Target recognition in synthetic aperture radar image based on PCANet

Baogui Qi1, Haitao Jing2, He Chen1,*, Yin Zhuang1, Zhuo Yue1, Chonglei Wang1
1Beijing Key Laboratory of Embedded Real-time Information Processing Technology, Beijing Institute of Technology, Beijing 100081, People’s Republic of China
2Shanghai Institute of Satellite Engineering, Shanghai 200240, People’s Republic of China
*E-mail: chenhhe@bit.edu.cn

Abstract: Automatic targets recognition (ATR) for synthetic aperture radar (SAR) image is very important. ATR can be used in traffic management, national frontier safety, and so on. Traditional algorithms for SAR ATR is composed of extraction features and training classifier. The features are essential for the classification accuracy. However, choosing good features by hand is a hard task. The deep convolutional neural networks (CNNs) which can learn features automatically have got a great performance in natural images. However, the CNNs have many parameters and need a lot of data to train such networks. The remote-sensing data of SAR is limited. Then, the authors need a simple network which needs not much data and easy to train. The principal component analysis network (PCANet) is a shallow network that performs well in the recognition task and needs no hand features choosing. Though this network has produced a wide application in the natural images, it is rarely used in the SAR images. The experimental result of the moving and stationary target acquisition and recognition (MSTAR) dataset shows that the PCANet can achieve over 99% accuracy on ten-class targets. This result is better than traditional algorithms and is very close to the results of deep-learning methods.

1 Introduction

With the development of remote-sensing technology, the amount of remote-sensing data are increased fast [1]. Target recognition based on remote-sensing data is also increasing. Automatic targets recognition (ATR) for SAR image is used in traffic management, national frontier safety, and many fields. The synthetic aperture radar (SAR) can get high-resolution images in all-weather and all-day [2]. So, SAR images are wildly used in the target detection and classification. A good ATR method can help us to more effectively deal with many works.

The traditional method of classification is to extract some handcrafted features and train a classifier [3]. Wang et al. proposed a classification scheme based on geometric and backscattering characteristics [4]. Ding et al. proposed a target recognition method by exploiting the azimuth sensitivity information (ASI) [5]. The ASI is constructed based on the original SAR images, which can describe the azimuth sensitivity of a certain target class at a special azimuth. Zhu et al. proposed an improved shape contexts method which can describe the topology and intensity of the scattering points of the targets [6]. These methods need to use the features designed by researchers. These handcrafted feature extraction methods have some drawbacks. The first is that we do not know which features are best to train the classifier. The second is that it needs to change features for different targets type. Different targets have different effective features. Then, we could not know if there are more features which can improve the performance of recognition.

With the development of deep-learning method, many image-recognition fields have achieved great performance. The deep convolutional neural networks (CNNs) can automatically learn the feature from a large dataset. It is challenging to utilise deep CNNs for SAR target classification because the CNNs have many parameters to train and require much training data. However, it is hard to get enough SAR data and train a CNN. Some researchers had made some efforts to cope with this problem. Lin et al. proposed the convolutional highway unit to train deeper networks with limited SAR data [7]. Chen et al. presented a new all-convolutional networks (A-ConvNets), which only consists of sparsely connected layers, without fully connected layers being used [8]. Huang et al. proposed a transfer learning-based method, making knowledge learned from sufficient unlabelled SAR scene images transferrable to labelled SAR target data [9]. These methods need less data than the usual CNN, while such networks are also complicated and not easy to train.

Therefore, we need a simple network which has fewer parameters and is easy to train. The principal component analysis network (PCANet) is a shallow network that presents great performance in the recognition mission and needs no hand-crafted features choosing [10]. Though this network has produced a wide application in the natural images, it is rarely used in the SAR images.

In this letter, we attempt to use the PCANet which can learn powerful high-level feature representations with few SAR training data. Then, the network has fewer parameters than CNNs. The experimental result of the moving and stationary target acquisition and recognition (MSTAR) dataset shows that the PCANet can achieve over 99% accuracy on ten-class targets. Then, this result is better than traditional hand-crafted features methods.

The rest of this paper is organised as follows. Section 2 introduces the PCANet. Experimental results are provided in Section 3. Conclusions and future work are provided in Section 4.

2 Description of PCANet

In this section, we describe the PCANet. The structure of PCANet is illustrated in Fig. 1. The number of training images is \( M \). \( I_i \) is the \( i \)th image. The size of each image is \( I \times r \). The patch (or 2D filter) size is \( s_1 \times s_2 \). The method of learning filters from training images is as follows:

For each pixel in an image, we take \( s_1 \times s_2 \) patch and collect all patches of the \( i \)th image. We use \( p_{ij} \) to denote the \( j \)th patch in \( I_i \). After subtracting patch mean from each patch, we obtain \( P_i = [p_{i1}, p_{i2}, \ldots, p_{ik}] \), where \( p_{ij} \) is a mean-removed patch. Then, we get

\[
P = [P_1, P_2, \ldots, P_M].
\]
Table 1 Number of training and testing images

| class  | Train (depression: 17°) | Test (depression: 15°) |
|--------|-------------------------|------------------------|
|        | no. images              | no. images             |
| BTR60  | 256                     | 195                    |
| 2S1    | 299                     | 274                    |
| BRDM2  | 298                     | 274                    |
| D7     | 299                     | 274                    |
| T62    | 299                     | 273                    |
| ZL131  | 299                     | 274                    |
| ZSU234 | 299                     | 274                    |
| BMP2   | 233                     | 196                    |
| BTR70  | 233                     | 196                    |
| T72    | 232                     | 196                    |

Table 1, which are captured at 15° and 17° depression angles.

If the filter number in stage $i$ is $L_i$, PCA minimises the reconstruction error within a family of orthonormal filters, i.e.

$$
\min \| P - VV^T P \|_F^2, \quad \text{s.t.} \quad VV^T = I_{L_i},
$$

(2)

where $I_{L_i}$ is identity matrix of size $L_i \times L_i$. The solution is the $L_i$ principal eigenvector of $PP^T$. The PCA filters are expressed as

$$
W_i = \text{mat}_{2N \times 2N}(q(PP^T)),
$$

(3)

where $\text{mat}_{2N \times 2N}(x)$ maps $x$ to a matrix $W_i$, and $q(PP^T)$ denotes the $i^{th}$ principal eigenvector of $PP^T$.

The second stage shares the same progress with the first stage. The only difference is the input data. The input of the second stage is the output of the first stage. The $l^{th}$ filter output can be expressed as

$$
L_i^l = L_i \times W_i^l, i = 1, 2, ..., N,
$$

(4)

where $*$ denotes 2D convolution. After the same progress, we get the output of the second stage, which is

$$
O_i^l = \{L_i^l \times W_i^j\}_{j=1}^{2L_i^2}.
$$

(5)

The output number of the first stage is $L_1$, and every output of the first stage has $L_2$ outputs, and the output number of the second stage is $L_1 \times L_2$.

Then, we binarise the outputs and get $H(O_i^l)$, where $H(\cdot)$ is a Heaviside step function. Then, we convert the $L_2$ outputs in $O_i^l$ into a single integer-valued image:

$$
T_i^l = \sum_{j=1}^{L_2} 2^{l-1}H(O_i^j).
$$

(6)

For each of the $L_1$ images $T_i^l$, we partition it into $B$ blocks. We compute the histogram of the decimal values in each block, and concatenate all the $B$ histograms into one vector and denote as $Bhist(T_i^l)$. Then the feature of the input image is defined to be set of block-wise histograms

$$
\mathbf{f}_i = \left[ Bhist(T_i^1), \ldots, Bhist(T_i^{L_1}) \right]^T.
$$

(7)

Then, we can train a support vector machine (SVM) classifier use these features. Then, we can use this PCANet and SVM to recognise a new target.

3 Experimental results

The experiment data used in this letter were collected by the Sandia National Laboratory SAR sensor platform. The data is called MSTAR data set. Those images were collected using an X-band HH-polarisation SAR sensor, in a 0.3 m*0.3 m resolution spotlight mode, full aspect coverage. The ten target classes are listed in Table 1, which are captured at 15° and 17° depression angles. Examples of ten target classes of SAR images and optical images are shown in Fig. 2. The MSTAR data set is widely used to test the performance of SAR recognition algorithms.

In our experiment, the PCANet is evaluated on the ten targets classification problem. The train data are images that captured at 17° depression angle and the test data are images that captured at 15° depression angle. We do not do any particular pre-processing algorithm for those SAR images. To find out the best performance of different parameters. We test some examples. When $s_1 = s_2 = 7, L_1 = L_2 = 8$, we test the influence of train data number. The results are shown in Table 2. The average accuracy is better when the data is more. Then, we test the influence of different size of $s_1$. The results are shown in Table 3. With the size of $s_1$ becoming larger, the accuracy first becomes higher and then lower. This is related to the targets sizes. We also change the number of $L_1$, the result is shown in Table 4. Some essential parameter values of PCANet in

![Detailed block diagram of PCANet](http://creativecommons.org/licenses/by/3.0/)

Fig. 1 Detailed block diagram of PCANet

This is an open access article published by the IET under the Creative Commons Attribution License
the experiment are as follows, $s_1 = s_2 = 10$, $L_1 = 10$, $L_2 = 8$. Table 5 shows the confusion matrix of the ten targets classification. Each row in the confusion matrix denotes the actual target class, and each column represents the class predicted by the algorithm. ZIL131, ZSU234, and BTR70 can be recognised with a probability of 100%. The average accuracy of all targets is 99.22%.

We also compare the performance of PCANet with some state-of-the-art methods. These methods include SVM [11, 12], AdaBoost [12, 13], modified polar mapping classifier (M-PMC) [14], ASI [5], and A-ConvNets [8]. The classification accuracies of these algorithms are listed in Table 6. As all those algorithms use the same data, we just cite the results published in papers. We can see that the PCANet has a better performance even when compared with other CNNs methods.

### Table 2

| Number | 300 | 675 | 900 | 1350 | 2700 |
|--------|-----|-----|-----|------|------|
| average accuracy(%) | 73.15 | 86.14 | 90.56 | 93.03 | 97.40 |

### Table 3

| Size of $s_1$, $s_2$ | 7,7 | 10,10 | 12,12 | 15,15 | 17,17 |
|---------------------|-----|-------|-------|-------|-------|
| average accuracy(%) | 93.03 | 96.78 | 97.61 | 98.23 | 97.80 |

### Table 4

| Number of $L_1$ | 4 | 6 | 8 | 10 | 12 |
|-----------------|---|---|---|----|----|
| average accuracy(%) | 93.98 | 93.11 | 93.03 | 92.93 | 92.74 |

### Table 5

| Class | BTR60 | 2S1 | BRDM2 | D7 | T62 | ZIL131 | ZSU234 | BMP2 | BTR70 | T72 | P (%) |
|-------|-------|-----|-------|----|-----|--------|--------|------|-------|-----|-------|
| BTR60 | 192   |     | 2     |    |     | 1      |        |      |       |     | 98.46 |
| 2S1   | 269   |     | 1     | 1  |     | 1      | 1      | 1    | 1     | 2   | 98.18 |
| BRDM2 |     | 271 |       | 2  |     | 2      | 1      | 1    | 1     |     | 99.27 |
| D7    | 1     | 272 |       | 1  |     | 1      |        |      | 1     | 2   | 98.91 |
| T62   |     |     | 271   | 1  | 1   |        | 274    |      | 1     | 2   | 99.27 |
| ZIL131|     |     |       |    | 1   | 1      |        |      | 1     | 1   | 99.27 |
| ZSU234|     |     |       |    |     | 274    |        |      | 1     | 1   | 100  |
| BMP2  |     |     |       |    |     |        | 194    |      | 1     | 1   | 98.99 |
| BTR70 |     |     |       |    |     |        |        |      | 196   | 1   | 100  |
| T72   |     |     |       |    |     |        |        |      |       | 194 | 98.99 |
| average|     |     |       |    |     |        |        |      |       |     | 99.22 |

### Conclusion and future work

In this paper, we use PCANet to cope with the SAR target recognition task. PCANet is a simple network which has fewer parameters and is very easy to train. It can automatically learn and extract hierarchical features like CNNs. The experimental result of the MSTAR data set shows that the PCANet can achieve over 99% accuracy on ten-class targets. Then, this network performs better than the traditional methods of hand-crafted features extraction. In the future work, we will do some modification to make sure this network would be more suitable for SAR target. Then, we will realise a hardware system which can be used in a real-time recognition task.

### Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant no. 91438203, the Chang Jiang Scholars Program under Grant T2012122, and the Hundred...
Table 6 Classification accuracy of PCANet and some state-of-the-art methods

| Method      | Probability of correct classification (%) |
|-------------|-------------------------------------------|
| SVM         | 90                                        |
| AdaBoost    | 92                                        |
| M-PMC       | 98.8                                      |
| ASI         | 95.03                                     |
| A-ConvNets  | 99.13                                     |
| PCANet      | 99.22                                     |

Leading Talent Project of Beijing Science and Technology under Grant Z141101001514005.

6 References

[1] Zhu, X.X., Tuia, D., Mou, L., et al.: ‘Deep learning in remote sensing: A comprehensive review and list of resources’, IEEE Geosci. Remote Sens. Mag., 2017, 5 (4), pp. 8–36

[2] Ouchi, K.: ‘Recent trend and advance of synthetic aperture radar with selected topics’, Remote Sens.-Basel, 2013, 5 (2), pp. 716–807

[3] El-Darymli, K., Gill, E.W., McGuire, P., et al.: ‘Automatic target recognition in synthetic aperture radar imagery: a state-of-the-art review’, IEEE. Access., 2016, 4, pp. 6014–6058

[4] Wang, C., Zhang, H., Wu, F., et al.: ‘A novel hierarchical ship classifier for COSMO-SkyMed SAR data’, IEEE Geosci. Remote Sens. Lett., 2014, 11, (2), pp. 484–488

[5] Ding, B., Wen, G., Huang, X., et al.: ‘Target recognition in SAR images by exploiting the azimuth sensitivity’, Remote. Sens. Lett., 2017, 8, (9), pp. 821–830

[6] Zhu, J.W., Qiu, X.L., Pan, Z.X., et al.: ‘An improved shape contexts based ship classification in SAR images’, Remote Sens.-Basel, 2017, 9, (2)

[7] Lin, Z., Ji, K., Kang, M., et al.: ‘Deep convolutional highway unit network for SAR target classification with limited labeled training data’, IEEE Geosci. Remote Sens. Lett., 2017, 14, (7), pp. 1091–1095

[8] Chen, S.Z., Wang, H.P., Xu, F., et al.: ‘Target classification using the deep convolutional networks for SAR images’, IEEE Trans. Geosci. Remote Sens., 2016, 54, (8), pp. 4806–4817

[9] Huang, Z., Pan, Z., Lei, B: ‘Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data’, Remote Sens.-Basel, 2017, 9, (9), p. 907

[10] Chan, T.H., Jia, K., Gao, S., et al.: ‘PCANet: A simple deep learning baseline for image classification?’, IEEE Trans. Image Process., 2015, 24, (12), pp. 5017–5032

[11] Zhao, Q., Principe, J.C: ‘Support vector machines for SAR automatic target recognition’, IEEE Trans. Aerosp. Electron. Syst., 2001, 37, (2), pp. 643–654

[12] Srinivas, U.: ‘In SAR automatic target recognition using discriminative graphical models’, IEEE Int. Conf. on Image Processing, ICIP 2011, Brussels, Belgium, September 2014, pp. 33–36

[13] Sun, Y., Liu, Z., Todorovic, S., et al.: ‘Adaptive boosting for SAR automatic target recognition’, IEEE Trans. Aerosp. Electron. Syst., 2007, 43, (1), pp. 112–125

[14] Park, J.J., Kim, K.T.: ‘Modified polar mapping classifier for SAR automatic target recognition’, IEEE Trans. Aerosp. Electron. Syst. Aes, 2014, 50, (2), pp. 1092–1107