Journal of New Media
DOI: 10.32604/jnm.2022.029360
Article

Deep Feature Bayesian Classifier for SAR Target Recognition with Small Training Set

Liguo Zhang[1,2], Zilin Tian[1], Yan Zhang[3,*], Tong Shuai[4], Shuo Liang[4] and Zhuofei Wu[5]

[1]College of Computer Science and Technology, Harbin Engineering University, Harbin, 150001, China
[2]Xidian University, Xi'an, 710000, China
[3]China Industrial Control Systems Cyber Emergency Response Team, Beijing, 100040, China
[4]The 54th Research Institute of CETC, Shijiazhuang, 130100, China
[5]University of Bologna, Bologna, 00013, Italy
*Corresponding Author: Yan Zhang. Email: zhangyan_5933@163.com
Received: 01 March 2022; Accepted: 11 April 2022

Abstract: In recent years, deep learning algorithms have been popular in recognizing targets in synthetic aperture radar (SAR) images. However, due to the problem of overfitting, the performance of these models tends to worsen when just a small number of training data are available. In order to solve the problems of overfitting and an unsatisfied performance of the network model in the small sample remote sensing image target recognition, in this paper, we uses a deep residual network to autonomously acquire image features and proposes the Deep Feature Bayesian Classifier model (RBnet) for SAR image target recognition. In the RBnet, a Bayesian classifier is used to improve the effect of SAR image target recognition and improve the accuracy when the training data is limited. The experimental results on MSTAR dataset show that the RBnet can fully exploit effective information in limited samples and recognize the target of the SAR images more accurately. Compared with other state-of-the-art methods, our method offers significant recognition accuracy improvements under limited training data. Noted that the RBnet is moderately difficult to implement and has the value of popularization and application in engineering application scenarios in the field of small-sample remote sensing target recognition and recognition.

Keywords: Bayesian classifier; limited data; synthetic aperture radar (SAR); target recognition

1 Introduction

With the advancement of remote sensing technology, obtaining high-quality remote sensing photographs has become increasingly simple, allowing for more thorough observations of specific information and changes on the earth's surface. Data labeling is getting increasingly time-consuming and labor-intensive as the volume of remote sensing data grows, and it frequently requires talent from the application area. It is challenging to provide rich training samples for traditional deep learning-based models, leading to the overfitting issue in the model training process, and the performance of target recognition is significantly reduced. As a result, even though learning with a limited amount of samples is a typical problem, research efforts for SAR target recognition with a limited training set are still needed.

Traditional machine learning algorithms have been widely used in remote sensing image automatic processing and a lot of research achievements have been achieved, such as [1] used a monogenic signal to acquire wide spectrum information and maximize SAR images spatial localization, and [2] proposed projection features to realize SAR target recognition. The convolutional neural networks (CNNs) is used by [3] which automatically learn the hierarchies of features from a substantial amount of training data. However, To get good recognition performance, deep learning-based methods require a large amount of training data. It is challenging for the model to obtain strong recognition performance when the labeled data used to train it is insufficient. In order to alleviate the contradiction between "the demand of massive training samples" and "the reality of limited annotation samples", researchers at home and abroad have proposed a small sample learning method, that is, to obtain more accurate target recognition results on test datasets through a small amount of training data.

The previous work on SAR image target recognition with limited training samples can be divided into model improvement, customized transfer learning and data augmentation techniques. Zhang et al. [4] offer an updated CNN model to overcome the limited sample issue by feature augmentation using optimally selected convolutional layers and ensemble learning techniques, which is subsequently utilized to replace the original softmax layer with an ensemble learning-based classifier. In [5], a multiple feature-based lightweight CNNs (MFCNNs) model for SAR target recognition with variable training data ratios is described without the use of a separate preprocessing method or posture information. In [6,7], the idea is also based on CNNs and model improvement. As depicted in [8], rather than upgrading the model, another approach is to apply tailored transfer learning, which use an auto-encoder to gain knowledge from a sufficient number of unlabeled SAR images and transfer it to a labeled SAR data set. Data augmentation is the most straightforward method when data is insufficient. Zhang et al. [9] used a pre-trained significance graph network to extract foreground and background information of samples. Then, independent encoders are reconstructed in feature space through a fusion network to generate samples containing more new concepts. Chen et al. [10] mentioned a meta-learner containing an image deformation subnetwork. Image distortion can lead to reduced image features, but the distorted image can still provide critical semantic information. Although the above-improved algorithms achieved some recognition effects, there are still some limitations. Simultaneously, when the number of accessible training data per category is less than 20 samples, the effectiveness of all of these approaches suffers dramatically.

To further improve discriminative performance for target recognition and address the issue that obtaining SAR images and doing human labeling on a wide scale is typically time-consuming and arduous, we propose new methods based on a limited sample. At present, research on limited sample learning mainly focuses on classification tasks, especially image classification tasks. At the same time, there are few types of research on the limited sample in a target recognition, for the important reason that it is a relatively difficult task to transfer. In order to solve the problem of sample limitation, Bayesian learning is proposed in this paper. Although Bayesian learning has been applied to small sample deep learning algorithms due to its unique probability inference ability, applying image information to Bayesian learning is still a critical problem. To overcome the challenges differently, In this article, the Bayesian is combined with the residual network (ResNet) to design a deep network, namely RBNet, to improve SAR target recognition accuracy.

The main contribution of this article can be summarized as follows:

- In the proposed RBNet, transfer learning is employed to transmit the effective information of the image to the Bayesian classifier for small sample learning tasks. Transfer learning is an efficient method for training an extensive network with sparse training data without overfitting. The suggested solution, as compared to existing methods, eliminates the complexity and inefficiencies of data tagging. The superiority of our solution may be proven by comparing the performance of the old scheme with the proposed strategy.
- The residual network we used in RBNet can fit the classification function better than the simple multi-layer network, that is, find the parameter value of the desired function.
- The suggested technique outperforms the state-of-the-art CNN-based method in SAR target recognition with a limited sample, which is a bottleneck barrier in SAR target recognition.

2 Related Work

This section provides a brief overview of prior SAR target identification experiments utilizing CNNs, followed by a brief introduction to Few-shot recognition, transfer learning, and open-set classification.

2.1 SAR Target Recognition with CNNS

With the continuous development of deep learning technology, its application in SAR image target recognition has a good prospect. Chen et al. [11] introduced Tiny Yolo-Lite, a lightweight ship detector achieved by self-designed backbone network topology and network pruning, and which employs knowledge extraction approach to compensate for network performance deterioration induced by pruning. Liang et al. [12] presented an adaptive hierarchical ship recognition approach based on the coarse-to-fine mechanism and built an enhanced visual attention mechanism that included image and frequency domains. To properly recognize the target, non-parametric BKDE (Block Kernel Density Estimation) is utilized. Wang et al. [13] utilized SSD and FPN to increase the recognition accuracy of SAR image ships considerably. Wang et al. [14] investigated the impact of coherent speckle in SAR images on the use of CNN in SAR target recognition. On this premise, they suggested a bipolar coupled CNN structure. The denoising subnetwork was utilized initially for denoising, and then the classification subnetwork was used to learn residual speckle properties and target information. This topology can increase the network's noise resilience. In [15], researchers conducted experiments on various CNNs with classification accuracy as the evaluation criterion. The experiment results showed that ResNet has the highest classification accuracy.

2.2 Few-shot Recognition

Few-shot classification aims to use limited training data to achieve better learning results. For deep learning algorithms to grasp and generalize certain abstract ideas and, eventually, to get a better classification result, a large quantity of data is typically required for training. Because of this, its performance is directly related to the quantity of data it can handle. Humans, on the other hand, are capable of copying, learning, and achieving excellent results even when there aren't enough instances to learn from. Bayesian learning is a kind of thinking that falls within this category. In order to acquire the posterior distribution, it makes use of previous knowledge and information from small samples, which results in obtaining the overall distribution. Bayesian learning is a method of learning and inference that makes use of probability to represent the associated uncertainty and to realize the process of learning and inference. Due to its greater resemblance to the human learning mindset, Bayesian learning is considered to be more appropriate for small sample learning. In reality, Bayesian learning has gradually been applied in small sample learning. For example, Hierarchical Bayesian Program Learning (HBPL) is proposed to explain the structure of characters and complete the classification and recognition of characters [16]. The HBPL introduced the concepts of combination and causality, which proved to be very effective in tackling small sample learning issues. Following that, a unique Bayesian learning framework is developed to mimic human learning processes and distinguish fundamental components from existing characteristics [17]. The framework may then generate new characters based on these components, rather than just classifying and recognizing existing characters. Some scholars advocate using the Bayesian meta-learning technique to learn the posterior distribution of the prototype vectors of relations. When a graph neural network is applied to the global relation graph, it is possible to parameterize the initial prior of prototype vectors [18]. The method combines Bayesian learning and meta-learning to improve performance.

2.3 Transfer Learning

However, despite the fact that classical machine learning technology has achieved significant success and has been effectively implemented in a wide range of practical applications, it still has certain limits. In many instances, it is often difficult, costly, and time-consuming to acquire adequate training data. In certain cases, it is almost impossible. To address the aforementioned issue, as mentioned in [19], the concept of "transfer learning" has been put forward, which focuses on knowledge transfer across domains. Aim of transfer learning is to transfer information from a large dataset known as the source domain to a smaller dataset known as the target domain in order to overcome the challenge of gathering sufficient training data to construct models [20]. An existing model may be used as the basis for a new model on another problem in the Transfer learning approach. When used in the feature extraction step, it may lower the cost of training the final job by training the first task first. Transfer learning with CNNs is a method that has been widely employed in a variety of fields. Three types of transfer learning issues exist: transductive, inductive, and unsupervised transfer learning [21]. These three categories may be described in terms of labeling from a labeling perspective. Transductive transfer learning, in its broadest sense, applies to circumstances in which the label information is derived only from the domain of the source. Depending on whether or not the target domain instances include label information, the situation may be classified as inductive transfer learning. It is referred to as unsupervised transfer learning when there is no label information available for both the source and the destination domains. Oquab et al. [22] demonstrated the effectiveness of a basic transfer learning approach on smaller benchmark datasets. They also reused mid-level features collected from ImageNet-trained CNNs. In order to improve transfer learning effect, some researchers also combined the transfer learning and incremental learning [23]. Meanwhile, a novel transfer learning algorithm [16] is proposed. It can directly predict the weight parameters of the classifier [24]. This method trains a parameter predictor. The pre-trained network in a large dataset can map the model to a test set with few samples by directly predicting parameters from activation values.

3 Methodology

3.1 RBNet Network Architecture

Inspired by Zhuang et al. [19] and Guan et al. [25], we present a novel network architecture for SAR target recognition. In this article, the residual network (ResNet) is combined with the Bayes classifier to improve the few-shot SAR target recognition accuracy. We use the Bayes classifier rather than the normal Linear classifier to improve the classification accuracy on SAR target recognition. Moreover, skip connections are added to enhance the network's ability to utilize features of different scales in the final prediction, as shown in Fig. 1. Since the model has a smaller number of learnable parameters, a network design like this may help to mitigate the overfitting issue.

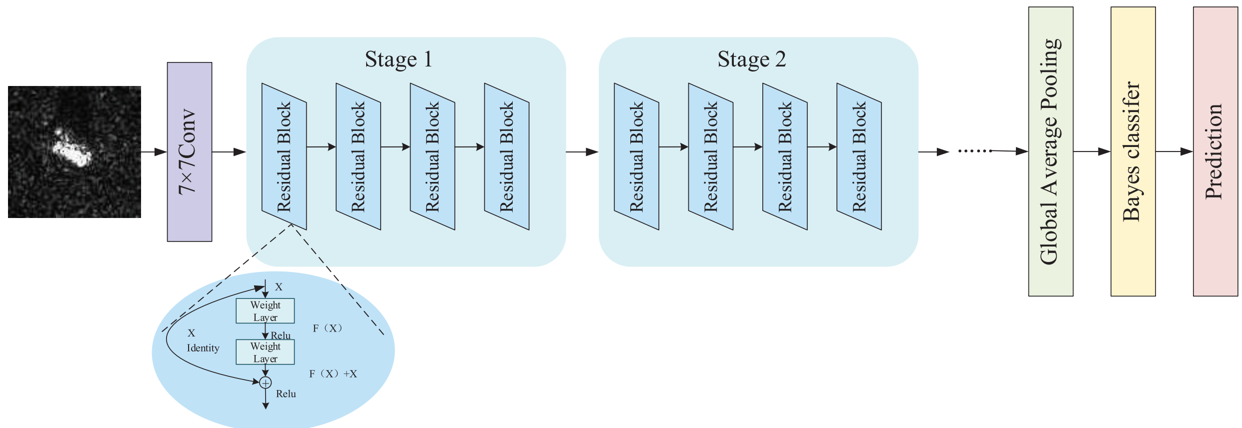

Figure 1: Framework of the proposed RBNet

The network used in this paper mainly refers to the ResNet-34 [26]. In a similar vein, we selected a 34-layer structure and made additional tweaks to make it more appropriate for our data. Each stage in this model contains 8 residual blocks, and 64 convolution kernels of size 3×3 follow each Residual block to reduce speckle noise and increase the number of channels. In our network, the size of all convolution kernels is set to 3×3. The parameters setting of our network is shown in Tab. 1.

Table 1: The parameters setting of our network

Layer name	Layer type	Output size
Conv1	7×7	112×112
Stage1	$\begin{pmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{pmatrix} \times 3$	56×56
Stage2	$\begin{pmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{pmatrix} \times 3$	28×28
Stage3	$\begin{pmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{pmatrix} \times 3$	14×14
Stage4	$\begin{pmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{pmatrix} \times 3$	7×7
Average pool, Fc, softmax		1×1

Resnet was first proposed by He et al. [27]. This article found that if a K layer of the network f is the current optimal network, there must be a deeper network f' whose the last few layers are the output of the Identity Mapping of K layer of the network f. Furthermore, the result of f' is consistent with f. So K is not the so-called "optimal number of layers", and compared with shallow networks, deeper networks should not perform worse.

However, if the network is deepened directly, there will be gradient explosion or gradient disappearance problems, which will cause the network performance to decline. For this type of problem, normalization can be used to ensure that the stochastic gradient descent is used in the backpropagation to achieve the convergence effect. However, the normalization operation is only applicable to dozens of layers of networks. When the network is deepened, the problem of model degradation will still occur. Moreover, ResNet introduces a jump link line, which can directly skip one or more layers through an identity shortcut key to complete the data transfer, as shown in Fig. 2. When the depth of the neural network increases and the gradient disappears, identity mapping is performed on the original network to ensure the integrity of the information, extract more detailed information, and avoid a rapid decline in network performance.

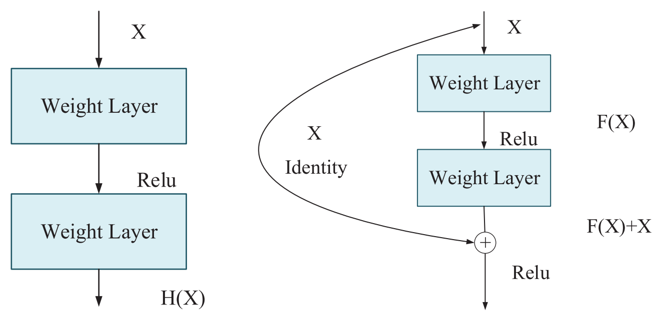

Figure 2: Ordinary convolution block(L). Residual block(R)

The residual network contains two 3 $*$ 3 convolution layers. The residual function F(X) is superimposed over the upper output in the second layer to guarantee that the output y has the same vector dimension as the upper and lower module outputs, as seen in Fig. 2. By shifting the learning objective from full output to residual, ResNet reduces the issue of information loss and kernel loss inherent in classic convolutional networks and ensures the integrity of data by transferring the input directly to the output. Additionally, simplifying learning objectives decreases the difficulty of learning.

$$F(X) = \mathcal{W}_2\sigma(\mathcal{W}_1 X) \tag{1}$$

$$y = F(X, \{\mathcal{W}_i\}) + \mathcal{W}_s X. \tag{2}$$

Following the completion of the ResNet-34 network's final full connection layer, we implement the shared features. The activation function, as shown in Eq. (3), extends the noisy ReLU by include the noise extension.

$$\alpha_{ReLU}(t) = \max(0, t + \varphi), \text{ with } \varphi \sim \mathcal{N}(0, \sigma(t)) \tag{3}$$

3.2 Bayes Classifier

In few-shot classification, transfer learning is used to make the model obtain auxiliary knowledge from related fields in advance and improve the classification accuracy on new classes. We first use the base classes data to train the model to update its parameters, θ^l where l = 1, 2, . . . , L is the index of convolutional layers, and determine the threshold parameter, φ, of the Bayes classifier. In our network, L is empirically set to 18. When training the model on base classes, we fix φ and update θ^l of the convolutional layers. After updating the convolutional layer parameters, we fix the parameters of the convolutional layers and adjust the threshold parameter so that the network model obtains the best classification accuracy on base classes. At the beginning of the experiments, we empirically set the initial value of the threshold parameter in the Bayes classifier to 0.1, then adjust the parameters within the upper and lower ranges to make the network model achieve the best classification performance on the base classes. In our experiments, the threshold parameter of the Bayes classifier is set in the 0.0500–1.1182 range. After the parameter transfer, the threshold parameter does not need to be updated on new classes. In the model, we use the cross-entropy as the loss function, and adopt Adam [28] as the optimization algorithm. In summary, the Bayesian decision strategy perform Bayesian statistical analysis so as to realize the identification of known and unknown classes.

After training the network model, it performs feature statistics on all samples in the training set. Our model first classifies the input sample according to the Bernoulli naive Bayesian classification. The class corresponding to the maximum posterior probability is seen as a prediction result for the sample. The Bayesian formula applied in this network model is as follows:

$$P(y_i|X) = \frac{P(X|y_i)\,P(y_i)}{\sum_j P\left(X|y_j\right) P\left(y_j\right)} \tag{4}$$

In the Eq. (4), $P(y_i|X)$ represents the probability that the y_i when the feature is X, and $y_1, y_2, \cdots, y_n$ are exhaustive events, namely $\bigcup_{i=1}^{n} y_i = \Omega, y_i y_j = \phi, P(y_i) > 0$, $P(y_i)$ is a priori probability that does not need to consider many factors related to X and represents the occurrence probability of class y_i. P (X) is called a normalized constant. X is the feature vector of an input sample and is extracted by the network, and y refers to the corresponding label of the sample. After a sample is an input into the model, the model will perform classification based on known classes, and we will obtain its posterior probability $P(y_i|X)$. In this paper, we use $P(Xy_i)$ as an important reference for whether the sample image belongs to unknown classes. Its means the probability when a sample has feature vector X and belongs to the known class y_i. In summary, the Bayesian decision strategy performs Bayesian statistical analysis to realize the identification of known and unknown classes.

4 Experiments

In this section, we conduct a series of experiments to test and validate the proposed method. Section 4.1 describes the dataset and environment utilized in the experiments. In Section 4.2, we compare our findings to those obtained via the use of other well acknowledged methods. In the last portion of Section 4.3, we conduct ablation experiments to verify our method.

4.1 Data Set and Experimental Setup

In the experiment, the data sets are part of the MSTAR data sets which is open. Each SAR slice images in the MSTAR collection has a single vehicle target in the image's center. Vehicle targets are classified into ten categories: infantry fighting vehicles (BMP2), armored transport vehicles (BTR 70, BTR 60), self-propelled howitzers (2S1), tanks (T72, T62), armored reconnaissance vehicles (BRDM 2), bulldozers (D7), cargo trucks (ZIL131), and self-propelled anti-aircraft gun (SPAA) (ZSU234). Each kind of vehicle target in the MSTAR dataset has photos with azimuths ranging from 0° to 360°, however in actuality, the azimuths of each target sample are around 1° to 5° apart in the published portion of the dataset. According to the acquisition circumstances, the data in the MSTAR dataset may be classified into two categories: Standard Operating Condition (SOC) and Extended Operating Condition (EOC). Under SOC circumstances, ten kinds of ground vehicle targets are covered. The data for the training set were obtained at a 17° imaging pitch angle, whereas the data for the test set were taken at a 15° imaging pitch angle. The optical views of many vehicle targets and their associated SAR images under SOC circumstances are shown in Fig. 3. In our trials, we collect 10 distinct kinds of data under SOC conditions, as stated in Tab. 2.

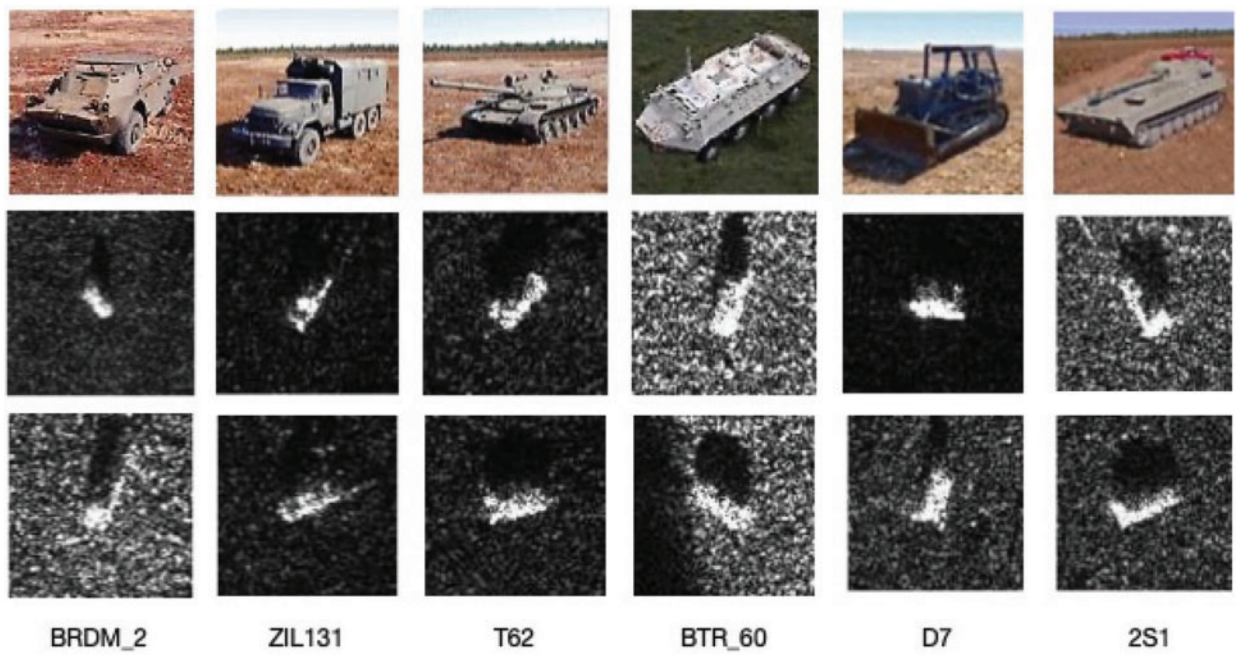

Figure 3: Optical images of some vehicle targets and their corresponding SAR images, the first row is the optical image, the second row and the third row are the SAR images

In all the experiments, an aggressive but straightforward data augment strategy is applied to to use each available sample fully. Our experimental platform configuration is shown in Tab. 3.

Table 2: Data statistics

Object classes	2S1	BRDM_2	BTR_60	D7	SN_132	SN_9563	SN_C71	T62	ZIL131	ZSU _23_4
Training set (Amplitude)	275	275	196	275	233	234	234	274	275	275
Validation set (Amplitude)	300	299	257	300	196	195	196	300	300	300

4.2 Performance Comparison

Our model is evaluated in comparison to state-of-the-art approaches, including MFCNNs [6], CNN-TL-bypass [8], and TMDC-CNNS [25]. The performance of our technique on the small sample dataset is evaluated by randomly picking 100, 50, 20, and 10 samples from each category in the whole training dataset. The results are summarized in Tab. 4, and it is clear that our model beats other strategies on an overall basis.

Table 3: Experimental platform configuration

Configuration	Version
LINUX	UBUNTU 18.04
CPU	INTEL CORE I7–9700K 3.60 GHZ
GPU	NVIDIA GEFORCE RTX 2080TI 11G
MEMORY	128G
PYTORCH	PYTORCH 1.3.1
CUDA	CUDA9.0

Table 4: Accuracy of different methods and different number of training samples

Model	Number of samples in each class				
	10	20	50	100	all
MFCNNs	70.67	88.36	93.63	95.06	98.12
CNN-TL-bypass	78.43	88.59	95.59	97.19	98.96
TMDC-CNNS	79.89	89.15	97.23	97.89	99.02
RBNet	80.18	90.23	97.54	98.12	99.36

Table 5: Ablation studies with accuracy (%) of different network structures

Model	Number of samples in each class				
	10	20	50	100	all
ResNet-18(baseline)	71.56	87.53	93.59	97.19	98.96
ResNet-34	73.93	88.89	94.63	97.69	98.98
ResNet-18+Bayes classifier	75.19	89.12	95.47	98.05	99.02
RBNet	80.18	90.23	97.54	98.12	99.36

When this method has only 10, 20, 50, and 100 training samples for the new target, the average recognition accuracy is 80.18%, 90.23%, 97.54%, and 98.12%, respectively. Compared with the MFCNNs network, the average recognition accuracy is increased by 9.51%, 1.87%, 3.91%, and 3.06%. Compared with the CNN-TL-bypass network, the average recognition accuracy is increased by 1.75%, 1.64%, 1.95%, and 0.93%. Compared with the TMDC-CNNS network, the average recognition accuracy is increased by 0.29%, 1.08%, 0.31%, and 0.34%, respectively, all of which have achieved higher advantages. Experimental results show that there are only a small number of image decorations in each category, and the RBNet can still get a good target recognition accuracy rate, indicating that this method is effective for small sample target recognition tasks.

To further illustrate the improvement and effectiveness of our RBNet, confusion matrices of RBNet are given in Figs. 4 and 5. The number of horizontal and vertical coordinates in confusion matrices corresponds to the numeric label assigned to each target category. The confusion matrix in Fig. 4 is for 15 training sets, with 15 pictures from each category randomly chosen for training purposes. Fig. 5 depicts the confusion matrix for all training sets. What can be clearly seen in this figures is the effectiveness for SAR target recognition with small training set. Following analysis, it is expected that the RBNet approach may successfully mitigate the overfitting issue associated with limited data.

4.3 Ablation Studies

As our RBNet are derived from the residual network and Bayes classifier, we also perform ablation studies with the ResNet-34, Bayes classifier, and the most commonly used architecture: ResNet-18 [27], to illustrate our efforts on the design of our proposed architecture even more precisely.

In order to show the performance of various models, different numbers of samples in each category are employed, and the results are presented in the following Tab. 5. As can be observed, when compared to the baseline ResNet-18 [27], the ResNet-34, the ResNet-18 with Bayes classifier, and our network architecture RBNet all outperform the baseline for a limited number of training samples, with our proposed RBNet model outperforming the baseline by the most significant margin. As shown in the Tab. 3, the RBNet model we proposed in this article also achieves high recognition accuracy when the number of training data is restricted.

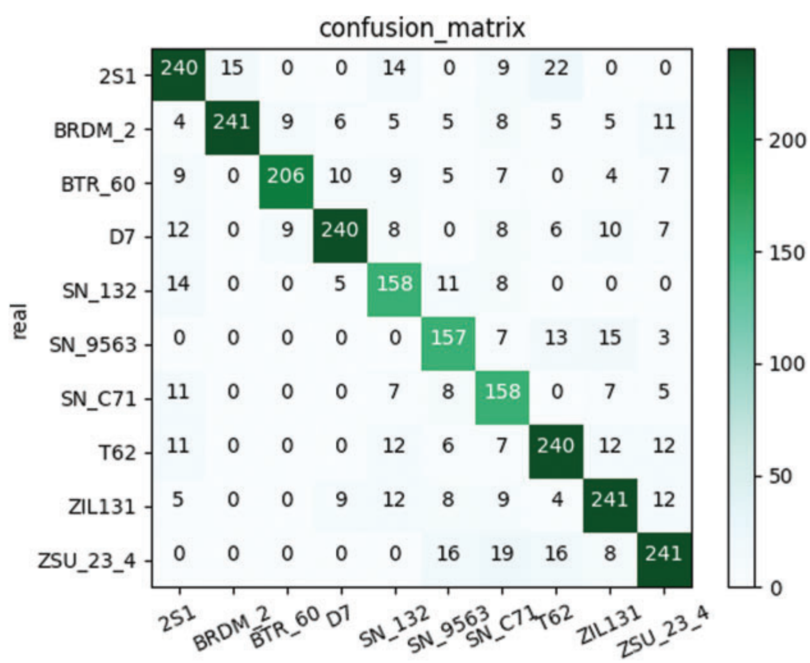

Figure 4: Confusion matrices of RBNet in 10 training samples

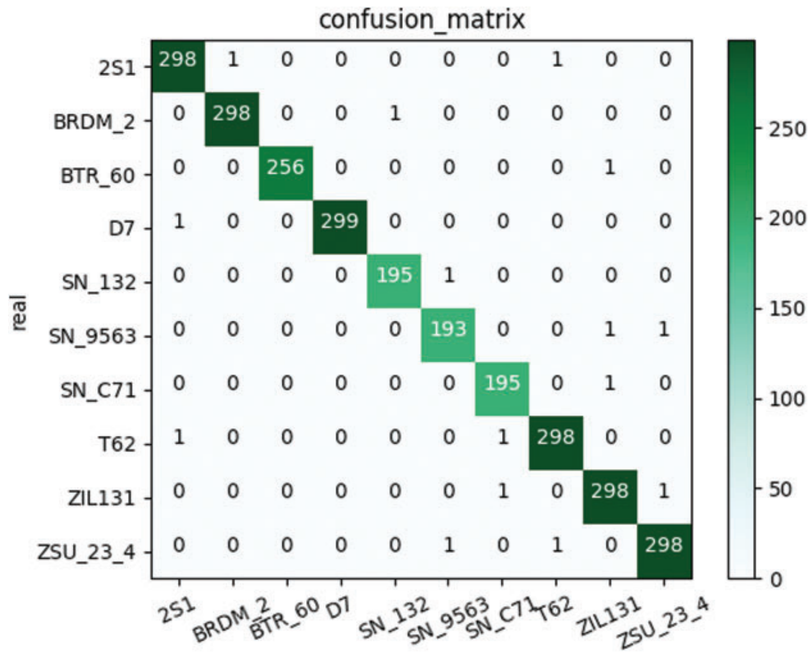

Figure 5: Confusion matrices of RBNet in all training samples

5 Citations

RBNet was used in this study to improve the performance of SAR target recognition when just a limited number of training samples were available. To accomplish the intended outcomes, this model employs a novel network architecture based on the Bayes classifier and the Resnet network. Experiments on RBNet demonstrated that our suggested model outperformed a variety of state-of-the-art approaches, despite the fact that we only had a small quantity of training data. These discoveries may be beneficial to other researchers who are attempting to achieve effective outcomes in their own investigations. Apart from these, there are several more exciting projects to work on, such as improving the way of intelligent data augmentation techniques and how to enhance this framework and improve its recognition and localization performance.

Acknowledgement: We express our heartfelt thanks to Velickovic P, Cucurull G, Casanova A, et al. for providing the open source.

Funding Statement: Our research fund is funded by the National Key R&D Program of China (2021YFC3320302).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] G. Dong and G. Kuang, "SAR target recognition via sparse representation of monogenic signal on grassmann manifolds," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 3, pp. 1308–1319, 2016.

[2] M. Liu, S. Chen, J. Wu, F. Lu, J. Wang *et al.,* "Configuration recognition via class-dependent structure preserving projections with application to targets in SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 6, pp. 2134–2146, 2018.

[3] R. Min, H. Lan, Z. Cao and Z. Cui, "A gradually distilled CNN for SAR target recognition," *IEEE Access*, vol. 7, pp. 42190–42200, 2019.

[4] F. Zhang, Y. Wang, J. Ni, Y. Zhou and W. Hu, "SAR target small sample recognition based on CNN cascaded features and AdaBoost rotation forest," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 6, pp. 1008–1012, 2020.

[5] J. H. Cho and C. G. Park, "Multiple feature aggregation using convolutional neural networks for SAR image-based automatic target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 12, pp. 1882–1886, 2018.

[6] B. Ding, G. Wen, X. Huang, C. Ma and X. Yang, "Data augmentation by multilevel reconstruction using attributed scattering center for SAR target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 6, pp. 979–983, 2017.

[7] Z. Lin, K. Ji, M. Kang, X. Leng and H. Zou, "Deep convolutional highway unit network for SAR target classification with limited labeled training data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1091–1095, 2017.

[8] Z. Huang, Z. Pan and B. Lei, "Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data," *Remote Sens.*, vol. 9, no. 9, pp. 907, 2017.

[9] H. Zhang, J. Zhang and P. Koniusz, "Few-shot learning via saliency-guided hallucination of samples," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, Long Beach, CA, USA, pp. 2770–2779, 2019.

[10] Z. Chen, Y. Fu and Y. Wang, "Image deformation meta-networks for one-shot learning," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, Long Beach, CA, USA, pp. 8680–8689, 2019.

[11] S. Chen, R. Zhan and W. Wang, "Learning slimming sar ship object detector through network pruning and knowledge distillation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 1267–1282, 2020.

[12] Y. Liang, K. Sun and Y. Zeng, "An adaptive hierarchical recognition method for ship targets in high-resolution SAR images," *Remote Sensing*, vol. 12, no. 2, pp. 303, 2020.

[13] J. Wang, C. Lu and W. Jiang, "Simultaneous ship recognition and orientation estimation in SAR images based on attention module and angle regression," *Sensors*, 2018, vol. 18, no. 9, pp. 2851.

[14] J. Wang, T. Zheng and P. Lei, "Ground target classification in noisy SAR images using convolutional neural networks," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018, vol. 11, no. 11, pp. 4180–4192.

[15] G. Sapijaszko and W. B Mikhael, "An overview of recent convolutional neural network algorithms for image recognition," in *2018 IEEE 61st Int. Midwest Symp. on Circuits and Systems (MWSCAS)*, IEEE, Windsor, ON, Canada, pp. 743–746, 2018.

[16] B. M. Lake, R. Salakhutdinov and J. B. Tenenbaum, "One-shot learning by inverting a compositional causal process," in *Proc. of the 26th Int. Conf. on Neural Information Processing Systems*, Lake Tahoe Nevada, vol. 2, pp. 2526–2534, 2013.

[17] B. M. Lake, R. Salakhutdinov and J. B. Tenenbaum "Human-level concept learning through probabilistic program induction," *Science*, 2015, vol. 350, no. 6266, pp. 1332–1338.

[18] M. Qu, T. Gao and L. P. Xhonneux, "Few-shot relation extraction via Bayesian meta-learning on relation graphs," in *Int. Conf. on Machine Learning*, PMLR, Virtual Event, pp. 7867–7876, 2020.

[19] F. Zhuang, Z. Qi and K. Duan, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2020.

[20] Z. Huang, Z. Pan and B. Lei, "Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data," *Remote Sensing*, vol. 9, no. 9, pp. 907, 2017.

[21] S J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[22] M. Oquab, L. Bottou and I. Laptev, "Learning and transferring mid-level image representations using convolutional neural networks," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Columbus, OH, USA, pp. 1717–1724, 2014.
[23] H. Qi, M. Brown and D. G. Lowe, "Low-shot learning with imprinted weights," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 5822–5830, 2018.
[24] S. Qiao, C. Liu and W. Shen, "Few-shotimage recognition by predicting parameters from activations," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 7229–7238, 2018.
[25] J. Guan, J. Liu and P. Feng *et al.*, "Multiscale deep neural network with Two-stage loss for SAR target recognition with small training set," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2021.
[26] X. Z. Han and R. Jin, "A small sample image recognition method based on ResNet and transfer learning," in *2020 5th Int. Conf. on Computational Intelligence and Applications (ICCIA)*, IEEE, Beijing, China, pp. 76–81, 2020.
[27] K. He, X. Zhang and S. Ren, "Deepresidual learning for image recognition," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, pp. 770–778, 2016.
[28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.