A novel semi-supervised method based on active sample proposal and modified self-learning for SAR target discrimination

Yan Wang[a,b*], Dan Li[a,c], Xiaobo Deng[a], Feng Xu[b]
[a] AVIC Leihua Electronic Technology Research Institute, Wuxi, China; [b] Key Lab for Information Science of Electromagnetic Waves, Fudan University, Shanghai, China; [c] Electronic Information College, Northwestern Polytechnical University, Xi'an, China

ABSTRACT

Due to the visual difference between the synthetic aperture radar (SAR) image and optical image, it is difficult to accurately label the SAR image, leading to the fact that there are only few labeled images with lots of unlabeled images. In this situation, the self-learning, which utilizes both the labeled samples and unlabeled samples to learn an optimal classifier, performs well. For self-learning, the performance of the initial classifier has a great influence on the following learning. During the self-leaning procedure, the classifier might easily get incorrect predictions of unlabeled samples provided by itself especially in original rounds when the accuracy of classifier is undesirable. Therefore, the performance of self-learning is usually unstable. Based on the active sample proposal and modified self-learning, this paper gives a novel semi-supervised method for SAR target discrimination. Firstly, an undirected graph is constructed by using all the unlabeled samples and the most informative samples are selected to be labeled by man. Secondly, confidence of the classifier's prediction on unlabeled samples is achieved by both the discrimination result and local geometrical information. Experimental results on the measured SAR dataset illustrate the proposed semi-supervised discrimination method can still obtain good discrimination performance with few labeled samples.

Keywords: SAR (synthetic aperture radar) image, self-learning, semi-supervised method, target discrimination

1. INTRODUCTION

For its capacity of generating high resolution images free from lighting and weather conditions synthetic aperture radar (SAR) has found wide applications both in civilian and military fields, such as earth measurement, urban planning and intelligence reconnaissance[1]. Meanwhile, SAR automatic target recognition (ATR) possesses great military potential and therefore draws much more attentions in the last decades. Lincoln lab proposes three stage process framework for SAR ATR based on attention mechanism[2, 3], including detection, discrimination and classification. Target discrimination aims at decreasing the classification cost by removing the false alarms produced during the detection process. As the linking stage between detection and classification, discrimination has a huge effect on the performance of the whole SAR ATR system[3, 4] and therefore attracts increasing attention in the past decades.

The traditional SAR target discrimination is generally accomplished in a supervised way where all the training samples are needed to be labeled. The good performance of supervised methods is guaranteed by sufficient labeled samples. However, in real application, to obtain enough labeled samples is usually difficult, high-cost and time consuming. The coherent imaging mechanism of SAR image makes the situation worse. The lack of labeled samples makes it difficult to properly capture the underlying pattern of the image. Due to these restrictions of the supervised learning, the semi-supervised learning, which utilizes both the labeled samples and unlabeled samples for training[5], occurs and is applied to various applications.

Classical semi-supervised learning methods contain self-learning[6], co-training, and transductive SVM (TSVM)[7]. Self-learning is a simple but effective semi-supervised method which first weakly trains a classifier with limited labeled samples and then choose a certain number of unlabeled samples with the most confident labels provided by the classifier to update the labeled training samples. This process repeatedly proceeds until some criterion is satisfied. The initial classifier trained with few true labeled samples has great effect on the accuracy of self-learning. If the labeled samples are not representative, the discrimination performance of the final classifier will greatly degrade. Moreover, in each repetition, the standard self-

* wangyanpx@163.com

Third International Conference on Computer Science and Communication Technology (ICCSCT 2022)
edited by Yingfa Lu, Changbo Cheng, Proc. of SPIE Vol. 12506, 1250662
 · 0277-786X · doi: 10.1117/12.2661824

learning only utilizes the classifier's posteriori probability outputs as the measurement of labeling confidence on unlabeled samples, which makes the labeling confidence away from errorless. Therefore, when using the self-learning method in SAR target discrimination, the discrimination is unstable.

In this paper, propose a novel semi-supervised SAR target discrimination based on active sample proposal and modified self-learning to improve the SAR target discrimination performance. The major contributions of the proposed method are two-fold. Firstly, the initial few labeled samples are actively selected for man to be labeled rather than randomly selected. Specially, all the unlabeled samples are used to construct an undirected graph with the samples as its nodes and the similarity between them as the edge weights. The most informative samples are selected through the extended Nyquist-Shannon sampling theory[8] to be labeled by man as the initial labeled samples. Secondly, in each self-learning round, the classifier's posteriori probability and local geometrical information are jointly utilized to reliably determine the labeling confidence on unlabeled samples.

The rest of this paper is organized as follows. In section II, we present the improved semi-supervised discrimination method. In, section III, we perform experiments on the measured SAR dataset and analyze the results. Finally, we provide some conclusions on this paper in section IV.

2. METHODOLOGY

As shown in Figure 1, the proposed semi-supervised method for SAR target discrimination mainly includes two stages: the training stage and test stage. Meanwhile, the training stage mainly comprises three parts: (1) the Lincoln feature extraction; (2) the active samples proposal; (3) modified self-learning on Lincoln features. The training stage aims at learning an optimal classifier using a small amount of labeled SAR chips and a large amount of unlabeled SAR chips. The test stage is similar with the training stage except that the optimal SVM classifier obtained in the training stage is directly utilized to discriminate the test SAR chips. Since this paper is mainly concentrated on the training stage, we minutely present the process of training stage in the following.

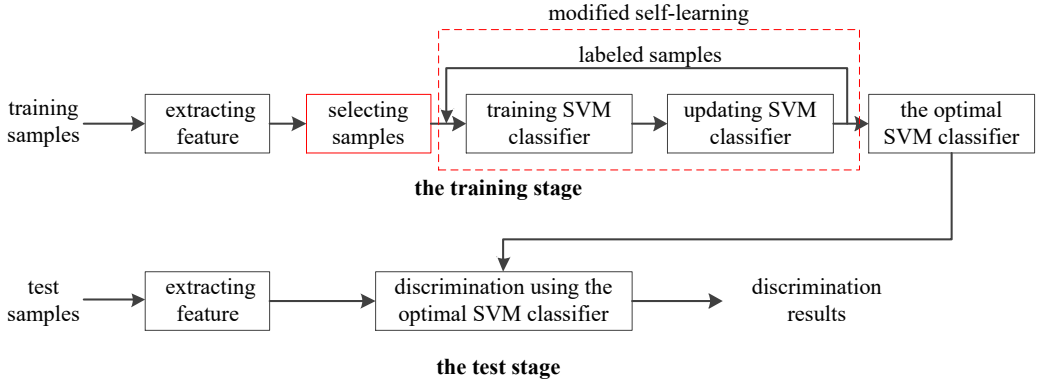


Figure 1. Framework of the semi-supervised discrimination method.

2.1 Lincoln feature extraction

For realizing SAR target discrimination, the Lincoln Lab designed 23 Lincoln features to capture the difference between the target and clutter, which has been used in the real SAR target recognition system. In the proposed semi-supervised target discrimination task, two texture-based features are chosen as the discrimination features since they can effectively discriminate target from clutter.

(1) The log standard deviation feature: The standard deviation of an image measures the fluctuation of pixel intensities. Usually, the intensities of a target are more greatly fluctuated than those of the clutter and therefore have the larger standard deviation. To alleviate the effect of speckle noise in SAR images, the log standard deviation σ is utilized in the proposed method and can be obtained by

$$\sigma = \sqrt{\frac{S_2 - \frac{S_1}{T}}{T-1}} \quad \text{where}$$

$$S_1 = \sum_{x,y \in region} 10\, log_{10}\, I\,(x,y)$$
$$S_2 = \sum_{x,y \in region} [10\, log_{10}\, I\,(x,y)]^2 \quad (1)$$

$I(x,y)$ is the intensity value of the pixel located at (x,y), and T represents the number of pixels in the SAR image.

(2) The fractal dimension feature: Another widely used Lincoln feature is the fractal dimension, which measures how the pixels with large intensity value in a binary image distribute in the spatial space. It describes the spatial dimension of an object. Generally speaking, the bright pixels of a target chip usually concentrated around the target while those of clutter are randomly distributed. Therefore, we use the fractal dimension to separate the target from clutter. The fractal dimension is defined as the Hausdorff dimension of N brightest pixels in a binary image and calculated as follows:

$$H_d = \frac{log_{10}\, N_1 - log_{10}\, N_2}{log_{10}\, d_2 - log_{10}\, d_1} \quad (2)$$

where N_1 and N_2 denotes the number of 1-pixel-by-1-pixel boxes and 2-pixel-by-2-pixel boxes that are needed to cover the bright region in the binary image, respectively; d_1 and d_2 represent the size of boxes, and $d_1 = 1, d_2 = 2$. Since the target is more densely distributed than the clutter, $1 < H_d < 2$ for the target and $H_d < 1$ for the clutter.

2.2 Active samples proposal

After extracting the Lincoln feature, a graph is constructed with all the unlabeled samples and the most informative samples are selected. These selected samples will be labeled by man and utilized to train the initial classifier. In this paper, the theoretical finding of Akshay Gadde et al.[8] on graph signal sampling is adopted to facilitate the sample selection.

Concretely, we construct an undirected neighborhood graph $G = (V, E)$ with N unlabeled samples as its nodes and edges $E = \{(x_i, y_i, w_{ij})\}, x_i, x_j \in V, i = 1, \cdots, N, j = 1, \cdots, N$, where (x_i, x_j, w_{ij}) represents an edge with weights w_{ij} between nodes x_i and x_j, denoting the similarity between the respective nodes. A graph is eminently suitable for describing the relationship among elements in the dataset. There are three matrices widely used in the graph, including the degree matrix D, the adjacency matrix W and the combinatorial Laplacian matrix L. $D = diag\{d_1, d_2, \ldots, d_N\}$ is a diagonal matrix, where d_i is the degree of the node x_i and obtained by summing the weights of edges connected to node x_i. W is an $N \times N$ matrix with $W_{ij} = w_{ij}$ and $L = D - W$. In this paper, the symmetric normalized form of the Laplacian matrix $\mathrm{L} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}$ is used, which is a symmetric positive semi-definite matrix with a set of real eigenvalues and a corresponding orthogonal set of eigenvectors.

In the field of graphs, a graph signal is defined as a scalar-valued discrete mapping $f\colon V \to \mathbb{R}$, and the value of the signal on x_i is denoted by $f(x_i)$. We suppose $S \in V$ is a subset of the graph and $S^c = V \backslash S$ is its complement set. By reserving the value of signal on a subset of nodes S, i.e., the sampling set, we can realize sampling a graph signal f onto S.

In the realm of traditional signal processing, the Nyquist-Shannon sampling theorem provides an upper bound for a signal's bandwidth, by which the signal could be accurately reconstructed given a certain sampling rate. Analogously, in the field of graph, a concept of frequency is also needed. The eigenvalues and eigenvectors of the Laplacian matrix L can provide such a spectral interpretation. Meanwhile, the eigenvalues can be regarded as frequencies and imply the change of eigenvector. That is to say when an eigenvalue is high, its corresponding eigenvector will violently change9. The orthogonal eigenvectors can span a subspace in $\mathbb{R}^N$ and the projection of a signal onto the eigenvectors is regarded as its Graph Fourier Transform (GFT), analogous with the Fourier transform in traditional signal processing.

Let $L_2(S^c)$ be the space of all graph signals that are zero everywhere except possibly on the nodes in S^c, i.e., $\forall \phi \in L_2(S_c), \phi(S) = 0$. Moreover, suppose $\omega(\phi)$ is the bandwidth of a graph signal ϕ, i.e., the maximum non-zero frequency of the signal. According to the Theorem Anis *et al.* have proved[9], for a set S, the maximum cut-off frequency can be found by searching the bandwidth $\omega(\phi^*)$ of the smoothest possible signal $\phi^* \in L_2(S^c)$. From the work by Antonio[9], the true cut-off frequency $\omega_c(S)$ can be relaxed by $\Omega_k(S)$, where $k > 0$ is a given integer parameter. The larger k is, the better the estimate of cut-off frequency is. $\Omega_k(S)$ and ϕ_k^* can be numerically obtained from the smallest eigenpair $(\sigma_{1,k}, \psi_{1,k})$ of the reduced matrix $\left(\mathrm{L}^{\,k}\right)_{S^c}$:

$$\Omega_k(S) = \sigma_{1,k} \quad (3)$$

$$\phi_k^*(S^c) = \psi_{1,k}, \phi_k^*(S) = 0 \tag{4}$$

Then the optimal sampling set S_{opt} can be found according to algorithm 1.

Algorithm 1. Greedy process for searching S_{opt}

Input: undirected graph $G = \{V, E\}$ on unlabeled samples, Laplacian matrix $\mathcal{L}$, target size L, parameter $k \in \mathbb{Z}^+$, $S = \{\emptyset\}$

if $|S| < L$ do

For S, obtain the smoothest signal $\phi_k^* \in L_2(S^c)$ by Eq. (3) and Eq.(4).

Update v with $argmax_{x_i}[(\phi_k^*(x_i)^2)]$.

Update S with $S \cup v$

end if

$S_{opt} \leftarrow S$

2.3 Modified self-learning

The standard self-learning is the simplest semi-supervised learning paradigm. It firstly trains a classifier with limited labeled samples and then predicts the labels on the unlabeled samples. The most confident predictions will be regarded as labeled samples to update the labeled training set. During the train stage, this process proceeds repeatedly until some criterion is achieved. As the classifier is not absolutely correct, its predictions on unlabeled samples are usually away from errorless.

In order to guarantee the reliability of labels on unlabeled samples, we incorporate the classifier's posteriori probability and local geometrical information to evaluate the confidence of whether the label is correct. According to the manifold assumption, a sample usually has the same label with most of its neighboring samples, which can be used to estimate the reliability of the labels on unlabeled samples. For self-learning, the initial labeled samples are completely correct. Therefore, the local geometrical relationship between the unlabeled samples and true labeled samples can be exploited as the basis of determining the confidence of unlabeled samples.

Let $Pr(y = 1)$ denote the posteriori probability of a sample being positive, which is estimated by the classifier. Then, the labeling confidence of each sample is estimated based on the following equation:

$$J_p = Pr(y = 1) \cdot \left(\frac{p_c}{c}\right) \tag{5}$$

where p_c represents the number of positive sample in the c neighbors.

Let L be M labeled training samples selected by the method in 2.2 and U be N unlabeled samples. The SVM classifier is chosen in the SAR target discrimination for its outstanding performance. Thus, the sketch of the modified self-learning algorithm can be depicted as algorithm 2:

Algorithm 2. Modified Self-learning

Input: labeled training set L, unlabelled training set U;

Output: optimal classifier F.

Initialize classifier F based on L.

Construct a pool U' of unlabeled samples by randomly selecting u samples from U.

if length(U)>0 do

Exploit classifier F to give labels of unlabeled samples on the portion u in U'.

Estimate the confidence of each samples using Eq. (5).

Select p samples with high labeling confidence as positive samples and n samples with low-labeling confidence as negative samples to expand the labeled training set.

Update F using the augmented labeled training set L.

Randomly select $2p + 2n$ samples from U to supplement U'.

end if

3. EXPERIMENTS

3.1 Dataset description and experiment set

In order to verify the effectiveness of the proposed method, we perform experiments on the miniSAR and FARAD SAR dataset, which is released by Sandia National Laboratories of America in 2005 and 2015, respectively. These two SAR datasets arewidely used to demonstrate the SAR target detection and discrimination performance. The miniSAR dataset totally contains twenty SAR images whose resolution is 0.1mx0.1m and size is 1638x2501. Seven images which include targets are employed in our experiments. The FARAD SAR dataset totally contains 175 SAR images, with with 4-inch resolution, and we choose 7 images in our experiments. There exist different types of objects in these SAR images, not only including the interested targets such as cars and helicopters, but also the clutters such as roads, trees, building and so on. In Figure 2, we respectively give an example image in the miniSAR and FARAD SAR dataset, where the interested targets are circled with red rectangle.

The following two experiments are performed using a PC with Intel Core i7-6700 CPU of 3.3 GHz and memory of 16 GB. The program codes are written in MATLAB R2016a.

Firstly, the log-normal-based CFAR and clustering method[10] are utilized to detect the seven SAR image and obtain the candidate regions which may contain the targets, i.e., chips. Tables 1 and 2 give the detection results of miniSAR and FARAD SAR datasets, respectively, from which we can see that the number of clutters is double that of targets and it is necessary to remove the false alarm from candidate regions to alleviate the classification cost. For discrimination experiments, we regard the detection results of three SAR images as the training set and the rest SAR images as the test set.

(a) (b)

Figure 2. Example image from the miniSAR and FARAD SAR dataset. (a) miniSAR. (b) FARAD SAR.

Table 1. Number of chips of the detection results of miniSAR dataset.

	Number of clutters	Number of targets	Number of total chips
Training set	313	156	469
Test set	273	153	426

Table 2. Number of chips of the detection results of FARAD SAR dataset.

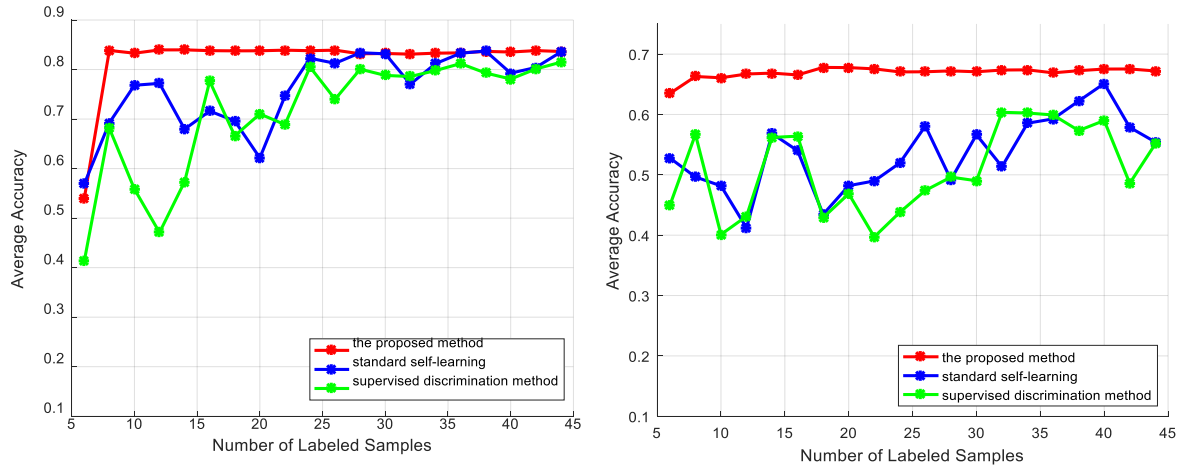
	Number of clutters	Number of targets	Number of total chips
Training set	242	157	399
Test set	154	66	220

3.2 Experimental results

In this part, the performance of the proposed method is compared with those of the standard self-learning and supervised discrimination method. We carry out a series of experiments under different configurations of initial labeled samples. For compared methods, the initial labeled samples are randomly chosen from the unlabeled training set and labeled by man, which is different from the proposed method where the initial labeled samples are automatically chosen by the algorithm and labeled by man. Figure 3 shows how the average accuracy of three discrimination methods varies under the same configurations of initial labeled samples. To guarantee the reliability of discrimination results, the experiments are repeated for 10 independent runs and the accuracy is obtained by averaging over these repetitions.

From Figure 3, it can be seen that as the number of initial labeled samples increases, the performance of three discrimination methods firstly becomes better and then remains stable. This results from that the discrimination performance mainly relies on the valuable information provided by the samples, rather than the number of labeled samples. Also, our proposed method performs better and is more stable than the standard self-learning and the supervised method with the same number of labeled samples. This is mainly because: (1) the proposed method actively selects the most informative samples to be labeled by man, making sure the initial classifier is with high performance; (2) the proposed method selects the most reliable samples to update the classifier during the iterative learning.

From Figure 3, it can be seen that as the number of initial labeled samples increases, the performance of three discrimination methods firstly becomes better and then remains stable. This results from that the discrimination performance mainly relies on the valuable information provided by the samples, rather than the number of labeled samples. Also, our proposed method performs better and is more stable than the standard self-learning and the supervised method with the same number of labelled samples. This is mainly because: (1) the proposed method actively selects the most informative samples to be labeled by man, making sure the initial classifier is with high performance; (2) the proposed method selects the most reliable samples to update the classifier during the iterative learning.



(a) miniSAR (b) FARAD SAR dataset

Figure 3. Average accuracy of different discrimination methods changes as the number of initial labeled samples increases for miniSAR and FARAD SAR dataset.

4. CONCLUSION

In this paper, we propose a novel semi-supervised method based on active sample proposal and modified self-learning for

SAR target discrimination. Firstly, only given the unlabeled samples, the Lincoln features are extracted. Then, the most informative samples are selected for man to be labeled through the sampling theory. Finally, we modify the self-learning algorithm by incorporating the classifier's posteriori probability and local geometrical information to estimate the label's confidence on unlabeled samples. We perform experiments on the real SAR datasets and the results illustrate that under the same configurations of initial labeled samples the proposed semi-supervised method performs better than the standard self-learning and supervised discrimination method.

REFERENCES

[1] Xiang, D., Tang, T., Zhao, L. and Su, Y., IEEE Geosci. Remote Sens. Lett., 10, 1414-1418(2013).
[2] Oliver, C. J. and Quegan, S., [Understanding Synthetic Aperture Radar Images], Artech House, Boston, MA, (1998).
[3] Du, L., Wang, Y. and Xie, W., IEEE Trans. Geosci Remote Sens., 58, 666-679(2020).
[4] Wang, Z., Du, L., Zhang, P., Wang, F., Xu, S. and Su, H., IEEE Trans. Geosci. Remote Sens., 14, 1855-1872(2017).
[5] Zhu, X., [Semi-supervised Learning Literature Survey], Univ. Wisconsin, Madison, USA, Technical Report, (2006).
[6] Blei, D. M. and Jordan, M. I., Proc. 26th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, Toronto, ON, Canada, 127-134(2003).
[7] Wang, D. Y. and Mori, G., IEEE Trans. Pattern Anal. Mach. Intell., 31, 1762-1774(2009).
[8] Gadde, A., Anis, A. and Ortega, A., Proc. 20th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD'14), 492-501(2014).
[9] Anis, A., Gadde, A. and Ortega, A., IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), (2014).
[10] Gao, G., Liu, L., Zhao L., Shi, G. and Kuang, G., IEEE Trans. Geosci. Remote Sens., 47, 1685-1697(2009).