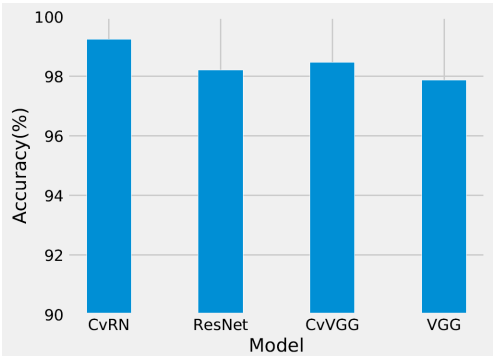


FIGURE 12. Comparison of the validation accuracy of the four models.

For both RadioML2018.10a and Drone RF datasets, an epoch is the number of iterations in which the total number of samples chosen is equal to the size of the training set. The networks are trained over 500 epochs with a batch size of 512. We find that adaptive learning rate annealing is helpful, and thus reduce the learning rate by 0.1 after every10 epochs, where the model fails to improve its validation accuracy. We also allow early stopping once the validation accuracy does not improve over 30 epochs. The results we report in this section represent the best generalization accuracy over 500 epochs or if the early stopping criterion is met. This is a standard practice for obtaining unbiased error estimations when facing significant computational requirements for training such networks and exploring different architectures, configurations, and datasets. We evaluate both SGD and the Adam optimizer, while Adam is usually superior than SGD. Real-valued layers are initialized using the uniform distribution [48], while complex-valued layers leverage complex weight initialization.

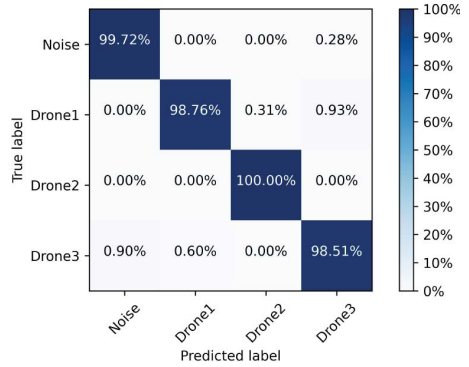
C. CLASSIFICATION ACCURACY

After training the models, we obtain the classification accuracy results using the RadioML2018.10a and drone RF datasets. The results are presented in Figs. 10, 11, 12 and13.

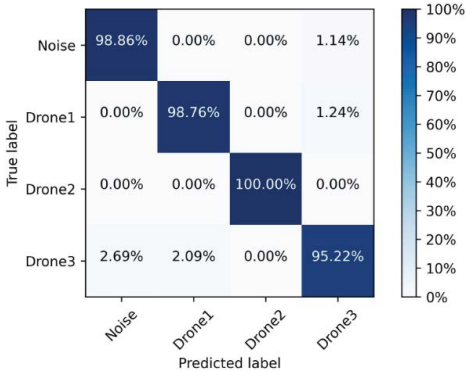
1) THE RADIOML 2018.10A DATASET

We test CVRN, ResNet, CVVGG, and VGG at every SNR level and the validation RN v GO accuracy results are plotted in Fig.10. For better visualization of the results, we also present the confusion matrix results for the four models when SNR is 10dB in Fig. 11. From Figs. 10 and 11, we have the following observations.

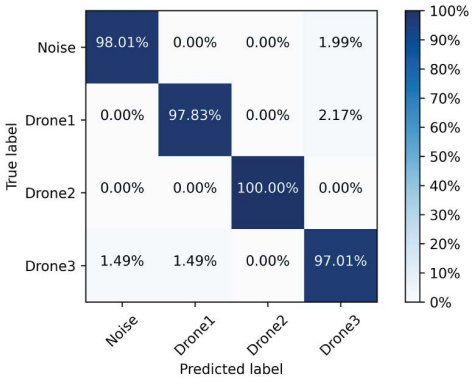
- When the SNR is less than 0dB (i.e., when the SNR is medium or low), both the CvRN and CvVGG curves are close to the corresponding real-valued curves, respectively. This is because the I/Q coherence information in the signal has been severely polluted when the SNR is not high.
- When the SNR is higher than 0dB (i.e., when the highly coherent region), itis evident that both CVRN and CVVGG G outperform their real-valued counterparts, respectively. This is because CvNNs can better 1earn the



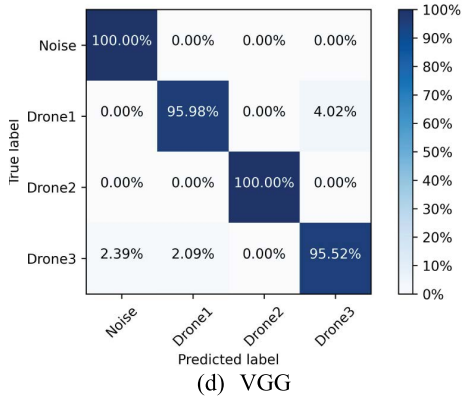
(a) CvRN



(b) ResNet



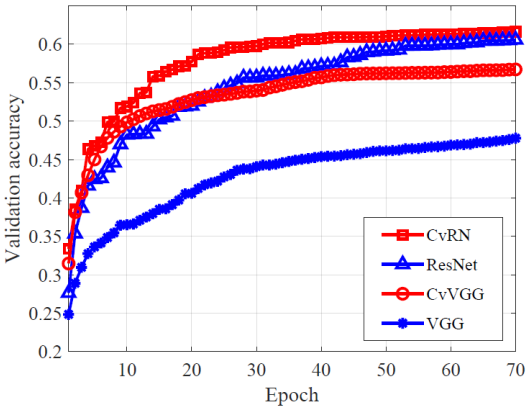
(c) CvVGG



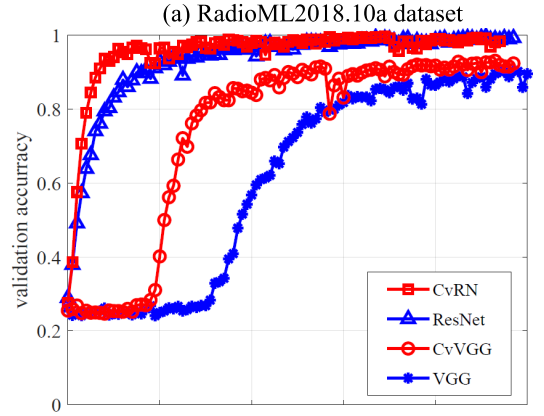
(d) VGG

FIGURE 13. The confusion matrices obtained with the Drone RF dataset by (a) CvRN;(b) ResNet; (c) CvVGG; and (d) VGG.

l/Q coherence information in the signal, which is less polluted.



(a) RadioML2018.10a dataset



(b) Drone RF dataset

FIGURE 14. The learning curves for (a) the RadioML2018.10a dataset and (b) the drone RF dataset.

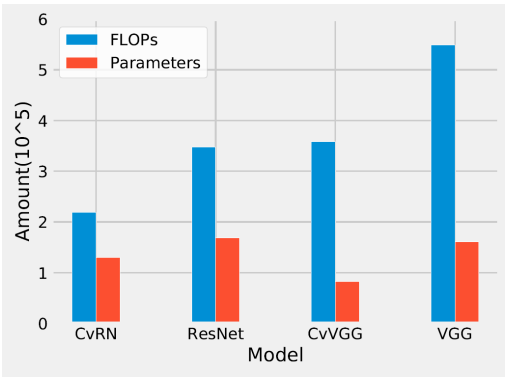


FIGURE 15. Comparison of model FLOPs and parameter for CvRN, ResNet, CvVGG and VGG.

2) THE DRONE RF DATASET

We also investigate the CyNN performance with the drone RF dataset. The average classification accuracy results obtained by the four models are presented in Fig.12. To provide more insight of the identification result, the confusion matrices of the four schemes are presented in Fig. 13. From the comparisons in Figs. 12 and 13, we can conc-lude that the two CvNN models performs well in identifying RF I/Q signals, and outperform their real-valued counterparts, respectively.

D. LEARNING SPEED

In this experiment, we use learning curves to measure the learning speed of the four schemes. The results obtained with the RadioML2018.10a dataset and the drone RF dataset are presented in Fig. 14. We can see that both CvNN models learn faster than their real-valued counterparts, respectively. We conjecture that the reason are as follows.

- Signal coherence: Due to the signal coherence information in the raw 1/Q data, CvNN ww can learn faster since it can rely on more data sources than real-valued models.
- Degree of freedom: As discussed in Section IV, CvNN has fewer degrees of freedom since its weights can be represented by amplitude and phase. When the degree of freedom is reduced, the arbitrariness of the solution will also be reduced. Thus the CvNNS can learn faster than their real-valued counterparts, respectively.

E. COMPUTATIONAL COMPLEXITY AND NUMBER PARAMETERS

Floating point operations per second (FLOPS) is used as an approximation so as to calculate the number of operations using the model [49]. Also, the number of parameters is a metric of capacity or the ability to approximate functions. When there are too many parameters, the neural network tends to overfit the data. On the other hand, with too few parameters, the neural network tends to underfit the data.

We conduct an experiment for CvRN, ResNet, CvVGG, and VGG to compare these models' FLOPs and number of parameters. The experimental results, obtained from the TensorFlow API, are presented in Fig. 15. Comparing the computation complexity and parameters, we can conclude that CvRN is a better choice than ResNet, and similarly, CvVGG performs better than VGG. This is because they have fewer parameters and therefore require less FLOPs and storage space than their real-valued counterparts, respectively.

V. CONCLUSION

In this paper, we demonstrated the effectiveness of the proposed CvNNs, which are powerful models for wireless signal recognition. Specifically, we presented a system that can identify signal coherence information inherent in each modulated signal's raw I/Q waveforms. One of our main motivations was to explore the feasibility of building a DL-aided system that is able to rapidly, robustly, and in real-time recognize signal modulation from raw signal I/Q waveforms. The proposed models are amenable for life-long learning and defensing against adversarial attacks due to their fast learning speed and high confidence about the classification result. Our experimental study using two public datasets showed that such a system could be readily deployed and operate within realistic environments, without constraints on any prior knowledge of the underlying protocol or software implementations.

Our experimental study also quantified the effect of different factors across multiple dimensions of the testing and training processes. While our approach is robust and effective under realistic conditions, we believe that our work is an important exploratory step within a vast and challenging new space, such as transfer learning and continual 1ifelong learning with several interesting future directions identified. In addition, combining traditional feature engineering, knowledge graphs and deep learning for multimodal fusion learning is also a potential idea.

REFERENCES

[1] J. Xiong, R. Bi, Y. Tian, X. Liu, and D. Wu, "Toward lightweight, privacy-preserving cooperative object classification for connected autonomous vehicles," *IEEE Internet Things J.*, vol. 9, no. 4, pp. 2787–2801, Feb. 2022.

[2] J. Xiong, R. Bi, M. Zhao, J. Guo, and Q. Yang, "Edge-assisted privacy-preserving raw data sharing framework for connected autonomous vehicles," *IEEE Wireless Commun.*, vol. 27, no. 3, pp. 24–30, Jun. 2020.

[3] S. Harinee and A. Mahendran, "Secure ECG signal transimission for smart healthcare," *Int. J. Performability Eng.*, vol. 17, no. 8, pp. 711–721, Aug. 2021.

[4] Y. Tian, T. Li, J. Xiong, M. Z. A. Bhuiyan, J. Ma, and C. Peng, "A blockchain-based machine learning framework for edge services in IIoT," *IEEE Trans. Ind. Informat.*, vol. 18, no. 3, pp. 1918–1929, Mar. 2022.

[5] D. Li, W. E. Wong, M. Chau, S. Pan, and L. S. Koh, "A survey ofNFC mobile payment: Challenges and solutions using blockchain and cryptocurrencies," in *Proc. 7th Int. Conf. Dependable Syst. Appl. (DSA)*, Nov. 2020, pp. 69–77.

[6] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 416–429, Mar. 2000.

[7] 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.

[8] Y. Lin, X. Zhu, Z. Zheng, Z. Dou, and R. Zhou, "The individual identification method of wireless device based on dimensionality reduction and machine learning," *J. Supercomput.*, vol. 75, no. 6, pp. 3010–3027, Jun. 2019.

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

[10] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features Off-the-shelf: An astounding baseline for recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2014.

[11] R. Wu, "Abnormal information identification and elimination in cognitive networks," *Int. J. Performability Eng.*, vol. 14, no. 10, p. 2271, 2018.

[12] H. Wang, J. Li, L. Guo, Z. Dou, Y. Lin, and R. Zhou, "Fractal complexity-based feature extraction algorithm of communication signals," *Fractals*, vol. 25, no. 4, Aug. 2017, Art. no. 1740008.

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

[14] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3072–3108, 2019.

[15] C. Hou, X. Zhang, and X. Chen, "Electromagnetic signal feature fusion and recognition based on multi-modal deep learning," *Int. J. Performability Eng.*, vol. 16, no. 6, pp. 941–949, Jun. 2020.

[16] X. Wang, X. Wang, and S. Mao, "CiFi: Deep convolutional neural networks for indoor localization with 5 GHz Wi-Fi," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[17] X. Wang, X. Wang, and S. Mao, "Deep convolutional neural networks for indoor localization with CSI images," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 1, pp. 316–327, Jan. 2020.

[18] Y. Tu, Y. Lin, H. Zha, J. Zhang, Y. Wang, G. Gui, and S. Mao, "Large-scale real-world radio signal recognition with deep learning," *Chin. J. Aeronaut.*, pp. 1–14, Oct. 2021, doi: 10.1016/j.cja.2021.08.016.

[19] T. J. O'shea and N. West, "Radio machine learning dataset generation with gnu radio," in *Proc. GNU Radio Conf.*, vol. 1, no. 1, pp. 1-6, 2016.

[20] 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.

[21] M. Wang, Y. Lin, Q. Tian, and G. Si, "Transfer learning promotes 6G wireless communications: Recent advances and future challenges," *IEEE Trans. Rel.*, vol. 70, no. 2, pp. 790–807, Jun. 2021.
[22] Y. Lin, Y. Tu, and Z. Dou, "An improved neural network pruning technology for automatic modulation classification in edge devices," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5703–5706, May 2020.
[23] Y. Lin, H. Zhao, X. Ma, Y. Tu, and M. Wang, "Adversarial attacks in modulation recognition with convolutional neural networks," *IEEE Trans. Rel.*, vol. 70, no. 1, pp. 389–401, Mar. 2021.
[24] Z. Bao, Y. Lin, S. Zhang, Z. Li, and S. Mao, "Threat of adversarial attacks on DL-based IoT device identification," *IEEE Internet Things J.*, early access, Oct. 14, 2021, doi: 10.1109/JIOT.2021.3120197.
[25] Y. Dong, X. Jiang, H. Zhou, Y. Lin, and Q. Shi, "SR2CNN: Zero-shot learning for signal recognition," *IEEE Trans. Signal Process.*, vol. 69, pp. 2316–2329, 2021.
[26] S. Peng, H. Jiang, H. Wang, H. Alwageed, Y. Zhou, M. M. Sebdani, 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.
[27] Y. Lin, Y. Tu, Z. Dou, L. Chen, and S. Mao, "Contour stella image and deep learning for signal recognition in the physical layer," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 1, pp. 34–46, Mar. 2021.
[28] S. Zhang, L. Yun, T. Ya, and S. Mao, "Electromagnetic signal modulation recognition technology based on lightweight deep neural network," *J. Commun.*, vol. 41, no. 11, p. 12, 2020.
[29] 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.
[30] 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, 2018.
[31] T. Wang, Z. Zheng, M. H. Rehmani, S. Yao, and Z. Huo, "Privacy preservation in big data from the communication perspective—A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 753–778, Aug. 2018.
[32] Z. Zheng, A. K. Sangaiah, and T. Wang, "Adaptive communication protocols in flying ad hoc network," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 136–142, Jan. 2018.
[33] 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.
[34] H. Wang, L. Guo, Z. Dou, and Y. Lin, "A new method of cognitive signal recognition based on hybrid information entropy and D-S evidence theory," *Mobile Netw. Appl.*, vol. 23, no. 4, pp. 677–685, Aug. 2018.
[35] A. Hirose and S. Yoshida, "Generalization characteristics of complex-valued feedforward neural networks in relation to signal coherence," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 4, pp. 541–551, Apr. 2012.
[36] A. Hirose, "Complex-valued neural networks: The merits and their origins," in *Proc. Int. Joint Conf. Neural Netw.*, Jun. 2009, pp. 1237–1244.
[37] Y. Tu, Y. Lin, C. Hou, and S. Mao, "Complex-valued networks for automatic modulation classification," *IEEE Trans. Veh. Technol.*, vol. 69, no. 9, pp. 10085–10089, Sep. 2020.
[38] A. Y. H. Al-Nuaimi, M. F. Amin, and K. Murase, "Enhancing MP3 encoding by utilizing a predictive complex-valued neural network," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jun. 2012, pp. 1–6.
[39] Z. Zhang, H. Wang, F. Xu, and Y.-Q. Jin, "Complex-valued convolutional neural network and its application in polarimetric SAR image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 7177–7188, Dec. 2017.
[40] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[41] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1–9.
[42] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[43] X. Wang, X. Wang, and S. Mao, "ResLoc: Deep residual sharing learning for indoor localization with CSI tensors," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–6.
[44] E. Ollila, V. Koivunen, and H. V. Poor, "Complex-valued signal processing—Essential models, tools and statistics," in *Proc. Inf. Theory Appl. Workshop*, Feb. 2011, pp. 1–10.
[45] H.-S. Choi, J.-H. Kim, J. Huh, A. Kim, J.-W. Ha, and K. Lee, "Phase-aware speech enhancement with deep complex U-Net," in *Proc. Int. Conf. Learn. Represent.*, 2018, pp. 1–20.
[46] M. F. Al-Sa'd, A. Al-Ali, A. Mohamed, T. Khattab, and A. Erbad, "RF-based drone detection and identification using deep learning approaches: An initiative towards a large open source drone database," *Future Gener. Comput. Syst.*, vol. 100, pp. 86–97, Nov. 2019.
[47] S. Cioni, G. Colavolpe, V. Mignone, A. Modenini, A. Morello, M. Ricciulli, A. Ugolini, and Y. Zanettini, "Transmission parameters optimization and receiver architectures for DVB-S2X systems," *Int. J. Satell. Commun. Netw.*, vol. 34, no. 3, pp. 337–350, May 2016.
[48] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," *Proc. 30th Int. Conf. Artif. Intell. Statist.*, 2010, pp. 249–256.
[49] J.-H. Luo, H. Zhang, H.-Y. Zhou, C.-W. Xie, J. Wu, and W. Lin, "ThiNet: Pruning CNN filters for a thinner net," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 10, pp. 2525–2538, Oct. 2019.

JIE XU received the B.E. and M.E. degrees from the China University of Mining and Technology, Xuzhou, China, in 2010, where he is currently pursuing the Ph.D. degree in information and communication engineering with the School of Information and Control Engineering. Since 2013, he has been with the No. 36 Research Institute of CETC, Jiaxing, China, where he is also a Senior Engineer. His current research interests include specific emitter identification and digital signal processing.

CHENGYU WU received the B.E. degree from the Zhejiang University of Science and Technology, Hangzhou, China, in 2005, the M.E. degree from the North University of China, Taiyuan, China, in 2008, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2016. He is currently a Lecturer with the School of Information Science and Technology, Zhejiang Sci-Tech University. He has published more than ten academic papers in important journals and conferences, including seven papers in the journals collected by SCI. His current research interests include radio resource management, digital signal processing, and machine learning.

SHUANGSHUANG YING received the B.E. degree from Zhejiang Normal University, Jinhua, China, in 2009, and the M.E. degree from the China University of Mining and Technology, Xuzhou, China, in 2012. She is currently a Lecturer with the Taizhou Vocational & Technical College, Taizhou, China. Her current research interests include specific emitter identification, signal processing, and deep learning.

HUI LI received the B.E. and M.E. degrees from Harbin Engineering University, Harbin, China, in 2014. He is currently a Senior Engineer with the No. 36 Research Institute of CETC, Jiaxing, China. His current research interests include specific emitter identification and signal processing.