A Novel Few-Shot Learning Method for Synthetic Aperture Radar Image Recognition

Zhenyu Yue [a], Fei Gao [a, *], Qingxu Xiong [a], Jinping Sun [a], Amir Hussain [b], Huiyu Zhou [c]

[a] School of Electronic and Information Engineering, Beihang University, Beijing 100191, China

[b] Cyber and Big Data Research Laboratory, Edinburgh Napier University, EH11 4BN Edinburgh, U.K.

[c] Department of Informatics, University of Leicester, LE1 7RH Leicester, U.K.

[*] Corresponding author

Abstract

Synthetic aperture radar (SAR) image recognition is an important stage of SAR image interpretation. The standard convolutional neural network (CNN) has been successfully applied in the SAR image recognition due to its powerful feature extraction capability. Nevertheless, the CNN requires numerous labeled samples for satisfactory recognition performance, while the performance of the CNN decreases greatly with insufficient labeled samples. Aiming at improving the SAR image recognition accuracy with a small number of labeled samples, a new few-shot learning method is proposed in this paper. We first utilize the attention prototypical network (APN) to calculate the average features of the support images from each category, which are adopted as the prototypes. Afterwards, the feature extraction is performed on the query images using the attention convolutional neural network (ACNN). Finally, the feature matching classifier (FMC) is adopted for calculating the similarity scores between the feature maps and the prototypes. We embed the attention model SENet to the APN, ACNN, and FMC, which effectively enhances the expression of the prototypes and the feature maps. Besides, the loss function of our method consists of cross-entropy and prototype-separability losses. In the training process, this loss function increases the separability of different prototypes, which contributes to higher recognition accuracy. We perform experiments on the Moving and Stationary Target Acquisition and Recognition (MSTAR) and the Vehicle and Aircraft (VA) datasets. It has been proved that our method is superior to the related state-of-the-art few-shot image recognition methods.

Keywords: synthetic aperture radar, image recognition, few-shot learning

1 Introduction

Synthetic aperture radar (SAR) can acquire high resolution images all day and all weather by

virtue of its penetrating capability [1, 2]. SAR image recognition, which is an important stage in the SAR image interpretation, aims at obtaining the target information contained in the SAR images [3, 4]. Traditional machine learning methods for SAR image recognition include support vector machine (SVM) [5], dictionary learning [6], sparse representation [7], and so on. Nevertheless, the performance of these methods depends on the quality of hand-crafted features. In addition, the speckle noise in the SAR images increases the difficulty of feature extraction [8, 9]. Thus, an effective feature extraction method is the key to SAR image recognition.

Recently, deep learning techniques have received great attention in the field of image processing [10-12]. As an important deep learning model, convolutional neural network (CNN) is widely used because of its superior feature extraction capabilities [13, 14]. CNN is a multi-layer network which mainly consists of convolution, pooling, and fully connected layers. The convolution layer adopts the convolution kernels to extract image features, while the pooling layer reduces the dimension of the features through pooling operation, thereby reducing the risk of overfitting [15, 16]. The image features are further integrated by the fully connected layer and finally mapped to the predicted labels. In comparison with the traditional methods, CNN is a more effective feature extraction method, and it has been applied to SAR image recognition [17-19].

The CNN needs numerous labeled samples to train the model parameters. However, the sample annotation for SAR images is difficult because of the speckle noise. With insufficient labeled samples, it is difficult for the CNN model to achieve satisfactory recognition accuracy [20, 21]. On the basis of prior knowledge, humans can correctly classify images by learning a small number of labeled samples [22]. Inspired by this, researchers propose few-shot learning methods [23, 24]. In few-shot learning tasks, the sample categories are divided to the basic and new categories, wherein the former contains a large number of labeled samples while the latter only contains limited labeled samples. The goal of few-shot learning is to correctly recognize the unlabeled samples of the new categories. Nevertheless, the labeled samples in the new categories are insufficient. Therefore, few-shot learning models utilize the numerous labeled samples in the basic category during the training process. Once having been trained, these models can be generalized to the new categories and used to recognize the corresponding unlabeled samples [25-27].

To improve the SAR image recognition accuracy with limited labeled data, this paper presents a new few-shot learning method which consists of the attention prototypical network (APN),

attention convolutional neural network (ACNN), and feature matching classifier (FMC). We first utilize the APN to generate prototypes. The prototypes represent the cluster centers for different categories, which are obtained by calculating the average features of the images from the same category. Afterwards, the feature maps of the query images are extracted by the ACNN. Finally, we adopt the FMC to generate the similarity scores between the feature maps and the prototypes. The innovations of our method are:

(1) We design the FMC for generating the similarity scores. The FMC is a neural network model with trainable parameters, thus it is capable of learning an appropriate similarity metric function for SAR data in a data-driven way.

(2) The APN, ACNN, and FMC are embedded with the SENet. As an attention model, SENet is capable of assigning weights for the feature maps, which enhances the representation of the prototypes and feature maps.

(3) By combining the cross-entropy and the prototype-separability losses, we design a new loss function which is capable of increasing the separability of different prototypes. Since the recognition process is performed by calculating the similarity scores between the feature maps and the prototypes, this loss function can effectively improve the recognition performance of our method.

The structure of the rest of our paper is as follows. Section 2 reviews the related works. Section 3 introduces the detailed principles of our method. In Section 4, we verify the effectiveness of our method by performing abundant experiments. Finally, this paper is summarized in Section 5.

2 Related Work

2.1 Few-shot Learning Methods

Recently, few-shot learning methods have become the research hotspot in the image recognition. Koch et al. present the siamese neural network, which is capable of discriminating the similarities between the input images [28]. The siamese neural network consists of two CNN models with shared parameters to simultaneously extract the features of two input images. Then the distance function is adopted for calculating the similarities between the two feature maps, thereby judging whether the two images are from the same category. The matching network is composed of a neural-network-based feature extractor and a distance-function-based classifier [29]. In the training process, the matching network employs the episodic training mechanism, which ensures that the conditions of training and testing processes are matched. In addition, the attention and memory mechanisms

are utilized to enable rapid learning. The prototypical network proposed in [30] is a simple yet effective few-shot learning method. This method first utilizes the standard CNN model to extract image features, then the prototypes of different categories are obtained using the feature averaging operation. Finally, the classification of the query images is performed by comparing the Euclidean distances between the query images features and different prototypes. Sung et al. design the relation network algorithm which contains an embedding model for feature extraction and a relation model for feature similarity comparison [31]. In contrast with the distance-function-based similarity comparison methods, the relation model consists of neural network, which enables an end-to-end training scheme. Different from other few-shot learning methods, deep nearest neighbor neural network (DN4) adopts the local descriptors instead of the image-level features [32]. Following the descriptor extraction, the KNN method is utilized for calculating the similarities of different images by measuring the cosine distance between the local descriptors. Simon et al. design the deep subspace networks (DSN) wherein the limited samples are used for constructing dynamic classifiers [33]. The experiments demonstrate that DSN yields competitive results and is robust against perturbations. Ye et al. modify the feature extraction network via the self-attention mechanism and present the few-shot embedding adaptation transformer (FEAT) method wherein the extracted image features are discriminative and task-specific [34].

2.2 Few-shot Learning in SAR image recognition

Inspired by the superiority of the above few-shot learning methods, the researchers have applied them to SAR image recognition with limited labeled samples. Tang et al. present a modified siamese network which contains a CNN model, a similarity discriminator and a classifier [35]. Compared with the original siamese network, the modified method adopts a classifier instead of a similarity discriminator to generate label predictions. The hybrid inference network in [36] first uses the embedding model to extract image features, then the transductive and inductive inferences are respectively utilized to classify the input images. Finally, the two classification results are combined to obtain the final result. Fu et al. design a new meta-learning method consisting of a meta-learner and a base-learner [37]. During the training process, three transfer learning strategies are used for exploiting the prior information. Besides, the difficult task mining strategy is adopted so that the proposed method pays more attention to difficult tasks in the training process, thereby promotes the efficiency of meta-learning. In [38], Yang et al. improve the performance of relation network by

replacing the relation model with a graph neural network (GNN). Because of the strong relation extraction ability of GNN, the modified network can effectively measure the relationship between the input images.

As can be seen, most of the few-shot image recognition methods consist of two stages: feature extraction and classification, which are the major concerns of these methods. Although the above few-shot image recognition methods have made some promising results, most of these methods utilize the distance-based similarity metric functions in the classification stage. Since the parameters in the distance-based functions are fixed, it is necessary to manually select appropriate distance functions for different types of data. To automatically learn a suitable similarity metric function for SAR data, we design the FMC to replace the distance-based functions. The FMC consists of neural network, which is capable of learning a suitable similarity metric in a data-driven way. In addition, compared with the above methods, we embed the SENet in our model to enhance the representation of image features in the feature extraction stage. Moreover, our method follows the thought of prototypical network [30]. A key novelty is that we present the prototype-separability loss and combine it with the cross-entropy loss in the training process. This new loss contributes to a higher recognition accuracy by increasing the separability between different prototypes.

3 Proposed method

3.1 Few-Shot Learning Formulation

The few-shot image recognition task contains three sets: a support set S, a query set Q, and an auxiliary set A. S and Q consists of the samples of new categories, while A consists of the samples of basic categories. S includes N categories, and each category involves K labeled images. Therefore, the few-shot recognition task is named "N-way K-shot". The images in Q are the unlabeled ones which share the same category space with S. A contains more categories with numerous labeled images. However, the category space of A is non-intersecting with those of S and Q.

The few-shot recognition methods aim at classifying the unlabeled images in Q based on the learning with S. Since the labeled images in S are insufficient, the set A is utilized in the training process. To guarantee that the conditions in the training and testing processes are matched, many few-shot learning methods employ the episodic training mechanism. To be specific, during

each training episode, the auxiliary support set A_S (similar to set S) and the auxiliary query set A_Q (similar to set Q) are formed by randomly selecting samples from set A. Afterwards, the parameters in the few-shot learning models are updated through the forward and backward propagation using the images in A_S and A_Q.

3.2 Overall Framework

The training and testing processes of the proposed method are shown in Figure 1. We utilize set A to learn a matching method for the support and query images, then the trained model is applied to recognize the query images of new categories in set Q based on a small number of support images in set S.
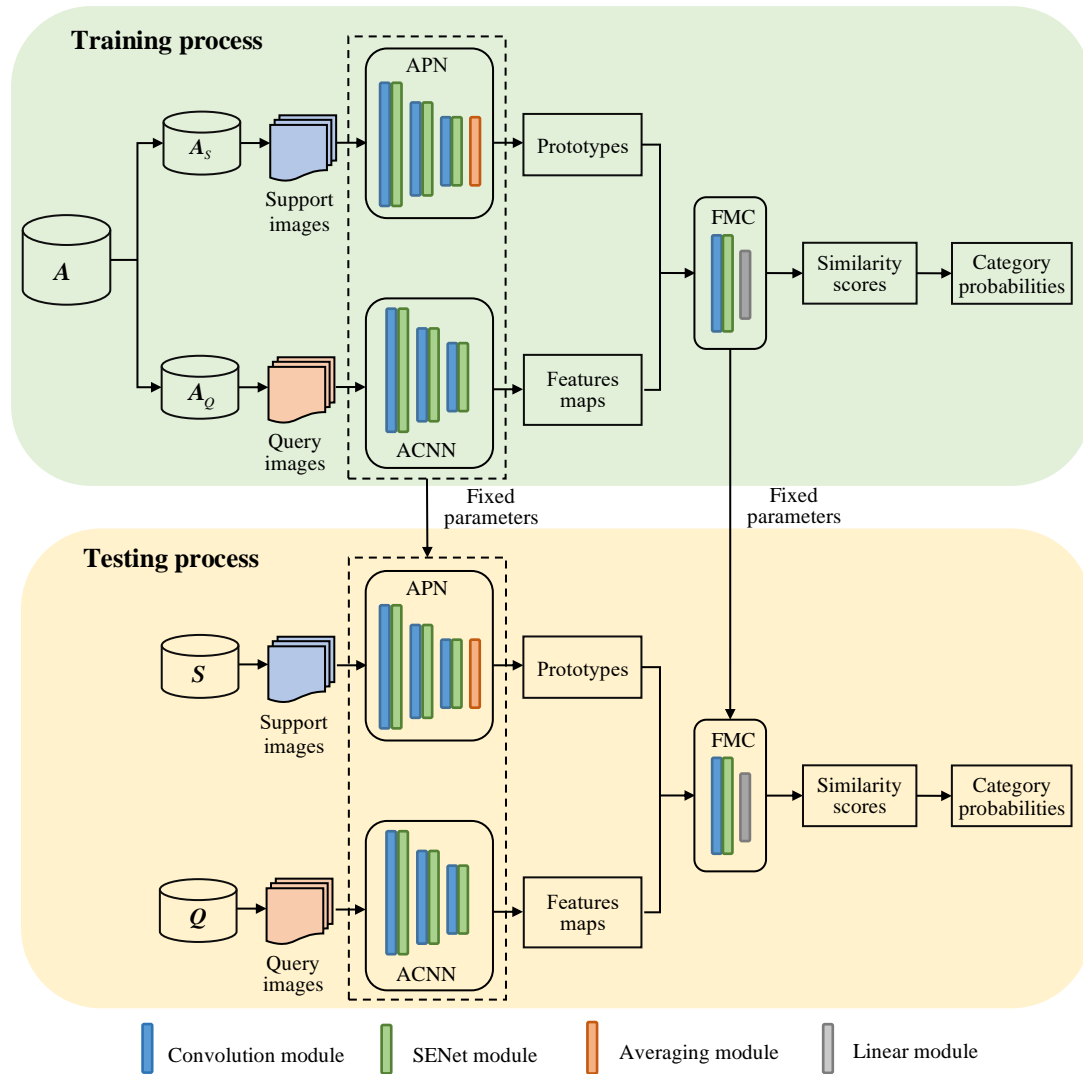

Figure 1. Training and testing processes of the proposed few-shot SAR image recognition method.

In the training process, we adopt the episodic training mechanism. During each training episode,

the images in set A are randomly selected to form sets A_S and A_Q. Then the support images in

A_S are fed into the APN to generate the prototypes of different categories. The APN consists of

convolution, SENet, and averaging modules. Afterwards, the ACNN model extracts the features of

the query images in A_Q based on the convolution and SENet modules. As can be seen, the APN

(except for the averaging module) and ACNN are both feature extractors which are utilized for

extracting the features of SAR images. Hence, we utilize the weight-sharing strategy between APN

(except for the averaging module) and ACNN. Finally, the FMC is utilized for calculating the

similarity scores between the extracted feature maps and different prototypes. The category

probabilities of query images are obtained by performing the softmax operation on the similarity

scores. The model parameters are updated by minimizing the loss function which includes cross-

entropy and prototype-separability losses.

In the testing process, we utilize the support images in S to calculate the prototypes of the

new categories. Then the ACNN is used for extracting the feature maps of the query images in Q.

Finally, the similarity scores between the feature maps and prototypes are obtained by the FMC. We

calculate the category probabilities of the query images using the softmax operation, and the

category corresponding to the maximum probability is adopted as the predicted category of the

query images. Next, the APN, FMC and loss function of our method are described in detail.

3.3 Attention Prototypical Network

We design the APN model to calculate the prototypes. As shown in Figure 1, the APN consists

of convolution, SENet, and averaging modules. The convolution module adopts convolution and

pooling operations to extract the image features. The SENet module is utilized for optimizing the

features obtained by the convolution module. In the averaging module, we respectively calculate

the average features of the images from the same category, thereby obtaining the prototypes. The

detailed process is as follows:

Suppose S_i contains all the support images whose labels are i, which is the subset of S. In

the prototype calculation, the images in S_i are fed into convolution modules for extracting the

feature maps $F \in \mathbb{R}^{C \times H \times W}$, where C, H and W represent the channel number, height, and width

of F. Besides, we adopt the attention models for feature optimization. In the feature extraction,

the attention models calculate the attention values which are used for assigning the weights for the feature maps, thus the important features are emphasized while the unimportant ones are restrained [39, 40]. As a light-weight attention model, the SENet optimizes the image features without increasing the computational complexity [41]. Therefore, we employ the SENet and embed it in the APN model. The core idea of SENet is that the importance of the features in different channels is different. Thus, the feature maps can be optimized by assigning weights for different channels. The operations performed in the SENet are as follows:

As shown in Figure 2, the SENet adopts global pooling module to shrink F in the spatial dimensions $H \times W$, then the global pooling feature map $F_{gp} \in \mathbb{R}^{C}$ is obtained. $F_{gp}(n) \in F_{gp}$ is calculated using Equation (1):

$$F_{gp}(n) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} F(n,i,j) \tag{1}$$

where $F(n,i,j) \in F$. Afterwards, F_{gp} is forwarded to the linear module for generating the attention value $F_{A} \in \mathbb{R}^{C}$. Finally, the weights for different channels in the original feature map F are assigned by F_{A}. Concretely, we multiply F with F_{A} in the channel dimension:

$$F' = F \otimes F_{A} \tag{2}$$

where $\otimes$ is the multiplication operation and $F' \in \mathbb{R}^{C \times H \times W}$ denotes the optimized feature map. As can be seen, the SENet optimizes the feature maps in the channel dimension. By assigning weights for different channels, the SENet focuses on the features of important channels, while restrains those of the unimportant channels. Hence, the SENet is capable of concentrating on key features.

We utilize a series of convolution and SENet modules to process the support images. Then the averaging module is adopted to obtain the prototypes by calculating the average features of the images from the same category. Suppose $f_{\Phi}(\cdot)$ represents the functional expression of APN model and Φ is the parameter matrix, the calculation of prototypes is expressed in Equation (3)

$$c_i = f_{\Phi}(S_i) \tag{3}$$

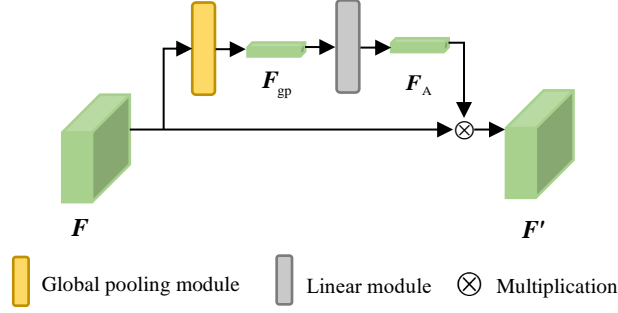where c_i denotes the prototype of category i.

Figure 2. The structure of SENet.

3.4 Feature Matching Classifier

To recognize the query images, we use the ACNN model to extract the feature maps. Except for the averaging module, ACNN shares the structures and parameters with APN, thus the structure of ACNN is not introduced in detail. The prototypical network method in [30] adopts the distance-based similarity metric functions to calculate the similarity scores between the feature maps and prototypes. Then the category probability $p \in \mathbb{R}^{M \times N}$ is obtained by the softmax function, where M and N respectively represent the number of images and categories in the query set. $p_{i,j} \in p$ is generated by:

$$p_{i,j} = \frac{\exp\left(-d\left(F_Q(i), c_j\right)\right)}{\sum_{n=1}^{N} \exp\left(-d\left(F_Q(i), c_n\right)\right)} \tag{4}$$

$\exp(\cdot)$ and $d(\cdot)$ denote the exponential and distance functions, respectively. F_Q is the set which contains the feature maps of query images, and $F_Q(i)$ denotes the feature map of the ith query image.

Due to the fixed parameters in the distance-based similarity metric functions, it is necessary to manually select appropriate distance functions for different types of data. We design a new feature matching classifier (FMC) which consists of neural network. Compared with the distance-based functions, the FMC is a neural network model with trainable parameters, thus it is capable of learning an appropriate similarity metric function for SAR data based on the training dataset. To obtain the category probability of query images, we separately combine its feature map with each prototype. Then the FMC is utilized to process the combined feature maps. As shown in Figure 1, FMC contains the convolution, SENet, and linear modules. The combined feature maps are first

forwarded to convolution and SENet modules for further processing and optimization. Then the

similarity score $z \in \mathbb{R}^{M \times N}$ between the query image and different prototypes is output by the linear

module. $z_{i,j} \in z$ is calculated by:

$$z_{i,j} = f_{\Psi}(\boldsymbol{F}_Q(i), \boldsymbol{c}_j) \tag{5}$$

where $f_{\Psi}(\cdot)$ represents the functional expression of the FMC model and Ψ is the parameter

matrix. The category probabilities are obtained by performing the softmax operation on the

similarity scores:

$$p_{i,j} = \frac{\exp(z_{i,j})}{\sum_{n=1}^{N} \exp(z_{i,n})} \tag{6}$$

3.5 Loss Function

Many few-shot image recognition algorithms adopt the cross-entropy loss function that

measures the differences between the recognition results and the corresponding ground-truth labels.

Suppose $y \in \mathbb{R}^M$ denotes the ground-truth labels of the query images, and $\overline{y} \in \mathbb{R}^{M \times N}$ is the one-

hot encoding of y. M and N respectively represent the numbers of images and categories in

the query set. The expression of the cross-entropy loss is as follows:

$$\text{loss}_{\text{ce}} = -\frac{1}{M} \sum_{i=1}^{M} \sum_{j=1}^{N} \overline{y}_{i,j} \log p_{i,j} \tag{7}$$

In the proposed method, we calculate the similarity scores between the feature maps and the

prototypes, thereby obtaining the recognition results. Therefore, the separability between different

prototypes affects the recognition performance of our method. We design the prototype-separability

loss as follows:

$$\text{loss}_{\text{ps}} = \frac{1}{\sum_{i \neq j} \left\| \boldsymbol{c}_i - \boldsymbol{c}_j \right\|_2^2} \tag{8}$$

where $i, j \in [1, N]$. In the training process, optimizing loss_{ps} can effectively increase the

separability between different prototypes, which helps to improve the performance of our method.

We utilize loss1 and loss2 to denote the cross-entropy and the prototype-separability losses,

respectively. By combining these two losses, the loss function of our method is shown in equation

(9):

$$\text{loss=loss1+loss2} \tag{9}$$

Algorithm 1 describes the training process of our method in detail.

Algorithm 1: Training process of the proposed few-shot SAR image recognition method.

Input: auxiliary set A. N_{iter} is the number of training iterations. N denotes the number of categories in the auxiliary set. A_i denotes the subset of A containing all the samples whose labels are i. A_S represents the auxiliary support set and A_Q is the auxiliary query set. A_{S_i} and A_{Q_i} respectively denote the subset of A_S and A_Q containing the samples whose labels are i. N_S and N_Q are the numbers of the randomly selected samples.

Output: learned parameter matrixes Φ and Ψ. Φ and Ψ denote the parameter matrixes of APN and FMC, respectively.

   for n in $\{1, \cdots, N_{iter}\}$ **do**

     for i in $\{1, \cdots, N\}$ **do**

- randomly select N_S samples from A_i to form the subset A_{S_i}.
- calculate the prototype of A_{S_i} using Equation (3).
- randomly select N_Q samples from ($A_i \setminus A_{S_i}$) to form the subset A_{Q_i}.

    end for

- $A_Q = A_{Q_1} \cup \cdots \cup A_{Q_N}$, use ACNN model to extract the features of each image from A_Q and calculate the category probabilities according to Equations (5) and (6).
- calculate $loss_{ce}$ and $loss_{ps}$ according to Equations (7) and (8).
- acquire the loss of our method by combing $loss_{ce}$ and $loss_{ps}$.
- update parameter matrixes Φ and Ψ by minimizing the combined loss.

   end for

---

4 Experiments

In this part, we first introduce the datasets, the evaluation metrics, and the detailed implementation of our method. Afterwards, we compare the recognition performance of our method with that of other methods. Finally, the proposed method is further discussed, including effectiveness verification, influence of the auxiliary set, computational efficiency, and the new loss function.

4.1 Preliminary

4.1.1 Datasets Description

The experiments are performed on two SAR datasets: the Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset and the Vehicle and Aircraft (VA) dataset.

(1) MSTAR Dataset

The resolution of the images in the MSTAR dataset is 0.3 m, and the image size is of 64×64

pixels. The imaging radar works on the X-band with HH polarization, and the depression angles include 15° and 17°. This dataset contains ten categories of targets, namely T62, T72, BMP2, D7, ZIL131, BTR60, BTR70, 2S1, ZSU234, and BRDM2. The optical and SAR images are shown in Figure 3. Table 1 shows the numbers of images contained in the ten categories.
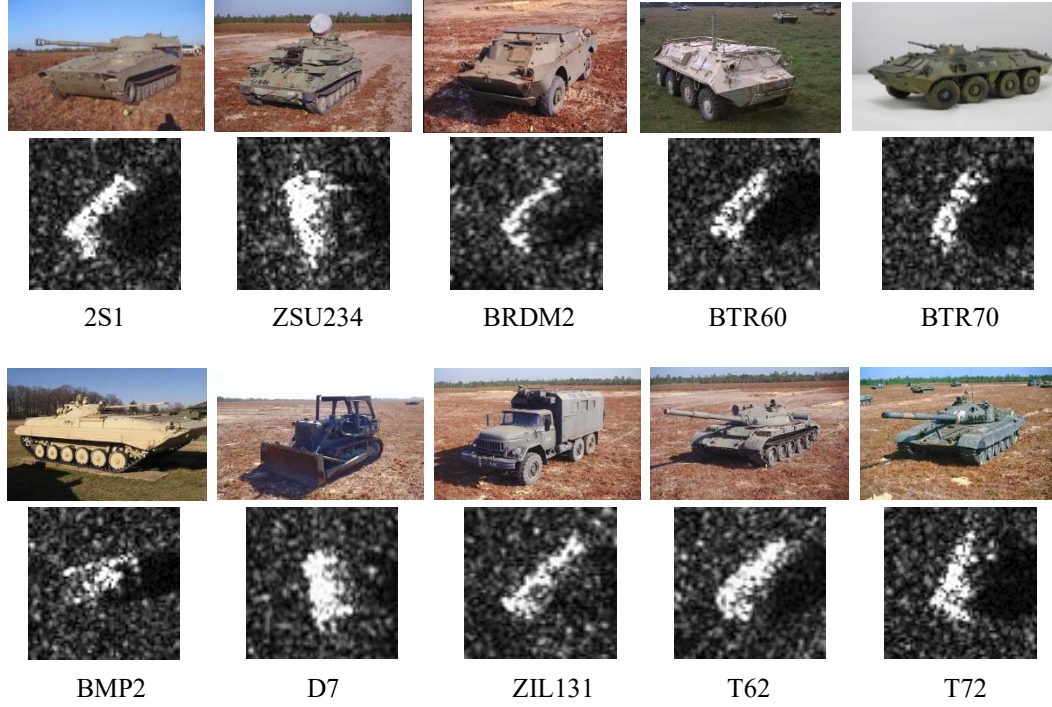

| 2S1 | ZSU234 | BRDM2 | BTR60 | BTR70 |

| BMP2 | D7 | ZIL131 | T62 | T72 |

Figure 3. The optical and SAR images in the MSTAR dataset.

Table 1. The numbers of images contained in the MSTAR dataset.

Category	Depression	
	17°	15°
2S1	299	274
ZSU234	299	274
BRDM2	298	274
BTR60	256	195
BTR70	233	196
BMP2	233	195
D7	299	274
ZIL131	299	274
T62	299	273
T72	232	196
Sum	2747	2425

We divide the MSTAR dataset into auxiliary, support, and query sets. As shown in Table 2, we randomly select six categories (BRDM2, BTR60, BTR70, ZIL131, T62, and T72) and the corresponding images are used to form the auxiliary set. Next, we respectively select K images

of the other four categories (2S1, BMP2, D7, and ZSU234) with the depression of 17° to construct

the support set. Finally, the images from the same four categories with 15° depression are utilized

to form the query set. As a result, the support and query sets share the same label space, which is

non-intersect with that of the auxiliary set.

Table 2. The numbers of images in the MSTAR auxiliary, support and query sets.

Auxiliary set			Support set			Query set		
Category	Depression	Number	Category	Depression	Number	Category	Depression	Number
BRDM2	17°&15°	572	2S1	17°	K	2S1	15°	274
BTR60	17°&15°	451	BMP2	17°	K	BMP2	15°	195
BTR70	17°&15°	429	D7	17°	K	D7	15°	274
T62	17°&15°	572	ZSU234	17°	K	ZSU234	15°	274
T72	17°&15°	428						
ZIL131	17°&15°	573						
Sum	3025		Sum	4×K		Sum	1017	

(2) VA Dataset

The VA dataset contains 5 categories of vehicles and 2 categories of aircrafts: sedan car,

delivery truck, bus, business car, fire truck, airliner, and helicopter. Figure 4 shows the

corresponding optical and SAR images. The resolution of the images is 0.5 m. The images sizes of

the vehicles and aircrafts are 64×64 and 128×128, respectively. In the experiments, we resize the

aircraft images to 64×64, which is consistent with the vehicle images. The imaging radar works on

the C-band with HH polarization, and the depression angle is 45°. The numbers of images contained

in different categories are shown in Table 3.

In Table 4, the VA dataset is divided to auxiliary, support, and query sets. The auxiliary set

consists of the images of sedan car, delivery truck, and helicopter. The support set is constructed by

respectively selecting K images from the other four types (bus, business car, fire truck, and

airliner). Since the number of images in VA dataset is much less than that of the MSTAR dataset,

the maximum value of K is set to 5. Finally, the remaining images in the four types are used to
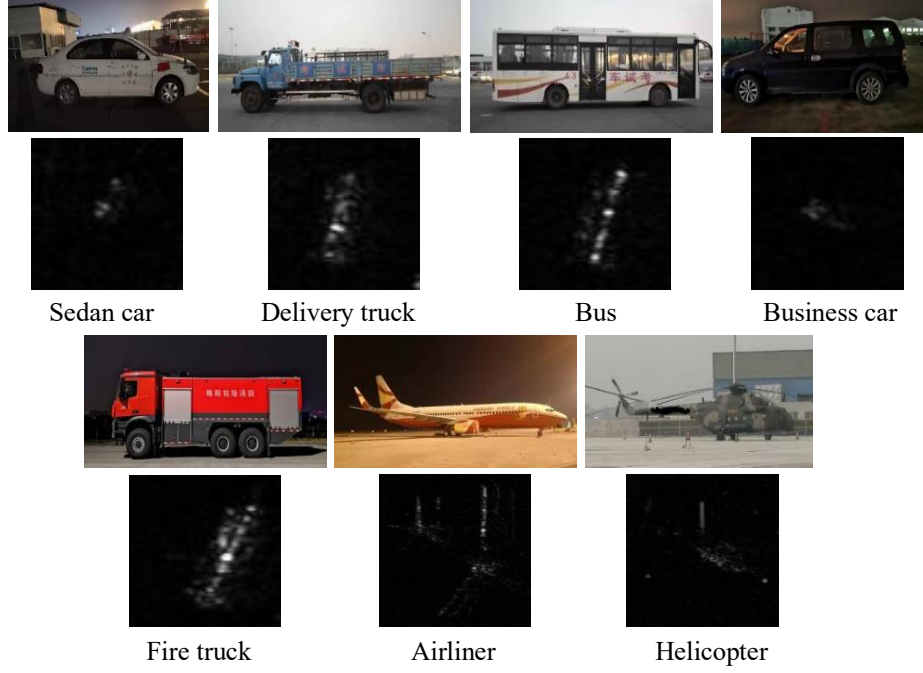
form the query set.

Sedan car   Delivery truck   Bus   Business car

Fire truck   Airliner   Helicopter

Figure 4. The optical and SAR images in the VA dataset.

Table 3. The numbers of images contained in the VA dataset.

Category	Number
Sedan car	35
Delivery truck	35
Bus	35
Business car	35
Fire Truck	35
Airliner	38
Helicopter	70
Sum	283

Table 4. The numbers of images in the VA auxiliary, support and query sets.

Auxiliary set		Support set		Query set	
Category	Number	Category	Number	Category	Number
Sedan car	35	Bus	K	Bus	30
Delivery truck	35	Business car	K	Business car	30
Helicopter	70	Fire Truck	K	Fire Truck	30
		Airliner	K	Airliner	33
Sum	140	Sum	$4\times K$	Sum	123

4.1.2 Evaluation Metrics

In the experiments, the overall precision (OP), overall accuracy (OA), average F1 (AF1) score, and kappa score are utilized for evaluating the recognition performance of different methods. OP and OA are calculated based on the confusion matrix of the recognition results:

$$\text{OP} = \frac{1}{M} \sum_{i=0}^{N-1} \frac{\left(\sum_{j=0}^{N-1} e_{ij}\right) e_{ii}}{\sum_{j=0}^{N-1} e_{ji}} \tag{10}$$

$$\text{OA} = \frac{1}{M} \sum_{i=0}^{N-1} e_{ii} \tag{11}$$

where M and N denote the numbers of images and categories in the query set, respectively. e_{ij} is the element of the confusion matrix at coordinate (i, j). $\sum_{j=0}^{N-1} e_{ij}$ and $\sum_{j=0}^{N-1} e_{ji}$ represent the element sums of the confusion matrix in the ith row and i^{th} column, respectively. F1 score is the harmonic average of the precision and recall values. The F1 score of category i is obtained using equation (12).

$$\text{F1}_i = 2 \times \frac{\text{P}_i \times \text{R}_i}{\text{P}_i + \text{R}_i} \tag{12}$$

P_i and R_i denote the precision and recall values for category i:

$$\text{P}_i = \frac{e_{ii}}{\sum_{j=0}^{N-1} e_{ji}} \tag{13}$$

$$\text{R}_i = \frac{e_{ii}}{\sum_{j=0}^{N-1} e_{ij}} \tag{14}$$

The AF1 score is the weighted average F1 scores of different categories:

$$\text{AF1} = \frac{1}{M} \sum_{i=0}^{N-1} \left(\sum_{j=0}^{N-1} e_{ij}\right) \text{F1}_i \tag{15}$$

Kappa score measures the consistency between the predicted results and the ground-truth labels:

$$\text{Kappa} = \frac{p_o - p_e}{1 - p_e} \tag{16}$$

p_o and p_e are calculated by equations (17) and (18).

$$p_o = \frac{1}{M} \sum_{i=0}^{N-1} e_{ii} \tag{17}$$

$$p_e = \frac{\sum_{i=0}^{N-1} \left(\sum_{j=0}^{N-1} e_{ij}\right)\left(\sum_{j=0}^{N-1} e_{ji}\right)}{M^2} \tag{18}$$

4.1.3 Implementation Details

In the structure of our method, the APN includes three convolution modules, three SENet modules, and an averaging module. The ACNN shares the structures and parameters with the APN except for the averaging module. The FMC consists of a convolution module, a SENet module, and

a linear module. The convolution module contains convolution, batch normalization, and pooling layers. The kernel size in convolution layers is 3×3, and the pooling layers adopt the average pooling operation with a size of 2×2. In the APN and ACNN, the kernel number for the convolution module is 64. The FMC adopts 128 kernels in the convolution module. The averaging module employs the averaging operation. The linear module in the FMC consists of two linear layers, and the numbers of neurons in these two layers are 256 and 1, respectively. The SENet is composed of a global pooing module and a linear module. The global pooling module contains a pooling layer which employs the global pooling operation. The linear module of SENet in APN and ACNN contains two linear layers, and the neuron numbers for these two layers are respectively set to 8 and 64. In the FMC, the linear module of SENet also includes two linear layers, the neuron numbers of which are 16 and 128, respectively.

4.2 Recognition Performance Comparison

We compare the proposed method with siamese neural network [28], matching network [29], prototypical network [30], relation network [31], DN4 [32], DSN [33], and FEAT method [34]. The Siamese neural network includes two CNN models for simultaneously extracting the features of two input images. Then the similarities of the two images are measured by a distance function. The matching network employs the attention and memory mechanisms for rapid learning. The prototypical network calculates the Euclidean distance between the query images and prototypes, thereby obtaining the label predictions. The relation network designs the relation model which adopts the neural network to compare the similarities of images. The DN4 method abandon the image-level features and adopt local descriptors to represents the local regions of the image. The DSN method utilizes the limited samples to construct dynamic classifiers. Based on the self-attention mechanism, FEAT method is capable of extracting discriminative and task-specific image features.

4.2.1 Experiments on the MSTAR Dataset

The MSTAR support set includes 4 categories, each of which contains K labeled images. In the experiments of most few-shot learning method, the typical settings of K are 1 and 5. Hence, we conduct our experiments under the typical settings. In other words, we compare the recognition performance of different methods under two conditions: "4-way 5-shot" and "4-way 1-shot".

(1) 4-way 5-shot

The experimental results under the "4-way 5-shot" condition is shown in Table 5. The performance of our method outperforms that of the Siamese, FEAT, and DN4 algorithms. This because these comparison methods employ the distance-based similarity metric functions in which the parameters are fixed. In our method, we design the FMC which consists of neural network. Different from the distance-based functions, the FMC is capable of learning a suitable similarity metric function for SAR data in a data-driven way. The OP of our method achieves 98.73%, which outperforms that of the prototypical network by 6.35%. The reason is that we design the prototype-separability loss and combine it with the cross-entropy loss in the training process. This new loss is capable of increasing the separability between different prototypes, which contributes to a higher recognition accuracy. Compared with the Relation and DSN methods, we improve the feature extraction ability of our model by embedding the SENet in the APN, ACNN, and FMC. As an attention model, the SENet emphasizes the important features while restrains the unimportant ones, thus the recognition performance of our method is effectively improved.

Table 5. The recognition performance of different methods on the MSTAR dataset with "4-way 5-shot" condition.

Methods	OP	OA	AF1	Kappa
Siamese	85.92%	82.89%	82.27%	77.05%
FEAT	89.24%	88.20%	88.35%	84.17%
Prototypical	92.38%	92.23%	92.24%	89.58%
DN4	92.32%	91.35%	91.37%	88.39%
Matching	94.98%	94.79%	94.77%	93.01%
DSN	94.33%	94.20%	94.20%	92.22%
Relation	95.23%	95.08%	95.07%	93.41%
Ours	**98.73%**	**98.72%**	**98.72%**	**98.29%**

To intuitively compare the recognition performance, we utilize the t-Distributed Stochastic Neighbor Embedding (t-SNE) algorithm to visualize the output of the eight methods. As shown in Figure 5, different colors denote different categories. In Figure 5 (h), only a small number of the query images are misclassified, thus our method can effectively distinguish different categories. In the visual figures of the other seven methods, however, the between-category distances are much smaller, which decreases the recognition accuracy. Especially in Figure 5 (a) and Figure 5 (b), the images of different categories are seriously confused. Table 6 shows the recognition performance of our method for different categories. As can be seen, the precision, recall, and F1 score of BMP2 are all 100%. Due to the confusion in the recognition results, the performance of 2S1, D7, and

ZSU234 is inferior to that of the BMP2.

| (a) Siamese | (b) FEAT | (c) Prototypical | (d) DN4 |

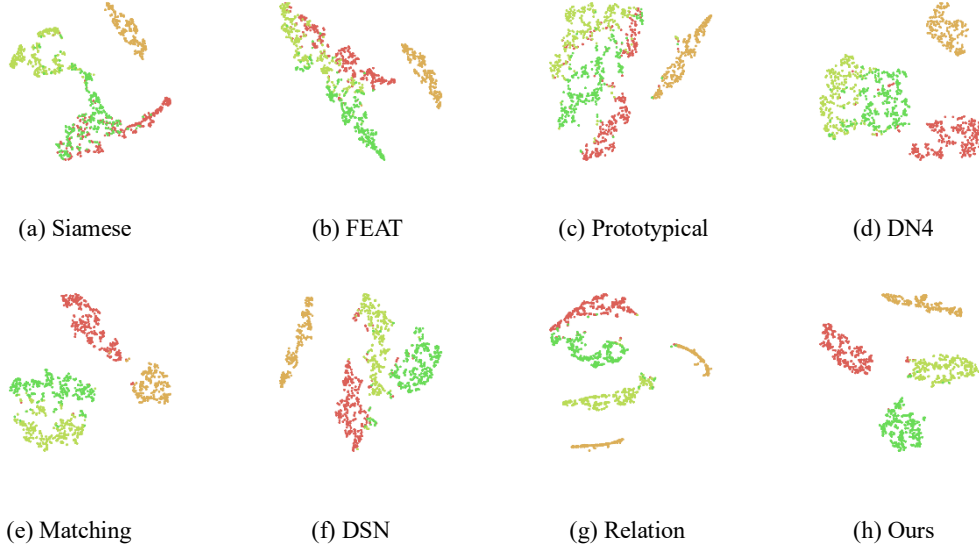| (e) Matching | (f) DSN | (g) Relation | (h) Ours |

Figure 5. The visual figures of the outputs obtained by different methods on the MSTAR query set. Different colors denote different categories.

Table 6. The recognition performance of our method for different categories in the MSTAR query set.

Category	Precision	Recall	F1-score
2S1	**100%**	98.54%	99.26%
BMP2	**100%**	**100%**	**100%**
D7	97.12%	98.54%	97.83%
ZSU234	98.18%	98.18%	98.18%

(2) 4-way 1-shot

Next, we compare the performance of different methods under the "4-way 1-shot" condition. As shown in Table 7, our method outperforms the other methods in OP, OA, AF1, and kappa score. The OP of our method reaches 97.08%, which demonstrates that our method is capable of obtaining satisfactory recognition performance with very few support images. Due to the weakness of distance-based similarity metric functions, the performance of Siamese, FEAT, and DN4 algorithms are inferior to that of our method. Since we adopt the SENet for enhancing the feature extraction and the prototype-separability loss for increasing the separability between different prototypes, the recognition accuracy of our method is superior to that of Matching, DSN, and Relation methods.

Table 7. The recognition performance of different methods on the MSTAR dataset with "4-way 1-shot" condition.

Methods	OP	OA	AF1	Kappa
Siamese	78.93%	76.79%	76.20%	68.86%
FEAT	84.37%	83.97%	83.97%	78.50%
Prototypical	89.22%	88.69%	88.66%	84.82%
DN4	85.84%	85.64%	85.69%	80.73%
Matching	85.73%	82.99%	83.38%	77.16%
DSN	87.96%	87.32%	86.97%	82.98%
Relation	92.91%	92.72%	92.69%	90.24%
Ours	**97.08%**	**97.05%**	**97.05%**	**96.04%**

4.2.2 Experiments on the VA Dataset

The VA support set includes 4 categories, and the number of labeled images in each category is K. Table 8 shows the recognition performance of different methods under the "4-way 1-shot" condition. Our method obtains the highest OP, OA, AF1, and kappa score, which demonstrates that our method can achieve satisfactory recognition performance on different SAR databases. The OP of our method (92.75%) is higher than that of the relation network (86.67%) and DSN method (88.68%), which proves the superiority of the SENet. Compared with the Siamese, FEAT, and DN4 algorithms, we adopt the FMC instead of the distance-based functions, thus a suitable similarity metric function for SAR data is learned in a data-driven way.

Table 8. The recognition performance of different methods on the VA dataset with "4-way 1-shot" condition.

Methods	OP	OA	AF1	Kappa
Siamese	73.09%	74.80%	73.61%	66.36%
FEAT	90.13%	88.62%	88.20%	84.82%
Prototypical	87.54%	84.55%	83.16%	79.39%
DN4	87.94%	85.37%	84.95%	80.48%
Matching	84.84%	84.55%	84.45%	79.36%
DSN	88.68%	87.80%	87.80%	83.73%
Relation	86.67%	86.18%	86.13%	81.54%
Ours	**92.75%**	**92.68%**	**92.69%**	**90.24%**

Then we utilize the t-SNE algorithm to visually compare the performance of different methods. In Figure 6 (a)-(g), there exists obvious confusion between different categories. Since the between-category distance in Figure 6 (h) is much larger, our method correctly recognizes most of the query images. The performance of our method for different categories is shown in Table 9. The precision, recall, and F1 scores of the business car and airliner are above 96%. However, the recognition

performance of the bus and fire truck is inferior on account of the confusion.

(a) Siamese (b) FEAT (c) Prototypical (d) DN4

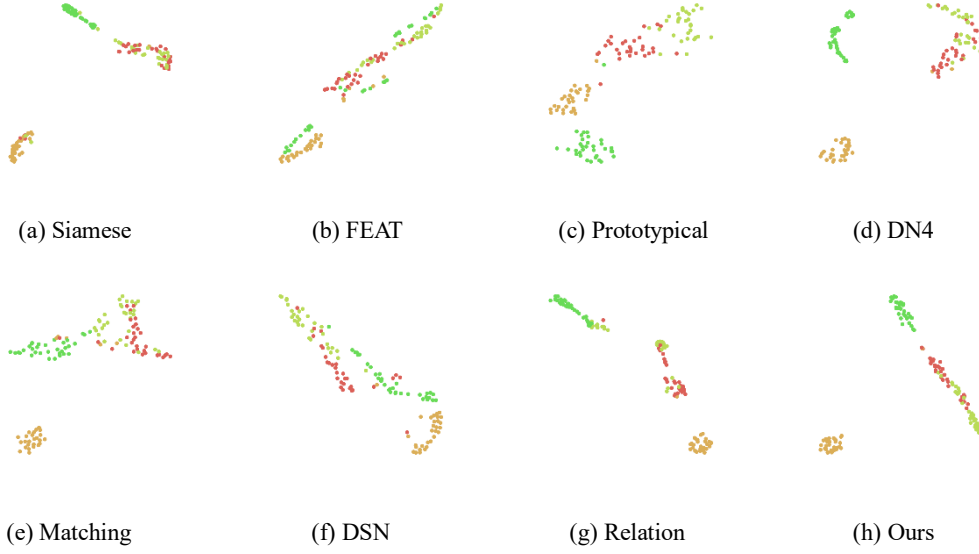(e) Matching (f) DSN (g) Relation (h) Ours

Figure 6. The visual figures of the outputs obtained by different methods on the VA query set. Different colors denote different categories.

Table 9. The recognition performance of our method for different categories in the VA query set.

Category	Precision	Recall	F1-score
Bus	83.87%	86.67%	85.25%
Business car	**100%**	96.67%	98.31%
Fire truck	89.66%	86.67%	88.14%
Airliner	97.06%	**100%**	**98.51%**

4.3 Discussion

4.3.1 Effectiveness Verification of the Proposed Method

Compared with the prototypical network in [30], we employ three strategies to improve the recognition performance. At first, we design the FMC to replace the distance-based similarity metric functions. Afterwards, we embed the SENet in the APN, ACNN, and FMC for feature optimization. Finally, a new loss function is designed by incorporating the cross-entropy and prototype-separability losses. Next, we verify the effectiveness of the FMC, the SENet, and the new loss function.

In Table 10, "baseline" denotes the prototypical network [30]. As can be seen, the OP, OA, AF1, and kappa score of these three strategies outperform those of the "baseline". The FMC automatically learns a suitable similarity metric function for SAR data, thus its performance is better than the distance-based functions where the parameters are fixed. The SENet utilizes attention values to assign weights for the feature maps. As a result, the representation of the prototypes and

query image features are enhanced. In the training process, the new loss function is capable of enlarging the separability between different prototypes, which contributes to the performance improvement. By combining the three strategies, the performance of our method is further improved.

Table 10. Effectiveness verification of FMC, SENet and new loss. Baseline denotes the prototypical network.

Strategies	OP	OA	AF1	Kappa
Baseline	92.38%	92.23%	92.24%	89.58%
FMC	95.98%	95.87%	95.89%	94.46%
SENet	96.32%	96.07%	96.06%	94.72%
New loss	95.88%	95.67%	95.68%	94.19%
FMC + SENet + New loss	**98.73%**	**98.72%**	**98.72%**	**98.29%**

4.3.2 Influence of the Auxiliary Set

The above experiments take advantage of all the images in the auxiliary set during the training process. In this section, we discuss the performance of our method trained by partial auxiliary set. At first, we respectively set the number of images from each category of the MSTAR auxiliary set to 20, 40, 60, 80, and 100. The experiments are performed under the "4-way 5-shot" condition. Figure 7 (a) shows the performance of our method trained by different numbers of images. As can be seen, the OP of our method is around 89.5% when each category contains 20 images. This because the training images are insufficient, which decreases the recognition accuracy. As the number of training images increases, the recognition performance of our method is gradually promoted. When each category contains 100 images, the OP of our method reaches around 97%.

Next, we discuss the impact of image categories contained in the auxiliary set on the performance of our method. As shown in Figure 7 (b), we randomly select partial categories from the MSTAR auxiliary set, and the experiments are performed under the "4-way 5-shot" condition. It is obvious that the performance of our method improves with the increase of the number of categories. The OP of our method is below 70% when the auxiliary set contains two categories of images. As the number of categories increases to six, the OP of our method rises to around 99%. This because the more categories contained in the auxiliary set, the more difficult it is to recognize the images during the training process, which helps to promote the performance of our method.

<center>(a)   (b)</center>
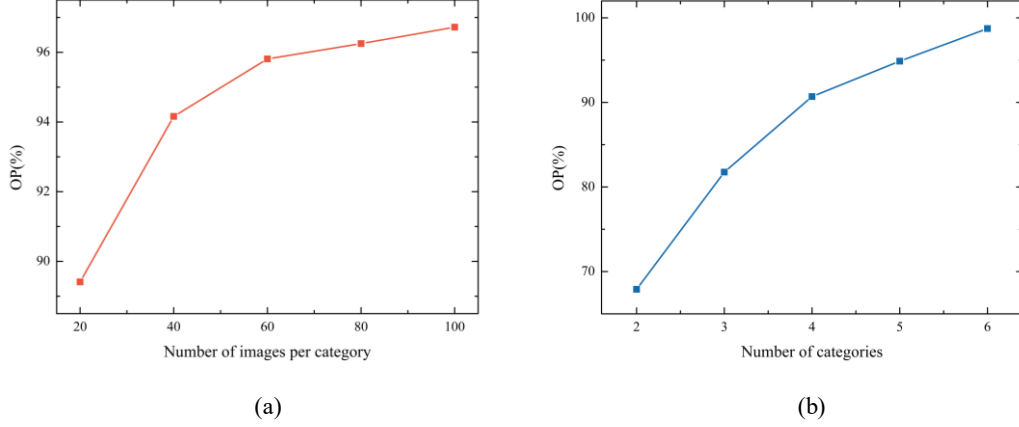
Figure 7. The influence of auxiliary set on the performance of our method. (a) The OP of our method trained by different numbers of images. (b) The OP of our method trained by different numbers of categories.

4.3.3 Computational efficiency

For evaluating the computational efficiency, we compare the running time of our method with that of the comparison methods. The experiments are performed on the MSTAR dataset with the "4-way 5-shot" condition. The GPU in our computer is GeForce RTX 2080 Ti with 11 GB memory. Table 11 shows the running time for different methods to recognize all the images in the query set. The running time of our method is 1.37 s, which is much shorter than that of the Siamese neural network (3.78 s) and the relation network (3.19 s). This because the Siamese and relation networks calculate the similarity scores between the query image and each support image. By contrast, our method first generates the prototypes of different categories, then the similarity scores between the query image and different prototypes are calculated, thus less running time is needed. Besides, our method requires more running time than the prototypical network. This due to we adopt the FMC to promote the recognition performance, which increases the computational complexity.

<center>Table 11. The running time required for recognizing all the images in the MSTAR query set.</center>

Methods	Running time
Siamese	3.78s
FEAT	1.31s
Prototypical	**1.09s**
DN4	2.62s
Matching	2.28s
DSN	2.47s
Relation	3.19s
Ours	1.37s

4.3.4 Discussion of the new loss function

In our method, the loss function consists of the cross-entropy and the prototype-separability losses. The cross-entropy loss is a common loss function in image recognition tasks, and it is used for measuring the difference between the predicted labels and the ground-truth labels. We design the prototype-separability loss to increase the separability between different prototypes, which is combined with cross-entropy loss to train our model.

The curves of the two losses over different iterations are drawn in Figure 8. We utilize loss1 and loss2 to denote the cross-entropy and the prototype-separability losses, respectively. As the iteration increases, loss1 drops from around 1.8 to 0. Whereas, loss2 drops from around 0.0011 to 0.0002. The magnitude of loss1 is much larger than that of loss2.

(a) cross-entropy loss   (b) prototype-separability loss
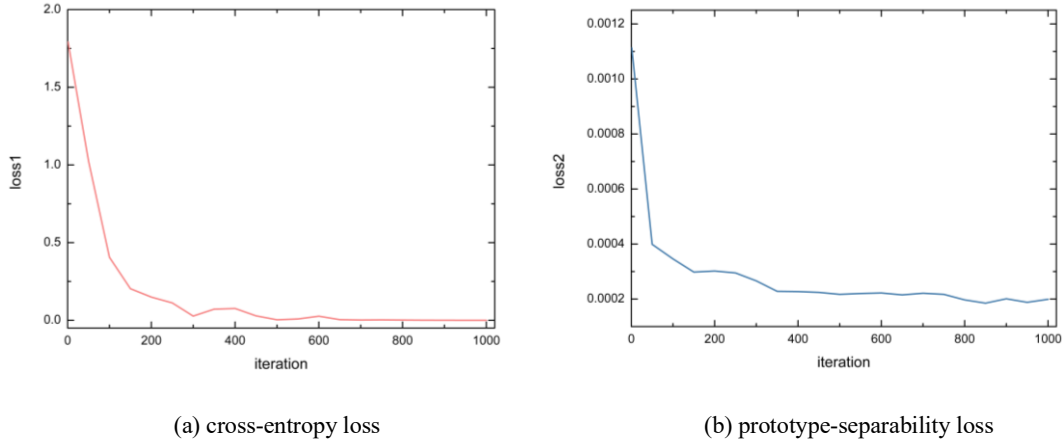
Figure 8. The curves of the cross-entropy and the prototype-separability losses over different iterations.

We utilize the weight parameter α to enlarge the magnitude of loss2, and try a new manner to combine these two losses:

$$loss=loss1+\alpha loss2 \tag{19}$$

To obtain an appropriate weight parameter, the α is set as 0, 1, 10, 100, 1000, and 10000, respectively. The OA of our method under different α is shown in Table 12:

Table 12. The OA of our method under different α

α	0	1	10	100	1000	10000
OA	97.05%	**98.72%**	98.23%	97.94%	96.26%	89.28%

When α is set as 0, the OA of our method is 97.05%. Whereas, the OA of our method increases to 98.72% if α is set as 1, which demonstrates the effectiveness of loss2. However, the OA of our method drops if α continues to increase. The reason is as follows:

The loss2 is designed to increase the separability between different prototypes. According to

the definition in equation (8), the loss2 does not measure the difference between the predicted labels and the ground-truth labels of the training images. If we separately use loss2 to train our model, the recognition accuracy of our method cannot be improved. In our method, the loss2 is utilized as an auxiliary loss which is combined with loss1 to train our model. As the weight of loss2 increases, the proportion of loss1 drops, thus the recognition accuracy of our method decreases.

In our method, we do not use the weight parameter to increase the magnitude of loss2. Although the magnitude of loss2 is much smaller than that of loss1, it is used as an auxiliary loss which is capable of improving the performance of our method.

5 Conclusion

We present a new few-shot SAR image recognition method which consists of APN, ACNN, and FMC. The APN calculates the prototypes of different categories. The ACNN is adopt for extracting the feature maps of query images. We design the FMC to generate the similarity scores between the feature maps and prototypes. As a neural network model, the FMC can automatically learn a suitable similarity metric function for SAR images. We embed the SENet in the structure of our method for enhancing the expression of prototypes and feature maps. In addition, the loss function for our method is a combination of the cross-entropy and prototype-separability losses. This new loss promotes the recognition performance by increasing the separability between different prototypes. Experiments on the MSTAR and VA datasets prove that our method obtains satisfactory accuracy in the few-shot SAR image recognition tasks. For instance, the OP of our method achieves 98.73% on the MSTAR dataset with the "4-way 5-shot" condition, which outperforms that of other few-shot learning methods such as DSN, DN4, matching network, and relation network. Our future work includes integrating the traditional features of SAR images into the few-shot recognition methods, and we believe that the finely-crafted traditional features will contribute to superior recognition performance.

References

1. Moreira A, Prats-Iraola P, Younis M, Krieger G, Hajnsek I, Papathanassiou K P. A tutorial on synthetic aperture radar [J]. IEEE Geoscience and Remote Sensing Magazine. 2013, 1(1): 6-43.

2. Ma F, Gao F, Wang J, Hussain A, Zhou H. A novel biologically-inspired target detection method based on saliency analysis for synthetic aperture radar (SAR) imagery [J]. Neurocomputing, 2020, 402: 66-79.

3. Chen H, Zhang F, Tang B, Yin Q, Sun X. Slim and efficient neural network design for resource-constrained SAR target recognition [J]. Remote Sensing. 2018, 10(10): 1618.

4. Gao F, Yue Z, Wang J, Sun J, Yang E, Zhou H. A Novel Active Semisupervised Convolutional Neural Network Algorithm for SAR Image Recognition [J]. Computational Intelligence and Neuroscience, 2017.

5. Anagnostopoulos G C. SVM-based target recognition from synthetic aperture radar images using target region outline descriptors [J]. Nonlinear Analysis, Theory, Methods and Applications, 2009, 71(12): 2934-2939.

6. Zhan X, Zhang R, Yin D, Huo C. SAR Image Compression Using Multiscale Dictionary Learning and Sparse Representation [J]. IEEE Geoscience and Remote Sensing Letters, 2013, 10(5): 1090-1094.

7. Dong G, Kuang G, Wang N, Wang W. Classification via Sparse Representation of Steerable Wavelet Frames on Grassmann Manifold: Application to Target Recognition in SAR Image [J]. IEEE Transactions on Image Processing, 2017, 26(6): 2892-2904.

8. Yue Z, Gao F, Xiong Q, Wang J, Hussain A, Zhou H. A novel attention fully convolutional network method for synthetic aperture radar image segmentation [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 4585-4598.

9. Gao F, Ma F, Wang J, Sun J, Yang E, Zhou H. Visual Saliency Modeling for River Detection in High-resolution SAR Imagery [J]. IEEE Access, 2017, 6: 1000-1014.

10. Zeng N, Li H, Wang Z, Liu W, Liu S, Alsaadi F E, Liu X. Deep-reinforcement-learning-based images segmentation for quantitative analysis of gold immunochromatographic strip [J]. Neurocomputing, 2021, 425: 173-180.

11. Zeng N, Wang Z, Zineddin B, Li Y, Du M, Xiao L, Liu X, Young T. Image-based quantitative analysis of gold immunochromatographic strip via cellular neural network approach [J]. IEEE transactions on medical imaging, 2014, 33(5): 1129-1136.

12. Zeng N, Wang Z, Zhang H, Kim K, Li Y, Liu X. An improved particle filter with a novel hybrid proposal distribution for quantitative analysis of gold immunochromatographic strips [J]. IEEE Transactions on Nanotechnology, 2019, 18: 819-829.

13. Lee H, Kwon H. Going Deeper With Contextual CNN for Hyperspectral Image Classification [J]. IEEE Transactions on Image Processing, 2017, 26(10): 4843-4855.

14. Ren P, Sun W, Luo C, Hussain A. Clustering-oriented Multiple Convolutional Neural Networks for Single Image Super-resolution [J]. Cognitive Computation, 2017, 10(1): 165-178.

15. Zhang F, Liu Y, Zhou Y, Yin Q, Li H. A lossless lightweight CNN design for SAR target recognition [J]. Remote Sensing Letters, 2020, 11(5): 485-494.

16. Dong H, Zhang L, Zou B. Densely Connected Convolutional Neural Network Based Polarimetric SAR Image Classification [C]. IEEE International Geoscience and Remote Sensing Symposium. 2019: 3764-3767

17. Zhang J, Xing M, Sun G C, Bao Z. Integrating the Reconstructed Scattering Center Feature Maps With Deep CNN Feature Maps for Automatic SAR Target Recognition [J]. IEEE Geoscience and Remote Sensing Letters (Early access). 2021.

18. Cho J H, Chan G P. Additional feature CNN based automatic target recognition in SAR image [C]. Proceedings of the 4th Asian Conference on Defence Technology. 2017: 1-4.

19. 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 [J]. Cognitive Computation. 2018, 11(6): 809-824.

20. 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 [J]. Cognitive Computation. 2019: 1-12.

21. Gao F, Yang Y, Wang J, Sun J, Yang E, Zhou H. A Deep Convolutional Generative Adversarial Networks (DCGANs)-Based Semi-Supervised Method for Object Recognition in Synthetic Aperture Radar (SAR) Images [J]. Remote Sensing, 2018, 10(6): 846.

22. Liu B, Yu X, Yu A, Zhang P, Wan G, Wang R. Deep Few-Shot Learning for Hyperspectral Image Classification [J]. IEEE Transactions on Geoscience and Remote Sensing. 2019, 57(4): 2290-2304.

23. Das D, Lee C. A Two-Stage Approach to Few-Shot Learning for Image Recognition [J]. IEEE Transactions on Image Processing, 2020, 29: 3336-3350.

24. Tang H, Li Y, Han X, Huang Q, Xie W. A Spatial–Spectral Prototypical Network for Hyperspectral Remote Sensing Image [J]. IEEE Geoscience and Remote Sensing Letters. 2020, 17(1): 167-171.

25. Yang X, Nan X, Song B. D2N4: A Discriminative Deep Nearest Neighbor Neural Network for Few-Shot Space Target Recognition [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(5): 3667-3676.

26. Wang D, Cheng Y, Yu M, Guo X, Zhang T. A hybrid approach with optimization-based and metric-based meta-learner for few-shot learning [J]. Neurocomputing, 2019, 349: 202-211.

27. Li X, Yu L, Fu C W, Fang M, Heng P. A. Revisiting metric learning for few-shot image classification [J]. Neurocomputing, 2020, 406: 49-58.

28. Koch G, Zemel R, Salakhutdinov R. Siamese Neural Networks for One-shot Image Recognition [C]. Proceedings of the 32nd International Conference on Machine Learning. 2015

29. Vinyals O, Blundell C, Lillicrap T, Wierstra D. Matching Networks for One Shot Learning [C]. Advances in Neural Information Processing Systems. 2016: 3637-3645.

30. Snell J, Swersky K, Zemel R S. Prototypical Networks for Few-shot Learning [C]. Advances in Neural Information Processing Systems. 2017: 4078-4088.

31. Sung F, Yang Y, Zhang L, Xiang T, Torr P H, Hospedales T M. Learning to Compare: Relation Network for Few-Shot Learning [C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2018: 1199-1208.

32. Li W, Wang L, Xu J, Huo J, Gao Y, Luo J. Revisiting local descriptor based image-to-class measure for few-shot learning[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 7260-7268.

33. Simon C, Koniusz P, Nock R, Harandi M. Adaptive subspaces for few-shot learning[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 4136-4145.

34. Ye H J, Hu H, Zhan D C, Sha F. Few-shot learning via embedding adaptation with set-to-set functions[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 8808-8817.

35. Tang J, Zhang F, Zhou Y, Yin Q, Hu W. A Fast Inference Networks for SAR Target Few-Shot Learning Based on Improved Siamese Networks [C]. IEEE International Geoscience and Remote Sensing Symposium. 2019: 1212-1215.

36. Wang L, Bai X, Gong C, Zhou F. Hybrid Inference Network for Few-Shot SAR Automatic Target Recognition [J]. IEEE Transactions on Geoscience and Remote Sensing. 2021, to be published.

37. Fu K, Zhang T, Zhang Y, Wang Z, Sun X. Few-Shot SAR Target Classification via Metalearning [J]. IEEE Transactions on Geoscience and Remote Sensing. 2021, to be published.

38. Yang R, Xu X, Li X, Wang L, Pu F. Learning Relation by Graph Neural Network for SAR Image Few-Shot Learning [C]. IEEE International Geoscience and Remote Sensing Symposium. 2020: 1743-1746.

39. Woo S, Park J, Lee J Y, Kweon I S. CBAM: Convolutional Block Attention Module [C]. Proceedings of the European Conference on Computer Vision. 2018, 3-19.

40. Li X, Wang W, Hu X, Yang J. Selective Kernel Networks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019, 510-519.

41. Hu J, Shen L, Sun G. Squeeze-and-Excitation Networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2020, 42(8): 2011-2023.

Zhenyu Yue received the B.S. degree in electronic and electrical engineering for civil aviation from the Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China, in 2016. He is currently pursuing the Ph.D. degree in the Beihang University. His research interests include radar signal processing, machine learning, image segmentation, and image recognition.

Fei Gao received the B.S. degree in industrial electrical automation, and the M.S. degree in electro-magnetic measurement technology and instrument from the Xi'an Petroleum Institute, Xi'an, China, in 1996 and 1999, respectively, and the Ph.D. degree in signal and information processing from the Beihang University, Beijing, China, in 2005. He is currently a Professor with the School of Electronic and Information Engineering, Beihang University. His research interests include target detection and recognition, image processing, deep learning for applications in remote sensing.

Qingxu Xiong received the PhD degree in electrical engineering from Peking University, Beijing, China, in 1994. From 1994 to 1997, he worked in the Information Engineering Department at the Beijing University of Posts and Telecommunications as a postdoctoral researcher. He is currently a professor in the School of Electrical and Information Engineering at the Beihang University. His research interests include high performance Switching, performance modeling of networks, and semantic communication networks. He is a member of the IEEE.
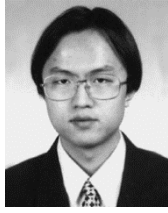
Jinping Sun received the M.S. and Ph.D. degrees from the Beihang University, Beijing, China, in 1998 and 2001, respectively. He is currently a Professor with the School of Electronic and Information Engineering, Beihang University. His research interests include high-resolution radar signal processing, image understanding, and robust beamforming.

Amir Hussain received the B.Eng (highest 1st class honours with distinction) and Ph.D degrees, from the University of Strathclyde, Glasgow, U.K., in 1992 and 1997, respectively. He has held Postdoctoral and academic positions with the University of West of Scotland, Glasgow, U.K., from 1996 to 1998, University of Dundee, Dundee, U.K., from 1998 to 2000, and University of Stirling, Stirling, U.K., from 2000 to 2018. He joined Edinburgh Napier University in Scotland, U.K., in 2018, as Founding Director of the Centre of AI and Data Science. He has coauthored three international patents and over 400 publications, including 170+ international journal papers, 20 books/monographs, and over 100 book chapters. His research interests include developing cognitive data science and AI technologies, to engineer the smart and secure systems of tomorrow.

Huiyu Zhou received the B.Eng. degree in radio technology from Huazhong University of Science and Technology, Wuhan, China, the M.S. degree in biomedical engineering from University of Dundee, Dundee, U.K., and the Ph.D. degree in radio technology, biomedical engineering, and computer vision from Heriot-Watt University, Edinburgh, U.K. He is currently a Professor with the School of Informatics, University of Leicester, Leicester, U.K. His research interests include medical image processing, computer vision, intelligent systems, and data mining.