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

JIE XU$^{1,2}$, CHENGYU WU$^2$; SHUANGSHUANG YING$^4$, HUI LI$^6$

1School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, China
2School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou, China
3No.36 Research Institute of CETC, Jiaxing, China
4Taizhou Vocational and Technical College, Taizhou, China

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

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) (CxVGG) model and a Complex-valued ResNet (CxRN) model, are proposed for handling raw signal IQ data. We examine the merit of complex-valued neural networks (CxNN) 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, CxNN 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 (CxNN), 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]. Identification 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 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, which 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.

Deep learning in the complex domain is challenging. Hirose et al. demonstrate that complex-valued networks are better than real-valued networks for denoising, coherent, 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.

![Graph](image)

**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**

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 REGULARIZATION

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:

$$\arg \max \|W \cdot X\|$$

where \(W\) denotes the convolutional kernel weight matrix and \(X\) 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 \(W_c\) and the input data \(X\), given by:

$$\arg \max \|W_c \cdot X\|$$

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

$$\arg \max \sum_{n}A_{\alpha\beta} \bar{A}_{\alpha\beta} e^{i(\alpha \mu + \beta \nu)}$$

where \(W_{\mu,\nu} = \bar{A}_{\mu,\nu} e^{i\mu} \) and \(X_{\mu,\nu} = \bar{A}_{\mu,\nu} e^{i\nu} \).

In (3), \(Z_{\mu,\nu}\) is multiplied by \(\bar{A}_{\mu,\nu}\) and rotated by \(\beta_{\mu,\nu}\). 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:** \( X \cdot A - Y \cdot B \)
- **Mixed channel:** \( X \cdot B + Y \cdot A \)

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).

### 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:

\[
z = (m, n)
\]

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:

\[
\begin{align*}
(m_1, n_1) + (m_2, n_2) & \equiv (m_1 + m_2, n_1 + n_2) \\
(m_1, n_1) \cdot (m_2, n_2) & \equiv (m_1 m_2 - n_1 n_2, m_1 n_2 + n_1 m_2)
\end{align*}
\]

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

\[
\begin{align*}
(m_1, n_1) + (m_2, n_2) & \equiv (m_1 + m_2, n_1 + n_2) \\
(m_1, n_1) \cdot (m_2, n_2) & \equiv (m_1 m_2 - n_1 n_2, m_1 n_2 + n_1 m_2)
\end{align*}
\]

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 \( A + jB \), the output corresponding to the input patch \( X + jY \) is given by:

\[
(X \cdot A - Y \cdot B) + j(X \cdot B + Y \cdot A)
\]

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 \([A, B]\) to represent the convolutional kernel and let the input data be \([X, Y]\). For traditional two real-valued channel kernel, the dot product between the kernel and input data is:

\[
X \cdot A + Y \cdot B
\]

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 \([A, -B]\) and \([B, A]\). 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 layer 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 independent 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:

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

where $V$ is a $2 \times 2$ covariance matrix given by:

$$
V = \begin{pmatrix}
V_{rr} & V_{ri} \\
V_{ri} & V_{ii}
\end{pmatrix}
$$

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

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

The real part and imaginary part of the shift parameter $\beta$ and $\gamma$; in the scaling parameter $\gamma$ will be initialized to zero, while $\gamma_{rr}$ and $\gamma_{ii}$ in the scaling parameter $\gamma$ 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 layer as:

$$
BN(\mathbf{X}) = \gamma \cdot \mathbf{X} + \tilde{\beta}
$$

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:

$$
\mathbb{C}\text{Relu} = \text{Relu}(\text{Re}(z)) + \text{Relu}(j\text{Im}(z))
$$

The surface plots for the real and imaginary parts of $\mathbb{C}$ReLU 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 $w = \tilde{a} + j\tilde{b}$ and a complex-valued input as $x = \tilde{x} + j\tilde{y}$. Similar to the complex-valued convolutional operation presented in Section III part A, we have:

$$
\tilde{w} \cdot \mathbf{x} = (\tilde{a} \cdot \mathbf{x} - \tilde{b} \cdot \mathbf{y}) + j(\tilde{b} \cdot \mathbf{x} + \tilde{a} \cdot \mathbf{y})
$$

This process is illustrated in Fig. 5.

E. COMPLEXED-VALUED VGG ARCHITECTURE

The proposed CvVGG architecture is shown in Fig. 6.
The VGG model [40] is based on AlexNet [41] and has several unique features. Instead of using large receptive fields like AlexNet (11 x11 with a stride of 4), VGG uses very small receptive fields (3x3 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 layers 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.

| Table I | The CVVGG Network Layout |
|---------|----------------------------|
| Layer   | Output dimension           |
| Input   | 1024 x 2                   |
| Complex-valued Convolution | 1024 x 64                   |
| MaxPooling | 512 x 64                   |
| Complex-valued Convolution | 512 x 64                   |
| MaxPooling | 256 x 64                   |
| Complex-valued Convolution | 256 x 64                   |
| MaxPooling | 128 x 64                   |
| Complex-valued Convolution | 128 x 64                   |
| MaxPooling | 64 x 64                    |
| Complex-valued Convolution | 64 x 64                    |
| MaxPooling | 32 x 64                    |
| Complex-valued Convolution | 32 x 64                    |
| MaxPooling | 16 x 64                    |
| Complex-valued Convolution | 16 x 64                    |
| MaxPooling | 8 x 64                     |
| FC/CRelu  | 128                        |
| FC/CRelu  | 128                        |
| FC/SoftMax | 24                         |

In this paper, we refer to [7] and design a VGG architecture for modulated signal I/O raw waveforms and replace the 2D Convolutional layer with the 1D Convolutional layer. 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 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 layer and the BN layer with their complex-valued counterparts, respectively.

| Table II | The CVRN Network Layout |
|----------|--------------------------|
| Layer    | Output dimension         |
| Input    | 1024 x 2                 |
| Residual Stack | 512 x 32                |
| Residual Stack | 256 x 32                |
| Residual Stack | 128 x 32                |
| Residual Stack | 64 x 32                 |
| Residual Stack | 32 x 32                 |
| Residual Stack | 32 x 16                 |
| FC/CRelu  | 128                      |
| FC/CRelu  | 128                      |
| FC/SoftMax | 24                      |

We also utilize the 1x1 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 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 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 noise. 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), 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.

**FIGURE 9.** Relatively Point Destiny Colorbar

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

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 every 10 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 [4, 8], 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 and 13.

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 higher than 0dB (i.e., when the highly coherent region), it is evident that both CVRN and CVVGG outperform their real-valued counterparts, respectively. This is because CvNNs can better learn the I/Q coherence information in the signal, which is less polluted.

- 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.
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. 11.

From the comparisons in Figs. 12 and 13, we can conclude that the two CvNN models perform 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 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 lifelong learning and defending 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 lifelong 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] Xiong, Jinbo, et al. “Towards lightweight, privacy-preserving cooperative object classification for connected autonomous vehicles.” IEEE Internet of Things Journal (2021).

[2] Xiong, Jinbo, et al. “Edge-assisted privacy-preserving raw data sharing framework for connected autonomous vehicles.” IEEE Wireless Communications 27.3 (2020): 24-30.

[3] H.S and A. Mahendran, “Secure ECG Signal Transmission for Smart Healthcare,” International Journal of Performability Engineering, vol. 17, no. 8, pp. 711-721, Aug. 2021.

[4] Tian, Youliang, et al. “A blockchain-based machine learning framework for edge services in IoIOT.” IEEE Transactions on Industrial Informatics 18.3 (2021): 1918-1929.

[5] D. Li, W.E. Wong, M. Chau, S. Pan, and L.S.Koh, “A survey of NFC Mobile Payment: Challenges and solutions using blockchain and cryptocurrencies,” 2020 7th International Conference on Dependable Systems and Their Applications (DSA), pp. 69-77, 2020.

[6] Swami, Ananthram, and Brian M. Sadler. “Hierarchical digital modulation classification using cumulants.” IEEE Transactions on communications 48.3 (2000): 416-429.

[7] O’Shea, Timothy James, Tamoguna Roy, and T. Charles Clancy. “Over-the-air deep learning based radio signal classification.” IEEE Journal of Selected Topics in Signal Processing 12.1 (2018): 168-179.

[8] Lin, Yun, et al. “The individual identification method of wireless device based on dimensionality reduction and machine learning.” The Journal of Supercomputing 75.6 (2019): 3010-3027.

[9] Wang, Yu, et al. “LightAMC: Lightweight automatic modulation classification via deep learning and compressive sensing.” IEEE Transactions on Vehicular Technology 69.3 (2020): 3491-3495.

[10] Sharif Razavian, Ali, et al. “CNN features off-the-shelf: an astounding baseline for recognition.” Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2014.

[11] Wu, Ruowu, et al. “Abnormal information identification and elimination in cognitive networks.” International Journal of Performability Engineering 14.10 (2018): 2271.

[12] Wang, Hui, et al. “Fractal complexity-based feature extraction algorithm of communication signals.” Fractals 25.04 (2017): 1740008.

[13] Zhang, Chaojun, Paul Patras, and Hamed Haddadi. “Deep learning in mobile and wireless networking: A survey.” IEEE communications surveys & tutorials 21.3 (2019): 2224-2287.

[14] Sun, Yaohua, et al. “Application of machine learning in wireless networks: Key techniques and open issues.” IEEE Communications Surveys & Tutorials 21.4 (2019): 3072-3108.

[15] C. Hou, X.Zhang, and X. Chen, “Electromagnetic Signal Feature Fusion and recognition based on multi-modal Deep learning,” International Journal of Performability Engineering, vol. 16, no. 6, pp. 941-949, June 2020.

[16] Wang, Xuyu, Xiangyu Wang, and Shiwen Mao. “CiFi: Deep convolutional neural networks for indoor localization with 5 GHz Wi-Fi.” 2017 IEEE International Conference on Communications (ICC). IEEE, 2017.

[17] Wang, Xuyu, Xiangyu Wang, and Shiwen Mao. “Deep convolutional neural networks for indoor localization with CSI images.” IEEE Transactions on Network Science and Engineering 7.1 (2018): 316-327.

[18] Ya, T. U., et al. “Large-scale real-world radio signal recognition with deep learning.” Chinese Journal of Aeronautics (2021).

[19] O’Shea, Timothy J., and Nathan West. “Radio machine learning dataset generation with gnuradio.” Proceedings of the GNU Radio Conference. Vol. 1. No. 1. 2016.

[20] O’Shea, Timothy, and Jakob Hoydis. “An introduction to deep learning for the physical layer.” IEEE Transactions on Cognitive Communications and Networking 3.4 (2017): 563-575.

[21] Wang, Meiyu, et al. “Transfer learning promotes 6G wireless communications: recent advances and future challenges.” IEEE Transactions on Reliability (2021).

[22] Lin, Yun, Ya Tu, and Zheng Dou. “An improved neural network pruning technology for automatic modulation classification in edge devices.” IEEE Transactions on Vehicular Technology 69.5 (2020): 5703-5706.

[23] Lin, Yun, et al. “Adversarial attacks in modulation recognition with convolutional neural networks.” IEEE Transactions on Reliability 70.1 (2020): 389-401.

[24] Bao, Zhiha, et al. “Threat of Adversarial Attacks on DL-based IoT Device Identification.” IEEE Internet of Things Journal (2021).

[25] Dong, Yihong, et al. “SR2CNN: zero-shot learning for signal recognition.” IEEE Transactions on Signal Processing 69 (2021): 2316-2329.
[26] Peng, Shengliang, et al. “Modulation classification based on signal constellation diagrams and deep learning.” IEEE transactions on neural networks and learning systems 30.3 (2018): 718-727.

[27] Lin, Yun, et al. “Contour stella image and deep learning for signal recognition in the physical layer.” IEEE Transactions on Cognitive Communications and Networking 7.1 (2020): 34-46.

[28] ZHANG, Sicheng, et al. “Electromagnetic signal modulation recognition technology based on lightweight deep neural network.” Journal on Communications 41.11 (2020): 12.

[29] Wang, Yu, et al. “Data-driven deep learning for automatic modulation recognition in cognitive radios.” IEEE Transactions on Vehicular Technology 68.4 (2019): 4074-4077.

[30] Tu, Ya, et al. “Semi-supervised learning with generative adversarial networks on digital signal modulation classification.” Comput. Mater. Continua 55.2 (2018): 243-254.

[31] Wang, Tao, et al. “Privacy preservation in big data from the communication perspective—A survey.” IEEE Communications Surveys & Tutorials 21.1 (2018): 753-778.

[32] Zheng, Zhigao, Arun Kumar Sangaiya, and Tao Wang. “Adaptive communication protocols in flying ad hoc network.” IEEE Communications Magazine 56.1 (2018): 136-142.

[33] Tang, Bin, et al. “Digital signal modulation classification with data augmentation using generative adversarial nets in cognitive radio networks.” IEEE Access 6 (2018): 15713-15722.

[34] Wang, Hui, et al. “A new method of cognitive signal recognition based on hybrid information entropy and DS evidence theory.” Mobile Networks and Applications 23.4 (2018): 677-685.

[35] Hirose, Akira, and Shotaro Yoshida. “Generalization characteristics of complex-valued feedforward neural networks in relation to signal coherence.” IEEE Transactions on Neural Networks and learning systems 23.4 (2012): 541-551.

[36] Hirose, Akira. “Complex-valued neural networks: The merits and their origins.” 2009 International joint conference on neural networks. IEEE, 2009.

[37] Tu, Ya, et al. “Complex-valued networks for automatic modulation classification.” IEEE Transactions on Vehicular Technology 69.9 (2020): 10085-10089.

[38] Al-Nuaimi, Ahmed Yanub H., Md Fajul Aman, and Kazuyuki Murase. “Enhancing MP3 encoding by utilizing a predictive complex-valued neural network.” The 2012 International Joint Conference on Neural Networks (IJCNN), IEEE, 2012.

[39] Zhang, Zhanjian, et al. “Complex-valued convolutional neural network and its application in polarimetric SAR image classification.” IEEE Transactions on Geoscience and Remote Sensing 55.12 (2017): 7177-7188.

[40] Simonyan, Karen, and Andrew Zisserman. “Very deep convolutional networks for large-scale image recognition.” arXiv preprint arXiv:1409.1556 (2014).

[41] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. “Imagenet classification with deep convolutional neural networks.” Advances in neural information processing systems 25 (2012).

[42] He, Kaiming, et al. “Deep residual learning for image recognition.” Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[43] Wang, Xuyu, Xiangyu Wang, and Shiwen Mao. “ResLoc: Deep residual sharing learning for indoor localization with CSI tensors.” 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), IEEE, 2017.

[44] Ollila, Esa, Visa Koivunen, and H. Vincent Poor. “Complex-valued signal processing—essential models, tools and statistics.” 2011 Information Theory and Applications Workshop. IEEE, 2011.

[45] Choi, Hyeong-Seok, et al. “Phase-aware speech enhancement with deep complex u-net.” International Conference on Learning Representations. 2018.

[46] Al-Sa'id, Mohammad F., et al. “RF-based drone detection and identification using deep learning approaches: An initiative towards a large open source drone database.” Future Generation Computer Systems 100 (2019): 86-97.