

ARTICLE

Metric-based Few-shot Classification in Remote Sensing Image

Mengyue Zhang^{1,2} Jinyong Chen^{1,2} Gang Wang^{1,2*} Min Wang^{1,2} Kang Sun¹

1. The 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang, Hebei, 050081, China

2. CETC Key Laboratory of Aerospace Information Applications, Shijiazhuang, 050081, Hebei, China

ARTICLE INFO

Article history

Received: 23 November 2021

Revised: 14 February 2022

Accepted: 21 February 2022

Published: 8 March 2022

Keywords:

Few-shot

Metric learning

Remote sensing

Target recognition

Episodic training

ABSTRACT

Target recognition based on deep learning relies on a large quantity of samples, but in some specific remote sensing scenes, the samples are very rare. Currently, few-shot learning can obtain high-performance target classification models using only a few samples, but most researches are based on the natural scene. Therefore, this paper proposes a metric-based few-shot classification technology in remote sensing. First, we constructed a dataset (RSD-FSC) for few-shot classification in remote sensing, which contained 21 classes typical target sample slices of remote sensing images. Second, based on metric learning, a k-nearest neighbor classification network is proposed, to find multiple training samples similar to the testing target, and then the similarity between the testing target and multiple similar samples is calculated to classify the testing target. Finally, the 5-way 1-shot, 5-way 5-shot and 5-way 10-shot experiments are conducted to improve the generalization of the model on few-shot classification tasks. The experimental results show that for the newly emerged classes few-shot samples, when the number of training samples is 1, 5 and 10, the average accuracy of target recognition can reach 59.134%, 82.553% and 87.796%, respectively. It demonstrates that our proposed method can resolve few-shot classification in remote sensing image and perform better than other few-shot classification methods.

1. Introduction

Remote sensing technology is a spatial observation technique, which captures remote sensing image of the earth's surface. It has numerous characteristics, including wide observation range, high efficiency and less restric-

tion to geographic conditions. Therefore, it is widely used in geological research, environmental monitoring, military defense, etc. Moreover, with the progress of aerospace technology, space-borne remote sensing detection technology is developing very rapidly. China's high-resolution satellites have imaging capability of panchromatic,

*Corresponding Author:

Gang Wang,

The 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang, Hebei, 050081, China; CETC Key Laboratory of Aerospace Information Applications, Shijiazhuang, 050081, Hebei, China;

Email: wanggg@tju.edu.cn

DOI: <https://doi.org/10.30564/aia.v4i1.4124>

Copyright © 2022 by the author(s). Published by Bilingual Publishing Co. This is an open access article under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License. (<https://creativecommons.org/licenses/by-nc/4.0/>).

multispectral and radar, which can achieve high-resolution earth observations and capture high-quality remote sensing images of the earth's surface. Recently, in deep learning, researchers have utilized massive remote sensing images to carry out various forms of research and applications, such as remote sensing image target recognition, and proposed neural networks such as Faster R-CNN^[1], YOLO^[2], R-FCN^[3] and SSD^[4], which have made significant progress in target recognition.

Target recognition based on deep learning is widely used in the field of remote sensing. This method can achieve promising target recognition performance through training and optimizing by a large number of labeled samples. Generally, it is believed that supervised deep learning algorithms can achieve acceptable performance upon 5000 labeled samples per class. When a data set of at least 10 million labeled samples is used for training, it will meet or exceed human performance. However, in some specific remote sensing scenarios, the number of available samples is very limited. For example, it is exceptionally difficult for dynamic targets such as ships to obtain sufficient samples of multiple types, conditions and sun incidence angles through remote sensing technology. Some ships, whose quantity is small and frequency of dispatches is low, there may be rare samples or no samples. In such a scenario, when samples are rare, the deep learning network will certainly have problems such as over-fitting. Besides, the reduction in the quantity of samples will cause the difference between targets of the same class to become larger, which may deteriorate the performance of target recognition. This situation is a typical few-shot problem.

To solve the few-shot problem, researchers performed Few-Shot Learning (FSL), that is, for a particular task T , the algorithm uses the limited training data D_{train} to learn experience E , thereby improving the performance of the algorithm on the test data D_{test} . Researchers have conducted few-shot target recognition research on natural image datasets, such as data augmentation, meta-learning based methods, metric-learning based methods and so forth. Augmentation of data attempts to create active samples to improve the size of data. Meta-learning based methods introduce meta-learning to make the meta-learner generalize to new tasks. Metric-learning based methods utilize metric learning to learn a mapping space, where targets can be separated easily. All of the methods as mentioned earlier have made great progress in few-shot classification.

The current methods mainly focused on data augmentation, optimizer, transfer learning, however, they did not pay sufficient attention to that in remote sensing scenarios, the scarcity of samples deteriorated the problem that intra-class distance is large and inter-class distance is small.

Thus, in this paper, we develop a new metric-based

method to work out few-shot target recognition. First, we proposed a k-nearest neighbor target recognition network, which used k-nearest neighbor^[5] to find remarkable features in the measurement module. By this way, the difficult problem caused by large intra-class distance and small inter-class distance can be effectively solved. Furthermore, we introduced episodic training to improve the generalization of the model.

To test our model, we prepared a remote sensing dataset of few-shot classification, which is called RSD-FSC, containing 21 classes of typical remote sensing targets. Experiments are conducted on this dataset. Compared to the Prototype network^[12], it improves the 1-shot, 5-shot and 10-shot accuracy from 56.62% to 59.13%, from 67.32% to 82.55% and from 71.18% to 87.80%, respectively.

2. Related Work

There are two main directions of few-shot learning from the existing literature. On one side, researchers use data generation methods to augment samples, such as using Generative Adversarial Networks (GANs) to generate highly realistic new samples through adversarial training of generative networks and judgment networks^[6-9]; AJ Ratner et al. introduced a geometric-changes-sequence model based on geometric changes such as horizontal flipping, rotation, scaling, and adding Gaussian noise to achieve random geometric changes to existing samples, complete the amplification of the number of samples and ensure the target in the image without serious distortion^[10]; Amit Alfassy et al. proposed Label-Set Operations networks (LaSO), which uses convolutional neural networks to extract the features of several images, and combines the features, such as intersection, union, difference, etc. to get the features of the hidden target in the image and generate new samples^[11]. On the other side, the investigator conducted research based on the structure of model and training methods. Prototype network, which is based on the metric learning, gets the feature tensor of images by convolutional neural networks. Then, it calculates the average feature tensor as a prototype for each class, and finally calculates the distance between query images and prototypes as the similarity to recognize the category of targets^[12]. Chen Lifu et al. introduced transfer learning to learn a supervised pre-trained convolutional neural network in the source domain, and then use a few samples in the target domain to retrain the model. This training method based on transfer learning improves the target recognition accuracy and convergence speed of the network^[13]. Task-Agnostic Meta-Learning (TAML) introduced an optimization strategy to prevent meta-learner bias in several tasks^[14].

3. The Proposed Method

3.1 K-nearest Neighbor Target Recognition Network

Metric learning can maximize the inter-class variations and minimize the intra-class variations in a mapping space, therefore, this paper proposed a metric-based k-nearest neighbor target recognition network. The essential idea is that in the feature mapping space, similar targets are close, and different targets are far away. The cosine distance between targets is used as the similarity to judge the category of targets. The flowchart of the k-nearest neighbor target recognition network is shown in Figure 1. It is mainly composed of two parts: a feature extractor and a feature measurement module, which correspondingly perform the extracting features of the targets in remote sensing images and the measurement of similarity between targets.

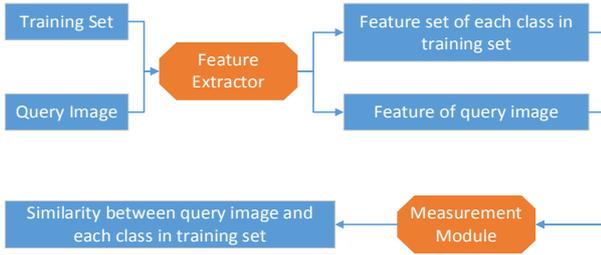


Figure 1. K-nearest neighbor target recognition network

1) Feature extractor

The feature extractor can be any suitable convolutional neural networks without the fully connected layer, which can extract deep local features of the image.

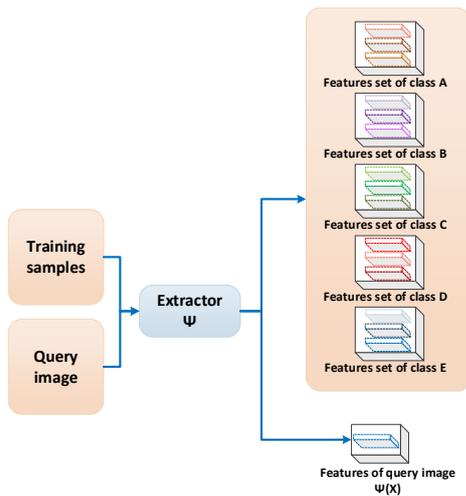


Figure 2. Feature extractor

As shown in Figure 2, the input of the feature extractor are images. The input image X passes through the feature extractor. The feature extractor outputs a feature tensor

$\Psi(X)$:

$$\Psi(X) = [x_1, x_2, \dots, x_n] \in R^{d \times n} \quad (1)$$

Where x_i is a feature kind of the target, which has d n-dimensional feature vectors. Thus, we can get the feature set of each class of targets:

$$\{\Psi(X)^1, \Psi(X)^2, \dots\} = \{[x_1, x_2, \dots, x_n]^1, [x_1, x_2, \dots, x_n]^2, [x_1, x_2, \dots, x_n]^3, \dots\} \quad (2)$$

Furthermore, when input a query image q , the feature extractor outputs its feature tensor $\Psi(q)$:

$$\Psi(q) = [q_1, q_2, \dots, q_n] \in R^{d \times n} \quad (3)$$

Therefore, the feature extractor turns images into deep local feature tensors, which can distinguish different classes of targets. Through the feature extractor, we get the feature tensor of the query image and the feature set of each class of targets of training samples. Besides, the feature extractor outputs the feature tensors to the measurement module to calculate the similarity between targets.

2) Measurement Module

In the measurement module, motivated by the prototype network^[12], using the distance between targets in the mapping space to recognize targets. Prototype network gets feature tensors of images by convolutional neural networks, and then calculates the average feature tensor as a prototype for each class, and finally calculates the distance between query images and prototypes as the similarity to recognize the category of targets. However, in few-shot learning, the inter-class variation is small, and the intra-class variation is large, which is problematical, thus, the average feature tensor can't represent each class well. Consequently, this paper uses the k-nearest neighbor to find representative samples to calculate the similarity between query images and every class of training sets.

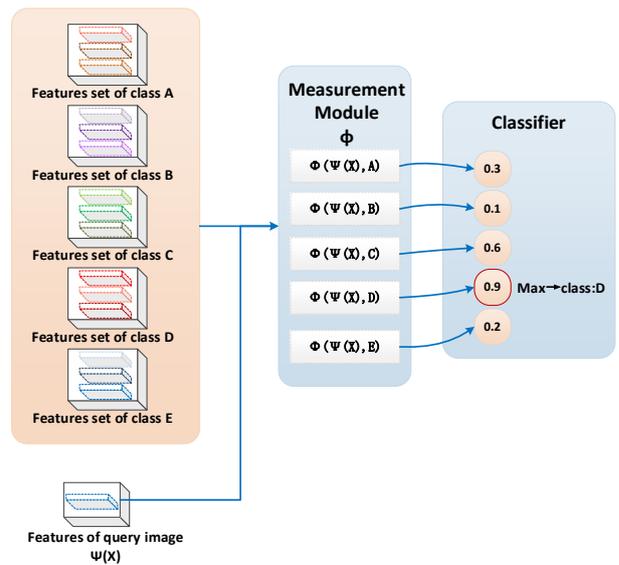


Figure 3. Measurement module

As shown in FIGURE 3, first, the measurement module gets the feature tensor $\Psi(q)$ of the query image and the features set $\{\Psi(X)^1, \Psi(X)^2, \dots\}$ of each class of targets from the feature extractor. Second, we use k-nearest neighbors to find the top-k feature vectors similar to the feature of the query image in the feature set of each class, and then calculate the cosine distance between the query image and each class of targets, respectively. This cosine distance is defined as the similarity between the query image and each class of targets. The smaller the distance, the larger will be the similarity.

The calculation process is as follows. For each feature component x_i of the query image, using k-nearest neighbor to find the top-k similar feature vectors from the features set of each class targets:

$$\text{Top-k similar feature vectors} = \{\hat{x}_i^1, \hat{x}_i^2, \dots, \hat{x}_i^k\} \quad (4)$$

Calculate the cosine distance between x_i and each \hat{x}_i for every feature component, and sum k distances up to get the similarity D_i between the query image and each class of targets in individually feature component:

$$\cos(x_i, \hat{x}_i) = \frac{x_i^T \hat{x}_i}{\|x_i\| \cdot \|\hat{x}_i\|} \quad (5)$$

$$D_i(\Psi(q), c) = \sum_{j=1}^k \cos(x_i, \hat{x}_i)$$

For each image, there are n feature components. So, for each image, we calculate n times distance D_i and sum up them to get the similarity D between the feature of the query image and the feature set of each class of targets:

$$D(\Psi(q), c) = \sum_{i=1}^n D_i(\Psi(q), c) \quad (6)$$

Here, c represents the feature set of various targets of class A, B, C, D, and E.

This metric module considered the challenge when the number of samples is rare, the difference between different kinds of targets is smaller and the distance between the same kind of targets is larger. Moreover, when the quantity of the training samples is small, calculating the similarity between the query image and training images and finding representative training images are possible.

3.2 Episodic Training

In transfer learning, researcher tunes the networks from the source domain with few training samples in the target domain. However, the networks from the source domain are trained with thousands of samples. Thus, it is extremely hard to adjust the source networks to the target domain due to over-fitting. Consequently, this paper adopts a training mode of episodic training. As shown in Figure 4,

the model succeeds in few-shot target recognition on the testing set by simulating similar few-shot target recognition on the training set. Concretely, we randomly sample from the training set to construct thousands of few-shot target recognition tasks to improve the model's ability of generalization. When facing new classes of targets, the model can recognize the targets through training with only a few samples.

In the training stage, we randomly collect samples from the training set to form lots of N-way K-shot learning tasks. In target recognition, the N-way K-shot task refers to learning to recognize N types of targets, and each class of target has K labeled samples that can be used as training data^[15]. Specifically, in this paper, we carried out 5-way 1-shot, 5-shot and 10-shot target recognition tasks in the training stage of the simulation experiment. Thus, we learnt to recognize 5 kinds of targets in each task, each class of targets had only 1, 5, or 10 labeled training samples. Each task was called 1 episode. We carried out 100,000 stages of training tasks, and the recognition result of each episode was fed back to the network to adjust the parameters of our network.

In the testing stage, where there is the real few-shot scene, we performed 5-way 1-shot, 5-shot and 10-shot target recognition tasks, respectively, to test whether the model can recognize targets with few samples when facing new classes of targets.

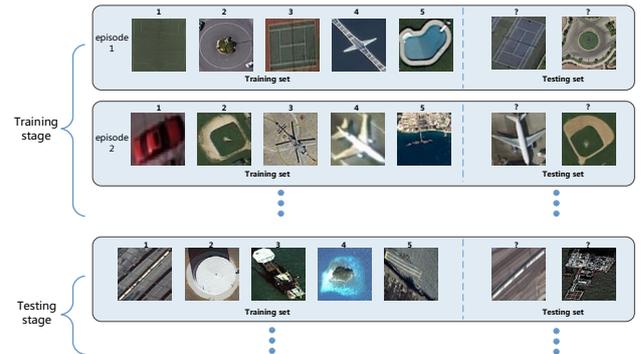


Figure 4. Episodic training

4. Experimental Results

4.1 Datasets

Mini-ImageNet^[16], Stanford Dogs^[17], Stanford Cars^[18], and etc. are datasets that are commonly used in few-shot learning. The pictures in these datasets only contain a single target without complex background, which is helpful for the classifier to learn features of target. However, these datasets mainly contain natural scene images, which are different from remote sensing images. In response to

this problem, this paper constructed a target slices dataset of remote sensing images and conducted few-shot remote sensing image target recognition on this dataset.

Based on the remote sensing image dataset of DOTA [19], NWPU VHR-10 [20-22], RSD-GOD [23] and some remote sensing images from Gaofen-2 satellite, this paper prepared a remote sensing dataset of few-shot classification, which is called RSD-FSC, for target recognition. There are 21 types of targets, including aircraft carrier, oil platform, oil tank, train, island, warship, harbor, missile, large vehicle, swimming pool, bridge, tennis court, airplane, small vehicle, helicopter, roundabout, basketball court, ground track field, baseball court, soccer ball field and airport. The samples are shown in Figure 5. There are about 100,000 remote sensing image target slices with multi-resolution and multi-scale. The number distribution is shown in Table 1.

Table 1. Statistics of RSD-FSC

| NO. | Categories | Instances |
|-----|--------------------|-----------|
| 1 | warship | 1,008 |
| 2 | harbor | 1,875 |
| 3 | missile | 9,916 |
| 4 | large vehicle | 10,063 |
| 5 | swimming pool | 576 |
| 6 | bridge | 10,131 |
| 7 | tennis court | 524 |
| 8 | airplane | 756 |
| 9 | small vehicle | 598 |
| 10 | helicopter | 127 |
| 11 | roundabout | 265 |
| 12 | basketball court | 159 |
| 13 | ground track field | 163 |
| 14 | baseball court | 390 |
| 15 | soccer ball field | 316 |
| 16 | airport | 4,182 |
| 17 | oil platform | 9,957 |
| 18 | aircraft carrier | 70 |
| 19 | island | 11,420 |
| 20 | oil tank | 45,211 |
| 21 | train | 6,733 |

4.2 Experimental Setting

Settings for k-nearest neighbor target recognition network

K-nearest neighbor target recognition network was conducted in a Linux system, using Python 3.7.2 and torch 0.4.1.

During the training stage, we chose aircraft carrier,

train, harbor, missile, large vehicle, swimming pool, tennis court, airplane, small vehicle, helicopter, roundabout, basketball court, ground track field, baseball court, soccer ball field and airport as the auxiliary training set. We conducted 100,000 few-shot target recognition tasks. For each task, we first randomly sampled 5 classes from 16 classes, which would be recognized. Then, we randomly chose 1,5 or 10 samples for each class as the training set and 10 or 15 samples as the testing set. The results of recognition were fed back to adjust the parameters of the network.

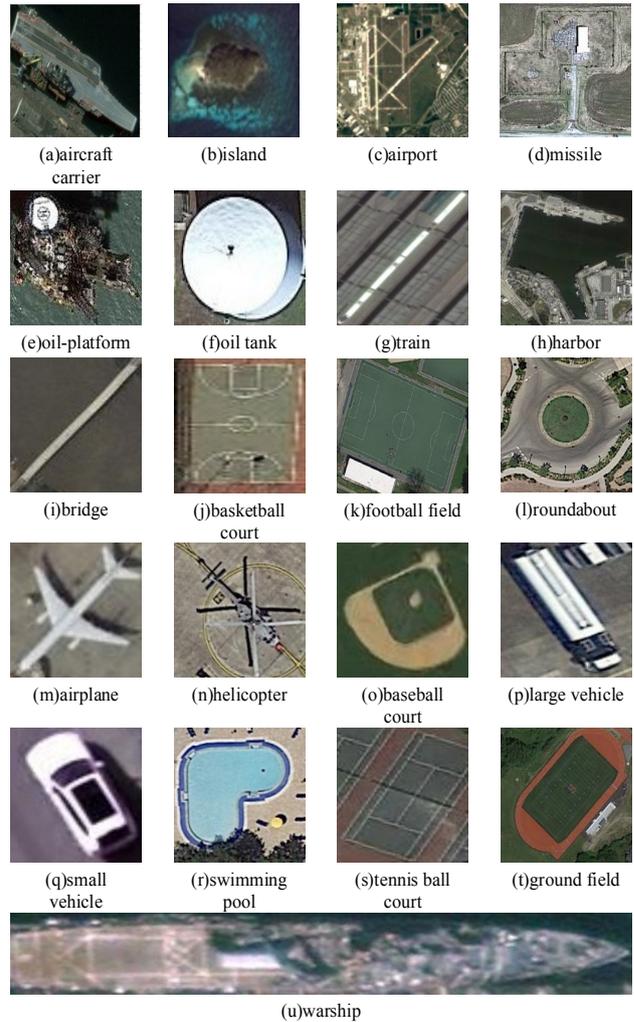


Figure 5. images from RSD-FSC

In the testing stage, to simulate few-shot scenario, we chose oil platform, oil tank, bridge, island and warship as a few-shot set. We conducted 3,000 few-shot target recognition tasks. For each task, based on the few-shot set, we randomly chose 1, 5 or 10 samples for each class as the training set and 10 or 15 samples as the testing set. The results of recognition were counted to calculate the average accuracy of recognition and not fed back to adjust the network’s parameters.

Comparison Methods

First, we trained *ResNet-256F*^[24], which is a deep neural network, with 1, 5, 10 and thousands of samples for each class from the few-shot set. The aim of the training is to verify whether the deep learning can work out few-shot target recognition.

Second, to prove the efficiency of k-nearest neighbor target recognition network, we conducted a comparative experiment with prototype network^[12] and matching network^[16] on RSD-FSC. Concretely, we conducted 5-way 1-shot, 5-shot or 10-shot tasks respectively to calculate the average accuracy of recognition.

4.3 Results

The results are shown in TABLE II. Our result showed that when training samples were sufficient, *ResNet-256F* performed efficiently, but when the quantity of training samples was decreased, it can't recognize targets accurately. Interestingly, the accuracy decreased when the quantity of training samples increased from 1 to 5, which was caused by overfitting.

However, our proposed model combining episodic training and metric learning generalized well when facing new emerged classes in the few-shot scenario. After thousands of episodes of all few-shot target recognition tasks, the proposed model overcame overfitting and well recognized the targets with few training samples. Furthermore, the proposed model used *k*-NN in measurement module to calculate the similarity between the query image and each class of targets, which overcame the problem that the inter-class distance was small and intra-class distance was large. Compared with prototype network and matching network, it achieved 7.52% and 10.75% improvements in the 5-way 1-shot classification task. In the 5-way 5-shot classification task, the model achieved 15.24% and 21.20% improvements. In the 5-way 10-shot classification task, our model achieved 16.62% and 18.71% improvements. This was because that these models utilized all the training samples to measure similarity, and didn't effectively limit the classifier due to the effect of smaller inter-class distances and larger intra-class distances.

Table 2. Average Precision of Recognition (%)

| Model | 1-shot (k=3) | 5-shot (k=3) | 10-shot (k=3) |
|------------------------------------|-----------------|-----------------|------------------|
| Ours | 59.134 | 82.553 | 87.796 |
| <i>ResNet-256F</i> ^[24] | 26.700 | 20.000 | 5.000 |
| Prototype network ^[12] | 51.617 | 67.315 | 71.181 |
| Matching network ^[16] | 48.384 | 61.350 | 69.091 |

Moreover, according to the recognition results, we found that compared methods tended to confuse oil-platform with oil tank, train with bridge and helicopter with airplane.

As shown in Figure 6, the helicopter is tend to be recognized as an oil tank by the compared methods. The reason is that the pattern composed by the apron and airscrew is similar to the pattern on the top of the oil tank. However, it can be efficiently recognized by our method.

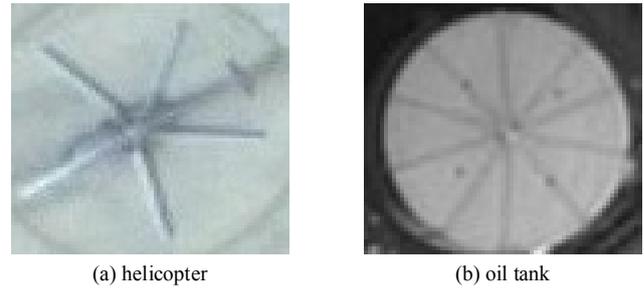


Figure 6. Comparison of helicopter and oil tank

Furthermore, as shown in Figure 7, it's easy to find that textures of bridge and train in remote sensing images are extremely similar, which leads to low recognition accuracy.

Similarly, some oil platforms tend to be confused with the warship, too.

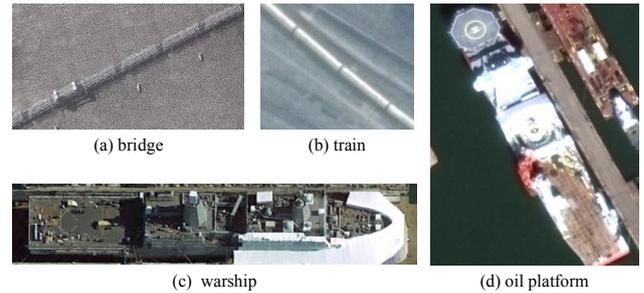


Figure 7. Confusion examples

5. Conclusions

In this paper, first, we constructed a dataset RSD-FSC for few-shot target recognition. Second, based on metric learning, we proposed a k-nearest neighbor target recognition network, which can overcome the deterioration of smaller inter-class distance and larger intra-class distance caused by few-shot. Third, we introduced episodic training, which can overcome overfitting and enable the model generalize to new classes of targets in few-shot target recognition. Finally, we carried out 5-way 1-shot, 5-shot and 10-shot experiments using the k-nearest neighbor target recognition network and contrast experiments on the dataset RSD-FSC. The experiment results show that in the few-shot target recognition, the k-nearest neighbor

target recognition network proposed in this paper is useful and superior to the compared networks. The method proposed in this paper is of great significance to promote the research of few-shot target detection and recognition technology in remote sensing images.

Funding

This work was supported in part by the CETC key laboratory of Aerospace Information Applications under Grant No.SXX19629X060.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] Ren, S., He, K., Girshick, R., et al., 2015. Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*. pp. 91-99.
- [2] Redmon, J., Divvala, S., Girshick, R., et al., 2016. You only look once: Unified, real-time object detection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. pp. 779-788.
- [3] Girshick, R., Donahue, J., Darrell, T., et al., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. pp. 580-587.
- [4] Liu, W., Anguelov, D., Erhan, D., et al., 2016. SSD: Single shot multibox detector. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. vol. 9905 LNCS. pp. 21-37. DOI: https://doi.org/10.1007/978-3-319-46448-0_2.
- [5] Boiman, O., Shechtman, E., Irani, M., 2008. In defense of nearest-neighbor based image classification. *26th IEEE Conf. Comput. Vis. Pattern Recognition. CVPR*. pp. 1-8.
- [6] Goodfellow, I.J., et al., 2014. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* 3, 2672-2680.
- [7] Radford, A., Metz, L., Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.
- [8] Mao, X., Li, Q., Xie, H., et al., 2017. Least Squares Generative Adversarial Networks. *Proceedings of the IEEE International Conference on Computer Vision*. pp. 2813-2821.
- [9] Durugkar, I., Gemp, I., Mahadevan, S., 2016. Generative multi-adversarial networks. *arXiv preprint arXiv:1611.01673*.
- [10] Ratner, A.J., Ehrenberg, H.R., Hussain, Z., et al., 2017. Learning to compose domain-specific transformations for data augmentation. *Advances in Neural Information Processing Systems*. pp. 3237-3247.
- [11] Alfassy, A., Karlinsky, L., Aides, A., et al., 2019. Laso: Label-set operations networks for multi-label few-shot learning. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. pp. 6541-6550.
- [12] Snell, J., Swersky, K., Zemel, R., 2017. Prototypical networks for few-shot learning. *Advances in Neural Information Processing Systems*. pp. 4078-4088.
- [13] Chen, L., Wu, H., Cui, X., et al., 2018. Convolution neural network SAR image target recognition based on transfer learning. *Chinese Space Science and Technology*. 38.6(2018), 46-51. (in Chinese)
- [14] Jamal, M.A., Qi, G.J., 2019. Task agnostic meta-learning for few-shot learning. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. pp. 11711-11719.
- [15] Wang, Y., Yao, Q., Kwok, J.T., et al., 2020. Generalizing from a Few Examples: A Survey on Few-shot Learning. *ACM Comput. Surv.* 53(3), 1-34.
- [16] Oriol, V., Charles, B., Tim, L., 2016. kavukcuoglu koray, and W. Daan. Matching Networks for One Shot Learning. *Advances in Neural Information Processing Systems*. pp. 3630-3638.
- [17] Khosla, A., Jayadevaprakash, N., Yao, B., et al., 2011. Novel dataset for fine-grained image categorization: Stanford dogs. *Proc. CVPR Workshop on Fine-Grained Visual Categorization (FGVC)*. 2(1).
- [18] Krause, J., Stark, M., Deng, J., 2013. 3D object representations for fine-grained categorization. *Proceedings of the IEEE International Conference on Computer Vision*. pp. 554-561.
- [19] Xia, G.S., Bai, X., Ding, J., et al., 2018. DOTA: A Large-Scale Dataset for Object Detection in Aerial Images. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. pp. 3974-3983.
- [20] Cheng, G., Han, J., Zhou, P., et al., 2014. Multi-class geospatial object detection and geographic image classification based on collection of part detectors. *ISPRS J. Photogramm. Remote Sens.* 98, 119-132.
- [21] Cheng, G., Han, J., 2016. A survey on object detection in optical remote sensing images. *ISPRS J. Photogramm. Remote Sens.* 117, 11-28.
- [22] Cheng, G., Zhou, P., Han, J., 2016. Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing

- Images. *IEEE Trans. Geosci. Remote Sens.* 54(12), 7405-7415.
- [23] Zhuang, S., Wang, P., Jiang, B., et al., 2019. A single shot framework with multi-scale feature fusion for geospatial object detection. *Remote Sens.* 11(5), 594.
- [24] He, K., Zhang, X., Ren, S., et al., 2016. Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, Las Vegas. 2016-December, 770-778.