IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 29, 2021, accepted February 5, 2021, date of publication February 9, 2021, date of current version February 17, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3058305

LPI Radar Waveform Recognition Based on Multi-Resolution Deep Feature Fusion

XUE NI[1], HUALI WANG[1], FAN MENG[2], JING HU[1], AND CHANGKAI TONG[1]
[1]College of Communications Engineering, Army Engineering University of PLA, Nanjing 210007, China
[2]Nanjing Marine Radar Institute, Nanjing 210007, China

Corresponding author: Huali Wang (huali.wang@ieee.org)

ABSTRACT Deep neural networks are used as effective methods for the Low Probability of Intercept (LPI) radar waveform recognition. However, existing models' performance degrades seriously at low Signal-to-Noise Ratios (SNRs) because the effective features extracted by the networks are insufficient under noise jamming. In this paper, we propose a multi-resolution deep feature fusion method for LPI radar waveform recognition. First, we apply the enhanced Fourier-based Synchrosqueezing Transform (FSST), which shows good performance at low SNRs, to convert radar signals into time-frequency images. Then, we construct a multi-resolution deep convolutional network to extract more deep features from each resolution channel. Next, we explore an interactive feature fusion strategy for deep feature fusion. By some down-sampling or up-sampling blocks, different resolution features are fused to generate new features. Finally, we apply a fusion algorithm to the fully connected layer to achieve classification fusion for better performance. Simulation experiments on twelve kinds of LPI radar waveforms show that the overall recognition accuracy of our method can reach 95.2% at the SNR of -8 dB. It is proved that our approach does indeed improve the recognition accuracy effectively at low SNRs.

INDEX TERMS Radar waveform recognition, multi-resolution, feature fusion, FSST, convolutional neural network.

I. INTRODUCTION

LPI radar, not easily intercepted by non-cooperative receivers, has been widely used in modern radar systems. They often use various modulation waveforms with variable parameters to improve anti-interception performance. Randomly altering one or more of the modulation parameters can confuse the intercept receivers, especially in a complex environment of high noise interference and multiple signals [1], [2]. Therefore, LPI radar waveform recognition has become a tough challenge in Electronic Warfare (EW).

Traditional methods relying on artificial experience to select features fail to recognize modulation waveforms in a complex environment [3], so they cannot meet the needs of modern intelligent recognition systems. Recently, with the successful application of deep learning in image classification and other fields [4], [5], scholars are devoted to converting one-dimensional radar signal into two-dimensional feature images by signal processing and using the deep learning models to extract image features automatically for classification.

In order to exhibit the difference among various radar signals, signal processing based on Time-Frequency (T-F) analysis are employed to handle modulated radar waveforms, such as Short-Time Fourier Transform (STFT), Wigner Ville Distribution (WVD) [6], Choi-William Distribution (CWD) [7]–[10], and Cohen class Time-Frequency Distribution (CTFD) [11]. CWD has better performance in cross-term suppression and higher quality in T-F images, but it has high algorithm complexity and performs poorly at low SNRs. The Fourier-based Synchrosqueezing Transform (FSST) [12] is an effective T-F analysis method for radar signals, which provides a sharp and concentrated feature representation and is robust to noise and small perturbations. However, the frequency squeezing makes few features available for identification.

Deep learning networks have been successfully applied in radar signal recognition tasks. After a series of image processing, T-F images are sent into the deep network model for automatic feature extraction and classification. In [13],

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

Zhang converted the T-F images to binary images with 32 × 32 pixels as inputs and adopted the Convolutional Neural Network (CNN) and Elman Neural Network (ENN) for eight radar signals recogniton. The recognition rate achieved 94.5% at SNR of -2 dB. In [11], Qu used 64 × 64 T-F images for the input layer of CNN, and explored a deep Q-learning network to improve the probability of successful recognition. In [8], the authors utilized CWD to obtain T-F images and constructed a six-layer CNN for LPI radar waveform recognition. Experiments verified the influence of different input sizes on recognition accuracy. The results showed that CNN with 128 × 128 inputs obtained a 93.58% recognition rate at SNR of -6 dB. Authors in [10] applied the ResNet for complex multiple radar waveforms. In addition to the CNN network, the Stacked Auto Encoder(SAE) model had also been applied to radar waveform classification [14]. Besides, some scholars used deep learning networks to extract the features of T-F images and selected the Tree-based Pipeline Optimization Tool (TPOT) or Support Vector Machine (SVM) for classification [9], [15]. All the above methods have made some progress in improving radar waveform recognition performance, but most methods showed severe recognition performance degradation when the SNR was less than -6 dB. Therefore, how to improve the radar waveform recognition performance at low SNRs is a challenging problem. We note that T-F images with different resolutions have different features that can be used for identification. T-F images with high resolution have more detailed features, while T-F images with low resolution contain more geometric features. Obviously, the features extracted from single-resolution inputs are insufficient to distinguish various radar signals at low SNRs, especially similar signals, making it difficult to improve recognition performance.

To solve the above problem, scholars have adopted some fusion strategies to obtain more effective features. In [16], a T-F image fusion strategy and a multi-feature fusion algorithm were used for radar signal recognition. In [17], the features extracted from CNN were fused with shallow features to obtain more features in image expression. In addition, designing a network structure to extract features with multiple resolutions could also improve recognition performance, which had proven to be an effective method in object detection. In our previous work [18], we designed a dual-channel CNN structure to extract more features. However, we only used simple feature fusion for two resolutions. Considering that the features at different resolutions have different contributions to the recognition performance, we will try to use T-F images with multiple resolutions for network training and improve the fusion strategy to extract deep features.

In this paper, we propose a multi-resolution deep feature fusion method for LPI radar waveform recognition, to address the problem of insufficient features extracted from deep learning networks under noise jamming. Our work mainly focuses on three aspects: to select a better T-F representation method, to construct a multi-resolution network structure, and to develop an effective fusion strategy. The major contributions to this paper are list as follows:

1) Apply an enhanced FSST for signal T-F analysis to convert radar signals into T-F images. The FSST with spectrum enhancement shows better anti-noise performance and faster operation than CWD at low SNRs.

2) Propose a deep convolutional network with a multi-resolution structure. We use down-sampling to obtain images of different resolutions in the first layer, then we construct CNNs to extract more deep features from each resolution channel. Compared with a single resolution input, the multi-resolution network structure can extract more useful features.

3) Explore an interactive feature fusion strategy instead of simple concatenation fusion. Different resolution features from each CNN are fused to generate new features by some down-sampling or up-sampling blocks. The fusion modules enable the features to be enhanced, which helps to identify similar signals.

4) Use a fusion algorithm to the Fully Connected (FC) layer with different resolutions to achieve classification fusion, and select the Error-Correcting Output Coding-Support Vector Machine (ECOC-SVM) as a classifier, which can improve the fault tolerance of the classification. It shows a significant improvement in recognition accuracy.

The paper is organized as follows. Section II introduces the classification system and waveforms. Section III presents signal processing, including time-frequency analysis (FSST) and image processing. Section IV describes the classification method in detail, and section V shows the simulation results. Finally, Section VI concludes the paper.

II. CLASSIFICATION SYSTEM AND WAVEFORMS

This section presents the definition of LPI radar waveforms considered in this paper and the structure of LPI radar classification system.

A. SYSTEM STRUCTURE

The classification system consists of three parts: signal processing, feature extraction and fusion, and classification, as shown in Figure 1. In the signal processing stage, the received LPI radar waveforms are first converted into T-F images by the enhanced FSST. Then the T-F images are processed by cutting, image resizing, and normalization. Next, a deep convolutional network with three resolutions structure is designed for feature extraction. An interactive feature fusion strategy is used for feature fusion. After that, we apply a fusion algorithm on FC features and use ECOC-SVM as a classifier to realize waveform recognition.

B. SIGNAL MODEL

Assuming that the received LPI radar signal is a pulse wave with Additive Gaussian White Noise (AWGN). Then, the signal model can be expressed as:

$$s(t) = x(t) + n(t), 0 \le t \le T \tag{1}$$

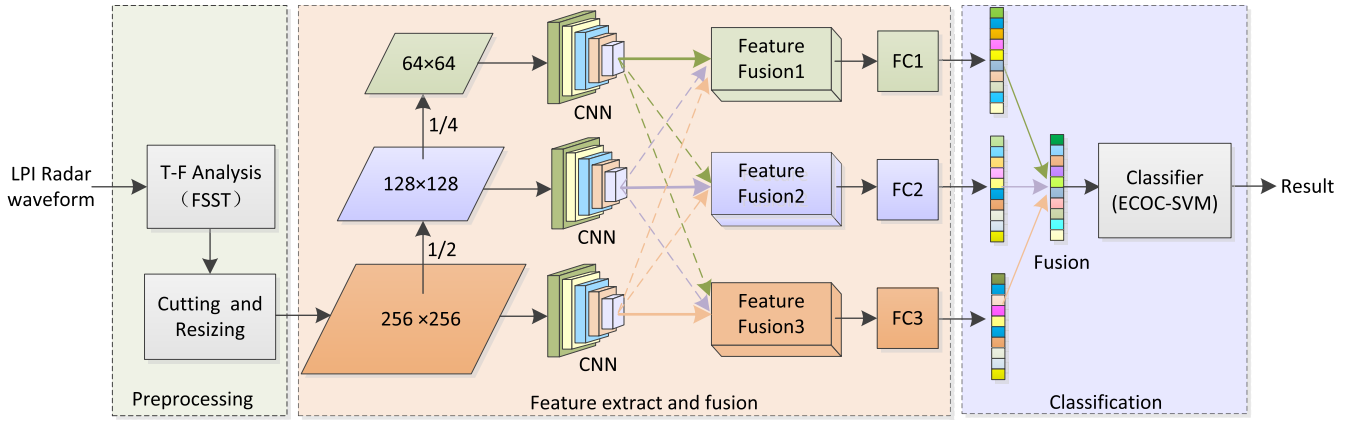

FIGURE 1. The proposed LPI radar waveform recognition system.

where

$$x(t) = A\exp(j(2\pi f(t)\cdot t + \phi(t)) \tag{2}$$

is the time complex LPI radar signal samples. A is amplitude, T is the pulse duration. Suppose that the signals have the same amplitudes, we set $A = 1$ here. $n(t)$ is white Gaussian noise with variance σ^2. The SNR of the signal is defined as $SNR = 10\log_{10}A^2/2\sigma^2$. $f(t)$ and $\phi(t)$ are the carrier frequency and phase function, respectively, which determine the modulation type of radar signal. In this paper, twelve types of intrapulse modulation LPI radar waveforms are considered [2]: BPSK, Costas, Frank, LFM, P1, P2, P3, P4, T1, T2, T3, and T4 codes.

III. SIGNAL PROCESSING

A. FSST FOR TIME-FREQUENCY ANALYSIS

The short-time Fourier based synchrosqueezing transform, introduced in [19], is a new time-frequency analysis method, which aims to provide the sharp and concentrated representation by allocating the STFT coefficient value to a different point in the T-F plane, and it has been proved to be robust to noise and small perturbations. For LPI radar waveforms, we use FSST to obtain a high-resolution T-F image, making it easier to distinguish between various received waveforms. We denote the short-time Fourier transform of signal $s(t)$ by $S(t,f)$:

$$S(t,f) = \int_{-\infty}^{+\infty} s(\tau)g(\tau - t)e^{-j2\pi f(\tau - t)}d\tau \tag{3}$$

where t and f are the time variable and the frequency variable, respectively. $g(t)$ is a window function. Limited by the size of the window, the frequency resolution of the short-time Fourier transform cannot be optimal. Starting from the STFT, FSST moves the coefficients $S(t,f)$ according to the map $(t,f) \to (t, \hat{w}(t,f))$, defined by

$$T(t,w) = \frac{1}{g(0)}\int_{-\infty}^{+\infty} S(t,f)\delta(w - \hat{w}(t,f))df \tag{4}$$

where

$$\hat{w}(t,f) = Re\left(\frac{1}{2\pi i}\frac{\partial_t(S(t,f))}{S(t,f)}\right) \tag{5}$$

is the local instantaneous frequency. δ is the Dirac delta function and $g(0)$ is the value of a sliding window $g(t)$ at time 0. In this paper, a Kaiser window of length 256 is applied.

FSST has improved the T-F feature resolution by squeezing the spectrum, but the squeezed spectrum is only with a few pixels width, which is not beneficial for waveform recognition at low SNRs. To enhance the signal T-F feature representation, we shift the T-F coefficients $T(t,w)$ in a small frequency range and add them together. The enhanced coefficients $\hat{T}(t,w)$ are expressed as follows

$$\hat{T}(t,w) = \sum_{w'=-w_0}^{w_0} T(t, w-w') \tag{6}$$

where w' is the value of frequency shift. In our paper, we set $w' = 1$ to keep noise and feature enhancement in balance. $\hat{T}(t,w)$ is a complex-valued bivariate. Here, we take the amplitude $\left|\hat{T}(t,w)\right|$ as an effective representation of the T-F images. Assuming that the length of the signal is L, we can obtain the T-F images with the size of $256\times L$. Compared with CWD, the enhanced FSST algorithm has low complexity and fast operation.

B. IMAGE PROCESSING

Before the images are sent into the classifier, image cutting and resizing are used to reduce the computational cost for the deep neural network. Considering that the T-F features of the signal are located in the upper of images, image cutting is applied to retain the upper half of the image. By the nearest neighbor technique, the T-F images are resized to 256 × 256.

The T-F images of twelve kinds of LPI radar waveforms with 256 × 256 pixels are shown in Figure 2. All the signals are generated in the condition of $SNR = 10dB$

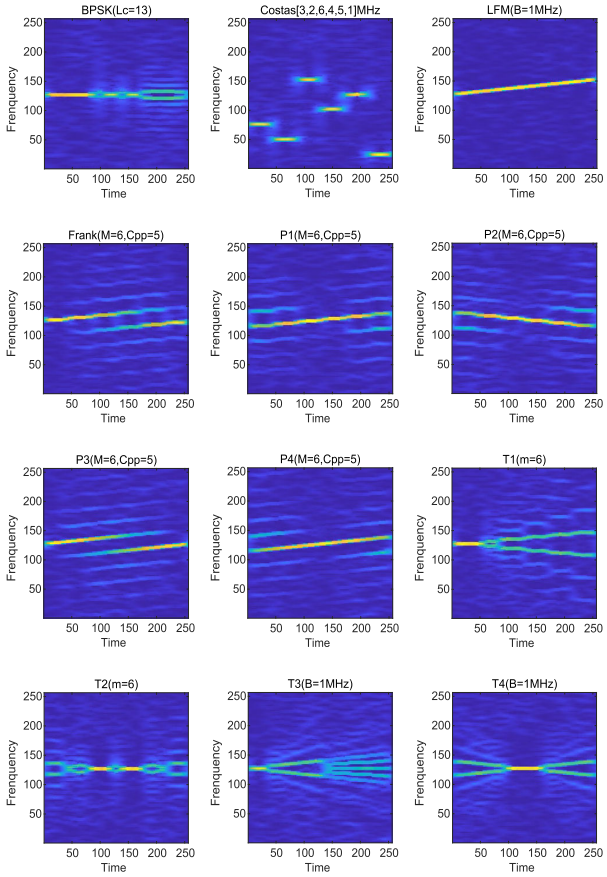

FIGURE 2. T-F images of twelve kinds of LPI radar waveforms (SNR = 10dB).

and $f_0 = 5MHz$. As shown, there are several differences between waveform categories in high SNR condition. It shows that the enhanced FSST is outstanding in terms of noise robustness and T-F feature representation. We also note that the polyphase codes have similar T-F features to LFM. Frank and P1 codes are approximately the stepped lines, while P1 and P4 are closer to the straight lines. The feature difference between polyphase codes becomes more obvious in our method, which is conducive to improve the probability of successful recognition, even at low SNRs.

IV. CLASSIFICATION METHOD

The classification method based on the deep convolutional network have four modules: multi-resolution network structure, interactive feature fusion module, classification fusion and the ECOC-SVM classifier. In this section, we will describe each module in detail.

Inspired by the feature pyramid structure [20], we design a multi-resolution structure based on CNN to extract deep features of different resolution images. The difference is that we do not directly obtain multi-resolution features from the single network's multiple layers. Figure 3 shows the

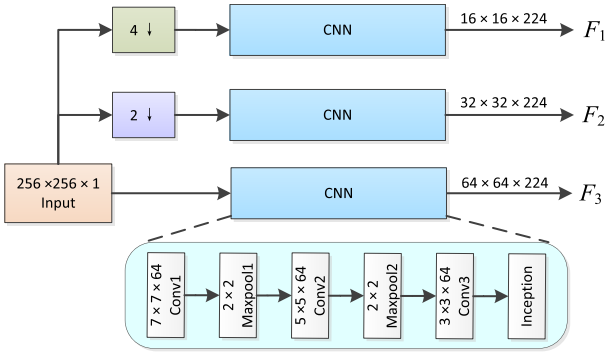

FIGURE 3. Multi-resolution network structure.

multi-resolution network structure. The input size is 256 × 256 × 1. Then, 2 × 2 and 4 × 4 maxpooling modules are used for down-sampling to generate 128 × 128 and 64 × 64 T-F images. Next, the designed CNN [18], which proved to be an effective network in our previous work, is applied to different resolution inputs.

The CNN consists of Conv1, Maxpool1, Conv2, Maxpool2, Conv3, and Inception. The convolutional kernel size for Conv1, Conv2 and Conv3 is designed as 7 × 7, 5 × 5, and 3 × 3, according to the width of the time-frequency curves in the T-F images. The numbers of the convolutional kernels are set to 64. Each convolutional layer is followed by Batch Normalization (BN) to speed up training and prevent the gradient from disappearing or exploding. We use Rectified Linear Unit (ReLU) as a nonlinear activation function. Maxpool1 and Maxpool2 are set to 2 × 2 filer size, stride size 2, and no zero-padding. We add the Inception architecture [21] after Conv3 to increase the depth of the network and extract more features at different scales.

Figure 4 shows the Inception module structure. It is composed of 1 × 1, 3 × 3, 5 × 5 convolutions and 3 × 3 avgpooling stacked upon each other. Here, the 5 × 5 convolution is dissolved is 1 × 5 and 5 × 1 convolution, which speeds up the calculation. The kernel number after concatenation is 224. We denote the feature maps for CNN outputs as $\{F_1, F_2, F_3\} \in \mathbb{R}^n$, which have the sizes of 16 × 16 × 224,

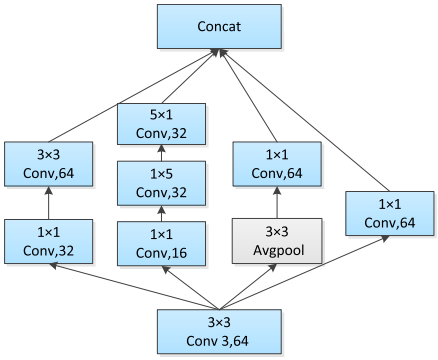

FIGURE 4. The Inception module structure.

$32 \times 32 \times 224$, $64 \times 64 \times 224$ with respect to three resolution inputs.

A. INTERACTIVE FEATURE FUSION (IFF)

F_1, F_2, and F_3 represent effective features with different resolutions, which have similarities and differences. Different from the former methods that fuse multi-resolution features using addition or concatenation, our fusion idea is to sum various resolution features interactively for each output module, realized by the down-sampling module and the up-sampling module. The strategy is shown in Figure 5.

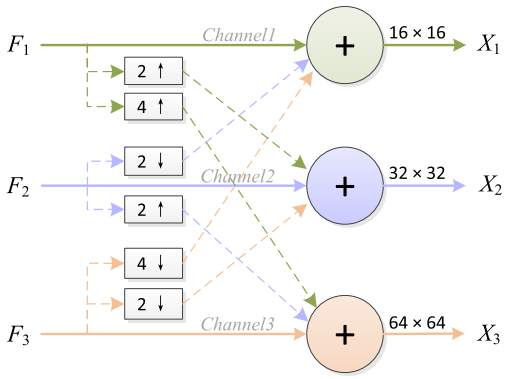

FIGURE 5. Interactive feature fusion structure.

As shown, we apply a 2×2 avgpooling module with a stride of 2 and a 4×4 avgpooling module with a stride of 4 to achieve $2\uparrow$ and $4\uparrow$ down-sampling. In contrast, for $2\downarrow$ and $4\downarrow$ up-sampling, we use a 2×2 deconvolution module with a stride of 2 and a 4×4 deconvolution module with a stride of 4. By addition, the fusion feature $\{X_1, X_2, X_3\}$ can be obtained as follows

$$\begin{aligned} X_1 &= F_1 + {F_2}^{2\downarrow} + {F_3}^{4\downarrow} \\ X_2 &= {F_1}^{2\uparrow} + F_2 + {F_3}^{2\downarrow} \\ X_3 &= {F_1}^{4\uparrow} + {F_2}^{2\uparrow} + F_3 \end{aligned} \tag{7}$$

B. FUSION FOR CLASSIFICATION

Following the interactive feature fusion module, the FC layers are connected to obtain the final classification feature vectors. By training the designed multi-resolution network, we can extract feature vectors $\{Y_1, Y_2, Y_3\}$ from FC layers. The feature vectors are closely related to the probability of classification. In order to reduce the recognition error caused by feature deviation, we apply the weighted average method for the classification fusion.

The weighted average method is to fuse different features proportionally by setting a weighted coefficient. It is described as follows:

$$Y = \alpha Y_1 + \beta Y_2 + \gamma Y_3 \tag{8}$$

where α, β, γ are the weighted coefficients that determine the proportion of each feature. Y is the fused feature vector. Here, we set $\alpha = \beta = \gamma = \frac{1}{3}$ for classification fusion.

C. CLASSIFIER

After feature extraction and fusion are finished, a final decision is made using a classifier. Support vector machine is a supervised learning algorithm that is widely used to solve classification problems due to its ability to deal with high dimensional data and efficiency in modeling diverse data [22]. The basic idea of the SVM is to construct an optimal hyperplane on the feature space that optimally separates each class. It is often designed for binary classification. Error-Correcting Output Codes (ECOC) is a successful framework to address the multiclass problem. It can improve the performance of classification because of its ability to correct the bias and variance errors of the base classifiers [23], [24]. In this paper, we apply ECOC-SVM to deal with multiple waveform classification.

The ECOC-SVM algorithm consists of two steps: encoding and decoding. Here, the one-versus-one strategy is designed for the coding matrix. Given a set of N classes to be learned, the ternary symbol-based coding matrix defines as $H \in \{-1, 0, 1\}^{N \times n}$, where n is the codeword length, $n = N(N-1)/2$. For the decoding stage, the minimum Hamming distance d_r among all pairs of rows can be defined as follows [24]

$$d_r = min \left\{ \sum_{j=1}^{n} (1 - sign(y_{i_1}^j \cdot y_{i_2}^j))/2 \right\} \tag{9}$$

For $i_1, i_2 \in \{1, \ldots, N\}, i_1 \neq i_2$, being $y_{i_1}^j$ the jth position of the codeword for class c_{i_1}.

V. SIMULATION RESULT AND DISCUSSION

In this section, we verify the performance of the proposed recognition model by simulation experiments. We generate radar waveform datasets for simulation and then compare the performance of the proposed algorithm with other methods at different SNRs. Besides, we discuss the effect of fusion module on recognition performance and detect the robustness of the system.

A. RADAR WAVEFORM DATASETS

In the experiment, twelve kinds of simulated LPI radar signals are generated for training and testing. In order to make the simulated signal similar to the real signal, the parameter values of the signal are randomly set within the specified range, and white Gaussian noise with different SNR are added to the signal. The parameters of each signal are shown in Table 1. The signal sampling rate $f_s = 50MHz$, and the length of each signal N is between 600 and 1200. Each signal has 2200 samples with the SNR value from -10 to 10 dB with a step of 2 dB. After these signals are generated, we apply the enhanced FSST to transform waveforms into T-F images and resize T-F images to 256×256. We divide the datasets into two parts, 70% of samples randomly selected for training, the remaining 30% samples for testing.

TABLE 1. Parameters of twelve types of LPI radar signals.

Modulation Type	Parameters	Value
LFM	Carrier frequency f_0	$[f_s/5, f_s/4]$
	Bandwidth B	$[f_s/20, f_s/8]$
Costas	Fundamental frequency $f_{\min}$	$f_s/40$
	Frequency-hopping number L	$[6, 10]$
BPSK	Carrier frequency f_0	$[f_s/5, f_s/4]$
	Barker codes length L_c	$\{7, 11, 13\}$
	Cycles per phase code C_{pp}	$[12, 20]$
Frank, P1, P2	Carrier frequency f_0	$[f_s/5, f_s/4]$
	Number of frequency steps M	$[6, 10]$
	Cycles per phase code C_{pp}	$[2, 5]$
P3, P4	Carrier frequency f_0	$[f_s/5, f_s/4]$
	Number of subcodes N_c	$\{36, 49, 64, 81, 100\}$
	Cycles per phase code C_{pp}	$[2, 5]$
T1, T2	Carrier frequency f_0	$[f_s/5, f_s/4]$
	Number of segments m	$[4, 6]$
T3, T4	Carrier frequency f_0	$[f_s/5, f_s/4]$
	Bandwidth B	$[f_s/20, f_s/8]$

B. PERFORMANCE COMPARISON

To demonstrate the recognition performance of the proposed algorithm, we compare the approach with CWD-CNN [8], CWD-MFCNN [18], CWD-ResNet-SVM [15], FSST-CNN [12], and FSST-SAE. The FSST-SAE method uses the enhanced FSST to obtain 64 × 64 T-F images and applies an SAE network [14] with two autoencoder layers for classification. For the network training, we select the Stochastic Gradient Descent with Momentum (SGDM) algorithm as the optimizer. The learning rate, epoch, and batch size are set to 0.01, 6, and 32, respectively. In order to ensure the reliability of the recognition performance, the experimental results are averaged by ten times. The CNNs are trained on a high computation PC with the central processing unit (CPU) i7-7820HK@2.9GHz, with 6GB RAM, and the Cuda enabled graphics processing unit (GPU) Nvidia GeForce GTX1080.

Figure 6 shows the overall recognition results of five methods at different SNRs. As shown in Figure 6, when the SNR > −2 dB, all five methods except CWD-ResNet-SVM have high recognition accuracy, more than 99% after six epochs of training. With the decline of SNR, the recognition curve has dropped significantly. The performance of CWD-ResNet-SVM is relatively poor, followed by CWD-CNN, and CWD-MFCNN has better performance than FSST-CNN and FSST-SAE. The recognition performance of FSST-SAE is slightly better than that of FSST-CNN, which benefits from the enhanced spectrum of FSST. Compared with CNN, SAE can achieve classification without data labels but takes longer training time. The proposed method shows the best recognition results, especially at low SNRs. When the SNR is -6 dB, our method's recognition accuracy is 97.9%, increasing over the best results of the other four methods by 1.0%.

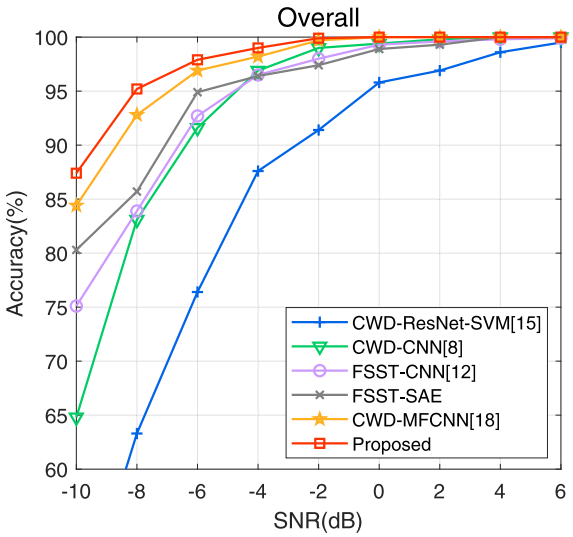

FIGURE 6. Performance comparison of six methods at different SNRs.

When the SNR is -8 dB, our method can still achieve 95.2% recognition accuracy, which exceeds CWD-MFCNN by 2.4% and FSST-CNN by 11.3%. The results demonstrate that the proposed approach has higher recognition accuracy and stronger anti-noise ability than the existing approach. The performance improvement benefits from two aspects: the enhanced FSST has better anti-noise performance, and the multi-resolution fusion structure is more conducive to extracting effective features for recognition.

Figure 7 shows the recognition accuracy of twelve radar signals at different SNRs. We find that Costas, LFM, and P2 codes with apparent differences on T-F images can be well recognized in the proposed method. P1, P4, T2, and T4 codes with similar features are relatively difficult to classify than other codes at low SNRs. It should be noticed that the recognition accuracy for P4 and T3 codes in this paper have been greatly improved than the results of other methods. The results demonstrate that the multi-resolution deep feature fusion method can extract more detailed features, which is helpful to identify similar signals.

Table 2 shows the confusion matrix of twelve kinds of waveforms at SNR = −8dB. It is clearly that the codes with similar T-F images are easily confused with each other, such as P1 and P4 codes. For P1 code, 22.3% is misidentified as P4 code, and for P4 code, 11.7% is misidentified as P1 code. In addition, 5.7% of T4 codes are identified as T2 code. The reason is that the severe noise makes the feature fuzzy and reduces the recognition accuracy. We can further consider radar signal denoising, which will greatly improve the recognition performance.

C. THE EFFECT OF FUSION MODULES ON RECOGNITION PERFORMANCE

In order to evaluate the effect of fusion modules on the recognition performance in our method, we drop the interactive fea-

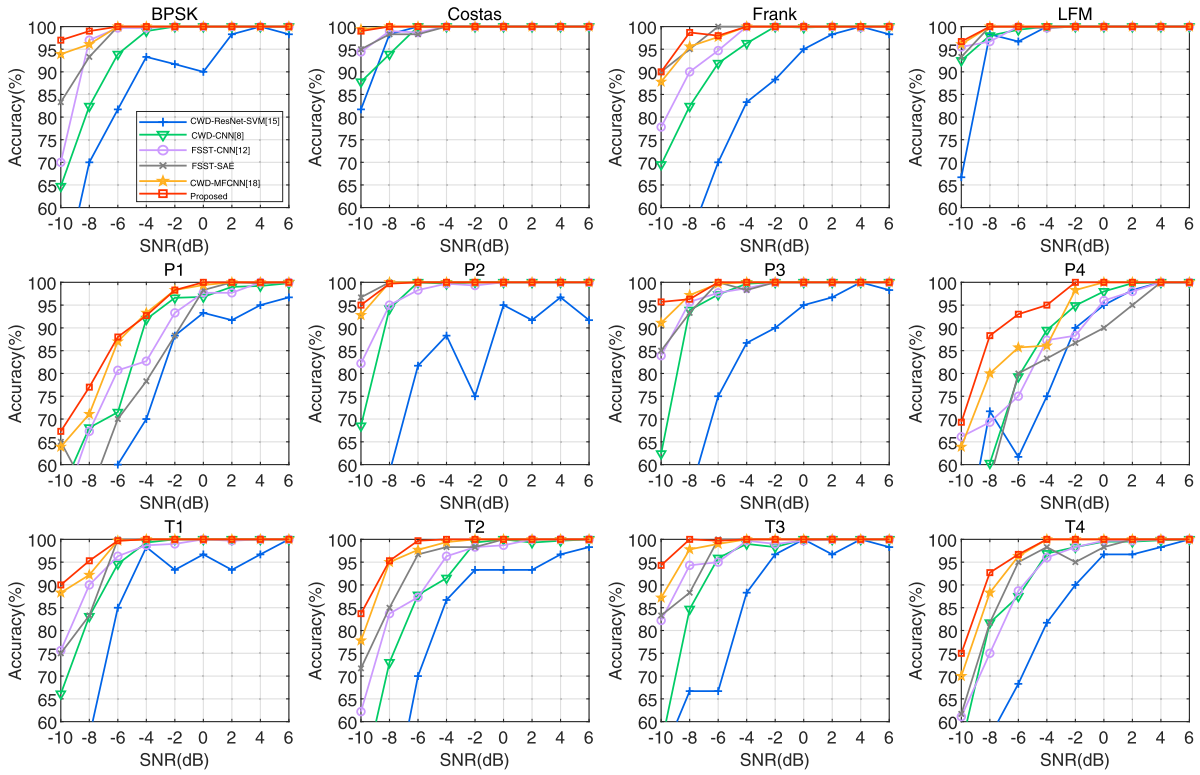

FIGURE 7. Recognition accuracy of twelve radar signals at different SNRs.

TABLE 2. Confusion matrix for waveform classification (SNR=-8dB).

	BPSK	Costas	Frank	LFM	P1	P2	P3	P4	T1	T2	T3	T4
BPSK	**99.0**	0	0	0	0	0	0	0	0.3	0	0.7	0
Costas	0	**100**	0	0	0	0	0	0	0	0	0	0
Frank	0	0	**98.7**	0	0.7	0	0.6	0	0	0	0	0
LFM	0	0	0	**100**	0	0	0	0	0	0	0	0
P1	0	0	0	0	**77.0**	0	0.7	22.3	0	0	0	0
P2	0	0	0	0	0	**99.7**	0	0	0	0.3	0	0
P3	0	0	1.0	0	0.7	0	**96.3**	2.0	0	0	0	0
P4	0	0	0	0	11.7	0	0	**88.3**	0	0	0	0
T1	3.3	0	0	0	0	0	0	0	**95.3**	0	1.0	0.4
T2	0.7	0	0	0	0	0	0	0	0	**95.3**	0.7	3.3
T3	0	0	0	0	0	0	0	0	0	0	**100**	0
T4	0	0	0	0	0	0.3	0	0	1.3	5.7	0	**92.7**

ture fusion module and the classification fusion module in the following experiments. We build three independent channels through the down-sampling module, with input resolutions of 64×64, 128×128, and 256×256. Each channel uses the same CNN structure and training parameters to train them separately. For performance comparison, we denote the three resolution recognition methods with no fusion modules as FSST-CNN 64, FSST-CNN 128, and FSST-CNN 256, here $X_i = F_i (i = 1, 2, 3)$. In addition, we denote the three resolution recognition methods with interactive feature fusion modules as FSST-CNN 64 + IFF, FSST-CNN 128 + IFF, and FSST-CNN 256 + IFF. In FSST-CNN 64 + IFF, the features $\{F_1, F_2, F_3\}$ from Channels 1, 2, and 3 are fused into X_1, as described in Equation (7). The output is obtained by classifying the FC vector Y_1. FSST-CNN 128 + IFF and FSST-CNN 256 + IFF have similar fusion strategy, but the module outputs are X_2 and X_3 respectively, the results are obtained by the FC vector Y_2 and the FC vector Y_3.

Table 3 shows the effect of fusion modules on recognition performance. As shown, the three methods with no fusion modules have different recognition results at low SNRs. At the same SNR, the recognition accuracy of FSST-CNN 128 exceeds that of FSST-CNN 64 and FSST-CNN 256. If the resolution is larger than 256×256, it will not improve the recognition accuracy but increase training parameters and training time. On the contrary, if the resolution is less than 64×64, the extracted detailed features are insufficient, resulting in limited recognition. So it is reasonable to select 256×256, 128×128, and 64×64 as inputs for fusion. By adding an interactive feature fusion module, the recognition performance of the three methods has been improved. For example, FSST-CNN 64 + IFF improves the recognition accuracy by 1.9% than FSST-CNN 164 at SNR of -10dB. With the addition of the feature and classification fusion modules, the proposed method improves by approximately 2.9% accuracy than FSST-CNN 128 + IFF when the SNR is -10 dB. These results demonstrate that the multi-resolution fusion strategy can improve performance more significantly at low SNRs.

TABLE 3. Effect of fusion strategy on recognition performance.

Method	Accuracy(%)		
	-10dB	-8dB	-6dB
FSST-CNN 64	82.8	92.4	96.6
FSST-CNN 128	84.7	92.9	96.7
FSST-CNN 256	80.6	89.5	96.2
FSST-CNN 64 + IFF	84.7	92.8	97.1
FSST-CNN 128 + IFF	84.9	93.1	97.1
FSST-CNN 256 + IFF	83.0	92.6	96.6
Proposed	**87.8**	**95.2**	**97.9**

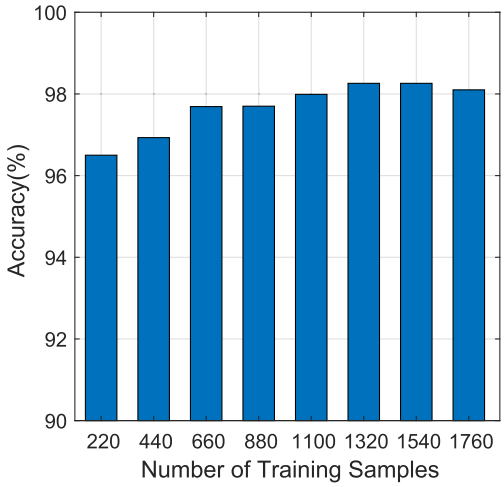

FIGURE 8. Recognition accuracy of different numbers of training samples.

D. EXPERIMENT WITH ROBUSTNESS

While the multi-resolution fusion method exhibits its effectiveness on a specific dataset composed of twelve types of radar signals, we wonder whether it would still work well on a new dataset and different training samples. To verify the robustness of the recognition system, we generate a new dataset with the same number of samples according to the signal parameters in Table 1, and add random Gaussian white noise with the SNR of -10dB $\sim$10dB to the waveforms. For each radar waveform, we select 220 samples for testing and set the number of training samples from 220 to 1760, with a step of 220.

As shown in Figure 8, when the number of training samples is more than 660, the accuracy curve tends to steady, which indicates that the system has good stability for classification. Furthermore, the recognition accuracy only slightly decreases when the sample number is less than 660. It also shows that the method can achieve excellent recognition performance in a small training set, which is more important in the actual environment for radar waveform recognition. This experiment shows that the proposed method has better robustness, maintaining outstanding performance in a new dataset and different training samples.

VI. CONCLUSION

In this paper, we have presented a strategy for LPI radar waveform recognition based on multi-resolution deep feature fusion. The approach can identify twelve kinds of LPI radar waveforms in a low SNR environment, including BPSK, Costas, Frank, LFM, P1, P2, P3, P4, T1, T2, T3, and T4. Simulation results show that the proposed technique has classified all the radar waveforms with 95.2% recognition accuracy at SNR of -8dB compared to the other methods. Furthermore, the algorithm also has good robustness. Compared with CWD, the enhanced FSST has a faster speed for signal time-frequency analysis and is more suitable in real waveform identification systems. The multi-resolution feature fusion strategy provides a feasible LPI radar waveform recognition solution at low SNR levels. It can be used in electronic reconnaissance to identify real-world radar signals.

REFERENCES

[1] A. Lokam, N. Murthy, and N. Sarma, "A novel method for recognition of modulation code of LPI radar signals," *Int. J. Recent Trends Eng.*, vol. 1, no. 3, p. 176, 2009.
[2] P. E. Pace, *Detecting and Classifying Low Probability of Intercept Radar*, 2nd ed. Norwood, MA, USA: Artech House, 2008.
[3] C. Xu, J. Zhang, Q. Zhou, and S. Chen, "Recognition of radar signals based on AF grids and geometric shape constraint," *Signal Process.*, vol. 157, pp. 30–44, Apr. 2019.
[4] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in *Proc. IEEE Int. Symp. Circuits Syst.*, Jun. 2010, pp. 253–256.
[5] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
[6] C. Wang, J. Wang, and X. Zhang, "Automatic radar waveform recognition based on time-frequency analysis and convolutional neural network," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 2437–2441.
[7] M. Zhang, M. Diao, and L. Guo, "Convolutional neural networks for automatic cognitive radio waveform recognition," *IEEE Access*, vol. 5, pp. 11074–11082, 2017.
[8] 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.
[9] J. Wan, X. Yu, and Q. Guo, "LPI radar waveform recognition based on CNN and TPOT," *Symmetry*, vol. 11, no. 5, p. 725, May 2019.
[10] 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.
[11] Z. Qu, C. Hou, C. Hou, and W. Wang, "Radar signal intra-pulse modulation recognition based on convolutional neural network and deep Q-learning network," *IEEE Access*, vol. 8, pp. 49125–49136, 2020.
[12] G. Kong and V. Koivunen, "Radar waveform recognition using Fourier-based syncrosqueezing transform and CNN," in *Proc. CAMSAP*, Dec. 2019, pp. 664–668.
[13] M. Zhang, M. Diao, L. Gao, and L. Liu, "Neural networks for radar waveform recognition," *Symmetry*, vol. 9, no. 75, pp. 1–20, 2017.
[14] M. Zhang, H. Wang, K. Zhou, and P. Cao, "Low probability of intercept radar signal recognition by staked autoencoder and SVM," in *Proc. 10th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2018, pp. 1–6.
[15] Q. Guo, X. Yu, and G. Ruan, "Lpi radar waveform recognition based on deep convolutional neural network transfer learning," *Symmetry*, vol. 11, no. 540, pp. 1–14, 2019.
[16] L. Gao, X. Zhang, J. Gao, and S. You, "Fusion image based radar signal feature extraction and modulation recognition," *IEEE Access*, vol. 7, pp. 13135–13148, 2019.
[17] Y. Xiao, W. Liu, and L. Gao, "Radar signal recognition based on transfer learning and feature fusion," *Mobile Netw. Appl.*, vol. 25, no. 4, pp. 1563–1571, Aug. 2020.

[18] X. Ni, H. Wang, Y. Zhu, and F. Meng, "Multi-resolution fusion convolutional neural networks for intrapulse modulation LPI radar waveforms recognition," *IEICE Trans. Commun.*, vol. E103.B, no. 12, pp. 1470–1476, 2020.
[19] T. Oberlin, S. Meignen, and V. Perrier, "The Fourier-based synchrosqueezing transform," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2014, pp. 315–319.
[20] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2117–2125.
[21] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.
[22] G. Vanhoy, T. Schucker, and T. Bose, "Classification of LPI radar signals using spectral correlation and support vector machines," *Anal. Integr. Circuits Signal Process.*, vol. 91, no. 2, pp. 305–313, May 2017.
[23] X. Gu, F. Deng, X. Gao, and R. Zhou, "An improved sensor fault diagnosis scheme based on TA-LSSVM and ECOC-SVM," *J. Syst. Sci. Complex.*, vol. 31, no. 2, pp. 372–384, Apr. 2018.
[24] S. Escalera, O. Pujol, and P. Radeva, "Separability of ternary codes for sparse designs of error-correcting output codes," *Pattern Recognit. Lett.*, vol. 30, no. 3, pp. 285–297, Feb. 2009.

XUE NI received the B.S. degree in communication engineering and the M.S. degree in communication information system from Hohai University, Changzhou, China, in 2006 and 2009, respectively. She is currently pursuing the Ph.D. degree with the College of Communications Engineering, Army Engineering University of PLA, Nanjing, China. Her research interest includes intelligent recognition of radar signals.

HUALI WANG received the Ph.D. degree in electronic engineering from the Nanjing University of Science and Technology, China, in 1997. He is currently a Professor with the College of Communications Engineering, Army Engineering University of PLA, Nanjing, China. His research interests include statistical and array signal processing.

FAN MENG received the M.S. degree in communication and information system from Hohai University at Changzhou, Changzhou, China, in 2009. He is currently a Senior Engineer with the Nanjing Marine Radar Institute, China. His research interest includes radar signal processing.

JING HU received the B.S. and M.S. degrees from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2003 and 2007, respectively. She is currently pursuing the Ph.D. degree with the College of Communications Engineering, Army Engineering University of PLA, Nanjing. Her current research interests include wireless communications, satellite communications, and communication anti-jamming.

CHANGKAI TONG received the B.S. degree from Naval Aeronautical and Astronautical University, Yantai, China, in 2009, and the M.S. degree from the Naval Research Academy, Beijing, China, in 2015, respectively. He is currently pursuing the Ph.D. degree with the College of Communications Engineering, Army Engineering University of PLA, Nanjing, China. His current research interest includes deep learning brain–computer interface.