



# Image Target Recognition Based on Deep Learning

Yao Yao\*, Liqiang Han, Ben Fan, Dan Wang and Wei Fan

Beijing Institute of Space Mechanics and Electricity Beijing, China

\*Corresponding author: Yao Yao, Beijing Institute of Space Mechanics and Electricity Beijing, Beijing, China, Email: Yuny1101@163.com

Research Article

Volume 3 Issue 1

Received Date: March 28, 2025

Published Date: April 14, 2025

DOI: [10.23880/oaja-16000160](https://doi.org/10.23880/oaja-16000160)

## Abstract

Target recognition image is of great significance to the acquisition of ground and sea targets in the synthetic aperture radar (SAR) field. It has become a hot issue to realize automatic target detection and improve the accuracy of target recognition. In order to accurately obtain target information in images and solve the problem of over-fitting in deep neural network training, this study applied SAR image iterative denoising based on non-local adaptive dictionary to process SAR images, and constructed CNN network to extract SAR image features. Experimental results show that the proposed method can effectively improve the recognition accuracy of SAR images from sample data, and the recognition rate reaches 97.25% on MSTAR data sets.

**Keywords:** Deep Learning; Image Denoising; Object Identification

## Abbreviations

SAR: Synthetic Aperture Radar; LST1M: Long short-term Memory; MSTAR: Moving and Stationary Target Acquisition and Identification; OMP: Orthogonal Matching Pursuit; SVM: Support Vector Machine.

## Introduction

In the scene of remote sensing images, the characteristics of remote sensing images pose challenges and difficulties to the target recognition task. Target detection based on deep learning with Synthetic Aperture Radar (SAR) as the scene can improve the accuracy of target recognition [1]. Synthetic Aperture Radar (SAR) is an active microwave imaging radar system which is completely different from the traditional optical imaging system such as infrared and visible light. Early research on SAR image target recognition and classification focuses on feature extraction and classifier design. For example, Wang H, et al. [2] applied multi-neighborhood

orthogonal basis to achieve multi-stage filtering sampling of SAR images, and obtained the spatial scale of SAR multi-stage Gauss differential images and applied it to SAR image feature extraction [2]. Then, the multi-scale kernel function is used to map the image features of different levels by using the multi-scale kernel support vector machine (SVM) model, and finally synthesized to realize the target recognition of SAR images. Liu H, et al. [3] proposed a decision-making method based on sparse representation fusion support vector machine [3]. F/Sr-c classifier is used to classify and detect SAR images, and the positions of non-zero elements in SR coefficients are detected to identify and classify SAR images. SVM-C was used to extract PCA features from images. Finally, the features extracted by SR-C and SVM-C are fused to achieve SAR image target classification.

Wu T, et al. [4] based on the method of cascade decision fusion of SVM and Sparse representation classification (SRC) [4], firstly used SVM to classify images and obtain posterior probability of images, then used threshold decision method

to obtain category images with high confidence, and finally used SRC to classify SAR images. The decision values of SVM and SRC are combined to achieve SAR image target recognition and classification.

In recent years, the emergence of deep learning [5] has made breakthrough progress in image recognition research. AlexNet [6], VGG [7], GoogleNet [8], Res Net [9] and other neural network models appeared successively and showed their edge in ImageNet competition. Therefore, the method of deep learning must be introduced into the research of SAR image target recognition and classification.

In the study of deep learning, Hu X, et al. [10] proposed a CM Net network model based on SAR image recognition [10]. In this model, a small convolution kernel is used to design four convolutional pooling layers to complete feature extraction, and Softmax loss and center loss are used to jointly supervise network training, so as to improve the generalization ability of network model and improve the accuracy of SAR image recognition. However, in the face of SAR images with complex scenes, shallow neural network has limited learning ability, poor generalization ability and limited improvement in SAR image recognition rate. In the preliminary experiment of this paper, transfer learning is combined with the Inception resnet- V2 network model. The network parameters of pre-trained simulated SAR images are migrated as the initial parameters of the target network, and then the SAR images are trained using the Inception Resnet-v2 target network. The feature extraction of target image is realized. Finally, the SAR image is recognized and classified by Softmax classifier. In this method, Inception-Resnet-V2 model with deep network layer is used to train SAR images to obtain deeper image features. Transfer learning is used to improve model generalization ability and solve the problem of small samples. Compared with literature, this method has enhanced network learning ability and significantly improved the recognition rate of SAR images. However, deep network training pays more attention to the semantic information of images, and the details of images are seriously lost in the training process, which has a certain impact on improving the accuracy of SAR image recognition. Wang K, et al. [11] combined transfer learning with VGG16 network model to complete the extraction of target image features by migrating the pretraining model of VGG16 network. In this paper, VGG16 deep neural network is selected, but the problem of image details loss still exists in the training process.

Because the SAR target is very sensitive to the change of the observation azimuth, the MSTAR data set gives the SAR image target with different angles. Therefore, the multi-angle recognition of the SAR image target is the focus of the current SAR image recognition algorithm research, and the

recognition results of SAR images with different imaging angles are different. Literature proposed a bidirectional Long short-term Memory (LSTM) recurrent neural network to recognize SAR image targets from different imaging angles [12]. First of all, the method to sort the different imaging Angle of SAR image target of different target sequence, and then realize the Gabor filter and three local binary pattern, the space characteristics were extracted, and then through the multilayer perceptron dimension, in the end, through a two-way LSTM features a multi-angle recursive neural network fusion, The classifier is further integrated to realize target recognition. 10 kinds of targets were identified, and the accuracy reached 99.9%. In addition, its anti-noise and anti-confusion performance is better than traditional deep learning-based methods. In literature [13,14], three SAR images of the same target at different azimuth angles were selected, and the SAR images were input into the network as the RGB three channel images of color images for target recognition. For multi-angle SAR image target recognition, the feature extraction network in literature [15] is mainly composed of a single image feature can operate on-board and transmit images of specific targets identified to the ground, thereby reducing satellite-to-ground transmissions. In order to enable the model to run on the satellite, a deep compression method is used to compress the model. Finally, the identification and classification effect of the designed CNN network model is verified by experiments.

## Sar Model Based on Deep Convolutional Neural Network

### Convolutional Neural Network

Convolutional neural network is a supervised learning neural network with hierarchical model in deep learning. Its core structure includes convolution layer, pooling layer and full connection layer. Convolutional neural network uses convolutional layer and pooling layer to realize feature extraction. Compared with traditional machine learning, convolutional neural network can automatically extract target features during training. At the same time, the number of model parameters is greatly reduced, and the model generalization ability is improved.

As the core part of the convolution layer, the convolution kernel performs matrix transformation calculation through neurons and transmits feature information to the next layer to achieve feature extraction. Generally, there are multi-scale convolution kernels in the neural network model, and feature information can be fully extracted by feature extraction of multi-scale convolution kernels one by one. The convolution process is as follows: extraction EfficientNet network and a BiGRU network for further multi-angle SAR image sequence feature extraction. Firstly, the method uses a set of sharing-

weight EfficientNet network to extract the spatial features (B, L, U) of a single image in the image sequence, in which B is the step length of model training, L is the number of images in the image sequence, and U is the dimension of spatial features of a single image. Then, feature dimensions (B, L, U) were obtained by dimensionally transforming the image feature sequences extracted from the image sequences, which were then sent into BiGRU network to extract the sequential features (B, V) of multi-angle SAR image sequences, where, V is the dimension of feature. Finally, the features extracted by BiGRU network are sent to the full connection layer to obtain the final output and get the classification category.

The recognition accuracy of this method is significantly higher than other multi-angle SAR image target recognition methods in MSTAR data. Literature [16] established an improved pooled CNN model for SAR target Angle sensitivity, which can improve the recognition performance of convolutional neural network at different azimuth angles without affecting the algorithm complexity.

$$Y_j^l = f \left( \sum_{i \in N_j} X_i^{l-1} \times M_{ij}^l + b_j^l \right) \quad (1)$$

Where,  $Y_j^l$  represents the j-th feature maps of l layer; M stands for convolution kernel; b represents the off-set term; f represents nonlinear activation of feature information.

This paper mainly studies the application of CNN in the field of SAR image target recognition, and designs a SAR image target recognition system based on CNN to realize the intelligent recognition of ground targets.

When feature extraction is completed at the convolutional layer, the extracted feature information is transmitted to the pooling layer to further reduce the feature matrix from the convolutional layer and optimize the extraction of feature information. At the same time, it can effectively reduce the feature dimension and reduce the computation of network model.

As the core structure of convolutional neural network, the full connection layer is equivalent to the "classifier" of neural network. The full connection layer distinguishes the learned feature information by weighted summation of the feature information learned from the model and mapping it to the label data in the sample space.

## SAR Image Denoising

As the SAR imaging system is affected by coherent radiation, the generated images have large speckle noise, and the processing and interpretation of SAR images have serious interference, which increases the difficulty of convolutional

neural network training to a certain extent and is not conducive to the maximum utilization of the performance of convolutional neural network. Therefore, K-SVD similar block matching and adaptive dictionary updating are used to filter and de-noise SAR images to minimize the impact of noise on neural network model learning process and ensure optimal network performance.

K-SVD algorithm is an image denoising algorithm based on super-complete sparse decomposition [17]. In this algorithm, the original over-complete dictionary is trained under the image sample through iterative process, and the atoms in the atom library are constantly adjusted by sparse decomposition coefficient, and finally the over-complete matrix which can reflect the image features more effectively is obtained. Assuming that the over-complete dictionary is  $D \in \mathbb{R}^{M \times k}$  and the noisy image  $I \in \mathbb{R}^{M \times N}$  is processed by sparse decomposition, the redundant sparse model is as follows:

$$\hat{\alpha} = \arg \min \|\alpha\|_0 \text{ s.t. } \|D\alpha - I\|_2^2 \leq \varepsilon \quad (2)$$

Where,  $\alpha$  is the redundant sparse representation coefficient of the image. In general,  $\|\alpha\|_0 \leq L \ll M$ , where L is the largest sparsity. Under certain  $\varepsilon$ ,  $\alpha$  can be obtained by (2).

In order to remove noise more effectively [18], (2) can be converted into:

$$\hat{\alpha} = \arg \min \|\alpha\|_0 \text{ s.t. } \|D\alpha - I\|_2^2 \leq T \quad (3)$$

Where, T is the threshold dependent on  $\varepsilon$  and  $\sigma$ , while  $\sigma$  is the standard deviation of noise contained in the image.

In order to better use orthogonal matching pursuit (OMP) algorithm to solve the sparse model of image block, transform (3) for the following models:

$$\hat{\alpha} = \arg \min_{\alpha} \|D\alpha - I\|_2^2 + \mu \|\alpha\|_0 \quad (4)$$

Where  $\mu$  is the penalty factor. In practice, noisy images usually satisfy the model  $Y = X + n$ . Where Y is the actual observed image containing noise, X is the ideal image without noise, and n is noise.

Since dictionary training is only effective for small image blocks, if the original noisy image is used for dictionary training directly, the sparsity of the image will be affected to some extent and the optimal sparse representation of the image cannot be obtained [9]. Therefore, the image can be partitioned to obtain small image blocks first, and then similar block groups can be obtained by non-local similar block matching, and the similar block groups can be processed by K-SVD algorithm based on OMP, which can not only obtain the optimal sparse representation of the image, but also improve the efficiency and accuracy of the algorithm.

In nonlocal similar block matching are used to get the similar block group, set the size of the reference block of Yi for  $\sqrt{m} \times \sqrt{m}$ , within the scope of region  $D \times D$  searching, the Euclidean distance between the matching block y and the reference block Yi is calculated and sorted, Q before taking an image block (average Q value of 7) similar to the reference blocks block set Yi.

If the noiseless image X satisfies the sparse model of Equation (3), the denoising model of the image can be obtained as:

$$\{\hat{\alpha}_{ij}, \hat{X}\} = \arg \min_{\alpha_{ij}, X} \lambda \|X - Y\|_2^2 + \sum_{i,j} \mu_{ij} \|\alpha_{ij}\|_0 + \sum_{i,j} \|D\alpha_{ij} - R_{ij}X\|_2^2 \quad (5)$$

Where,  $\lambda \|X - Y\|$  represents the approximation degree between the noisy image Y and the noise-free image X;  $\sum_{ij} \mu_{ij} \|\alpha_{ij}\|_0$  and  $\sum_{ij} \|D\alpha_{ij} - R_{ij}X\|_2^2$  are the prior conditions of sparsity and image decomposition consistency respectively.

In order to overcome edge blurring in image denoising, the dictionary update part is integrated into Bayesian denoising [8], thus the objective function of K-SVD algorithm can be obtained:

$$\{\hat{D}, \hat{\alpha}_{ij}, \hat{X}\} = \arg \min_{D, \alpha_{ij}, X} \lambda \|X - Y\|_2^2 + \sum_{i,j} \mu_{ij} \|\alpha_{ij}\|_0 + \sum_{i,j} \|D_{ij} - R_{ij}X\|_2^2 \quad (6)$$

The model can be easily solved by alternating iteration algorithm.

### SAR Model Design Based on Deep Convolutional Neural Network

The deep CNN architecture designed in this paper consists of two convolution layers, two downsampling layers, three Dropout layers, one Flatten layer, and two full connection layers, which can achieve target classification and recognition in SAR images.

- **Data Input Layer:** The input layer directly input the pixel value of the image, and the image size is  $1 \times 128 \times 128$ .
- **Conv1 Convolution Layer:** Thirty-two different convolution kernels, each with a size of  $3 \times 3$ , were adopted, that is, Conv1 convolution layer selected and learned features from the input data, and finally output-32 feature planes. Since each convolution kernel needs to train  $3 \times 3 = 9$  parameters and 1 bias parameter, a total of  $(3 \times 3 + 1) \times 32 = 320$  parameters need to be learned for 32 convolution kernels. The size of the experimental input image is  $128 \times 128$ , the sliding step of the convolution kernel convolution operation is 2 pixels, and the mapping size of the feature plane after convolution is  $((128 - 3 + 2)/2) \times ((128 - 3 + 2)/2) = 63$

$\times 63$ .

- **The Lower Sampling Layer of S1:** 32 feature planes with a size of  $63 \times 63$ , which were extracted from Conv1, were sampled at the maximum value without overlap, and the sampling area was  $2 \times 12$ . Therefore, the output of S1 layer was  $32 ((63/2) + 1) \times ((63/2) + 1) = 32 \times 32$  feature planes after dimensionality reduction, reducing the data volume to about a quarter of the original.
- **The Dropout Layer:** Due to the complex network structure and insufficient number of training samples, the deep network often has the phenomenon of overlearning and fitting. Researchers have come up with a new way of thinking about Dropout. Dropout means that during each training of the network model, the probability p is randomly output as 0 ( $p = 0.2$  in this layer), so that the weights of some nodes in the hidden layer of the network do not participate in the training. The weights retain the results of previous training and do not update temporarily. The physical significance of Dropout is equivalent to the disconnection of some nodes in the network. These disconnected nodes are also ignored in backpropagation learning.
- **Conv2 Convolution Layer S2 Undersampling Layer and Dropout Layer:** Further abstracting and generalizing the features of image data.
- **Flatten Full Connection Layer:** The Flatten layer is used to "Flatten" the input, that is, one-dimensional the multidimensional input, and is often used for the transition from the convolution layer to the fully connected layer
- **Dense Fully Connected Layer:** There are 512 nerve nodes in this fully connected layer, and each nerve node is connected to all nerve nodes in the upper layer. The output of each neuron can be expressed as follows:

$$y_j = \max\left(0, \sum x_i \cdot w_{i,j} + b_j\right) \quad (7)$$

- Where, xi is the value of neurons input at the full connection layer;  $w_{i,j}$  is the weight;  $b_j$  is the off-Set. ReLU function is still used as the activation function. The number of neurons in the full connection layer directly affects the fitting effect and training speed of the network.
- **Dropout Layer:** To further prevent overfitting, add a layer of Dropout,  $p = 0.5$ .
- **Dense Output Layer:** Using Softmax multi-classifier, The output value is between 0 and 1, representing the probability that the sample belongs to a certain type of label. In this paper, there are 11 types of ground object targets, so the expression of 11 Softmax neurons in the output layer is as follows:

$$p_i = \frac{\exp(y_i)}{\sum_{j=1}^n \exp(y_j)} \quad (8)$$

Where, yi is the result of the output of the i-th neuron; n

is the number of divine elements.

### Residual Network Module with Convolution Attention

The convolutional attention module filters the important features of the image in channel and space, so that the network model introduced by CBAM can extract the main features of the target more accurately. In this paper, CBAM is applied to the residual network. Four convolution attention modules are added to the residual nodes of ResNet101 network, and the image features learned by each residual module are analyzed. By assigning different weights to the feature images, the network is guided to extract the key feature information of the target image in the training process. Thus, the network model can improve the feature expression ability of SAR image targets.

## Experimental Setup and Result Analysis

### The SAR Data Sets

The experimental data set was based on the Moving and Stationary Target Acquisition and Identification (MSTAR) remote Sensing SAR database. This dataset uses a high-resolution spotlight synthetic aperture radar (SAR) with a resolution of 0.3 meters to collect SAR images of various vehicle targets. It operates in the X-band and uses the HH polarization mode. The MSTAR dataset is mainly composed of static SAR vehicle target images. In order to facilitate the study of the influence of different imaging angles on the recognition algorithm, this dataset includes vehicle target images from different imaging angles. To facilitate the training and prediction of the recognition algorithm, the dataset will process the collected data, extract target slice data with a pixel size of  $128 \times 128$ , and divide the data into a training set and a test set. The samples in the training set mainly include SAR images of three different types of vehicles with an imaging angle of  $17^\circ$ . The test set mainly covers three types of targets with an imaging angle of  $15^\circ$ , and the effectiveness of the recognition algorithm is verified through different imaging angles. The MSTAR dataset contains a batch of mixed target slice data. In terms of categories, the mixed target slice data mainly includes various types of wheeled vehicles, tracked vehicles, etc. These targets provide SAR images from different imaging angles, which are mainly used to carry out research on target recognition algorithms for SAR images at different imaging angles.

### Image Data Preprocessing

Since the SAR image training samples of MSTAR data set have different image resolutions and sizes and are insufficient in number, they cannot be directly used for deep convolutional neural network learning and are prone

to over-fitting. Therefore, data preprocessing and training sample base expansion are firstly carried out in this paper.

- **Standardized Data Set:** In order to ensure the consistency of image size and resolution, images in JPEG format with 95% resolution are extracted from the native Sun Floating point format, and the unified image size is  $128 \times 128$ . In the experiment, images of similar objects at two azimuth angles were shuffled, and 80% of each category was randomly selected as the training set and the remaining 20% as the test set.
- **Real-Time Data Enhancement:** During the training and learning of network model, real-time data promotion was carried out, including the Angle of random rotation of picture, the amplitude of horizontal deviation of picture, the amplitude of vertical deviation of picture, the shear transformation of counterclockwise direction, the amplitude of random scaling of Angle, etc.
- **SAR Image Simulation:** Simulated SAR images obtain SAR image sweep data through electromagnetic scattering model and scene model, and then obtain by combining time-frequency transformation and imaging algorithm [19]. The geometric relationship between ground and target scene is obtained by ray tracing method, and 3D SAR scene model is preliminarily established. Rough surface scattering theory and other techniques are used to model the rough characteristics of real scenes. The electromagnetic scattering model of ground and target scene is established by using ray bounce method and other methods, and the sweep frequency data of target in SAR image is obtained. Finally, the simulated SAR target image is obtained by time-frequency transform and imaging technology.

ImageNet data sets are used for pre-training in previous SAR image recognition studies. However, the similarity between image features in ImageNet data sets and SAR image features is not high, and the network parameters trained are not suitable for the initialization parameters of the training SAR image network model. Compared with SAR images provided by MSTAR dataset, simulated SAR images have higher feature similarity and smaller speckle noise, and are more suitable for pre-training samples than ImageNet images.

The SAR simulation dataset is used in the pre-training process [17], which contains seven SAR target images in multiple scenarios, totaling 22647 simulated SAR images.

## Experimental Result

The MSTAR training set and test set used in the experiment are shown in Table 1. In the recognition experiment of 11 kinds of targets. The training samples are

SAR image data of the variants at a 17° elevation, and the test samples are SAR image data of the targets above mentioned

variants at a 15° elevation.

Category	Model	Training set Angle	Test set Angle
Type I	9563	17°	15°
	9566	/	15°
	C21	/	15°
Type IIa	132	17°	15°
Type IIb	812	/	15°
Type IIc	S7	/	15°
Type III	C71	17°	15°
Type IV	k10yt7532	17°	15°
Type V	b01	17°	15°
Type VI	E71	17°	15°
Type VII	92VL13015	17°	15°
Type VIII	A51	17°	15°
Type IX	E12	17°	15°
Type X	d08	17°	15°
Type XI	/	17°	15°

**Table 1:** Sample Distribution of Training Set and Test Set in Mstar Standard Dataset.

Test objective	Identify Results											Accuracy/%
	Type I	Type II	Type V	Type VI	Type III	Type IV	Type VII	Type XI	Type VIII	Type IX	Type X	
Type Ia	191	1	1	0	0	0	0	0	0	0	0	98.45
Type Ib	170	19	0	0	3	0	0	0	0	0	0	89.95
Type Ic	168	13	0	0	1	11	0	0	0	0	0	86.57
Type IIa	1	121	0	0	0	0	0	2	0	0	0	98.53
Type IIb	6	195	0	0	2	0	0	0	0	0	0	96.06
Type IIc	7	183	0	0	1	0	0	0	0	0	0	95.31
Type V	1	0	286	0	0	0	1	0	0	1	0	98.96
Type VI	2	0	0	279	0	0	0	0	0	0	1	98.59
Type III	2	0	0	0	192	0	0	0	0	0	0	98.97
Type IV	0	1	0	0	12	276	0	0	0	0	0	95.17
Type VII	0	0	0	1	0	0	269	0	0	0	0	99.62
Type XI	0	0	0	0	0	0	0	270	0	0	0	100
Type VIII	0	0	0	0	0	0	0	0	271	0	5	98.18
Type IX	0	0	0	0	0	0	0	0	0	276	0	100
Type X	0	0	1	0	0	0	0	0	0	0	273	99.63
Overall recognition rate	97.25%											

**Table 2:** Mstar11 Sar Target Recognition Result.

The test classification results of the deep convolutional network model in this paper are shown in Table 2. As can be seen from Table 2, SAR image target classification results are good, with an average recognition rate of 97.25%, among them, the recognition rates of six types, are all greater than 99%. In addition, the SAR targets of the same type and different variants are also recognized well. Based on the convolutional neural network modeling method, the same MSTAR training set and test set were used for experimental testing, and the average recognition rate was 93.76%. It can be seen that the deep convolutional neural network method proposed in this paper is superior to the benchmark method to some extent, and the SAR target recognition rate is higher.

### Limitations of the Model and Directions for Improvement

The deep convolutional network model in this paper, after being trained with simulated SAR images obtained through preprocessing of the MSTAR dataset and expansion of the training sample library, has a relatively high recognition rate for targets based on the features of MSTAR targets. However, the result of such training will lead to a certain degree of dependence on this feature basis. When training with data having different imaging parameters and target features, the recognition accuracy may decrease. In order to avoid this problem, the best approach for future work is to perform intermediate feature fusion of SAR image data and data from other types of sensors at specific layers of the deep neural network, so as to improve the overall recognition rate by capturing complementary information.

In addition, the current research is based on the recognition rate results obtained by training the dataset on the ground. In the future, with the improvement of the computing power of on-board satellite equipment and the in-depth application of artificial intelligence, methods such as model lightweight design, including removing redundant parameters through pruning algorithms, can enable the model to meet the real-time operation requirements of on-board satellite equipment, reduce resource limitations, and satisfy higher real-time requirements.

### Conclusion

Aiming at the problem that the timeliness of target recognition in satellite SAR images cannot meet the rapid application, this paper proposes a fast and intelligent recognition of targets in SAR images using deep convolutional neural network. Simulation results show that the proposed method can quickly realize the intelligent recognition of spaceborne SAR image targets without manual intervention. Compared with existing deep neural network methods, this method effectively solves the technical problem of serious

loss of detail features of SAR images with small samples trained by deep neural network, provides theoretical and technical reference for the research and further application of SAR image target recognition, and finally realizes the purpose of real-time recognition and application of on-board imaging data. Provide effective technical support. The experimental results show that the proposed model can improve the target recognition rate obviously and has good recognition and classification effect.

### References

1. Dongxing Y, Guangyi S, Yukun Z (2024) Research Progress on High-Resolution Remote Sensing Image Scene Classification[J]. Spacecraft Recovery & Remote Sensing 45(4): 124-138.
2. Wang H, Sun F, Cai Y, Chen N, Pei D (2011) Sar image automatic target recognition based on local multi-resolution features[J]. Journal of Tsinghua University Science and Technology 51(8): 1049-1054.
3. Liu H, Li S (2013) Decision fusion of sparse representation and support vector machine for sar image target recognition[J]. Neurocomputing 113: 97-104.
4. Wu T, Xia J, Huang Y, et al. (2020) Target recognition method of sar images based on cascade decision fusion of svm and src[J]. Journal of Henan Polytechnic University (Natural Science) 39(4) 118-124.
5. Hinton GE, Osindero S, Teh YW (2006) A fast learning algorithm for deep belief nets[J]. Neural computation 18(7): 1527-1554.
6. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks[C]. Communications of the ACM 60(6): 84-90.
7. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
8. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, et al. (2015) Going deeper with convolutions[C]. IEEE conference on computer vision and pattern recognition, pp: 1-9.
9. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition[C]. IEEE conference on computer vision and pattern recognition, pp: 770-778.
10. Hu X, Yao Q, Hou B, Song H, Lei H (2019) Target recognition using convolution neural network for sar images[J]. Science Technology and Engineering 19(21): 228-232.
11. Wang K, Zhang G, Leung H (2019) Sar target recognition

- based on cross-domain and cross-task transfer learning[J]. IEEE Access 7: 153391-153399.
12. Zhang F, Hu C, Yin Q, Li W, Li H.-C, et al. (2017) Multi-aspect-aware bidirectional LSTM networks for synthetic aperture radar target recognition[J]. Ieee Access 5: 26880-26891.
  13. Zou H, Lin Y, Hong W (2018) Research on multi-aspect sar images target recognition using deep learning[J]. Journal of Signal Processing 34(5): 512-522.
  14. Pei J, Huang Y, Huo W, Zhang Y, Yang J, et al. (2017) Sar automatic target recognition based on multiview deep learning framework[J]. IEEE Transactions on Geoscience and Remote Sensing 56(4): 2196-2210.
  15. Zhao P, Huang L (2020) Multi-aspect sar target recognition based on efficientnet and gru[C]. IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium, pp: 1651-1654.
  16. Zhang F, Liu Y, Zhou Y, Yin Q, Li HC (2020) A lossless lightweight cnn design for sar target recognition[J]. Remote Sensing Letters 11(5): 485-494.
  17. Deka B, Bora PK (2013) Removal of correlated speckle noise using sparse and overcomplete representations[J]. Biomedical Signal Processing and Control 8(6): 520-533.
  18. Elad M, Aharon M (2006) Image denoising via sparse and redundant representations over learned dictionaries[J]. IEEE Transactions on Image processing 15(12): 3736-3745.
  19. Goodman JW (1976) Some fundamental properties of speckle[J]. JOSA 66(11): 1145-1150.