



Communication

SAR Target Detection Based on Domain Adaptive Faster R-CNN with Small Training Data Size

Yuchen Guo ¹ , Lan Du ^{2,*} and Guoxin Lyu ²

¹ Academy of Advanced Interdisciplinary Research, Xidian University, Xi'an 710071, China; ychguo@xidian.edu.cn

² National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China; lvgx_xd@163.com

* Correspondence: dulan@mail.xidian.edu.cn

Abstract: It is expensive and time-consuming to obtain a large number of labeled synthetic aperture radar (SAR) images. In the task of small training data size, the results of target detection on SAR images using deep network approaches are usually not ideal. In this study, considering that optical remote sensing images are much easier to be labeled than SAR images, we assume to have a large number of labeled optical remote sensing images and a small number of labeled SAR images with the similar scenes, propose to transfer knowledge from optical remote sensing images to SAR images, and develop a domain adaptive Faster R-CNN for SAR target detection with small training data size. In the proposed method, in order to make full use of the label information and realize more accurate domain adaptation knowledge transfer, an instance level domain adaptation constraint is used rather than feature level domain adaptation constraint. Specifically, generative adversarial network (GAN) constraint is applied as the domain adaptation constraint in the adaptation module after the proposals of Faster R-CNN to achieve instance level domain adaptation and learn the transferable features. The experimental results on the measured SAR image dataset show that the proposed method has higher detection accuracy in the task of SAR target detection with small training data size than the traditional Faster R-CNN.

Keywords: synthetic aperture radar; target detection; domain adaptation; generative adversarial network



Citation: Guo, Y.; Du, L.; Lyu, G. SAR Target Detection Based on Domain Adaptive Faster R-CNN with Small Training Data Size. *Remote Sens.* **2021**, *13*, 4202. <https://doi.org/10.3390/rs13214202>

Academic Editor: Salvatore Stramondo

Received: 23 August 2021
Accepted: 12 October 2021
Published: 20 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Synthetic aperture radar (SAR) is an active Earth observation system, which has the characteristics of all-weather, all-time, high resolution and strong penetration. Therefore, SAR can play a role that other remote sensing means cannot play. Automatic target recognition (ATR) of SAR image has become a key technology for processing massive SAR image data. A typical SAR image ATR system is usually divided into three stages: detection, discrimination and recognition. SAR image target detection technology is important for the SAR ATR because the performance of the detection stage will directly affect the accuracy of the subsequent processing.

The existing SAR image target detection methods can be divided into two types: non-learning target detection algorithm and learning based target detection algorithm. Constant false alarm rate (CFAR) [1] is a kind of traditional non-learning target detection algorithm which is widely used in SAR system for SAR target detection. With the rapid development of deep learning, learning based deep convolution neural network (CNN) has been widely and successfully applied in target detection [2–6] and achieves better target detection performance than non-learning algorithms, such as faster regions with convolutional neural network (Faster R-CNN) [3], which combs a regional proposal network based on CNN with a Fast R-CNN detector. Although the SAR target detection methods based on Faster R-CNN can achieve satisfying performance, a large number of labeled training

samples are required in most of these methods for model learning, which means that it will lead to a significant performance drop for small training dataset sizes.

Some studies expand the training samples by using artificially generated samples to deal with the small training data size problem. For example, several data generation methods based on generative adversarial network (GAN) models are proposed in [7–10] to generate realistic infrared images from optical images. Using data from the source domain to assist the interested target domain's task is also present in transfer learning. In particular, when the data distribution is assumed to change across the source and target domains and the learning task remains the same, we call this transfer learning as domain adaptation [11]. The domain adaptation [12,13] has been studied to improve the performance of target domain with the help of source domain that contains abundant samples. The purpose of domain adaptation is to map the data in source domain and target domain with different distribution to the same feature space, and find a measurement criterion to make them as close as possible in this space. Then the data in the target domain can be classified by directly using the classifier trained on the source domain [14,15]. Early domain adaptation methods minimize the maximum mean discrepancy (MMD) [12,13,16] to reduce domain discrepancy. Recent works [14,15,17,18] implement domain confusion with gradient reversal layer (GRL) or generative adversarial nets (GANs) loss for feature alignment through adversarial learning. In SAR target recognition, Huang et al. explored what, where, and how to transfer knowledge from other domains to the SAR domain through MMD constraint [19]. In target detection research, Inoue et al. achieved a two-step weakly supervised domain adaptation framework using conventional pixel-level domain adaptation methods and pseudo labeling [20]. Chen et al. proposed a domain adaptive Faster R-CNN model to learn domain-invariant features by aligning feature distribution at the image-level and instance-level through adversarial learning [21]. He et al. and Chen et al. proposed an importance-weighted GRL to re-weight source or interpolation samples [22,23].

In this study, a SAR target detection method based on domain adaptive Faster R-CNN is proposed, which can solve the problem of low detection accuracy due to small training data size of SAR images by transferring knowledge from optical remote sensing images of source domain to SAR images of target domain through domain adaptive learning. A large number of labeled optical remote sensing images and a small number of labeled SAR images can capture the similar scenes, such as the parking lots, which contain the interested target with similar high-level semantic features in both source and target domains. In the proposed method, domain adaptation constraint is added after the proposals of Faster R-CNN, which realizes instance level constraint and makes full use of label information to achieve accurate domain adaptation learning. More specifically, a GAN constraint is utilized as the domain adaptation constraint in the adaptation module to achieve domain adaptation and learn the transferable features. With the help of label information, the proposed model can accurately narrow the differences between the high-level features of interested target of the target domain and those of the source domain via domain adaptive learning, so as to assist the learning of target detection model in the target domain. The proposed model can alleviate the overfitting problem caused by the small training data size of SAR images, and effectively improve the SAR image target detection performance.

The remainder of this letter is arranged as follows. Section 2 represents the proposed domain adaptive Faster R-CNN framework. Section 3 shows the experimental results and analysis based on the measured SAR data and optical remote sensing data. Section 4 discusses the domain adaptation performance in our method. Finally, Section 5 concludes this letter.

2. Domain Adaptive Faster R-CNN

The SAR target detection methods are faced with the challenge of performance degradation when lacking labeled data for model training. However, the optical remote sensing images are much easier to be labeled, thus abundant labels can be obtained to learn the

features that perform well for detection. In this paper, we aim to transfer the abundant information of optical remote sensing images with a large number of labeled samples to SAR images with a small number of labeled samples to help SAR target detection. As mentioned in the introduction, although there are great differences between optical remote sensing images and SAR images, the interested target captured in them have similar high-level semantic features. Thus, it is possible to learn the transferable high-level features between the optical and SAR images via adaptation methods.

Let $D_s = \{(\mathbf{x}_i^s, \mathbf{y}_i^s)\}_{i=1}^{N_s}$ represent the source domain dataset with N_s samples, \mathbf{y}_i^s is the label of the i -th sample \mathbf{x}_i^s ; let $D_t = \{(\mathbf{x}_i^t, \mathbf{y}_i^t)\}_{i=1}^{N_t}$ represents the target domain dataset with N_t samples.

The proposed model consists of three parts: the feature extraction module, domain adaptation module and detection module. (1) The feature extraction module contains two base networks: truncated VGG16 network with the region proposal network (RPN) and the region-of-interest (ROI) pooling for the feature extraction of target domain; truncated VGG16 network with RPN and ROI pooling fixed with the pre-trained network by the source domain data for the feature extraction of source domain. (2) The domain adaptation module constrains the features of the target domain to be close to those of the source domain. (3) The detection module contains fully connected network for location and class prediction of the target. In the proposed method, the base networks in the feature extraction module and the detection module have the same architectures as those in traditional Faster R-CNN. For intuitive illustration, the architecture of the proposed domain adaptive Faster R-CNN is shown in Figure 1.

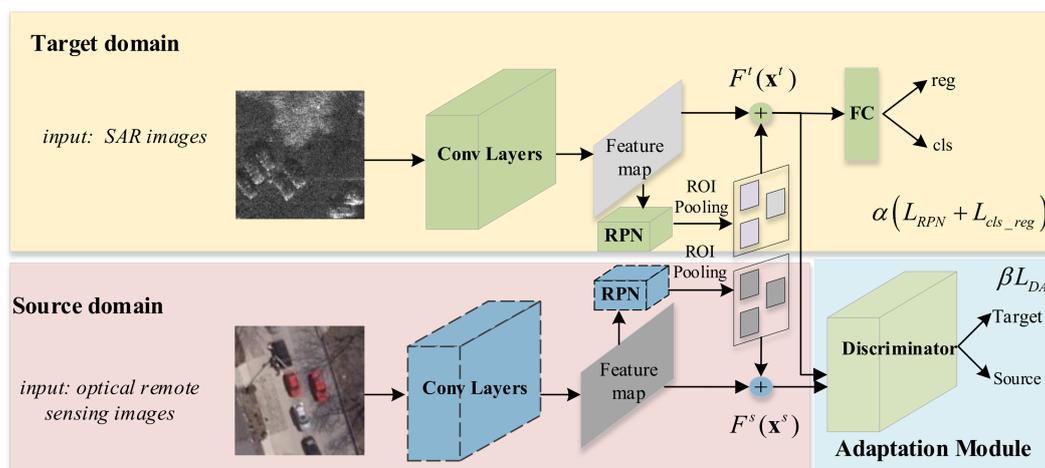


Figure 1. The architecture of the domain adaptive Faster R-CNN for source and target domain data, i.e., the optical remote sensing images and SAR images. In the domain adaptation module, the features of two domains are constrained with the GAN criterion.

The source domain has abundant samples to learn the features that depict the source domain data well, while the target domain cannot achieve promising performance with small training data size. We first train the source domain detection model with the source domain data, and then try to transfer the knowledge from the source domain to the target domain. First, we initialize the target domain detection model with the source domain detection model. For the source and target domain data, i.e., the optical images and SAR images, capturing the similar scene, we use the distributions over the high-level representations as the bridge for cross-domain data. Moreover, we assume that the marginal distributions over the high-level features across two domains should be similar. Specifically, to reduce the domain shift, we use the GAN constraint [24] in the domain adaptation module:

$$L_{DA} = E_{\mathbf{x}^s}[\log D(F^s(\mathbf{x}^s))] + E_{\mathbf{x}^t}[\log(1 - D(F^t(\mathbf{x}^t)))] \quad (1)$$

F^t presents the feature extraction module of target domain, $E_{x^t}[\log(1 - D(F^t(x^t)))]$ indicates the expectation that the discriminator D judges $F^t(x^t)$ as the feature belong to target domain, F^s presents the feature extraction module of source domain, $E_{x^s}[\log D(F^s(x^s))]$ indicates the expectation that the discriminator D judges $F^s(x^s)$ as the feature belong to source domain.

In our model, F^t and D are iteratively updated. Specifically D wants to maximize the L_{DA} loss and F^t wants to minimize the L_{DA} loss. The feature extraction module of source domain F^t and the discriminator D compete in a two-player minimax game. The discriminator tries to distinguish features of source domain from features of target domain; and the F^t tries to fool the discriminator. As a result, our model can force $F^t(x^t)$ to be indistinguishable from $F^s(x^s)$ and reduce the feature distance.

Then, we can give the objective function of the target domain Faster R-CNN:

$$\min_{F^t, FC} \max_D L = \alpha (L_{RPN} + L_{cls_reg}) + \beta L_{DA}, \quad (2)$$

$$L_{RPN} = L_{RPN_cls} + L_{RPN_reg}, \quad (3)$$

$$L_{cls_reg} = L_{cls} + L_{reg}, \quad (4)$$

where FC presents the detection module of target domain, L_{RPN_reg} and L_{RPN_cls} are respectively the anchor-location and anchor-classification loss terms of the RPN for the target proposals; L_{reg} and L_{cls} are respectively the box-regression and box-classification probability loss terms of the ROI for the target, α and β represent the weight coefficients of $L_{RPN} + L_{cls_reg}$ and L_{DA} , respectively.

3. Results

To validate the effectiveness of the proposed method, some experiments are conducted in this section. The description of the datasets is firstly presented. Then the experimental results and analysis on the measured data are shown via a domain adaptation scenario.

3.1. Description of the Datasets

One SAR image dataset and one optical remote sensing image dataset are adopted to conduct some experiments to verify the effectiveness of our method. In the following, the SAR image dataset is the miniSAR dataset [25], and the optical image dataset is the Toronto dataset [26].

The miniSAR dataset is acquired by U.S. Sandia National Laboratories with the resolution of 0.1 m \times 0.1 m. In detail, miniSAR dataset contains nine images with size 1638 \times 2510 pixels, seven of which are used for training and the other two for testing. Two SAR images of the miniSAR dataset are shown in Figure 2. In addition, the Toronto dataset covers the city of Toronto with a color depth of 24 bits per pixel (RGB) and 0.15 m spatial resolution. The original image has the size of 11,500 \times 7500 pixels, thus the original images are segmented into several subareas. Twelve subarea images are selected as the training set and 11 subarea images are selected as test images, respectively. In Figure 3, we present two images of the Toronto dataset. As we can see from Figures 2 and 3, there is some artificial clutter, including roads and buildings, and some natural clutter such as grasslands and trees.

It is not suitable to directly use the original images as the input of network, because the size of the original images is very large. In the following experiments, the original training images of the two datasets are cropped into multiple 300 \times 300 sub-images, and then these sub-images are utilized for network training. Moreover, the original test images are also cropped into 300 \times 300 sub-images by sliding window repeatedly. Here the size of the repetition sliding window is set as 100 pixels. Then the detection results of these test sub-images can be restored to the original test SAR image. Finally, the final detection results can be obtained by employing the non-maximum suppression (NMS) deduplication algorithm.



Figure 2. Two images of miniSAR dataset: (a) image A. (b) image B.



Figure 3. Two images of Toronto dataset: (a) image A. (b) image B.

3.2. Evaluation Criteria

The detection performance of different methods can be quantitatively evaluated via precision, recall, and F1-score [27], and the calculation formulas of them are as:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (5)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (6)$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (7)$$

where the number of correctly detected targets is represented as TP, the number of false alarms is represented as FP, and the number of miss alarms is represented as FN. The precision and recall measure the fraction of true positives over all detected results and the fraction of true positives over the ground truths, respectively. By calculating the harmonic mean between precision and recall, the main reference F1-score can be obtained to evaluate the detection performance comprehensively.

3.3. Performance Comparison

In this subsection, the proposed domain adaptive Faster R-CNN method is compared with conventional SAR target detection methods, such as the Gaussian-CFAR [1], Faster R-

CNN [3], and the Faster R-CNN initialized with the source domain detection model. Except for the CFAR method, which does not require training samples, the number of training samples for each method in the target domain is 110 firstly. The intuitional target detection results on miniSAR dataset of the proposed method and three comparison methods are shown in Figure 4. To make the experimental results clear and intuitive, the correctly detected target chips are represented as green rectangles, false alarms are represented as red rectangles and missing alarms are represented as blue rectangles. Since the performance of the CFAR methods depend on the detection threshold, all the following results of SAR target detection methods are the best by setting different thresholds.

As shown in Figure 4a,b, the Gaussian CFAR cannot locate targets accurately and has a lot of false alarms. Since the ground scene is quit complex in the miniSAR data, which contains vehicles as the targets to be detected and a lot of buildings, trees, roads and grasslands as the clutter, the CFAR methods are hard to gain a satisfying result by selecting a suitable clutter statistical model. As shown in Figure 4a,b, the Gaussian CFAR has a lot of false alarms and cannot locate targets accurately. Figure 4c,d show the detection results of Faster R-CNN. Compared with the unsupervised method CFAR, Faster R-CNN has a better detection performance with labeled information and deep convolutional network. Figure 4e,f show the detection results of Faster R-CNN initialized with the source domain detection model, which proves that simple transfer method has little help for the SAR detection task. Figure 4g,h show the detection results of our method. Both the missing alarms and false alarms in the results are reduced compared with the abovementioned methods because our method make use of the knowledge transferred from optical images.

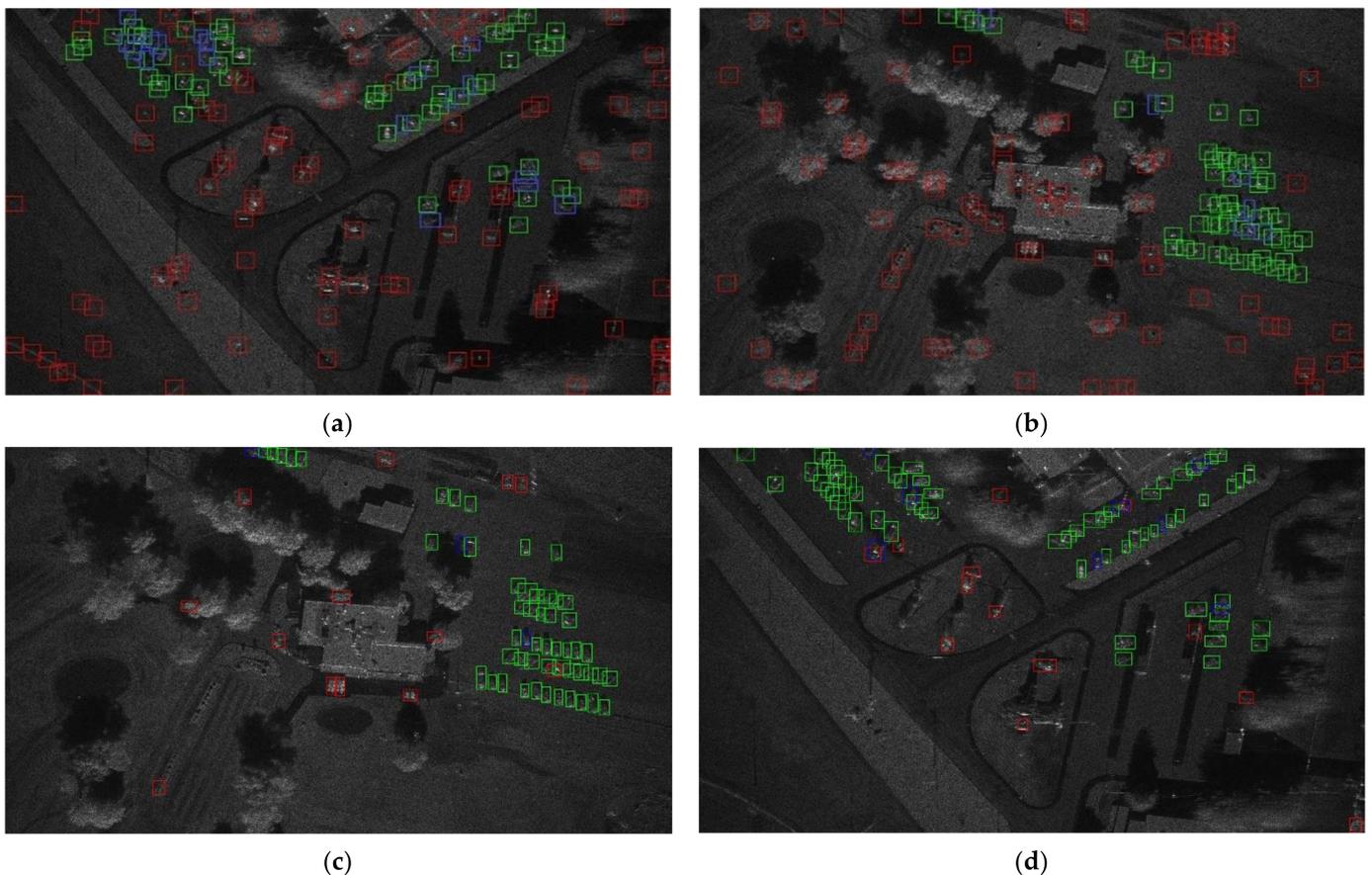


Figure 4. Cont.

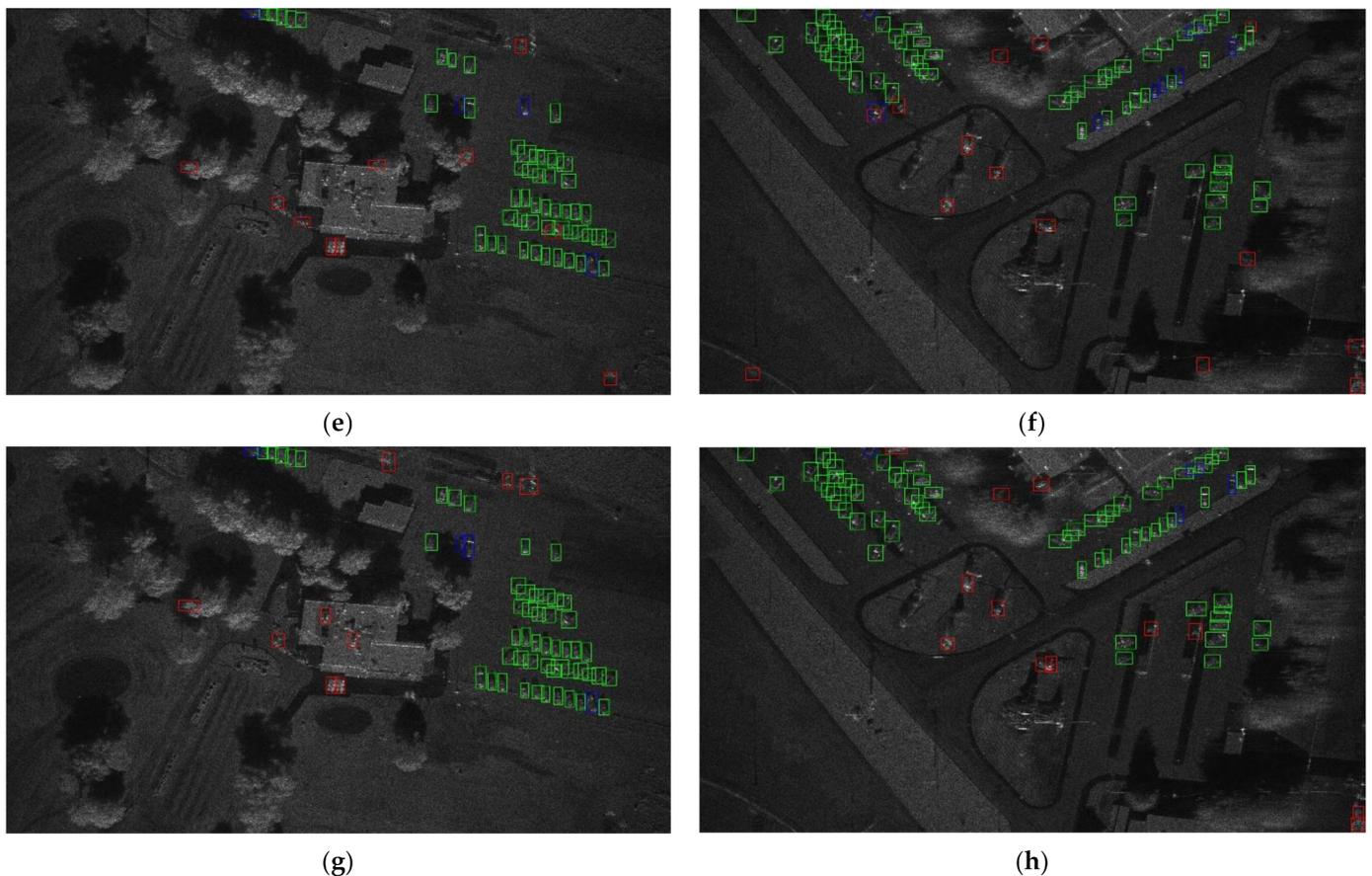


Figure 4. The SAR target detection results of the proposed method comparing with the comparisons for the two test images of miniSAR dataset. (a,b) are the results of Gaussian-CFAR; (c,d) are the results of Faster R-CNN; (e,f) are the results of Faster R-CNN with initialized operation; (g,h) are the results of our method.

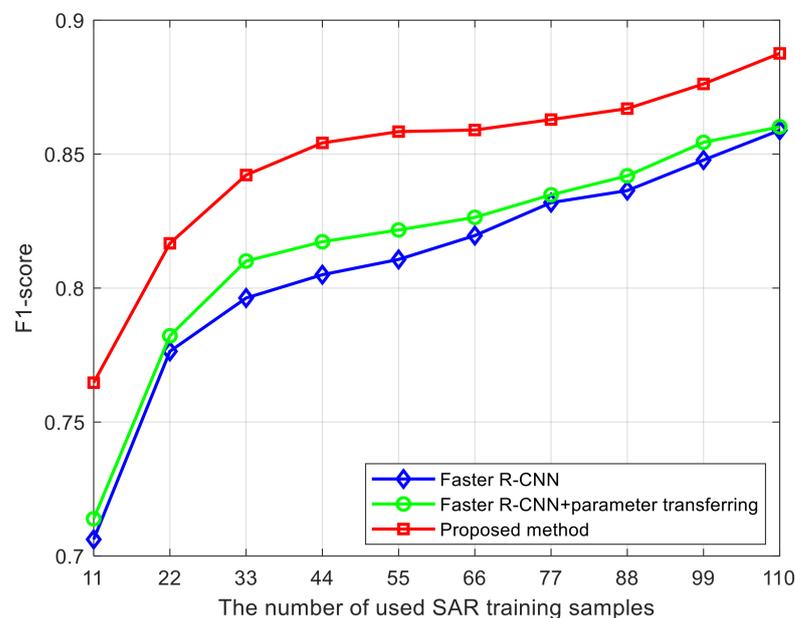
For quantitative evaluation and analysis, the missing alarms and the false alarms of all the methods are counted from the correctly detected target chips to verify the performance. Further, to quantitatively analyze the overall SAR target detection results, we calculate precision, recall and F1-score shown in Table 1 as the evaluation criteria. Compared with other SAR target detection methods, the proposed domain adaptive Faster R-CNN can correctly detect more targets. Meanwhile, the missing alarms and the false alarms of the proposed method are both the fewest. It can be clearly seen from the quantitative evaluation criteria that our domain adaptive Faster R-CNN achieves the highest scores on precision, recall and F1-score. Specifically, the proposed method is at least 2.81%, 2.44% and 2.61% higher in terms of precision, recall and F1-score than the other compared SAR target detection methods. It can be concluded that the performance of the proposed domain adaptive Faster R-CNN method significantly superior to that of the other three SAR target detection methods.

Further, the experimental results of each method are also given in Table 1 when the number of training samples in the target domain are reduced to 11. The proposed method yields a nearly 6% improvement in F1-score compared to Faster-RCNN. It can be found that our method has a greater performance improvement with a smaller number of target domain training samples.

Table 1. Overall evaluation of different conventional target detection methods.

Method	SAR Training Samples	Missing Alarms	False Alarms	Precision	Recall	F1-Score
Gaussian CFAR	×	24	177	0.3587	0.8048	0.4962
Faster R-CNN	110	11	26	0.8116	0.9106	0.8582
Faster R-CNN + parameter transferring		11	25	0.8175	0.9106	0.8615
Proposed method		8	21	0.8456	0.9350	0.8876
Faster R-CNN	11	21	30	0.7561	0.6667	0.7086
Faster R-CNN + parameter transferring		20	30	0.7561	0.6732	0.7122
Proposed method		15	24	0.8049	0.7347	0.7682

Then, the experimental results of Faster-RCNN, Faster-RCNN + parameter transferring and the proposed method with different numbers of used SAR training samples in the target domain are presented in Figure 5. As shown in Figure 5, the more the SAR training samples are used, the better performance the original Faster R-CNN and the other two methods obtain. Moreover, compared to the results of the original Faster R-CNN, the parameter transferring and the domain adaptation module both can contribute to better detection performance in the proposed method. In particular, when the used SAR training samples are fewer, the superiority of the domain adaptation module is more obvious. Thus the detection performance of the proposed method is much better than the original Faster R-CNN.

**Figure 5.** The experimental results with different numbers of used SAR training samples in the target domain.

4. Discussion

In order to yield an intuitive visualization of the domain adaptation results, t-distribution stochastic neighbor embedding (t-SNE) is employed to map features in multidimensional space to a two-dimensional space. Then the two-dimensional t-SNE visualizations from the features of Toronto and miniSAR samples can be shown. Since the miniSAR and Toronto datasets respectively contain 110 and 2600 training samples, we randomly choose 110 examples from the Toronto dataset and all 110 samples from the miniSAR dataset to

show the two-dimensional t-SNE results of the features for the 220 samples. The features of the 220 samples from two datasets are respectively obtained via the source and target base networks and then they are jointly mapped to a two-dimensional space with t-SNE, and the projection results are presented in Figure 6. As we can see from Figure 6a, before adaptation, the feature space between Toronto and miniSAR samples are clearly separable, which shows there is great domain discrepancy between the two datasets. In Figure 6b, after feature-level adaptation via our method, the features distribution of miniSAR samples is similar to that of Toronto samples, which illustrates our method can effectively align the features distribution and reduce the domain shift.

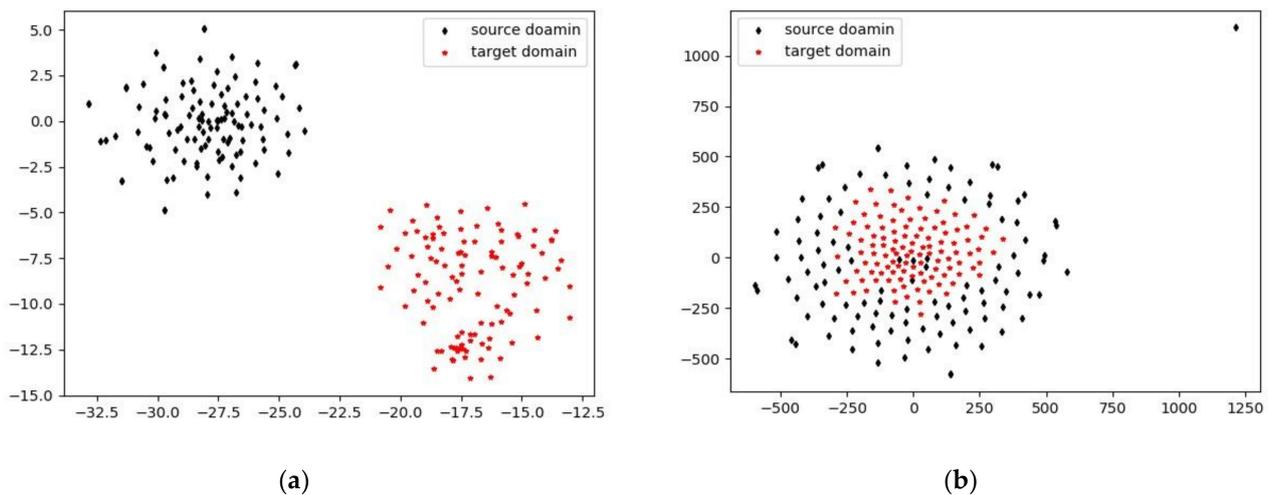


Figure 6. Two-dimensional t-SNE visualizations from the features of the Toronto and miniSAR samples. Black: 110 samples from Toronto dataset; Red: 110 samples from miniSAR dataset. (a) Feature visualization of Toronto and miniSAR samples without adaptation; (b) Feature visualization of Toronto and miniSAR samples via our proposed method.

5. Conclusions

This paper presents a novel SAR target detection method based on domain adaptive Faster R-CNN, which takes advantage of the domain adaptation to solve the problem of small training data size. The proposed approach is built on the Faster R-CNN model. The instance level domain adaptation is incorporated to learn the transferable features that bridge the cross-domain discrepancy between the optical remote sensing images and SAR images. In this way, the optical remote sensing images are utilized to help learn the features of SAR images that are useful for detection. Detailed experimental comparisons are given to confirm the effectiveness of the proposed method.

Author Contributions: Conceptualization, Y.G. and L.D.; methodology, G.L. and Y.G.; software, G.L. and Y.G.; validation, G.L.; investigation, G.L.; resources, L.D.; writing—original draft preparation, Y.G.; writing—review and editing, L.D.; visualization, G.L.; supervision, L.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by the National Science Foundation of China, Grannumber: 61771362, and in part by the 111 Project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gao, G.; Liu, L.; Zhao, L.; Shi, G.; Kuang, G. An adaptive and fast CFAR algorithm based on automatic censoring for target detection in high-resolution SAR images. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1685–1697. [[CrossRef](#)]
2. Girshick, R. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
3. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards realtime object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 91–99. [[CrossRef](#)] [[PubMed](#)]

4. Du, L.; Liu, B.; Wang, Y.; Liu, H.; Dai, H. Target detection method based on convolutional neural network for SAR image. *J. Electron. Inf. Technol.* **2016**, *38*, 3018–3025.
5. Li, J.; Qu, C.; Shao, J. Ship detection in SAR images based on an improved faster R-CNN. In *SAR in Big Data Era: Models, Methods and Applications (BIGSAR DATA)*; IEEE: Piscataway, NJ, USA, 2017.
6. Kang, M.; Ji, K.; Leng, X.; Lin, Z. Contextual region-based convolutional neural network with multilayer fusion for SAR ship detection. *Remote Sens.* **2017**, *9*, 860. [[CrossRef](#)]
7. Yuan, X.; Tian, J.; Reinartz, P. Generating artificial near infrared spectral band from rgb image using conditional generative adversarial network. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *3*, 279–285. [[CrossRef](#)]
8. Uddin, M.S.; Li, J. Generative Adversarial Networks for Visible to Infrared Video Conversion. In *Recent Advances in Image Restoration with Applications to Real World Problems*; IntechOpen: London, UK, 2020.
9. Yun, K.; Yu, K.; Osborne, J.; Eldin, S.; Nguyen, L.; Huyen, A.; Lu, T. Improved visible to IR image transformation using synthetic data augmentation with cycle-consistent adversarial networks. In *Pattern Recognition and Tracking XXX*; International Society for Optics and Photonics: Bellingham, WA, USA, 2019; Volume 10995, p. 1099502.
10. Uddin, M.S.; Hoque, R.; Islam, K.A.; Kwan, C.; Gribben, D.; Li, J. Converting Optical Videos to Infrared Videos Using Attention GAN and Its Impact on Target Detection and Classification Performance. *Remote Sens.* **2021**, *13*, 3257. [[CrossRef](#)]
11. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A comprehensive survey on transfer learning. *Proc. IEEE* **2020**, *109*, 43–76. [[CrossRef](#)]
12. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; pp. 3320–3328.
13. Long, M.; Cao, Y.; Wang, J.; Jordan, M. Learning transferable features with deep adaptation networks. In Proceedings of the International Conference on Machine Learning, Lille, France, 1–9 July 2015; pp. 97–105.
14. Saito, K.; Watanabe, K.; Ushiku, Y.; Harada, T. Maximum classifier discrepancy for unsupervised domain adaptation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 3723–3732.
15. Tzeng, E.; Hoffman, J.; Saenko, K.; Darrell, T. Adversarial discriminative domain adaptation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7167–7176.
16. Sun, B.; Saenko, K. Deep coral: Correlation alignment for deep domain adaptation. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Cham, Switzerland, 2016; pp. 443–450.
17. Gain, Y.; Lempitsky, V. Unsupervised domain adaptation by backpropagation. In Proceedings of the International Conference on Machine Learning, Lille, France, 1–9 July 2015; pp. 1180–1189.
18. Ganin, Y.; Ustinova, E.; Ajakan, H.; Germain, P.; Larochelle, H.; Laviolette, F.; Marchand, M.; Lempitsky, V. Domain-adversarial training of neural networks. *J. Mach. Learn. Res.* **2016**, *17*, 1–35.
19. Huang, Z.; Pan, Z.; Lei, B. What, where, and how to transfer in SAR target recognition based on deep CNNs. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 2324–2336. [[CrossRef](#)]
20. Inoue, N.; Furuta, R.; Yamasaki, T.; Aizawa, K. Cross-domain weakly-supervised object detection through progressive domain adaptation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 5001–5009.
21. Chen, Y.; Li, W.; Sakaridis, C.; Dai, D.; Van Gool, L. Domain adaptive faster r-cnn for object detection in the wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 3339–3348.
22. He, Z.; Zhang, L. Multi-adversarial faster-r-cnn for unrestricted object detection. In Proceedings of the IEEE International Conference on Computer Vision, Seoul, Korea, 27 October–3 November 2019; pp. 6668–6677.
23. Chen, C.; Zheng, Z.; Ding, X.; Huang, Y.; Dou, Q. Harmonizing Transferability and Discriminability for Adapting Object Detectors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 8869–8878.
24. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 417–432.
25. Gutierrez, D. MiniSAR: A Review of 4-Inch and 1-Foot Resolution Ku-Band Imagery. 2005. Available online: <https://www.sandia.gov/radar/Web/images/SAND2005-3706P-miniSAR-flight-SAR-images.pdf> (accessed on 9 August 2021).
26. Chen, Z.; Wang, C.; Luo, H.; Wang, H.; Chen, Y.; Wen, C.; Yu, Y.; Cao, L.; Li, J. Vehicle detection in high-resolution aerial images based on fast sparse representation classification and multiorder feature. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 2296–2309. [[CrossRef](#)]
27. Du, L.; Li, L.; Wei, D.; Mao, J. Saliency-guided single shot multibox detector for target detection in SAR images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 3366–3376. [[CrossRef](#)]