RESEARCH ARTICLE

LPI Radar Signals Modulation Recognition Based on ACDCA-ResNeXt

XUDONG WANG, GUIGUANG XU [ID], HE YAN [ID], DAIYIN ZHU [ID], YING WEN, AND ZEHU LUO
College of Electronics and Information Engineering/College of Integrated Circuits, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Corresponding author: Xudong Wang (xudong@nuaa.edu.cn)

ABSTRACT For low probability of intercept (LPI) radar waveform identification accuracy (ACC) problem at low Signal-to-Noise Ratios (SNRs), an approach based on time-frequency analysis (TFA) and Asymmetric Dilated Convolution Coordinate Attention Residual networks (ACDCA-ResNeXt) is proposed to recognize twelve kinds of LPI radar signals automatically. First, we apply Choi-Williams distribution (CWD), which shows superior performance at low SNRs, to transforming radar signals into time-frequency images (TFI). Then, in order to obtain the high-quality TFIs, a series of image processing techniques, including 2D Wiener filtering, image cutting, and image resize, are used to remove the background noise and redundant frequency bands of the TFI and obtain a fixed-size gray scale image containing main morphological features of the TFI. Finally, the TFIs are input into ACDCA-ResNeXt network that can extract and learn deep features to recognize radar waveforms. Furthermore, a fusion loss function, which is composed of a soft-label smoothed cross entropy loss function and a center loss function, improves the generalization capability performance of network and achieves a better clustering effect. Experimental results demonstrate that, for twelve kinds of LPI radar waveforms, the overall recognition ACC of the proposed approach achieves 97.94% when SNR is −8 dB.

INDEX TERMS Radar waveform recognition, time-frequency analysis (TFA), asymmetric convolution (AC), dilated convolution, coordinate attention (CA) mechanism.

I. INTRODUCTION

Special modulation is usually adopted for the low probability of interception (LPI) radar to prevent interception and detection by non-cooperative intercept receivers. Therefore, LPI radar waveform recognition is a part of the key technologies of radar countermeasure system [1]. In recent years, a large number of approaches based on modulation features have been proposed to recognize radar signals such as spectral correlation analysis [2], modulation domain analysis [3] and joint Holder coefficient feature [4]. These approaches have effectively enhanced the ACC of LPI radar waveform recognition, but they have some shortcomings and thus fail to meet the recognition requirements of automation, informatization and intelligence of the system.

The associate editor coordinating the review of this manuscript and approving it for publication was Hasan S. Mir.

Since the time-frequency analysis(TFA) method can effectively establish the time-frequency correspondence of non-smooth signals, this method has become an effective tool to study non-smooth signals in non-ideal environments [5], [6] [7], such as short-time Fourier transform (STFT), Wigner-Ville distribution (WVD) and Choi-William distribution (CWD). In recent years, combined with deep learning methods, they have been applied to LPI radar signal recognition and achieved good recognition performance. Zhang et al. [8] proposed an approach based on TFA and convolutional neural network (CNN), which could recognize eight types of LPI radar modulation signals. ACC of recognizing eight types of radar signals achieved 94.5% when SNR was −2 dB. However, four types of ploy-phase coded signals (P1-P4) in the LPI radar signals were not considered in this approach except Frank code. In addition, binary phase shift keying (BPSK) and T1 code had similar features

in time-frequency images (TFI) at a low SNR, which was not conducive to recognize these signals. The waveform of multiple continuous pulses was applied to TFI of BPSK, but the TFIs of other signals only contained a single pulse, which made dissimilar features generate and restrict the applications of this approach in practice. Kong et al. [9] uses CNN and CWD to identify twelve LPI radar signals (BPSK, Costas, LFM, Frank, P1 P4 and T1 T4), and experimentally verifies the effects of different image sizes, the size and number of filters, and the number of neurons on the recognition accuracy. Qu et al. [10] utilized Cohen time-frequency distribution, image processing and convolutional neural network to recognize twelve types of radar modulation signals. Although the recognition effect was satisfactory at a low SNR, binary image was still used as CNN input in this approach. Ma et al. [11] proposed a multi-feature image for joint decision-making model for the non-stationary characteristics of most LPI radar signals. The short-time autocorrelation image, double short-time autocorrelation images and TFIs were simultaneously input into the hybrid model classifier, and the overall recognition rate of this approach achieved 87.7% when the SNR was −6 dB. Qin et al. [12] proposed a recognition approach based on CWD and dilated residual network. When the SNR was −6 dB, the overall recognition ACC of linear frequency modulation (LFM) signals with similar TFI features exceeded 95%. Ni et al. [13] proposed a multi-resolution depth feature fusion approach based on the FSST, which was utilized for LPI radar waveform recognition. The overall recognition ACC of this approach was 95.2% at −8 dB. In addition, some researchers extracted features of TFI through transfer learning and conducted classification by the support vector machine (SVM) [14].

The above approaches made some progress in the ACC of radar waveform recognition. However, they did not take into account the effect of redundant frequency bands in the TFIs on recognition ACC, and the built network did not consider the learning capability of space and channel. As the size of the feature map used for classification was too small, the subtle features of TFIs were neglected. In addition, Softmax faced the problems of insufficient generalization capability and poor clustering effect. In this paper, we propose a novel approach for LPI radar signal modulation recognition which uses TFA, image processing and ACDCA-ResNeXt to address the above problems under noise jamming. We study the recognition of twelve types of LPI radar signals including BPSK, LFM, Costas, five polyphase codes (such as Frank, P1, P2, P3 and P4) and four polytime codes (such as T1, T2, T3 and T4).

The main contributions of this paper can be summarized as follows.

1) We propose a new signal preprocessing technique to eliminated background noise and redundant frequency bands in TFIs by using CWD, 2D Wiener filtering, image cutting and image resize.

2) To further enhance the location information extraction capability and channel feature extraction capability of the ResNeXt network, we help it learn more expressive and directional parameters during training by adding a coordinate attention (CA) block to the network.

3) We introduce the Asymmetric convolution (AC) block, i.e., we add the horizontal and vertical convolution operations to the square convolution to improve the extraction capability of the ordinary convolution for spatial location features.

4) A fusion loss ($\mathcal{L}_F$) function is proposed, which consists of soft-label smoothed cross entropy loss ($\mathcal{L}_{SLSCE}$) function, and center loss ($\mathcal{L}_C$) function. The $\mathcal{L}_{SLSCE}$ function is used to regularize the network to boost signal classification performance, while the $\mathcal{L}_C$ function is used to increase the distance between different types of signals and reduces the distance within the same category.

The rest of the paper is organized as follows. Section II introduces the system structure and the signal model. In Section III, TFA and image preprocessing are entered into detail. Section IV illustrates ACDCA-ResNeXt and $\mathcal{L}_F$ function in detail; Section V provides the simulation results. The conclusion is presented in Section VI.

II. SYSTEM STRUCTURE AND SIGNAL MODEL

In this section, we present the structure of LPI radar classification system and the definition of LPI radar signals considered in this paper.

A. SYSTEM STRUCTURE

As Figure. 1 in shown, the classification system consists of three main parts: signal preprocessing, feature extraction and classification. In the signal preprocessing stage, the received LPI radar waveform is firstly converted to TFI by CWD, which can reflect the instantaneous frequency of the signal and has good anti-noise performance at the same time. For modulation identification of radar waveforms, we pay more attention to the morphological characteristics of the TFIs, so 2D Wiener filtering is used to smooth the TFI to reduce noise interference, and image cutting is used to remove redundant frequency bands. Next, the size of the TFI is adjusted to reduce the training cost of network. Then, a ACDCA-ResNeXt is designed for feature extraction. Finally, we apply full connection layer and use Softmax as a classifier to realize waveform recognition.

B. SIGNAL MODEL

Assuming that the received LPI radar signal is a pulse wave with additive white Gaussian noise (AWGN). Then, the signal model can be expressed as

$$y[k] = x[k] + n[k], 0 \leq kT_s \leq T \quad (1)$$

where

$$x[k] = A[k]e^{j\theta[k]} \quad (2)$$

is the discrete time complex LPI radar signal samples, k is the sample index increasing every T_s for a sampling frequency f_s, $A[k]$ is the ideal sampling signal of instant envelope, T is
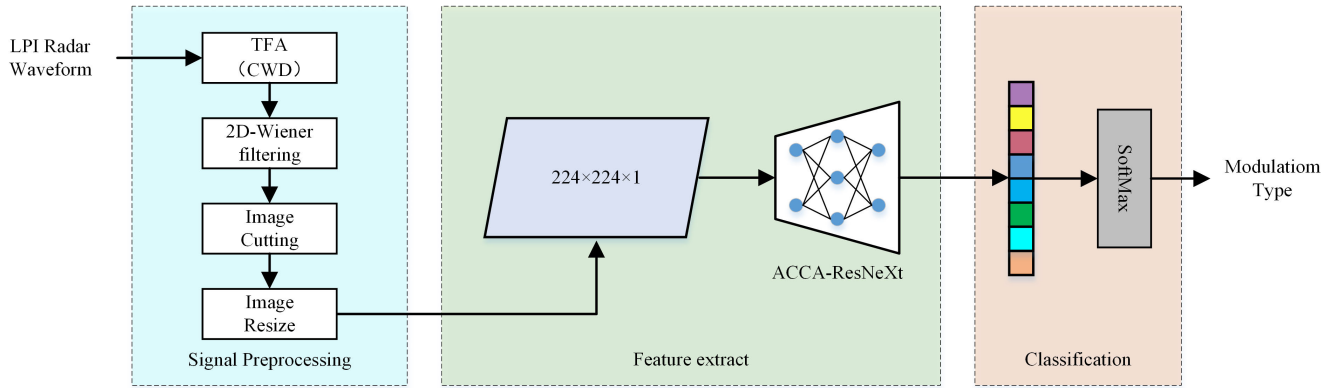
FIGURE 1. The proposed LPI radar waveform recognition system.

the pulse duration, T_s is the sampling interval, $n[k]$ is white Gaussian noise, and $\theta[k]$ is the instantaneous phase of the ideal sampling signal.

Here, $\theta[k]$ can be computationally expressed by instantaneous frequency $f[k]$ and the instantaneous phase offset $\varphi[k]$ as below

$$\theta[k] = 2\pi f[k](kT_s) + \varphi[k] \tag{3}$$

which determine the modulation type of radar signal.

In this subsection, we define twelve LPI radar waveforms are shown in Table 1, including LFM, Costas, BPSK, five polyphase code (such as Frank, P1, P2, P3 and P4 codes), and four polytime codes (such as T1, T2, T3 and T4 codes) [9].

TABLE 1. LPI radar signal modulation parameters.

Signal	$f[k]$[Hz]	$\varphi[k]_{i,j}$[rad] for subcode
LFM	$f_c + B\dfrac{kT_s}{T}$	constant
Costas	f_{min}	constant
BPSK	constant	0 or π
Frank	constant	$\frac{2\pi}{\Upsilon}(i-1)(j-1)$
P1	constant	$-\frac{\pi}{\Upsilon}[\Upsilon-(2j-1)][(j-1)\Upsilon+(i-1)]$
P2	constant	$-\frac{\pi}{2\Upsilon}[2i-1-\Upsilon][2j-1-\Upsilon]$
P3	constant	$\frac{\pi}{\rho}(i-1)^2$
P4	constant	$\frac{\pi}{\rho}(i-1)^2 - \pi(i-1)$
T1	constant	$\mathrm{mod}\{\frac{2\pi}{n}\lfloor[s(kT_s)-jT]\frac{jn}{T}\rfloor, 2\pi\}$
T2	constant	$\mathrm{mod}\{\frac{2\pi}{n}\lfloor\frac{n(s(kT_s)-jT)}{2}(\frac{2j-s+1}{T})\rfloor, 2\pi\}$
T3	constant	$\mathrm{mod}\{\frac{2\pi}{n}\lfloor\frac{nB(kT_s)^2}{2T}\rfloor, 2\pi\}$
T4	constant	$\mathrm{mod}\{\frac{2\pi}{n}\lfloor\frac{nB(kT_s)^2}{2T} - \frac{nBkT_s}{2}\rfloor, 2\pi\}$

In Table 1, B, f_c, f_{min}, Υ, ρ, n and s represent the modulation bandwidth, center frequency, fundamental frequency, number of encoded phases, length of encoded sequence, number of phase states and number of step frequency segments, respectively. i, j represent an iterative integer value from 1 to Υ. $\lfloor\cdot\rfloor$ denotes the integer downward function. $\mathrm{mod}(a, b)$ represents the remainder function.

III. SIGNAL PREPROCESSING
In this section, we introduce the various methods used to process the received radar signal in detail, which include CWD, 2D Wiener filtering used to remove background noise of the TFI, image cutting used to remove redundant frequency bands and image resize.

A. CHOI-WILLIAMS DISTRIBUTION
The most widely used methods to extract the TFI, include STFT, WVD and so on. Once the window function and its length are chosen, the time-frequency resolution of STFT is fixed. The optimal time position and the best frequency resolution cannot be achieved simultaneously. Therefore, STFT is only suitable for analyzing quasi-stationary signals that are stationary at the scale of the short time window. The WVD does not involve any window function and thus it has very high time-frequency resolution, but its applications are inevitably hindered by the cross-term interference. CWD is able to obtain desired characteristics such as higher resolution and removal of cross-terms by smoothing the WVD through time and frequency offsets using kernel function [7]. The mathematical definitions are shown as

$$TFR_x(t, \omega) = \iint A_s(\tau, \upsilon)\phi(\tau, \upsilon)e^{-j(\upsilon t + \omega\tau)}d\tau d\upsilon \tag{4}$$

$$A_s(\tau, \upsilon) = \frac{1}{2\pi}\int x(t + \frac{\tau}{2})x^*(t - \frac{\tau}{2})e^{-j2\pi\upsilon t}dt \tag{5}$$

$$\phi(\tau, \upsilon) = e^{\frac{-\upsilon^2\tau^2}{\varepsilon}} \tag{6}$$

where $A_s(\tau, \upsilon)$ is the ambiguity function of input signal $x(t)$. τ and υ are the time delay and frequency shift, respectively. $\phi(\tau, \upsilon)$ is kernel function and $\varepsilon > 0$ is scaling factor.

The TFIs of the twelve kinds of LPI radar signals when SNR=10 dB are illustrated in Figure. 2. The TFI of LFM with $f[k]$ increasing or decreasing linearly from the initial frequency f_c to $f_c + B$ with the pulse interval T is characterized by a diagonal straight line, as shown in Figure. 2(a). Costas has pseudo-random frequency hopping pattern over a wide signal bandwidth B and its TFI is shown Figure. 2(b). The phase offset φ of BPSK has only two states (i.e., 0 and π),

and its TFI shows a "<" shape, as shown in Figure. 2(c). The TFIs for the polyphase codes (i.e.), Frank, P1, P2, P3 and P4 codes) are shown in Figure. 2(d)-2(h), respectively. The TFIs of Frank, P1 and P2 codes has staircase patterns, whereas the TFIs of P3 and P4 codes are linear frequency variation. Among the polytime modulations, the TFIs of T1 and T3 codes have "<" shapes, as shown in Figure. 2(i) and 2(k). On the other hand, the TFIs of T2 and T4 codes are an "X" shape as shown in Figure. 2(j) and 2(l). With the decrease of SNR, the stepped and linear time-frequency characteristics of multiphase modulation signals and LFM signals are weakened to some extent. As the TFIs of BPSK and T1, T2 and T4 codes are similar to it, it will be confused easily during recognition [9].

B. 2D WIENER FILTERING
Although CWD can suppress the cross terms, a lot of noise still exists in the TFIs at a low SNR. The noise will affect the subsequent recognition effect. Therefore, the reduction of noise interference is necessary. 2D Wiener filter is an adaptive one, which can adjust the effect of the filter according to the local variance of the TFI.

$I_{m \times n}$ is used to represent a TFI of $m \times n$ pixels, and each pixel can be expressed as $I(i,j), i = 1, \ldots, m, j = 1, \ldots, n$. $\eta_{a \times b}$ is the local neighborhood of each pixel in $I_{m \times n}$, and the size of $\eta_{a \times b}$ is generally set as 40×40. The neighborhood mean value ϵ and variance ς^2 of each pixel are calculated with (7) and (8), respectively, and the image $T(i,j)$ after Wiener filtering is calculated with (9).

$$\epsilon = \frac{1}{a \times b} \sum_{i,j \in \eta} I(i,j). \tag{7}$$

$$\varsigma^2 = \frac{1}{a \times b} \sum_{i,j \in \eta} I^2(i,j) - \epsilon^2. \tag{8}$$

$$T(i,j) = \epsilon + \frac{\varsigma^2 - \kappa^2}{\varsigma^2}(I(i,j) - \epsilon). \tag{9}$$

In (9), κ^2 is the noise variance, which can be expressed by the mean value of neighborhood variance.

C. IMAGE CUTTING AND IMAGE RESIZE
After 2D wiener filtering, the noise in the TFI is effectively smoothed out. To further highlight the differences between LPI radar signals with different modulation types, we used marginal frequency distribution (MFD) to remove redundant regions, and then resized the images by bilinear interpolation.

The MFD provides a way to obtain the region containing the signal energy in the TFI under low SNR [15]. It sums the time values of each frequency in the smoothed TFI and stores them as column vector $\mathcal{H}$. It is defined as

$$\mathcal{H}(\omega) = \sum_{t=0}^{N-1} TFR_x(t, \omega). \tag{10}$$

A histogram of the normalized MFD is created, containing 100 subgroups, and the appropriate group number N_g is
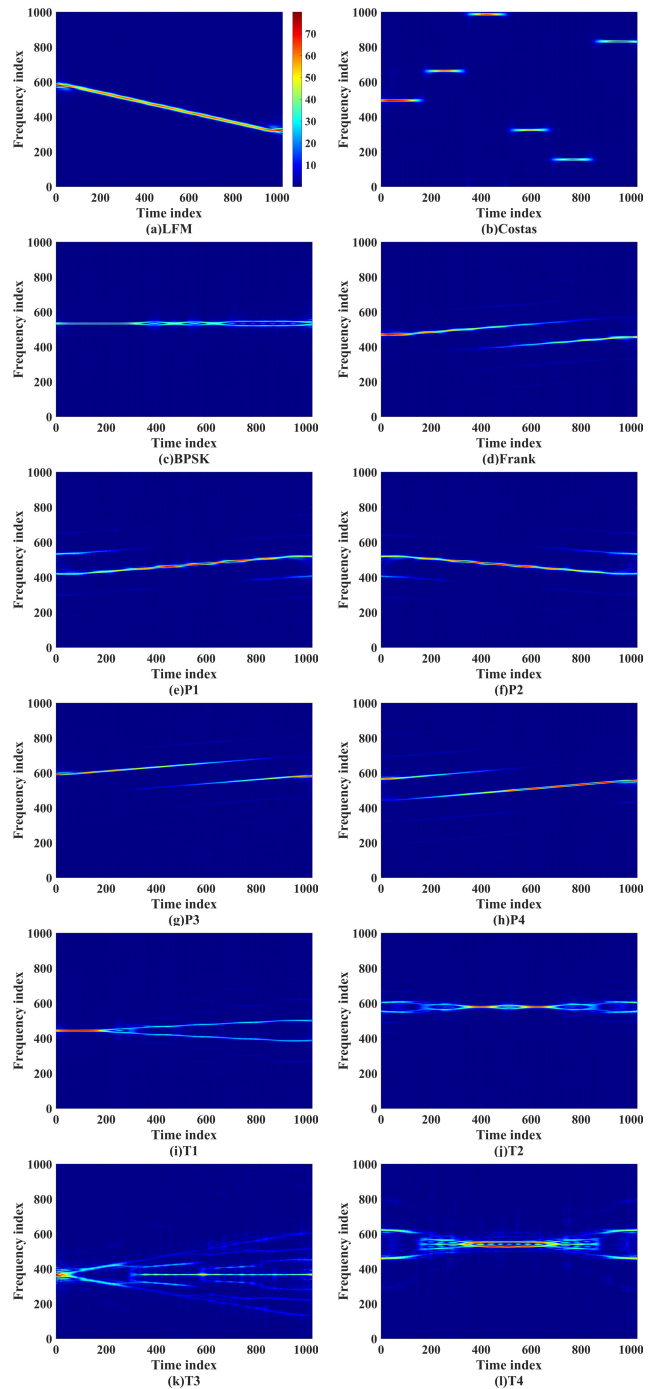

FIGURE 2. TFIs of twelve kinds of LPI radar waveforms (SNR=10 dB).

selected to determine first detection threshold. However, even if there is no signal distribution at a low SNR, the threshold will sometimes be exceeded. When N_g is great, it will be lower than the threshold even if there is signal distribution. Therefore, we adopted the "secondary threshold". That is, the marginal frequency is considered to have a signal when it is continuously higher than the threshold for ξ times, and it is considered to have no signal when it is continuously lower

than the threshold for λ times. The effective interval of signal energy distribution is determined [16].

After determining the starting and ending positions of the valid information region in the TFI, the region is intercepted and scaled to 224×224 by image resize (e.g., bilinear interpolation)to reduce the training cost of the network. According to the analysis to signal above, the method the signal prepressing is adopted to transform the received signals. As shown in Figure. 3, we choose $N_g = 45$, corresponding to a normalized energy $\mathcal{H} = 0.5786$ as the threshold, and $\lambda = \xi = 5$. The final TFI can well reflect the morphological features of original TFI.
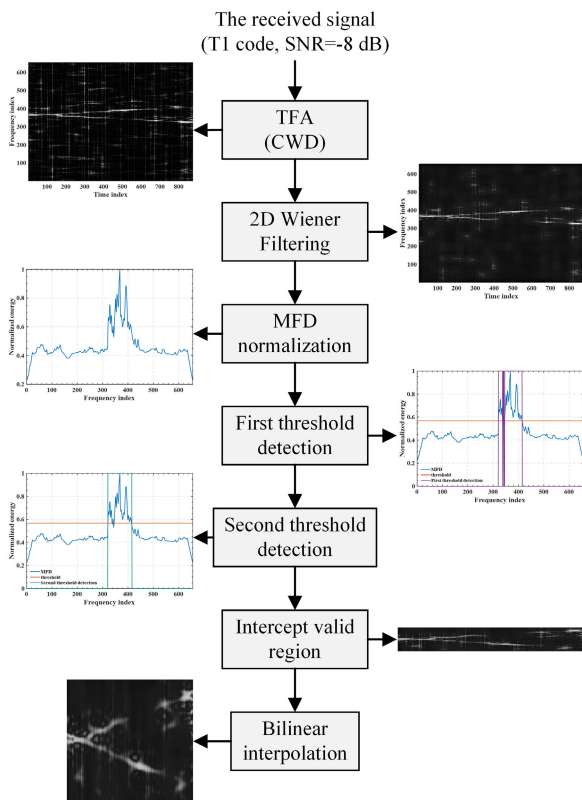

FIGURE 4. Relationship between effective interception and SNR.

In order to analyze the generalization capability of the network, the SNR of the training set ranges from -8 dB to 8 dB for each signal, 200 signals are generated every 2 dB, and there are totally 21,600 samples. The SNR of the test set ranges from -10 dB to 10 dB, 100 signals are generated every 2 dB, and there are totally 13,200 samples [9].

TABLE 2. Parameters of twelve kinds of LPI radar waveforms.

Signal	Parameters	Ranges
LFM	f_c	$U(1/10,1/3)$
	B	$U(1/10,1/5)$
Costas	N_h	3,4,5,6
	f_{min}	$U(1/24,1/20)$
BPSK	f_c	$U(1/10,1/3)$
	L_c	7,11,13
Frank, P1,P2	f_c	$U(1/10,1/3)$
	Υ	[8,12], Υ of P2 is even
P3,P4	f_c	$U(1/10,1/3)$
	ρ	36,64,81,100
T1,T2	f_c	$U(1/10,1/3)$
	n	[4,6]
T3,T4	f_c	$U(1/10,1/3)$
	n	[4,6]
	B	$U(1/20,1/8)$

FIGURE 3. Preprocessing of T1 signal (SNR = -8 dB).

Figure. 4 shows the ACC of effective interception of T1 signal obtained from 1,000 Monte Carlo experiments at different SNRs. It can be seen that the ACC can achieve 100% when it is -8 dB and above.

D. DESCRIPTIONS OF SIMULATION DATA SET
In the experiment, twelve types of simulated LPI radar signals in Section I are generated for training and testing the recognition system and their specific parameters are shown in Table 2. Here, N_h is number of frequency hopping, $U(\cdot)$ denotes a uniform distribution based on the sample rate. The length of signals is set between 1,024 and 2,048 randomly. In order to make the simulated signal similar to the real signal, Gaussian white noise with different SNRs is added to the signal.
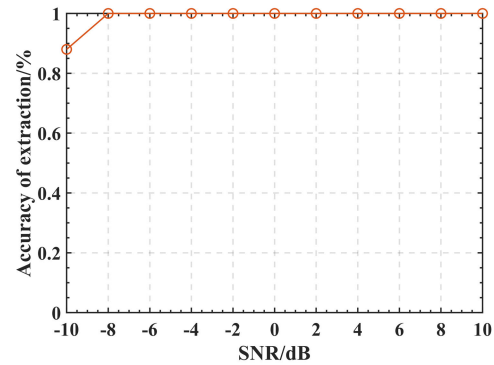
IV. METHODOLOGY
This section provides detailed information for ACDCA-ResNeXt, including its structure and parameters. Then, we combine $\mathcal{L}_C$ and $\mathcal{L}_{SLSCE}$ function to propose a $\mathcal{L}_F$ function to achieve better clustering effect and enhances the generalization performance of the network.

A. ACDCA-ResNeXt
ACDCA-ResNeXt is composed of AC block [17], CA block [18] and ResNeXt block [19], which strengthens capability of the ResNeXt block to obtain spatial, channel and location information and improves the receptive field of the model through dilated convolution.

1) AC BLOCK
AC block uses 1D asymmetric convolution kernels to strengthen the square convolution kernel and improve the

spatial information extraction capability of the convolution layer and the robustness of flipping and rotating images. We assume that $\mathcal{F} \in \mathbb{R}^{K \times K \times C}$ is 3D convolution kernel and its number is D, the input of a convolutional layer is $\mathcal{M} \in \mathbb{R}^{U \times V \times C}$ which is a feature map with a size of $U \times V$ and C channels, the output is $\mathcal{O} \in \mathbb{R}^{R \times T \times D}$ with D channels, respectively. The j of the output feature map channel is

$$\mathcal{O}_{:,:,j} = \sum_{k=1}^{C} \mathcal{M}_{:,:,k} \odot \mathcal{F}_{:,:,k}^{(j)}, \tag{11}$$

where $\odot$ represents 2D convolution operator, $\mathcal{M}_{:,:,k}$ is the channel k of $\mathcal{M}$ expressed in the $U \times V$ matrix form, and $\mathcal{F}_{:,:,k}^{(j)}$ is the channel k of $\mathcal{F}^{(j)}$, namely the 2D kernel of $K \times K$.

Generally, the batch normalizations are conducted to reduce overfitting and accelerate the training process of the network, and then the expression capability is enhanced by nonlinear mapping. Therefore, compared to (11), the output channel j becomes

$$\mathcal{O}_{:,:,j} = (\sum_{k=1}^{C} \mathcal{M}_{:,:,k} \odot \mathcal{F}_{:,:,k}^{(j)} - \mu_j)\frac{\gamma_j}{\sigma_j} + \beta_j, \tag{12}$$

where μ_j and σ_j are the mean value and standard deviation of batch standardization, γ_j and β_j are the scaling factor and bias term, respectively.

In order to strengthen the features extracted by 3D convolution kernel, the 3×3 convolution is split into three parallel convolutions including square 3×3, asymmetric 1×3 and 3×1 convolution kernels, and then the multi-scale features extracted by convolution kernels with different sizes are added together to enrich the feature spatial information. Therefore, the output channel j by AC block is

$$\tilde{\mathcal{O}}_{:,:,j} = \check{\mathcal{O}}_{:,:,j} + \bar{\mathcal{O}}_{:,:,j} + \hat{\mathcal{O}}_{:,:,j} \tag{13}$$

where $\check{\mathcal{O}}_{:,:,j}$, $\bar{\mathcal{O}}_{:,:,j}$ and $\hat{\mathcal{O}}_{:,:,j}$ represent the outputs of 3×3, 1×3 and 3×1 branch, respectively.

2) DILATED CONVOLUTION
Traditional convolutional neural networks are made of convolutional layers, and pooling layers are stacked continuously to deepen the network layers and improve the performance. However, as the network continues to deepen, the presence of pooling layers makes the image size smaller and smaller, and inevitably some information will be lost, and thus the accuracy of the final result will be affected. Based on this problem, we introduce dilated convolution to improve its deficiency [18]. The dilated convolution effectively extends the field of perception without using stride and without losing the resolution of the TFIs, and does not add an additional burden to the network training. A 2D dilated convolution with filter $\mathcal{X}$ of size $K \times K$ is defined as

$$\mathcal{P}_{:,:,c} = \sum_{k_1}^{K} \sum_{k_2}^{K} \mathcal{X}_{:+d \times k_1, :+d \times k_2, c} \times \mathcal{W}_{k_1, k_2} \tag{14}$$

where $\mathcal{X}_{:,:,c}$ represents a feature map and $\mathcal{P}_{:,:,c}$ is the corresponding output. k_1, k_2 denote the location in the filter $\mathcal{W}$. The dilation rate d corresponds to the stride. By changing the dilation rate, the receptive field size of its filters can be effectively enlarged.

3) ResNeXt BLOCK
Then, the feature $\tilde{\mathcal{O}} \in \mathbb{R}^{R \times T \times D}$ extracted by AC block is used as input of the ResNeXt block, and ResNeXt extends residual learning from single-path convolution to multi-path group convolution. The data $\tilde{\mathcal{O}}$ is sent to multiple paths whose number is a branch base, and convolution calculation is conducted independently for each path, and the calculation result is connected based on channels. Therefore, the ResNeXt block output is

$$\mathcal{K} = F_{skip}(\tilde{\mathcal{O}}) + \mathcal{U} \tag{15}$$

where

$$\mathcal{U} = \sum_{i=1}^{G} \mathcal{T}_i(\tilde{\mathcal{O}}) \tag{16}$$

is an aggregated transformations in Figure. 5(a), $\mathcal{T}_i$ is a bottleneck structure consisting of 1×1 convolution, dilation convolution, and 1×1 convolution, G is the same number of structures in the residual block, $F_{skip}(\cdot)$ is 1×1 convolution function, and $\mathcal{K}$ is the output of ResNeXt block.

4) CA BLOCK
The feature maps $\mathcal{U}$ act as the input to CA block. CA block is divided into embedding of coordinate information and generation of coordinate attention to capture the relationship between location information and channel information. An example is the coordinate attention mechanism in Figure. 5(b).

In the embedding of coordinate information, for an input $\mathcal{U} \in \mathbb{R}^{Q \times P \times N}$, pool kernels of size $(Q, 1)$ and $(1, P)$ are used to encode information of each channel along the horizontal and vertical directions, respectively. The specific processes of pooling are given as follow

$$z_c^h(h) = \frac{1}{P} \sum_{j=1}^{P} u_c(h, j). \tag{17}$$

$$z_c^w(w) = \frac{1}{Q} \sum_{j=1}^{Q} u_c(i, w). \tag{18}$$

where z_c is the output of the c channel, Q and P denote the height and width of the input feature map, respectively, u_c is the input of the c channel.

The above two transformations aggregate $\mathcal{U}$ along two spatial directions, respectively, to obtain a pair of directional perceptual feature maps $z^h \in \mathbb{R}^{1 \times 1 \times N}$ and $z^w \in \mathbb{R}^{1 \times 1 \times N}$, which help the network to locate the target of interest more accurately.

In the generation of coordinate attention, we concatenate z^h and z^w and then transform them with $F_{skip}(\cdot)$, yielding

$$\boldsymbol{v} = \delta(F_{skip}([\boldsymbol{z}^h, \boldsymbol{z}^w])), \tag{19}$$

where [,] is the concatenation operation along the spatial dimension, δ is a non-linear activation function, and $\boldsymbol{v} \in \mathbb{R}^{N/\zeta \times (Q+P)}$ is feature map that encodes spatial information in the horizontal and vertical directions. Here, N is the number of channels, and ζ is the reduction rate.

The $\boldsymbol{v}$ is decomposed into two independent tensors $\boldsymbol{v}^h \in \mathbb{R}^{N/\zeta \times Q}$ and $\boldsymbol{v}^w \in \mathbb{R}^{N/\zeta \times P}$ along the spatial dimension, and then use $F_{skip}(\cdot)$ to transform $\boldsymbol{v}^h$ and $\boldsymbol{v}^w$ into tensors input with the same number of channels as $\mathcal{U}$, respectively, and finally use the sigmoid activation function σ to process to obtain $\boldsymbol{g}^h$ and $\boldsymbol{g}^w$, which are define as

$$\boldsymbol{g}^h = \sigma(F_{skip}(\boldsymbol{v}^h)), \tag{20}$$

$$\boldsymbol{g}^w = \sigma(F_{skip}(\boldsymbol{v}^w)), \tag{21}$$

As shown in Figure. 5(c), the output $\mathcal{Z} \in \mathbb{R}^{H \times W \times N}$ of ACDCA-ResNeXt block is

$$\mathcal{Z} = F_{skip}(\tilde{\mathcal{O}}) + \mathcal{U} \times \boldsymbol{g}^h \times \boldsymbol{g}^w. \tag{22}$$

In summary, we improve the spatial information extraction ability of the convolutional layer through the AC block, and enhance its information extraction ability in position and channel features by introducing the CA block into the ResNeXt block. Meanwhile, we replace the normal convolution in the top layer of ResNeXt block with the dilated convolution to increase its perceptual field to obtain the finer features of TFI. As shown in Figure. 6, the proposed network consists of four ACDCA-ResNeXt blocks. To input the pre-processed TFIs, ACDCA-ResNeXt network is used to extract and learn the deep features of different modulation types, and then Softmax classifier is used to classify them.

B. FUSION LOSS FUNCTION

Due to the high similarity between the different classes of samples in this paper. In order to improve the recognition ACC and generalization ability of the network, the gap between the classes is increased and the gap within the classes is reduced so that the samples can achieve better clustering effect. Therefore, we propose an effective $\mathcal{L}_F$ function combining $\mathcal{L}_{SLSCE}$ function [21], [22] and $\mathcal{L}_C$ function [23].

Suppose that there are l samples $(\boldsymbol{x}^{(1)}, y^{(1)}), \ldots, (\boldsymbol{x}^{(l)}, y^{(l)})$ in a training batch, in which $\boldsymbol{x}^{(i)}$ is a training sample i, $y^{(i)}$ is the real label of $\boldsymbol{x}^{(i)}$. The cross-entropy loss ($\mathcal{L}_{CE}$) function is as follows

$$\mathcal{L}_{CE} = -\frac{1}{l} \sum_{i=1}^{l} \sum_{j=1}^{J} y_j^{(i)} \log(p_j^{(i)}) \tag{23}$$

$$p_j^{(i)} = \frac{e^{r_j}}{\sum_{n=1}^{J} e^{r_n}} \tag{24}$$

where $\boldsymbol{y}_j^i$ represents one-hot encoding vector converting the real label of the sample i into the J-dimension; J is the classification number of LPI radar signals; r_j represents the prediction probability that the sample i falls into the j class; $p_j^{(i)}$ is the prediction probability of the sample i.

$\mathcal{L}_{SLSCE}$ function reduces the risk of overfitting the network training by adding the labels of noise smoothing real labels, and avoids the loss of information between similar classes to a certain extent by smoothing the $p_j^{(i)}$. The $\mathcal{L}_{SLSCE}$ function is as follows:

$$\mathcal{L}_{SLSCE} = -\frac{1}{l} \sum_{i=1}^{l} \sum_{j=1}^{J} \bar{y}_j^{(i)} \log(p_j^{'(i)}) \tag{25}$$

$$p_j^{'(i)} = \frac{e^{(\frac{r_j}{S})}}{\sum_{n=1}^{J} e^{(\frac{r_n}{S})}} \tag{26}$$

$$\bar{\boldsymbol{y}}_j^{(i)} = (1 - \vartheta) \boldsymbol{y}_j^{(i)} + (1 - \boldsymbol{y}_j^{(i)}) \frac{\vartheta}{J} \tag{27}$$

where ϑ is label smoothing coefficient, and $\bar{\boldsymbol{y}}_j^{(i)}$ is the real label after label smoothing. The greater the soft coefficient S, the smoother the probability distribution

In order to narrow the intra-class gap and maintain the gap between the classes, $\mathcal{L}_C$ function is adopted in this paper. The formula for it is as follows

$$\mathcal{L}_C = \frac{1}{2} \sum_{i=1}^{l} \| f(\boldsymbol{x}^{(i)}) - C_{y^{(i)}} \|_2^2 \tag{28}$$

where $f(\boldsymbol{x}^{(i)})$ represents the feature vector of the sample i before the fully connected layer, and $C_{y^{(i)}}$ represents the center feature of the category to which the sample i belongs.

Specifically, the $\mathcal{L}_F$ is formulated as follows

$$\mathcal{L}_F = \mathcal{L}_{SLSCE} + \alpha \mathcal{L}_C \tag{29}$$

where α is used to balance the two losses.

V. EXPERIMENTS AND RESULTS

In this section, we analyze the performance of the proposed network in detail and the computational complexity and then compare the performance of the proposed algorithm with other approaches at different SNRs. In addition, we visualize the regions of TFIs that are of interest to different approaches by Grad-CAM.

A. EXPERIMENTAL PARAMETERS

In the training of ACDCA-ResNeXt, SDG and Adam were used as optimizers for center loss and cross entropy loss. The initial learning rate was 0.1, the batch of training batches was 64, and the total number of iterations was 90. When the loss of test and verification set did not decrease for 3 consecutive times, the learning rate would be reduced by 30%. All networks were implemented on the PyCharm Community 2018 Platform using PyTorch 1.8.1 and CUDA 10.1. The computer configuration was Intel® Core 17-10700 2.9 GHz, 16 GB RAM, and NVIDIA GeForce GTX 1060 with 6GB hardware capabilities.
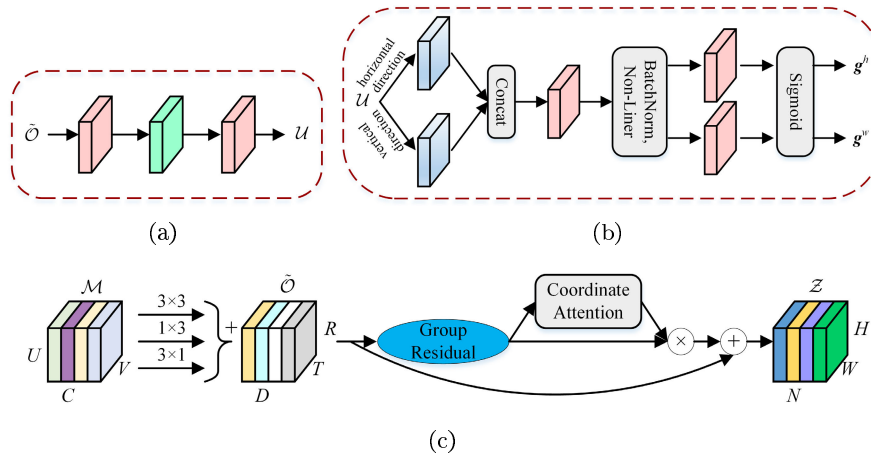
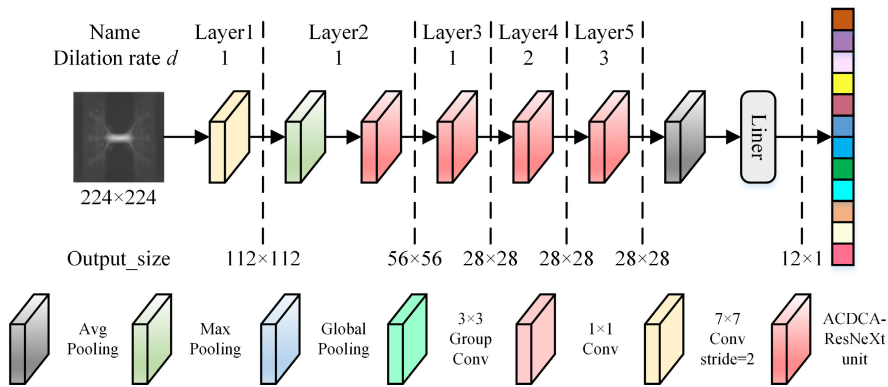**FIGURE 5.** The block of (a) Group Residual, (b) Coordinate Attention, and (c) proposed ACDCA-ResNeXt.

FIGURE 6. The structure of the ACDCA-ResNeXt.

B. THE EFFECTS OF TFI PREPROCESSING ON RECOGNITION ACC

We compare the effects of different preprocessing operations on recognition ACC. Qu et al. [10] binarized and scaled the TFI after 2D Wiener filtering, which improved the recognition effect. However, at a low SNR, the binary image information was seriously lost, resulting in low classification ACC [10]. After the preprocessing method proposed in this paper, the noise interference and the position difference in the frequency dimension in TFI can be reduced. The data after preprocessing mentioned above were used to train ACDCA-ResNeXt separately, and the overall recognition rate was shown in Figure. 7. Also, the TFIs after 2D Wiener filtering were also used as the network input. The experimental results demonstrated that the overall recognition ACC of LPI radar signal by the preprocessing approaches proposed in this paper was higher than other similar preprocessing approaches at different SNRs.

In addition, we discussed the effect of the input layer sizes (56×56, 112×112, 224×224 and 448×448) of ACDCA-ResNeXt on recognition effect. As shown in Figure. 8, the smaller the size (e.g., 56×56), the more detailed
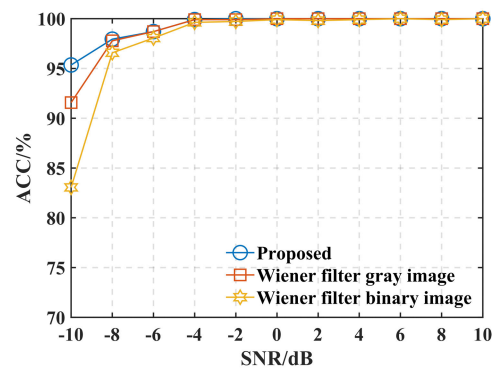
FIGURE 7. Comparison of recognition ACC of different preprocessing methods.

information (e.g., frequency hopping at the end of the TFI of BPSK signal) lost in the TFI. This will make it difficult to recognize the signals with similar time-frequency morphological characteristics under the noise effect. On contrary, if the input layer of the network is too large (e.g., 448×448), it increases the parameters of the network while keeping the size of the

convolutional kernel constant, which leads to overfitting and thus reduces the performance. Therefore, 224×224 was reasonably selected as the input size of the network.
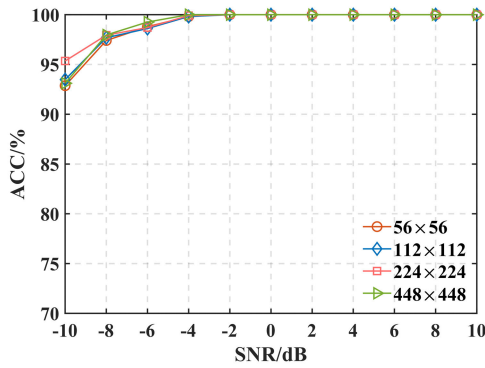

FIGURE 8. The effect of the input layer size of the ACDCA-ResNeXt on recognition effect.

C. COMPARISON WITH DIFFERENT LOSS FUNCTIONS

The effect of different S values on the classification ACC of ACDCA-ResNeXt was determined through experiments. As shown in Figure. 9, when there were 2100 training samples in each category and the S value was between 1.5 and 5, the best classification ACC was obtained if S was 3. Therefore, the optimal S could be determined by the parameter search approach. In the same way, the weights of $\vartheta = 0.3$ and $\alpha = 9$ could be determined by this approach.
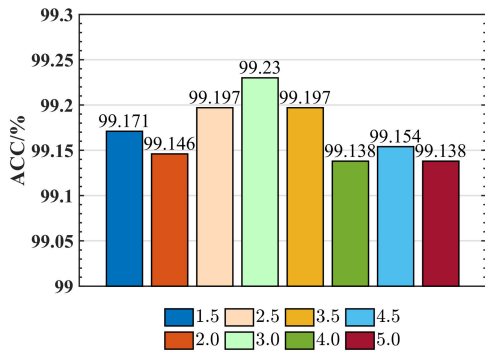
FIGURE 9. Classification results of the ACDCA-ResNeXt under different S values.

To better understand the classification effect of the proposed network with different loss functions, 90 samples were randomly selected from each class of test dataset, and t-distributed stochastic neighbor embedding (t-SNE) [24] was used to reduce the dimensionality of inputs and features extracted by ACDCA-ResNeXt with $\mathcal{L}_{CE}$, ACDCA-ResNeXt with $\mathcal{L}_{SLSCE}$, and ACDCA-ResNeXt with $\mathcal{L}_F$ to 2D representation. The results were shown in Figure. 10. It could be seen that $\mathcal{L}_F$ could significantly improve the generalization ability of the network and achieved better clustering effect.
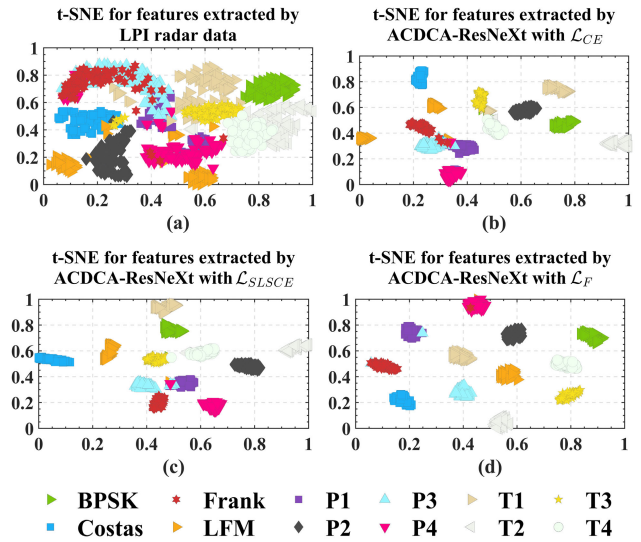

FIGURE 10. Visualize of twelve signals categories in of different loss functions.

D. RECOGNITION PERFORMANCE OF THE PROPOSED APPROACH

To demonstrate the recognition performance of the proposed approach, we compared the approach with CNN [10], dilated residual network (DRNet) [12] and ResNet [25]. Meanwhile, ResNeXt was taken a control to verify the improvement of ResNeXt recognition performance by dilated convolution, AC block and CA block. Literature [10] used CNN to extract the features of the binary TFI after 2D Wiener filtering. Literature [12] and Literature [25] adopted the preprocessing operations of graying, opening operation and bicubic interpolation, and then used DRNet and ResNet to extract the features. Finally, they all used the Softmax classifier to complete the LPI radar signals classification.
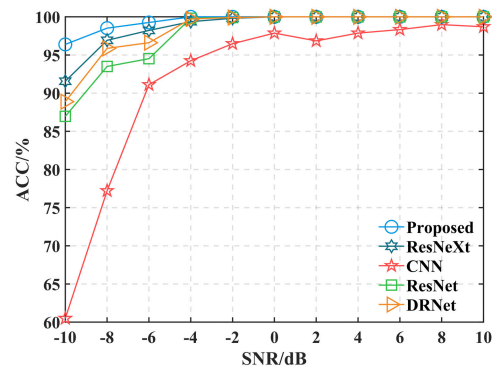

FIGURE 11. Performance comparison of five approaches at different SNRs.

Figure. 11 showed the overall ACC of the above five approaches at different SNRs. When SNR>-4 dB, the ACC of other approaches except CNN exceeded 97%. With the decrease of SNR, the recognition curve of the control approach declined significantly, while the recognition ACC

of the proposed approach could still achieve 97.94% when SNR=−8 dB, which were 20.72%, 4.44%, 2.07% and 1.02% higher than CNN, ResNet, DRNet and ResNeXt, respectively. When SNR=−10 dB, the proposed approach could still achieve high recognition ACC for the test data that were not involved in training, and its generalization capability outperformed other approaches.

The following verifies the respective recognition ACC of different approaches for twelve types of LPI radar signals at different SNRs, as shown in Figure. 12. With the decrease of SNR, the identification ACC of our approach is significantly improved compared with other approaches for LFM, P1, P3, P4, BPSK, and T2. Figure. 13 showed the recognition confusion matrix at SNR=−8 dB. Recognition errors mainly occurred in the P1 and P3 codes, and the P2, P3 and Frank codes for both groups of signals, which is due to the fact that the step characteristics in the TFIs of the two groups of signals became vague or even disappeared at low SNRs, resulting in confusion between the signals.

E. VISUALIZATION RESULTS OF DIFFERENT APPROACHES

In order to analyze the recognition performance of different networks in the previous section, the three networks (CNN, ResNet, and DRNet) were trained with the same data set, and the heap map of the confusing signal samples (P1, P3, P4 and Frank) in the test set were generated by the visualization approach Grad-CAM [24], as shown in Figure. 14. When CNN and ResNet recognized the above test samples, they were easily affected by background noise and thus scattered and incomplete attention areas appeared. DRNet, based on the residual block, introduced dilated convolution to increase the receptive field. It focused on the edge area of signal time-frequency energy, but the part of interest was still missing. Under the premise of introducing dilated convolution, we analyzed the focus areas of D-ResNeXt, ACD-ResNeXt and CAD-ResNeXt when recognizing the four types of confusing signals, as shown in Figure. 15. To sum up, the network proposed could pay attention to the edge features of energy distribution in TFIs by using dilated convolution, AC block and CA block. It was not easily affected by background noise, and the attention was more evenly distributed.

F. THE COMPUTATIONAL OF OTHER ATTENTION MECHANISMS

As TFI was rich in details, we should pay attention to the more useful parts when recognizing specific samples. In order to obtain a better classification effect, attention mechanism was introduced into the network, which was usually used to tell the network what content and attention should be concerned. Starting from the visualization and computational complexity and based on ACD-ResNeXt, we analyzed the performances of three different attention mechanisms, squeeze-and-excitation (SE) [27], convolutional block attention module (CBMA) [28] and CA.
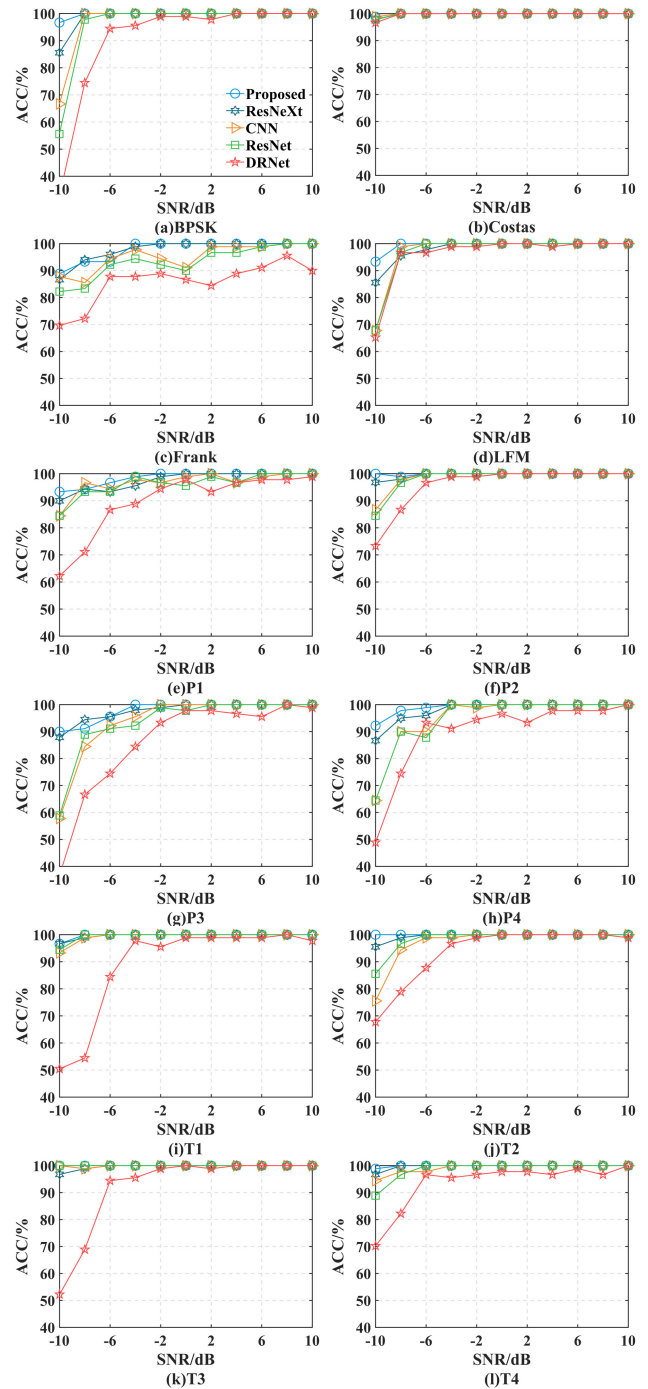

FIGURE 12. Recognition ACC of twelve radar signals at different SNRs.

1) VISUALIZATIONS RESULTS OF DIFFERENT ATTENTION MECHANISMS

As shown in Figure. 16, as SE only considers the encoding of features between the channels, it neglected the importance of location information. CBAM only obtained the local correlation between location and channel by using a large pooling layer. The CA used in this paper encoded the information of different channels along the horizontal and vertical directions and integrated the location and channel information more

FIGURE 13. Confusion matrix for waveform classification (SNR=−8 dB).
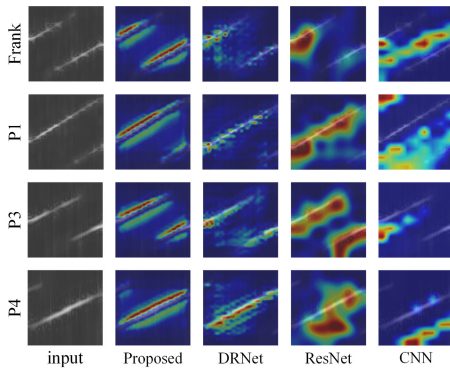

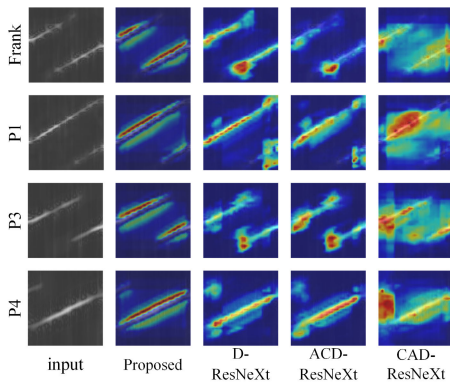**FIGURE 14.** Grad-CAM visualization results obtained using different approaches.

FIGURE 15. Grad-CAM visualization results obtained using the ResNeXt block with different blocks.

FIGURE 16. Grad-CAM visualization results obtained using the ACD-ResNeXt with different attention mechanisms.

convolutional layers and linear layers are as follows

$$FLOPs_c = \sum_{m=1}^{D} 2M_m^2 E_m^2 L_{m-1} L_m \qquad (30)$$

$$FLOPs_l = \sum_{m=1}^{D} (2L_{m-1} - 1)L_m \qquad (31)$$

where $FLOPs_c$ and $FLOPs_l$ represent the $FLOPs$ of convolution and linear layers, respectively, L and D are the numbers of layers and channels, respectively, and M and E denote the lengths of feature maps and kernels of the l layer.

In order to verify the performances of three different attention mechanisms, the ratios of the increased ACC to the increased parameters and $FLOPs$ were defined as $Score_1$ and $Score_2$ by comparing with the baseline approach. The $Score_1$ and $Score_2$ are shown as

$$Score_1 = \frac{ACC_c - ACC_b}{P_c - P_b}, \qquad (32)$$

$$Score_2 = \frac{ACC_c - ACC_b}{F_c - F_b}, \qquad (33)$$

where ACC_c and ACC_b denote the overall ACC of current and baseline approaches, respectively, P_c and P_b denote the parameters of current and baseline approaches, respectively, and F_c and F_b denote the $FLOPs$ of current and baseline approaches, respectively.

When the $Score$ was greater, it would indicate that this approach could sacrifice a few parameters/computational complexities to achieve higher ACC than the baseline approach.

It could be seen from Table 3 that 1) the embedding attention mechanism could effectively improve ACC. The SE and CBAM blocks were increased by 0.05% and 0.2% respectively when compared with the baseline approach without any attention mechanism; 2) by using CA block, the location and channel information were integrated more harmoniously, and the recognition ACC achieved 99.23%; 3) the ACDCA-ResNeXt network witnessed the highest $Score_1$ and $Score_2$, which indicated that CA block could sacrifice the

harmoniously, which would have more advantages in the extraction of TFI features.

2) COMPARISONS OF DIFFERENT ATTENTION MECHANISMS

Floating point of operations (FLOPs) and learning parameters are widely applied in the computational complexity of convolutional neural networks. The FLOPs calculations of
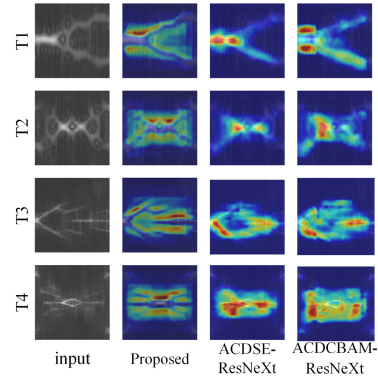
least model parameters and algorithm complexity to obtain a higher ACC.

TABLE 3. Result comparisons under different attention mechanisms.

Settings	$Params$	$FLOPs$	ACC	$Score_1$	$Score_2$
ACD-ResNeXt	2.09M	1783.80M	98.66%	-	-
+SE	2.29M	1790.91M	98.71%	0.25	0.007
+CBAM	2.29M	1790.91M	98.86%	1.00	0.028
+CA	2.36M	1798.14M	99.23%	2.11	0.040

VI. CONCLUSION

In this paper, we propose a recognition approach of LPI radar modulation signal based on TFI preprocessing and ACDCA-ResNeXt network. Firstly, in order to eliminate noise interference and redundant frequency band interference in TFI, we propose a new TFI preprocessing. Secondly, to strengthen the learning capability of space, location and channel features of network, we introduce AC block and CA block based on ResNeXt and increase the receptive field by using dilated convolution. Finally, joint training of center loss and soft label smoothing is proposed to overcome the problems of insufficient generalization capability of cross entropy loss and unsatisfactory clustering effect. Experimental results demonstrate that, for twelve kinds of LPI radar waveforms, the overall recognition ACC of the proposed approach achieves 97.94% when the SNR is -8 dB. Compared with the existing approaches, the proposed approach gets a better adaptability to low SNR and a stronger anti-confusion capability.

Finally, taking into account that Wiener filtering requires manual parameter adjustment and is difficult to be applied in practical engineering applications, future research can explore the use of deep learning networks to achieve blind denoising of time-frequency images and further suppress noise in images.

REFERENCES

[1] P. E. Pace, *Detecting and Classifying Low Probability of Intercept Radar*, 2nd ed. Norwood, MA, USA: Artech House, 2009.

[2] P. Misans and M. Terauds, "CW Doppler radar based land vehicle speed measurement algorithm using zero crossing and least squares method," in *Proc. 13th Biennial Baltic Electron. Conf.*, Oct. 2012, pp. 161–164.

[3] X. Teng, P. Tian, and H. Yu, "Modulation classification based on spectral correlation and SVM," in *Proc. Int. Conf. Wireless Commun., Netw. Mobile Comput.*, Oct. 2008, pp. 1–4.

[4] C. Shiwen, W. Gongming, X. Xiaopeng, and H. Jie, "A method of radar signal feature extraction based on fractional Fourier transform," in *Proc. IEEE 4th Int. Conf. Signal Image Process. (ICSIP)*, Jul. 2019, pp. 583–587.

[5] N. Lopac, F. Hrzic, I. P. Vuksanovic, and J. Lerga, "Detection of non-stationary GW signals in high noise from Cohen's class of time-frequency representations using deep learning," *IEEE Access*, vol. 10, pp. 2408–2428, 2022.

[6] A. Gao, Y. Zhu, W. Cai, and Y. Zhang, "Pattern recognition of partial discharge based on VMD-CWD spectrum and optimized CNN with cross-layer feature fusion," *IEEE Access*, vol. 8, pp. 151296–151306, 2020.

[7] Z. Feng, M. Liang, and F. Chu, "Recent advances in time–frequency analysis methods for machinery fault diagnosis: A review with application examples," *Mech. Syst. Signal Process.*, vol. 38, no. 1, pp. 165–205, Jul. 2013.

[8] M. D. M. Zhang and L. Guo, "Convolutional neural networks for automatic cognitive radio waveform recognition," *IEEE Access*, vol. 5, pp. 11074–11082, 2017.

[9] S.-H. Kong, M. Kim, L. M. Hoang, and E. Kim, "Automatic LPI radar waveform recognition using CNN," *IEEE Access*, vol. 6, pp. 4207–4219, 2018.

[10] Z. Qu, X. Mao, and Z. Deng, "Radar signal intra-pulse modulation recognition based on convolutional neural network," *IEEE Access*, vol. 6, pp. 43874–43884, 2018.

[11] Z. Ma, Z. Huang, A. Lin, and G. Huang, "LPI radar waveform recognition based on features from multiple images," *Sensors*, vol. 20, no. 2, p. 526, Jan. 2020.

[12] X. Qin, J. Huang, X. Zha, L. P. Luo, and D. X. Hu, "Radar emitter signal recognition based on dilated residual network," *Acta Electronica Sinica*, vol. 48, no. 3, pp. 456–462, Mar. 2020.

[13] X. Ni, H. Wang, F. Meng, J. Hu, and C. Tong, "LPI radar waveform recognition based on multi-resolution deep feature fusion," *IEEE Access*, vol. 9, pp. 26138–26146, 2021.

[14] Q. Guo, X. Yu, and G. Ruan, "LPI radar waveform recognition based on deep convolutional neural network transfer learning," *Symmetry*, vol. 11, no. 4, p. 540, Apr. 2019.

[15] D. Hu, "Research on denoise and detection technology of low probability radar radiation source signal," M.S. thesis, Dept. Electron. Eng., Hit Univ., Harbin, China, 2017.

[16] X. Wang and L. Yu, "Multi-channel self-correlation signal detection algorithm and its FPGA implementation," *Chin. J. Sci. Instrum.*, vol. 28, no. 5, pp. 875–881, May 2007.

[17] S. Xie, R. Girshick, P. Dollar, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recongnit.*, Honolulu, HI, USA, Dec. 2017, pp. 1492–1500.

[18] X. Ding, Y. Guo, G. Ding, and J. Han, "ACNet: Strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jun. 2019, pp. 1911–1920.

[19] Q. Hou, D. Zhou, and J. Feng, "Coordinate attention for efficient mobile network design," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2021, pp. 13713–13722.

[20] F. Yu, V. Koltun, and T. Funkhouser, "Dilated residual networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recongnit.*, Dec. 2017, pp. 472–480.

[21] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2818–2826.

[22] G. Shao, Y. Chen, and Y. Wei, "Deep fusion for radar jamming signal classification based on CNN," *IEEE Access*, vol. 8, pp. 117236–117244, 2020.

[23] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 499–515.

[24] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.

[25] X. Qin, X. Zha, J. Huang, and L. Luo, "Radar waveform recognition based on deep residual network," in *Proc. IEEE 8th Joint Int. Inf. Technol. Artif. Intell. Conf. (ITAIC)*, May 2019, pp. 892–896.

[26] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 336–359.

[27] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Feb. 2018, pp. 7132–7141.

[28] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19.

XUDONG WANG received the B.S., M.S., and Ph.D. degrees in communication engineering from the Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, in 2001, 2004, and 2008, respectively. He is currently an Associate Professor. His research interests include radar signal processing and FPGA hardware implementation.

GUIGUANG XU received the B.S. degree in electronic information engineering from the East China University of Technology, Jiangxi, China, in 2018. He is currently pursuing the M.S. degree with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronatics (NUAA), Nanjing. His current research interests include radar signal recognition and machine learning.

YING WEN received the B.S. degree in electronic information engineering from Jiangxi Normal University, Jiangxi, China, in 2021. She is currently pursuing the M.S. degree with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China. Her current research interests include radar signal recognition and machine learning.

HE YAN received the B.S. degree from Jilin University, Changchun, China, in 2008, and the Ph.D. degree in communication and information systems from the University of Chinese Academy of Sciences, Beijing, China, in 2013.

From 2013 to 2014, he was an Engineer with the 14th Research Institute of China Electronics Technology Group Corporation (CETC-14), Nanjing, China. Since 2015, he has been an Associate Professor with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing. His main research interests include synthetic aperture radar-ground moving target indication (SAR-GMTI), wide area surveillance-ground moving target indication (WAS-GMTI), video synthetic aperture radar-ground moving target indication (ViSAR-GMTI), and ocean current measurements based on spaceborne SAR.

DAIYIN ZHU received the B.S. degree in electronic engineering from Southeast University, Nanjing, China, in 1996, and the M.S. and Ph.D. degrees in electronics from the Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, in 1998 and 2002, respectively.

From 1998 to 1999, he was a Guest Scientist with the Institute of Radio Frequency Technology, German Aerospace Center, Oberpfaffen-hofen, Germany, where he was involved in the field of SAR interferometry. In 1998, he joined the Department of Electronic Engineering, NUAA, where he is currently a Professor. He has developed algorithms for several operational airborne SAR systems. His current research interests include radar imaging algorithms, SAR/ISAR autofocus techniques, SAR ground moving target indication (SAR/GMTI), and SAR interferometry.
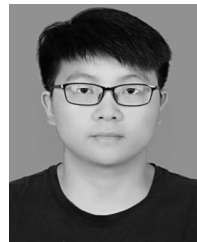
ZEHU LUO received the B.S. degree in communication engineering from the Anhui University of Technology, Maanshan, China, in 2019. He is currently pursuing the M.S. degree with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China. His current research interests include radar signal processing and machine learning.

• • •