IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 29 July 2023, accepted 22 September 2023, date of publication 26 September 2023, date of current version 2 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3319385

RESEARCH ARTICLE

Cyclostationary Feature Based Modulation Classification With Convolutional Neural Network in Multipath Fading Channels

LIYAN YIN, XIN XIANG, AND YUAN LIANG
Department of Electronic Engineering, Aviation Engineering College, Air Force Engineering University, Xi'an 710038, China
Corresponding author: Liyan Yin (13379547517@126.com)

This work was supported in part by the Natural Science Foundation of Shaanxi Province under Grant 2021JM-220.

ABSTRACT Modulation classification has been widely studied in recent years. However, few studies focus on the performance degradation in multipath fading channels, whose impact is non-negligible. In this paper, a convolutional neural network (CNN) employing cyclostationary (CS) feature, which maintain the essential characteristics in fading channels, is proposed for robust modulation classification. Our method can be implemented in two approaches, referred as CASE1 and CASE2. In CASE1, a single-structured CNN is designed for learning hybrid CS features to perform classification. And in CASE2, we present a CNN model based on a hierarchical structure to perform two-stage classification. Specifically, the coarse classification is performed by learning the second-order CS features with the first-level CNN. Next, another CNN can be selectively activated to learn from high-order CS features for fine classification within the subclass. In this way, our method uses CS features to provide favorable guidance for the learning process of CNN, thus improving the classification performance in fading channels. The experimental results demonstrate the advantages of the proposed method in terms of classification accuracy and computational complexity.

INDEX TERMS Modulation classification, cyclostationary features, convolutional neural network, multipath fading channels.

I. INTRODUCTION

Modulation classification is a key technology in modern wireless communication, and it receives widespread attention while facing numerous challenges. In spectrum monitoring, modulation classification often needs to provide more information than traditional spectrum sensing. Generally, spectrum sensing algorithms based on energy detection, matched filtering, and covariance-based detection only need to determine whether a channel is available or not [1], [2], [3]. Then, cognitive radio users can use these channels free of charge, and they can share access with licensed users without interfering with them [4], [5]. Besides searching for unoccupied channels like in traditional spectrum sensing algorithms, the modulation classifier also needs to distinguish different modulation types in the channel. Cognitive nodes can exploit this information to obtain more appropriate statistical data, make decisions based on detected signals, and distinguish their network from other transmissions. In the military field, modulation classification needs to obtain the modulation information of wireless signals in the battlefield environment with little or no prior information, thereby implementing communication interference and interference countermeasures in a targeted manner [6], [7]. However, an important issue in practical applications of modulation classification is its susceptibility to the impact of signal transmission environment [8], [9], [10]. In scenarios such as wideband wireless communication and mobile communication, multipath fading channels greatly degrades the classification performance. As in stationary signal analysis, signals passing through fading channels make their features close to those of noise, thereby reducing its separability.

The cyclostationary (CS) feature based method has considerable potential to solve the above problem [11], [12].

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

In a short observation time, both signal and fading channel can be modeled as cyclostationary process. On this premise, the designed cyclostationary features are insensitive to the channel and differ significantly from the features of noise. In addition, CS features have an significant advantage is that it is insensitive to unknown signal shaping and have the ability to preserve the phase information [13], [14], so there is no need for a rigorous pre-processing. C. M. Spooner was the first to exploit the difference in cyclic cumulants (CCs) to determine the modulation type in unknown signals [15]. Subsequently, the classification capacity of CC of each order was analyzed, as reported in [16], [17], and [18]. Like et al. combined multi-antenna technique with CS features to distinguish various signals such as ASK, PSK, and CDMA in multipath channels [19], [20]. Gardner et al. was the first to analyze the spectral correlation function (SOF) of various frequency band modulation signals, and it was pointed out that the classification of signals can be completed by using the spectral correlation characteristics of signals [13], [21], [22]. Then, Headley et al. used spectrum-correlated cyclic frequency slices and combined them with a 4-layer linear perception network to distinguish QPSK, MSK, and AM signals in fading channels [23].

In recent years, with its powerful feature representation and classification capacity, deep learning (DL) technique has been widely applied in many fields. As for modulation classification, the methods with DL tend to replace traditional methods with feature extraction. O'Shea first applied DL to modulation classification, used convolutional neural networks (CNNs) to train radio time series, and obtained good results [24], [25]. The long short-term memory (LSTM) neural network has unique advantages in processing time series signals, which can further enhance the performance of modulation classification [26], [27], [28]. Wei et al. compared the modulation classification performance of three types of models, including convolutional long short-term deep neural network (CLDNN), residual network (ResNet), and densely connected network (DenseNet). Among them, the best model was nearly 13.5% more accurate than a 6-layer CNN model under the same conditions [29]. In fact, the end-to-end learning with DL from signal discrete sequence often requires samples of high quality and quantity; in contrast, traditional methods with feature extraction are difficult for pre-processing and have poor robustness to environmental changes. Therefore, a combined scheme, i.e., extracting shallow features artificially, and guiding the deep network to learn deep features and then make decisions, is a reliable and stable means of approach.

Based on the above facts, this paper proposes a novel modulation classification method based on CS features with CNN in multipath fading channels. Specifically, we select features based on SOF and CC to make the classifier input robust to the channel. Then, modulation classification is achieved from the aspects of two cases. In CASE 1, a CNN with a single structure is designed to learn the hybrid features, including the profiles of SOF and CCs with zero delay, And, in CASE 2, a two-level tree-structured CNN model is designed. The first level CNN learns the profiles of SOF to perform coarse classification. Then the classifier selectively activates the second-level CNN, and uses the zero-delay CCs to perform fine classification in the subclass. In this way, the proposed method can achieve classification in fading channels without using prior knowledge such as carrier frequency and symbol rate. Hence, the main contributions of this paper include:

i) *Innovation in classification method:* With the help of powerful feature extraction and learning capabilities of CNN, adaptive decision-making, rather than statistical decision-making, is used for modulation classification.

ii) *Innovation in network input:*To combat the adverse effects of multipath fading channels, CS features instead of signal discrete sequence are used as input to the network. In addition, the form of low complexity input is studied.

iii) *Innovation in numerical studies*: We investigate the classification performance of the proposed method in various channels, and further analyze the contribution of feature extraction and classifier design on improving the performance in multipath fading channels.

The rest of this paper is organized as follows. Section II presents the signal model and derives the candidate CS features for classification. Section III proposes the CNN model for modulation classification and provides the implementation method. Section IV conducts numerical analysis. Finally, the conclusion of this paper is given in Section V.

II. ANALYSIS OF CS FEATURES IN A FREQUENCY SELECTIVE FADING CHANNELS

A. A SIGNAL MODEL

The received signal passing through the fading channel can be represented as:

$$y(t) = e^{j2\pi f_c t} e^{j\phi} x(t) \otimes h(t) + w(t), \tag{1}$$

where $x(t)$ and $y(t)$ represent the transmitted and received signals, respectively. $\otimes$ is the convolution operation, $w(t)$ is the additive Gaussian noise, f_c and ϕ represent the carrier frequency and phase offset, respectively. In this paper, the channel is modeled as a multipath fading channel, so the channel impulse response $h(t)$ can be written as:

$$h(t) = \sum_{l=0}^{L-1} \alpha_l e^{j\varphi_l} \delta(t - t_l), \tag{2}$$

where L is the number of channel taps, a_l, φ_l, and t_l are the amplitude, phase, and tap delay of the channel, respectively. If $x(t)$ is a digital modulated signal, (1) can be further expressed as:

$$y(n) = e^{j2\pi f_c n} e^{j\phi} \sum_{l=1}^{L} a_l e^{j\varphi_l} \sum_k s_k g(n - k - k_l) + w(n), \tag{3}$$

where $g(n) = p_{rx}(n) \otimes h(n) \otimes p_{tx}(n)$ represents a combination of the impulse responses of receiver, channel and transmitter

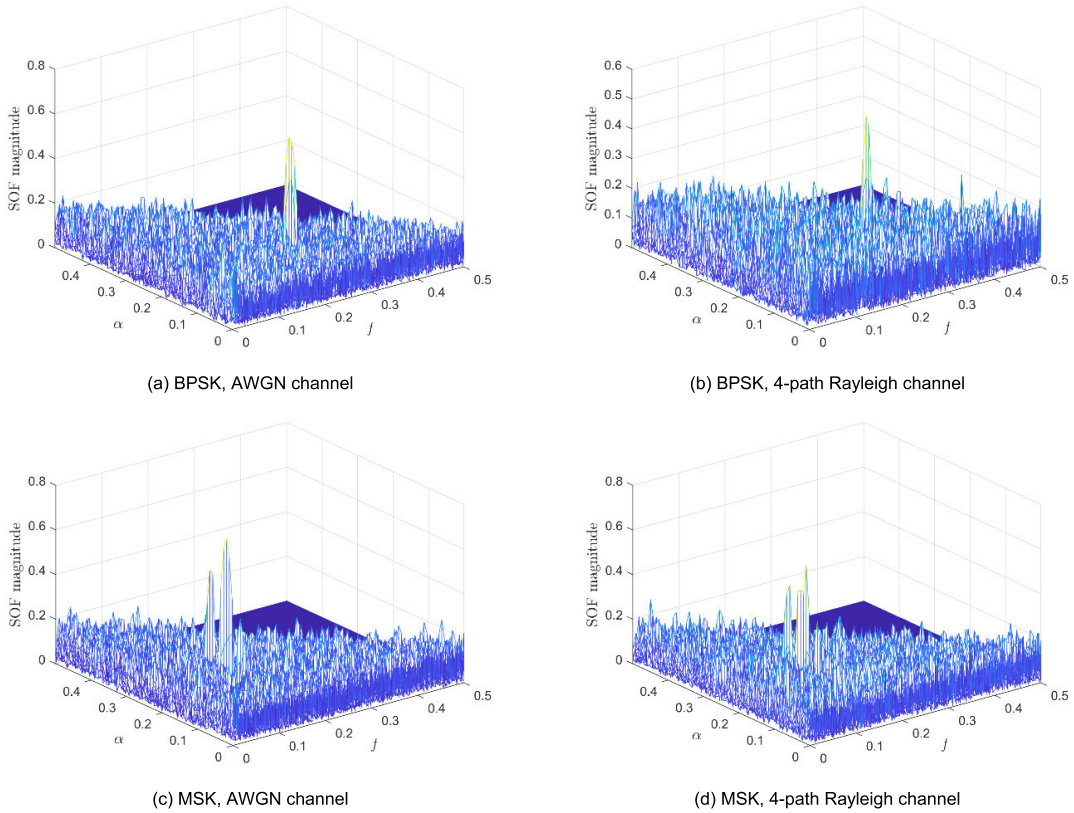

FIGURE 1. SOF of two modulated signal in the AWGN channel and 4-path Rayleigh channel, where the SNR is 5 dB, the symbol period is $1/T_b = f_s/5$, the carrier frequency is $f_c = f_s/4$, and the number of samples 4096.

filter, $k_l = t_l/T_s$ and T_s represents the index of time delay and the symbol duration, respectively. Besides, s_k is within the range of $n \in (k - 1/2, k + 1/2)$, and assumed to be a zero-mean normally distributed random vector.

B. SECOND-ORDER CS FEATURES

Under the premise of cyclic periodicity, the spectral correlation function (SCF) of the signal $x(n)$ can be expressed as:

$$S_x^\alpha(f) = \lim_{U\to\infty}\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} X_U(n, f+\frac{\alpha}{2})X_U^*(n, f-\frac{\alpha}{2}), \quad (4)$$

$$X_U(n,f) = \sum_{n-U/2}^{n+U/2} x(v)e^{j2\pi fv}, \quad (5)$$

where α is the Fourier component.

Here, it can be seen that $S_x^\alpha(f)$ is a measurement of the correlation between the spectral components of $x(t)$. A significant advantage using SCF is its insensitivity to AWGN. Since the spectral components of AWGN are uncorrelated, it does not affect the SCF for any value of $\alpha \neq 0$.

The SOF acts as a normalized version of SCF, which can be defined as follows:

$$C_x^\alpha(f) = \frac{S_x^\alpha(n,f)}{\sqrt{S_x^0(n, f+\alpha/2)S_x^{0*}(n-\alpha/2)}}. \quad (6)$$

It can be seen that, the value of $C_x^\alpha(f)$ is normalized between 0 and 1. Next, we consider the convolution of the signal and channel, which can be regarded as a linear time invariant (LTI) system transformation:

$$x_h(n) = x(n) \otimes h(n). \quad (7)$$

The SCF of the input and output of LTI system is defined by:

$$S_{x_h}^\alpha(f) = H(f+\frac{\alpha}{2})S_x^\alpha(f)H^*(f-\frac{\alpha}{2}), \quad (8)$$

where $H(f)$ is the channel frequency response. By substituting (11) into the definition of SOF in (9), it is given that:

$$C_{x_h}^\alpha(f) \equiv C_x^\alpha(f). \quad (9)$$

If the additive noise term in (1) is ignored, the influence of frequency, phase shift, and time delay in the received signal can be eliminated by taking the amplitude of SOF, which can be described as:

$$\left|C_y^\alpha(f)\right| \equiv \left|C_x^\alpha(f)\right|. \quad (10)$$

Therefore, the amplitude of SOF is robust to the channel effect.

Fig. 1 shows the SOF amplitude of the modulated signal under AWGN channel and 4-path Rayleigh fading channel. It can be found that the amplitude of SOF can effectively resist the influence of channel effects, where the amplitude

of SOF of the modulated signal passing through the multipath fading channel is approximately equal to that of the original undistorted signal. As long as the frequency of the signal of interest is not completely offset by the channel, the amplitude of SOF can retain reliable signal characteristics. Therefore, robust modulation classification in fading channels can be achieved by checking the pattern in the amplitude of SOF.

C. HIGH-ORDER CS FEATURES

Although SOF generates different amplitudes for different modulation types, certain modulation types (e.g., PSK/QAM-type modulation) can generate the similar pattern. This leads to the actual loss of distinguishability for signals in the modulation pool. To solve this problem, higher-order cyclic statistics (HOCS) is introduced. Concerning this, the p-order/q-conjugate temporal cumulant (TC) is used, which is obtained through:

$$\begin{aligned} C_x(n,s)_{p,q} &= \mathrm{Cum}\{x^{(*)_1}(n+s_1)\cdots x^{(*)_p}(n+s_p)\} \\ &= \sum_{\{P_n\}}(-1)^{Z-1}(Z-1)!\prod_{z=1}^{Z} m_x(n,s_z)_{p_z,q_z}, \end{aligned} \tag{11}$$

where m_x is the moment of the signal, $\{P_n\}$ is the set of different partitions of $\{1, 2, \ldots, p\}$; s_z is a delay vector with an index specified by z; p_z and q_z correspond to the number of elements and the number of conjugate terms in the subset P_z, respectively. When calculating TC, the impact of the lower-order moments is effectively alleviated, leaving only the residual impact of the current order. Note that the TC is also a periodic function of a CS signal, whose Fourier components are given by:

$$C_x^{\gamma}(s)_{p,q} = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} C_x(n,s)_{n,q}e^{-j2\pi\lambda nT_s}, \tag{12}$$

where $C_x^{\gamma}(s)_{p,q}$ is defined as the CC of $x(n)$.

By substituting (3) into (11) and (12), the set of values of the cyclic frequency for which CC is not constant to 0 can be expressed as:

$$\kappa_{n,q} = \{\gamma \,|\, \gamma = \beta + (p-2q)f_c,\ \beta = i/T_s, i \text{ integer}\}. \tag{13}$$

At this time, the value of CC is:

$$\begin{aligned} C_y^{\gamma}(s)_{p,q} &= C_{n,q}T_s^{-1}e^{j2\pi f_c\sum_{m=1}^{n}(-)_m s_m}e^{j(n-2q)\phi} \\ &\quad\times\sum\nolimits_n\prod_{m=1}^{n} g^{(*)_m}(n+s_m)e^{-j2\pi\beta nT_s} + C_w^{\gamma}(s)_{n,q}, \end{aligned} \tag{14}$$

where $C_{p,q}$ is the p-order/q-conjugate cumulant of the stationary discrete data sequence, and the possible negative sign $(-)$ comes from one of the q-conjugate $(*)_q$. Thus, the value of CC of the received signal is proportional to $C_{p,q}$, whose value is well known for the common modulation types, which is listed in Tab. 1.

From (13) and (14), we can see that the spectral peaks of the CCs of different modulations appear at different cyclic frequencies in frequency selective fading channels. Therefore, robust modulation classification can be achieved by checking the cyclostationary of the CCs of the modulated signals on the cyclic frequency axis.

TABLE 1. Theoretical values of cumulants.

$C_{p,q}$	BPSK	QPSK	16QAM	64QAM	OFDM
$C_{4,0}$	-2	1	-0.68	-0.62	0
$C_{4,1}$	-2	0	0	0	0
$C_{4,2}$	-2	-1	-0.68	-0.62	0
$C_{6,0}$	16	0	0	0	0
$C_{6,1}$	16	-4	2.08	1.80	0
$C_{6,2}$	16	0	0	0	0
$C_{6,3}$	16	4	2.08	1.80	0
$C_{8,0}$	-272	-34	-13.98	-11.50	0
$C_{8,1}$	-272	0	0	0	0
$C_{8,2}$	-272	34	-13.98	-11.50	0
$C_{8,3}$	-272	0	0	0	0
$C_{8,4}$	-272	-34	-13.98	-11.50	0

III. CS FEATURE BASED MODULATION CLASSIFICATION WITH CNN

A. FEATURE SELECTION AND SAMPLE CONSTRUCTION

Due to the fact that the amplitude of SOF is a three-dimensional (3-D) image, this provides an unreasonable amount of data for real-time operation of the classifier. Therefore, it is necessary to reduce the dimension of feature input. In previous studies, it has been proved that in the case of minimum computational complexity and maximum information retention, the profiles of cyclic frequency and frequency can be used to create a pseudo-3D image of SOF, which can perform classification with significantly higher reliability [30]. Therefore, the candidate features based on second-order CS used for classification in this paper are classified as cyclic frequency profile:

$$\Theta(\alpha) = \max_f \left|C_x^{\alpha}(f)\right|, \tag{15}$$

and frequency profile:

$$\Xi(f) = \max_{\alpha} \left|C_x^{\alpha}(f)\right|. \tag{16}$$

For high-order CS features in (20), the noise term is 0, and taking its absolute value can eliminate the correlation with carrier frequency, phase, and timing offset:

$$\begin{aligned} \Gamma_y^{\gamma}(s)_{p,q} &= \left|C_{p,q}\right| T_s^{-1}\left|\sum\nolimits_n\prod_{m=1}^{n} g^{(*)_m}(n+s_m)e^{-j2\pi\beta nT_s}\right|, \\ \gamma &= \beta + (p-2q)f_c, \beta = i/T_s, i \text{ integer}. \end{aligned} \tag{17}$$

It can be noted that, (17) is proportional to the following function:

$$\Phi(p,q,s,\gamma,r) = \left|\sum\nolimits_n\prod_{m=1}^{n} g^{(*)_m}(n+s_m)e^{-j2\pi\beta nT_s}\right|. \tag{18}$$

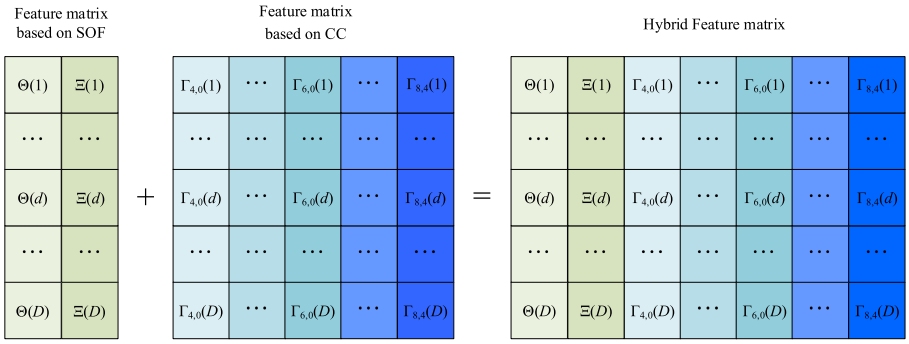

FIGURE 2. Form of the hybrid feature matrix.

It can be seen that this function depends on the order of CC p, the number of conjugates q, delay vector s, cyclic frequency γ and roll-off coefficient r. If the raised cosine shaping pulse is assumed, the maximum value of the obtained function $\Phi(p, q, s, \gamma, r)$ appears at $s = \vec{0}_n$, where $\vec{0}_n$ is the n-dimensional zero vector. And at that time, the function decreases as γ increases. Therefore, a vector with zero delay, denoted as $\Gamma_y^{\gamma}(\overrightarrow{0}_n)_{n,q}, n = 2, 4, 6, 8, q \leq n/2, q$ integer, is used as a candidate feature. It is worth noting that, our classification method uses deep learning instead of statistical decision-making. Therefore, there is no need to calculate non-zero discrete values at specific cycle frequencies, and thus $\Gamma_{p,q}^{\gamma}$ is directly input as feature vector into the classifier.

In addition, subcarriers in OFDM systems can be modeled as independently modulated signals, where each signal has its own second-order CS features. However, due to the ''destructive interference'' between overlapping CS features, their bandwidth overlap actually reduces the discrimination of the observed SOF between OFDM and single carrier modulation, especially at low SNRs [31]. Besides, OFDM exhibits significant distinguishability from single carrier modulation in high-order CS features. Therefore, our method uses features based on CC to distinguish OFDM signals from single-carrier modulated signals.

Next, the selected features need to be constructed as samples suitable for classifier input, and the two types of feature vectors mentioned above are rearranged to form a hybrid feature matrix. The feature values obtained from each calculation are arranged in rows, and different feature vectors are distributed in columns. It should be noted that the same number of indexes D are used during the calculation of two types of feature vectors. Finally, we merge the two feature matrices into a mixed feature matrix by column, and the construction process is shown in Fig. 2.

B. CNN MODEL FOR MODULATION CLASSIFICATION

Our purpose is to that using the CS features to classify nine modulated signals, including 2ASK, 2FSK, 4FSK, MSK, BPSK, QPSK, 16QAM, 64QAM, and OFDM. To this end, we consider a novel modulation classification method based on CNN, which has two different approaches, and the corresponding network architectures are shown in Fig. 3.

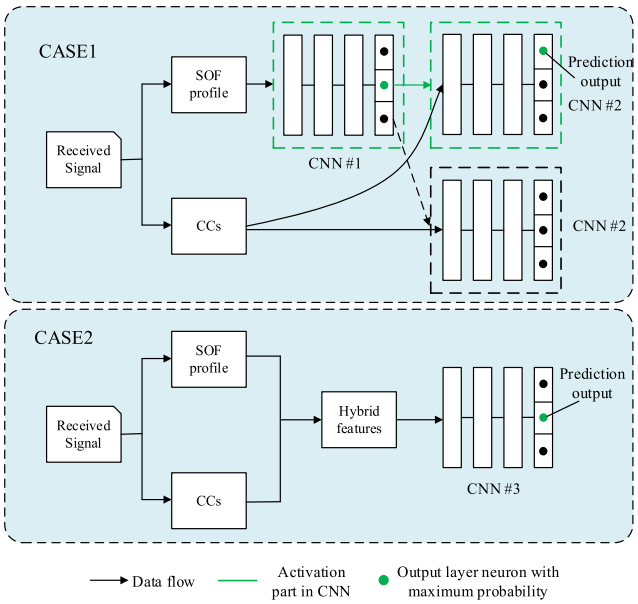

FIGURE 3. Illustration of two different approaches of the proposed method.

CASE1: A single-structured CNN is trained using all possible modulation types. The input of the network is in the form of the hybrid feature matrix described in Fig. 2. Then, the hybrid features are fed to the CNN classifier, which is pre-trained with the features of nine possible modulations. Finally, the network outputs the predicted results.

CASE2: A hierarchical CNN architecture is used to perform two-stage classification, which is shown in Fig. 4. In the first stage, CNN#1 using SOF profile feature input is trained to perform coarse classification in the given modulation, and the prediction outputs include 2ASK, 2FSK, 4FSK, MSK, and PSK/QAM/OFDM; if the coarse classifier predicts PSK/QAM/OFDM, CNN#2 using CC feature input is activated to perform the fine classification in the subclass, and the prediction outputs include BPSK, QPSK, 16QAM, 64QAM, OFDM.

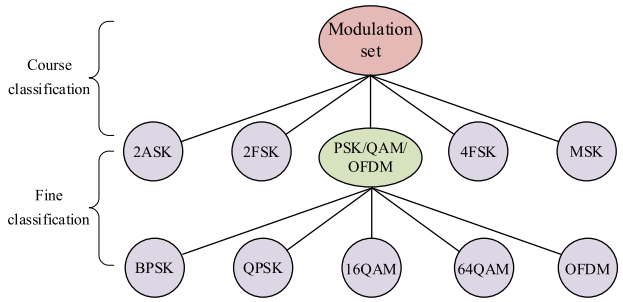

FIGURE 4. Illustration of tree structure of the two-stage classification.

Generally speaking, the performance of modulation classification is measured by using an accuracy indicator. First, the accuracy P_{CASE1} of CASE1 is defined as:

$$P_{\text{CASE1}} = P(\hat{\chi}_k = \chi_k), k \in [0, 8], \tag{19}$$

where χ_k represents the label sequence of the modulation type of the transmitted signal, and $\hat{\chi}_k$ is the predicted label sequence. The value of k ranges from 0 to 8, representing the modulation types of BPSK, QPSK, 16QAM, 64QAM, OFDM, 2ASK, 2FSK, 4FSK, and MSK, respectively. P_{CASE1} represents the accuracy of the nine types of modulation classification. For CASE2, two independent accuracy functions need to be defined, namely, coarse classification accuracy $P^{\text{C}}_{\text{CASE2}}$ and fine classification accuracy $P^{\text{F}}_{\text{CASE2}}$:

$$P^{\text{C}}_{\text{CASE2}} = \sum\nolimits_n P(\hat{\chi}_{\text{C}} = n|H_n), n \in [0, 4], \tag{20}$$

$$P^{\text{F}}_{\text{CASE2}} = P(\hat{\chi}_k = \chi_k|H_0), k \in [0, 4], \tag{21}$$

where H_n represents the hypothesis of the presence of 2ASK/2FSK/4FSK/MSK in the coarse classification, $\hat{\chi}_{\text{C}}$ is the prediction that the transmitted signal is χ_{C}, and $\hat{\chi}_k$ represents the prediction of the fine classification process in CASE2. For the transmitted modulated signal, χ_{C} can be defined as:

$$\chi_{\text{C}} = \begin{cases} 0, & k \in [0, 4] \\ k - 4, & k \in [5, 8] \end{cases} \tag{22}$$

The overall accuracy P_{CASE2} of CASE2 is shown in (29):

$$P_{\text{CASE2}} = P^{\text{C}}_{\text{CASE2}} P^{\text{F}}_{\text{CASE2}}. \tag{23}$$

C. IMPLEMENTATION OF THE CNN MODEL

The design and implementation of the CNN model used in this paper is based on the open-source learning library Keras [32]. As illustrated in Fig. 5, the proposed model consists of three convolutional layers, three pooling layers and two dense layers in sequence. The three convolutional layers have 64, 128, and 64 filters, respectively. There are two dense layers at the end of the network. The first hidden layer contains 128 neurons; the second hidden layer is determined by the number of output probabilities, consisting of 11, 8, and 5 neurons, applied to CNN#1, CNN#2, and CNN#3, respectively. A Leaky ReLU activation function with an alpha value of 0.1 is used in each convolutional layer to extract discriminative features. Our purpose in choosing

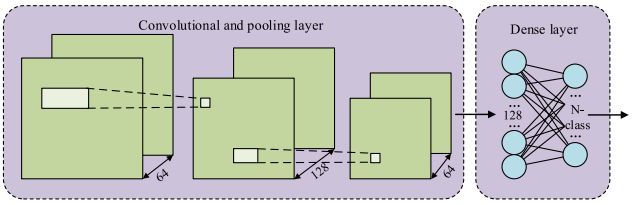

FIGURE 5. Illustration of the proposed CNN model.

Leaky ReLU is that, different from ReLU, it can map larger negative values to the smaller ones through a mapping line with a small slope. In each convolutional layer, 3×3 filters are used. Besides, 2×2 maximum pooling is used to reduce dimensionality and training time. A dense layer is then formed with 128 neurons and a Leaky ReLU activation function. After the dense layers, the probability class of each layer is calculated by the softmax activation function. In addition, we use the adaptive moment estimation (ADAM) optimizer, whose parameters are belta1 $= 0.5$, belta2 $= 0.999$, epsilon $= 1 \times 10^{-8}$, and the learning rate 1×10^{-4}. During the training stage, an early stopping mechanism is employed to prevent the model from over-fitting. Patience is selected as 10 epochs as the early stopping function, and validation loss is monitored during training. If the validation loss converges to a level and remains at this level for 10 epochs, the training is terminated and the weights at the end of the training stage are used in the test stage.

TABLE 2. Layout of the proposed CNN model.

Layers	Output dimension		
	CNN#1	CNN#2	CNN#3
Input	14×256	12×256	2×256
Conv1	64×14×256	64×12×256	64×2×256
Leaky ReLU1	64×14×256	64×12×256	64×2×256
Max Pool1	64×7×128	64×6×128	64×1×128
Conv2	128×7×128	128×6×128	128×1×128
Leaky ReLU2	128×7×128	128×6×128	128×1×128
Max Pool2	128×4×64	128×3×64	128×1×64
Conv3	64×4×64	64×3×64	64×1×64
Leaky ReLU3	64×4×64	64×3×64	64×1×64
Max Pool3	64×2×32	64×2×32	64×1×32
Flatten	4096	4096	2048
Dense1	128	128	128
Dense2	11	8	5

Tab. 2 shows the layout configuration of the proposed CNN model. As for the motivation behind this design, our first consideration is to use 64 filters with a size of 3×3 in the first convolutional layer to extract information about changes in the local region of the mapping output. Here, a smaller filter is selected for capturing peaks in the feature matrix

which creates local differences in the cycle frequency and latency axis [33]. After the cyclic properties of all local differences are determined, the second convolutional layer examines the attributes including size and location related to these cyclic properties. Herein, the better learning of these properties can be achieved by increasing the filter number to 128. As for the last convolutional layer, all attributes are converted to the mean of all information collected. Then, the number of neurons in the first dense layer is set to 128 so that enough information is acquired without over-fitting. Then, after flattening 3-D feature vectors into 1-D feature vectors through the flatten layer, two dense layers with 128 and 11 neurons are fed, and the number of neurons in the last layer is determined by the number of output probability vectors.

IV. SIMULATION EXPERIMENTS

A. SIMULATION SETTINGS

This section simulates based on the modulation types mentioned earlier. Except that the shaping of 2FSK uses the rectangular function, each modulated signal is modeled with an intermediate frequency (IF) uniformly distributed between sampling frequencies of 0.23 to 0.27 times, a symbol rate uniformly distributed between sampling frequencies of 0.16 and 0.24 times, and a raised cosine pulse shape with 50% residual bandwidth. The OFDM signal employs 32 subcarriers (ensure sufficient number of subcarriers so that they can be modeled as independent Gaussian signals), modulated using BPSK with no cyclic prefix and the same raised cosine filter as other signals. In addition, it is assumed that the receiving filter is an ideal low-pass filter that can completely remove out-of-band noise. As the prior information is assumed to be unknown in this study, signals are sampled at a constant rate, rather than being sampled at integer multiples of the symbol rate.

The input to the classifier is two types of feature, where the SOF profile is calculated by 400 symbols (corresponding to 4096 discrete sampling points of IF signals). The CCs uses 2000 symbols, corresponding to 20526 sample points. As mentioned above, the same number of indexes is used in the calculation of features, that is, each feature vector has the same length of 256. Thus, we have 3×10^4 feature matrices (SOF profile feature matrix, CC feature matrix and hybrid feature matrix) of each type of modulation in the dataset. Then, the dataset is divided into training and testing sets in a 7:3 ratio. Also, assuming that the frequency-selective fading channel is implemented using independent equal-power taps with random Rayleigh amplitude and uniformly distributed phase, and the SNR range is from 0 dB to 15 dB.

B. RESULTS AND ANALYSIS

The proposed method is first tested in AWGN channel. Fig. 6 shows the performance comparison of CAES1, CASE2, CASE2 coarse classification, and CASE2 fine classification under various SNRs. Specifically, the classification sets of CAES1 and CASE2 are the entire modulation set, while the

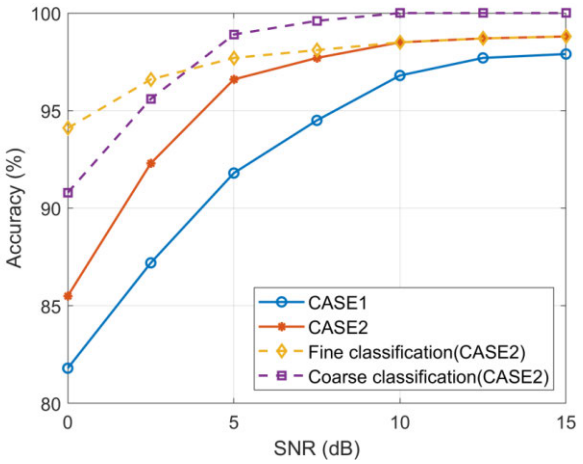

FIGURE 6. Comparison of classification performance of different approaches/stages of the proposed method in AWGN channel.

classification sets of CASE2 coarse classification and fine classification are the subset of the modulation set described in Section III. It can be seen from that in AWGN channel, CASE2 can achieve an accuracy of more than 95% when the SNR is above 5 dB, and the classification accuracy is close to 100% when the SNR is greater than 10 dB. In contrast, the performance of CASE1 is slightly worse than that of CASE2, e.g., to achieve the same level of accuracy as CASE2, the SNR needs to be increased by about 2 to 3 dB. This demonstrates the advantage of hierarchical classification in terms of accuracy. Moreover, the accuracy of CASE2 coarse classification can reach 100% when the SNR is higher than 10 dB, and it is about 85% when the SNR is 0 dB. In contrast, CASE2 fine classification performs better at a low SNR, reaching a classification accuracy of about 90% at the SNR of 0 dB. This is because the fine classification uses CCs as the feature, which reduces the impact of noise.

Next, the CASE2 implementation based on the proposed algorithm is simulated and the contribution of the features involved to modulation classification in frequency selective fading channels is analyzed. Fig. 8 shows the effect of using SOF profile features with different numbers of channel taps on coarse classification. It can be found that under the low SNR and severe multipath effects, the accuracy is less than 65%. However, the accuracy increases with the increase of SNR. Under the condition that the SNR is greater than 5 dB, the proposed method can achieve a satisfactory accuracy of more than 80% in coarse classification, and has only a small performance loss compared with that in AWGN channel. This proves that the proposed method using SOF profile features is robustness to the multipath effect on the channel under high SNR.

Fig. 9 shows the effect of using CC features with different numbers of channel taps on fine classification. It can be found that similar to the coarse classification, the accuracy increases with the increase of the SNR. Under the condition that the SNR is greater than 5 dB, the proposed method can

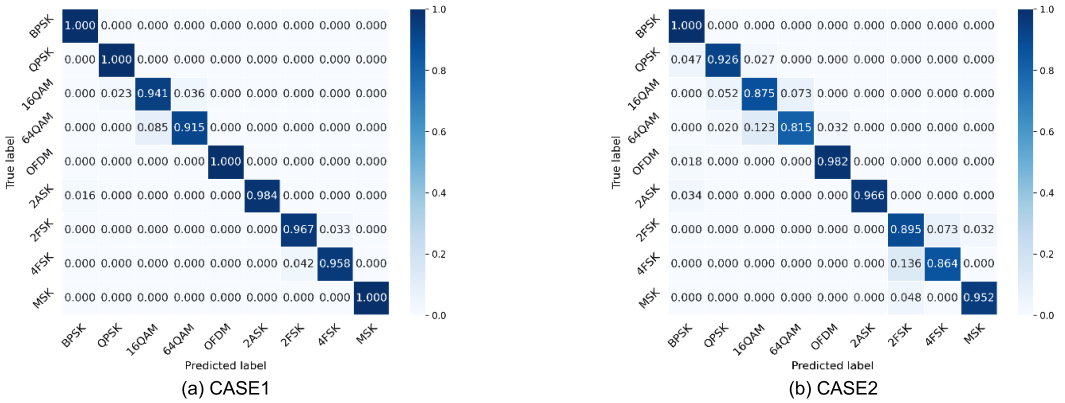

FIGURE 7. Confusion matrices of different approaches of the proposed method at the SNR of 5 dB.

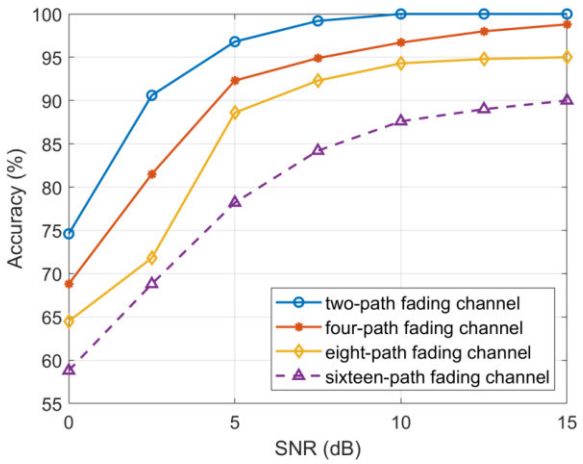

FIGURE 8. Comparison of classification performance using SOF profile features under different number of channel taps.

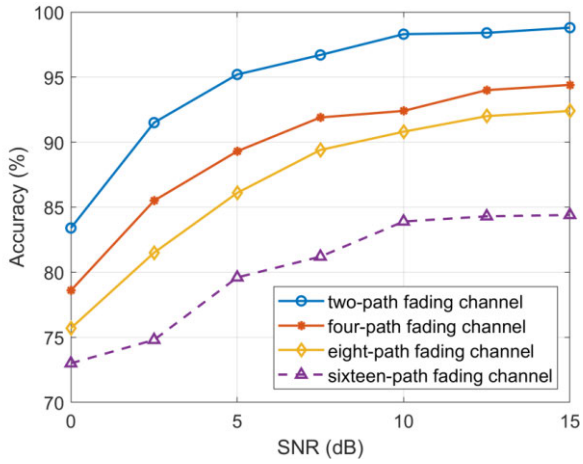

FIGURE 9. Comparison of classification performance using CC features under different number of channel taps.

also achieve a satisfactory accuracy of more than 80% when distinguishing sub-categories in the modulation sets. The difference is that the CC feature can achieve low performance attenuation at low SNR, with approximately 73% accuracy at 0 dB under the most severe multipath effects. In fact, this is due to the noise suppression of high-order CCs. Therefore, this proves that proposed method using CC features is also robustness to the multipath effect on the channel.

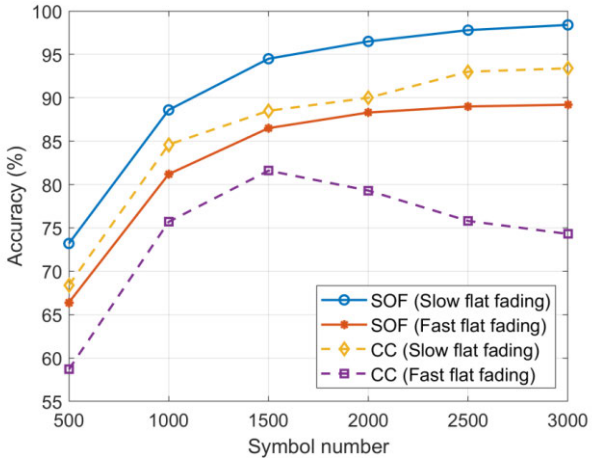

FIGURE 10. Comparison of classification performance using different features under different channel conditions.

Although the theoretical derivation does not involve fast fading channels, relevant channel scenarios are set up in simulation experiments. Each tap of the fast fading channel involved maintains a coherent value of 0.9 on approximately 500 samples equal to 50 symbols. Figure 10 shows the performance comparison of modulation recognition using SOF profile features and CC features under different channel conditions at a signal-to-noise ratio of 10 dB. It can be found that using two types of features for classification in fast fading channels results in performance degradation compared to slow fading channels, and using CC features is more severe compared to SOF profile features. This is because CC features require more symbols for calculation, and are therefore more affected by fast fading.

Fig. 11 shows the performance comparison of modulation recognition for CASE2 under broader channel conditions,

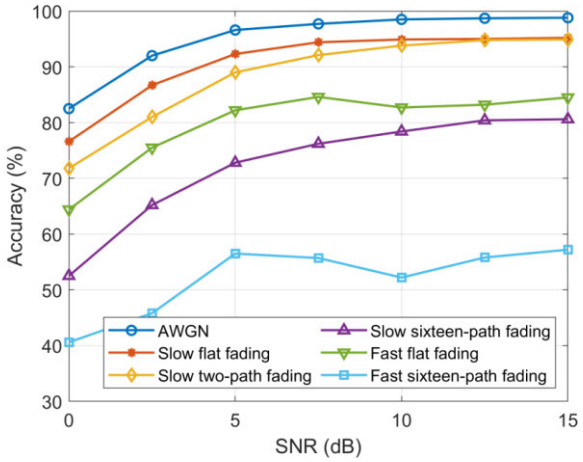

FIGURE 11. Comparison of classification performance using different features under different channel conditions.

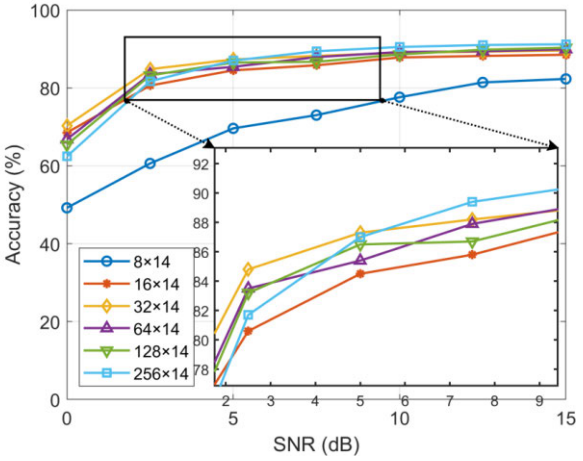

FIGURE 12. Comparison of classification performance of the proposed deep learning model between different input feature dimensions.

including AWGN channel, slow flat fading channel, slow two-path fading channel, slow sixteen-path fading channel, fast flat fading channel, and fast sixteen-path fading channel. The channel multipath and fast fading settings are the same as those in the simulation experiments mentioned above. It can be observed that under the condition of the SNR greater than 5 dB, the proposed method has almost the same performance in both slow flat fading channels and slow two-path fading channels, achieving an accuracy of over 90%. In slow sixteen-path fading channels with severe multipath effects, the proposed method can achieve an accuracy of over 70% even above 5 dB. This proves the robustness of the proposed method to multipath effects under high SNR. In addition, it can be observed that the performance of the proposed method significantly degrades under fast fading conditions. The accuracy decreases by more than 10% in fast flat fading channels compared to that in slow flat fading, and the overall accuracy is less than 50% under the worst channel conditions (fast sixteen-path fading channels). Therefore, the method is recommended to run in slow fading channel environments.

Next, the accuracy and complexity of the proposed deep learning model based on the implementation of CASE1 are evaluated in the slow four-path fading channel. Firstly, we analyze how the input feature dimension affects the performance of the proposed model through simulation experiment. Gradually reducing the dimensionality of feature vectors (i.e. reducing the number of indices used for various features, D), Fig. 12 shows the classification accuracy under different SNRs. The results indicate that the classification performance is not well enough when using the feature matrix input with 8-rows, but satisfactory performance is achieved when the number of rows is between 16 and 256 (power of 2) in the feature matrix. Considering the accuracy and training time under low SNR conditions, 32×14 is the most suitable input feature dimension for the proposed model, achieving an accuracy of 87.3% at 5 dB. Moreover, for the comparison of deep learning models in the following part, this size of feature input is used.

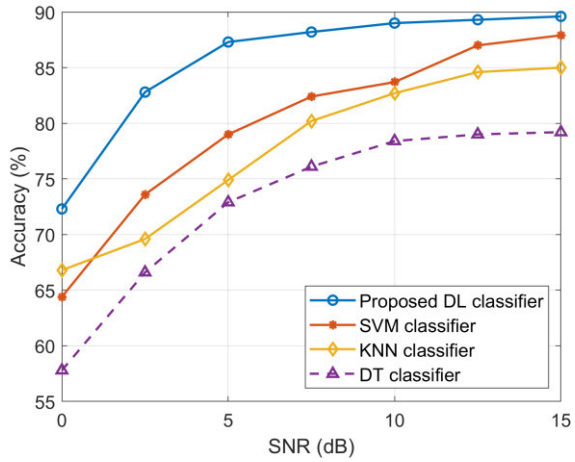

FIGURE 13. Comparison of classification performance between the proposed deep learning method and the traditional classification algorithm based on SVM.

Then, in order to investigate the performance gain of the proposed DL-based classifier compared to machine learning (ML)-based classifiers, we demonstrate the classification accuracy using the same feature input in Fig. 13. The involved ML classification algorithms are support vector machine (SVM) [34], K-Nearest Neighbor (KNN) [35], and decision tree (DT) [36]. It can be observed that at an accuracy level of 80%, the SNRs required by the SVM, KNN, and DT classifier are about 4 dB, 5.5 dB, and 10.5 dB higher than the proposed DL-based classifier, respectively. This proves that the DL-based classifier has better classification performance under the same feature input in fading channels. In addition, SVM requires a SNR estimator to create and train a model for each SNR level when used in reality. On the contrary, the proposed classifier is independent of SNR, and therefore has certain advantages in computational complexity.

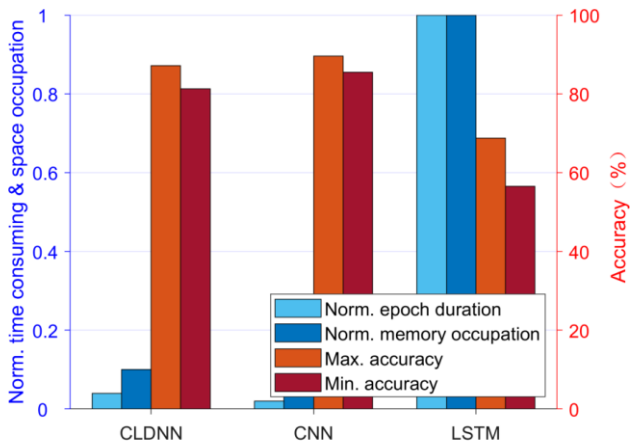

FIGURE 14. Comparison of performance of various deep learning models in terms of complexity, memory, and accuracy.

Finally, we compare the performance of the proposed CNN model with classical deep learning models, namely CLDNN [37] and LSTM [26]. As shown in Fig. 14, the proposed model has advantages in low time consumption (i.e. epoch duration), effective memory occupation and high classification accuracy. In this experiment, the epoch size of all models remains the same. The memory occupation and training time of each model are normalized using those of LSTM, because the maximum values are observed in the LSTM model. The average accuracy is the average of the accuracy within the range of SNR between 0 dB and 15 dB. Therefore, Fig. 13 actually provides results unrelated to computer performance. It should be noted that early stopping is used in all model training processes, and the least epochs are used. The experimental results show that the proposed CNN based deep learning model has a maximum accuracy of 89.6%, an average accuracy of 85.5%, a normalized epoch duration of 0.02, and a memory footprint of 0.05. This indicates that the proposed model is more robust and efficient than other classical deep learning models.

V. CONCLUSION

In this study, a CS feature based method with CNN was proposed for modulation classification in multipath fading channels. Two approaches were provided to implement our method. One was to directly learn from the hybrid CS features and perform classifications through a single-structured CNN; the other used a hierarchical CNN classifier to perform coarse classification based on SOF and fine classification based on CC, respectively. The simulation experiments indicated that, benefiting from feature extraction based on CS and classification based on CNN, the proposed method can achieve satisfactory performance under various channel conditions. Besides, the hierarchical approach achieves better performance than that through one-stage. Moreover, the proposed classifier with CNN was proven to be superior to the method with ML and other DL classifiers. All these results demonstrate the applicability of CS feature-based method with CNN in the rapidly changing channel environment of wireless communication networks.

However, in our method, the fading channel is modeled as a stationary or quasi stationary channel, and simulation also shows poor performance once the channel experiences fast fading. In future research, we will further explore the potential of CS features and DL models, aiming to solve the problems of modulation classification in dynamic channels.

REFERENCES

[1] S. Surekha and M. Z. U. Rahman, "Cognitive energy-aware spectrum sensing with improved throughput for medical sensor networks," *IEEE Sensors Lett.*, vol. 6, no. 6, pp. 1–4, Jun. 2022, doi: 10.1109/LSENS.2022.3180629.
[2] S. Subray, S. Tschimben, and K. Gifford, "Towards enhancing spectrum sensing: Signal classification using autoencoders," *IEEE Access*, vol. 9, pp. 82288–82299, 2021, doi: 10.1109/ACCESS.2021.3087113.
[3] F. Shen, G. Ding, Z. Wang, and Q. Wu, "UAV-based 3D spectrum sensing in spectrum-heterogeneous networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 5711–5722, Jun. 2019, doi: 10.1109/TVT.2019.2909167.
[4] H. Ding, X. Li, Y. Ma, and Y. Fang, "Energy-efficient channel switching in cognitive radio networks: A reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 12359–12362, Oct. 2020, doi: 10.1109/TVT.2020.3006471.
[5] S. Guo and X. Zhao, "Deep reinforcement learning optimal transmission algorithm for cognitive Internet of Things with RF energy harvesting," *IEEE Trans. Cognit. Commun. Netw.*, vol. 8, no. 2, pp. 1216–1227, Jun. 2022, doi: 10.1109/TCCN.2022.3142727.
[6] A. Abbas, V. Pano, G. Mainland, and K. Dandekar, "Radio modulation classification using deep residual neural networks," in *Proc. MILCOM IEEE Mil. Commun. Conf. (MILCOM)*, Nov. 2022, pp. 311–317, doi: 10.1109/MILCOM55135.2022.10017640.
[7] W. H. Clark, V. Arndorfer, B. Tamir, D. Kim, C. Vives, H. Morris, L. Wong, and W. C. Headley, "Developing RFML intuition: An automatic modulation classification architecture case study," in *Proc. MILCOM IEEE Mil. Commun. Conf. (MILCOM)*, Nov. 2019, pp. 292–298, doi: 10.1109/MILCOM47813.2019.9020949.
[8] Z. Xing and Y. Gao, "A modulation classification algorithm for multipath signals based on cepstrum," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 4742–4752, Jul. 2020, doi: 10.1109/TIM.2019.2955535.
[9] A. P. Hermawan, R. R. Ginanjar, D.-S. Kim, and J.-M. Lee, "CNN-based automatic modulation classification for beyond 5G communications," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1038–1041, May 2020, doi: 10.1109/LCOMM.2020.2970922.
[10] T. Huynh-The, T.-V. Nguyen, Q.-V. Pham, D. B. da Costa, G.-H. Kwon, and D.-S. Kim, "Efficient convolutional networks for robust automatic modulation classification in OFDM-based wireless systems," *IEEE Syst. J.*, vol. 17, no. 1, pp. 964–975, Mar. 2023, doi: 10.1109/JSYST.2022.3207377.
[11] T. V. R. O. Câmara, A. D. L. Lima, B. M. M. Lima, A. I. R. Fontes, A. D. M. Martins, and L. F. Q. Silveira, "Automatic modulation classification architectures based on cyclostationary features in impulsive environments," *IEEE Access*, vol. 7, pp. 138512–138527, 2019, doi: 10.1109/ACCESS.2019.2943300.
[12] A. Serbes, H. Cukur, and K. Qaraqe, "Probabilities of false alarm and detection for the first-order cyclostationarity test: Application to modulation classification," *IEEE Commun. Lett.*, vol. 24, no. 1, pp. 57–61, Jan. 2020, doi: 10.1109/LCOMM.2019.2947043.
[13] W. Gardner, W. Brown, and C.-K. Chen, "Spectral correlation of modulated signals: Part II-digital modulation," *IEEE Trans. Commun.*, vol. COM-35, no. 6, pp. 595–601, Jun. 1987, doi: 10.1109/TCOM.1987.1096816.
[14] A. Napolitano and C. M. Spooner, "Cyclic spectral analysis of continuous-phase modulated signals," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 30–44, Jan. 2001, doi: 10.1109/78.890336.
[15] C. M. Spooner, "Classification of co-channel communication signals using cyclic cumulants," in *Proc. Conf. Rec. 29th Asilomar Conf. Signals, Syst. Comput.*, Oct. 1995, pp. 531–536, doi: 10.1109/ACSSC.1995.540605.
[16] S. Majhi, R. Gupta, W. Xiang, and S. Glisic, "Hierarchical hypothesis and feature-based blind modulation classification for linearly modulated signals," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 11057–11069, Dec. 2017, doi: 10.1109/TVT.2017.2727858.

[17] O. A. Dobre, Y. Bar-Ness, and W. Su, "Higher-order cyclic cumulants for high order modulation classification," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2003, pp. 112–117, doi: 10.1109/MILCOM.2003.1290087.
[18] P. Marchand, C. Le Martret, and J.-L. Lacoume, "Classification of linear modulations by a combination of different orders cyclic cumulants," in *Proc. IEEE Signal Process. Workshop Higher-Order Statist.*, Banff, AB, Canada, Jul. 1997, pp. 47–51, doi: 10.1109/HOST.1997.613485.
[19] E. Like, V. Chakravarthy, P. Ratazzi, and Z. Wu, "Signal classification in fading channels using cyclic spectral analysis," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, no. 1, Dec. 2009, Art. no. 879812, doi: 10.1155/2009/879812.
[20] E. Like, V. Chakravarthy, R. Husnay, and Z. Wu, "Modulation recognition in multipath fading channels using cyclic spectral analysis," in *Proc. IEEE GLOBECOM Global Telecommun. Conf.*, Nov. 2008, pp. 1–6, doi: 10.1109/GLOCOM.2008.ECP.584.
[21] W. A. Gardner, "The spectral correlation theory of cyclostationary time-series," *Signal Process.*, vol. 11, no. 1, pp. 13–36, 1986, doi: 10.1016/0165-1684(86)90092-7.
[22] W. Gardner, "Spectral correlation of modulated signals: Part I-analog modulation," *IEEE Trans. Commun.*, vol. COM-35, no. 6, pp. 584–594, Jun. 1987, doi: 10.1109/TCOM.1987.1096820.
[23] W. C. Headley, J. D. Reed, and C. R. C. M. da Silva, "Distributed cyclic spectrum feature-based modulation classification," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Las Vegas, NV, USA, Mar. 2008, pp. 1200–1204, doi: 10.1109/WCNC.2008.216.
[24] N. E. West and T. O'Shea, "Deep architectures for modulation recognition," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, Mar. 2017, pp. 1–6, doi: 10.1109/DySPAN.2017.7920754.
[25] 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, doi: 10.1109/JSTSP.2018.2797022.
[26] 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, doi: 10.1109/TCCN.2018.2835460.
[27] Y. Guo, H. Jiang, J. Wu, and J. Zhou, "Open set modulation recognition based on dual-channel LSTM model," 2020, *arXiv:2002.12037*.
[28] Z. Ke and H. Vikalo, "Real-time radio technology and modulation classification via an LSTM auto-encoder," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 370–382, Jan. 2022, doi: 10.1109/TWC.2021.3095855.
[29] X. Wei, W. Luo, X. Zhang, J. Yang, G. Gui, and T. Ohtsuki, "Differentiable architecture search-based automatic modulation classification," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Nanjing, China, Mar. 2021, pp. 1–6, doi: 10.1109/WCNC49053.2021.9417449.
[30] Z. Wu, E. Like, and V. Chakravarthy, "Reliable modulation classification at low SNR using spectral correlation," in *Proc. 4th IEEE Consum. Commun. Netw. Conf.*, Jan. 2007, pp. 1134–1138, doi: 10.1109/CCNC.2007.228.
[31] P. D. Sutton, K. E. Nolan, and L. E. Doyle, "Cyclostationary signatures for rendezvous in OFDM-based dynamic spectrum access networks," in *Proc. 2nd IEEE Int. Symp. New Frontiers Dyn. Spectr. Access Netw.*, Apr. 2007, pp. 220–231, doi: 10.1109/DYSPAN.2007.37.
[32] F. Chollet. (2015). *Keras*. [Online]. Available: https://github.com/fchollet/keras
[33] K. Tekbiyik, Ö. Akbunar, A. R. Ekti, A. Görçin, G. K. Kurt, and K. A. Qaraqe, "Spectrum sensing and signal identification with deep learning based on spectral correlation function," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10514–10527, Oct. 2021, doi: 10.1109/TVT.2021.3109236.
[34] K. Tekbiyik, Ö. Akbunar, A. R. Ekti, A. Görçin, and G. Karabulut Kurt, "Multi–dimensional wireless signal identification based on support vector machines," *IEEE Access*, vol. 7, pp. 138890–138903, 2019, doi: 10.1109/ACCESS.2019.2942368.
[35] M. W. Aslam, Z. Zhu, and A. K. Nandi, "Automatic modulation classification using combination of genetic programming and KNN," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2742–2750, Aug. 2012, doi: 10.1109/TWC.2012.060412.110460.
[36] S. Luan, Y. Gao, W. Chen, N. Yu, and Z. Zhang, "Automatic modulation classification: Decision tree based on error entropy and global-local feature-coupling network under mixed noise and fading channels," *IEEE Wireless Commun. Lett.*, vol. 11, no. 8, pp. 1703–1707, Aug. 2022, doi: 10.1109/LWC.2022.3175531.
[37] X. Liu, D. Yang, and A. E. Gamal, "Deep neural network architectures for modulation classification," in *Proc. 51st Asilomar Conf. Signals, Syst., Comput.*, Oct. 2017, pp. 915–919, doi: 10.1109/ACSSC.2017.8335483.

LIYAN YIN received the B.S. and M.S. degrees from Air Force Engineering University, in 2017 and 2019, respectively, where he is currently pursuing the Ph.D. degree. His research interests include aeronautical communications and signal identification.

XIN XIANG received the B.S. and M.S. degrees from the Aeronautics Engineering College, Air Force Engineering University, Xi'an, China, and the Ph.D. degree in signal and information processing from Xidian University, China, in 2009. Currently, he is a Professor with the Aeronautics Engineering College, Air Force Engineering University. His current research interests are radio communications, software-defined radio, and navigation technology.

YUAN LIANG received the B.S., M.S., and Ph.D. degrees in information and communication engineering from Air Force Engineering University, Xi'an, China, in 2012, 2014, and 2018, respectively. Currently, he is a Lecturer with the Aeronautics Engineering College, Air Force Engineering University. His research interests include intelligent signal processing, wireless communications, secure communications, and optimization and intelligent algorithm.