Article

A Novel Multi-Input Bidirectional LSTM and HMM Based Approach for Target Recognition from Multi-Domain Radar Range Profiles

Fei Gao [1],[†] [ID]**, Teng Huang [1],*,[†]** [ID]**, Jun Wang [1],*, Jinping Sun [1]** [ID] **and Amir Hussain [2] and Huiyu Zhou [3]** [ID]

[1]   School of Electronic and Information Engineering, Beihang University, Beijing 100191, China; feigao2000@163.com (F.G.); sunjinping@buaa.edu.cn (J.S.)

[2]   School of Computing, Edinburgh Napier University, Edinburgh EH10 5DT, Scotland, UK; a.hussain@napier.ac.uk

[3]   Department of Informatics, University of Leicester, Leicester LE1 7RH, UK; hz143@leicester.ac.uk

*   Correspondence: huangteng1220@buaa.edu.cn (T.H.); wangj203@buaa.edu.cn (J.W.); Tel.: +86-010-82317240 (J.W.)

†   These authors contributed equally to this work.

check for
updates

Abstract: Radars, as active detection sensors, are known to play an important role in various intelligent devices. Target recognition based on high-resolution range profile (HRRP) is an important approach for radars to monitor interesting targets. Traditional recognition algorithms usually rely on a single feature, which makes it difficult to maintain the recognition performance. In this paper, 2-D sequence features from HRRP are extracted in various data domains such as time-frequency domain, time domain, and frequency domain. A novel target identification method is then proposed, by combining bidirectional Long Short-Term Memory (BLSTM) and a Hidden Markov Model (HMM), to learn these multi-domain sequence features. Specifically, we first extract multi-domain HRRP sequences. Next, a new multi-input BLSTM is proposed to learn these multi-domain HRRP sequences, which are then fed to a standard HMM classifier to learn multi-aspect features. Finally, the trained HMM is used to implement the recognition task. Extensive experiments are carried out on the publicly accessible, benchmark MSTAR database. Our proposed algorithm is shown to achieve an identification accuracy of over 91% with a lower false alarm rate and higher identification confidence, compared to several state-of-the-art techniques.

Keywords: automatic target recognition; human–machine interaction; recurrent neural network; deep learning

1. Introduction

Radar is an active detection sensor. It is not disturbed by natural light, fog, or rainy weather [1]. It can recognize targets of interest by transmitting and receiving electromagnetic waves. Therefore, radar automatic target recognition technology plays an important role in human–machine interaction designs of various intelligent devices such as autonomous driving [2–4], intelligent wheelchair [5,6], and so on.

High-resolution range profile (HRRP) is a one-dimensional signal that can represent the geometric shape of the target in the direction of the radar line of sight. Recognition of the target represented in HRRP is an important approach for radars to monitor targets of interest. Compared with two-dimensional imaging of radars [7–9], HRRP has the advantages of easy acquisition, processing, and storage; hence using HRRP to achieve radar target recognition is more suitable for human–machine

interaction designs of intelligent machine. However, HRRP has shortcomings such as target aspect, translation and amplitude sensitivity [10,11]. Therefore, radar target recognition based on HRRP is a complex and nonlinear classification problem.

As one-dimensional signals, the target information provided by HRRP is very limited. In this paper, we are interested to construct a new classifier based on the limited information. In response to this problem, many researchers have conducted extensive research and proposed many landmark methods. These are usually based on: (1) the distribution features of HRRP; (2) the data domain of HRRP; (3) different types of HRRP.

The recognition algorithms based on HRRP distribution features use statistical theory as the main analysis tool to identify the target of interest. For example, it was reported that the HRRP signal has specific distributions under some conditions. These distributions included β distribution [12], Gaussian distribution [13], γ distribution [14], double distribution composite model [15] and so on. Subsequently, the target of interest can be identified by extracting statistical features such as center distance and invariant moment. In addition, some researchers attempted to construct statistical modeling of HRRP from the target. When the position of the radar observation is unchanged, the aspect changes of the target were considered to be a random non-stationary process, given that the angular changes of these aspects exceed a certain range. Otherwise, they can be approximated as a stationary process. These processes can be described by a Hidden Markov Model (HMM), which has been widely used in target recognition based on multi-aspect HRRP sequences [16–18]. The recognition algorithms based on HRRP data domain mainly explore their characteristics from different data domains of HRRP. These data domains include time, frequency, and time-frequency domain, and result in different classification outcomes. For example, Liao et al.[16] used the time-domain amplitude features of HRRP to train HMMs, and obtained 82% recognition accuracy in 3° aperture data, and 92% recognition accuracy in 6° aperture data. Albrecht et al.[17] used the frequency-domain power spectrum features derived from the time-domain HRRP to train HMMs. Due to the translation invariance property of the power spectrum features, 94% recognition accuracy was obtained in [17]. Zhang et al.[18] used time-frequency domain (T-F) features to train HMMs. With more abundant features, 95% recognition accuracy was obtained in [18]. The recognition algorithms based on multi-look HRRP are also constructed from the aspect angle of the target of interests. However, unlike the recognition algorithms based on HRRP distribution features, they mainly perform non-coherent superposition of multiple HRRPs within a certain aspect angle range of the target, to suppress noise and fuse multi-look aspect features to improve the recognition performance. There are many strategies for multi-look processing, which can generally be performed on the sample level [19–21] or the feature level [22,23]. Sample level multi-look processing generally uses the method of averaging multiple samples, while feature level multi-look processing first needs to extract features from individual samples separately, and then fuse these features. Compared with single-look algorithms, multi-look algorithms improve the recognition accuracy to a certain extent, in both sample and feature levels. For example, in [20], a 97% correct identification rate was achieved using the multi-look method, while only a 37% accuracy rate was obtained using the single-look method.

The above traditional recognition algorithms explore HRRP signals from different perspectives and have achieved promising success. However, these algorithms usually produce single features, which makes it difficult to further improve the recognition performance. In state-of-the-art recognition algorithms based on HRRP data domain, combining the advantages of several data domains has not been well investigated. On the other hand, in recognition algorithm based on multi-look HRRP, sample level multi-look processing makes it difficult control the number that needs to be averaged. If these quantities are too large, the processing time of the recognition algorithms will increase, and conversely, it is difficult to improve the recognition performance. Feature level multi-look processing has the potential to improve recognition performance, but the implementation of such algorithms is difficult.

"End-to-End" feature learning using deep learning provides a solution to the problems of feature extraction and fusion in conventional algorithms mentioned above. The essence is to use multi-layer

neural networks to automatically extract basic features of the original data, or use multi-branch neural networks to automatically fuse multiple features [24]. In recent years, many researchers have proposed several deep learning models based on different applications, such as Deep Convolutional Neural Network (CNN)[25], Stacked Autoencoders (SAE)[26], Restricted Boltzmann machines(RBMs) [27], and Long Short-Term Memory (LSTM) Recurrent Neural Network [28]. CNN is considered one of the best models to solve the "perception" problem, but is mainly used to extract two-dimensional data features [29]. The SAE and RBMs models are unsupervised learning methods, which have been used for HRRP target recognition[30,31]. However, due to lack of prior knowledge, these unsupervised learning methods are unable to maintain high recognition accuracy when dealing with multi-class recognition problems. LSTM is an improved recurrent neural network. It solves the problem of gradient disappearance of traditional recurrent neural networks using long-span prior information. Thus, LSTM has unique advantages in processing sequential data [32]. At present, LSTM is mainly used in speech recognition [33] and natural language processing [34]. Some researchers [35,36] used HRRP as a one-dimensional signal, and applied LSTM to HRRP target recognition. In dealing with different types of targets, Jithesh et al. and Bin et al.[35,36] have achieved good classification performance. These algorithms provide insights into exploring the sequence feature of HRRP. Further, some researchers combined LSTM with CNN by a fully connected layer with sequence characteristics [37,38]. Such combined algorithms can, on the one hand, exploit advantages of both LSTM and CNN; and on the other hand, effectively and automatically fuse the features that are irrelevant in a physical sense, thereby greatly improving the recognition performance. These algorithms also provide ideas for exploring the feature fusion of different data domains in HRRP. Furthermore, in recent years, some researchers e.g., [39] have proposed a bidirectional LSTM (BLSTM) algorithm based on the LSTM. The current features of sequential data were not only dependent on the past information, but also related to future information. This context-dependent relationship can further improve the ability of LSTM to process sequential data.

In summary, the construction of an optimal classifier based on limited HRRP signals to correctly extract, learn, and fuse HRRP features for improving target recognition, is still an open question. In contrast to traditional recognition algorithms, we first extract 2-D sequence features from the formation process of HRRP, and then generate different data domains from these 2-D sequence features. Based on these sequence features, a novel target identification method is presented by combining bidirectional Long Short-Term Memory (BLSTM) and a Hidden Markov Model (HMM). The proposed algorithm first learns and fuses the sequence features of different data domains, and then combines these with the sequence features of the target multi-aspect, to achieve recognition. It consists of a shallow CNN, a multi-input BLSTM, and multiple HMMs. The shallow CNN is used for learning and dimensionality reduction of two-dimensional features in the time-frequency domain; the multi-input BLSTM is used for feature fusion of different data domains; and HMMs are used for target multi-aspect sequence feature learning and final target recognition.

The rest of this paper is organized as follows. Section 2 gives a brief review of the approaches involved. Section 3 describes the proposed method in detail. Section 4 presents a pipeline of transforming the real synthetic aperture radar (SAR) image to HRRP and then generating multi-domain HRRP sequences. Section 5 reports comparative experimental results for radar target recognition. Finally, concluding remarks are given in Section 6.

2. LSTM and BLSTM

LSTM is a special recurrent neural network model. Through a special gate structure, it can store and retrieve information for a long time. Figure 1a shows an LSTM storage module. It comprises three gate structures (input i_n, forget f_n, and output o_n), a memory unit controller cell, two input and output activation units, and three peepholes connections. The input and output gates are used to control the block input and output of the cell, and the forgetting gates are used to control the memory and forgetting state of the cell. The peephole is connected with status information before

all the doors, to obtain Constant Error Carousel information [40] that allows the cell to record more sequential information. Finally, the block output information is recurrent, and connects to the block input and all other gates, which enables LSTM to model complex and long-term dynamic features, and solves the gradient disappearance problem caused by long sequences in traditional recurrent neural networks [41]. The forward mechanism of the LSTM can be expressed by the following equation [42]:

$$
\begin{aligned}
I_n &= h(W_I x_n + R_I O_{n-1} + b_I) \\
i_n &= \sigma(W_i x_n + R_i O_{n-1} + p_i \odot c_{n-1} + b_i) \\
f_n &= \sigma(W_f x_n + R_f O_{n-1} + p_f \odot c_{n-1} + b_f) \\
c_n &= i_n \odot I_n + f_n \odot c_{n-1} \\
o_n &= \sigma(W_o x_n + R_o y_{n-1} + p_o \odot c_n + b_o) \\
O_n &= o_n \odot h(c_n)
\end{aligned}
\tag{1}
$$

where $I_n, i_n, f_n, c_n, o_n,$ and O_n represent the *block input*, *input gate*, *forget gate*, *memory cells*, *output gate*, and *block output* respectively. n is the number of sequential data. x_n is the input feature at the nth sequence. W is the weight matrix. R is the recurrent weight matrix. b is the bias vector. p is the peephole weight vector, and the subscript I, i, f, o respectively represent the *block input*, *input gate*, *forget gate*, and *output gate*. σ is the logistic sigmoid activation function. h is the hyperbolic tangent activation function, and $\odot$ denotes point-wise product with the gate value.

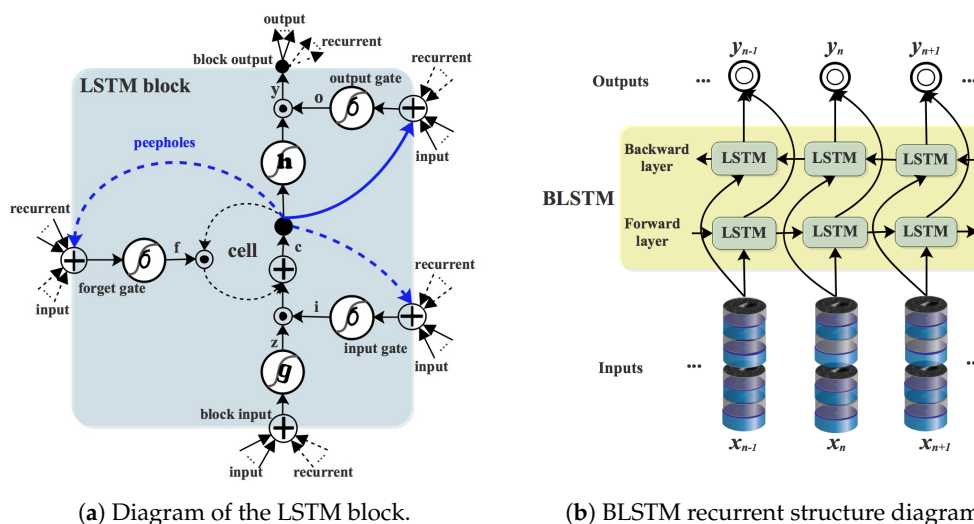(**a**) Diagram of the LSTM block.　　　　　　(**b**) BLSTM recurrent structure diagram.

Figure 1. LSTM and BLSTM model.

The LSTM has a disadvantage that it can only get past information but not future information. Compared with the LSTM, BLSTM solves this problem. As shown in Figure 1b, there are two independent LSTM networks in a BLSTM module. These two LSTM networks have different directions, one is a forward LSTM, and the other is a reverse LSTM. The forward LSTM is mainly used to extract the future information of sequence data, while the reverse LSTM is mainly used to extract past information of the sequence data. Finally, their results are connected to the same output unit, and future and past features are fused to produce the output. In this way, the BLSTM is able to extract and fuse future and past features of the sequence data, which can be expressed by the following formulas [43] :

$$
\begin{aligned}
\overrightarrow{O}_n &= \Gamma(W_{\overrightarrow{I}} x_n + W_{\overrightarrow{O}} O_{n-1} + b_{\overrightarrow{O}}) \\
\overleftarrow{O}_n &= \Gamma(W_{\overleftarrow{I}} x_n + W_{\overleftarrow{O}} O_{n+1} + b_{\overleftarrow{O}}) \\
y_n &= W_{\overrightarrow{y}} \overrightarrow{O}_n + W_{\overleftarrow{y}} \overleftarrow{O}_n + b_y
\end{aligned}
\tag{2}
$$

where $\overrightarrow{O}_n$ is the forward hidden sequence, $\overleftarrow{O}_n$ is the backward hidden sequence, and Γ is implemented by Equation (1).

In this paper, the BLSTM is used to extract future and past features of the sequence data to implement learning of multi-domain HRRP sequence features. However, in high-dimensional sequence data, the BLSTM is computationally inefficient and has limited ability to extract features. Therefore, choice of appropriate feature extraction and dimensionality reduction algorithms, combined with the BLSTM, are key to solving this problem.

3. Proposed Methods

The proposed target recognition framework, termed MIBL-HMM, is illustrated in Figure 2, including four processing steps, specifically: generating a multi-domain HRRP sequence, reducing feature dimensionality using a shallow CNN, fusing multi-domain sequence features with the multi-input BLSTM, and determining the category of each target sample via a HMM classifier.
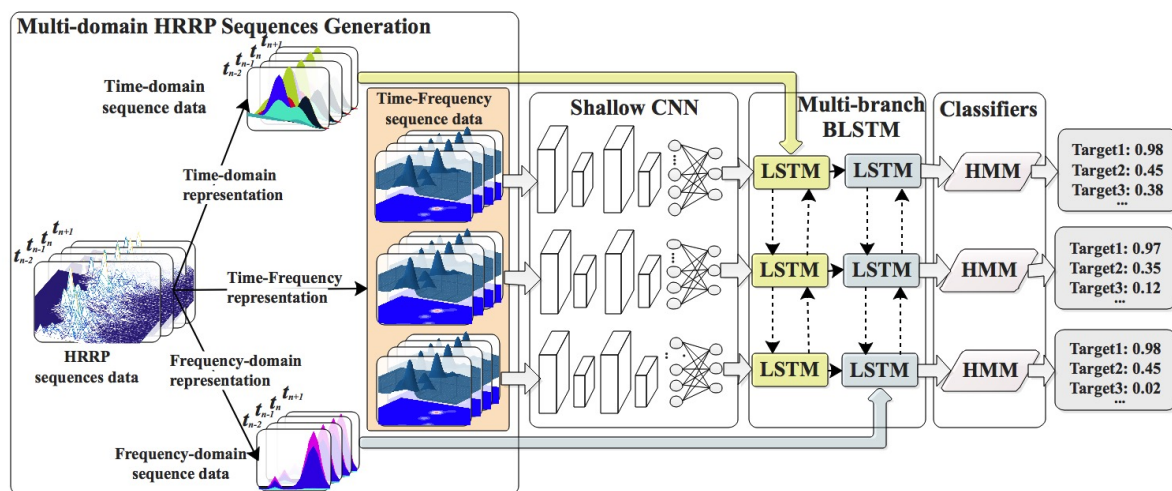

Figure 2. A flowchart of the proposed HRRP ATR framework based on BLSTM and HMM.

3.1. Multi-Domain HRRP Sequence Generation

The generation of a multi-domain HRRP sequence is divided into two steps: (1) extraction of the HRRP sequence; (2) Based on the HRRP sequence, multiple data domains are generated, termed the multi-domain HRRP sequence. Details are given next.

(1)　Extraction of the HRRP sequence:

HRRP is a one-dimensional signal that reflects different scatters of the target in the direction of the line of sight radar. One HRRP represents the target's response to one high range resolution radar pulse. In one HRRP, the amplitude of each range unit represents the intensity of the electromagnetic wave reflected by different scatters of the target. In a certain range of observation angles, the order of these intensities reflects the order of different scatters of the target in the direction of line of sight radar. Figure 3a shows one HRRP generated by a tank in response to one pulse of the radar. In this HRRP, these range units present the features of tank components in the order of the line of sight radar, specifically: first, the head of the tank; secondly, the body of the tank; and finally, the tail of the tank. Therefore, the range units contained in one HRRP have obvious sequence characteristics. This conclusion is consistent with [35,36]. In contrast to the literature, as shown in Figure 3b, we believe multiple HRRPs also possess timing sequence characteristics. This is because the electromagnetic pulses emitted by the radar have obvious timing sequence characteristics, so the HRRPs collected continuously also exhibit time-series characteristics. Subsequently, integrated with the sequence characteristics of a single HRRP, multiple HRRPs continuously acquired by the radar, will have obvious 2-D sequence characteristics. In this paper, we will extract HRRP sequences from these multiple HRRPs with 2-D sequence characteristics.
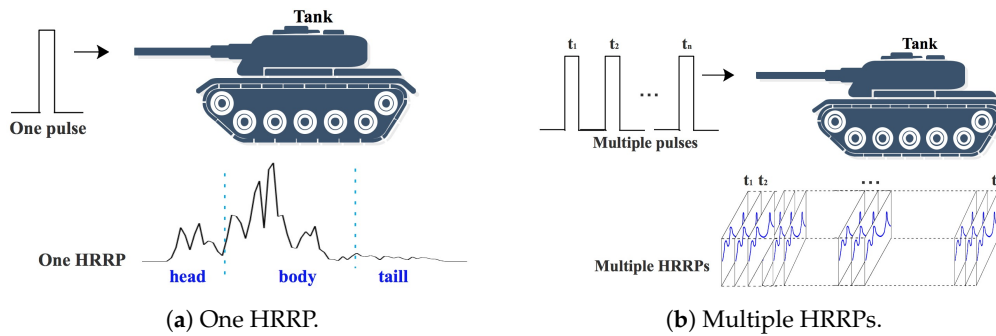
(**a**) One HRRP. (**b**) Multiple HRRPs.

Figure 3. Extraction of the HRRP sequence.

For a period of time, there is little difference, in theory, between multiple HRRPs continuously collected by the radar in a certain pose of the same target. However, due to various factors such as the environment, the difference between these HRRPs is relatively large. To alleviate these gaps and remove noise, the HRRPs collected during this period are arbitrarily divided into *n* groups which are not intersecting with each other. Next, each group is averaged by the multi-look processing strategy at the sample level. Under the condition that the original order is unchanged, a new set of *n* HRRPs is obtained, which is called a HRRP sequence. Such a HRRP sequence contains *n* elements, and the amount of information contained in each element is related to the average number of HRRPs. The larger the number, the more the amount of information. The information contained in each HRRP can be expressed by the radar aperture angle. For example, the radar aperture angle corresponding to each HRRP is $0.01°$. When the average number of HRRPs is 10, the information contained in each element in the HRRP sequence is $0.1°$ aperture. It should be noted that this average number is limited for different purposes of HRRP. When HRRPs are used as training sets, the average number can take a large value, which is beneficial for the learning of target features. However, when HRRPs are used as test sets, this average number generally does not take a larger value. In this case, it is only necessary to select the corresponding number in each group of the split, to generate a test HRRP sequence. Currently, although each sequence of the training set and the test set contain the same elements, each element of the test set sequence contains less information than each element of the training set sequence. This is to save time of data processing in the recognition algorithm and speed up recognition speed.

(2)   Generation of multi-domain HRRP sequence:

The various data domains considered in this paper include time domain, frequency domain, and time-frequency domain. Therefore, it is only necessary to generate these three data domains for each element of the HRRP sequence to obtain the multi-domain HRRP sequence. For the time domain and frequency domain features, the amplitude, value and power spectrum of the HRRP signal are used, respectively. For the time-frequency domain features, there are many extraction methods, such as short-time Fourier transform (STFT), Wigner-Ville distribution (WVD), and Adaptive Gaussian representation (AGR). STFT is the most commonly used time-frequency domain analysis method. It represents the signal characteristics at a certain moment, by a segment of the signal in the time window. However, the precision of time-frequency domain features obtained by the STFT method in time resolution and frequency resolution, is not available at the same time. The WVD method makes up for the shortcomings of the STFT method, but it leads to the problem of "cross-term interference". Compared with these time-frequency domain feature extraction methods, the AGR method can decompose the echo signal into time-frequency center and local resonance of time-frequency domain resolution, and adjust the corresponding parameters using a Gaussian basis function to optimize the relationship between them. It can not only solve the "cross-term interference" problem faced by the WVD method, but also extract the adaptive spectrogram features effectively in the time-frequency plane. To-date, it has been successfully applied to ISAR imaging and radar target recognition [44,45]. Therefore, in this work, we use the AGR method to extract the time-frequency features of HRRP, as in [18].

3.2. Feature Dimensionality Reduction Using a Shallow CNN

A key step in radar target recognition based on time-frequency data is the dimensionality reduction of time-frequency features. Dimensionality reduction of data generally faces such dilemma: on the one hand, we hope to reduce the dimensionality of data as much as possible to avoid redundancy and complexity; on the other hand, we hope to retain as many features of data as possible to expand the learning space for extracting effective features. To alleviate this problem, this paper uses a shallow CNN to reduce the dimensionality of time-frequency features.

The shallow CNN comprises an input layer, two convolution layers, and a fully connected layer. In the input layer, the number of neurons is equal to the dimension of time-frequency features. In the fully connected layer, the number of neurons is equal to the dimension of the BLSTM's neural network input layer. To train the weights between the input layer and the hidden layer, we add the *softmax* classifier on the fully connected layer. The number of the classifiers is equal to the number of target categories. The shallow CNN model can be expressed as follows:

$$f(y_i|X) = softmax[b^{(3)} + W^{(3)}(flatten(R(b^{(2)} + W^{(2)} \otimes (R(b^{(1)} + W^{(1)} \otimes X))))]\tag{3}$$

where X is the input feature vector of a time-frequency matrix, y_i indicates class i, $f(\cdot)$ is the probability of X belonging to class i, *softmax* represents the softmax function, $b(k)$ is the bias of the kth layer, $W(k)$ represents weights between the kth and $(k+1)^{th}$ layers, $\otimes$ represents the convolution operation, $flatten(\cdot)$ is a function which connects the convolution layer and the fully connected layer by collapsing an array into one dimension, and R indicates the Rectified Linear Unit (ReLU) activation function.

In the training stage, the weights of the convolution layer and the fully connected layer are initialized with random values, and then continuously adjusted by the back-propagation algorithm, under the condition of decreasing the cross-entropy loss. The training process is repeated till the learning error falls below a moderate tolerance level. Then, the dimensionality-reduced features x can be obtained using the feed-forward CNN network:

$$x(W^{(1)}, X) = R(b^{(1)} + W^{(1)} \otimes X)\tag{4}$$

Unlike common dimensionality reduction methods (such as principal component analysis, linear discriminant analysis, etc.), the shallow CNN has obvious advantages. It can not only reduce the dimensionality of data, but also extract spatial features of data by connecting with the fully connected layer in a non-linear way, which effectively enhances the feature preservation of data after dimensionality reduction.

3.3. Multi-Domain Features Fusing with Multi-Input BLSTM

The aim of the multi-input BLSTM is to learn and fuse the features of multi-domain HRRP sequences. It consists of two parts:

(1) Multi-input, i.e., the algorithm has multiple input interfaces. These are divided into a main input and branch inputs. There is only one main input, which is located in the first layer, and there are many branch inputs, which are in subsequent layers other than the first layer. In terms of the importance level of features, the most important features are input from the main input interface, and secondary features are input from the branch input interface. The deeper the branch interface, the less important feature input is received. The design of the multi-input, on the one hand side, ensures the most important features are learned using the whole neural network, i.e., by extracting the most abstract and easily recognizable features; and on the other hand, in the process of learning the most important features, it gradually fuses other features in order to suppress the neural network over-fitting, and finally improves the recognition performance of the algorithm.

Figure 4 shows the multi-input BLSTM used in our algorithm. It contains three input interfaces: one main input and two branch inputs. The main input interface receives time-frequency sequence

features after dimensionality reduction, the first branch interface receives the time-domain sequence features, and the second branch interface receives the frequency-domain features. The important levels of these three different data domain features will be discussed in the experimental section.
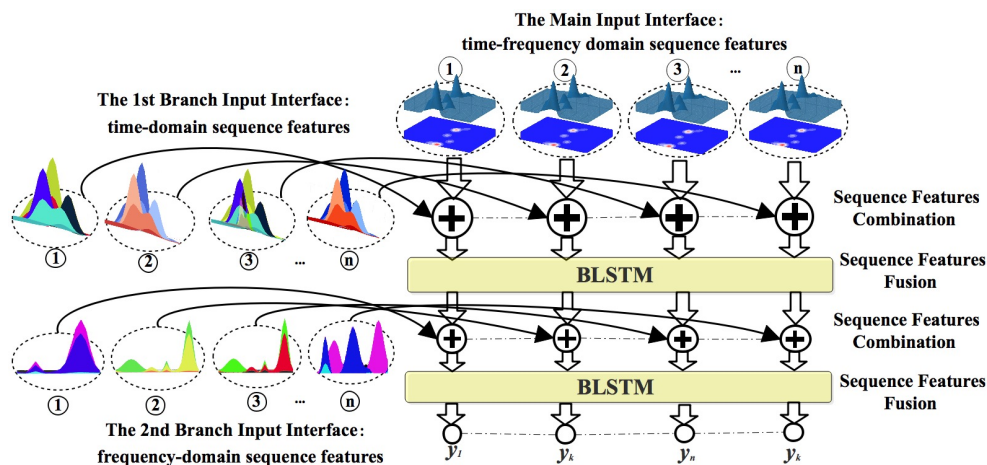

Figure 4. The proposed multi-input BLSTM.

(2) Features combination and fusion. In the process of learning the main input features, the multi-input BLSTM needs to gradually fuse features. At this time, it will face the problem of combining and fusing the features of two different data domains. There are many ways to combine sequence data, such as element-wise addition, subtraction, multiplication, or concatenation. Here, we use element-wise addition, which is the addition of the corresponding elements, as shown in Figure 4. In this way, simple linear addition does not destroy the features of different data domain sequences; and, the value of each range bin in HRRP is small, and can be appropriately increased by the addition with highlighted features. After the combination, the sequence features will go into the BLSTM network to automate the fusion of different data domain sequence features. Since each BLSTM network consists of two recurrent neural networks, one is forward processing data, the other is the backward processing data, both of which are connected to the same output layer so that each BLSTM network has the ability to learn the past and future information from data. Therefore, in the process of sequence feature fusion, the multi-input BLSTM cannot only learn and fuse the sequence features of different data domains at the current time, but also fuse the past and future information of these sequence features.

In conclusion, compared with the standard multi-layer BLSTM, the multi-input BLSTM enhances the ability of the standard multi-layer BLSTM with multi-input functionality. Compared with traditional HRRP target recognition methods, the multi-input BLSTM can deeply learn rich features contained in the HRRP data. Here, we use a multi-input BLSTM network consisting of three layers: input, hidden, and output layer. The input layer is three inputs mentioned above, the hidden layer includes two BLSTM networks, and the output layer is a fully connected layer with sequence features. The multi-input BLSTM is trained in the same way as the shallow CNN. Finally, the multi-domain HRRP sequences go into the trained multi-input BLSTM to produce the fused feature sequences, which are sent to a HMM classifier to learn and classify the multi-aspect features.

3.4. Multi-Aspect Features Learning and Classification with HMM

Multi-domain HRPP sequences are fused by the multi-input BLSTM to obtain a sequence feature. Each sequence feature corresponds to an aspect angle of the target. Before entering the HMM classifier, these sequence features need to be processed as follows.

In the HMM training stage, these sequence features are connected in series to form a matrix. As shown in Figure 5, the horizontal direction represents the aspect angle and the vertical direction represents the sequence features. Here, we use this feature matrix to construct the HMM model, i.e., to

establish the state of HMM and the mode of transition of these states, and then calculate the probability π and transition probability matrix *A* of the initial state of the corresponding HMM model. Since the details of the HMM-based multi-aspect feature classification algorithm used here have been described in several papers [16–18], we briefly summarize the basic concepts in this section.
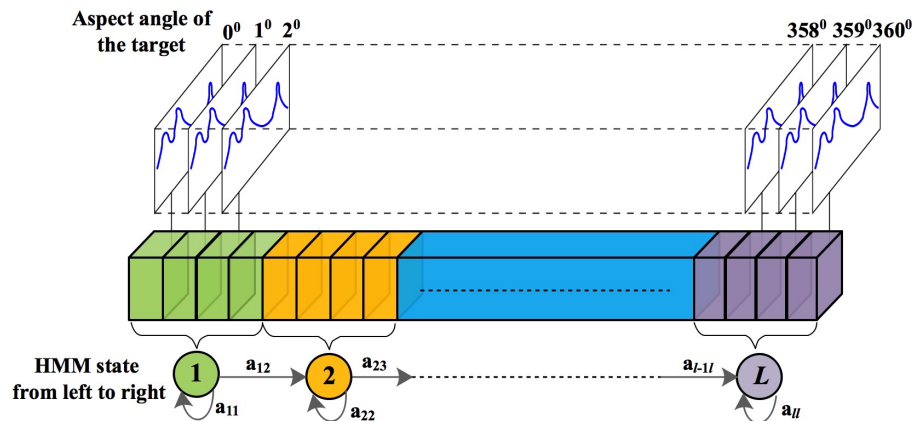

Figure 5. The basic process of HMM model construction.

(1) Establishing the state of the HMM model and the mode of these state transitions. Using the average aspect angle, the omni-directional angular space of the radar observation target, namely 360 degrees, is divided into *L* states, and the angular extent of each state is $360/L$ degrees. In addition, to reduce the parameters of the HMM model, the initial transition mode of the HMM model is from left to right, as shown in Figure 5.

(2) Calculating the probability π and transition probability matrix *A* of the initial state of the HMM model. When the radar observation position is constant, let $\delta\theta$ represent the range of the target aspect change when the radar continuously measures the target. Let θ_i represent the range of the state *i* (angle change). Regarding the state transition of the HMM, the state *i* is subjected to $\delta\theta < \theta_i$. Assuming that $A = \{a_{ij}\}$, a_{ij} represents the probability of state *i* to the next state *j*, then the initial value of *A* is as follows:

$$
\begin{aligned}
a_{i,i-1} = a_{i,i+1} = \delta\theta/2\theta_i, \\
a_{i,i} = (\theta_i - \delta\theta)/\theta_i
\end{aligned}
\tag{5}
$$

Moreover, the initial orientation of the target is generally considered to be uniformly distributed, so the probability of the initial state can be calculated:

$$
\pi_i = \theta_i / \sum_{i=1}^{N} \theta_i
\tag{6}
$$

As discussed above, Equations (5) and (6) constitute the initial estimates of *A* and π. These parameters can be better estimated by the Baum-Welch method. Through the above methods, the HMM model can be established and trained. Each type of target corresponds to an HMM model, i.e., the number of HMM models depend on the type of targets. In addition, it should be noted that this paper adopts the continuous HMM model. This is because when HRRP is vectorized to obtain discrete signals, HRRP will be distorted, which reduces the recognition performance of the HMM.

In the HMM test phase, the sequence features obtained by the multi-input BLSTM are directly go into the trained HMM model, and the likelihood estimates in the output of each HMM model are obtained. If the *i*-th HMM yields the largest likelihood, then we declare the sequence features are associated with the *i*-th target type. For example, if we obtain the following aspect sequence

$o = \{o_1, o_2, o_3, \ldots, o_M\}$ from the unknown target type T, then the probability of the observation sequence o is given by summing the joint probability $P(o|q, T)P(q|T)$ over all possible state paths q.

$$P(o|T) = \sum_{all\ q} P(o|q, T)P(q|T) \tag{7}$$

If the target type T_i gives the maximum likelihood for the observation sequence o, i.e.,

$$P(o|T_i) \geq P(o|T_k), \ \forall T_k \tag{8}$$

Then we declare the sequence o belongs to the target type T_i.

Compared with the traditional HMM-based HRRP target recognition algorithm, this paper builds the aspect sequence of HRRP based on the fusion of multi-domain sequence features, so the HMM can fully learn the features of HRRP to provide the recognition outcome.

4. Benchmark MSTAR Dataset

Due to the lack of open HRRP real data, we use the benchmark real data of moving and stationary targets for acquisition and recognition (MSTAR) SAR images published by the US Department of Defense to test the proposed algorithm. MSTAR is the standard database for evaluating SAR image target recognition algorithms. According to the header information provided with the database, the radar frequency of the collected data is 9.599000 GHz. We know that the radar target's characteristic dimension is 1 m if we use the high frequency band of 30 MHz–300 MHz, and 0.1 m if using the high frequency band of 300 MHz–3 GHz, and about 0.03 m by using the high frequency band of 9.599000 GHz. Therefore, the HRRP signals generated from the MSTAR database contain sufficient information about vehicle targets such as tanks, artillery, or trucks, which can be used for target recognition research [44–48]. The MSTAR database includes 10 types of military vehicle targets: BMP2, BRT70, T72, BTR60, 2S1, BRDM2, D7, T62, ZIL131, and ZSU234. Among them, the BMP2 contains three different variants: BMP2_9563, BMP2_9566 and BMP2_c21; the T72 also contains three different variants: T72_132, T72_812 and T72_s7. Although these variants have the same design blueprint, they come from different manufacturers and still have some differences in color and shape. In addition to the military vehicle targets, MSTAR also includes a type of man-made target: SLICY. The SLICY is often used as the interference target to test the generalization ability of the proposed recognition algorithm. The optical images of these 11 types of targets are shown in Figure 6.

The MSTAR data is acquired by the X-band spotlight mode SAR. At a certain depression angle, the SAR circles the target several times, and collects images from the target multiple-aspect angles in the range of 360°. The aspect angle change intervals are not uniform. Even the same target is different at different depression angles, which will increase the recognition difficulty of the proposed algorithm. In addition, in the experiments, the observation data at 17° depression angles are generally used for training, while 15° are used for testing. Table 1 shows the acquisition of SAR images in the MSTAR database.

Figure 6. Images of the MSTAR targets.

Table 1. Target Types in MSTAR Dataset.

Target Types	No. of SAR Images		Pixel Size of SAR Image	No. of HRRPs Extracted per SAR Image
	15°	17°		
BMP2_9563	195	233	128*128	101
BMP2_9566	196	233	128*128	101
BMP2_c21	196	233	128*128	101
T72_132	196	232	128*128	101
T72_812	195	232	128*128	101
T72_s7	191	232	128*128	101
BTR70	196	233	128*128	101
BTR60	195	256	128*128	101
BRDM2	274	298	128*129	101
ZSU234	274	299	158*158	120
T62	273	299	172*173	136
ZIL131	274	299	192*193	152
2S1	274	299	158*158	120
D7	274	299	177*178	136
SLICY	274	298	54*54	48

4.1. Inversing HRRP from the SAR Image

Figure 7 shows the basic process of inversing HRRP from the SAR image, which comprises three steps:

Firstly, we remove the zero padding and the Taylor window. The MSTAR images are formed by taking a 2-D inverse fast Fourier transform (IFFT) of the Taylor-windowed and zero-padded phase data on a rectangular grid. The 2-D FFT is undertaken at first. Then the transformed signal is shifted so that the low frequencies occur in the center. Figure 7 shows the resulting 2-D signal of an example of the MSTAR SAR images. A noticeable band of near-zero values appears at the border of the 2-D signal, as shown in Figure 7. This band is the zero-padding result, which must be removed from the border of the signal. Next, we remove the Taylor window based on its parameters (35 dB sidelobe suppression level).

Secondly, we remove the clutter. Once zero padding and the Taylor window have been removed, a 2-D inverse FFT is applied to produce a de-convolved and Nyquist-sampled SAR image. Then the target segmentation procedure is undertaken to achieve the clutter removal.

Finally, we extract the HRRP. Before extracting the HRRP, the segmented SAR image requires zero padding to restore the original image size. Then, the zero-padding SAR image is transformed using the FFT, and the HRRP is extracted. Table 1 shows the number of HRRP that can be extracted from each SAR image for each type of target.
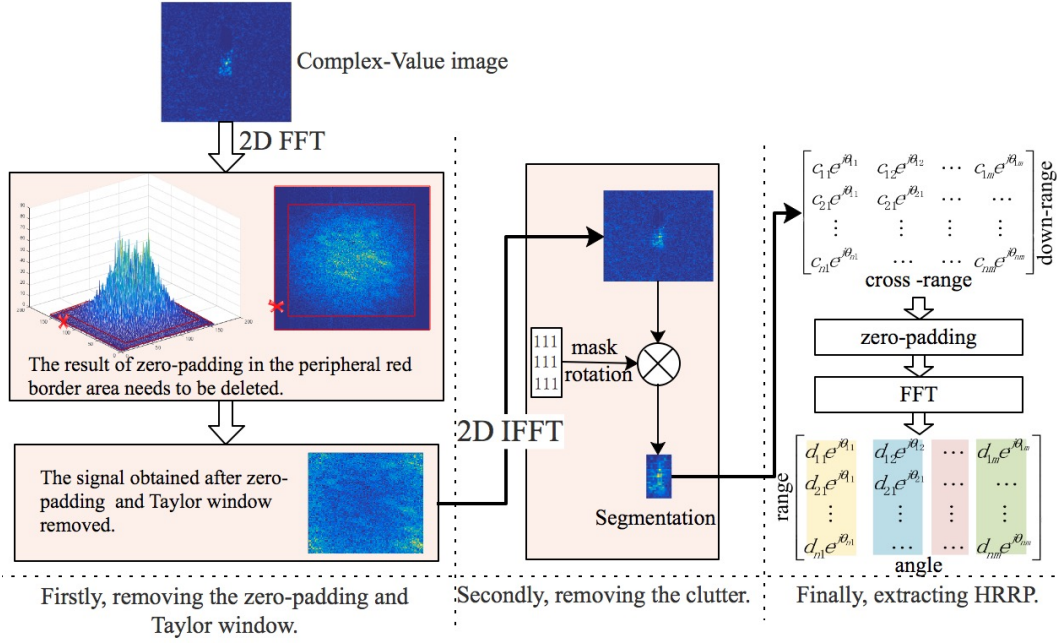

The result of zero-padding in the peripheral red border area needs to be deleted.

The signal obtained after zero-padding and Taylor window removed.

Firstly, removing the zero-padding and Taylor window.

Secondly, removing the clutter.

Finally, extracting HRRP.

Figure 7. The procedure from a SAR image to HRRPs.

In Table 1, the number of HRRPs extracted from different sized SAR images is different. However, since all SAR images have the same down-range and cross-range resolutions, i.e., $\Delta r_d = \Delta r_c = 0.305m$, the radar aperture angle represented by HRRP generated from each SAR image is equal, i.e.,

$$\Delta\theta = \frac{c}{2f\Delta r_c} \approx 3.0°$$

(9)

where $f = 9.6$ GHz, the center frequency of the radar waveform, and c is the speed of light.

4.2. Generating a Multi-Domain HRRP Sequence

According to the description in Section 3.1, the construction of the MSTAR multi-domain HRRP sequence goes through two steps.

First, the HRRP sequence is extracted. The HRRPs obtained from a slice of SAR image are taken as a set of HRRPs collected continuously by the radar over a period of time. Figure 8 shows the basic process of extracting a HRRP sequence from the BMP2 target. 101 HRRPs are divided into N groups. For ease of calculation, if 101 is not divisible by N, the remaining HRRP will be deleted. If $N = 4$, an HRRP sequence containing four elements is generated. In the training set, each element is the average result of 25 HRRPs, which contains $0.75°$ aperture information. In the test set, if not averaged, then each element is selected from one of the original 25 HRRPs, which contains $0.03°$ aperture information. Alternatively, an appropriate averaging number can be selected. For example, we select the amount of information for each element to be $0.3°$ aperture, i.e., 10 of the original 25 HRRPs for each element. This ensures that each element of the test set contains a certain amount of information to improve the recognition accuracy. Each element of the test set in this paper takes $0.3°$ aperture.
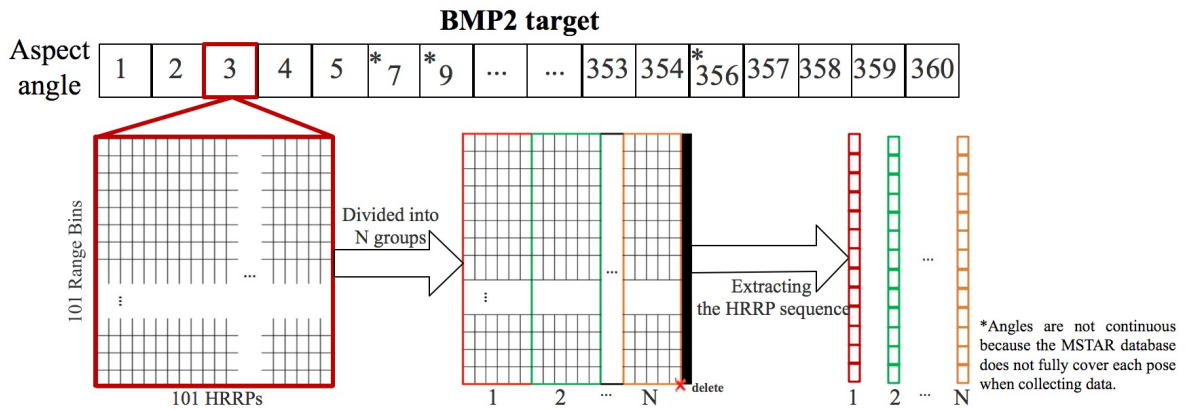
Figure 8. Extracting the HRRP sequence.

Secondly, we generate multi-domain HRPP data. The data domains in this paper include: time-, frequency-, and time-frequency (T-F)-domains. The amplitude in the time domain and the power spectrum in the time-frequency domain are used as target features for identification. For the T-F domain, we adopt AGR to extract T-F features of HRRP. In Figure 8, if $N = 4$, the 4-element multi-domain HRPP sequence in these three domains will be generated respectively, as shown in Figure 9. For each subfigure, the top-left part is the power spectrum in the frequency domain, the top-right part is the amplitude in the time domain, the middle part is the 3-D T-F feature generated from the power spectrum and the amplitude, and the middle-lower part shows 2-D T-F features.

Figure 9. Multi-domain HRPP sequence.

5. Experiments

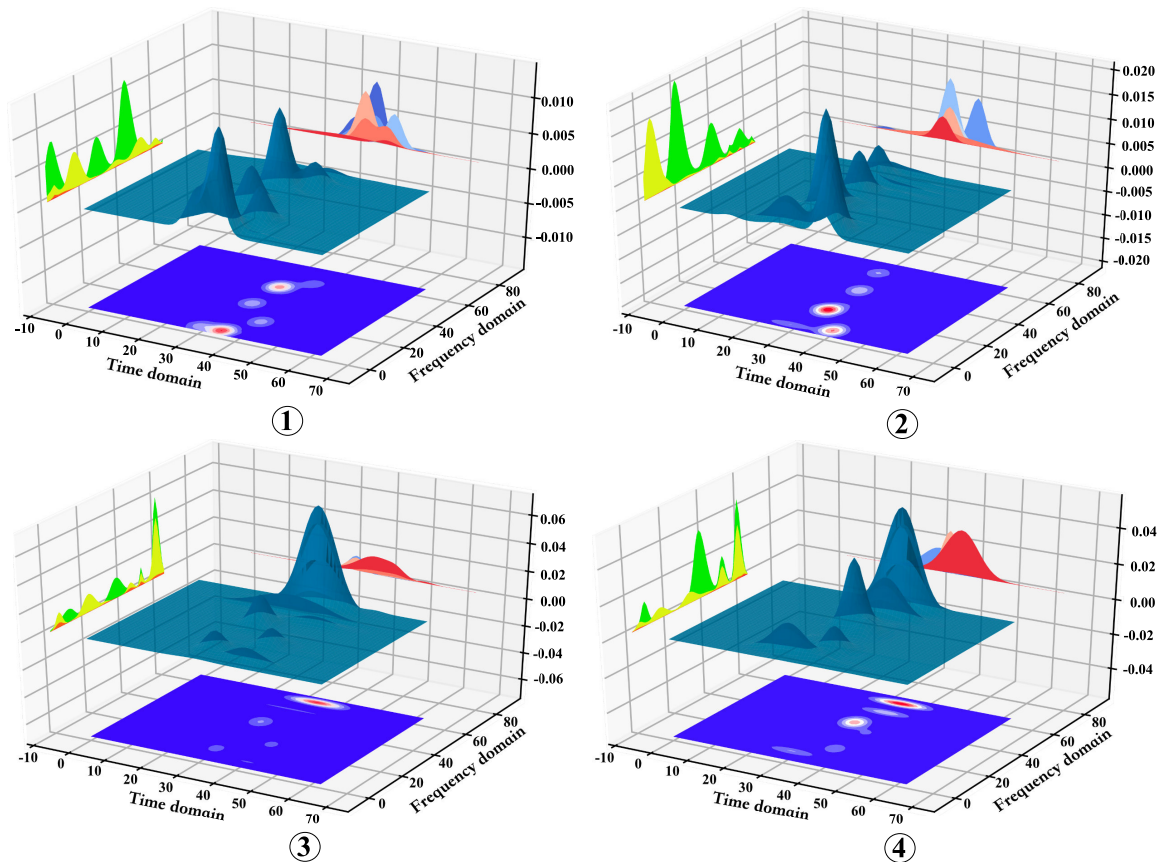To test the validity and generalization capability of the proposed algorithm, we designed a series of experiments under extended operating conditions (EOC) [49,50]. The so-called EOC is to test the algorithm according to various conditions in reality. These test conditions are different from those of the training algorithms. Specific EOC are set as follows:

(1) The test set is very different from the training set. Generally, the data of the 17° depression angle is selected as the training set, and the 15° depression angle is used as the test set. More stringently, one of the variants of BMP2 and T72 is used as the training set, while the other two variants are used as the test set.

(2) The aspect information of the target in the test set is assumed to be unknown and covers 360° aspect during the test. This hypothesis is meaningful in reality. First, it reduces dependence of the radar on other resources. For example, an auxiliary device such as a moving target indicator, a tracker, or the like is not required to acquire the azimuth information of the target. The azimuth coverage of the target is generally required to be 360°. However, this will increase the difficulty of the recognition algorithm, and will represent a significant challenge to our algorithm.

(3) The target in the test set is considered to be a non-cooperative target. This means that the target type that appears in the test set may not be present in the training set. This situation is inevitable in reality. Because of the incompleteness of the database, a new test sample may not find its similarity in the database. Therefore, the recognition algorithm needs evaluation criteria to determine whether the test sample belongs to the library (in-targets) or not (out-targets).

In this paper, three groups of experiments are set up under the above EOC: baseline experiments, validity verification experiments, and robustness evaluation experiments. These are described next.

5.1. Baseline Experiments

To verify the validity of the proposed multi-domain HRRP sequence features, we designed two sets of comparative experiments:

(1) B1 experiment: Comparing the validity of the HRRP random data and the sequence data in a single data domain;

(2) B2 experiment: Comparing the validity of the HRRP sequence data in the same data domain with different data domains.

5.1.1. B1 Experiment

To compare the validity of the HRRP random data and sequence data in a single data domain, we constructed two similar test models, namely single-domain random data model (SRDM) and single-domain sequential data model (SSDM). The design of these two models is shown in Figure 10. Figure 10a represents the SRDM model and Figure 10b shows the SSDM model.

	SRDM	T/F domain	T-F domain
1	Conv	32@3	32@3*3
2	Conv	32@3	32@5*5
3	F-C	9	
4	Conv	32@3	
5	Classifier	Softmax	

	SSDM	T/F domain	T-F domain
1	Conv	32@3	32@3*3
2	Conv	32@3	32@5*5
3	F-C	9	
4	BLSTM	32@(-1,N,64)	
5	Classifier	Softmax	

(**a**) SRDM model. (**b**) SSDM model.

Figure 10. Two similar single-domain HRRP test models.

In Figure 10a,b, the first column shows the number of layers of the model; the second column shows the name of the model and its structure; the third and fourth columns show which data domains are entered and configuration of the model in the current data domain. Overall, the two models have similar structures and configurations. They comprise five layers; the first four layers are for data preprocessing and the fifth layer is the classifier. *Conv*, *BLSTM*, and *F-C* represent a convolutional layer, a bidirectional LSTM layer, and a fully connected layer, respectively. Since HRRP is one-dimensional data in the time and frequency domain, but two-dimensional data in time-frequency domain, *Conv* has two configurations: 32@3 represents one-dimensional convolution of convolution kernel 3, which results in 32 features; 32@3 ∗ 3 represents a two-dimensional convolution of the convolution kernel 3 ∗ 3, and also results in 32 features, after this convolution operation has been implemented. In the *BLSTM* configuration, 32@(−1, *N*, 64) means that 32 features are obtained after the processing of this layer, −1 indicates that the number of the input sequence samples is not limited, *N* denotes that each sequence sample contains *N* elements, 64 implies each element is a 64-dimensional vector. In these two models, the configuration of the first three layers is identical, while the configuration of the fourth layer is different. In the fourth layer, the SRDM model uses a convolution layer, while the SSDM model uses the *BLSTM* layer. This is because, unlike the SRDM model, the SSDM model needs to extract sequence features. After the first four layers are set up for preprocessing, the two models eventually use the same classifier for recognition. The SRDM and SSDM models are designed in this way to ensure their experimental results are comparable. In the test results of the SRDM model, the validity of HRRP in time domain, frequency domain, and time-frequency domain can be compared under random conditions. In the test results of the SSDM model, the validity of HRRP in these domains can be compared under ordered conditions. In the test results of these two models, the effectiveness of HRRP in the same data domain under random and sequential conditions can be compared.

From Table 1, the training and test sets containing the "out-targets" are selected to test the SRDM and SSDM models. In the training set, we choose 9 types of targets; in the test set, because of the non-cooperative characteristics of the targets, we choose 2 types as the "out-targets" and the remaining as "in-targets". These target types are shown in Table 2. In Table 2, the targets except "Others" in the first line, are the targets of the training set, i.e., 9 types of "in-targets". In addition, the targets in the first column are the targets of the test set, in total 10 types, including 2 types of "out-targets". In addition, to better demonstrate the validity of the data, the HRRP data of these targets are respectively generated into four sets of data to train and test the two models, according to the description in Section 4. In these four sets of data, each sequence of the training set contains 5, 4, 2, and 1 elements, which correspond to the aperture angles of 0.6°, 0.75°, 1.5° and 3° respectively. The aperture angle of each element in the test set is 0.3°. The order of the HRRP sequence is ignored in the SRDM model. Under such conditions, the experimental results of the SRDM and SSDM models are shown in Figure 11a,b, respectively. The horizontal axis represents the aperture angle of the training set corresponding to different experiments, and the vertical axis represents the recognition rate of the model. ◆ , □ and ○ represent the recognition rate curves in time, frequency, and time-frequency domain, respectively.

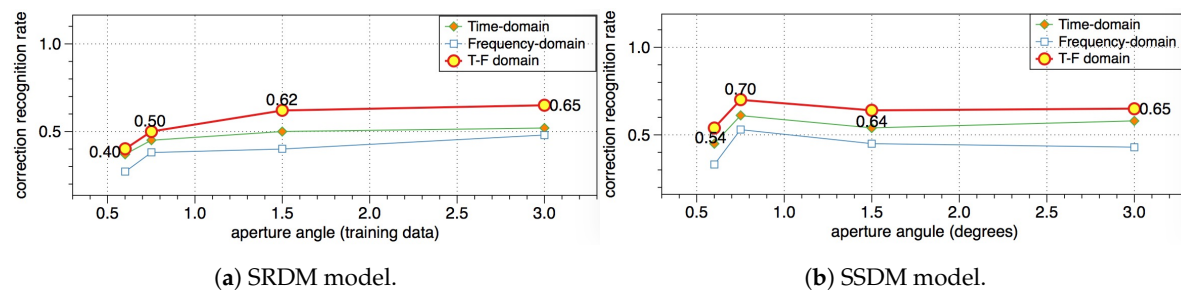**(a)** SRDM model.   **(b)** SSDM model.

Figure 11. Test results of four sets of single-domain HRRP data in SRDM and SSDM models respectively.

Table 2. Test results of the SSDM model trained in time-frequency domain with 0.75° aperture.

Target Types	BMP2_c21	T72_132	BTR70	SLICY	BTR60	BRDM2	ZSU234	T62	ZIL131	Others
BMP2_c21	**138**	5	3	10	4	9	2	7	2	16
T72_132	0	**133**	8	3	1	3	10	5	8	25
BTR70	11	7	**115**	10	12	6	8	3	2	22
BTR60	7	9	2	11	**122**	5	8	2	5	24
BRDM2	14	7	14	12	16	**150**	12	17	7	25
ZSU234	3	4	9	8	12	4	**188**	11	12	23
T62	8	10	5	6	10	11	14	**174**	12	23
ZIL131	7	8	6	9	15	11	10	12	**174**	22

The number of the "in-targets" HRRP sequences in the test set = 1878.
out-targets

2S1	27	32	8	8	9	13	24	31	27	**95**
D7	12	20	10	19	12	32	21	17	18	**113**

The number of the "out-targets" HRRP sequences in the test set = 548.
Correct recognition rate of the " in-targets", P_{cid} = 1194/1698 = 0.70;
Detection rate of the "in-targets" , P_d = 1698/1878 = 0.90;
False alarm rate of the "out-targets", FAR = 340/548 = 0.62.

In terms of the data domain, Figure 11a,b show that the feature recognition results in the time-frequency domain are the best, followed by the time-domain features, and finally the frequency-domain features. The recognition rate of the sequence features shown in Figure 11b is higher than that of the corresponding random features shown in Figure 11a. This shows that the HRRP sequence proposed in this paper is effective, which can improve the recognition rate of the targets. Figure 11a shows that the target recognition rate increases with the increase of aperture angles. When the aperture angle is 3.0°, i.e., all the HRRPs extracted from each SAR image in the training set are processed by the multi-look scheme at the sample layer to obtain one HRRP, where the target recognition rate is the highest. This is consistent with other conclusions reported in the literature [51]. Unlike Figure 11a, the target recognition rates shown in Figure 11b do not increase with the increase in aperture angle. This shows that the improved target recognition rate is not caused by the increase in aperture angles, and is, in fact, due to the HRRP sequence proposed here.

In the above comparative experiments, the SSDM model trained with the HRRP sequence in the time-frequency domain with 0.75° aperture produces the highest recognition rate. Table 2 shows the recognition results in the form of a confusion matrix. Since there are the "out-targets" in the test set, we need to use reasonable evaluation criteria.

The introduction of the "Others" type in the first line of Table 2 is the strategy adopted in this paper. Given a test sample, we first determine whether or not it belongs to the "in-targets". If so, we continue to determine which types it belongs to; if not, it is considered to belong to the "out-targets" and we classify it as the "Others" type. Ideally, the "in-targets" in the test set are correctly identified and the "out-targets" are correctly classified as the "Others". However, in the actual test, the "in-targets" may be incorrectly recognized as the "Others" type, and the "out-targets" may be incorrectly recognized as the "in-targets". Therefore, we need three evaluation criteria to evaluate the experimental results. The three criteria are defined as:

(1) Correct recognition rate of "in-targets":

$$P_{cid} = \frac{N_{cc}}{N_t - N_o} \tag{10}$$

where N_t is the total number of "in-targets" in the test set, N_o is the total number of "in-targets" identified as "Others" type in the test set, and N_{cc} is the total number of "in-targets" correctly identified in the test set.

(2) Detection rate of "in-targets":

$$P_d = \frac{N_t - N_o}{N_t} \tag{11}$$

(3) False alarm rate of "out-targets":

$$FAR = \frac{N_{\text{int}}}{N_{nt}} \tag{12}$$

where N_{int} is the total number of "out-targets" identified as "in-targets" in the test set, and N_{nt} is the total number of "out-targets" in the test set.

In Table 2, we set P_d to 0.9, which means that 10% of "in-targets" in the test set are judged as "out-targets" when they are less than a certain threshold. In this way, on the one hand, the reliability of identification of "in-targets" in the test set can be improved; and on the other hand, a reference threshold can be found to judge the "in-targets" as "out-targets" for subsequent experiments. Under such conditions, $P_{cid} = 0.70$, $FAR = 0.62$ are obtained in Table 2, which is still a long way from the ideal results ($P_{cid} = 1$, $FAR = 0$). Therefore, in the next set of experiments, we explore the fusion of multi-domain HRRP sequence features to further improve the target recognition rate.

5.1.2. B2 Experiment

In Section 5.1.1, both the SRDM and SSDM models have only one input, so these two models belong to the single-look processing type. The experimental results shown in Section 5.1.1 show that the recognition rates of these two models are not satisfactory. Compared to the single-look processing model, the multi-look processing model uses multiple inputs to extract and fuse multi-view features of data, which can improve recognition performance. Normally, the input data of the multi-look processing model are from the same data domain. Here, we will build a multi-input BLSTM model, and then, on the one hand, we use the same data domain HRRP sequence data for testing; and on the other hand, we use different data domains' HRRP sequence data for testing. The purpose is to compare the validity of sequence features in the same data domain and in different data domains.

Figure 12 shows the multi-input BLSTM model built in this section. The middle column represent the structure of this model, and the left and right sides are the corresponding configurations under different input data. The model comprises six layers, where the first five layers represent data preprocessing, and the sixth layer is the classifier, with use of *softmax*. Therefore, we use MIBL-softmax to represent the model. The MIBL-softmax model has three input interfaces: one main input interface and two branch input interfaces. The main input interface is at Layer 1, the first branch input interface is at Layer 4, and the second branch input interface is at Layer 5. When the input data is from the same data domain HRRP sequence, these data do not need to distinguish the importance level, and directly feed in the input interface of the MIBL-softmax model. When the input data is from different data domains, we send them to the primary and secondary input interfaces of the MIBL-softmax model, from large to small, according to the importance level.
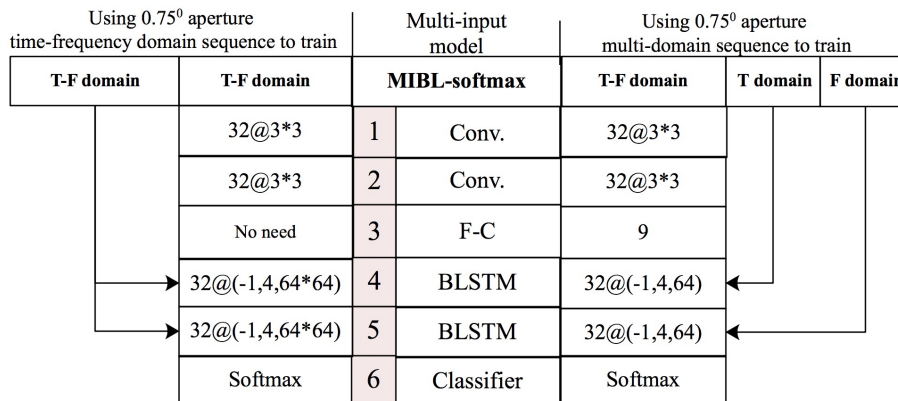
Using 0.75⁰ aperture time-frequency domain sequence to train		Multi-input model		Using 0.75⁰ aperture multi-domain sequence to train		
T-F domain	**T-F domain**	**MIBL-softmax**		**T-F domain**	**T domain**	**F domain**
	32@3*3	1	Conv.	32@3*3		
	32@3*3	2	Conv.	32@3*3		
	No need	3	F-C	9		
	32@(-1,4,64*64)	4	BLSTM	32@(-1,4,64)		
	32@(-1,4,64*64)	5	BLSTM	32@(-1,4,64)		
	Softmax	6	Classifier	Softmax		

Figure 12. MIBL-softmax model.

Looking at the experimental results shown in Section 5.1.1, when the HRRP sequence is at 0.75° aperture, the time-frequency domain features are the best, followed by the time domain features, and finally the frequency domain features. Therefore, the HRRP sequence of 0.75° aperture in the time-frequency domain is used and sent to three input interfaces of the MIBL-softmax model. On the other hand, we use the time-frequency domain, time domain, and frequency domain HRRP sequence of 0.75° aperture as the data of the different data domains, and use the main input interface, the first branch input interface and the second branch input interface of the MIBL-softmax model respectively. The test results of these two sets of data in the MIBL-softmax model are represented by the confusion matrix as shown in Tables 3 and 4.

Table 3. Test results of the same data domain HRRP sequence in the MIBL-softmax model.

Target Types	BMP2_c21	T72_132	BTR70	SLICY	BTR60	BRDM2	ZSU234	T62	ZIL131	Others
BMP2_c21	**155**	3	1	3	2	4	3	2	2	21
T72_132	4	**148**	2	8	2	3	4	3	2	20
BTR70	5	0	**155**	7	4	4	2	1	0	18
BTR60	2	4	3	6	**150**	3	5	0	6	16
BRDM2	7	8	7	6	10	**182**	5	12	6	31
ZSU234	7	4	6	4	5	6	**210**	4	6	22
T62	5	6	3	7	8	3	6	**211**	4	20
ZIL131	0	5	6	9	7	9	10	15	**188**	25

The number of "in-targets" HRRP sequences in the test set = 1878.
out-targets

2S1	16	15	20	43	15	14	15	11	14	**111**
D7	13	11	21	48	17	13	14	11	10	**116**

The number of the "out-targets" HRRP sequences in the test set = 548.
Correct recognition rate of the " in-targets", P_{cid} = 1399/1705 = 0.8205;
Detection rate of the "in-targets", P_d = 1705/1878 = 0.9078;
False alarm rate of the "out-targets", FAR = 321/548 = 0.5857.

Table 3 shows test results of the same data domain HRRP sequence in the MIBL-softmax model. The correct recognition rate P_{cid} of the MIBL-softmax model is 0.8205, which is significantly higher than P_{cid} = 0.70 shown in Table 2. Under the same threshold as shown in Section 5.1.1, the detection rate P_d of the MIBL-softmax model reaches 0.9078, which is higher than P_d = 0.900 shown in Table 2. Regarding the "out-targets", the FAR of the MIBL-softmax model is reduced to 0.5857, which is significantly lower than FAR = 0.62 shown in Table 2. This shows that, compared to the single-look processing model, the multi-look processing model has a better confidence level and the recognition rate for "in-targets", and a better distinction between "out-targets" and "in-targets", for the traditional case of using the same data domain HRRP sequence as input. This is because the multi-look processing model extracts the features of the HRRP sequence several times and continuously fuses them to obtain the best discriminative features. This in turn, improves the recognition performance of the targets

to some extent. However, due to the limitation of the HRRP sequence in the same data domain, the recognition performance of the MIBL-softmax model is difficult to be further improved.

Table 4. Test results of different data domain HRRP sequence in the MIBL-softmax model.

Target Types	BMP2_c21	T72_132	BTR70	SLICY	BTR60	BRDM2	ZSU234	T62	ZIL131	Others
BMP2_c21	**162**	1	0	4	1	3	3	2	1	19
T72_132	6	**152**	2	3	1	2	5	2	5	18
BTR70	1	2	**148**	3	6	2	5	7	1	21
BTR60	2	1	0	10	**160**	1	2	1	2	16
BRDM2	3	1	5	8	3	**220**	2	7	3	22
ZSU234	9	0	2	5	7	5	**213**	2	7	24
T62	0	2	0	2	1	1	4	**234**	5	24
ZIL131	0	0	7	3	2	0	7	8	**227**	20
The number of "in-targets" HRRP sequences in the test set = 1878.										
out-targets										
2S1	7	11	14	34	5	11	11	10	7	**164**
D7	12	10	11	27	12	20	7	17	15	**143**

The number of the "out-targets" HRRP sequences in the test set = 548.
Correct recognition rate of the "in-targets", P_{cid} = 1516/1714 = 0.8844;
Detection rate of the "in-targets", P_d = 1714/1878 = 0.9126;
False alarm rate of the "out-targets", FAR = 241/548 = 0.4397.

Unlike Table 3, Table 4 shows the test results of different data domain HRRP sequence in the MIBL-softmax model. Compared to Table 3, the three evaluation criteria shown in Table 4 have been significantly improved. The correct recognition rate P_{cid} of the "in-targets" increases from 0.8205 to 0.8844; the detection rate P_d of the "in-targets" increases from 0.9078 to 0.9126; and the false alarm rate FAR of the "out-targets" decreases from 0.5857 to 0.4397. This indicates that in the same multi-look processing model, the different data domain HRRP sequence is more favorable for target recognition than the same data domain HRRP sequence. Specifically, the data with the highest importance level, i.e., the main input data, can be better extracted through the whole network; secondly, the data with the highest importance level are seen to be the non-linear transformation of input data from each branch, i.e., the time-frequency domain data is the non-linear transformation of the time-domain data and the frequency-domain data. When the multi-look processing model fuses the input features, then, firstly, the data of each branch can provide similar features to the system, which can enhance these features in the fusion process, and thus improve the target recognition rate. Secondly, they can also add the features of different data domains in the process of non-linear transformation, which makes the multi-look processing model exhibit better generalization ability and lower false alarm rate. Therefore, in the process of multi-look sequence recognition, the HRRP sequence features in different data domains are more effective than those in the same data domain.

5.2. Validity Verification Experiments

In the baseline experiment of Section 5.1, we validated the validity of the multi-domain HRRP sequence features. In this section, we will verify the validity of the proposed MIBL-HMM algorithm for multi-domain HRRP sequence feature learning. Here, we continue to use the 0.75° aperture multi-domain HRPP sequence shown in Table 4 as the training data, and set up two sets of comparative experiments:

(1) V1 experiment: the comparative experiment between the MIBL-HMM algorithm and the MIBL-softmax algorithm. For the case of the multi-input BLSTM algorithm, the MIBL-HMM algorithm uses the HMM as the classifier, while the MIBL-softmax algorithm uses the *softmax* as the classifier. To verify the effectiveness of the MIBL-HMM in further learning multi-aspect features contained in the multi-domain HRRP sequences, the MIBL-HMM and the MIBL-softmax are compared under the same design and configuration, except for the different classifiers.

(2) V2 experiment: comparisons between MIBL-HMM and MIL-HMM algorithms. The MIL-HMM algorithm is an algorithm in which the multi-input LSTM algorithm uses the HMM as a classifier. To validate the effectiveness of the MIBL-HMM in learning multi-domain HRPP sequence features, the MIBL-HMM and the MIL-HMM are compared under the same design and configuration, except for the different numbers of the directions of LSTM.

Before starting the comparison, we need to train the HMM models. The algorithms reviewed in Section 3.4 are first used to construct the HMM models. These are then trained using the forward and reverse HRRP aspect sequences. The forward sequence is the aspect angle of the HRRP sequence sorted from small to large, i.e., $0° \rightarrow 1° \rightarrow... \rightarrow 359° \rightarrow 360°$, while the reverse sequence is arranged in the opposite direction. These sequences contain all HRRP aspect information and state statistics, and are able train HMM models quickly. The training set used in Section 5.1 contains 9 types of targets. Therefore, 9 HMM models need to be trained, with each target type corresponding to one HMM model. In the training process, the Baum-Welch algorithm is used to estimate the parameter *A* and the state density function of the HMM, so that the HMM can better reflect the scattering characteristics of the targets.

5.2.1. V1 Experiment

In the test phase of the MIBL-HMM algorithm, the multi-domain fusion sequence features of the unknown targets are fed into the 9 trained HMM models, and the corresponding likelihood values are generated, respectively. If the maximum of these likelihood values is less than the threshold set shown in Section 5.1.1, the unknown target is identified as an "out-targets". Otherwise, it is identified as "in-targets", and its specific type is the target type represented by the HMM model which obtains the maximum likelihood value. Table 5 shows the recognition results of the MIBL-HMM algorithm. The correct recognition rate of the "in-targets" is seen to be $P_{cid} = 0.9132$, the detection rate of the "in-targets" is $P_d = 0.9211$, and the false alarm rate of the "out-targets" is $FAR = 0.3467$. For the same test set, compared with the test results of the MIBL-softmax algorithm ($P_{cid} = 0.8844$, $P_d = 0.9126$, $FAR = 0.4397$, Table 4), the MIBL-HMM algorithm test results can be seen to have improved. The results show that the proposed method can further learn the multi-aspect features contained in the HRRP multi-domain sequence features to improve the recognition performance.

Table 5. Test results of the MIBL-HMM algorithm trained with 0.75° aperture multi-domain HRRP sequence.

Target Types	BMP2_c21	T72_132	BTR70	SLICY	BTR60	BRDM2	ZSU234	T62	ZIL131	Others
BMP2_c21	**155**	4	3	3	0	2	5	6	3	15
T72_132	1	**161**	3	2	4	2	1	2	0	20
BTR70	0	2	**170**	1	2	0	0	1	1	19
BTR60	1	2	0	0	**172**	2	1	0	0	17
BRDM2	0	3	2	6	3	**232**	3	5	3	17
ZSU234	2	4	2	3	5	1	**230**	5	3	19
T62	1	2	4	3	5	2	8	**213**	13	22
ZIL131	0	0	0	0	0	0	0	8	**247**	19

The number of "in-targets" HRRP sequences in the test set = 1878.
out-targets

2S1	5	8	11	5	5	6	8	10	11	**205**
D7	5	10	6	4	7	2	13	8	15	**204**

The number of "out-targets" HRRP sequences in the test set = 548.
Correct recognition rate of the "in-targets", $P_{cid} = 1580/1730 = 0.9132$;
Detection rate of the "in-targets", $P_d = 1730/1878 = 0.9211$;
False alarm rate of the "out-targets", $FAR = 190/548 = 0.3467$.

It should be noted that the HMM model is sensitive to the number of hidden states and the weight of the Gaussian probability density function. Changing these configurations will lead to additional complexity in the HMM model for sequential data processing. Figure 13 shows the recognition rate

of the HMM model using the MIBL-HMM algorithm under different hidden state numbers S and weights w of the Gaussian probability density function. When $S = 60$ and $w = 2$, the MIBL-HMM algorithm produces the best recognition outcome. The HMM model in the above experiments is set up in the same way. Therefore, in the following experiments, the parameters of the HMM model are set to $S = 60$, $w = 2$.
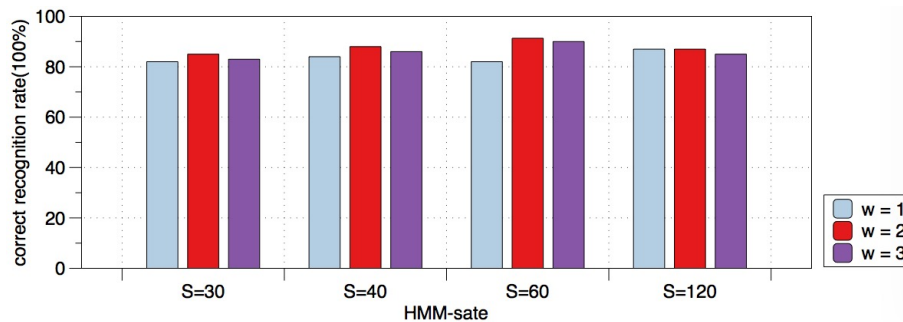

Figure 13. Different configuration of HMM model in MIBL-HMM algorithm obtains different recognition rates.

5.2.2. V2 Experiment

Similar to the testing of the MIBL-HMM algorithm, Table 6 shows the recognition results of the MIL-HMM algorithm. The correct recognition rate of the "in-targets" is seen to be $P_{cid} = 0.8501$, and the detection rate of the "in-targets" is $P_d = 0.9132$. The false alarm rate of the "out-targets" is $FAR = 0.4525$. Under the same test set, compared with the test results of the MIBL-HMM algorithm ($P_{cid} = 0.9132$, $P_d = 0.9211$, $FAR = 0.3467$, Table 5), the MIBL-HMM algorithm test results are significantly higher than the MIL-HMM algorithm. The results show that the proposed method can further learn the context features contained in the HRRP multi-domain sequence features to improve the recognition performance.

Table 6. Test results of the MIL-HMM algorithm trained with 0.75° aperture multi-domain HRRP sequence.

Target Types	BMP2_c21	T72_132	BTR70	SLICY	BTR60	BRDM2	ZSU234	T62	ZIL131	Others
BMP2_c21	**115**	8	2	10	11	9	5	6	8	22
T72_132	4	**121**	0	16	3	5	10	5	11	21
BTR70	0	0	**172**	1	0	3	3	0	1	16
BTR60	0	1	2	0	**167**	0	1	2	0	22
BRDM2	5	6	7	16	5	**191**	5	7	2	30
ZSU234	0	7	3	4	5	6	**193**	14	10	32
T62	2	0	1	3	0	1	0	**252**	2	12
ZIL131	2	2	3	6	4	0	2	0	**247**	8

The number of "in-targets" HRRP sequences in the test set = 1878.
out-targets

| 2S1 | 6 | 9 | 37 | 11 | 11 | 15 | 10 | 12 | 9 | **154** |
| D7 | 15 | 12 | 40 | 5 | 9 | 12 | 11 | 4 | 20 | **146** |

The number of the "out-targets" HRRP sequences in the test set = 548.
Correct recognition rate of the " in-targets", $P_{cid} = 1458/1715 = 0.8501$;
Detection rate of the "in-targets", $P_d = 1715/1878 = 0.9132$;
False alarm rate of the "out-targets", $FAR = 248/548 = 0.4525$.

In addition, we compare the proposed algorithm with several state-of-the-art methods. In general, the recognition performance of the template matching method is used as the baseline for HRRP target recognition [52]. As can be seen from Table 5, the correct recognition rate of our algorithm is 91.32%, which is better than 90% of the template matching method presented in [52]. Moreover, unlike the template matching method, our algorithm does not require aspect information of the target.

The algorithm proposed in this paper is also compared with the multi-look processing methods mentioned in Section 1. [20] used a multi-look processing method to identify the same 9 types of

targets as those used in this paper, and obtained 97% recognition accuracy under the 3° aperture test set. In the HRRP sparse feature-based target recognition, [16] also used the 3° aperture test set, and obtained 92% recognition accuracy. In time-frequency domain feature recognition, Zhang et al.[18] used the 6° aperture test set, and obtained 95.62% recognition accuracy. Although the recognition accuracy of these algorithms is relatively high, they are all obtained at the expense of computational time of the algorithms. To compare with these algorithms, our proposed algorithm is also tested on a 3° aperture test set, and the recognition accuracy is found to be 96.2%, as shown in Table 7. This recognition accuracy is higher than the results reported in literature, which demonstrates the superiority of this algorithm.

Table 7. Test results of MIBL-HMM testing with 3° aperture multi-domain HRRP sequence.

Target Types	BMP2_c21	T72_132	BTR70	SLICY	BTR60	BRDM2	ZSU234	T62	ZIL131	Others
BMP2_c21	**165**	0	1	1	0	2	5	0	3	19
T72_132	0	**174**	3	2	0	2	1	2	0	12
BTR70	0	2	**175**	1	2	0	0	1	1	14
BTR60	1	2	0	0	**172**	2	1	2	0	15
BRDM2	0	0	2	0	3	**252**	0	1	3	13
ZSU234	2	0	2	3	0	1	**250**	0	3	13
T62	1	2	0	1	0	2	1	**247**	3	16
ZIL131	0	0	0	0	0	0	0	0	**259**	15

Total number of test input sequences = 1878
confuser-targets(confusers)

2S1	5	8	11	5	5	6	8	10	11	**205**
D7	5	10	6	4	7	2	13	8	15	**204**

Total number of test input sequences = 548
Overall correct identification rate, P_{cid} = 1694/1761 = 0.962;
Overall "in-target" detection rate, P_d = 1761/1878 = 0.939;
Overall false alarm rate, FAR = 139/548 = 0.2536

5.3. Robustness Evaluation Experiments

To verify the generalization ability of the proposed algorithm in multi-domain HRRP sequence feature recognition, we designed the following two sets of comparative experiments, with rigorous experimental conditions for robustness evaluation.

(1) R1 experiment: One of the variant types of BMP2 and T72 targets is used as the training set, while the other three variants are used as test sets. Under such conditions, the robustness of the recognition algorithms is evaluated and compared.

(2) R2 experiment: The training set is the same as the training set of the R1 experiment, while the targets of the test set are set to "out-targets". Under such conditions, the robustness of the recognition algorithms is evaluated and compared.

In the experiments reported in Section 5.2, although our MIBL-HMM algorithm outperforms the MIBL-softmax algorithm on three evaluation criteria, the comparative recognition accuracy has little difference. We will further compare the two algorithms in robustness evaluation experiments described next.

5.3.1. R1 Experiment

The 0.75° aperture HRRP multi-domain sequences are used to train the MIBL-softmax and MIBL-HMM algorithms respectively. Comparative recognition results are shown in Tables 8 and 9, respectively.

Table 8. Test results of the MIBL-softmax algorithm under the condition that test set targets are "in-targets".

Target Types	BMP2_c21	BTR70	T72 _132	Others
BMP2_9563	**175**	2	3	15
BMP2_9566	**182**	2	0	12
BMP2_c21	**185**	0	0	11
BTR70	0	**180**	0	16
T72_132	0	0	186	10
T72_812	2	5	**170**	18
T72_s7	1	0	**173**	17

The number of "in-targets" HRRP sequences in the test set = 1365.
Correct recognition rate of the "in-targets", P_{cid} = 1251/1266 = 0.9881;
Detection rate of the "in-targets", P_d = 1266/1365 = 0.9274.

Table 9. Test results of the MIBL-HMM algorithm under the condition that the test set targets are "in-targets".

Target Types	BMP2_c21	BTR70	T72_132	Others
BMP2_9563	**185**	1	2	7
BMP2_9566	**186**	0	1	9
BMP2_c21	**190**	0	0	6
BTR70	0	**185**	0	11
T72_132	2	2	**184**	8
T72_812	4	2	**180**	9
T72_s7	1	0	**185**	5

The number of the "in-targets" HRRP sequences in the test set = 1365.
Correct recognition rate of the "in-targets", P_{cid} = 1295/1310 = 0.9885;
Detection rate of the "in-targets", P_d = 1310/1365 = 0.9597.

Under the same threshold conditions, as can be seen from Tables 8 and 9, the correct recognition rate P_{cid} of MIBL-softmax and MIBL-HMM algorithms for the "in-targets" has reached 98%. This shows that both the MIBL-softmax and the MIBL-HMM algorithms can learn the common features of targets from 0.75° aperture multi-domain HRRP sequence to accurately identify the targets. It can also be seen by comparing the "in-targets" detection rate P_d, the MIBL-HMM algorithm obtains an accuracy of 0.9597, while the MIBL-softmax algorithm achieves an accuracy of 0.9274. The detection rate of the MIBL-HMM algorithm for "in-targets" is thus significantly higher than that of the MIBL-softmax algorithm. This shows that the MIBL-HMM algorithm is more robust than the MIBL-softmax algorithm based on the same multi-domain HRRP sequence. This is attributed to the ability of the MIBL-HMM algorithm to further learn the multi-aspect features contained in the multi-domain HRRP sequence, to enhance the generalization ability of target recognition.

5.3.2. R2 Experiment

The training set of the R2 experiment is the same as that of the R1 experiment, while the targets of the test set are the seven other types of targets shown in Table 1. The latter are also used as "out-targets" in the R2 experiment. Further, the MIBL-softmax and the MIBL-HMM algorithms are still trained using the 0.75° aperture multi-domain HRRP sequence. Comparative recognition results are shown in Tables 10 and 11, respectively.

Tables 10 and 11 show both, the total false alarm rate of "out-targets" in the whole test set, and the false alarm rate of each type of target. Overall, the total false alarm rate of the MIBL-softmax algorithm is 0.288, while that of the MIBL-HMM algorithm is approximately 10% less, at 0.191. This shows that the MIBL-HMM algorithm performs better in rejecting the "out-targets" test samples. In addition, it can also be seen that the MIBL-HMM algorithm can enhance the generalization capability of target recognition by more deeply learning the multi-aspect features contained in the multi-domain HRRP sequence. The MIBL-HMM algorithm more effectively rejects the "out-targets" compared to the

MIBL-softmax algorithm. In Table 11, the MIBL-HMM algorithm achieves an ideal false alarm rate ($FAR < 0.100$) on the exclusion of three types of "out-targets": D7, T62, and ZIL131, where D7 is a bulldozer, T62 is a main battle tank, and ZIL131 is a freight truck. In addition, it can be seen in Table 11, that the MIBL-HMM algorithm does not demonstrate much improvement in detecting false alarm rate of the BTR60 target, which can be attributed to the test samples of BTR60 comprising abnormal samples [53]. Therefore, it is reasonable to expect that the MIBL-HMM algorithm will not reduce the false alarm rate detection of BTR60.

Table 10. Test results of the MIBL-softmax algorithm under the condition that test set targets are "out-targets".

Target Types	BMP2_c21	BTR70	T72_132	Others	FAR
BRDM2	14	90	10	**160**	114/274 = 0.416
BTR60	10	150	5	**30**	165/195 = 0.846
D7	13	19	11	**231**	43/274 = 0.157
T62	14	15	18	**226**	47/273 = 0.172
2S1	21	12	20	**221**	53/274 = 0.193
ZIL131	16	17	28	**213**	61/274 = 0.222
ZSU234	13	18	15	**228**	46/274 = 0.168

The number of "out-targets" HRRP sequences in the test set = 1838.
Overall false alarm rate of the "out-targets", FAR = 529/1838 = 0.288.

Table 11. Test results of the MIBL-HMM algorithm under the condition that test set targets are "out-targets".

Target Types	BMP2_c21	BTR70	T72_132	Others	FAR
BRDM2	12	50	8	**204**	70/274 = 0.255
BTR60	7	145	5	**38**	157/195 = 0.805
D7	7	5	11	**251**	23/274 = 0.084
T62	8	4	6	**255**	18/273 = 0.066
2S1	11	12	10	**241**	33/274 = 0.120
ZIL131	4	7	8	**255**	19/274 = 0.069
ZSU234	7	8	15	**244**	30/274 = 0.109

The number of the "out-targets" HRRP sequences in the test set = 1838.
Overall false alarm rate of the "out-targets", FAR = 350/1838 = 0.191.

In summary, we extract the sequence of HRRP and on this basis, generate multi-domain sequences which are effective for target recognition. At the same time, the results show our proposed algorithm can deeply learn and more effectively fuse features of HRRP sequences in multiple data domains.

6. Conclusions

The construction of an optimal classifier based on limited HRRP signals to correctly extract, learn, and fuse HRRP features, in order to improve the target recognition ability is an open research problem. In this paper, we propose to learn features contained in the HRRP data, by exploiting the acquisition process of HRRP to extract a HRRP sequence, and generate multi-domain sequence features in time-, frequency- and time-frequency domains. A novel algorithm for HRRP target recognition based on BLSTM and HMM is proposed, which, on the one hand, extracts and fuses the multi-domain HRRP sequence features effectively, using a multi-input approach with primary and secondary branches. On the other hand, our proposed algorithm improves recognition performance by combining multi-aspect sequence features contained in multi-domain HRRP sequences, with the standard HMM model.

In benchmark recognition tasks using the MSTAR database with 10 non-cooperative targets, our proposed algorithm achieves 91% correct recognition rate. Compared with other state-of-the-art methods, our approach exhibits enhanced recognition performance, lower false alarm rate, and higher

confidence. In addition, our proposed algorithm is simple to implement and, in the future, can be explored for implementing real-time human–machine interaction designs.

Abbreviations

The following abbreviations are used in this manuscript:

AGR	adaptive Gaussian representation
B1 experiment	the 1st baseline experiment
B2 experiment	the 2nd baseline experiment
BLSTM	bidirectional long short-term memory
CNN	convolutional neural network
Conv	convolutional layer
F domain	frequency domain
F-C	fully connected layer
FAR	False alarm rate
HMM	hidden Markov model
HRRP	high-resolution range profile
LSTM	long short-term memory
MIBL-HMM	the algorithm based on the multi-input BLSTM and HMM
MIBL-softmax	the algorithm based on the multi-input BLSTM and softmax
MIL-HMM	the algorithm based on the multi-input LSTM and HMM
MSTAR	moving and stationary target acquisition and recognition database
RBMs	restricted Boltzmann machines
R1 experiment	the 1st robustness evaluation experiment
R2 experiment	the 2nd robustness evaluation experiment
SAE	stacked autoencoders
SAR	synthetic aperture radar
SRDM	single-domain random data model
SSDM	single-domain sequential data model
STFT	short-time Fourier transform
T domain	time domain
T-F domain	time-frequency domain
V1 experiment	the 1st validity verification experiment
V2 experiment	the 2nd validity verification experiment
WVD	Wigner-Ville distribution

References

1. Golovachev, Y.; Etinger, A.; Pinhasi, G.; Pinhasi, Y. Millimeter Wave High Resolution Radar Accuracy in Fog Conditions—Theory and Experimental Verification. *Sensors* **2018**, *18*, 2148. [CrossRef]
2. Baker, C.J.; Smith, G.E.; Balleri, A.; Holderied, M.; Griffiths, H.D. Biomimetic Echolocation With Application to Radar and Sonar Sensing. *Proc. IEEE* **2014**, *102*, 447–458. [CrossRef]

3.   Gruber, A.; Gadringer, M.; Schreiber, H.; Amschl, D.; Bosch, W.; Metzner, S.; Pflugl, H. Highly scalable radar target simulator for autonomous driving test beds. In Proceedings of the Radar Conference, Oklahoma City, OK, USA, 23–27 April 2018; pp. 147–150.

4.   Tan, Q.J.O.; Romero, R.A. Ground vehicle target signature identification with cognitive automotive radar using 24–25 and 76–77 GHz bands. *IET Radar Sonar Navig.* **2018**, *12*, 1448–1465. [CrossRef]

5.   Gurbuz, S.Z.; Soraghan, J.; Balleri, A.; Clemente, C. Micro-Doppler Based In-Home Aided and Unaided Walking Recognition with Multiple Radar and Sonar Systems. *IET Radar Sonar Navig.* **2017**, *11*, 107–115. [CrossRef]

6.   Saho, K.; Sakamoto, T.; Sato, T.; Inoue, K.; Fukuda, T. Pedestrian classification based on radial velocity features of UWB Doppler radar images. In Proceedings of the International Symposium on Antennas and Propagation, Nagoya, Japan, 29 October–2 November 2012; pp. 90–93.

7.   Yue, Z.; Gao, F.; Xiong, Q.; Wang, J.; Huang, T.; Yang, E.; Zhou, H. A Novel Semi-Supervised Convolutional Neural Network Method for Synthetic Aperture Radar Image Recognition. *Cogn. Comput.* **2019**. [CrossRef]

8.   Gao, F.; Ma, F.; Wang, J.; Sun, J.; Yang, E.; Zhou, H. Visual Saliency Modeling for River Detection in High-Resolution SAR Imagery. *IEEE Access* **2018**, *6*, 1000–1014. [CrossRef]

9.   Fan, Z.; Yao, X.; Tang, H.; Qiang, Y.; Hu, Y.; Lei, B. Multiple Mode SAR Raw Data Simulation and Parallel Acceleration for Gaofen-3 Mission. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2018**, *11*, 2115–2126.

10.   Hudson, S.; Psaltis, D. Correlation filters for aircraft identification from radar range profiles. *IEEE Trans. Aerosp. Electron. Syst.* **2002**, *29*, 741–748. [CrossRef]

11.   Liao, X.; Bao, Z.; Xing, M. On the aspect sensitivity of high resolution range profiles and its reduction methods. In Proceedings of the Record of the IEEE 2000 International Radar Conference [Cat. No. 00CH37037], Alexandria, VA, USA, 12 May 2000; pp. 310–315.

12.   Beckman, D.; Frame, S. Comparison of features from SAR and GMTI imagery of ground targets. In Proceedings of the International Society for Optics and Photonics Algorithms for Synthetic Aperture Radar Imagery X, Orlando, FL, USA, 21 April 2003; Volume 5095, pp. 224–233.

13.   Jacobs, S.P.; O'Sullivan, J.A. Automatic target recognition using sequences of high resolution radar range-profiles. *IEEE Trans. Aerosp. Electron. Syst.* **2000**, *36*, 364–381. [CrossRef]

14.   Copsey, K.; Webb, A. Bayesian gamma mixture model approach to radar target recognition. *IEEE Trans. Aerosp. Electron. Syst.* **2003**, *39*, 1201–1217. [CrossRef]

15.   Lan, D.; Liu, H.; Zheng, B.; Zhang, J. A two-distribution compounded statistical model for Radar HRRP target recognition. *IEEE Trans. Signal Process.* **2006**, *54*, 2226–2238. [CrossRef]

16.   Liao, X.; Runkle, P.; Carin, L. Identification of ground targets from sequential high-range-resolution radar signatures. *IEEE Trans. Aerosp. Electron. Syst.* **2002**, *38*, 1230–1242. [CrossRef]

17.   Albrecht, T.W.; Gustafson, S.C. Hidden Markov models for classifying SAR target images. *Proc. SPIE* **2004**, *5427*, 302–308.

18.   Zhang, X.; Liu, Z.; Liu, S.; Li, G. Time-Frequency feature extraction of HRRP using AGR and NMF for SAR ATR. *J. Electr. Comput. Eng.* **2015**, *2015*, 36. [CrossRef]

19.   Williams, R.; Westerkamp, J.; Gross, D.; Palomino, A. Automatic target recognition of time critical moving targets using 1D high range resolution (HRR) radar. *IEEE Aerosp. Electron. Syst. Mag.* **2000**, *15*, 37–43. [CrossRef]

20.   Wong, S. Multi-look fusion identification: A paradigm shift from quality to quantity in data samples. *Proc. Spie Int. Soc. Opt. Eng.* **2009**, *7335*.

21.   Gross, D. A neural network ATR for high range resolution radar signature recognition of moving ground targets. In Proceedings of the IEEE Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 24–27 October 1999; Volume 2, pp. 1235–1239.

22.   Kahler, B.; Blasch, E. Robust multi-look HRR ATR investigation through decision-level fusion evaluation. In Proceedings of the 2008 IEEE 11th International Conference on Information Fusion, Cologne, Germany, 30 June–3 July 2008; pp. 1–8.

23.   Kittler, J.; Hatef, M.; Duin, R.P.; Matas, J. On combining classifiers. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 226–239. [CrossRef]

24.   Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436. [CrossRef]

25.   Gao, F.; Huang, T.; Wang, J.; Sun, J.; Hussain, A.; Yang, E. Dual-Branch Deep Convolution Neural Network for Polarimetric SAR Image Classification. *Appl. Sci.* **2017**, *7*, 447. [CrossRef]

26. Vincent, P.; Larochelle, H.; Lajoie, I.; Bengio, Y.; Manzagol, P.A. Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. *J. Mach. Learn. Res.* **2010**, *11*, 3371–3408.

27. Hinton, G.E. A Practical Guide to Training Restricted Boltzmann Machines. *Momentum* **2012**, *9*, 599–619.

28. Tai, K.S.; Socher, R.; Manning, C.D. Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks. *Comput. Sci.* **2015**, *5*, 36.

29. Huang, Y.L.; Xu, B.B.; Ren, S.Y. Analysis and pinning control for passivity of coupled reaction-diffusion neural networks with nonlinear coupling. *Neucom.* **2018**, *272*, 334–342. [CrossRef]

30. Zhao, F.; Liu, Y.; Huo, K.; Zhang, S.; Zhang, Z. Radar HRRP Target Recognition Based on Stacked Autoencoder and Extreme Learning Machine. *Sensors* **2018**, *18*, 173. [CrossRef] [PubMed]

31. Peng, X.; Gao, X.; Zhang, Y.; Li, X. An Adaptive Feature Learning Model for Sequential Radar High Resolution Range Profile Recognition. *Sensors* **2017**, *17*, 1675. [CrossRef]

32. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735. [CrossRef]

33. Soltau, H.; Liao, H.; Sak, H. Neural speech recognizer: Acoustic-to-word LSTM model for large vocabulary speech recognition. *arXiv* **2016**, arXiv:1610.09975.

34. Wen, T.H.; Gasic, M.; Mrksic, N.; Su, P.H.; Vandyke, D.; Young, S. Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems. *Comput. Sci.* **2015**, 1711–1721. [CrossRef]

35. Jithesh, V.; Sagayaraj, M.J.; Srinivasa, K.G. LSTM recurrent neural networks for high resolution range profile based radar target classification. In Proceedings of the International Conference on Computational Intelligence & Communication Technology, Ghaziabad, India, 9–10 February 2017; pp. 1–6.

36. Bin, X.U.; Bo, C.; Liu, H.; Lin, J. Attention-based Recurrent Neural Network Model for Radar High-resolution Range Profile Target Recognition. *J. Electron. Inf. Technol.* **2016**. [CrossRef]

37. Xu, Z.; Li, S.; Deng, W. Learning temporal features using LSTM-CNN architecture for face anti-spoofing. In Proceedings of the 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), Kuala Lumpur, Malaysia, 3–6 November 2015; pp. 141–145.

38. Ma, X.; Hovy, E. End-to-end sequence labeling via bi-directional lstm-cnns-crf. *arXiv* **2016**, arXiv:1603.01354.

39. Graves, A.; Schmidhuber, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw. Off. J. Int. Neural Netw. Soc.* **2005**, *18*, 602–610. [CrossRef]

40. Schmidhuber, J.; Gers, F.A. Recurrent Nets that Time and Count. In Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium, Como, Italy, 27 July 2007; Volume 03, p. 3189.

41. Weninger, F.; Bergmann, J.; Schuller, B. Introducing CURRENNT: The Munich Open-Source CUDA Recurrent Neural Network Toolkit. *J. Mach. Learn. Res.* **2015**, *16*, 547–551.

42. Liu, Q.; Zhou, F.; Hang, R.; Yuan, X. Bidirectional-Convolutional LSTM Based Spectral-Spatial Feature Learning for Hyperspectral Image Classification. *Remote Sens.* **2017**, *9*, 1330. [CrossRef]

43. Zhang, F.; Hu, C.; Yin, Q.; Li, W.; Li, H.; Hong, W. SAR Target Recognition Using the Multi-aspect-aware Bidirectional LSTM Recurrent Neural Networks. *arXiv* **2017**, arXiv:1707.09875.

44. Chen, V.C.; Ling, H. *Time-Frequency Transforms for Radar Imaging and Signal Analysis*; Artech House: Norwood, MA, USA, 2002.

45. Lui, H.S.; Persson, M.; Shuley, N.V. Joint time-frequency analysis of transient electromagnetic scattering from a subsurface target. *IEEE Antennas Propag. Mag.* **2012**, *54*, 109–130.

46. Sullivan, R. *Radar Foundations for Imaging and Advanced Concepts*; Scitech Publishing: Boston, MA, USA, 2004.

47. Lui, H.S.; Shuley, N. Joint time-frequency analysis on UWB radar signals. In Proceedings of the International Conference on Signal Processing and Communication Systems, Gold Coast, Australia, 17–19 December 2007.

48. Lui, H.S.; Shuley, N.V. Evolutions of partial and global resonances in transient electromagnetic scattering. *IEEE Antennas Wirel. Propag. Lett.* **2008**, *7*, 435–439.

49. Ross, T.D.; Velten, V.J.; Mossing, J.C. Standard SAR ATR evaluation experiments using the MSTAR public release data set. *Proc. SPIE Int. Soc. Opt. Eng.* **1998**. *3370*, 566–573.

50. Mossing, J.C.; Ross, T.D. Evaluation of SAR ATR algorithm performance sensitivity to MSTAR extended operating conditions. *Proc. SPIE Int. Soc. Opt. Eng.* **1998**, *3370*, 13.

51. Xing, M.; Bao, Z.; Pei, B. Properties of high-resolution range profiles. *Opt. Eng.* **2002**, *41*, 493–504. [CrossRef]

52. Dudgeon, D.E.; Lacoss, R.T. An overview of automatic target recognition. *Linc. Lab. J.* **1993**, *6*, 3–9.

53. Gao, F.; Huang, T.; Sun, J.; Wang, J.; Hussain, A.; Yang, E. A New Algorithm of SAR Image Target Recognition Based on Improved Deep Convolutional Neural Network. *Cogn. Comput.* **2018**, 1–16. [CrossRef]