IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 6, 2020, accepted October 22, 2020, date of publication October 30, 2020, date of current version November 11, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3034942

Automatic Modulation Classification in Time-Varying Channels Based on Deep Learning

YU ZHOU, (Student Member, IEEE), TIAN LIN, (Student Member, IEEE), AND YU ZHU, (Member, IEEE)

Key Laboratory for Information Science of Electromagnetic Waves (MoE), Department of Communication Science and Engineering, Fudan University, Shanghai 200433, China

Corresponding author: Yu Zhu (zhuyu@fudan.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61771147.

ABSTRACT Automatic modulation classification (AMC) is an important technology in military signal reconnaissance and civilian communications such as cognitive radios. Most of the existing works focused on the AMC in additional white Gaussian noise channels, but the AMC in time-varying wireless channels is more practical and challenging. In this article, we investigate the AMC in time-varying channels by using the deep learning method for high classification accuracy. Specifically, we take the modulation constellation diagram (CD) as the key feature and propose a slotted constellation diagram (slotted-CD) scheme in order to extract the feature of the time-evolution of the CD due to channel variation. We then develop an advanced neural network for modulation classification, where the output sub-images from the slotted-CD feature extractor are first processed separately by a number of parallel convolutional neural networks and then further processed by a recurrent neural network for exploring their time relationship. Experimental results show that the proposed AMC scheme achieves higher classification accuracy in both slow and fast fading channels when compared with the traditional deep learning based AMC schemes. Such performance improvement can be clearly illustrated by visualizing the outputs of the convolutional layers of the classifier. We also show that visualization can help optimize the parameters of the AMC neural networks.

INDEX TERMS Automatic modulation classification, constellation diagram, time-varying, convolutional neural network, bidirectional long short-term memory network.

I. INTRODUCTION

A. MOTIVATION AND BACKGROUND

With the explosive development of wireless communication technology and its ubiquitous applications, automatic modulation classification (AMC) has become an essential technique in both military and civilian fields [1]–[3]. For example, in military electronic countermeasures, in order to monitor enemy information, it is necessary to identify the modulation format and then demodulate the signal and decrypt the message. In a cognitive radio system, AMC can assist the secondary user to detect the existence of the primary user, especially at a low signal-to-noise ratio (SNR) [4], [5]. In addition, if the secondary user is able to recognize the modulation format of the primary user, it can then select a suitable modulation format for transmission to reduce its interference to the primary user [5], [6].

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

In literature, there are two main AMC methods: likelihood-based method and feature-based method [2]. The former treats AMC as a hypothesis testing problem based on the likelihood function with the received signal, and aims to maximize the likelihood function among multiple hypotheses [7]–[9]. However, it requires some prior information that may not always be available in practical applications, for example, the channel state information of a time-varying channel [10]. The latter is a statistical pattern recognition method and less subject to prior information. It usually consists of a feature extractor and a classifier. For the feature extractor, some specific features from the received signal are extracted, such as instantaneous amplitude, phase and frequency [11]–[13], higher order cumulants [14], [15], cyclostationary features such as the spectral correlation function [16]–[18], short-time Fourier transform [19]–[21], and constellation diagram (CD) [22], [23], etc. For the classifier, the extracted feature is then classified into different modulation formats. Traditional classifiers relied on artificial decisions, such as decision

trees [14], [24]. Recently, compressed sensing [25], [26] and classical machine learning techniques such as support vector machine (SVM) [27]–[30], clustering [29], and K-nearest neighbor (KNN) [31], have been applied to the design of classifier.

More recently, because of the superior capability of big data processing and classification, deep learning has been successfully applied in image classification, machine translation, and natural language processing [32]–[37], etc. Deep learning has also been considered in the classifier design of AMC with different network structures such as deep neural network (DNN) [15], [38], recurrent neural network (RNN) [39], and convolutional neural network (CNN) [40]. Specifically, because of the advantages of CNN in processing images and according to the fact that different modulation formats have different characteristics that can be demonstrated into images, researchers have proposed different CNN-based AMC schemes by taking different features, which include baseband signal waveform such as the sampled in-phase and quadrature (IQ) signals [40]–[42], eye diagram [43], Choi-Williams distribution [44], cyclic spectrum diagram [16], and CD related images such as the enhanced color image generated by using constellation density [23] or the one by masking the constellation density matrix with a proper filter [45], etc. Most of the above works assumed the additional white Gaussian noise (AWGN) or time-invariant channel model. However, in practical wireless systems, the propagation channel between the two communication sides varies due to the relative motion between them and/or the changes in the environment. Thus, the investigation of AMC in time-varying fading channels is more meaningful and challenging.

B. CONTRIBUTION

In this article, we investigate the AMC by using the deep learning method. In contrast to the previous designs which commonly assumed the AWGN channel, we particularly aim at the AMC design for time-varying channels. Our contributions can be summarized as follows.

- *Feature Extraction*: We propose a new feature extraction scheme for the AMC in time-varying channels, which is referred to as slotted constellation diagram (slotted-CD). In particular, the entire sampling time window is divided into consecutive time slots, and all the sampled IQ signals are divided into consecutive segments accordingly. The IQ signals are converted to a number of two dimensional (2D) gray images. By doing so, not only the trajectory of the rotation and scaling of the CD within each time slot, but also the evolution among the consecutive images can be utilized for better classification.
- *Design of Neural Network*: We propose an advanced classifier for the AMC in time-varying channels. In contrast to the conventional deep learning based classifiers, which normally consist of a single CNN or a fully connected DNN, the proposed classifier is a combined network of three basic neural networks. The first part employs several parallel CNNs and seamlessly connects to the slotted-CD extractor, where each sub-image corresponding to a time slot is processed by a parallel CNN. The second part is a bidirectional long short-term memory network (BLSTM) which processes the outputs of the parallel CNNs in the order of time slots, so as to exploit the time relationship among them for better classification accuracy. The last part is a DNN consisting of three fully connected dense layers for outputting the final result. It is due to the three-part structure that the proposed classifier can make full use of the output of the slotted-CD extractor and achieve high classification accuracy.
- *Experiments and Visualization*: We provide various experimental results to verify the effectiveness of the proposed AMC scheme. For example, we show that the proposed AMC scheme can achieve a high classification accuracy rate of almost 100% among 8 modulation formats in a slow fading channel with a moving speed region of 1.5m/s $\sim$ 4.5m/s and a 97% accuracy rate in a fast fading channel with a speed region of 22.5m/s $\sim$ 45m/s, respectively. Furthermore, we visualize the proposed neural networks. The visualization results verify the performance improvement of the proposed AMC scheme and also provide a means in optimizing the parameters of the AMC neural networks.

The rest of this article is organized as follows. Section II introduces the system model. Section III introduces the basic scheme for extracting the feature of CD and then proposes the slotted-CD scheme for time-varying channels. Section IV proposes a new classifier for the AMC in time-varying channels. Section V and VI provide experimental results and visualization results, respectively. Finally, conclusions are drawn in Section VII.

II. SYSTEM MODEL

Consider a communication system with AMC in Fig. 1. It is assumed that eight modulation formats, including three phase-shift keying (PSK) formats and five quadrature amplitude modulation (QAM) formats, are possibly taken at the transmitter and need to be classified at the receiver, which are binary PSK (BPSK), 4PSK, 8PSK, 16QAM, 32QAM, 64QAM, 128QAM and 256QAM. Similar to that in [23], it is assumed that the receiver has perfect synchronization with the transmitter. According to [46], the relationship between the transmitted symbol, $x[n]$, and the received baseband sampled signal, $y[n]$, at time instant n, can be represented in the following discrete-time baseband equivalent model

$$y[n] = h[n] \cdot x[n] + w[n], \tag{1}$$

where $w[n]$ denotes the AWGN signal and $h[n]$ denotes the time-varying channel coefficient at time n. Usually, $h[n]$ can be described by some statistical models such as Clarke's model [46], where it is regarded as a stationary random process with a Doppler spectrum. To characterize how the channel varies over time, the coherence time T_c is defined

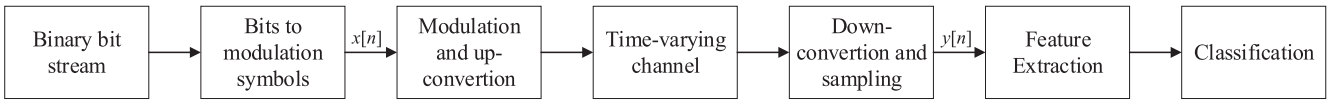

FIGURE 1. The system diagram for the AMC in time-varying wireless channels.

as the interval over which $h[n]$ is highly correlated and is inversely proportional to the Doppler spread. For example, according to [47], if T_c is defined as the interval over which the time correlation of $h[n]$ is above 0.5, then it can be approximated as

$$T_c \approx \frac{9}{16\pi f_m}, \tag{2}$$

where $f_m = v/\lambda$ is the maximum Doppler shift with v denoting the speed and λ denoting the carrier wave length.

As can be seen from (1), besides the noise effect, the channel fading $h[n]$ makes the modulation formats more challenging to recognize. As shown in Fig. 1, the AMC at the receiver normally consists of two steps. The first step is mainly to pre-process the received signals by extracting their features so that more suitable early-stage data can be obtained for the convenience of subsequent signal processing. With the extracted features, the second step is then to classify them through a neural network and obtain the classification decision. In the following two sections, we elaborate our design in these two steps, respectively.

III. FEATURE EXTRACTION OF CD

A. BASIC IDEA

For digital modulation, as the baseband signal is modulated in both the amplitude and phase of the carrier, different modulation formats have distinct CDs. More interestingly, for a typical time-varying fading channel, the CD changes with the channel coefficient in terms of rotation and scaling, which could be used for the classification of modulation formats.

With this motivation, the next question is how to efficiently generate a 2D image representing the CD for the classifier with the original raw baseband sampled signals. The main difficulties are two fold. *First*, the complex plane containing the CD has infinite resolution. However, the images sent to the classifier should have limited resolution, and from the perspective of computational complexity: the lower resolution, the better. *Second*, because of the channel variation, there will be the rotation and scaling trajectory in the CD. How to extract the trajectory, and more importantly, extract the evolution of the trajectory when generating the images is another difficulty.

To deal with these difficulties, we propose to complete the task in two levels and introduce them in the following two subsections, respectively. In Section III-B, we introduce a basic method to convert IQ sampled signals into a 2D gray image. In Section III-C, considering the time variation effect, we propose the slotted-CD scheme by dividing the sampling window into multiple time slots and generating a series of 2D sub-images over these slots.

B. GENERATING A GRAY IMAGE FOR CD

For ease of presentation, take the transmission of 16QAM in an AWGN channel with bit SNR = 10dB as an example.

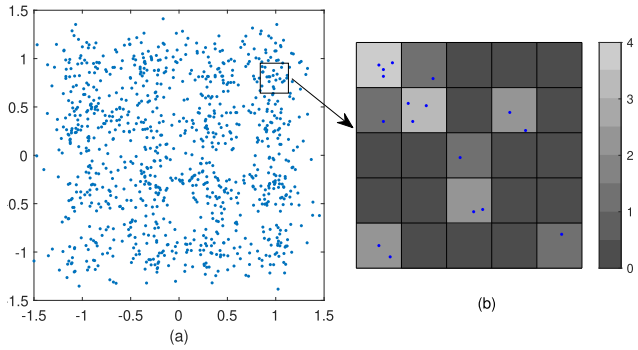

FIGURE 2. Conversion from a complex plane containing the IQ sampled signal points to a gray image. (a) The original complex plane containing the baseband IQ sampled signal points; (b) The gray image generated by first dividing the complex plane into cells and mapping the number of sampled signal points in each cell to a gray value.

The sampled baseband IQ signals are first converted into pixel points and are drawn in the complex plane as shown in Fig. 2(a). In theory, the complex plane has infinite resolution. However, the output image should have a limited resolution. Thus, we equally split the complex plane containing the CD into $R \times R$ cells corresponding to a predefined resolution level. Then, the number of dots falling into each cell is counted and mapped to a gray value of the associated pixel for that cell, as shown in Fig. 2(b).

However, in the time-varying fading channel, what can be observed in the complex plane at the receiver is much different from that in the AWGN channel. Take the transmission of 16QAM in a time-varying channel as an example, where the channel follows the Clark's model with a normalized channel power. Assume a carrier frequency $f_c = 2$GHz, a bandwidth $W = 100$kHz, SNR = 18dB, and a moving speed $v = 1.5$m/s. By ignoring the effect of pulse shaping and assuming that the sampling rate is equal to the symbol rate and the bandwidth for simplicity, there are 36000 IQ sampled signal points in a sampling time window of 360ms, as shown in Fig. 3(a).

It can be clearly seen that besides the noise influence, the CD is rotated and scaled with the channel variation. A proper boundary should be determined in the complex plane to convert it into a gray image for AMC. In the traditional case for the AWGN channel, since most of the amplitudes of the received signal points are within a limited range, the CD could be cut with a fixed boundary in the complex plane [23]. However, in a time-varying channel, as can be seen from (1), the received signal is affected by both the channel variation and the Gaussian noise. If the boundary is still set to a fixed value, there is always a probability that the time-varying channel coefficient makes the received signal points exceed the boundary, as shown by the IQ sampled points outside the red colored boundary in Fig. 3(a). Those points cannot be utilized for AMC. To deal with this

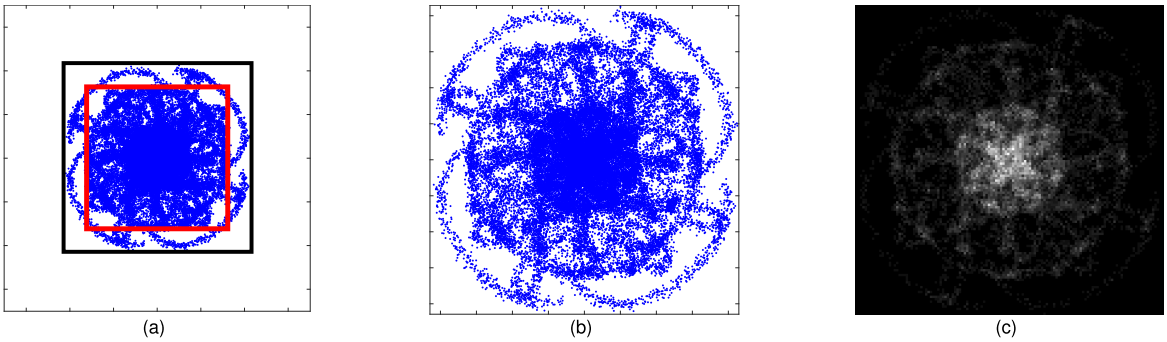

FIGURE 3. Converting sampled signal points into a gray image for the AMC in time-varying channels. (a) The whole complex plane containing all IQ baseband sampled signal points; (b) The area of the complex plane within the black boundary in (a); (c) The final 2D gray image.

problem, considering the fact that for a finite number of IQ sampled points, the maximum value of the signal amplitude is finite, we then set the boundary to 101% of the maximum amplitude of IQ sampled signals. Then, all the sampled signals are included in the image, as shown in Fig. 3(b), which is also the area within the black boundary in Fig. 3(a). Then, by mapping the number of sampled points in each cell in Fig. 3(b) to a gray value as similar to that in Fig. 2, a gray image with $R = 128$ is obtained, as shown in Fig. 3(c).

C. SLOTTED-CD SCHEME

Although the gray image generated according to the basic method in Section III-B shows the trajectory of the CD rotation and scaling in time-varying channels, it may not bring enough information for better AMC because of the following two limitations.

- If the channel fading is relatively fast and the number of sampled points is large, there would be many points that overlap each other. Unlike that in AWGN channels, where the sampled points are concentrated around the transmitted symbols, in time-varying fading channels, the sampled points falling into one particular cell may correspond to different constellation points in different fading states. That is, due to rotation and scaling, different transmitted constellation points can be received with very similar IQ sampled points. Thus, it becomes difficult to associate the sample points with the original transmitted constellation points for AMC.
- Although the gray image shows the rotation and scaling trajectory, it does not reflect how it evolves. The IQ signals are actually a pair of functions of time. With only a single image, we can only see the result of a process over a time period but not the whole evolution of the process. This is like using an oscilloscope to display signals. Normally it has two modes: volts-versus-time display mode and volts-versus-volts display mode (or X-Y mode). In the time model, it shows the two input signals as functions of time, while in the X-Y mode, it shows one input signal against the other. The CD corresponds to the display of the I and Q signals in the X-Y mode. However, with only the X-Y mode, the information is not completely provided.

Thus, it is necessary to find some sophisticated feature extraction scheme for the AMC over time-varying channels. In this study, we propose the slotted-CD scheme, where the entire sampling window, T_{w}, is divided into a number of time slots with time length T_{st}, and a gray sub-image is generated for each time slot according to the method in Section III-B. Not only the trajectory of the CD rotating and scaling within one time slot, but also the evolution among the consecutive images can be utilized for the classifier. Then, the key here is how to set the slot length. One requirement is that the channel should be highly correlated (described by coherence time [47]) so that there is no significant overlap in the trajectory of the CD within one time slot. This is similar to that in setting the block length for channel estimation, where the channel in the pilot part and the date part in a transmission block should be highly correlated [48]. For another example, for the transmit precoding with channel state information feedback from the receiver, the channel should be highly correlated within one estimation-feedback-precoding process [49]. For a system with a particular bandwidth, the slot length is then represented equivalently by the number of samples in a slot, denoted by N_{st}. There are several criteria that can be considered when setting the slot length.

1) To guarantee that the sub-image of each slot has enough sampled points to cover almost all the original constellation points. For example, take twice or three times the number of original constellation points as a lower limit.
2) To guarantee that each sub-image contains the sampled points in a limited number of coherence time periods. For example, take one to three coherence time periods as an upper limit.
3) If the above two criteria cannot be satisfied simultaneously, then the first one has a higher priority.

It can be seen that the first criterion actually sets the minimum value of T_{st} (or equivalently N_{st}) such that the basic feature of a CD can be extracted in each sub-image. The second criterion sets the maximum value of T_{st} such that the effect of the trajectory overlapping problem due to the channel time variation is limited.

To illustrate the result of the proposed slotted-CD scheme, consider the same example system like that in Fig. 3 with a carrier frequency $f_{\mathrm{c}} = 2$GHz, a bandwidth $W = 100$kHz,

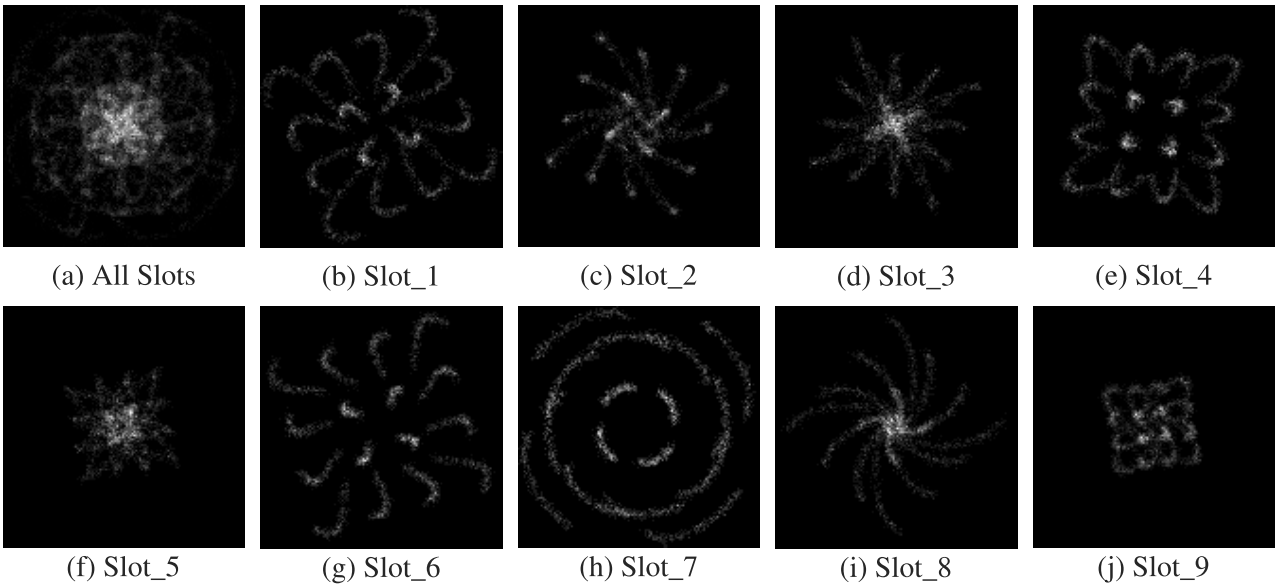

FIGURE 4. Result of the slotted-CD feature extraction scheme. (a) Result of the basic feature extraction scheme without slot division, which is same as Fig. 3(c); (b)-(j) Nine sub-images corresponding to the nine time slots.

SNR = 18dB, and a speed $v = 1.5$m/s. The coherence time is $T_c = 18$ms according to (2). According to Criterion 1, it is better to set N_{st} to be larger than 512 when taking 256QAM as the highest modulation order. According to Criterion 2, three times of the T_c corresponds to about 5400 samples, which can be taken as an upper limit of N_{st}. Assume that the total sampling time window is 360ms, which contains a total of 36000 samples. According to the above criteria, the window can be divided into $L = 9$ slots with each slot containing $N_{st} = 4000$ samples.

Fig. 4 demonstrates the result of the slotted-CD scheme, where Fig. 4(a) shows the gray image of the basic processing without slot division, which is the same as Fig. 3(c). Figs. 4(b)-(j) show the nine sub-images in the nine time slots, respectively. It can be seen that compared with Fig. 4(a), the trajectory overlapping effect is greatly reduced. The evolution of the CD can be clearly seen by taking the nine sub-images together.

IV. DESIGN OF CLASSIFIER

In this section, we first introduce a basic classifier, where the network structure is similar to the conventional AMC network for AWGN channels, but the parameters are optimized by training with the samples in time-varying channels. We then propose an advanced classifier for the AMC in time-varying channels. Finally, we introduce some common configurations of the two classifiers, such as loss function, activation function, etc.

A. BASIC CLASSIFIER: SCDN

We start with a basic classifier in Fig. 5. Similar to the conventional AMC network for AWGN channels [23], the basic classifier consists of a single CNN and a DNN in series (we refer to this classifier as SCDN). The input to the SCDN is a single $R \times R$ gray image, which is generated by the basic feature extraction scheme in Section III-B. The parameters of the SCDN are fine-tuned by training in time-varying channels, and thus, the SCDN scheme can be regarded as a benchmark for the AMC in time-varying channels.

As shown in Fig. 5, the CNN part of the SCDN consists of four convolutional layers with 8, 16, 32, 16 filters and the kernel size of 3×3, 3×3, 2×2, 2×2, respectively, and three max-pooling layers with the same size of 2×2 for layer one to three. A flatten layer is deployed to connect the CNN and DNN with the output dimension of $R^2/4$. The DNN part consists of four fully connected dense layers with 1024, 256, 64, 8 neurons, respectively.

B. ADVANCED CLASSIFIER: MCBLDN

The basic classifier SCDN in Section IV-A has two limitations in the AMC for time-varying channels. First, the input is a single gray image containing all the received IQ signal points over a time period. As shown in Section III-C, it does not provide enough information about how the CD rotates and scales with the channel variation. Second, there is no specific design in the network structure of SCDN, for dealing with the channel time variation effect.

In order to overcome these limitations, we first take the proposed slotted-CD feature extractor in Section III-C, and then design an advanced classifier to process the multiple sub-images from the slotted-CD feature extractor. Assume that there are totally L time slots and then L associated sub-images. As shown in Fig. 6, the classifier consists of three parts in series. The first part consists of multiple parallel CNNs with the same structure, each of which processes a sub-image separately. As shown in Fig. 6(b), each CNN consists of three convolutional layers with the kernel size of 3×3, 3×3, 2×2, and N_F, $2N_F$, and N_F convolutional filters, respectively. The pool size of each max-pooling layer is 2×2. The dimension of the output of each CNN is $N_F \times (R/8) \times (R/8)$.

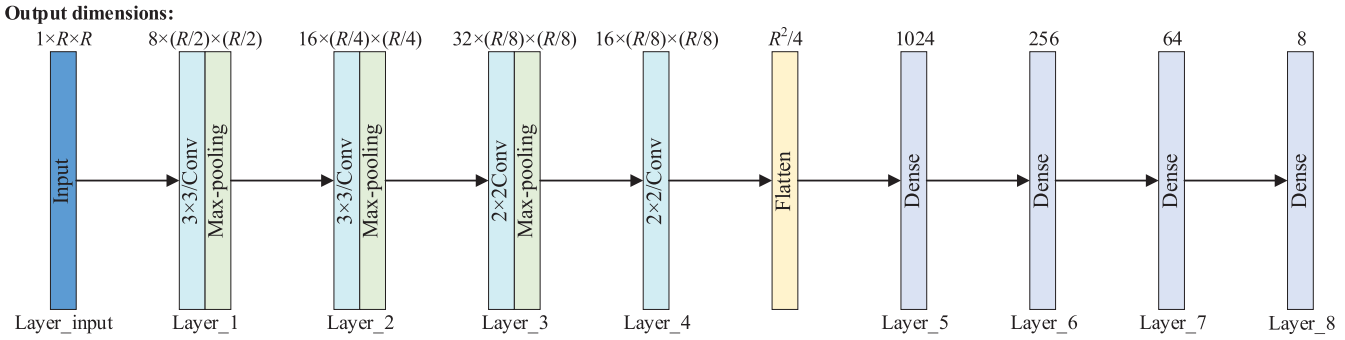

FIGURE 5. Network structure of the basic SCDN classifier.

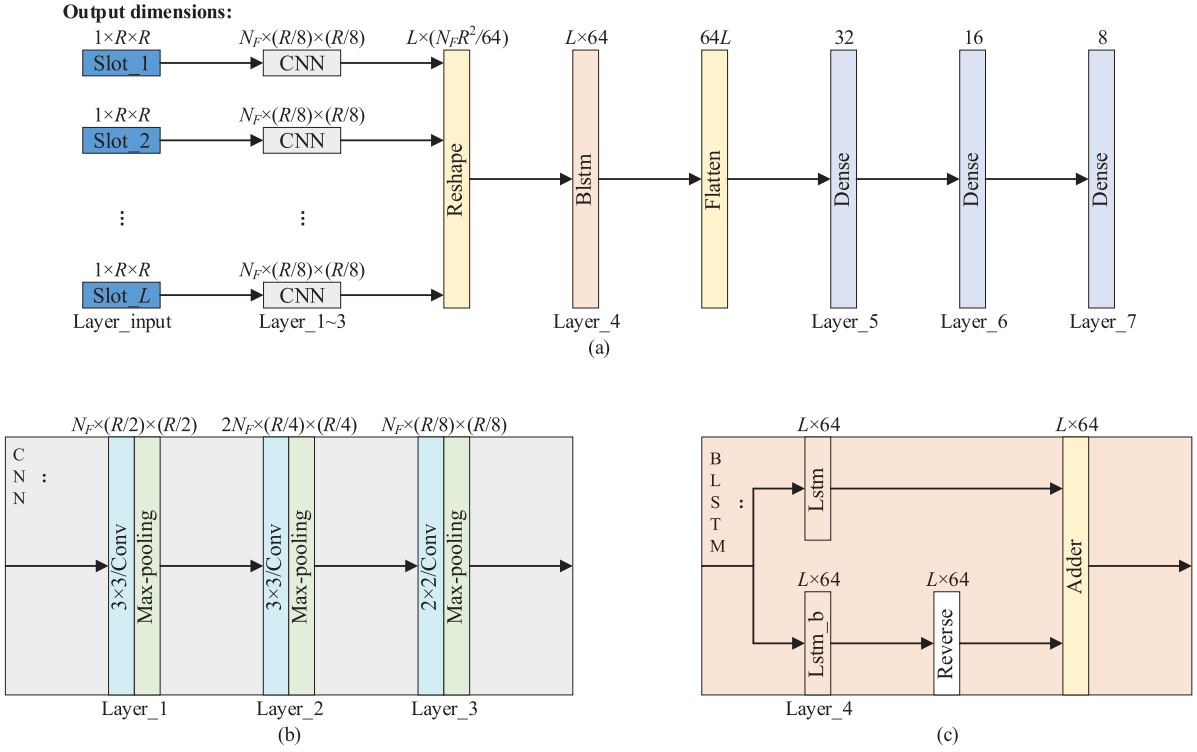

FIGURE 6. (a) Network structure of the proposed MCBLDN classifier; (b) The detailed structure of the CNN block in (a); (c) The detailed structure of the BLSTM block in (a).

It should be mentioned that these sub-images are related in time, which inspires us to think about the RNN architecture. Therefore, the second part of the advanced classifier is BLSTM, whose detailed structure is shown in Fig. 6(c). Assuming that there are totally L time slots, the outputs of the preceding L parallel CNNs are first concatenated and reshaped with a data size of $L \times N_F R^2/64$ and then sent into the BLSTM. Note that in this reshaping operation, the original time order of L sub-images from the slotted-CD feature extractor is retained in the input of BLSTM so that the time relationship among these sub-images is utilized for AMC. The BLSTM consists of two LSTMs, where the upper one in Fig. 6(c) processes the input data in the positive order and the lower one in the reverse order. The operation of either LSTM consists of L steps. For each time step, the input dimension is $N_F R^2/64$ and the output dimension is 64. For better utilization of the time sequence information, the outputs of all steps are retained. Therefore, the output dimension of either LSTM is $L \times 64$. Finally, the output of the lower LSTM is reordered, added to the output of the upper LSTM, and converted into the data with the dimension of $64L$ through a flatten layer.

The third part of the advanced classifier is a DNN which consists of three fully connected dense layers with 32, 16, 8 neurons, respectively. As the whole advanced classifier consists of multiple CNNs in the first part, and a BLSTM network and a DNN in the other two parts, we refer to this classifier as MCBLDN.

C. COMMON CONFIGURATIONS OF THE TWO CLASSIFIERS

Some common configurations of the above two classifiers are introduced as follows. One-hot coding is used for the output of both the basic SCDN and the proposed MCBLDN.

TABLE 1. System setup for the simulation over slow time-varying channels.

Parameters	Value
Carrier frequency f_c	2GHz
Bandwidth W	100kHz
Window length N_w	36000
Moving velocity v	1.5m/s to 4.5m/s
SNR	$0 \sim 23$dB
Number of training data sets	76.8K
Number of testing data sets per SNR	3K

The loss function is the categorical cross-entropy function. All the layers adopt the rectified linear units (ReLU) activation function except that the last output layer adopts the softmax function. Finally, in order to improve the fitting ability of the two classifiers, the dropout operation is adopted to the fully connected dense layers with the dropout rate set to 0.2.

V. EXPERIMENTS AND RESULTS

In this section, we first compare different AMC schemes in terms of the classification accuracy as a function of SNR in time-varying channels. We then provide some experimental results of the classification accuracy for different channel time variations. Finally, we show the effect of the resolution of the gray image on the classification performance and the network complexity.[1]

A. CLASSIFICATION ACCURACY VERSUS SNR

For slow fading channels, the system setup is the same as that in Fig. 4 and the parameters are summarized in Tab. 1, where $f_c = 2$GHz, and $W = 100$kHz. The time-varying channels satisfy the Clark's model with v uniformly distributed between 1.5m/s to 4.5m/s for the slow fading channels realization. The length of the sampling window is set to 360ms so that a total of 36000 IQ sampled signal points are used for AMC. Three AMC schemes are considered for comparison. For the based SCDN scheme, a single 2D gray image with $R = 128$ is generated according to the method in Section III-B for feature extraction, and the SCDN in Fig. 5 is employed as the classifier. In the proposed advanced scheme, i.e., the MCBLDN scheme for the AMC in the time-varying channels, the slotted-CD scheme in Section III-C is applied for feature extraction, where the sampling window is divided into $L = 9$ time slots. For comparison, we also simulated the AMC scheme in [41], where a residual neural network (ResNet) was developed to directly process the raw IQ data for classification. All of the three AMC schemes use the same training and testing data sets. In the training stage, 76800 training data sets were generated for the 24 SNR values from 0dB to 23dB with 3200 data sets for each SNR. In the testing stage, the final accuracy was obtained by averaging the results over 3000 testing data sets for each SNR.

[1] All source codes about producing source data sets and designing classifiers are provided openly in https://github.com/zhouyu9712/AMC-TVC-DL.

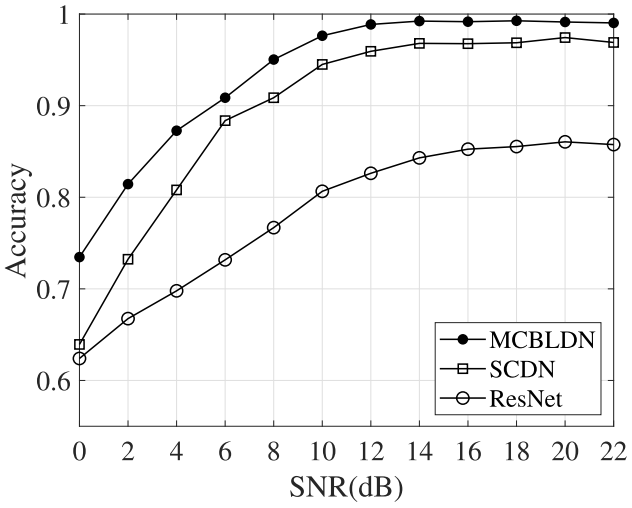

FIGURE 7. AMC accuracy versus SNR for three classifiers over the slow fading channels.

Fig. 7 shows the average AMC accuracy performance versus SNR for the basic SCDN, the proposed MCBLDN, and the conventional ResNet [41] in the slow fading channels. It can be seen that for all the schemes, the classification accuracy increases as the SNR increases, and it increases rapidly at the low SNR region but slowly at the high SNR region. In particular, the ResNet, which directly takes the raw IQ data for classification without any feature extraction, has not shown good performance compared with the CD feature based schemes. This shows that the feature extraction is necessary for the classification in time-varying channels. Fig. 7 also shows that the SCDN scheme always has a performance gap to the MCBLDN scheme over the whole SNR range since it does not exploit specific design to deal with the time-varying effect. By taking the slotted-CD scheme for feature extraction and designing more specific neural networks, the MCBLDN scheme achieves higher accuracy and even approaches almost 100% accuracy at high SNRs.

Figs. 8(a)-(c) demonstrate the detailed classification accuracy in Fig. 7 for the three schemes when SNR is fixed at 10dB, where each row illustrates the detailed classification decisions to all the eight possible modulation formats when a particular modulation format is applied at the transmitter. Obviously, the darker the diagonal element, the higher the accuracy. Comparing the three sub-figures, it can be seen that most of the classification errors happened at high-order modulation formats, where 64QAM and 128QAM are the two most difficult formats to recognize. In particular, the conventional ResNet scheme does not perform well in distinguishing different QAM formats, and the proposed MCBLDN scheme achieves the best accuracy among the three AMC schemes.

B. IMPACT OF CHANNEL VARIATION

In this subsection we evaluate the impact of channel variation on different AMC schemes. Fig. 9 shows the average AMC accuracy performance versus SNR for the three schemes in

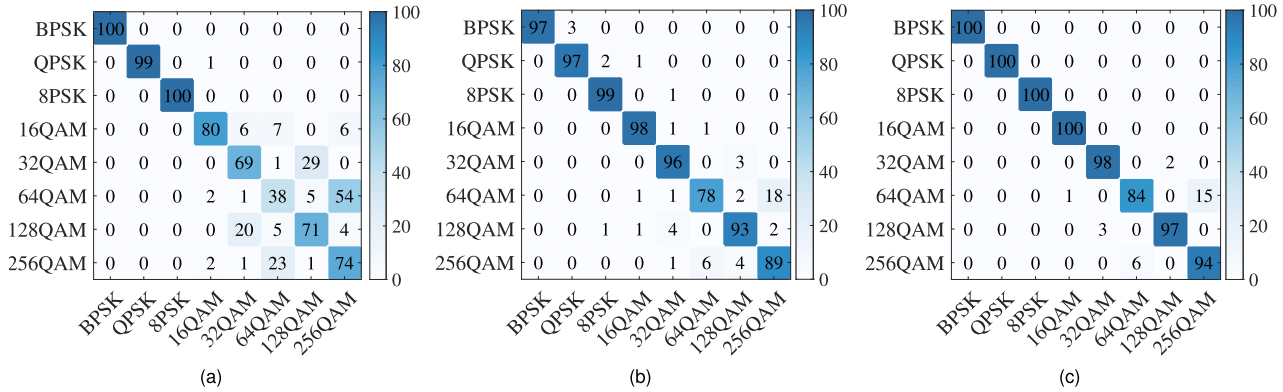

FIGURE 8. The detailed classification accuracy among eight modulation formats. (a) ResNet; (b) SCDN; (c) MCBLDN.

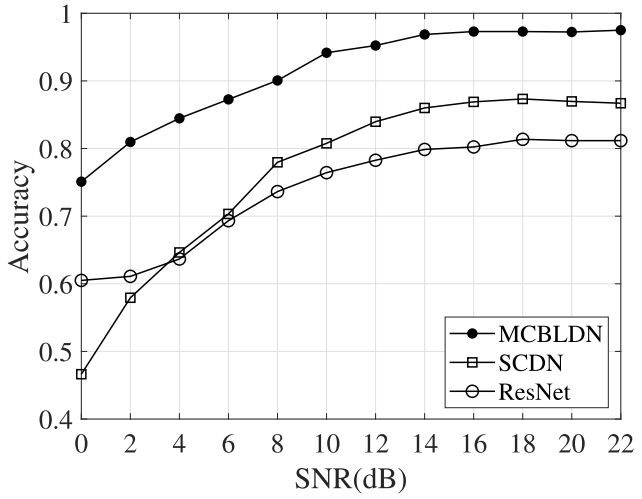

FIGURE 9. AMC accuracy versus SNR for three classifiers over the fast fading channels.

fast fading channels, where all the system parameters are the same as those in the slow fading case except that the moving velocity v is uniformly distributed between 22.5m/s to 45m/s. As the channel coherence time becomes shorter, for the proposed MCBLDN with the slotted-CD scheme, the sampling window is then divided into $L = 30$ time slots. Fig. 9 shows that the proposed MCBLDN scheme can still maintain a high classification accuracy as it not only utilizes the CD variation trajectory within each sub-image but also exploits the evolution among the sub-images via the BLSTM network. However, the classification performance of the basic SCDN scheme is significantly reduced in the fast fading channels, and the performance of the conventional ResNet is also not as good as the MCBLDN scheme.

Next, we evaluate the impact of slot length (data segment) to the classification performance. The system setup is the same as that in Section V-A except that the SNR is fixed at 10dB, and the total number of training data sets is 48000. Two moving speed regions are considered, i.e., 1.5m/s $\sim$ 4.5m/s for the slow fading channels and 22.5m/s $\sim$ 45m/s for the relatively fast fading channels. Furthermore, we take the slot length (or equivalently the number of sampled points per slot $N_{\rm st}$) as a parameter for testing when the total sampling

TABLE 2. Total parameters of different classifiers under slot length $N_{\rm st}$.

$N_{\rm st}$	MCBLDN	SCDN	ResNet
$N_{\rm st} = 400$	341,170	1,334,360	232,040
$N_{\rm st} = 600$	271,540	1,334,360	232,040
$N_{\rm st} = 1200$	201,910	1,334,360	232,040
$N_{\rm st} = 2000$	174,058	1,334,360	232,040
$N_{\rm st} = 4000$	153,169	1,334,360	232,040
$N_{\rm st} = 6000$	146,206	1,334,360	232,040
$N_{\rm st} = 9000$	141,564	1,334,360	232,040

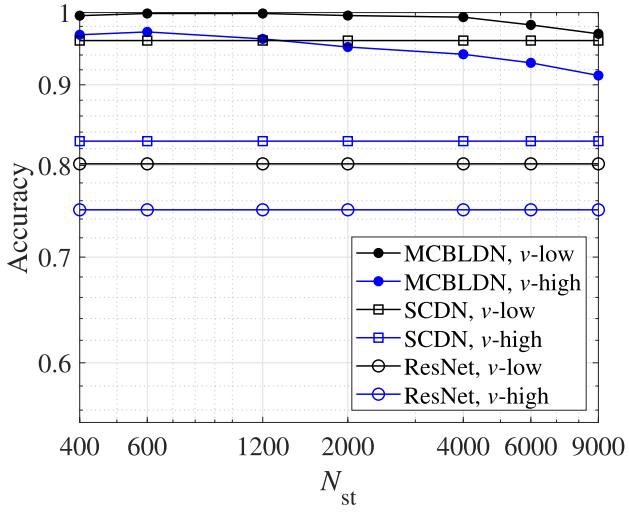

FIGURE 10. The AMC accuracy versus $N_{\rm st}$ for three classifiers over slow and fast time-varying channels.

window (or the total number of sampled signal points $N_{\rm w}$) is fixed. Tab. 2 shows the number of the hyper-parameters of the three different classifiers with $R = 64$. Note that as there is no slot division in the SCDN and ResNet schemes, the number of their hyper-parameters does not change with $N_{\rm st}$.

Fig. 10 shows the average accuracy as a function of $N_{\rm st}$ for the three classification schemes in both slow and fast fading channels when SNR $=$ 10dB and $N_{\rm w} = 36000$. As there is no slot division in the ResNet and SCDN schemes, their performance does not change with $N_{\rm st}$ and is consistent with that in Fig. 7 and Fig. 9. It can be seen from the Fig. 10 that the

TABLE 3. Number of hyper-parameters of different classifiers with different image resolutions.

Resolution $R \times R$	MCBLDN	SCDN	ResNet
128×128	448, 081	4, 480, 088	232, 040
64×64	153, 169	1, 334, 360	232, 040
32×32	79, 441	547, 928	232, 040

performance of the proposed MCBLDN scheme is sensitive to N_{st}. According to our discussion on the criteria of setting the slot length in the slotted-CD scheme in Section III-C, on one side, N_{st} cannot be too large in time-varying channels. Otherwise, there will be significant overlap in the trajectory of the rotation and scaling of the CD in a number of channel coherence time periods. On the other side, N_{st} should be large enough to contain a sufficient number of sampled points in each sub-image.

This is clearly reflected in Fig. 10. In the case of slow fading, the MCBLDN scheme reaches a peak accuracy of 99.5% when $N_{\text{st}} = 2000$, while in the case of fast fading, a peak accuracy of 97% when $N_{\text{st}} = 600$. This is consistent with the variation of the channel coherence time. When the channel varies faster, the coherence time becomes shorter and smaller N_{st} needs to be set. We also notice from Fig. 10 that when N_{st} moves left from the peak point, the classification accuracy does not change too much as long as N_{st} is larger than the lower limit (e.g., twice of the number of constellation points of the maximum modulation order). However, when N_{st} moves right from the peak point, the accuracy gradually decreases as more CD trajectories overlap each other within a sub-image.

C. CLASSIFICATION ACCURACY VERSUS COMPLEXITY

Network complexity is an important issue in some applications that require hardware and power efficient solutions, such as cognitive radio systems. The deep learning based AMC method should be hardware and power-efficient. Since the hardware cost and power consumption are proportional to the network size, we investigate some key parameters in the proposed scheme which determine the network size and provide the trade-off between performance and complexity. Note that one property of the proposed scheme is that no matter how many samples there are in each time slot, the dimension (resolution) of the output gray sub-image from the feature extractor is the same, i.e., $R \times R$. Thus, the smaller the value of R, the smaller the network size and the less the hardware cost and power consumption. According to the detailed network structure of the basic SCDN scheme in Fig. 5 and that of the proposed MCBLDN scheme in Fig. 6, the number of network parameters in the CNN part in these two schemes is proportional to R^2. Thus, the network size and complexity can be greatly reduced with the decrease of R, as shown in Tab. 3.

Fig. 11 further demonstrates the classification accuracy for the basic SCDN scheme and the proposed MCBLDN scheme with different values of R. The system setup is the same as that in Tab. 1. From this figure, we can see that when R is reduced from 128 to 64, the classification accuracy of both

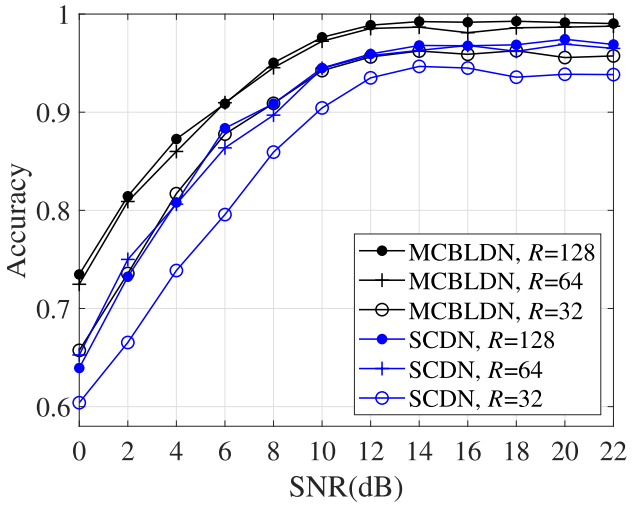

FIGURE 11. AMC accuracy versus SNR with different R for MCBLDN and SCDN over time-varying channels.

the SCDN and the MCBLDN decreases only a little. When R is further reduced to 32, there is an obvious performance degradation. This shows that there may exist a threshold in the resolution of the constellation gray image to the classification performance. By combining the results in Tab. 3 and Fig. 11, we can see that the proposed AMC scheme provides a flexible performance-complexity trade-off by adjusting R according to the performance and complexity requirement. Meanwhile, the proposed MCBLDN scheme can achieve higher performance with less number of hyper-parameters than the conventional SCDN scheme for the AMC over time-varying channels.

VI. VISUALIZATION

Visualization has been shown as a useful tool to illustrate the network performance, and even help optimize the network parameters. In this section, we apply this tool to explain why the proposed MCBLDN scheme outperforms the basic SCDN scheme. We also show that visualization can help optimize the parameters of the classifiers.

A. VISUALIZATION OF SCDN

We show the visualization result of the basic scheme SCDN in this subsection, where the system set up is exactly the same as that in Fig. 7 except that the SNR is fixed at 10dB. Randomly take just one data sample for visualization when the transmitter employs 16QAM. Fig. 12(a) shows the output gray image of the basic feature extractor (also the input of the classifier) and Figs. 12(b)-(e) visualize the outputs of four convolutional layers of the SCDN classifier, respectively. The number of small images in each sub-figure corresponds to the number of output features calculated by each convolutional layer. For example, as the output dimension is $8 \times 128 \times 128$ for the first convolutional layer (see Fig. 5), there are eight features from this layer and thus eight small images in Fig. 12(b).

From Figs. 12(b)-(e) we can see that the features calculated by a convolutional layer is highly related to the features generated by its preceding layer. In general, the low-order

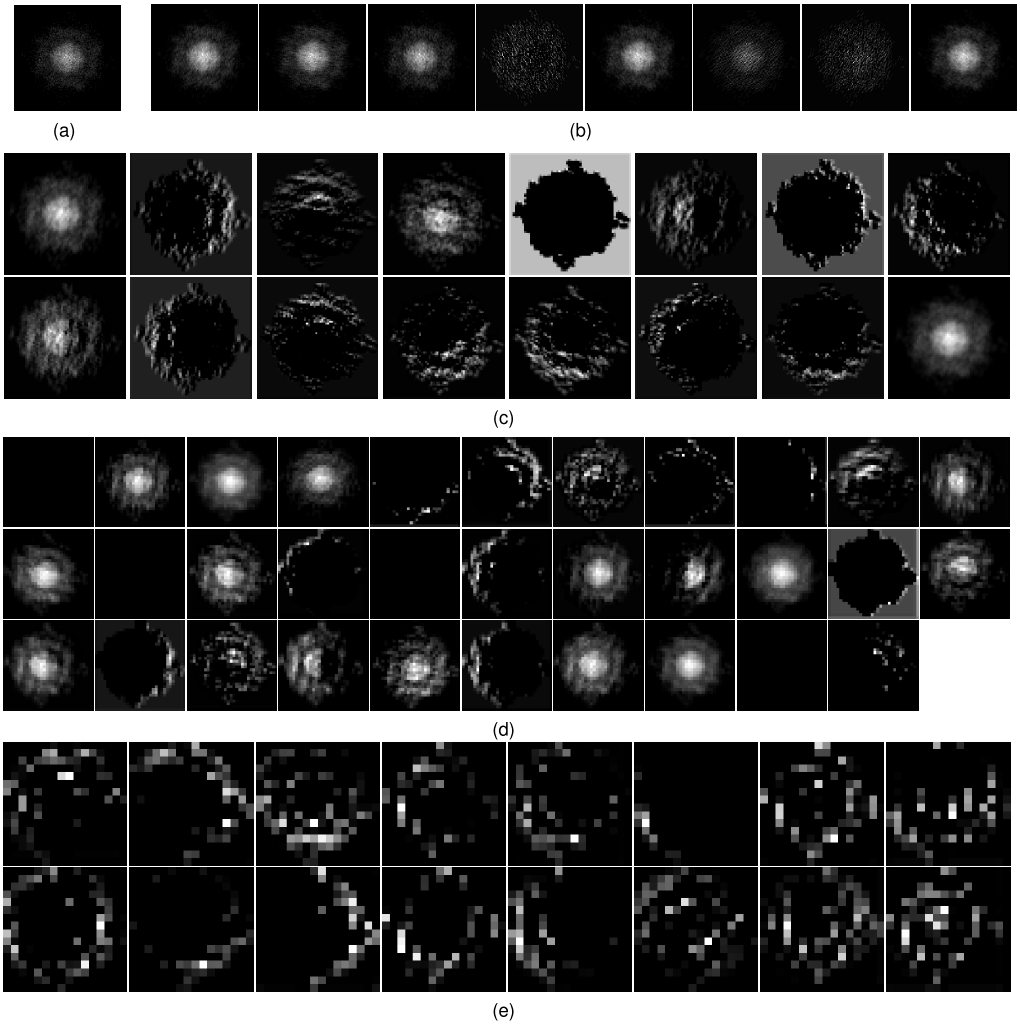

FIGURE 12. Visualization of the SCDN classifier. (a) The input layer; (b) The first convolutional layer; (c) The second convolutional layer; (d) The third convolutional layer; (e) The fourth convolutional layer.

convolutional layers are more concerned on the information like the outline of the input image (e.g., the small image at the first row and the fifth column of Fig. 12(c)), or the edge of the input image (e.g., the small image at the second row and the fourth column of Fig. 12(d)). On the other hand, the features of the high-order layers are more abstract.

These sub-figures also show that for a particular layer, its output features are generally different from each other. Thus, it can be inferred that each output feature focuses on a specific characteristic of the layer input. For example, different small images in Fig. 12(c) focus on different areas of their input images in Fig. 12(b). In the experiment, we also found that the classification accuracy cannot be further improved by adding more layers or more filters in each layer. Actually, when further increasing the number of filters, we observed duplicate feature information in visualization. Therefore, visualization provides a means for setting network parameters such as the number of layers and the number of filters in each layer.

B. VISUALIZATION OF MCBLDN

In this subsection, we demonstrate the visualization result of the first part of the MCBLDN classifier, which consists of $L = 9$ parallel CNNs, with the same data sample as that in Section VI-A. Fig. 13(a) shows the L sub-images with $R = 128$ generated by the proposed slotted-CD feature extractor, and Figs. 13(b)-(d) visualize the outputs of the three convolutional layers of the L parallel CNNs, respectively. According to Fig. 6(b) and Fig. 7, N_F is set to 3, and there are respectively 3, 6, and 3 output features from the three convolutional layers of each CNN. We demonstrate the output features of all the L CNNs column-by-column in Figs. 13(b)-(d). That is, the output features of the three convolutional layers of the l-th CNN are demonstrated in the l-th column in Fig. 13(b), the $(2l-1)$-th and $2l$-th columns in Fig. 13(c), and the l-th column in Fig. 13(d), respectively.

From Figs. 13(b)-(d) we can see that the features calculated by a convolutional layer are highly related to the features generated by its preceding layer. As the L parallel CNNs separately process the L sub-images, the output features of different CNNs are generally different from each other. By comparing Fig. 12 with Fig. 13, the reason why the proposed MCBLDN scheme outperforms the basic SCDN scheme is clear. In the SCDN, as all the IQ baseband signals over a number of coherence time periods are contained in

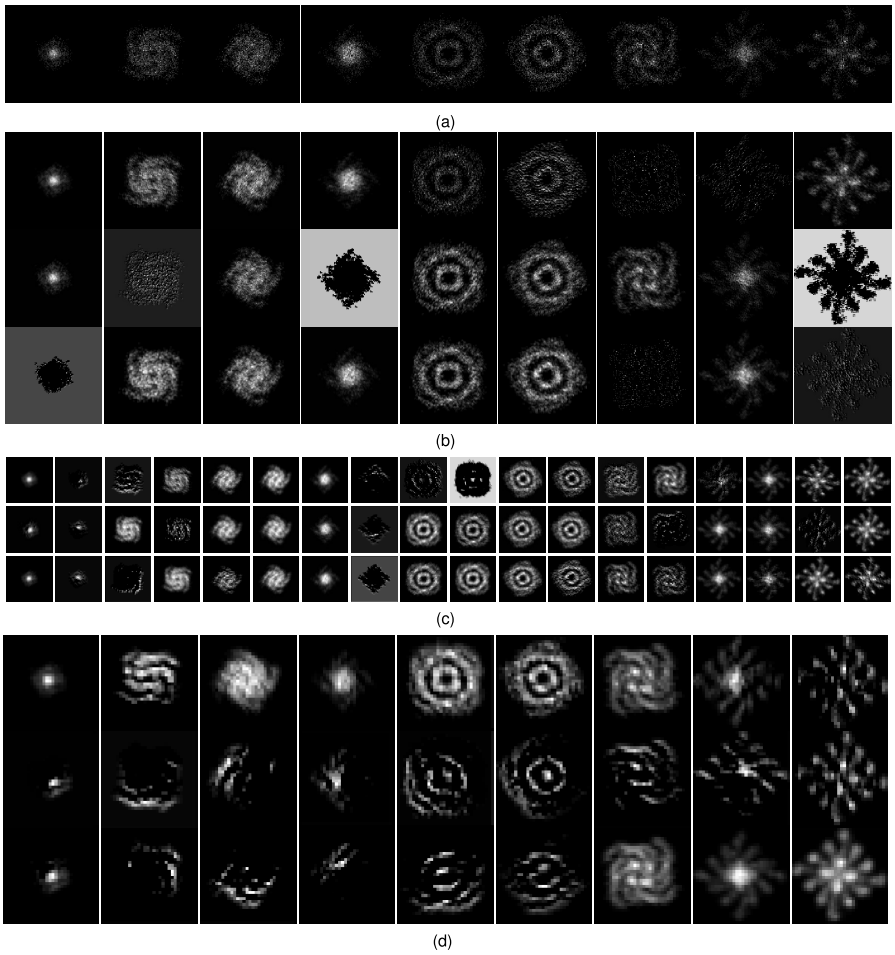

FIGURE 13. Visualization of the MCBLDN classifier with $N_F = 3$. (a) The input layer; (b) The first convolutional layer; (c) The second convolutional layer; (d) The third convolutional layer.

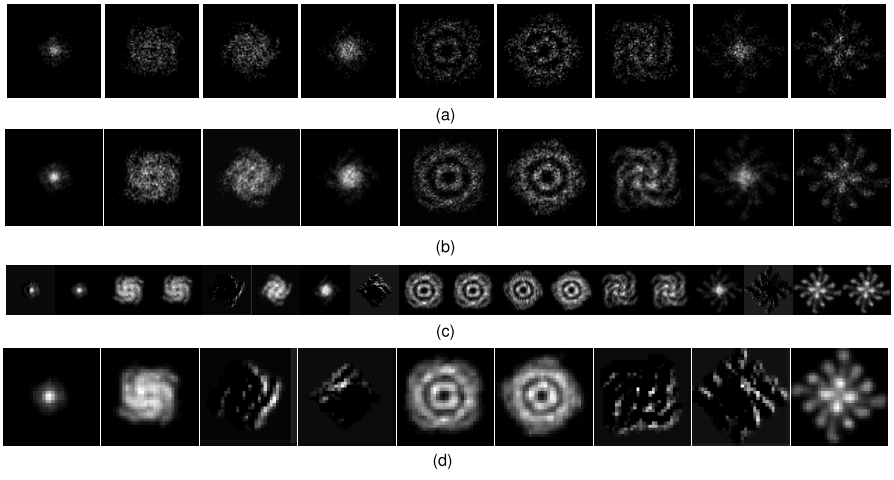

FIGURE 14. Visualization of the MCBLDN classifier with $N_F = 1$. (a) The input layer; (b) The first convolutional layer; (c) The second convolutional layer; (d) The third convolutional layer.

a single image by the basic feature extractor, the SCDN classifier could hardly learn the trajectory of the rotation and scaling of the CD. In contrast, in the MCBLDN scheme, the slotted-CD feature extractor divides the original image into a series of sub-images that could guide the parallel CNNs to learn the information of the variation of the CD. In other words, if the slot length is properly selected, each parallel CNN could learn the trajectory of the rotation and scaling of the CD within each sub-image, and then outputs finer features for the succeeding network.

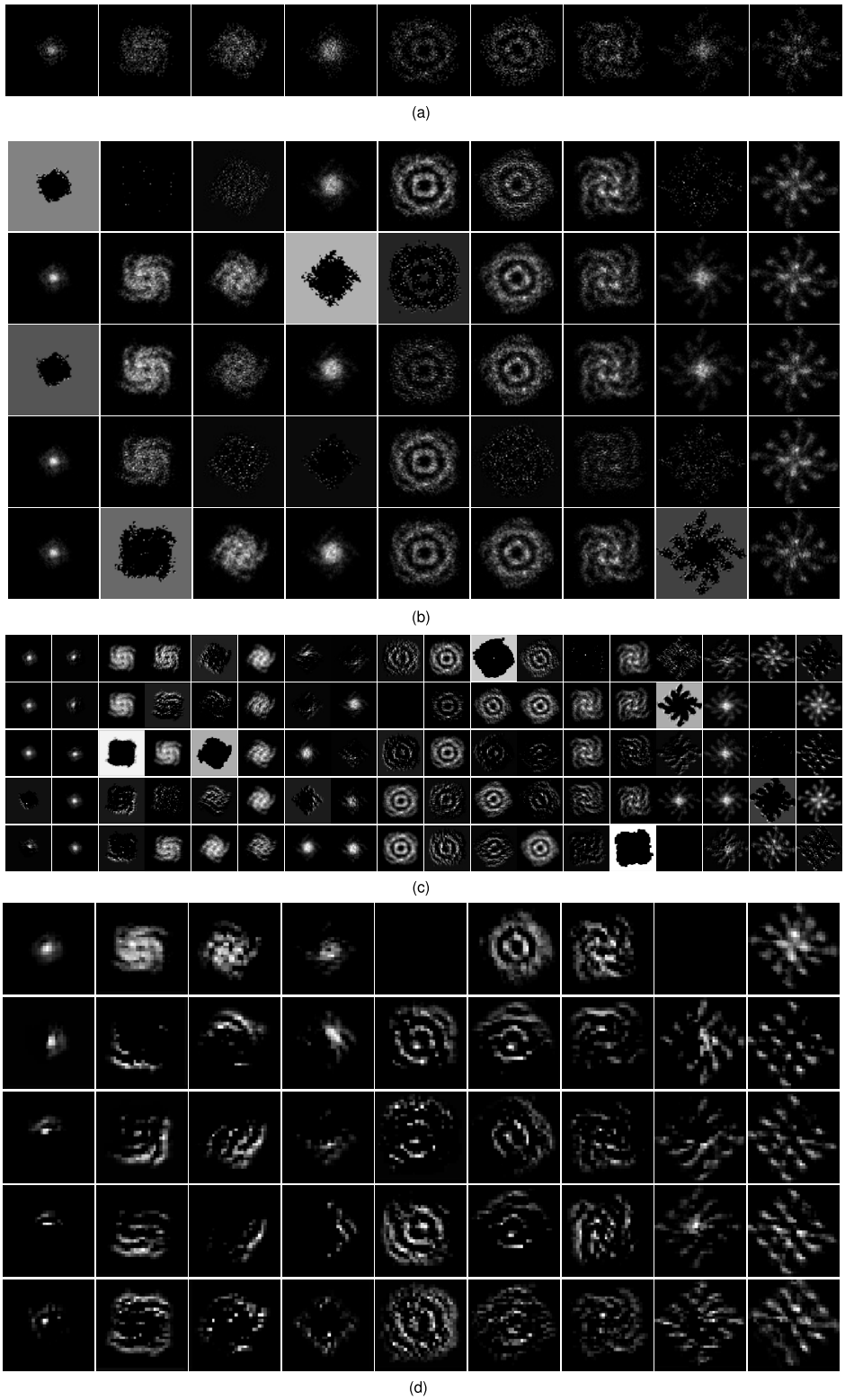

FIGURE 15. Visualization of the MCBLDN classifier with $N_F = 5$. (a) The input layer; (b) The first convolutional layer; (c) The second convolutional layer; (d) The third convolutional layer.

C. PARAMETERS SETTING OF MCBLDN BY VISUALIZATION

It is universally acknowledged that neural networks are often treated as black boxes as it is very hard to analyze them theoretically. In this subsection, we show that from visualization results, the setting of some hyper-parameters can be optimized. We focus on setting the number of convolutional filters in each layer of the parallel CNN part in the proposed MCBLDN scheme as it is a significant network parameter related to the classification performance. In particular, as shown in Fig. 6(b), it is the parameter N_F that

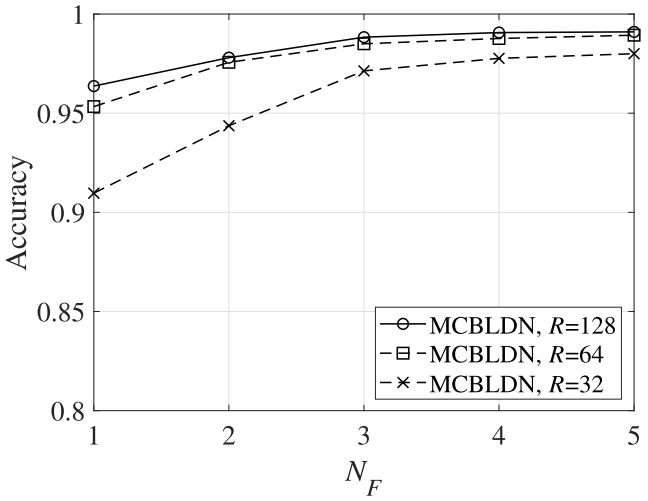

FIGURE 16. Classification accuracy versus N_F for the proposed MCBLDN scheme.

decides the number of convolutional filters. We investigate the influence of different values of N_F on the classification accuracy via experiments and visualization. The system set up is the same as that in Fig. 7 except that the SNR is fixed at 10dB with 48000 training data sets.

Fig. 14 and Fig. 15 demonstrate the visualization results of the three convolutional layers of the CNN part of the MCBLDN for $N_F = 1$ and 5, respectively, where the meaning of each sub-figure is the same as that in Fig. 13 for $N_F = 3$. That is, sub-figure (a) shows the input images of all $L = 9$ parallel CNNs, and sub-figures (b)-(d) show the output features of the three convolutional layers of all CNNs column-by-column. That is, the output features of the three convolutional layers of the l-th CNN are demonstrated in the l-th column in sub-figure (b), the $(2l-1)$-th and $2l$-th columns in sub-figure (c), and the l-th column in sub-figure (d), respectively.

From Fig. 14 for $N_F = 1$, we can see that the output features of different CNNs are different from each other, which shows that the features extracted by the CNNs with $N_F = 1$ are probably not enough. From Fig. 15 for $N_F = 5$, we can see that the CNNs may extract duplicate features and even some features with useless information. There are some completely black images in the output, which means the pixel values of these images are all 0. For example, the one at the first row and the fifth column and the one at the first row and the eighth column of Fig. 15(d). This phenomenon shows that there is probably redundancy in the network structure. Finally, comparing these two figures to Fig. 13 for $N_F = 3$, we can see that there is almost no duplicate feature information or all-black feature when $N_F = 3$. Thus, from the above results, $N_F = 3$ is the best choice to balance the classification accuracy and network complexity.

This conclusion is further verified by Fig. 16, which shows the classification accuracy performance as a function of N_F. It can be seen that the classification accuracy increases as N_F increases. When N_F increases from 1 to 3, the classification accuracy increases rapidly. However, there is no significant performance improvement when N_F is larger than 3. In Fig. 16, we also provide the results for $R = 64$ and $R = 32$, where similar trend is observed. Note that R determines the size of the input sub-images and in turn determines the network complexity as shown in Tab. 3. Thus, by combining the results in Tab. 3 and Fig. 16, we can see that $R = 64$ may be the best choice considering the performance and complexity trade-off.

VII. CONCLUSION

The channel time variation in practical systems greatly affects the AMC accuracy rate and makes the AMC problem more difficult to deal with. In this article, we have taken the CD as the key feature and tried to utilize the rotation and scaling trajectory of CD in time-varying channels for better classification performance. Specifically, we have first proposed the slotted-CD feature extraction scheme, which outputs a series of sub-images that can reflect the CD time variation both within each sub-image and between the consecutive sub-images. We have then proposed an advanced classifier, the MCBLDN, by using the deep learning technology. The classifier first employs multiple parallel CNNs to process the sub-images of the slotted-CD extractor separately and then employs the BLSTM to exploit the relationship among the sub-images for modulation classification. Simulation results have shown that the MCBLDN scheme can achieve much higher classification accuracy than the traditional deep learning based AMC schemes in time-varying channels. In particular, the MCBLDN scheme can achieve an accuracy rate of 97% in fast fading channels. Furthermore, we have shown that such performance improvement can be explained via the visualization of the outputs of the convolutional layers. We have also shown that some key hyper-parameters such as N_F can be optimized from the visualization results.

As for future work, we note that besides CD there are also some other features which can be used for AMC, for example, the cyclostationary feature [17], [18], which is sensitive to frequency variation and suitable to classify frequency modulations. Thus, if the modulation set for AMC not only includes bandwidth efficient modulations such as high-order PSK and QAM, but also power efficient ones such as frequency-shift keying modulation, a more sophisticated AMC scheme taking both the CD and the cyclostationary feature can be developed, which is a promising design approach for future work.

ACKNOWLEDGMENT

The authors would like to thank the editor and the reviewers for their valuable suggestions and comments for improving the paper.

REFERENCES

[1] Z. Zhu and A. K. Nandi, *Automatic Modulation Classification: Principles, Algorithms and Applications*. Hoboken, NJ, USA: Wiley, 2015.

[2] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: Classical approaches and new trends," *IET Commun.*, vol. 1, no. 2, pp. 137–156, Apr. 2007.

[3] Y. Ma, Y. Gao, Y.-C. Liang, and S. Cui, "Reliable and efficient sub-nyquist wideband spectrum sensing in cooperative cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 10, pp. 2750–2762, Oct. 2016.
[4] B. A. Fette, *Cognitive Radio Technology*. Amsterdam, The Netherlands: Elsevier, 2006.
[5] A. Tsakmalis, S. Chatzinotas, and B. Ottersten, "Automatic modulation classification for adaptive power control in cognitive satellite communications," in *Proc. 13th Signal Process. Space Commun. Workshop*, Sep. 2014, pp. 234–240.
[6] A. Tsakmalis, S. Chatzinotas, and B. Ottersten, "Centralized power control in cognitive radio networks using modulation and coding classification feedback," *IEEE Trans. Cognit. Commun. Netw.*, vol. 2, no. 3, pp. 223–237, Sep. 2016.
[7] J. L. Xu, W. Su, and M. Zhou, "Likelihood-ratio approaches to automatic modulation classification," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 41, no. 4, pp. 455–469, Jul. 2011.
[8] M. Abu-Romoh, A. Aboutaleb, and Z. Rezki, "Automatic modulation classification using moments and likelihood maximization," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 938–941, May 2018.
[9] W. Chen, Z. Xie, L. Ma, J. Liu, and X. Liang, "A faster maximum-likelihood modulation classification in flat fading non-Gaussian channels," *IEEE Commun. Lett.*, vol. 23, no. 3, pp. 454–457, Mar. 2019.
[10] H. Zhang, Y. Wang, L. Xu, T. Aaron Gulliver, and C. Cao, "Automatic modulation classification using a deep multi-stream neural network," *IEEE Access*, vol. 8, pp. 43888–43897, 2020.
[11] A. K. Nandi and E. E. Azzouz, "Algorithms for automatic modulation recognition of communication signals," *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 431–436, Apr. 1998.
[12] M. A. Azza, A. E. Moussati, and O. Moussaoui, "Implementation of an automatic modulation recognition system on a software defined radio platform," in *Proc. Int. Symp. Adv. Electr. Commun. Technol. (ISAECT)*, Rabat, Morocco, Nov. 2018, pp. 1–4.
[13] Y. Ettefagh, M. H. Moghaddam, and S. Eghbalian, "An adaptive neural network approach for automatic modulation recognition," in *Proc. 51st Annu. Conf. Inf. Sci. Syst. (CISS)*, Baltimore, MD, USA, Mar. 2017, pp. 1–5.
[14] F. Benedetto, A. Tedeschi, and G. Giunta, "Automatic blind modulation recognition of analog and digital signals in cognitive radios," in *Proc. IEEE 84th Veh. Technol. Conf.*, Montreal, QC, Canada, Sep. 2016, pp. 1–5.
[15] W. Xie, S. Hu, C. Yu, P. Zhu, X. Peng, and J. Ouyang, "Deep learning in digital modulation recognition using high order cumulants," *IEEE Access*, vol. 7, pp. 63760–63766, 2019.
[16] R. Li, L. Li, S. Yang, and S. Li, "Robust automated VHF modulation recognition based on deep convolutional neural networks," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 946–949, May 2018.
[17] G. J. Mendis, J. Wei, and A. Madanayake, "Deep learning-based automated modulation classification for cognitive radio," in *Proc. IEEE Int. Conf. Commun. Syst. (ICCS)*, Shenzhen, China, Dec. 2016, pp. 1–6.
[18] G. J. Mendis, J. Wei-Kocsis, and A. Madanayake, "Deep learning based radio-signal identification with hardware design," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 5, pp. 2516–2531, Oct. 2019.
[19] S. Jeong, U. Lee, and S. C. Kim, "Spectrogram-based automatic modulation recognition using convolutional neural network," in *Proc. 10th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Prague, Czech Republic, Jul. 2018, pp. 843–845.
[20] Q. Zhang, Z. Xu, and P. Zhang, "Modulation recognition using wavelet-assisted convolutional neural network," in *Proc. Int. Conf. Adv. Technol. Commun. (ATC)*, Ho Chi Minh City, CO, USA, Oct. 2018, pp. 100–104.
[21] Y. Zeng, M. Zhang, F. Han, Y. Gong, and J. Zhang, "Spectrum analysis and convolutional neural network for automatic modulation recognition," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 929–932, Jun. 2019.
[22] 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.
[23] 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.
[24] T. Bo, J. Tang, and C.-K. Chan, "Modulation format recognition for optical signals using connected component analysis," *IEEE Photon. Technol. Lett.*, vol. 29, no. 1, pp. 11–14, Jan. 1, 2017.
[25] L. Xie and Q. Wan, "Automatic modulation recognition for phase shift keying signals with compressive measurements," *IEEE Wireless Commun. Lett.*, vol. 7, no. 2, pp. 194–197, Apr. 2018.
[26] L. Xie and Q. Wan, "Cyclic feature-based modulation recognition using compressive sensing," *IEEE Wireless Commun. Lett.*, vol. 6, no. 3, pp. 402–405, Jun. 2017.
[27] S. Hassanpour, A. M. Pezeshk, and F. Behnia, "Automatic digital modulation recognition based on novel features and support vector machine," in *Proc. 12th Int. Conf. Signal-Image Technol. Internet-Based Syst. (SITIS)*, Naples, Italy, 2016, pp. 172–177.
[28] X. Zhang, T. Ge, and Z. Chen, "Automatic modulation recognition of communication signals based on instantaneous statistical characteristics and SVM classifier," in *Proc. IEEE Asia–Pacific Conf. Antennas Propag. (APCAP)*, Auckland, Zealand, Aug. 2018, pp. 344–346.
[29] F. Yang, L. Yang, D. Wang, P. Qi, and H. Wang, "Method of modulation recognition based on combination algorithm of K-means clustering and grading training SVM," *China Commun.*, vol. 15, no. 12, pp. 55–63, Dec. 2018.
[30] X. Zhang, J. Sun, and X. Zhang, "Automatic modulation classification based on novel feature extraction algorithms," *IEEE Access*, vol. 8, pp. 16362–16371, 2020.
[31] M. Waqar Aslam, Z. Zhu, and A. Kumar Nandi, "Automatic modulation classification using combination of genetic programming and KNN," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2742–2750, Aug. 2012.
[32] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
[33] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
[34] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11–26, Apr. 2017.
[35] G. Hinton, L. Deng, D. Yu, G. Dahl, A.-R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath, and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.
[36] A. Ioannidou, E. Chatzilari, S. Nikolopoulos, and I. Kompatsiaris, "Deep learning advances in computer vision with 3D data: A survey," *ACM Comput. Surveys*, vol. 50, no. 2, pp. 1–38, Jun. 2017.
[37] R. Socher, Y. Bengio, and C. D. Manning, "Deep learning for NLP (without magic)," in *Proc. Tutorial Abstr. ACL*, 2012, p. 5.
[38] W. Shi, D. Liu, X. Cheng, Y. Li, and Y. Zhao, "Particle swarm optimization-based deep neural network for digital modulation recognition," *IEEE Access*, vol. 7, pp. 104591–104600, 2019.
[39] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 3, pp. 433–445, Sep. 2018.
[40] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *Proc. Int. Conf. Eng. Appl. Neural Netw.*, Aberdeen, U.K., 2016, pp. 213–226.
[41] 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.
[42] Y. Wang, J. Wang, W. Zhang, J. Yang, and G. Gui, "Deep learning-based cooperative automatic modulation classification method for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4575–4579, Apr. 2020.
[43] D. Wang, M. Zhang, Z. Li, J. Li, M. Fu, Y. Cui, and X. Chen, "Modulation format recognition and OSNR estimation using CNN-based deep learning," *IEEE Photon. Technol. Lett.*, vol. 29, no. 19, pp. 1667–1670, Oct. 1, 2017.
[44] M. Zhang, M. Diao, and L. Guo, "Convolutional neural networks for automatic cognitive radio waveform recognition," *IEEE Access*, vol. 5, pp. 11074–11082, 2017.
[45] Y. Kumar, M. Sheoran, G. Jajoo, and S. K. Yadav, "Automatic modulation classification based on constellation density using deep learning," *IEEE Commun. Lett.*, vol. 24, no. 6, pp. 1275–1278, Jun. 2020.
[46] D. Tse and P. Viswanath, *Fundamentals Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[47] T. S. Rappaport, *Wireless Communication: Principle and Practice*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.
[48] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, Jul. 1998.
[49] L. Zhao, D. W. K. Ng, and J. Yuan, "Multi-user precoding and channel estimation for hybrid millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1576–1590, Jul. 2017.

YU ZHOU (Student Member, IEEE) received the B.S. degree in computer science and engineering from Northeastern University, China, in 2018. He is currently pursuing the M.S. degree in communication science and engineering with Fudan University. His current research interests include automatic modulation classification, deep learning, and millimeter wave signal processing.

TIAN LIN (Student Member, IEEE) received the B.Eng. degree (Hons.) in communication science and engineering from Fudan University, in 2017, where he is currently pursuing the Ph.D. degree. His current research interests include hybrid beamforming for massive MIMO systems, millimeter wave signal processing, passive beamforming and channel estimation for intelligent reflecting surface-aided MIMO systems, and deep learning for physical layer communication.

YU ZHU (Member, IEEE) received the B.Eng. degree (Hons.) in electronics engineering and the M.Eng. degree (Hons.) in communication and information engineering from the University of Science and Technology of China, in 1999 and 2002, respectively, and the Ph.D. degree from the Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, in 2007.

Since 2008, he has been with Fudan University, where he is currently a Professor with the School of Information Science and Technology. His current research interests include broadband wireless systems and networks and signal processing for communications. He received the Shanghai Pujiang Scholar Award, in 2008, and the Fudan Zhuoxue Award, in 2012. He served as the Physical Track Co-Chair of the IEEE WCNC 2013, the Co-Chair of the Wireless Communications Systems Symposium of the IEEE ICCC 2014, IEEE ICCC 2019, and the Workshop Co-Chair of the Asia–Pacific Conference on Communications, in 2017. He served as an Editor for the IEEE WIRELESS COMMUNICATIONS LETTERS. He is also an Editor of the *Journal of Communications and Information Networks*.