IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 29, 2019, accepted January 21, 2020, date of publication January 27, 2020, date of current version January 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2969450

Recognizing the HRRP by Combining CNN and BiRNN With Attention Mechanism

JINWEI WAN, BO CHEN, (Member, IEEE), YINGQI LIU, YIJUN YUAN, HONGWEI LIU, (Member, IEEE), AND LIN JIN

National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China

Corresponding author: Bo Chen (bchen@mail.xidian.edu.cn)

This work was supported in part by the Young Thousand Talent by Chinese Central Government, in part by the 111 Project under Grant B18039, in part by the National Natural Science Foundation of China (NSFC) under Grant 61771361, in part by the NSFC for Distinguished Young Scholars under Grant 61525105, and in part by the Shaanxi Innovation Team Project.

ABSTRACT In this paper, we integrate the advantages of convolutional neural network (CNN) and bidirectional recurrent neural network (BiRNN) with attention mechanism, and propose a CNN-BiRNN based method to recognize the individual high resolution range profile (HRRP). In the proposed method, the CNN is utilized to explore the spatial correlation of raw HRRP data and extract expressive features followed by a BiRNN taking the full consideration of temporal dependence between range cells. Furthermore, in order to enhance the robustness to misalignment, an attentional mechanism is employed after BiRNN to allow the CNN-BiRNN model to focus on the discriminative target area. The combination of CNN and BiRNN with attention mechanism makes the extracted features are not only efficient, but also strongly resistant to the time-shift sensitivity. Experimental results on measured HRRP data demonstrate the effectiveness and the robustness to misalignment of the proposed method.

INDEX TERMS Radar automatic target recognition (RATR), high-resolution range profile (HRRP), convolutional neural networks (CNNs), recurrent neural networks (RNNs), attention mechanism.

I. INTRODUCTION

Radar High-resolution range profile (HRRP), is composed of the amplitude of the coherent summations of the complex returns from target scatterers in each range cell, which represents the projection of the complex returned echoes from the target scattering centers onto the radar line-of-sight (LOS). Since HRRP contains abundant target structure signatures, such as target size, scatterer distribution, etc., HRRP-based radar automatic target recognition (RATR) has received intensive attention from the RATR community [1]–[24].

Emphasis of HRRP-based RATR mainly laid on how to extract discriminative features from data robustly and effectively, and these efforts can be roughly divided into the following two categories. The first category is physical mechanism based methods, and they usually accomplish the recognition tasks in the learned feature subspace with some specific physical meanings. In [4], [5], RELAX-based algorithms are employed to extract the location information of predominant scatterers from HRRP data as features for recognition tasks. The researchers investigate the recognition methods based on kinds of spectra features in [6]–[8]. Generally, these methods perform well in practice but heavily rely on personal experiences and prior knowledge. The second category, data-driven methods, have attracted increasing attention in past years due to their ability to learn discriminative features from the dataset automatically. In [9], the principal components analysis (PCA)-based feature subspace is constructed to minimize reconstruction error for RATR. Given the statistical correlation of HRRP, the researchers in [10]–[14] employ statistical models, such as factor analysis (FA) to project and recognize HRRP in a learned low dimensional latent feature space. Considering the sparsity within the HRRP, Feng *et al.* [15] and Zhou [16] apply sparse constraint on the feature vectors and solve the problems via l_0-minimization. However, all those methods build linear and shallow architectures that limit their capability to represent the complicated HRRP data. Advanced data-driven methods, such as deep neural networks, have been successfully employed in HRRP-based RATR in recent years due to their powerful expressive capability. Feng *et al.* [18] employ the average profile as the correction terms and stack a series of corrective autoencoders (CAE) to extract features from HRRP. In view of the temporal dependence in HRRP, Xu *et al.* [19] divide

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

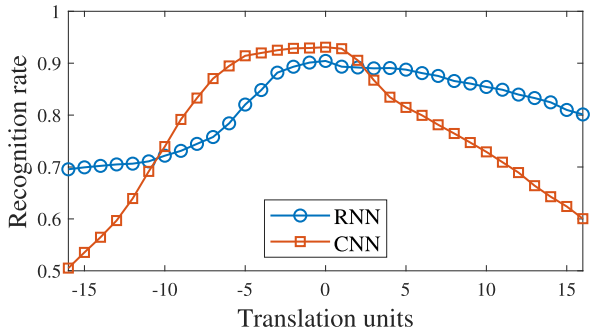


FIGURE 1. Variation of the recognition rate with the different extents of misalignment, via CNN and RNN.

the HRRP sample into multiple overlapping sequential feature and use recurrent neural network (RNN [25]) to learn latent representation for the recognition task. To explore the correlation between range cells and extract the structured discriminative features in HRRP, Wan *et al.* [20] employ convolutional neural networks (CNNs [26]) to address the HRRP-based RATR task.

Although the deep learning based models have shown better performance than the shallow or linear ones, there are still some defects. For instance, the CNN model in [20] mainly focuses on the local structure information without modeling the global structure or temporal correlation of HRRP. As a result, their CNN model is sensitive to the misalignment between HRRPs. As for RNN, the authors in [19] model the raw HRRP data with fully-connected based RNN directly, which makes it difficult to extract efficient features and causes a worse recognition results compared with CNN [20]. In Fig. 1, we exhibit the recognition results of CNN and RNN models on measured radar HRRP dataset. Note that, the structure of CNN follows [20], which can be regarded as an extended LeNet-5 [35], and it is composed of three convolutional with kernel size 1×9, pooling with size 1×2, batch normalization layers, one fully connected layer with size 300 and one classification layer. The structure of RNN follows [19], which is an RNN encoder containing one hidden layer with size 200, and the last hidden state is utilized for the recognition task. As same as [19], in RNN, the HRRP sample is divided into multiple overlapping sequential feature with length 32 and overlap 16. Apparently, as shown in Fig. 1, without misalignment (*Translation* $= 0$), the CNN, 92.5%, performs 2% better than the RNN, 90.1%. The recognition rates of CNN are almost constant with the misalignment falls in $[-5, 2]$, as the higher-layer representation of CNN for small translations of the input within the pooling size is invariant. However, the CNN is not always robust to the time misalignment, and when the misalignment exceeds the pooling size, e.g. $[-16, 16]$, the performance of CNN drops significantly with the increase of the misalignment units as shown in Fig. 1. By the contrast, that of the RNN model dose not change a lot as it utilizes the temporal dependence to dynamically grasp and deliver global discriminative information.

To address the above issues and combine the advantages of CNN and RNN models, in this paper, we integrate CNN and bidirectional RNN (BiRNN) with attention mechanism to propose a CNN-BiRNN based method for the single HRRP recognition. A CNN model is employed to extract expressive features from the raw HRRP data followed by a BiRNN with the full consideration of temporal dependence between range cells in the single HRRP. Furthermore, in order to enhance the robustness to misalignment, an attention mechanism is introduced after BiRNN to allow the model to focus on the discriminative target area automatically. All the components are trained jointly in an end-to-end fashion. Through the combination of CNN, BiRNN and attention mechanism, our method sufficiently exploits the correlation between range cells and the temporal dependence in an individual HRRP, which improves the recognition performance and introduces the robustness to time-shift sensitivity. We evaluate the proposed method on the measured HRRP data. The comprehensive comparisons with several existing methods demonstrate that the proposed method achieves better performance under aligned data and remarkably improves the performance under misalignment.

The remainder of this article is organized as follows. In Section II, the signal model of HRRP, as well as the fundamental of the CNN and RNN model, is introduced. We detail the CNN-BiRNN and compare it with other related works in Section III and IV, respectively. In Section V, the experimental results on measured HRRP data are shown. Section VI draws the conclusion.

II. PRELIMINARIES

In this section, we briefly review the HRRP, following which the architecture of CNN and RNN are provided.

A. HRRP

Generally, the range resolution of high-resolution radar (HRR) is much smaller than the targets size. Thus, the HRR can effectively divide the complex targets such as an aircraft into many range "cells". According to [9], [13], the t-th complex HRRP can be written as

$$\mathbf{x}_C(t) = e^{j\theta(t)}\left[\tilde{x}_1(t), \tilde{x}_2(t), \ldots, \tilde{x}_L(t)\right] \tag{1}$$

where $\theta(t)$ stands for the initial phase of the t-th returned echo, and $\tilde{x}_l(t) = \sum_{i=1}^{V_l} \sigma_{li} e^{j\phi_{li}(t)}$ denotes the echo of l-th range cell, which is composed of V_l scatterers of strength σ_{li} and phase $\phi_{li}(t)$. Then, the t-th time domain HRRP $\mathbf{x}(t)$ is obtained by taking absolute value of $\mathbf{x}_C(t)$, which can be represented as

$$\begin{aligned}\mathbf{x}(t) &= \left[\,|\tilde{x}_1(t)|, |\tilde{x}_2(t)|, \ldots, |\tilde{x}_L(t)|\,\right] \\ &= \left[\left|\sum_{i=1}^{V_1}\sigma_{1i}e^{j\phi_{1i}(t)}\right|, \left|\sum_{i=1}^{V_2}\sigma_{2i}e^{j\phi_{2i}(t)}\right|,\right. \\ &\quad \left. \ldots, \left|\sum_{i=1}^{V_l}\sigma_{li}e^{j\phi_{li}(t)}\right|, \ldots, \left|\sum_{i=1}^{V_L}\sigma_{Li}e^{j\phi_{Li}(t)}\right|\right]\end{aligned} \tag{2}$$

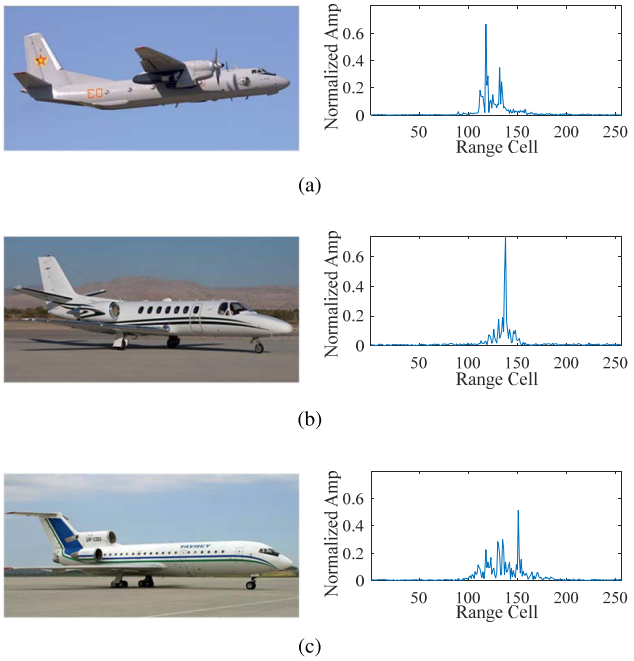


FIGURE 2. Examples of three aircraft and their corresponding time domain HRRP samples. (a) An-26; (b) Cessna Citation S/II; and (c) Yark-42.

where $|\cdot|$ means taking absolute value. The right side of Fig. 2 presents the time domain HRRP samples from the corresponding three airplane targets, i.e., An-26, Cessna Citation S/II, and Yark-42. For notational convenience, we will replace the notation $\mathbf{x}(t)$ with $\mathbf{x}$ in the following sections.

B. CNN AND RNN

1) CNN

A typical CNN architecture [26]–[28] generally comprises of alternate layers of convolution and pooling followed by one or more fully connected layers at the end. Detailedly, we suppose $\mathbf{X} \in \mathbb{R}^{H\times W\times C}$ is a multichannel input, where each dimension of $\mathbf{X}$ represents the height, width, and the number of channels, respectively. We also assume that the CNN consists of L convolutional layers, and layer $l \in \{1, \ldots, L\}$ has K_l filters. For the k-th filter at layer 1, a convolution operation applies filter $\mathbf{W}^{(k,1)} \in \mathbb{R}^{h^{(1)}\times w^{(1)}\times C}$ to $\mathbf{X}$, which yields feature map

$$\mathbf{y}^{(k,1)} = f(\mathbf{X} * \mathbf{W}^{(k,1)} + \mathbf{b}^{(k,1)}) \tag{3}$$

where $*$ denotes the convolutional operator, $\mathbf{b}^{(k,1)}$ is the bias for the k-th feature map, and $f(\cdot)$ is the RELU nonlinear activation function. To increase the content covered by a convolutional kernel and the sparsity of the hidden units, a convolutional layer is usually followed by a pooling layer. The resolution of the feature map is reduced by pooling over the local neighborhood on the feature map. After the convolution and the pooling, k-th feature map at layer 1, $\mathbf{h}^{(k,1)} = f_p(\mathbf{y}^{(k,1)}) \in \mathbb{R}^{H^{(h1)}\times W^{(h1)}\times 1}$ is yielded, where $H^{(h1)}, W^{(h1)}$ are the height and width of $\mathbf{h}^{(k,1)}$, and $f_p(\cdot)$ defines the type of pooling operation. After applying K_1 filters, we have the feature maps of layer 1, $\mathbf{H}^{(1)} \in \mathbb{R}^{H^{(h1)}\times W^{(h1)}\times K_1}$.

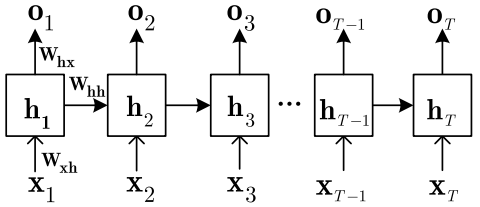


FIGURE 3. Illustration of the RNN model.

With this convolutional-pooling operation repeated in sequence for L layers, the feature map of last convolutional layer, $\mathbf{H}^{(L)}$ is obtained, which is usually fed into some fully connected layers to produce the final classification feature $\mathbf{z}$. In the end, the feature $\mathbf{z}$ goes into a softmax classifier for the recognition task.

2) RNN

A basic RNN model, as shown in Fig. 3, includes a input layer, a hidden layer and a output layer. Given a sequential sample of length T, $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_{T-1}, \mathbf{x}_T]$, the RNN reads $\mathbf{x}$ in order from $\mathbf{x}_1$ to $\mathbf{x}_T$, and calculates the hidden state $\mathbf{h}_t$ and the output $\mathbf{o}_t$ by iterating the following equations from $t = 1$ to T

$$\mathbf{h}_t = f(W_{xh} \cdot \mathbf{x}_t + W_{hh} \cdot \mathbf{h}_{t-1}), \quad t = 1 \ldots T \tag{4}$$
$$\mathbf{o}_t = f(W_{hx} \cdot \mathbf{h}_t), \quad t = 1 \ldots T \tag{5}$$

where $W_{xh} \in \mathbb{R}^{m\times d}$, $W_{hh} \in \mathbb{R}^{m\times m}$ and $W_{hx} \in \mathbb{R}^{m\times n}$ respectively represent the weights of input-hidden, hidden-hidden and hidden-output, and d, m, n are the size of input, hidden and output. $f(\cdot)$ refers to the activation function which operates on each elements. Usually, the last output $\mathbf{o}_T$ is used as the feature of the softmax classifier to predict the label, as it contains information of the whole sequential data $\mathbf{x}$. Due to the excellent modeling of sequential data, RNN [25] achieves state-of-the-art performance on different tasks, such as speech recognition [29], machine translation [30], [31], HRRP recognition [19] and so on.

In this paper, we would like to make full use of the temporal dependence in HRRP. Hence, we propose to utilize BiRNN in our model, which is an extended RNN model that includes forward and backward RNNs.

III. METHOD

The proposed method, as shown in Fig. 4, mainly consists of three parts: the input HRRP data $\mathbf{X}$ at the bottom, a CNN-BiRNN hybrid model in the middle, and a softmax classifier on the top. Without loss of generality, we take the time domain HRRP (2) as input. The proposed method includes the following three steps: the construction of the CNN-BiRNN model, the training and the test of the model.

A. THE CNN-BIRNN HYBRID MODEL

As shown in the middle of Fig. 4, the CNN-BiRNN hybrid model contains two components: a CNN with L convolutional layers at the bottom half and a BiRNN with attentional

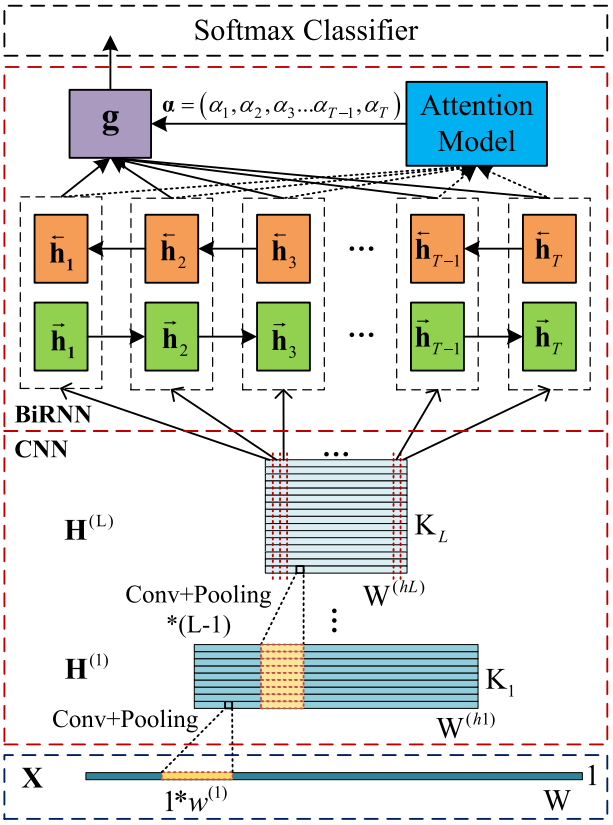


FIGURE 4. An overview of the proposed method based on RNN-BiCNN. A deep CNN is employed to extract expressive feature from **X**, the resultant feature map, $\mathbf{H}^{(L)}$, is considered as a sequence and passed into an attentional BiRNN for sequential modeling, and a softmax classifier receives the features from the attentional BiRNN and outputs the classification results.

mechanism [31] on the top half. The components of the model are trained jointly in an end-to-end fashion.

A single HRRP sample $\mathbf{X} \in \mathbb{R}^{1\times W\times 1}$ is a real-valued vector (W is the length of range dimension), thus a 1-dimensional convolution CNN model is employed, where the convolution operation only takes place at the range dimension. For the first layer, as shown at the bottom of Fig. 4, convolution operations with stride length 1 apply K_1 filters, $\mathbf{W}^{(1)} \in \mathbb{R}^{1\times w^{(1)}\times 1}$, to $\mathbf{X}$, resulting in feature map of layer 1, $\mathbf{H}^{(1)} \in \mathbb{R}^{1\times W^{(h1)}\times K_1}$. For the following $L-1$ convolution layers, convolution-pooling operations repeatedly apply K_l filters, $\mathbf{W}^{(l)} \in \mathbb{R}^{1\times w^{(l)}\times K_{l-1}}$, to $\mathbf{H}^{(l-1)}$, and feature map $\mathbf{H}^{(L)} \in \mathbb{R}^{1\times W^{(hL)}\times K_L}$ is gained after L convolutional layers. Each convolutional layer is followed by a pooling layer, thus the time (range) dimension is shortened and the temporal dependence grows as the convolutional layer increases. It should also be noted that the discriminative features extracted by the CNN also preserve the sequence order that exist in HRRP, which can be naturally applied to BiRNN for sequential modeling.

Afterwards, to sufficiently utilize the temporal dependence in HRRP, we pass the $\mathbf{H}^{(L)}$ into a BiRNN for a further processing. In particular, after dropping the singleton dimension, $\mathbf{H}^{(L)} \in \mathbb{R}^{W^{(hL)}\times K_L}$ can be considered as a sequence of length $W^{(hL)}$ with K_L feature vector at each time step. In the following, we replace the time steps $W^{(hL)}$ with T for notational convenience. The forward RNN reads $\mathbf{H}^{(L)}$ in its original order and generates a hidden state $\mathbf{fh}_t$ at each time step, and the backward RNN reads $\mathbf{H}^{(L)}$ in its reverse order and produces $\mathbf{bh}_t$, i.e.,

$$\mathbf{fh}_t = f(W_{fxh}\cdot \mathbf{H}_t^{(L)} + W_{fhh}\cdot \mathbf{fh}_{t-1}), \quad t = 1\ldots T \tag{6}$$

$$\mathbf{bh}_t = f(W_{bxh}\cdot \mathbf{H}_t^{(L)} + W_{bhh}\cdot \mathbf{bh}_{t-1}), \quad t = T\ldots 1 \tag{7}$$

where $W_{fxh} \in \mathbb{R}^{m\times K_L}$ and $W_{bxh} \in \mathbb{R}^{m\times K_L}$ are the input-hidden weights, $W_{fhh} \in \mathbb{R}^{m\times m}$ and $W_{bhh} \in \mathbb{R}^{m\times m}$ denote the weights that connect the hidden state, m is the dimensionality of hidden state and $f(\cdot)$ refers to the sigmoid activation function. Then, the concatenation of the forward state $\mathbf{fh}_t$ and backward state $\mathbf{bh}_t$ creates $\mathbf{h}_t$, i.e., $\mathbf{h}_t = [\mathbf{fh}_t, \mathbf{bh}_t]$. As a result, each hidden state $\mathbf{h}_t$ contains information of the whole target, with strong focus on the parts surrounding the area at step t.

Generally, in BiRNN, the information at each time step may gradually lose along the forward and backward propagation. In order to avoid the information loss and automatically focus on the discriminative time steps, the attention mechanism is introduced, which, as a byproduct, is able to relax the misalignment issue. In our method, we adopt a multilayer perceptron to calculate the attention weight based on the hidden states and denote $\mathbf{g}$ as the invariant feature vector, which is the weighted sum of $\mathbf{h}_t$, $t = 1\ldots T$, i.e.,

$$\mathbf{g} = \sum_{t=1}^{T} \alpha_t \cdot \mathbf{h}_t. \tag{8}$$

The weight α_t is computed by

$$\alpha_t = \frac{\exp(e^t)}{\sum_l \exp(e^l)} \tag{9}$$

$$e^t = U_a^\top \tanh(W_a \cdot \mathbf{h}_t) \tag{10}$$

where $U_a \in \mathbb{R}^{1\times n}$, $W_a \in \mathbb{R}^{n\times 2m}$ are the parameters of the attentional model, and weight α_t is the coefficient that scores the matching degree between the hidden state $\mathbf{h}_t$ and the recognition task. The invariant feature vector $\mathbf{g}$ integrates the information at all time steps according to the discrimination of the hidden state at each time step.

Given the invariant feature vector $\mathbf{g}$, we adopt the softmax function to predict the label vector of the input sample $\mathbf{X}$, i.e.,

$$p(c|\mathbf{g};\theta) = \frac{\exp(\theta^{(c)\top}\mathbf{g})}{\sum_{j=1}^{C}\exp(\theta^{(j)\top}\mathbf{g})} \tag{11}$$

where C represents the number of categories, $p(c|\mathbf{g};\theta)$ denotes the probability of $\mathbf{g}$ belonging to the c-th category, and θ is the parameter of the softmax classifier.

B. TRAINING AND TEST PROCEDURE

The process of HRRP target recognition with the CNN-BiRNN hybrid model contains two stages, the training stage and the test stage. In the training stage, we use the training

Algorithm 1 Training Procedure of the CNN-BiRNN Hybrid Model

Obtain the training data $\mathbf{D} = \{\mathbf{X}, t\}$. Set mini-batch size m and the number of convolutional layer L.
Initialize the CNN parameter Θ_{CNN}, the RNN parameter Θ_{RNN}, the attention model parameter Θ_{ATT} and the softmax classifier parameter Θ_{CLS}, $\Theta = \{\Theta_{CNN}; \Theta_{RNN}; \Theta_{ATT}; \Theta_{CLS}\}$.
for *iter* $= 1, 2, 3 \ldots$ **do**
    Randomly select a mini-batch of m samples from $\mathbf{D}$ to form a subset $\{\mathbf{X}, t\}^{(1,m)}$;
    Obtain $\mathbf{H}^{(L)}$ by applying L layers convolutional-pooling operations to $\{\mathbf{X}\}^{(1,m)}$, $\mathbf{H}^{(L)} = f_{CNN}(\{\mathbf{X}\}^{(1,m)}; \Theta_{CNN})$;
    Calculate the forward and backward states, $\mathbf{fh}_t$ and $\mathbf{bh}_t$, according to (6) and (7), $\mathbf{fh}_t, \mathbf{bh}_t = f_{BiRNN}(\mathbf{H}^{(L)}; \Theta_{RNN})$;
    Concatenate $\mathbf{fh}_t$ and $\mathbf{bh}_t$ to obtain $\mathbf{h}_t$, $\mathbf{h}_t = [\mathbf{fh}_t, \mathbf{bh}_t]$;
    Calculate the invariant feature vector $\mathbf{g}$ according to (8), $\mathbf{g} = f_{ATT}(\mathbf{h}_t; \Theta_{ATT})$;
    Predict the label by (11), $y = \arg\max p(c|\mathbf{g}; \Theta_{CLS})$;
    Calculate $\nabla_\Theta \mathcal{J}(\Theta)$ according to (13), and update Θ using (14);
end for

data to learn the parameters Θ by calculating the gradient of the cost function $\mathcal{J}$,

$$\Theta = \arg\min \mathcal{J}(\Theta) \tag{12}$$

where Θ includes all parameters to be trained, and $\mathcal{J}(\Theta)$ is the cross-entropy loss function. Given the ground-truth label t of $\mathbf{X}$, $\mathcal{J}(\Theta)$ is defined as

$$\mathcal{J}(\Theta) = -\frac{1}{m}\sum_{n=1}^{m}\sum_{c=1}^{C} t^{(n)} \ln p(c|\mathbf{X}^{(n)}; \Theta) \tag{13}$$

where m is the sample number of each mini-batch. The stochastic gradient descent (SGD) and the back propagation method are utilized to iteratively optimize the problem in (10) by

$$\Theta(i+1) = \Theta(i) - \eta(i) \cdot \nabla_\Theta \mathcal{J} \tag{14}$$

where $i = 1, 2, \ldots$ denotes the iteration number, $\eta(i)$ is the learning rate at ith iteration, and ∇_Θ is the gradient operation operator with respect to Θ. The complete training procedure is described in Algorithm 1.

In the test stage, the test sample $\mathbf{X}$ goes through the model and the label is obtained by

$$y = \arg\max p(c|\mathbf{X}^{(n)}; \Theta). \tag{15}$$

IV. RELATED WORK

As introduced in section I, among the Radar HRRP target recognition methods based on deep neural networks (DNNs), the plain CNN [20] and the target-aware recurrent attentional network (TARAN) [19] are the closest to our work. In [20], a 1-dimensional CNN model is employed to extract structured discriminative features from HRRP and shows competitive results on the recognition task. Different from their work, we also consider the temporal dependence in HRRP and replace the fully-connected layers with an attentional BiRNN to model this dependence, which makes our method robust to the time-shift sensitivity, and more flexible and practical than the plain CNN.

An attention RNN model, called TARAN, is proposed in [19] for HRRP-based target recognition, which employs the attentional RNN to encode the raw HRRP sequence into hidden states and weights the hidden state at each time step to focus on the discriminative regions. Our work is different from the TARAN in two ways. First, to model the temporal correlation between range cells, the TARAN builds the attentional RNN model directly from the raw HRRP sample, whereas the proposed model treats the feature extracted from CNN as sequential data to jointly learn the parameters of CNN and BiRNN. As shown in Fig. 10, features extracted from CNN are more discriminative for the recognition task than the raw HRRP data. Meanwhile, the attention mechanism in TARAN model only considers the forward state, which is insufficient to evaluate the contribution of the hidden states. In contrast, our BiRNN employs both forward and backward states to calculate the attention weights, which offers a comprehensive consideration of temporal correlation in an individual HRRP. In addition, as a preprocessing procedure, TARAN has to manually segment a single HRRP sample into multiple time steps before their RNN model where the segment parameters need to be carefully tuned. On the contrary, CNN is integrated to extract features and make our method as an end-to-end style without any manual segmentation.

V. EXPERIMENTS

To show the effectiveness and the robustness of the proposed method, we compare it to other counterparts on the measured data in this section. The measured HRRP data and the detailed experiment settings used in our experiments will be introduced firstly. Then, several detailed analysis and discussions about the recognition performance of our method are presented and studied, respectively. Meanwhile, the influence of some model parameters is analyzed and discussed during the experiments.

A. MEASURED DATA

The results presented in this paper are based on three measured airplane data, An-26, Cessna Citation S/II and Yark-42, which are shown in the left side of Fig. 2. An-26 is a medium-sized propeller aircraft, Cessna Citation S/II is a small-sized jet aircraft and Yark-42 is a large and medium-sized jet aircraft. The radar works on C-band with a bandwidth

TABLE 1. Parameters of radar and planes.

Radar parameters	Center freq.	5520 MHz	
	Bandwidth	400 MHz	
Planes	**Length(m)**	**Width(m)**	**Height(m)**
Yark-42	36.38	34.88	9.83
An-26	23.80	26.20	9.83
Cessna Citation S/II	14.40	15.90	4.57

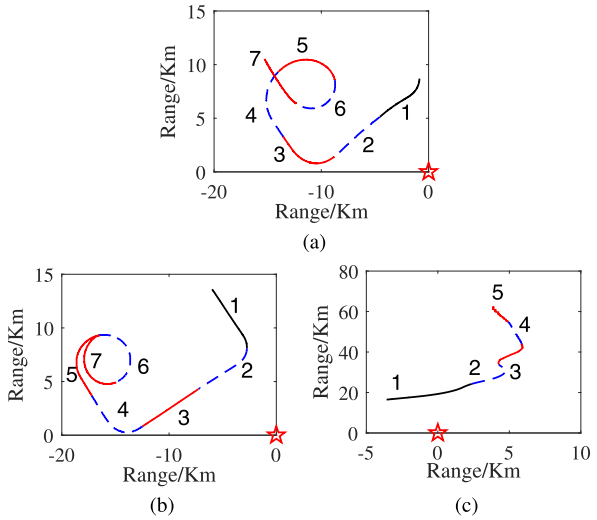


FIGURE 5. Projections of target trajectories onto the ground plane. (a) An-26; (b) Cessna Citation S/II; and (c) Yark-42. The radar locates on (0, 0).

of 400 MHz and the range resolution is about 0.375 m. The parameters of the radar and airplane targets are shown in Table 1, and the projections of target trajectories onto the ground plane are shown in Fig. 5, from which the target-aspect angle of an airplane can be estimated according to its relative position to radar.

From Fig. 5 we can find that the measured data are segmented into several parts. We choose the training and test datasets based on the following two principles: (*i*) the training dataset should cover almost all of the target-aspect angles; (*ii*) the elevation angles of the training and the test dataset are different. Thus, the 5-th and the 6-th segments of An-26, the 6-th and the 7-th segments of Cessna Citation S/II and the 2-rd and the 5-th segments of Yak-42 are selected as training samples, and the rest segments are taken as test samples. More concretely, there are totally 140, 000 training samples and 5, 200 test samples involved in our experiments.

As discussed in [22], it is a prerequisite for radar target recognition to deal with the target-aspect, time-shift, and amplitude-scale sensitivity. Similar to the previous study [10], [22], HRRP training samples should be aligned by the time-shift compensation techniques in ISAR imaging [32] to avoid the influence of time-shift sensitivity. Each HRRP sample is normalized by l_2-normalization algorithm to avoid the amplitude-scale sensitivity. In the following experiments, unless otherwise stated, all of the HRRPs are assumed to be aligned and normalized, and at high signal-to-noise ratio (SNR).

TABLE 2. Average recognition rates of proposed method with various number of convolutional layers based on aligned test samples.

Method	Architecture	Accuracy(%)
CNN-BiRNN-1	32-BiRNN	90.5
CNN-BiRNN-2	32-32-BiRNN	91.6
CNN-BiRNN-3	32-32-64-BiRNN	92.7
CNN-BiRNN-4	32-32-64-64-BiRNN	92.9
CNN-BiRNN-5	32-32-64-64-128-BiRNN	93.1
CNN-BiRNN-6	32-32-64-64-128-128-BiRNN	93.3

B. EXPERIMENTAL SETUP

As shown in section III, our CNN-BiRNN based method consists of L convolutional layers and a fixed-structure bidirectional RNN with attention mechanism. Since the longer kernels perform better in [20], following their work, we apply kernel size of 1×9 in the convolutional layers. Each convolutional layer is followed by a 1×2 with stride 2 pooling layer. The size of the hidden state in BiRNN and the size of attention in the attention mechanism are both set to 200. To balance the learning efficiency and stability of the model, through grid search, we perform RMSProp optimizer with a learning rate of 0.00001 and a mini-batch size of 100. All weights are initialized from a normal Gaussian distribution with its standard deviation set to 0.01. The recognition results are obtained by averaging classification rates from each category. All experiments are performed on a desktop computer equipped with an Intel Core i7-4770 CPU, 32 GB RAM, and a graphics card NVIDIA GeForce GTX 1070.

As the HRRP has the characteristic of time-shift sensitivity, the position of target areas in a HRRP sample may change during measuring. Thus, besides exhibiting the average recognition rate of the proposed methods with aligned test samples, we also show that with unaligned test samples to evaluate the robustness to the time-shift sensitivity. Without loss of generality, in the experiments of testing time-shift sensitive, we align each training sample with the centroid of sample and shift the aligned test samples from -16 to 16 range cells with a step size 1 to simulate the time-shift.

C. IMPACT OF CONVOLUTIONAL LAYERS

Since each convolutional layer is followed with a pooling layer, the sequence length of $\mathbf{H}^{(L)} \in \mathbb{R}^{W^{(hL)} \times K_L}$ shrinks as the number of convolutional layers L grows. To explore the relationship between sequence length and recognition result, and find the optimal network structure, we conduct several experiments of the proposed method with different numbers of convolutional layers with aligned and unaligned test samples. The results are shown in Table 2 and Fig. 6, respectively. For notational convenience, we denote CNN-BiRNN-L to represent the recognition method of CNN-BiRNN with L convolutional layers. Note that we set the output channels of layers 1 and 2 to 32, double the output channels of layers 3 and 4, and double again at layers 5 and 6.

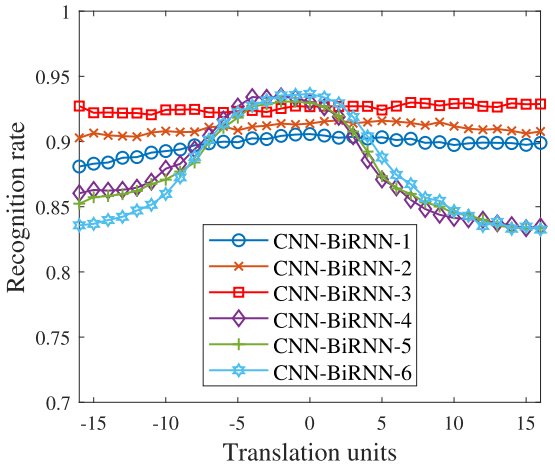


FIGURE 6. Variation of the recognition rate with the different extents of misalignment, via CNN-BiRNN-1, CNN-BiRNN-2, CNN-BiRNN-3, CNN-BiRNN-4, CNN-BiRNN-5 and CNN-BiRNN-6.

From Table 2, we could know that deeper models with aligned test samples achieve better results, because they have more powerful nonlinear expressiveness. However, in Fig. 6, the deeper the better is not always true, as the ability to resist time-shift is not benefiting from enlarging the depth of the model. The reason is that the deeper convolutional network has the larger receptive field at the top layer. More detailedly, in our CNN part, since the kernel size is 1×9, the receptive fields at the top layer of CNN-BiRNN-4, CNN-BiRNN-5 and CNN-BiRNN-6 are 72, 144 and 288, which, compared with the lengths/widths of the targets falling in $40 \sim 100$ range cells, are too large and make the sequential of the top layer representation ambiguous. In Fig. 7, we show the top layer features learned from CNN-BiRNN-3 and CNN-BiRNN-5 via a test sample with shift units of -16, 0 and 16, respectively. We can find that the representations at the top layer of CNN-BiRNN-3 are nearly unchanged owing to a relatively proper receptive field, while that of CNN-BiRNN-5 vary greatly. It results in a rapid decline on recognition rate with shifted test samples, as the models are trained based on the aligned data. Consequently, to consider both of the recognition performance and the robustness to the time-shift sensitivity, we choose CNN-BiRNN-3 for further comparisons in the following experiments.

D. RECOGNITION PERFORMANCE

To evaluate the efficiency of the proposed method, we compare CNN-BiRNN-3 with several existing HRRP recognition methods, including three traditional shallow methods: Maximum correlation coefficient (MCC) [7], Adaptive Gaussian classifier (AGC) [10] and Linear support vector machine (LSVM), and four deep learnin-based methods: Deep Belief Networks (DBN) [33], Denoising Autoencoders (DAE) [34], Convolutional neural networks (CNN) [20] and Target-aware recurrent attentional network (TARAN) [19]. The MCC and AGC are two classic statistical recognition models that use Gaussian distribution to model the amplitude of HRRP.

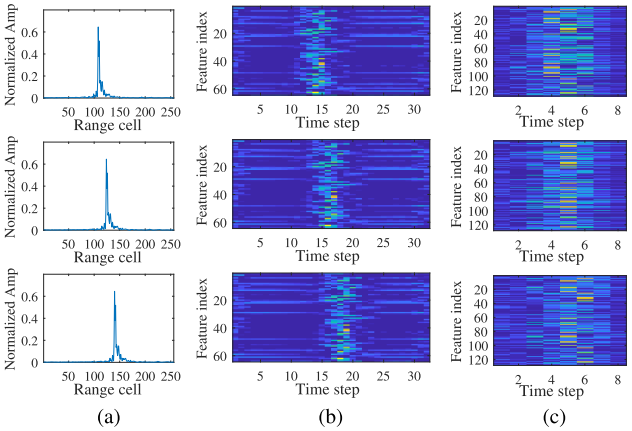


FIGURE 7. (a) an HRRP sample shifted by -16 (top), 0 (middle) and 16 (bottom) units; (b) the top convolutional feature learned from CNN-BiRNN-3; (c) the top convolutional feature learned from CNN-BiRNN-5.

TABLE 3. Average Recognition Rates of the Proposed Method and the Counterpart Methods based on Aligned Test Samples.

No.	Method	Accuracy(%)
1	MCC	62.42
2	AGC	85.63
3	LSVM	86.70
4	DBN-3	89.29
5	DAE-3	90.42
6	TARAN	90.10
7	CNN-3	92.57
8	**CNN-BiRNN-3**	**92.66**

The LSVM is an efficient machine learning algorithm designed to minimize structural risk for good generalization performance. The DBN and DAE are two typical fully-connected neural network based models while the CNN is an effective local structural feature extractor. Since the three-layer DBN, DAE and CNN achieve better classification accuracy in [18] and [20], we only exhibit the results of these methods with three hidden layers, denoted as DBN-3, DAE-3 and CNN-3. The TARAN, which contains one hidden layer, is an RNN with an attention mechanism to model the temporal dependence and find the informative areas in HRRP.

1) HIGH SNR SCENARIO

We first evaluate the recognition performance of the methods under raw test data, i.e., high SNR scenario, from two aspects: the recognition rate with and without the alignment operation on test samples. The setup of the experiments is the same as section V-B and the results are shown in Table 3 and Fig. 8, respectively. Apparently, in the scenario of aligned test samples, i.e., Table 3, the recognition rates of five deep methods are better than that of the shallow ones, since the deep models usually have stronger expressive ability. The CNN-BiRNN-3 and CNN get a relatively higher recognition rate than the other deep fully-connected networks, as the convolutional operation can extract more discriminative features from HRRP. The best recognition result is delivered by our

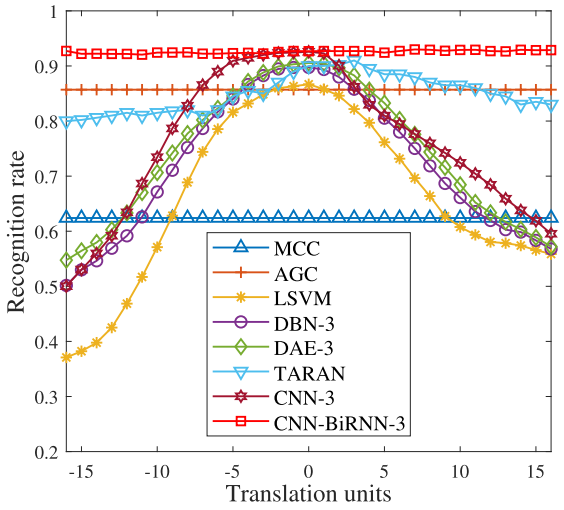


FIGURE 8. Variation of the recognition rate with the different extents of misalignment, via MCC, AGC, LSVM, DBN-3, DAE-3, TARAN, CNN-3 and CNN-BiRNN-3.

CNN-BiRNN-3 as it introduces not only a three layer CNN, but also an extra BiRNN for feature extraction.

However, the results are quite different in the scenario of misalignment test samples, i.e., Fig. 8, and the recognition rates of MCC, AGC, TARAN and CNN-BiRNN-3 are nearly unchanged during the translations. Although the phenomenon of time-shift invariance of these methods is similar, the reasons behind it are not all the same. The MCC and AGC are capable of resisting time-shift sensitivity because they align the test data with the stored templates during the test phase. For the CNN-BiRNN-3 and TARAN, they both employ RNN with attentional mechanism to model the temporal dependence of HRRP, which makes their feature extraction focus on the target region. For a further investigation, we exhibit the attention weights learned from CNN-BiRNN-3 via a test sample with different time-shift units in Fig 9. The envelopes of a HRRP sample with shift units of -16, 0 and 16 are shown in Fig 9(a), which have different target regions because of the time-shift. With the help of the data-dependent attention mechanism, as shown in Fig 9(b), the attention weights learned from CNN-BiRNN-3 are focused on and varied with the target area of HRRP even with the test sample having different extents of misalignment. As a result, the recognition results remain unchanged during the translations.

In addition, our CNN-BiRNN-3 performs at least 2% better than TARAN, as it adopts 3 convolutional layers for feature extraction rather than modeling the HRRP data directly. In Fig. 10, we show the first two principal components of the three aircrafts based on different features, where Fig. 10(a) is the raw HRRP data and Fig. 10(b) is the feature learned from the convolutional layer of CNN-BiRNN-3. It should be clarified that in Fig. 10(b), we perform the principal component analysis based on the output of the top convolution layer, which is also the input of BiRNN. Clearly, compared to the raw HRRP data (Fig. 10a), the features extracted from the convolutional layer of CNN-BiRNN-3 (Fig. 10b) are

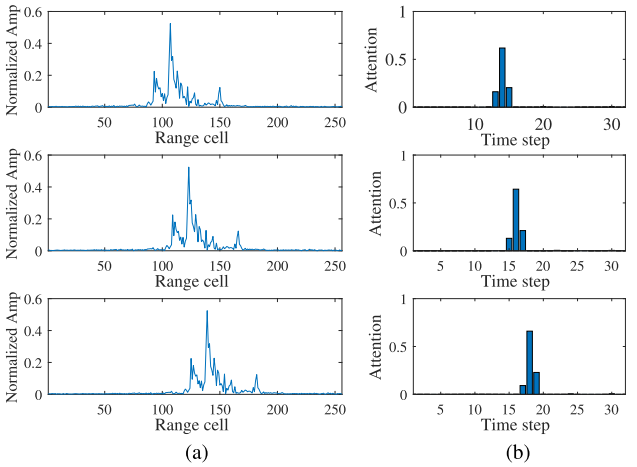


FIGURE 9. Illustration of the CNN-BiRNN-3 resist time-shift sensitivity; (a) an HRRP sample with translation units of -16 (top), 0 (middle) and 16 (bottom); (b) the attention weights learned by CNN-BiRNN-3.

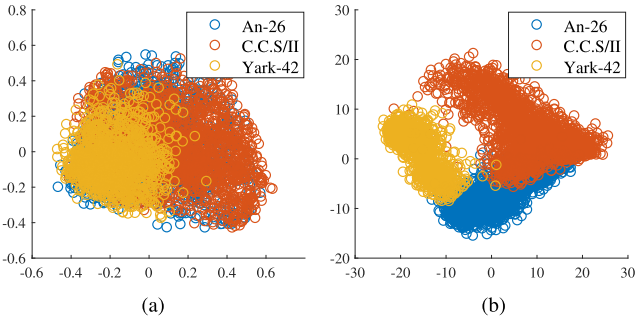


FIGURE 10. Distribution of the test samples (a) and their corresponding representations learned from CNN-BiRNN-3 (b) on the first two principal components.

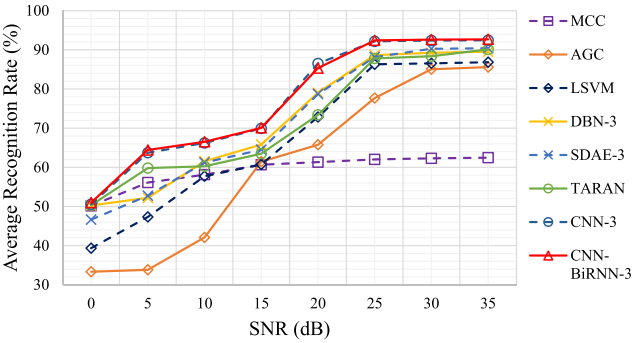


FIGURE 11. Variation of the recognition rates with SNR, via MCC, AGC, LSVM, DBN-3, DAE-3, TARAN, CNN-3 and CNN-BiRNN-3.

more separable between classes and centralized within them, which intuitively explains why our method achieves a better recognition rate than TARAN.

In summary, compared with the existing HRRP recognition methods, the proposed method not only achieves better recognition rate but also has strong ability to resist time-shift sensitivity in high SNR scenario.

2) NOISE ROBUSTNESS

In the real application, the noise level of a test sample is usually different from those of the training samples. Thus, to

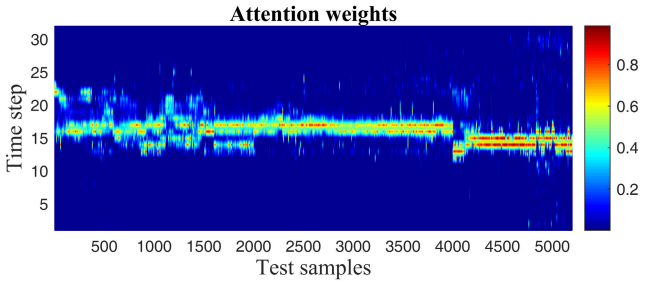


FIGURE 12. Attention weights α_t of the test samples learned from CNN-BiRNN-3. The x-axis 1 ∼ 2000, 2001 ∼ 4000, and 4001 ∼ 5200 respectively represent the test sample index of An-26, Cessna Citation S/II and Yark-42, while the y-axis corresponds to the time steps.

comprehensively evaluate the proposed method, we also exhibit the recognition rates under different SNR scenarios to verify the noise robustness of our method. According to [12], [13], we add simulated Gaussian white noise to the inphase and quadrature component of the high SNR raw test data to create low SNR test data. The SNR is defined as

$$SNR = 10 \times \log_{10}\left(\frac{P_x}{P_{Noise}}\right) = 10 \times \log_{10}\left(\frac{\sum_{l=1}^{L} P_{x_l}}{L \times P_{Noise}}\right) \quad (16)$$

where P_x, P_{x_l} respectively denote the average power of HRRP and the power of the original echo per range cell, L denotes the number of range cells (here $L = 256$), and P_{Noise} denotes the power of noise. Without loss of generality, we set the SNR ranging from 0dB to 35dB with the step size of 5dB, and the recognition rates are shown in Fig. 11. Generally, the recognition performance of all methods decreases with the increasing of the noise. When $SNR \geq$ 25dB, our method outperforms other models. When SNR is low, say $SNR <$ 25dB, our method and CNN-3 perform similarly and leading among all models thanks to CNN's robustness to noise [36].

E. VISUALIZATION

The proposed method provides an intuitive way to inspect the importance of time steps (regions) for HRRP samples. This is done by visualizing the attention weights α_t from (9), as shown in Fig. 12. Each column in Fig. 12 indicates the weights associated with each HRRP time steps, and the brighter part is the area that is more relevant to the recognition task. We can notice that the attention mechanism focuses on the middle areas and puts relative small weights on the others. Specifically, the attention weights of Cessna Citation mainly focus on the time step around 16 while An-26 and Yark-42 concentrate on larger regions. This phenomenon coincides with the characteristic of the targets, which describes that different target area sizes, as introduced in V-A, are equivalent to different time steps.

VI. CONCLUSION

In this paper, in order to extract efficient and time-shift invariant features from HRRP, we combine the CNN and BiRNN, and propose a CNN-BiRNN based method for the HRRP target recognition problem. In the proposed method, a CNN is employed to explore the spatial correlation and extract expressive features from raw HRRP data, and the resultant feature maps are fed into a BiRNN with the full consideration of temporal dependence. Furthermore, to enhance the robustness to misalignment, an attentional mechanism is employed after BiRNN to allow the method to focus on the discriminative target area. Experiments based on measured radar data showed that with the help of the efficient discriminative feature from CNN and the precise sequential modeling of BiRNN with attention mechanism, the proposed method not only achieved better recognition rate but also had strong capability of resisting time-shift sensitivity.

REFERENCES

[1] A. Zyweck and R. Bogner, "Radar target classification of commercial aircraft," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 598–606, Apr. 1996.
[2] R. Mitchell and J. Westerkamp, "Robust statistical feature based aircraft identification," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 35, no. 3, pp. 1077–1094, Jul. 1999.
[3] S. P. Jacobs and J. A. O'Sullivan, "Automatic target recognition using sequences of high resolution radar range-profiles," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 364–381, Apr. 2000.
[4] X. Liao, P. Runkle, and L. Carin, "Identification of ground targets from sequential high-range-resolution radar signatures," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 4, pp. 1230–1242, Oct. 2002.
[5] F. Zhu, X.-D. Zhang, Y.-F. Hu, and D. Xie, "Nonstationary hidden Markov models for multiaspect discriminative feature extraction from radar targets," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2203–2214, May 2007.
[6] X.-D. Zhang, Y. Shi, and Z. Bao, "A new feature vector using selected bispectra for signal classification with application in radar target recognition," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 1875–1885, Sep. 2001.
[7] L. Du, H. Liu, Z. Bao, and M. Xing, "Radar HRRP target recognition based on higher order spectra," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2359–2368, Jun. 2005.
[8] J. Chai, H. Liu, and Z. Bao, "Combinatorial discriminant analysis: Supervised feature extraction that integrates global and local criteria," *Electron. Lett.*, vol. 45, no. 18, p. 934, 2009.
[9] L. Du, H. Liu, Z. Bao, and J. Zhang, "Radar automatic target recognition using complex high-resolution range profiles," *IET Radar Sonar Navigat.*, vol. 1, no. 1, p. 18, 2007.
[10] L. Du, H. Liu, and Z. Bao, "Radar HRRP statistical recognition: Parametric model and model selection," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1931–1944, May 2008.
[11] L. Shi, P. Wang, H. Liu, L. Xu, and Z. Bao, "Radar HRRP statistical recognition with local factor analysis by automatic Bayesian Ying-Yang harmony learning," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 610–617, Feb. 2011.
[12] L. Du, P. Wang, H. Liu, M. Pan, F. Chen, and Z. Bao, "Bayesian spatiotemporal multitask learning for radar HRRP target recognition," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3182–3196, Jul. 2011.
[13] L. Du, H. Liu, P. Wang, B. Feng, M. Pan, and Z. Bao, "Noise robust radar HRRP target recognition based on multitask factor analysis with small training data size," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3546–3559, Apr. 2012.
[14] X. Zhang, B. Chen, H. Liu, L. Zuo, and B. Feng, "Infinite max-margin factor analysis via data augmentation," *Pattern Recognit.*, vol. 52, pp. 17–32, Apr. 2016.
[15] B. Feng, L. Du, H. W. Liu, and F. Li, "Radar HRRP target recognition based on K-SVD algorithm," in *Proc. IEEE CIE Int. Conf. Radar*, Piscataway, NJ, USA, vol. 1, Oct. 2011, pp. 642–645.
[16] D. Zhou, "Radar target HRRP recognition based on reconstructive and discriminative dictionary learning," *Signal Process.*, vol. 126, pp. 52–64, Sep. 2016.
[17] M. Pan, J. Jiang, Q. Kong, J. Shi, Q. Sheng, and T. Zhou, "Radar HRRP target recognition based on t-SNE segmentation and discriminant deep belief network," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 9, pp. 1609–1613, Sep. 2017.

[18] B. Feng, B. Chen, and H. Liu, "Radar HRRP target recognition with deep networks," *Pattern Recognit.*, vol. 61, pp. 379–393, Jan. 2017.
[19] B. Xu, B. Chen, J. Wan, H. Liu, and L. Jin, "Target-aware recurrent attentional network for radar HRRP target recognition," *Signal Process.*, vol. 155, pp. 268–280, Feb. 2019.
[20] J. Wan, B. Chen, B. Xu, H. Liu, and L. Jin, "Convolutional neural networks for radar HRRP target recognition and rejection," *EURASIP J. Adv. Signal Process.*, vol. 2019, no. 1, p. 5, 2019.
[21] B. Chen, H. Liu, J. Chai, and Z. Bao, "Large margin feature weighting method via linear programming," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 10, pp. 1475–1488, Oct. 2009.
[22] L. Du, H. Liu, Z. Bao, and J. Zhang, "A two-distribution compounded statistical model for radar HRRP target recognition," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2226–2238, Jun. 2006.
[23] B. Chen, H. Liu, L. Yuan, and Z. Bao, "Adaptively segmenting angular sectors for radar HRRP automatic target recognition," *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 1, 2008, Art. no. 641709.
[24] H.-W. Liu, B. Chen, B. Feng, and L. Du, "Radar high-resolution range profiles target recognition based on stable dictionary learning," *IET Radar, Sonar Navigat.*, vol. 10, no. 2, pp. 228–237, Feb. 2016.
[25] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997.
[26] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.* Red Hook, NY, USA: Curran Associates, 2012, pp. 1097–1105.
[27] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Piscataway, NJ, USA, Jun. 2015, pp. 1–9.
[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Piscataway, NJ, USA, Jun. 2016, pp. 770–778.
[29] A. Graves, A. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 6645–6649.
[30] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 3104–3112.
[31] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, May 2015, pp. 1–15. [Online]. Available: http://arxiv.org/abs/1409.0473
[32] J. Walker, "Range-Doppler imaging of rotating objects," *IEEE Trans. Aerosp. Electron. Syst.*, vols. AES-16, no. 1, pp. 23–52, Jan. 1980.
[33] G. E. Hinton, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.
[34] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. 25th Int. Conf. Mach. Learn. (ICML)*, New York, NY, USA, 2008, pp. 1096–1103.
[35] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
[36] Y. Qian, M. Bi, T. Tan, and K. Yu, "Very deep convolutional neural networks for noise robust speech recognition," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 24, no. 12, pp. 2263–2276, Dec. 2016.

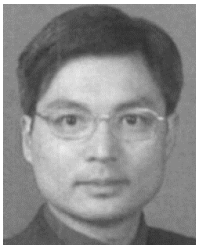
JINWEI WAN received the B.S. degree from the Guilin University of Electronic Technology, Guilin, China, in 2009, and the M.S. degree from Xidian University, Xi'an, China, in 2012, where he is currently pursuing the Ph.D. degree with the National Laboratory of Radar Signal Processing. His research interests include machine learning and radar automatic target detection and recognition.

BO CHEN (Member, IEEE) received the B.S. and Ph.D. degrees in electronic engineering from Xidian University, Xi'an, China, in 2003 and 2008, respectively. From 2008 to 2013, he was a Research Scientist with the Department of Electrical and Computer Engineering, Duke University. In 2014, he was selected to Young Thousand Talents Program and worked as a Professor at Xidian University. His research interests include statistical machine learning, statistical signal processing, and radar automatic target detection and recognition.

YINGQI LIU was born in 1995. He is currently pursuing the master's degree with Xidian University. His research interests include radar automatic target recognition and machine learning.

YIJUN YUAN was born in 1995. He is currently pursuing the master's degree with Xidian University. His research interests include radar automatic target recognition and adversarial example.

HONGWEI LIU (Member, IEEE) received the M.S. and Ph.D. degrees in electronic engineering from Xidian University, Xi'an, China, in 1995 and 1999, respectively. He worked at the National Laboratory of Radar Signal Processing, Xidian University. From 2001 to 2002, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA. He is currently a Professor with the National Laboratory of Radar Signal Processing, Xidian University. His research interests are radar automatic target recognition, radar signal processing, and adaptive signal processing.



LIN JIN is currently a Professor with the National Laboratory of Radar Signal Processing, Xidian University. His research interests are radar signal processing, radar system design, and radar automatic target recognition.