A SAR Target Recognition Based on Guided Reconstruction and Weighted Norm-Constrained Deep Belief Network

JIAN WANG1,2, (Member, IEEE), JIE LIU1, PING REN1, AND CHUN-XIA QIN1

1Electronic and Information College, Northwestern Polytechnical University, Xi’an 710129, China
2No. 365 Institute, Northwestern Polytechnical University, Xi’an 710065, China

Corresponding author: Jian Wang (jianwang@nwpu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61671383; in part by the Shaanxi Key Industry Innovation Chain Project under Grant 2018ZDCXL-GY-03-05, Grant 2019ZDLGY14-02-02, and Grant 2019ZDLGY14-02-03; and in part by the Northwestern Polytechnical University Graduate Innovation Fund under Grant Z2017144.

ABSTRACT The deep learning has some new vitality to the synthetic aperture radar (SAR) automatic target recognition (ATR). By introducing different constraints, deep belief network (DBN) has become apply to SAR target recognition recent years, but the existing DBN algorithms have some questions including the high training epochs, low recognition rate and complex structure. Therefore, an algorithm based on guided reconstruction and weighted norm-constrained DBN is proposed. Firstly, in order to reduce the dimension of the image output feature, increase the speed of preprocessing, generate a one-dimensional image vector and normalized, a SAR target classification algorithm with two-scale fusion character based on guided filter reconstruction algorithm is introduced. Then, the sparse feature extraction of SAR image is carried out by weighted norm-constrained DBN. By the regularization constraint of the probability distribution, the algorithm can minimize the joint probability distribution of visual layer and hidden layer by samples. The low-dimensional feature is further improved based on the generalized optimization of norm-constrained RBM. Finally, by a regularized Softmax which can classify the targets and obtain output results. The experimental results show that the SAR target recognition algorithm based on the guided reconstruction and weighted constrained deep confidence network not only improves the target recognition performance and generalization ability, but also reduces the output feature dimension and network training times, and the recognition performance of the algorithm is further improved.

INDEX TERMS Guided filter, deep belief network, weighted norm-constrained, synthetic aperture radar (SAR), target recognition.

I. INTRODUCTION

With its all-weather, all-time capabilities, synthetic aperture radar (SAR) is widely used in different military applications such as target positioning, tracking, precision strike and other applications [1]. SAR image automatic target recognition (ATR) is to find the target of interest in the high resolution and wide swath SAR image and determine its label [2]–[4]. Because SAR target features do not conform to the human visual system, target feature extraction has always been a hot topic in ATR. It is two key factors of feature extraction and classifier design that affect classification accuracy in SAR image target recognition. Researchers begin to apply deep learning methods to SAR image target recognition, but there are still two kinds of problems to be solved. ☸ Compared with the visible image, SAR image has stronger speckle noise in the optimization design of depth model; ☹ In order to ensure the generalization ability of the model, a large number of training data are needed to train the model parameters due to the lack of training samples and more parameters of deep convolution neural network. In recent years, scholars have introduced some optimization design of deep learning methods with the ability to actively extract target features in the research of SAR image recognition problems, and designed...
a variety of different networks to improve recognition performance [5]–[8]. Deep convolutional networks use sparsely connected layers instead of fully connected layers, reducing the number of free parameters [5]. Later a SAR automatic target classification method based on wavelet scattering convolutional network is proposed [6]. By the space-fixed and space-varying scattering feature joint learning method, the algorithm can get the multi-aspect SAR target recognition [7]. And a bidirectional long short-term memory (LSTM) algorithm recursive neural networks as well as Softmax classifiers can achieve target recognition [8]. Although these algorithms in the above references can achieve higher recognition performance, there are some issues hindering its applicability, e.g., the classification speed and higher complexity.

In the case of limited samples, feature enhancement has proven to be one of the effective methods to solve the problem of small samples. Through different feature enhancement methods, the combination of features extracted from different network structural layers could improve the recognition performance of the algorithm. Based on transfer learning algorithm, it can transfer the learning knowledge of tagged SAR image to a new SAR target recognition task [9]. By the deep convolutional generative adversarial networks (DCGAN) which adopted a semi-supervised learning method can reduce the number of free parameters and the negative impact of mislabeled samples on network performance [10]. In addition, in the case of limited samples, the convolutional neural network (CNN) and the multi-scale feature extraction module are used for SAR image target recognition [11]. However, when CNN model training is performed in a CNN algorithm, as the number of layers increases, some input information often will disappear when it reaches the back-end network after many operations such as convolution and pooling [3], [5], [6], [11]. Therefore, the strategy of joint training to generative adversarial networks (GAN) is introduced for SAR target recognition [12]. Based on the deep belief networks (DBN), the RVIN prediction model (RVIN detector) is trained to map the feature vector to the noise type label [13]. In order to better extract the discriminative features of the samples, the autoencoder and the Euclidean distance constraint were efficiently combined to reduce the distance between intra-class samples while increasing the distance between different types of samples [14]. At the same time, to alleviate the overfitting problem of the network, the dropout operation is introduced, and the classifier uses support vector machine (SVM) as the output. By designing an effective feature fusion network, it can pre-train the sparse autoencoder (SAE) using greedy layer-wise training, and use Softmax to complete the target classification [15]. Based on the idea that the intra-class spacing of homogeneous data is small and that of non-homogeneous data is large, the similarity constraint is introduced into the Restricted Boltzmann Machines (RBM) and judgment information is added to the learning process of RBM [16]. Adopt multi-flow regularized low-rank approximation (MRLA) for SAR feature extraction, and the local sparse representation (LSR) is used as the classification decision output [17]. Use SAR-HOG operator for SAR feature extraction, and use supervised decision dictionary learning and sparse representation (SDDLSR), support vector machine (SVM), k-nearest neighbor classifier (KNN), sparse representation classification (SRC) and the label consistent kernel singular value decomposition (LCK-SVD) as output of the classifier [18]. The algorithm completes DBN network recognition of SAR targets by introducing norm regularization constraints [1] and similarity constraints [12]. High recognition accuracy could be achieved, even in the case of large differences in appearance between target categories. But the model structure is relatively complex, the network convergence is slow and the recognition rate is not high, the limitation of the number of training samples also causes overfitting of the network. In addition, for similar targets with different military configurations or small differences between targets, the above algorithm needs to resolve the coherent speckle noise of SAR images and the slight difference between targets, and its average recognition accuracy is low.

Aiming at the above problems, in order to make full use of the original information of the input image and maximize the information between networks, this article proposes a SAR target classification algorithm based on guided reconstruction and weighted constrained deep belief network. First, a SAR image is enhanced with two-scale data using a guided reconstruction algorithm to generate a one-dimensional image vector and normalized in order to reduce the dimensionality of image output features and improve the speed of preprocessing; then, the weighted norm-constrained DBN algorithm performs deep sparse feature extraction on SAR targets. Through the rational design of the model structure, redefining the generalized optimization problem of the restricted Boltzmann machine with weighted norm constraints, it can minimize the joint probability distribution of the samples in the visible and hidden layers. With the regularization constraint of the probability distribution of the hidden layer, the sparsity of low-dimensional features is further improved. Finally, a regularized Softmax is connected to classify the target in order to obtain the output result.

II. SAR IMAGE RECONSTRUCTION WITH GUIDED FILTER

For SAR target recognition, because the differences between different and same categories targets are relatively small in spatial characteristics, in order to make full use of the original information of the input image, a relevant algorithm needs to be used to reconstruct and preprocess the SAR original image to fully highlight the differences between heterogeneous target images. This article uses a feature map based on guided filter (GF) to perform two-scale image reconstruction and feature enhancement on the SAR image [19].

First, take the absolute value of the high-frequency image after Laplacian filtering for the original image $I_n$, and perform local average to obtain the reconstructed feature map $S_n$:

$$ S_n = |I_n \times L| \times g_{r,\sigma} \quad (1) $$
In the formula, \( I_n \) is the \( n \)-th source image and \( L \) is a \( 3 \times 3 \) Laplacian filter. \( g_{r,\sigma} \) is a Gaussian low-pass filter with a standard deviation \( \sigma \) of a window \((2r+1)\times(2r+1)\), and both \( r \) and \( \sigma \) are set to 5 [15]. The significance level of the image high-frequency detail component can be effectively represented by \( S_n \). The weight map \( P_n \) of \( S_n \) is as follows:

\[
P_n^k = \begin{cases} 
1 & \text{if } S_n^k = \max(S_{n1}^k, S_{n2}^k, \ldots, S_{nN}^k) \\
0 & \text{otherwise}
\end{cases}
\]  

(2)

In the formula, \( N \) is the number of input source images, and \( S_n^k \) is the feature map value when the number of pixels of the \( n \)-th input image is \( k \). After GF processing for each \( P_n \) and \( I_n \), the weighted graphs are as follows:

\[
W_n^B = GF_{r_1,\varepsilon_1}(P_n, I_n)
\]  

(3)

\[
W_n^D = GF_{r_2,\varepsilon_2}(P_n, I_n)
\]  

(4)

In the formula, \( r_1, \varepsilon_1, r_2 \) and \( \varepsilon_2 \) are parameters of weighted GF, \( W_n^B \) and \( W_n^D \) are weighted graphs of low frequency approximate components and high frequency detail components respectively. The parameter set is \( r_1 = 45, \varepsilon_1 = 0.3, r_2 = 7, \varepsilon_2 = 10^{-6} \), and \( W_n^B, W_n^D \) are normalized.

Finally, the low-frequency approximate component \( I_n^L \) and high-frequency detail component \( I_n^H \) of \( I_n \) are reconstructed by weighted average to obtain \( I_n^B \) and \( I_n^D \) respectively, and the reconstructed image \( I_n^R \) is:

\[
I_n^R = I_n^B + I_n^D = \sum_{n=1}^{N} \left( W_n^B(I_n \times G) + W_n^D(I_n - I_n \times G) \right)
\]  

(5)

In the formula, the low-frequency approximate component \( I_n^B \) is obtained by performing a two-scale decomposition on the source image using a window \((2r+1)\times(2r+1)\) (where \( r = 15 \)) smoothing filter \( G \), and the high-frequency detail component \( I_n^D \) is obtained from \( I_n \) minus low frequency approximation components.

### III. FEATURE EXTRACTION OF WEIGHTED CONSTRAINED DEEP BELIEF NETWORK

#### A. RESTRICTED BOLTZMANN MACHINE AND DEEP BELIEF NETWORK TRAINING MODEL

DBN can extract feature information from high-dimensional complex data, but it is easy to fall into the problem of local minima by using the traditional shallow network structure [20]. Every 2 adjacent layers in the DBN network are initialized as Restricted Boltzmann Machine (RBM). The RBM is composed of several neurons in the visible layer and the hidden layer and its neurons in the same layer are not connected to each other. In RBM, the weight of each neuron is represented by a neuron bias \( B \) in the visible layer and a neuron bias \( C \) in the hidden layer. The weights of any two visible layers connected to the hidden layer neurons are represented as \( W^{vh} \) for its connection strength, \( V (V = \{v_1, v_2, \ldots, v_J, \ldots, v_D\}) \) and \( H (H = \{h_1, h_2, \ldots, h_j, \ldots, h_F\}) \) are binary random variables of visible layer and hidden layer, respectively. In RBM, the output unit is conditionally independent of the input state, so given an input, an unbiased sample of the posterior distribution can be obtained. This sample can be expressed as:

\[
P (h_j | V) = \sigma \left( b_j + \sum_l W_{lj}^v x_l \right)
\]  

(6)

\[
P (v_i | H) = \sigma \left( c_i + \sum_j W_{ij}^h h_j \right)
\]  

(7)

where \( \sigma \) is the sigmoid function, \( W_{lj}^v \) is the weights of the neurons in the \( i \)-th visible layer and the neurons in the \( j \)-th hidden layer. So the \( W \) define the feature of the sample. Based on this, the energy function can be defined as:

\[
E (V, H) = -B^T V - C^T H - V^T W^{vh} H
\]  

\[
= - \sum_{i=1}^{d_x} b_i v_i - \sum_{j=1}^{d_y} c_i h_j - \sum_{i=1}^{d_x} \sum_{j=1}^{d_y} W_{ij}^{vh} v_i h_j
\]  

(8)

Among them, \( W^{vh} = \{w_{ij}^{vh}\} \) represents the weight basis matrix, \( v_i \) represents the neuron in the \( i \)-th visible layer \( V \), \( h_j \) represents the neuron in the \( j \)-th hidden layer \( H \), \( d_x \) represents the number of neurons in the visible layer, and \( d_y \) represents the number of neurons in the hidden layer.

It can be known from the energy model of the RBM that the visible layer neurons reconstructed by the hidden layer satisfy the Gibbs distribution, but the weight parameters \( W^{vh} \) and the biases \( B \) and \( C \) in the energy model are unknown. These parameters need to be learned through learning. The training process of RBM is to solve the minimization of the energy function to update the weight between two adjacent layers and the deviation between the layers.

For a sample of data \( x_n \), the parameters are updated using a Contrastive Divergence (CD) algorithm. The purpose of training the RBM network is to initialize the weight parameters of the regression network, thereby overcoming the shortcomings of the regression network which is easy to fall into the local optimum and the training time is too long due to the random initialization of the weight parameters. When training RBM, first, the input sample vector is input into the network through the visible layer to obtain a vector \( V \) of the visible layer, and the vector \( V \) passes the output of the network to the hidden layer \( H \) through a non-linear mapping; then the input original signal is reconstructed without distortion through an input method of randomly selecting visual layers; finally, the new neurons randomly selected in the visible layer are sampled by Gibbs and mapped non-linearly to forward transfer to reconstruct the activated neurons of the hidden layer, gaining \( H \). These backward and forward steps to randomly extract some neurons are Gibbs sampling operations, and the core basis for supporting the update weight is the correlation difference between the activated
neurons in the hidden layer and the input of the visual layer.

B. LOW-DIMENSIONAL FEATURE EXTRACTION OF WEIGHTED NORM-CONSTRAINED DEEP BELIEF NETWORKS

Given a set of \( N \) training data \( \{x_1, x_2, \ldots, x_N\} \), the generalized optimization problem of norm-constrained Boltzmann machines can be described as:

\[
\min_{W_{1h}, c_j} \sum_{n=1}^{N} \sum_{i} \ln p(x_n, h_n; \theta) + \lambda \sum_{n=1}^{N} f(p(h_n|x_n))
\]  

(9)

Among them, \( p(h_n|x_n) = \frac{1}{1 + \exp(-\sum_{j=1}^{d_h} W_{ji} h_n - c_j)} \) represents the joint probability distribution of the \( j \)-th neuron of the \( n \)-th sample \( p(h_n|x_n) = \frac{1}{1 + \exp(-\sum_{j=1}^{d_h} W_{ji} h_n - c_j)} \) is activated in the visible layer and hidden layer. \( f(p(h_n|x_n)) = \sum_{j=1}^{d_h} |p(h_n = 1|x_n)| \), \( f(p(h_n|x_n)) = \sum_{j=1}^{d_h} |p(h_n = 1|x_n)| \) are functions related to \( f(p(h_n|x_n)) = \sum_{j=1}^{d_h} |p(h_n = 1|x_n)| \).

In a SAR target image, the texture feature of the target is the external representation of the position and relationship between the scattering centers. Compared with the target itself, the scattering center is sparse. Therefore, it can be speculated that the accuracy and effectiveness of target feature representation in SAR ATR can be improved by adding these methods. The common methods include adding \( L_1 \), \( L_2 \) and \( L_{1/2} \) norm constraints to the objective function \( [12], [20], [21] \). So the regularization constraint is applied to the probability distribution of the hidden layer to improve the efficiency of the network. The item on the right is expressed in the form of norm as:

\[
\left( f(p(h_j|x)) \right)_{L_1} = \sum_{j=1}^{d_h} |p(h_j = 1|x)|
\]

(10)

\[
\left( f(p(h_j|x)) \right)_{L_2} = \frac{1}{2} \sum_{j=1}^{d_h} |p(h_j = 1|x)|^2
\]

(11)

\[
\left( f(p(h_j|x)) \right)_{L_{1/2}} = \sum_{j=1}^{d_h} |p(h_j = 1|x)|^{1/2}
\]

(12)

Applying regularization constraints to the probability distribution of the hidden layer can improve the efficiency of the network. The gradients of the regularized representations of the three norms can be obtained by derivation.

\[
\left( \frac{\partial \sum_{j=1}^{d_h} \sum_{n=1}^{N} |p(h_{nj} = 1|x_n)|}{\partial c_j} \right)_{L_1} = \sum_{n=1}^{N} (p(h_{nj} = 1|x_n) - p^2(h_{nj} = 1|x_n))
\]

(13)

\[
\left( \frac{\partial \sum_{j=1}^{d_h} \sum_{n=1}^{N} |p(h_{nj} = 1|x_n)|^2}{\partial c_j} \right)_{L_2} = 2 \sum_{n=1}^{N} (p^2(h_{nj} = 1|x_n) - p^3(h_{nj} = 1|x_n))
\]

(14)

\[
\left( \frac{\partial \sum_{j=1}^{d_h} \sum_{n=1}^{N} |p(h_{nj} = 1|x_n)|^{1/2}}{\partial c_j} \right)_{L_{1/2}} = \frac{1}{2} \sum_{n=1}^{N} (p^{1/2}(h_{nj} = 1|x_n) - p^{3/2}(h_{nj} = 1|x_n))
\]

(15)

For the regular term on the right side of formula (9) which can be expressed as a linear superposition of \( n \) different norms, the generalized optimization problem of norm constrained Boltzmann machine is used to deal with.

\[
f(x) = \sum_{i=1}^{n} \lambda_i \|X\|_{p_i}
\]

(16)

Among them, \( \lambda_i \) is the weight of different forms of norm regularization constraints, \( \|X\|_{p_i} \) represents different \( p \) norm forms, the optimal network model and its coefficient \( \theta = \{W_{vh}, b, c\} \) and the low-dimensional feature vector corresponding to each training sample \( h_n \) \( [n = 1, 2, \ldots, N] \) can be obtained through multiple rounds of training.

IV. SAR TARGET RECOGNITION BASED ON GUIDED RECONSTRUCTION AND WEIGHTED CONSTRAINT DBN

According to the above analysis, in order to be able to identify the target SAR image containing coherent speckle noise, it should be designed to get or enhance these nuances from the input image, so as to improve the target recognition accuracy. The algorithm of this article is shown in FIGURE 1. First, using a guided reconstruction algorithm can reduce the dimensionality of the output features and increase the speed of pre-processing, and the recognition performance is better; then, by using the weighted norm-constrained DBN to perform deep sparse feature extraction on the SAR target, under the condition of regularization constraint of the probability distribution of the hidden layer, by minimizing the joint probability distribution of the samples in the visible layer and the hidden layer, the sparsity of low-dimensional features is further improved on the basis of the generalized optimization of norm-constrained RBM; finally, the regularized Softmax classifier is used to classify SAR targets. However, SAR target recognition using the weighted constrained deep belief network method, apart from designing a reasonable network structure and adjusting and optimizing parameters for the network, network optimization is also needed. During the optimization process, one of the most common problems is overfitting. Because the training sets and test set data have some unique features in addition to their common features. If the network parameters matching the training set data are more accurate, the recognition performance of the test set is worse, that is to say, the generalization performance of
the network is very poor. When the network parameters are trained, if the training data is overfitting, it will reduce the fitting ability and recognition performance of the test set, resulting in poor generalization performance of the network. The regularization method is introduced to solve the problem of overfitting. Therefore, it is necessary to solve the overfitting problem of model training caused by limited SAR data set [24]. In this regard, the following methods are adopted in this article: ① The regularization method is introduced for the network to solve the overfitting problem. The regularization method uses three kinds of norm regularization $L_{1/2}$, norm regularization $L_1$ and norm regularization $L_2$ [24], [23]. ② Use the Dropout operation to eliminate and weaken the joint adaptability of hidden layer neurons to achieve better generalization ability [5]. Through the experimental verification, this algorithm uses the dropout operation before the Softmax layer, and discards the upper neurons when the dropout of the hidden layer network node is set to 0.5, which can improve the generalization ability of the model during the training stage.③ Normalize the output of each layer of the network, and add a linear transformation layer to the network, so that the input of the next layer approaches to the Gaussian distribution, in order to improve the speed of the gradient descent method to solve the optimal network parameters [24] and termination of cross-validation [25].

The regularization method is used to solve the problem of overfitting in order to restrict or limit the parameters to be optimized. During the test, it is necessary to adjust and optimize the network parameters according to the test conditions of the test samples. In our actual test, when the test sample data is used for testing, the overfitting model may appear and cause the poor accuracy. Therefore, in view of the poor performance of test samples and the lack of good generalization, sometimes it is necessary to adjust the model algorithm in the test process. By adding regularization constraints, the solution space can be further reduced, and even in the individual regularization mode, the solution obtained becomes sparse, which further improves the generalization performance of the algorithm model.

V. EXPERIMENTAL SIMULATION AND VERIFICATION ANALYSIS

In order to further verify the performance of the algorithm in this article, three (TABLE 1) and ten classes (TABLE 2) of SAR target sample databases in the MSTAR project data set are used. The training sample set uses the data of pitch angle of 17°, and the test sample set uses the data of pitch angle of 15°. Three classes of targets include: BMP-2 infantry fighting vehicle (including SN-9563, SN-9566, and SN-C21), T-72 tank (including SN-132, SN-812, and SN-S7), and BTR-70 armored transport vehicle (including SN-C71). The training and test sample set is shown in TABLE 1 and FIGURE 2. Based on the three classes of targets, the ten classes of targets also include: BTR60, 2S1, BRDM2, D7, T62, ZIL131, ZSU23/4, and the training and test sample set is shown in TABLE 2. At present, SAR target recognition classification scenarios mainly include two types of standard operating conditions (SOC) and extended operating condition (EOC). Test environment: Intel (R) Core (TM) i5-3230M CPU @ 2.6GHz, memory of 4GB, programmed with MATLAB R2016a, tested respectively under SOC and EOC.

A. ROBUSTNESS ANALYSIS OF STANDARD OPERATING CONDITIONS (SOC) ALGORITHM

In order to make full use of the original information of the input image and maximize the difference between samples, the size of the input image reconstructed by the guided filter has a size of $128 \times 128$. After the reconstruction preprocessing is completed, the algorithm further adopts the method of clipping the training image to ensure the recognition accuracy of the system and save the calculation cost. This method can reflect the performance of recognition algorithm and further reflect the influence of image resolution on recognition rate. Experimental tests show that when $128 \times 128$ input image is cropped to $64 \times 64$ pixel resolution under SOC condition, high recognition accuracy can be achieved, and increasing image resolution cannot further improve model recognition rate [11]. Therefore, after the input image of $128 \times 128$ training set is reconstructed in this algorithm, the training sets...
image is cropped $32 \times 32$ or $64 \times 64$ image blocks, the image blocks are vectorized and energy normalized into $1 \times 1024$ or $1 \times 4096$ image vectors, gaining the inputs of the five structures in TABLE 3.

Structure 1: The DBN network structure is relatively simple and does not consider the problem of network training over-fitting. The structure $1024 \times 98$ uses a three-layer forward neural network, with dropout set to 0 and a weighted momentum factor of 0.69. There is no norm-constrained optimization problem. The sigmoid function is used for the output activation function of each layer (including the output layer), and the model is trained 10 times. With the change of noise reduction parameters, the recognition results of three types of targets by three-layer depth neural network are shown in TABLE 3.

It can be seen from TABLE 3 that the recognition rate is not optimal when no noise is introduced into the data, further proving the difference between the training sample and the test sample. When different levels of noise are added to the data, the difference between the training sample and the test sample is reduced to some extent, and the robustness and the recognition performance of the algorithm is better. Because the different network models have different structure, different training methods and different fitting degree.
of samples. Therefore, the noise reduction coefficient of the optimal recognition performance is also different.

Structure 2: The DBN consists of two $1024 \times 100$ and $100 \times 100$ restricted Boltzmann machines. DBN is used to extract features from the target, and the deep neural network is used as the classification output. The initial weights are randomly generated by Gaussian distribution with mean value of 0 and variance of 0.05, the biases are all initialized as 0, and the learning rates of RBM and DBN are set to 0.0045. In the first three layers of the deep neural network, the weights and bias of each layer of each layer of the trained DBN are used. In order to reduce the overfitting effect of the network, the dropout is set to 0.5, the weighted momentum factor experimentally verified is 0.65. The output activation function of each layer (including the output layer) uses the sigmoid function, and the model is trained 25 times. The results of classification and recognition of three kinds of SAR targets by this network are shown in TABLE 4.

Structure 3: The DBN consists of a $1024 \times 98$ restricted Boltzmann machine, randomly generating initial weights with a Gaussian distribution with a mean value of 0 and a variance of 0.01, the biases are all initialized as 0, the learning rates of RBM and DBN are set to 0.0045, and the norm constraint coefficients are set to $\lambda_1=0$, $\lambda_2=0$, $\lambda_3=0.0045$. The dropout is set to 0.5 and the weighted momentum factor is set to 0.65. The activation function of each layer uses the sigmoid function, the output uses the Softmax classifier, and the model is trained 25 times. TABLE 6 describes the identification results of structure 4 with different degrees of noise added to the input, and TABLE 7 describes the classification results of structure 4 with different coefficient values of weighted norm constraints.

For further analysis, the algorithm in this article compares the recognition performance of these four different network structures, as shown in TABLE 8. From TABLE 8, it can be concluded that compared with structure 1, structure 3 (Network Structure $1024 \times 98$) has better recognition performance of deep belief network than forward neural network under the same network structure 1, with an increase of 1.36%; using a stacked Boltzmann machine, with the increase of the number of Boltzmann machine layers and neurons, the structure 2, with a small number of data samples, its network structure is

---

### TABLE 3. Identification results of network structure 1 with the change of noise parameters.

| Network Structure  | 1024-100 | 1024-98 | 1024-98 | 1024-98 | 1024-98 | 1024-98 | 1024-98 |
|--------------------|----------|---------|---------|---------|---------|---------|---------|
| Batch Training Samples | 24       | 29      | 29      | 29      | 29      | 29      | 29      |
| L2 Norm Constraint  | $10^{-4}$ | $10^{-6}$ | $10^{-5}$ | $10^{-5}$ | $10^{-5}$ | $10^{-5}$ | $10^{-5}$ |
| Noise Reduction Parameters | 0        | 0.05    | 0.1     | 0.15    | 0.2     | 0.25    | 0.3     |
| Learning Rate       | 1.4      | 1.35    | 1.19    | 1.34    | 1.24    | 1.36    | 1.38    |
| Learning Rate Change Factor | 1.2    | 1.21    | 1.21    | 1.35    | 1.35    | 1.4     | 1.4     |
| Recognition Rate    | 93.19%   | 93.02%  | 93.02%  | 94.21%  | 93.36%  | 93.56%  | 92.68%  |

### TABLE 4. Identification results of network structure 2 with the change of noise parameters.

| Batch Training Samples | 24 | 24 | 24 | 24 | 24 | 24 | 24 |
|------------------------|----|----|----|----|----|----|----|
| L2 Norm Constraint     | 10^{-5} | 10^{-5} | 10^{-5} | 10^{-5} | 10^{-5} | 10^{-5} | 10^{-5} |
| Noise Reduction Parameters | 0     | 0.05 | 0.1  | 0.15 | 0.2  | 0.25 | 0.3 |
| Learning Rate          | 6.2 | 6  | 6.2 | 6.4 | 6.2 | 6.8 | 6.6 |
| Learning Rate Change Factor | 1     | 1   | 1   | 1   | 1   | 1   | 1   |
| Recognition Rate       | 93.87% | 93.19% | 93.02% | 92.85% | 91.48% | 91.14% | 89.95% |
TABLE 5. Identification results of network structure 3 with the change of noise parameters.

|                              | 1024-98 | 1024-98 | 1024-98 | 1024-98 | 1024-98 | 1024-98 |
|------------------------------|---------|---------|---------|---------|---------|---------|
| Batch Training Samples       | 24      | 24      | 24      | 24      | 24      | 24      |
| L2 Norm Constraint           | $10^{-5}$ | $10^{-5}$ | $10^{-5}$ | $10^{-5}$ | $10^{-6}$ | $10^{-5}$ |
| Noise Reduction Parameters   | 0       | 0.05    | 0.1     | 0.15    | 0.2     | 0.25    |
| Learning Rate                | 6.9     | 6.9     | 5.6     | 6.9     | 6.9     | 6       |
| Learning Rate Change Factor  | 1       | 1       | 1       | 1       | 1       | 1       |
| Recognition Rate             | 93.36%  | 94.55%  | 94.04%  | 94.21%  | 95.57%  | 94.38%  | 94.38%  |

TABLE 6. Identification results of network structure 4 with the change of noise parameters.

|                              | 1024-100 | 1024-100 | 1024-100 | 1024-100 | 1024-100 | 1024-100 |
|------------------------------|----------|----------|----------|----------|----------|----------|
| Batch Training Samples       | 24       | 24       | 24       | 24       | 24       | 24       |
| L2 Norm Constraint           | $10^{-6}$ | $10^{-6}$ | $10^{-5}$ | $10^{-5}$ | $10^{-6}$ | $10^{-5}$ |
| Noise Reduction Parameters   | 0       | 0.05    | 0.1     | 0.15    | 0.2     | 0.25    |
| Learning Rate                | 1.5      | 1.65    | 2.2     | 1.82    | 2.75    | 2.2     |
| Learning Rate Change Factor  | 1       | 1       | 1       | 1       | 1       | 1       |
| Recognition Rate             | 95.2%    | 94.89%  | 96.42%  | 96.76%  | 96.42%  | 96.93%  | 97.27%  |

TABLE 7. The recognition result of our algorithm is different with the value of weighted norm constraint coefficient.

| Recognition Rate | 95.40% | 96.25% | 97.10% | 97.27% | 96.59% | 96.93% | 96.42% | 96.42% |
|------------------|--------|--------|--------|--------|--------|--------|--------|--------|
| $\lambda_1$     | 0      | 0.0045 | 0      | 0      | 0.0045 | 0      | 0      | 0      |
| $\lambda_2$     | 0      | 0      | 0.0045 | 0      | 0.0045 | 0      | 0      | 0.0045 |
| $\lambda_3$     | 0      | 0      | 0      | 0.0045 | 0      | 0.0045 | 0      | 0.0045 |

TABLE 8. Comparison of recognition performance of several different network structures in the algorithm of this article.

| Category             | Structure1 | Structure2 | Structure3 | Structure3 | Structure4 |
|----------------------|------------|------------|------------|------------|------------|
| Network Structure    | 1024×98    | 1024×100   | 1024×98    | 1024×100   | 4096×100   |
| Batch Training Samples| 29        | 24         | 24         | 24         | 24         |
| L2 Norm Constraint   | 10.5       | 10.5       | 10.5       | 10.5       | 10.5       |
| Noise Reduction Parameters | 0.15 | 0          | 0.2       | 0.3       | 0.3       |
| Learning Rate        | 1.34       | 6.2        | 6.9        | 2.8        | 2.2        |
| Learning Rate Change Factor | 1.35 | 1          | 1         | 1         | 1         |
| Recognition Rate     | 0.94208    | 0.93867    | 0.95571    | 0.97274    | 0.97615    |

more complex, the recognition performance is reduced and the risk of overfitting will be greatly increased. Based on structure 3 (Network Structure 1024×98), a weighted norm regularization constraint of structure 3 (Network Structure 1024×100) is introduced and the recognition performance is improved by 1.704%. Structure 4 combines weighted norm regularization constraints and Softmax output, and verifies that the combination of weighted norm regularization
constraints and Softmax can improve the recognition performance of DBN networks. The algorithm of this article sets a reasonable network structure, pre-processes the training samples by using a guided reconstruction algorithm, extracts deep sparse features by the weighted norm constrained deep belief network as well as using the Softmax classifier as the output layer of the recognition network to improve the recognition performance of the network.

To further test the performance of the algorithm in this article, it is compared with other papers. In TABLE 9, the similarity constrained DBN algorithm contains three layers of similarity-constrained Boltzmann machines, and the corresponding number of hidden layer neurons is 1000, 500, and 200, respectively, the dimensions of the \( u \) vectors used in each layer is 500, 500, and the maximum number of iterations is set to 300 [16]. Based on the L0.5 norm-constrained DBN algorithm, a two-layer recognition system is adopted. The number of neurons in the hidden layer is set to 300, the number of iterations is 200, and the number of nodes in the visible layer is set to 4096 [12]. The experimental results show that the network structures in the paper [12], [16] are relatively complex, with more iterations and slow convergence speed, and unsatisfactory recognition performance.

In this article, Gaussian distribution with mean value of 0 and variance of 0.01 is used to generate the initial connection weights of the network. Then the biases are initialized with 0. The learning rates of RBM and DBN are set to 0.0045, and the norm constraint coefficient is \( \lambda_1 = 0, \lambda_2 = 0, \lambda_3 = 0.0045 \). The network model is trained for 25 times, and the relatively good output results can be obtained.

It can be seen from TABLE 9 that on the basis of [16], the algorithm in this article uses the guided reconstruction algorithm to enhance SAR image, and overall recognition rate of the algorithm is improved by 5.075%. In this article, the algorithm uses the same norm-constrained DBN as well as the guided filter to the SAR image reconstruction, and compared with [12] and [14] algorithms the recognition rate is improved by 2.305% and 3.134% respectively. In order to further test the network performance, the SAR target is increased from three classes to ten classes. The number of hidden layer neurons in this algorithm is less, the number of iterations is greatly reduced, and the convergence speed is faster. Compared with the algorithm in reference, the recognition rate of this algorithm is still improved, and its recognition rate is improved by 2.185% [15]. Therefore, with the increase of SAR target types, this algorithm is still effective.

### B. ROBUSTNESS ANALYSIS OF THE EXTENDED OPERATING CONDITION (EOC) ALGORITHM

Compared with the standard operating conditions (SOC), the existence of the extended operating conditions (EOC) of the target images, the training set and test set have following characteristics such as the pitch angle changes greatly, the shape configuration changes greatly, the training set and the test set of the same target of each model target image difference, so it can further verify the generalization of the recognition network. To further verify the algorithm’s EOC performance and the generalization of recognition network, three classes of targets were tested for several different pitch angles: 2S1 self-propelled howitzer, BRDM-2 armored reconnaissance vehicle, ZSU-234 and self-propelled anti-aircraft gun. The training sample set uses SAR image data of three kinds of targets with pitch angle of 15° and the test sample set uses SAR image data of three types of targets with pitch angle of 30° and 45° respectively.

It can be seen from TABLE 10 that SAR-HOG operator and the algorithm in this article have better recognition performance than MLA-LSR algorithm when the test sample is 30° pitch angle. Softmax as a classifier output is better than SVM, KNN, SRC, LCK-SVD algorithm in classification performance. When the pitch angle changes, it is slightly worse than the SDDL algorithm in robustness, but the training speed is faster and the recognition efficiency is higher.
TABLE 10. Accuracy of different references when pitch angles and configuration change.

| Method              | Training Time (s) | Test Time (s) | 30-Degree Pitch Recognition Rate | 45-Degree Pitch Recognition Rate |
|---------------------|------------------|---------------|----------------------------------|----------------------------------|
| MLA-LSR[17]         | 56.62            | -             | 91.29%                           | 77.78%                           |
| SAR-HOG-SDDLSR[18]  | 1404.72          | 0.828         | 96.61%                           | 80.86%                           |
| SAR-HOG-SVM[18]     | 5.766            | 11.262        | 91.32%                           | 72.61%                           |
| SAR-HOG-KNN[18]     | 0.618            | 1.044         | 92.53%                           | 72.36%                           |
| SAR-HOG-SRC[18]     | 0.0362           | 1.182         | 93.05%                           | 73.02%                           |
| SAR-HOG-LCK-SVD[18] | 113.76           | 0.816         | 92.62%                           | 73.76%                           |
| Our Algorithm       | 0.2718           | 0.01934       | 92.64%                           | 76.04%                           |

TABLE 11. Training samples and change of pitch angles recognition performance.

|                      | DBN Batch Training Sample | NN Batch Training Samples | Recognition Rate |
|----------------------|---------------------------|---------------------------|------------------|
|                      | 20                        | 21                        | 24               | 28               | 30               | 35               | 40               | 42               |

TABLE 12. Parameter performance of change of pitch angles recognition.

| Test Sample Configuration | Norm Constraint | Noise Reduction | Parameter Learning Rate | Learning Rate Change Factor | Network Structure | Recognition Rate |
|--------------------------|-----------------|----------------|-------------------------|----------------------------|------------------|-----------------|
| 30 Degree Pitch          | $10^{-5}$       | 0.3            | 2.8                     | 1                          | 4096-100         | 92.64%          |
| 45 Degree Pitch          | $10^{-5}$       | 0.3            | 2.8                     | 1                          | 4096-100         | 76.04%          |

With the increase of the pitch angle of the test sample from 30° to 45°, the recognition rate of the algorithm in this article is reduced by 16.6%, which is slightly worse than that of MLA-LSR (13.51%) and HOG-SDDLSR (15.15%). However, it is better than the other algorithms which decrease 18.71%, 20.17%, 20.03% and 19.06% respectively. Although MLA-LSR and SAR-HOG-SDDLSR have better robustness under extended operation conditions, but require long training time. The algorithm in this article has both short training time and test time, and its performance is closed to optimal.

In order to further test the algorithm in this article, the relevant recognition parameter adjustment test is carried out for the pitch angle change of training samples, as shown in TABLE 11 and TABLE 12. As can be seen from the table, that the performance of this algorithm is the best when the pitch angle is 30° and the noise reduction parameter is 0.3, the learning rate is 2.8 and the change factor is 1.

VI. CONCLUSION

This article researches on a new SAR image target recognition strategy based on guided reconstruction and weighted norm-constrained deep belief Network. Firstly, the guided reconstruction algorithm can attain two-scale data enhancement on SAR images, generating and normalizing one-dimensional image vectors which can effectively reduce the dimension of image output features and improve the speed of preprocessing, and reducing the effect of coherent speckle noise on the model in SAR images. Then, the weighted norm constrained DBN is used to extract deep sparse features from SAR images. Under the regularization constraint of probability distribution of hidden layer, the joint probability distribution of samples is minimized in the visible layer and hidden layer, and the sparsity of low dimensional features is further improved on the basis of generalized optimization of norm constrained RBM; Finally, the regularized Softmax is used to classify the target, and the output results are obtained, so that the features of the previous layers are fully preserved in the high-level features used in the network classification. Compared with the classification performance of known models, the SAR target recognition algorithm in this article, on the basis of improving target recognition performance and generalization ability, can still reduce the output feature dimension and network training times. Therefore, the recognition performance of the algorithm will be further improved.

In this article, because DBN itself is an algorithm based on a probability model, the recognition rate under the fixed parameters is not the only one. Sometimes it is necessary to take the average multiple times, which cannot meet the requirements for the high performance and reliability recognition of SAR ATR. It is dependent on human experience in the process of setting parameters. In the future work, it should to study some adaptive parameter settings. According to the
characteristics of the sample library, it is necessary to search for and set the optimal parameters adaptively, and to find a better network structure so as to improve the performance of the SAR classification algorithm.

REFERENCES

[1] K. El-Darymli, E. W. Gill, P. Mcguire, D. Power, and C. Moloney, “Automatic target recognition in synthetic aperture radar imagery: A State-of-the-Art review,” IEEE Access, vol. 4, pp. 6014–6058, 2016.

[2] F. Gao, J. You, J. Wang, J. Sun, E. Yang, and H. Zhou, “A novel target detection method for SAR images based on shadow probability and saliency analysis,” Neurocomputing, vol. 267, pp. 220–231, Dec. 2017.

[3] Z. Zhang, H. Wang, F. Xu, and Y.-Q. Jin, “Complex-valued convolutional neural network and its application in polarimetric SAR image classification,” IEEE Trans. Geosci. Remote Sens., vol. 55, no. 12, pp. 7177–7188, Dec. 2017.

[4] H. Chen, F. Zhang, B. Tang, Q. Yin, and X. Sun, “Slim and efficient neural network design for resource-constrained SAR target recognition,” Remote Sens., vol. 10, no. 10, p. 1618, Oct. 2018.

[5] S. Chen, H. Wang, F. Xu, and Y. Q. Jin, “Target classification using the deep convolutional networks for SAR images,” IEEE Trans. Geosci. Remote Sens., vol. 54, no. 8, pp. 4806–4817, Aug. 2016.

[6] H. Wang, S. Li, Y. Zhou, and S. Chen, “SAR automatic target recognition using a rototranslational invariant wavelet-scattering convolution network,” Remote Sens., vol. 10, no. 4, p. 501, Mar. 2018.

[7] F. Zhang, Z. Fu, Y. Zhou, H. Xu, and W. Hong, “Multi-aspect SAR target recognition based on space-fixed and space-varying scattering feature joint learning,” Remote Sens. Lett., vol. 10, no. 10, pp. 998–1007, Oct. 2019.

[8] F. Zhang, C. Hu, Q. Yin, W. Li, H.-C. Li, and W. Hong, “Multi-Aspect-Aware bidirectional LSTM networks for synthetic aperture radar target recognition,” IEEE Access, vol. 5, pp. 26880–26891, 2017.

[9] Z. Huang, Z. Fan, and B. Lei, “Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data,” Remote Sens., vol. 9, no. 9, p. 907, Aug. 2017.

[10] F. Gao, Y. Yang, J. Wang, J. Sun, E. Yang, and H. Zhou, “A deep convolutional generative adversarial networks (DCGANs)-based semi-supervised method for object recognition in synthetic aperture radar (SAR) images,” Remote Sens., vol. 10, no. 6, p. 846, May 2018.

[11] Q.-C. Feng, D.-L. Pen, and G. Yu, “SAR target recognition with variants based on convolutional neural network,” J. Image Graph., vol. 24, no. 2, pp. 258–268, Feb. 2019.

[12] Z.-Y. Cui, “Theory and key techniques research on synthetic aperture radar target recognition,” Univ. Electron. Sci. Technol. China, Chengdu, China, Tech. Rep., 2015.

[13] S.-P. Xu, G.-Z. Zhang, and C. Li, “A fast random-valued impulse noise detection algorithm based on deep belief network,” J. Electron. Inf. Technol., vol. 41, no. 5, pp. 1130–1136, May 2019.

[14] S. Deng, L. Du, C. Li, J. Ding, and H. Liu, “SAR automatic target recognition based on Euclidean distance restricted autoencoder,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 10, no. 7, pp. 3323–3333, Jul. 2017.

[15] M. Kang, K. Jii, X. Leng, X. Xing, and H. Zou, “Synthetic aperture radar target recognition with feature fusion based on a stacked autoencoder,” Sensors, vol. 17, no. 1, p. 192, Jan. 2017.

[16] J. Ding, H.-G. Liu, B. Chen, B. Feng, and Y.-H. Wang, “Similarity constrained deep belief networks with application to SAR image target recognition,” J. Electron. Inf. Technol., vol. 38, no. 1, pp. 97–103, Jan. 2016.

[17] M. T. Yu, G. Y. Fan, and G. Y. Kuan, “SAR target recognition via local sparse representation of multi-manifold regularized low-rank approximation,” Remote Sens., vol. 10, no. 2, p. 211, Feb. 2018.

[18] S. Song, B. Xu, and J. Yang, “SAR target recognition via supervised discriminative dictionary learning and sparse representation of the SAR-HOG feature,” Remote Sens. Surv., vol. 8, no. 8, pp. 683, Aug. 2016.

[19] S. Li, X. Kang, and J. Hu, “Image fusion with guided filtering,” IEEE Trans. Image Process., vol. 22, no. 7, pp. 2864–2875, Jul. 2013.

[20] Z. Cui, S. S. Ge, Z. Cao, J. Yang, and H. Ren, “Analysis of different sparsity methods in constrained RBM for sparse representation in cognitive robotic perception,” J. Intell. Robotic Syst., vol. 80, no. 51, pp. 121–132, Dec. 2015.

[21] Z. Xu, X. Chang, F. Xu, and H. Zhang, “L1/2 regularization: A thresholding representation theory and a fast solver,” IEEE Trans. Neural Netw. Learn. Syst., vol. 23, no. 7, pp. 1013–1027, Jul. 2012.

[22] B. Ding and G. Wen, “Target recognition of SAR images based on multi-resolution representation,” Remote Sens. Lett., vol. 8, no. 11, pp. 1006–1014, Nov. 2017.

[23] J. Liang and R. Liu, “Stacked denoising autoencoder and dropout together to prevent overfitting in deep neural network,” in Proc. 8th Int. Congr. Image Signal Process. (CISP), Oct. 2015, pp. 697–701.

[24] P. Murugan and S. Durairaj, “Regularization and optimization strategies in deep convolutional neural network,” Tech. Rep., 2017.

[25] S. A. Wagner, “SAR ATR by a combination of convolutional neural network and support vector machines,” IEEE Trans. Aerosp. Electron. Syst., vol. 52, no. 6, pp. 2861–2872, Dec. 2016.