Article

An Integrated Counterfactual Sample Generation and Filtering Approach for SAR Automatic Target Recognition with a Small Sample Set

Changjie Cao ¹, Zongyong Cui ¹, Zongjie Cao ^{1,2,*}, Liying Wang ¹ and Jianyu Yang ¹

¹ School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China; 201811011923@std.uestc.edu.cn (C.C.); zycui@uestc.edu.cn (Z.C.); wly@std.uestc.edu.cn (L.W.); jyyang@uestc.edu.cn (J.Y.)

² The Center for Information Geoscience, University of Electronic Science and Technology of China, Chengdu 611731, China

* Correspondence: zjcao@uestc.edu.cn

Abstract: Although automatic target recognition (ATR) models based on data-driven algorithms have achieved excellent performance in recent years, the synthetic aperture radar (SAR) ATR model often suffered from performance degradation when it encountered a small sample set. In this paper, an integrated counterfactual sample generation and filtering approach is proposed to alleviate the negative influence of a small sample set. The proposed method consists of a generation component and a filtering component. First, the proposed generation component utilizes the overfitting characteristics of generative adversarial networks (GANs), which ensures the generation of counterfactual target samples. Second, the proposed filtering component is built by learning different recognition functions. In the proposed filtering component, multiple SVMs trained by different SAR target sample sets provide pseudo-labels to the other SVMs to improve the recognition rate. Then, the proposed approach improves the performance of the recognition model dynamically while it continuously generates counterfactual target samples. At the same time, counterfactual target samples that are beneficial to the ATR model are also filtered. Moreover, ablation experiments demonstrate the effectiveness of the various components of the proposed method. Experimental results based on the Moving and Stationary Target Acquisition and Recognition (MSTAR) and OpenSARship dataset also show the advantages of the proposed approach. Even though the size of the constructed training set was 14.5% of the original training set, the recognition performance of the ATR model reached 91.27% with the proposed approach.

Keywords: synthetic aperture radar (SAR); automatic target recognition (ATR); generative adversarial nets (GANs); counterfactual sample; small sample set

Citation: Cao, C.; Cui, Z.; Cao, Z.; Wang, L.; Yang, J. An Integrated Counterfactual Sample Generation and Filtering Approach for SAR Automatic Target Recognition with a Small Sample Set. *Remote Sens.* **2021**, *13*, 3864. <https://doi.org/10.3390/rs13193864>

Academic Editor: Dusan Gleich

Received: 30 July 2021

Accepted: 22 September 2021

Published: 27 September 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 important method for Earth observation sensors with a wide range of applications [1]. Among the many applications of SAR, automatic target recognition (ATR) technology is a critical means of SAR image interpretation [2–4]. ATR technology has been greatly improved during the past decade because of the vigorous development of machine learning algorithms.

Among machine learning algorithms, feature extraction and classifier algorithms have made the most significant contributions. Feature extraction algorithms, including principal component analysis [5–7], non-negative matrix factorization [8], and linear discriminant analysis [9,10], can extract numerous features with discriminative information. The support vector machine (SVM) [11,12] has successfully maximized the target classification margin by selecting support vectors, and it has also had great success in ATR technology. In addition, the popularity of deep learning algorithms has risen due to their

significant advantages when applied to recognition techniques. Based on the inspiration of AlexNet, Kechagias-Stamatis et al. proposed a high-performance SAR ATR model, which demonstrated excellent performance [13]. Pei et al. utilized deep neural networks for multi-view SAR ATR [14], and Zhang et al. extended deep neural networks to the field of character recognition [15].

Unfortunately, all of the aforementioned algorithms are data-driven methods, which often rely on the size of the dataset [16]. Generally speaking, the performance of state-of-the-art ATR algorithms is acceptable in scenarios where sufficient SAR target samples are available for training the recognition model [17]. However, the assumption of sufficient SAR target samples is not always satisfied when the above methods are adopted to solve SAR ATR tasks. Therefore, the above algorithms are difficult to achieve the desired performance under this premise. In general, the purpose of the ATR model is to minimize the expected risk of the recognition result to the target sample set. Since the posterior probability cannot be easily obtained, the ATR model usually obtains an empirical risk from the training sample set to replace it. However, the small sample dataset must be incomplete, and the empirical risk for a small sample dataset cannot be equated to the empirical risk for a complete dataset, which leads to misrecognition of certain target samples by the ATR model. In addition, this is a problem that the above data-driven algorithms cannot avoid. It is necessary to establish a method that can apply the ATR model even when the number of SAR target samples is insufficient.

There are usually two kinds of methods to solve the problem of insufficient SAR target samples [18]. First, it is feasible to develop an algorithm that can effectively extract SAR target sample features from extremely limited datasets. In this regard, the extensive application of transfer learning in the field of optics provides an alternative approach [19,20]. Huang et al. utilized the transfer learning method to perform ATR with limited labeled SAR target samples, but the method still uses other unlabeled SAR target sample information [21]. However, pre-training large-scale neural networks using homogeneous datasets is a challenge when implementing actual SAR ATR tasks. Second, an effective and direct approach would be to increase the size of the SAR target sample dataset by generating more data. This type of method is more suitable for situations involving extremely limited SAR target sample datasets without other information sources. A common method of increasing the size of the SAR target sample dataset is data augmentation, such as rotation, flipping, or random cropping. However, the algorithms involved always apply geometric transformations to expand the size of the dataset. Geometric transformations cannot improve the data distribution determined by high-level features, as they only lead to an image-level transformation through depth and scale [22]. Therefore, these methods still fail to improve the performance of the SAR ATR model to an acceptable level under the condition of a small sample set.

In recent years, a variety of methods for improving the performance of the ATR model using generated data has been proposed. Malmgren et al. used simulated SAR target samples through data augmentation to improve the performance of ATR models by transfer learning [23]. Zhong et al. also utilized data augmentation to perform person re-identification with a small sample set [24]. Guo et al. proposed an innovative simulation technology based on generative adversarial networks (GANs), which was proposed to address the lack of accuracy of computer-aided drawing (CAD) models [25]. The rapid development of GANs has led to an increasing number of applications in SAR ATR in the past few years. Gao et al. used the discriminator of deep convolutional generative adversarial networks (DCGAN) to implement the ATR task [26]. Schwegmann et al. proposed information maximizing generative adversarial networks (InfoGAN) to perform SAR ship recognition [27]. Although previous studies on GANs have addressed the performance of semi-supervised learning in the recognition task, research has yet to explore whether the generated SAR target samples can be moved from the generation architecture and used in other available ATR frameworks [28]. Cui et al. used a label-free SAR target sample generation method to increase the size of a small sample set. However, it is time

intensive to label a large number of unlabeled target samples [29]. Choe et al. sought to address this issue in the field of optics by utilizing the generated samples to solve face recognition problems with a small sample set [18]. Tang et al. also used generated samples to implement person re-identification [30]. All of the aforementioned literature has proven the boundless potential of GANs in relation to the SAR ATR task.

However, when GANs-based methods were used to generate SAR target samples, it is easy to create a counterfactual target sample due to the characteristics of SAR target data. This interesting phenomenon is shown in Figure 1. As shown in Figure 1, despite the fact that some generated sample looks no different from the real target sample, it may cause a significant decrease in the performance of the ATR model. We refer to such target samples as counterfactual samples in this paper. The effect of counterfactual target sample on the ATR model may be opposite to the effect perceived by the naked eye.

Figure 1. The schematic diagram of counterfactual target samples. Some generated sample looks no different from the real sample. However, it may causes a significant decrease in the performance of the ATR model. The effect of counterfactual target sample on the ATR model may be opposite to the effect perceived by the naked eye.

It means that the target samples generated by the GANs-based method cannot be directly used for the training of the ATR model. The generated target samples may damage the performance of the ATR model if the generated samples are not filtered. Although optical imagery has been successful in applying generated sample sets in other frameworks, the interpretability of SAR images is poorer than that of optical images, which increases the difficulty of using the SAR ATR model with small sample sets. Zhang et al. used a pre-filtering method to address the low interpretability problem with generated SAR target samples and demonstrated the necessity for pre-filtering these samples [31]. However, the premise of this pre-filtering method is still based on the assumption that many training samples are available, which is contrary to the assumption of the SAR ATR model with a small sample set. This shows that the current pre-filtering method has certain limitations for target samples generated by small sample sets. At the same time, it also means that some counterfactual samples that do not look the same as real samples were also ignored in the previous pre-filtering work. Moreover, there is little relevant research paying attention to the filtering of generated samples at present, especially for the filtering of counterfactual samples.

In this paper, we focus on solving the small sample set problem in the SAR ATR model via a counterfactual target sample. An integrated counterfactual sample generation and filtering approach is proposed to alleviate the impact of a small sample set. Counterfactual samples are target samples that can reasonably expand the data space and are difficult to find in the real world. The proposed method makes full use of the overfitting characteristics of GANs, which ensures the generation of counterfactual target samples. Then, the generated counterfactual samples that are beneficial to the ATR model are filtered by learning different ATR functions. As the proposed method continuously generates counterfactual target samples, it can also dynamically improve the recognition performance of the ATR model. More importantly, the proposed method adopts an innovative way of using

generated counterfactual target samples, which significantly improves the recognition performance of the SAR ATR model under the condition of a small sample set.

2. Related Work

2.1. The Generation of Counterfactual Target Samples

Counterfactual target samples can reasonably expand the data space, which means that these samples have a distribution similar to the real data. The purpose of GANs is to estimate the real data distribution and generate target samples with a similar distribution. This provides the basis for using GANs to generate counterfactual samples in this study. The generator G is trained to generate target samples by simulating noise z from the noise distribution $P_z(z)$ in the real SAR target sample distribution $P_r(x)$, where $x = G(z)$. The purpose of training the discriminator D is to determine whether the target sample generated by G is real or not. The G and D were originally represented by the multilayer perceptron. In summary, D and G appear to play the minimax game, with the application of the following function $V(D, G)$ [32]:

$$\min_G \max_D V(D, G) = E_{x \sim P_r(x)} [\log D(x)] + E_{x \sim P_g(x)} [\log(1 - D(G(x)))], \quad (1)$$

where $D(x)$ refers to the probability that D determines that the target sample x comes from the real sample distribution $P_r(x)$ rather than the generated sample distribution $P_g(x)$. In this minimax game, G attempts to generate target samples that are as real as possible to deceive D to the greatest extent. D needs to effectively function to distinguish between real target samples and generated samples. Then, realistic SAR target samples are gradually generated during the minimax adversarial process between two networks.

However, conventional GANs architecture is often unable to remain stable during the generation process. In this regard, deep convolutional generative adversarial networks (DCGANs) have been developed, which utilize certain training techniques to maintain the stability of the generation model [33]. Subsequently, Martin Arjovsky et al. demonstrated that conventional GANs utilize an unreasonable distance measurement to fit the real SAR target sample distribution $P_r(x)$ [34]. Therefore, they proposed the Wasserstein distance to replace the original distance measurement between $P_r(x)$ and $P_g(x_i)$:

$$W(P_r(x), P_g(x_i)) = \inf_{\gamma \in \Pi(P_r(x), P_g(x_i))} E_{(x, x_i) \sim \gamma} [|x - x_i|], \quad (2)$$

where $\Pi(P_r(x), P_g(x_i))$ is the set of all joint distributions $\gamma(x, x_i)$, and its marginal distributions are $P_r(x)$ and $P_g(x_i)$. The set of all of the joint distributions is the joint distribution between the real distribution $P_r(x)$ and all of the generated distributions $P_g(x_i)$ that appear in the above adversarial process.

However, when GANs use the Wasserstein distance to measure the distance between the generated distribution $P_g(x)$ and the real distribution $P_r(x)$, they are prone to gradient explosion during the training process. Therefore, Ishaan Gulrajani et al. enforced the 1-Lipschitz gradient penalty constraint on the Wasserstein distance [35]. The gradient penalty constraint effectively alleviates the drawback of GANs using the Wasserstein distance to describe the distance between $P_g(x)$ and $P_r(x)$.

Moreover, most ATR models are often based on supervised learning architecture. Therefore, it is not sufficient for the ATR model to stably generate SAR target samples. The generation model can generate SAR target samples by category, which is more suitable for the existing supervised learning architecture. Fortunately, the conditional GAN was proposed to control the category information in the training process [36]. The loss function of the conditional GAN is as follows:

$$\min_G \max_D V(D, G) = E_{x \sim P_r(x)} [\log D(x|y)] + E_{x \sim P_g(x)} [\log(1 - D(G(x|y)))], \quad (3)$$

where y refers to the category information. Then, the category information y can guide the categorization of target samples during the generation process.

Almost none of the architecture of the aforementioned GANs model are specific to the ATR model. So far, our recent work explored an SAR target sample generation method for the SAR ATR model, called LDGAN [37]. LDGAN explored an effective method of generating SAR target samples. At the same time, the experimental results of this work also showed that the filtered generated samples brought greater improvement to the performance of the ATR model. Moreover, the proposed simple filtering method also led to the appearance of counterfactual samples. Obviously, simply generating target samples for the ATR model is insufficient. However, there is little relevant research paying attention to the filtering of generated samples at present, especially for the filtering of counterfactual samples. Therefore, the proposed method intends to further explore the use of generated SAR target samples to improve the performance of ATR model. In the next subsection, related work on the filtering of generated counterfactual samples is introduced.

2.2. The Filtering of Counterfactual Target Samples

According to the discussion in the previous subsection, it may be feasible for GANs to stably generate counterfactual samples by category. However, the effect of counterfactual target sample on the ATR model may be opposite to the effect perceived by the naked eye. Moreover, the low interpretability of SAR target samples also increases the difficulty of judging those effects. Therefore, it is difficult for the naked eye to directly judge the effect of the generated counterfactual samples on the ATR model. In other words, all generated counterfactual samples need to be filtered before they are used in the training process of the ATR model. This phenomenon implies that the labeled counterfactual samples generated by proposed generation component need to be relabeled. At this time, the difficulty in the SAR ATR model with a small sample set is transformed into the following problem: a large number of counterfactual target samples and a small number of real SAR target samples are used to train the ATR model, where the counterfactual SAR target samples require filtering.

A variety of literature describes the use of pseudo-label assignment on unlabeled data as a way to improve ATR model performance. A notable example is a co-training algorithm, which utilizes two weak ATR models to produce pseudo-labels for unlabeled SAR target data to improve the ATR model recognition performance [38]. Theoretically, co-training can improve the recognition performance of a weak ATR model with arbitrary precision by using unlabeled SAR target data. However, the hypothesis of the co-training algorithm for the SAR target dataset is overly optimistic.

Co-training algorithm requires that each SAR target sample contains two “views”, called T_1 and T_2 , and each view must contain sufficient information to recognize the type of SAR target sample. Moreover, each SAR target sample can be recognized from T_1 and T_2 , and the two views must also be compatible, which means that the target types recognized from these two views are consistent. Additionally, the relationship between the two views must satisfy the conditional independence assumption; that is, the correlation between the two views cannot be too high. For example, when the target sample X_i is recognized as a positive sample from the view T_1 , it can also be recognized as a positive sample from the view T_2 . Moreover, the conditional independence assumption is also satisfied between views T_1 and T_2 . Obviously, these conditions are too strict for the current small sample SAR target dataset.

To appropriately relax the constraints of the co-training algorithm, Balcan et al. [39] define the concept of α -expansion:

$$P(V_1 \oplus V_2) \geq \alpha \min[P(V_1 \wedge V_2), P(V_2 \wedge V_1)], \quad (4)$$

where $V_i \subseteq T_i^+$. T_i^+ represents the positive sample set in view T_i . The probability that two views are correctly labeled as positive samples at the same time is represented by $P(V_1 \wedge V_2)$. The probability that only one view is correctly labeled as positive samples is

represented by $P(V_1 \oplus V_2)$. Balcan et al. also demonstrated that when the view satisfies the α -expansion, co-training can reduce the error rate of the initial ATR model after several iterations. Furthermore, the α -expansion can be applied to almost all datasets. Moreover, the α -expansion concept provides a reasonable explanation as to why the co-training algorithm is still effective in actual ATR tasks, even though it does not satisfy the conditional independence assumption.

However, the development of co-training for ATR tasks is still limited by the requirement for two “view” datasets. Wang et al. demonstrated that the recognition performance of the ATR model can be improved by the co-training algorithm if there are large differences between the two recognition models [40]. This reflects the essence of the co-training algorithm; that is, the difference between weak ATR models can be created by design. The method of effectively manufacturing the difference between two weak ATR models can replace the requirements for two “view” datasets. It also lays a further foundation for applying the co-training algorithm to the SAR ATR model.

Unfortunately, although the above methods relax the constraints of co-training on the dataset, they still cannot satisfactorily apply the co-training algorithm to an ATR model trained with a small sample set. When the ATR model is trained by a small sample set, it is almost impossible to ensure that each view contains sufficient information. Moreover, when there is an “insufficient view” due to a small sample set, the SAR ATR model struggles to recognize the SAR target sample types. Therefore, the optimal ATR model trained by each view results in the misclassification of some SAR target samples, which is contrary to the premise that different views conform to compatibility requirements. Thus, applying the co-training algorithm based on insufficient views is still a challenging issue. In this regard, Wang et al. proposed that the weak ATR model could additionally verify the recognition results based on providing additional recognition results, which could alleviate the difficulty of co-training with insufficient views [41]. This approach also provides an alternative way to solve the problem of insufficient views when applying co-training algorithms.

3. The Proposed Approach

3.1. Some Important Motivations about the Proposed Approach

General recognition algorithms tend to suffer from the problem of overfitting when the size of the training set is too small. The GAN is a complex model composed of a deconvolution network and a convolution network. When GANs are trained with limited samples, their risk of overfitting is likely to increase. Conversely, an ATR model with lower complexity trained with limited samples is prone to underfitting. Therefore, the motivation for implementing the proposed approach is introduced in this subsection. An interesting aspect of the proposed approach is that it takes advantages of the GANs tendency to overfit a small sample set to solve the limited samples problem. Figure 2 illustrates the above concept to provide an easier understanding.

Figure 2. The implementation motivation about the proposed approach.

GANs encounter overfitting, which means that the boundary curve of D is also overfitted under the condition of limited samples. The role of D is to determine whether the target sample generated by G is real or not. As shown in Figure 2, when GANs overfit the data, G generates some new samples, which include some counterfactual target samples. In addition, D also treats these generated counterfactual samples as real target samples. Therefore, GANs generate counterfactual target samples that they consider to be real samples. However, GANs generate counterfactual target samples with more information by overfitting them, and the overfitting also introduces more redundancy and erroneous counterfactual samples.

The proposed generation model generates many counterfactual target samples with different discrimination information. This means that some of the counterfactual samples are beneficial to the ATR model, while others may be harmful. Overall, this is the inevitable cost of increasing the discrimination information. However, the proposed method is designed to use counterfactual target samples that are beneficial to the ATR model to improve the model performance. This reflects the necessity of filtering the counterfactual target samples. It is also the core idea of the proposed integrated method. First, we make full use of the advantages of generation model overfitting to obtain counterfactual samples with more redundant information. Next, we need to filter counterfactual samples that are beneficial to the ATR model to improve its recognition performance. Finally, the ATR model can utilize these filtered counterfactual samples to compensate for part of the discrimination information missing from the small sample set.

The feature extraction module is also another important motivation for developing the method presented in this manuscript. The proposed generation model generates a large number of counterfactual samples with redundant information. Therefore, the feature extraction module is different from the general feature extraction process, as it is more similar to a filtering process. The feature extraction module needs to constantly update the discrimination information that can be used to improve the ATR model from the generated redundant information. The significance of the feature extraction module can be explained in conjunction with the desired filtering process.

In the case of limited information, the recognition model needs more discrimination information to effectively improve performance. The proposed generation model generates counterfactual samples with other redundant information. However, the redundant information contains valid discrimination information and erroneous information. Therefore, the updated discrimination information is different from the information accumulated after the sample gradually increases. It includes information filtering and information accumulation simultaneously. However, the feature extraction module cannot be independent of the proposed framework. The update of discriminant information requires the following steps to be realized: feature extraction $\rightarrow$ recognition $\rightarrow$ verification $\rightarrow$ selection $\rightarrow$ feature extraction. As shown in Figure 1, we use limited information to continuously train a new recognition model under the premise of utilizing different counterfactual samples. Although not all of these new ATR models learn the correct information, they can learn more discriminant information than the original model. The recognition models with more discriminative information can provide samples with higher-confidence recognition results. According to the recognition result, we rely on the proposed verification and selection strategy to obtain reliable counterfactual samples. Then, we realize the feature selection and feature accumulation of the counterfactual samples in this process. Finally, we move the correctional boundary curve of the ATR model closer to the desired boundary curve by using different counterfactual samples. The detailed architecture of the proposed method is introduced in the rest of this section.

3.2. The Generation of Counterfactual Target Samples

According to the aforementioned information, the proposed generation model should have the ability to stably generate counterfactual target samples by category. Any target sample input to D may come from the real SAR target sample distribution $P_r(x)$ or the

generated sample distribution $P_g(x)$. Moreover, the discriminator usually uses cross-entropy loss to calculate the loss of target samples from the two distributions. Therefore, the contribution of any sample to the generation model is as follows:

$$-P_r(x)[\log D(x)] - P_g(x)[\log(1 - D(G(x)))]. \quad (5)$$

assume that the derivative of Equation (5) with respect to $D(x)$ is 0. Then, the optimal discriminator $D(x)$ is obtained as follows:

$$D(x) = \frac{P_r(x)}{P_r(x) + P_g(x)}. \quad (6)$$

when the discriminator of the GANs is trained to reach the optimal status, the optimal discriminator is substituted into Equation (2). Then, the loss function of the GANs is transformed into the following form:

$$\begin{aligned} \min_G \max_D V(D, G) &= E_{x \sim P_r(x)} [\log \frac{2P_r(x)}{P_r(x) + P_g(x)}] + E_{x \sim P_g(x)} [\log \frac{2P_g(x)}{P_r(x) + P_g(x)}] - 2\log 2 \\ &= 2JS(P_r || P_g) - 2\log 2 \end{aligned} \quad (7)$$

At this time, the generation of SAR target samples is transformed into the problem of minimizing the Jensen-Shannon divergence between $P_r(x)$ and $P_g(x)$. When the support set of $P_r(x)$ and $P_g(x)$ is a low-dimensional manifold in a high-dimensional space, the probability that the measure of the overlap between $P_r(x)$ and $P_g(x)$ is 0 is 1 [35]. Moreover, the support set of $P_r(x)$ and $P_g(x)$ is obviously a low-dimensional manifold in X-dimensional space, where X is the number of pixel space dimensions of the target sample. Therefore, $P_r(x)$ and $P_g(x)$ barely overlap. The Jensen-Shannon divergence between $P_r(x)$ and $P_g(x)$ is $\log 2$ if $P_r(x)$ and $P_g(x)$ barely overlap, and this phenomenon often occurs in real ATR tasks. At this moment, it is worthless to utilize the optimization algorithm to derive a constant at. It is also easy to understand the reason that GANs cannot stably generate counterfactual target samples during the generation of the SAR target sample.

Therefore, the loss function of the generation model utilizes the Wasserstein distance to replace the Jensen-Shannon divergence as a distance measurement between $P_r(x)$ and $P_g(x)$. The loss function of the generation model is formulated as follows:

$$L \approx \max_{\omega: ||D_\omega||_L \leq K} E_{x \sim P_r(x)} [D_\omega(x)] - E_{x \sim P_g(x)} [D_\omega(G(x))], \quad (8)$$

where $||D_\omega||_L \leq K$ means that Equation (8) is established when the Lipschitz constant $||D_\omega||_L$ of D does not exceed the constant K . The parameters of D are represented by ω .

Moreover, although the selection of the Wasserstein distance measurement prevents the occurrence of a collapse mode, it still exposes the generation model to the risk of the gradient explosion problem. Therefore, the proposed generation model also needs to impose the 1-Lipschitz gradient penalty constraint on the Wasserstein distance. Then, the loss function of the generation model is formulated as follows:

$$L = E_{x \sim P_g} [D_\omega(G(x))] - E_{x \sim P_r} [D_\omega(x)] + \lambda E_{x \sim P_x} [(||\nabla_x D_\omega(x)||_2 - 1)^2], \quad (9)$$

where P_x is calculated by $P_x = \epsilon P_r + (1 - \epsilon) P_g$, and λ is a gradient penalty constant. Thus far, the proposed generation model can stably generate counterfactual target samples.

In addition, the proposed generation model also needs to generate counterfactual samples by category. This means that the proposed generation model also needs to use category information during the generation process. Therefore, the loss function adopted in the proposed generation model is formulated as follows:

$$L = E_{x \sim P_g} [D_\omega(G(x)|y)] - E_{x \sim P_r} [D_\omega(x|y)] + \lambda E_{x \sim P_x} [(||\nabla_x D_\omega(x|y)||_2 - 1)^2], \quad (10)$$

where the target category information is represented by y .

Next, the proposed architecture of the generation model is explained. The architecture overview of the proposed generation model is shown in Figure 3. The generator only consists of deconvolution layers, and no fully connected or pooling layers are utilized in G. Moreover, strided deconvolutions are used to replace the pooling layer to ensure that the generation model can generate images and of the correct size. The discriminator is composed of several convolutional layers and a fully connected layer. The fully connection layer is used to replace the sigmoid function in the last layer. Moreover, the original GANs utilizes the sigmoid function to judge whether or not the SAR target sample is real. However, the purpose of the discriminator in the proposed generation model has been changed to minimize the Wasserstein distance between $P_r(x)$ and $P_g(x)$. Therefore, it is reasonable to remove the sigmoid function for solving regression tasks.

Figure 3. The architecture overview of the proposed generation model. The generation model consists of a deconvolution network and a convolution network. The condition information y refers to the target label condition. z means the initialization noise when the deconvolution network is initialized. $G(z)$ indicates the generated counterfactual target samples.

3.3. The Filtering of Counterfactual Target Samples

The proposed generation model can stably generate counterfactual target samples by category. However, merely generating counterfactual samples cannot alleviate the difficulty in the SAR ATR model when it encounters a small sample set. The proposed generation model generates many counterfactual target samples with more discrimination information. It is worth noticing that the counterfactual samples with more discrimination information are obtained by the GANs overfitting a small sample set. Therefore, the proposed method also needs to be able to filter the counterfactual target samples that are beneficial to the ATR model. Moreover, the proposed filtering method needs to conform to the premise of a small sample set. By implementing the generation and filtering of counterfactual samples simultaneously, it becomes feasible to improve the performance of the ATR model through counterfactual target samples.

Though the counterfactual target samples are generated by the proposed generation model, the problem of using the SAR ATR model with a small sample set is not solved immediately. Among the generated counterfactual samples, some samples may improve the performance of the ATR model, and others may degrade it. At this time, the difficulty in the SAR ATR model with a small sample set is transformed into the following problem: a large number of counterfactual target samples and a small number of real SAR target samples are used to train the ATR model, where the counterfactual SAR target samples require filtering. In other words, all generated counterfactual samples that have been labeled need to be relabeled. Fortunately, the co-training algorithm uses a small number of

labeled samples and a large number of unlabeled samples to improve the performance of the recognition model.

When the co-training algorithm is used to improve the performance of the ATR model, its strict constraints on the dataset cannot be ignored. Therefore, it is necessary to use the existing dataset to construct application scenarios for the co-training algorithm. First, the SAR target sample data are generally recognized from the pixel perspective. Thus, it is difficult for the existing dataset to meet the co-training algorithm's requirement of two "views". At this point, the proposed filtering approach needs to manufacture differences between multiple weak ATR models to satisfy this requirement of co-training. Second, it is impossible to guarantee sufficient "view" information due to the influence of a small sample set. In other words, the different weak ATR models may result in the misrecognition of some SAR target samples, which is contrary to the premise that different views conform to the compatibility requirement. Therefore, the different weak ATR models need verification when they add pseudo-labels to the counterfactual target samples, which can alleviate the limitations of insufficient views. In this regard, the category information provided by the counterfactual target samples during the generation process can be used as verification.

According to the above requirements, the architecture of the proposed generation model is shown in Figure 4. The proposed approach integrates a generation model and a batch of SVMs for filtering. The generation model is capable of continuously generating counterfactual target samples and applying them to a subsequent filtering process. Then, a batch of SVMs is trained by different counterfactual target samples sets to manufacture the significant differences between them. The filtering process that utilizes and labels the counterfactual target samples is an important component of the proposed approach.

Figure 4. The architecture of the proposed approach. The proposed approach consists of a generation model and a batch of SVMs with differences. The yellow dashed line indicates the proposed filtering process that uses and labels the generated samples.

In a loop of the proposed filtering process, support vector machines are trained in order. However, different counterfactual target samples sets with pseudo-labels are not provided in order. The filtered counterfactual target samples are not the filter results of the last support vector machine. Regardless of whether the support vector machine or a filtered counterfactual target sample contain all of the knowledge accumulated during the iteration process, the previous SVM is first trained by the collected SAR target samples. Once the SVM is trained using the SAR target samples, it can filter the newly counterfactual target samples. Then, the filtered counterfactual target samples are used to update the existing training sample set, and the updated training sample set leads to the training of a new SVM. At this time, the new SVM can re-filter the previously filtered counterfactual target samples but not the original SAR target sample set, allowing it to be trained using a combination of the previous training sample set and the latest batch of filtered samples.

Thus, the recognition performance of the new SVM is higher than that of the previous SVM. The re-filtering process also improves recognition performance. Once the SAR target samples are filtered, they enter the next loop of the proposed filtering process, or the re-filtering process, where both the SVM and the training SAR target sample set are gradually improved and refined. Finally, when the recognition rate of the SVM no longer significantly improves with the change in the counterfactual target samples set, the proposed filtering process ends all loops.

In a word, the essence of the filtering process of counterfactual target samples is the framework of the co-training algorithm. Just as the co-training algorithm relied on the discrimination information of unlabeled data, the proposed filtering component also relied on the discrimination information provided by the generated target samples. The proposed approach continuously produced counterfactual target samples by generation models, and manufactured the difference between any two SVMs, which laid a foundation for improving the performance of the recognition model. The proposed filtering process is also a double iteration process. The favorable counterfactual target samples set trains support vector machines to achieve better recognition performance. In turn, a support vector machine with stronger recognition performance can filter the favorable counterfactual target samples more accurately. The favorable counterfactual target samples and the performance of the recognition model are improved synchronously in the above iteration process.

It is necessary to provide a special explanation of how to verify and select the favorable counterfactual target samples during the filtering process. The new recognition models with more discrimination information provide the counterfactual target samples with higher-confidence recognition results. However, under the condition of limited information, the discrimination information of the ATR model is only increased compared with the information before updating. In other words, the discrimination information possessed by the recognition model is still insufficient. The recognition model still faces the risk of making mistakes when it filters favorable counterfactual target samples. From the co-training algorithm perspective, the current dilemma is a single “view” training problem with insufficient information. In this regard, the proposed approach needs to verify the recognition results to alleviate the difficulty of insufficient discrimination information in the ATR model. At this time, the condition information used to generate the model also plays a verification role. There is no threshold comparison during the proposed filtering process. Such counterfactual target samples can only be selected to update the feature when the recognition results of counterfactual target samples are consistent with the condition information of the generation model. Moreover, the filtered counterfactual target samples are filtered again after the ATR model is updated. Because the discrimination information of the ATR model gradually increases in the process of updating features, different ATR models may have different perceptions of the same batch of counterfactual target samples.

4. Experimental Results

4.1. Experiment Arrangement and Experiment Requirements

The experiments were divided into two parts to be implemented separately in this study. First, ablation experiments using the proposed approach were performed to demonstrate that each component of the proposed approach has a specific role. Second, recognition verification experiments were performed to show the recognition performance of the proposed method when it faces a small sample set. The experiments were performed on the moving and stationary target acquisition and recognition (MSTAR) dataset. MSTAR consists of X-band and HH polarization SAR target samples with a 0.3 m resolution [42]. All target samples were captured at two different pitch angles: 15 degrees and 17 degrees. All target slices in Figure 5 are 64×64 RAW data. SAR target samples with a pitch angle of 17 degrees were used as the training set, and those with a pitch angle of 15 degrees were used as test samples. With the use of the proposed model, the experiment obtained favorable counterfactual target samples. These counterfactual target samples were used to expand the size of the training sample set and subsequently verify the recognition per-

formance of the proposed approach. The same test sample set was used in all recognition performance tests in subsequent experiments.

Under the premise of a small sample set, the aim of the proposed method is to improve the performance of the recognition model through the generation and filtering of counterfactual target samples. To establish the scenario in which the SAR ATR model is used with a small sample set, only 40 SAR target samples from each category were selected as the training set in the experiments, which means that the size of the constructed training set was only 14.5% of the original training set. This means that it is difficult to improve the recognition performance to an acceptable level using existing data-driven methods.

Figure 5. (Top) SAR images and (Bottom) optical images of ten categories of target samples. According to the order: 2S1, BMP2, BRDM2, BTR60, BTR70, D7, T62, T72, ZIL131, and ZSU23/4.

4.2. The Ablation Experiments of the Proposed Generation Component

The first ablation experiment demonstrated the effectiveness of the proposed generation component. The proposed generation component can stably generate counterfactual target samples by category. Therefore, it is necessary to compare the different generated counterfactual target samples based on the different architecture of conditional GANs. In the experiments, the generated samples of conditional GANs and conditional deep convolutional GANs were compared. The comparison results for different generation models trained by the constructed small sample dataset are shown in Figure 6. Moreover, real target samples that are similar to the generated results were used for comparison to verify the difference between the generated results and the real target samples.

Figure 6. The SAR images of ten categories target samples. According to the order: 2S1, BMP2, BRDM2, BTR60, BTR70, D7, T62, T72, ZIL131, and ZSU23/4. Row 1: the real SAR images of ten categories target samples. Row 2: The SAR images of ten categories target samples generated by the proposed model. Row 3: The SAR images of ten categories target samples generated by conditional GAN. Row 4: The SAR images of ten categories target samples generated by conditional deep convolutional GAN.

4.3. The Ablation Experiments of the Proposed Filtering Component

It is not sufficient to directly filter SAR target samples with the naked eye due to the low interpretability of such samples, which highlights the importance of implementing a filtering method. The second experiment explored the effectiveness of the proposed filtering process. However, it is also necessary to prove the necessity of filtering the generated counterfactual target samples before demonstrating the superiority of the proposed filtering process. These experiments used the SVM as the filtering tool. The filtering results of the three generated samples are shown in Table 1.

Table 1. The comparison results of different generated samples through the filtering model.

Efficiency Indicators	Conditional GAN	Conditional Deep Convolutional GAN	Proposed Generation Component
total number of generated samples	1000	1000	1819
the type of generated sample	unable to filter	unable to filter	need to filter
filtering pass rate	0%	0%	16.49%
total number of filtered samples	0	0	300

Besides, under the condition of a small sample set, subtle changes in the training set have little effect on the recognition performance of the ATR model [31]. Therefore, the experiment was also performed with a severely underfitting ATR model to prove the necessity of filtering the generated counterfactual target samples. In the experiments, 100 target samples were randomly selected from the original training set. Subsequently, four different types of target samples were added to the constructed training set in batches, and the ATR model was retrained to observe the performance of the recognition model. The experiments used 50 target samples in each batch. The same experimental steps were also performed after filtering the three generated target samples. However, it was difficult for the other generated samples to pass the filtering step of the pre-training recognition model, except for the samples generated by the proposed method. The experimental results obtained are shown in Figure 7.

Figure 7. The comparison chart of the recognition performance of the ATR model obtained through different training sample sets. Only the target samples generated by the proposed generation model can be filtered.

Next, experiments were performed to prove the superiority of the proposed filtering process. There are few studies on the filtering of generated samples under the condition

of a small sample set. Filtering methods using SVM have proved effective. Therefore, the second experiment compared the proposed filtering process with the filtering methods using SVM. The second experiment used two different filtering methods to filter the same batch of generated target samples, which produced an interesting experimental result. The counterfactual target samples filtered by the two different methods are shown in Figure 8.

Moreover, it is not sufficient for the generated counterfactual target samples to be filtered by the proposed filtering process. The experiments also needed to verify whether the filtered generated target samples contribute to the performance of the ATR model. Therefore, the two methods were used to filter the same batch of generated samples in the above experiment. Then, 683 generated counterfactual target samples were obtained by the proposed filtering process. At the same time, 947 generated counterfactual target samples were filtered by filtering methods using SVM.

Figure 8. The filtered counterfactual target samples of ten categories. According to the order: 2S1, BMP2, BRDM2, BTR60, BTR70, D7, T62, T72, ZIL131, and ZSU23/4. Row 1: The counterfactual target samples filtered by proposed filtering process. Row 2: The counterfactual target samples filtered by SVM.

Subsequently, two different types of generated counterfactual target samples were added to the constructed small sample set in separate experiments. The two sets of training samples were utilized to train the same convolutional neural networks, and their recognition performance was then compared. The comparison results of ATR model recognition performance are shown in Figure 9. Moreover, the experiment also investigated the recognition performance of SVM and CNN when using small sample sets. To compare the recognition results, when the number of training samples no longer increased, the best performance of the recognition model was used as its subsequent recognition result. Therefore, different recognition performance curves stopped growing at different times.

Figure 9. The ATR model recognition performance comparison results obtained by different filtering methods. All recognition models are identical to the architecture of the above ATR model.

4.4. The Recognition Performance of the Proposed Approach

The previous experiments prove the effectiveness of the various components of the proposed method. The recognition performance of the proposed approach using a small sample set is shown in this subsection. Convolutional neural networks (CNNs) and support vector machines (SVMs) are currently the two most widely used methods in the SAR ATR model. Therefore, the recognition performance of the proposed approach was compared with that of the two recognition methods to show its performance in solving the problem of a small sample set.

To prove the superiority of the proposed method, experiments were designed to show the negative influence of the small sample set on the ATR model. The experiments used the original training set to train the ATR models based on CNN and SVM. Then, a small sample set was constructed to train the above two ATR models separately. The recognition performance of different ATR models is shown in Figure 10. The subscript S indicates that the current model was trained by the constructed small sample set. The dotted line represents the recognition performance achieved by the two ATR models when facing the small sample set. The solid line indicates the recognition rate of ATR models when the number of SAR target samples is sufficient. Similarly, when the number of training samples no longer increased, the best performance of the recognition model was used as the subsequent recognition result. In other words, the recognition performance of the SAR ATR model no longer increased due to the limitation of the number of training samples. In this way, the experiments aimed to show the negative influence of a small sample set on existing popular recognition algorithms.

Figure 10. The recognition performance of different ATR models. The dotted line represents the recognition performance achieved by the two ATR models when facing the small sample set. The solid line indicates the recognition rate of ATR models when the number of SAR target samples is sufficient.

Subsequently, experiments were performed to prove that the proposed method is able to eliminate the negative influence. The recognition performance of data-driven ATR models is not ideal when the number of training samples is limited. Therefore, the proposed method expands the training set by generating counterfactual target samples. After expanding the training set with counterfactual target samples, these samples were used in experiments to prove the advantages of the proposed method.

The proposed approach can generate counterfactual target samples autonomously and filter them to improve the recognition performance of the ATR model. The counterfactual

target samples need to be filtered before they are used to update the classifier. Moreover, a small number of filtered counterfactual target samples has little impact on the classifier update if the number of generated SAR target samples is small. Therefore, the experiments accumulated a certain number of counterfactual target samples before starting the proposed filtering process. In this experiment, 100 samples were selected from the generated target samples every time the generation model results converged. A total of 4100 generated target samples generated from 41 batches were used in this experiment. A total of 683 available counterfactual target samples emerged after the end of the proposed filtering process loop. Then, those available counterfactual target samples were combined with 400 real target samples to form a new training sample set. The new training sample set was used to train the same convolutional neural networks. The results of comparing recognition performance between different recognition models are shown in Figure 11.

Figure 11. The recognition performance comparison results of different recognition models. The solid line indicates the recognition rate of different recognition models when the amount of SAR target samples is sufficient. The dotted line indicates the recognition rate of different recognition models when the ATR model solves the dilemma of a small sample set.

At the same time, it is well-known that MSTAR is a very friendly dataset to existing recognition methods. It is not enough for the proposed method to show excellent performance on a single dataset. Therefore, the experiments were also ready to verify the effectiveness of the proposed method on OpenSARship [43]. OpenSARship is ship data derived from the Sentinel-1. OpenSARship mainly included the following five types of target samples: bulkcarrier, cargo, containership, generalcargo, and tanker. The experiments cut out 1521 ship slices from the large scene image of OpenSARship. The details of the constructed ship dataset utilized in the experiment are shown in Table 2.

Table 2. The details of the constructed ship dataset.

Target Type	The Number of Train Set	The Number of Test Set
bulkcarrier	200	200
cargo	200	200
containership	80	80
generalcargo	205	204
tanker	76	76

The experiment randomly selected 40 ship samples from each type of target sample. We followed the same steps to carry out the recognition performance verification experiment shown in the manuscript. The comparing results of ship recognition performance between different recognition models are shown in Figure 12.

Figure 12. The comparing results of ship recognition performance between different recognition models. The solid line indicates the recognition rate of different recognition models when the amount of SAR target samples is sufficient. The dotted line indicates the recognition rate of different recognition models when the ATR model solves the dilemma of a small sample set.

5. Discussion

5.1. The Analysis of the Proposed Generation Component

Figure 6 reflected the characteristics of counterfactual target samples generated by different generation models. The conditional GANs only learned some contour features of SAR target samples due to the MLP architecture. In contrast to the generated results of conditional GANs, conditional deep convolutional GANs learned detailed features of SAR target samples. However, the generated results were inferior to those of the proposed generation model due to the different distance measurements between the above generation methods. Figure 6 also showed that the target samples generated by the proposed generation model are almost the same as the real SAR target samples. The experimental results also proved that the generated sample distribution was closest to the real data distribution with the proposed generation model.

5.2. The Analysis of the Proposed Filtering Component

As shown in Table 1, the filtering pass rate of target samples generated by conditional GANs and conditional deep convolutional GANs is almost zero. In contrast, some of target samples generated by the proposed method can be filtered. It showed that different generated samples also have obvious differences in the same recognition model. However, the ATR model requires independent and identically distributed target sample sets for training. This phenomenon reflected the necessity of filtering to generated target samples.

As shown in Figure 7, expanding the size of the training sample set by generating target samples may not necessarily improve the performance of the ATR model. If the target samples generated by the conditional GAN are not filtered, the performance of the ATR model is degraded. At the same time, other generated target samples have improved the performance of the ATR model to varying degrees. Among them, the target samples generated by the proposed generation component improve the performance of the ATR model by 18.49%. Moreover, when the generated samples are filtered, the recognition rate of the ATR model is improved by 28.18%. The recognition performance of ATR

improved by 9.69% compared with when the generated sample was not filtered. All of the experimental results illustrate the necessity of filtering the generated counterfactual target samples. Moreover, it also meant that the other two generated samples cannot be used to improve the performance of the ATR model. Therefore, the experiments did not use the other two generated samples to improve the performance of the ATR model in subsequent experimental results.

With observation of Figure 8, the judgment of counterfactual samples with the naked eye may be contrary to the results of the proposed filtering process. For example, although the generated samples in the second row of Figure 8 appear to be more similar to the real SAR target samples, they were not filtered by the proposed filtering process. Conversely, the generated samples in the first row were filtered by the proposed filtering process. However, these filtered SAR target samples are not consistent with any target samples with a pitch angle of 17 degrees. This experimental phenomenon proved the effectiveness of the proposed filtering component. Moreover, it also verified the additional discriminative information provided by the generation model from the visual results.

According to the experimental results shown in Figure 9, although the number of generated samples filtered by the proposed filtering method was smaller, it contributed more to the performance of the ATR model. The experimental results also showed that the proposed filtering method can increase the recognition rate by another 4.02% compared with the existing filtering method. It illustrated the importance of choosing filtering methods for generated target samples.

Moreover, we also made some inferences regarding why the generated counterfactual target samples improved the performance of the ATR model, which were derived from the experimental results in Figure 8. Some of the filtered counterfactual target samples were completely different from the original training set. However, the generation model is based on the distribution fitting. In other words, the data distributions of all filtered counterfactual target samples are identical to that of the target samples in the original training set. Although the counterfactual target samples do not allow us to accept it visually, these samples do contribute to the improvement of the performance of the ATR model. The ATR model utilized counterfactual target samples to achieve better recognition performance, which is different from the method of expanding the dataset from the target sample level. The algorithm expands the dataset at the distribution level. The proposed method can autonomously generate counterfactual target samples and provide effective filtering.

5.3. The Analysis of Recognition Performance

According to the results in Figure 10, CNN is a recognition model with strong feature expression ability. When the training samples were sufficient, the recognition rate of CNN reached 95.53%. However, this ability also increases the model complexity of CNN. Model complexity is positively correlated with the number of training samples. In other words, when the training samples were insufficient, a small sample set degraded the performance of the CNN recognition model. The requirements for larger datasets made CNN more susceptible to the reduced number of training samples. When the number of training samples was only 400, the CNN recognition performance degraded to 72.67%. Thus, the recognition performance of CNN dropped by 22.86% due to the impact of a small sample set. Similarly, the recognition performance of SVM degraded from 92.97% to 79.47% as the number of training samples decreased. All experimental results proved the negative impact of the small sample set on the ATR model. Moreover, the experimental results also showed that the recognition model based on CNN architecture is more affected by a small sample set than that based on SVM. This is also the reason why the proposed method uses SVM in the filtering process.

The above experiments show the negative impact of a small sample set on different recognition models. As shown in Figure 11, the proposed approach alleviated the negative influence of a small sample set on the SAR ATR model. The recognition rate of the ATR model with the use of the proposed approach reached 91.27% when the number

of SAR target samples was only 400. With the current dataset, this recognition rate was 18.6% and 11.8% higher than the recognition performance of CNN and SVM, respectively. More surprisingly, the recognition rate of the proposed model under the small sample set condition was 2.23% higher than that of SVM when the number of training samples was sufficient. Obviously, if the ATR model did not rely on the generated samples to provide additional discriminative information, it was difficult to improve the recognition performance by more than 10% without making any changes. The comparison results for recognition performance also illustrate the superiority of the proposed approach in dealing with the problem of a small sample set. Moreover, the CNN-based algorithms certainly cannot achieve the expected recognition performance under the condition of a small sample set. At the same time, the proposed method intended to improve the recognition performance of ATR model in this scenario. Therefore, compared with the expected performance of the CNN model (under the condition of sufficient samples), the proposed method still has a performance gap. The performance of the proposed method still has 2.04% gap with the expected performance of the CNN-based algorithm. Although proposed method can alleviate the impact of the small sample set on the ATR model, it cannot completely offset all the missing information in the small sample set.

According to the results in Figure 12, SVM and CNN both showed obvious performance degradation when facing OpenSARship. Although SVM and CNN of the same architecture can reach 92.97% and 95.53% on MSTAR, respectively, the performance of these two recognition models has dropped to 65.26% and 56.34% on OpenSARship. It proved that the OpenSARship was indeed a challenging dataset. However, the proposed method still increased the recognition rate of the ATR model by 4.59%. It demonstrated that the proposed method can show its effectiveness in solving the small sample set problem of different SAR target datasets.

6. Conclusions

Data-driven ATR algorithms are heavily dependent on the size of the training sample set. In this paper, we proposed an integrated counterfactual sample generation and filtering approach to alleviate the negative influence of a small sample set on the SAR ATR model by increasing the size of the training sample set. Ablation experiments were performed to demonstrate the effectiveness of the various components of the proposed method. The experiments showed that counterfactual target samples were generated by the proposed generation component. Furthermore, the proposed filtering component increased the recognition rate by another 4.02% compared with the existing filtering method. Moreover, the experiments demonstrated that the proposed approach helped to improve recognition performance when the ATR model encountered the small sample set problem. When the number of training samples was 14.5% of the original training sample set, the recognition rate of the ATR model reached 91.27% with the proposed approach. Under the conditions of the current dataset, this recognition rate was 18.6% and 11.8% higher than that of CNN and SVM, respectively. More surprisingly, the recognition rate of the proposed model under the small sample set condition was 2.23% higher than that of SVM when the number of training samples was sufficient. Moreover, the experimental results on OpenSARship demonstrated that the proposed method also showed its effectiveness in solving the small sample set problem of different SAR target datasets. All experimental results showed that the proposed approach has various advantages. Therefore, the proposed approach provides an alternative way to solve the problem of the SAR ATR model with a small sample set.

Author Contributions: Formal analysis, C.C.; Funding acquisition, Z.C. (Zongyong Cui), Z.C. (Zongjie Cao) and J.Y.; Investigation, L.W.; Methodology, C.C.; Project administration, J.Y.; Visualization, L.W.; Writing—original draft, C.C.; Writing—review & editing, Z.C. (Zongyong Cui), Z.C. (Zongjie Cao) and J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China under Grant 61801098 and 61971101, Shanghai Aerospace Science and Technology Innovation Fund SAST2018-079, and Science and Technology on Automatic Target Recognition Laboratory (ATR) Fund 6142503190201.

Data Availability Statement: All data sets used in this article are public data sets.

Acknowledgments: The authors would like to thank the OpenSARship for providing SAR images.

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

References

1. Sun, Y.; Liu, Z.; Todorovic, S.; Li, J. Adaptive boosting for SAR automatic target recognition. *IEEE Trans. Aerosp. Electron. Syst.* **2007**, *43*, 112–125. [\[CrossRef\]](#)
2. Jianxiong, Z.; Zhiguang, S.; Xiao, C.; Qiang, F. Automatic Target Recognition of SAR Images Based on Global Scattering Center Model. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3713–3729. [\[CrossRef\]](#)
3. O’Sullivan, J.; Devore, M.; Kedia, V.; Miller, M. SAR ATR performance using a conditionally Gaussian model. *IEEE Trans. Aerosp. Electron. Syst.* **2001**, *37*, 91–108. [\[CrossRef\]](#)
4. Srinivas, U.; Monga, V.; Raj, R.G. SAR Automatic Target Recognition Using Discriminative Graphical Models. *IEEE Trans. Aerosp. Electron. Syst.* **2014**, *50*, 591–606. [\[CrossRef\]](#)
5. Lu, C.; Feng, J.; Chen, Y.; Liu, W.; Lin, Z.; Yan, S. Tensor Robust Principal Component Analysis with a New Tensor Nuclear Norm. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *42*, 925–938. [\[CrossRef\]](#)
6. Deng, X.; Tian, X.; Chen, S.; Harris, C.J. Nonlinear Process Fault Diagnosis Based on Serial Principal Component Analysis. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *29*, 560–572. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Yang, J.; Zhang, D.; Frangi, A.F.; Yang, J.Y. Two-dimensional PCA: A new approach to appearance-based face representation and recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 131–137. [\[CrossRef\]](#)
8. Jia, Y.; Kwong, S.; Hou, J.; Wu, W. Semi-Supervised Non-Negative Matrix Factorization With Dissimilarity and Similarity Regularization. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *31*, 2510–2521. [\[CrossRef\]](#)
9. Hu, P.; Peng, D.; Sang, Y.; Xiang, Y. Multi-View Linear Discriminant Analysis Network. *IEEE Trans. Image Process.* **2019**, *28*, 5352–5365. [\[CrossRef\]](#)
10. Li, C.-N.; Shao, Y.-H.; Yin, W.; Liu, M.-Z. Robust and Sparse Linear Discriminant Analysis via an Alternating Direction Method of Multipliers. *IEEE Trans. Neural Networks Learn. Syst.* **2019**, *31*, 915–926. [\[CrossRef\]](#)
11. Richhariya, B.; Tanveer, M. A reduced universum twin support vector machine for class imbalance learning. *Pattern Recognit.* **2020**, *102*, 107150. [\[CrossRef\]](#)
12. Zhao, Q.; Principe, J.C. Support vector machines for SAR automatic target recognition. *IEEE Trans. Aerosp. Electron. Syst.* **2001**, *37*, 643–654. [\[CrossRef\]](#)
13. Kechagias-Stamatis, O.; Aouf, N.; Belloni, C. SAR Automatic Target Recognition based on Convolutional Neural Networks. In Proceedings of the International Conference on Radar Systems, Belfast, UK, 23–26 October 2017; doi: 10.1049/cp.2017.0437. [\[CrossRef\]](#)
14. Pei, J.; Huang, Y.; Huo, W.; Zhang, Y.; Yang, J.; Yeo, T.-S. SAR Automatic Target Recognition Based on Multiview Deep Learning Framework. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 2196–2210. [\[CrossRef\]](#)
15. Zhang, X.-Y.; Yin, F.; Zhang, Y.-M.; Liu, C.-L.; Bengio, Y. Drawing and Recognizing Chinese Characters with Recurrent Neural Network. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 849–862. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Liu, X.; Huang, Y.; Pei, J.; Yang, J. Sample discriminant analysis for SAR ATR. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 2120–2124.
17. Karjalainen, A.I.; Mitchell, R.; Vazquez, J. Training and Validation of Automatic Target Recognition Systems using Generative Adversarial Networks. In Proceedings of the 2019 Sensor Signal Processing for Defence Conference (SSPD), Brighton, UK, 9–10 May 2019; doi: 10.1109/sspd.2019.8751666. [\[CrossRef\]](#)
18. Choe, J.; Park, S.; Kim, K.; Park, J.H.; Kim, D.; Shim, H. Face Generation for Low-Shot Learning Using Generative Adversarial Networks. In Proceedings of the 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, Italy, 22–29 October 2017; doi: 10.1109/iccvw.2017.229. [\[CrossRef\]](#)
19. Shin, H.-C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans. Med. Imaging* **2016**, *35*, 1285–1298. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Xie, M.; Jean, N.; Burke, M.; Lobell, D.; Ermon, S. Transfer learning from deep features for remote sensing and poverty mapping. In Proceedings of the 30th AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; Volume 30.
21. Huang, Z.; Pan, Z.; Lei, B. Transfer Learning with Deep Convolutional Neural Network for SAR Target Classification with Limited Labeled Data. *Remote Sens.* **2017**, *9*, 907. [\[CrossRef\]](#)
22. Wong, S.C.; Gatt, A.; Stamatescu, V.; McDonnell, M.D. Understanding data augmentation for classification: When to warp. In Proceedings of the 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Gold Coast, QLD, Australia, 30 November–2 December 2016.

23. Malmgren-Hansen, D.; Kusk, A.; Dall, J.; Nielsen, A.A.; Engholm, R.; Skriver, H. Improving SAR Automatic Target Recognition Models With Transfer Learning From Simulated Data. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1484–1488. [\[CrossRef\]](#)
24. Zhong, Z.; Zheng, L.; Zheng, Z.; Li, S.; Yang, Y. CamStyle: A Novel Data Augmentation Method for Person Re-Identification. *IEEE Trans. Image Process.* **2018**, *28*, 1176–1190. [\[CrossRef\]](#)
25. Guo, J.; Lei, B.; Ding, C.; Zhang, Y. Synthetic Aperture Radar Image Synthesis by Using Generative Adversarial Nets. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1111–1115. [\[CrossRef\]](#)
26. Gao, F.; Yang, Y.; Wang, J.; Sun, J.; Yang, E.; Zhou, H. A Deep Convolutional Generative Adversarial Networks (DCGANs)-Based Semi-Supervised Method for Object Recognition in Synthetic Aperture Radar (SAR) Images. *Remote Sens.* **2018**, *10*, 846. [\[CrossRef\]](#)
27. Schwegmann, C.P.; Kleyhans, W.; Salmon, B.P.; Mdakane, L.; Meyer, R. Synthetic aperture radar ship discrimination, generation and latent variable extraction using information maximizing generative adversarial networks. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 2263–2266. [\[CrossRef\]](#)
28. Zheng, Z.; Zheng, L.; Yang, Y. Unlabeled Samples Generated by GAN Improve the Person Re-identification Baseline in Vitro. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 3774–3782. [\[CrossRef\]](#)
29. Cui, Z.; Zhang, M.; Cao, Z.; Cao, C. Image Data Augmentation for SAR Sensor via Generative Adversarial Nets. *IEEE Access* **2019**, *7*, 42255–42268. [\[CrossRef\]](#)
30. Tang, H.; Zhao, Y.; Lu, H. Unsupervised Person Re-Identification with Iterative Self-Supervised Domain Adaptation. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Long Beach, CA, USA, 16–17 June 2019; pp. 1536–1543.
31. Zhang, M.; Cui, Z.; Wang, X.; Cao, Z. Data Augmentation Method of SAR Image Dataset. In Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2018), Valencia, Spain, 22–27 July 2018; pp. 5292–5295.
32. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Bengio, Y. Generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems 27 (NIPS 2014), Montreal, QC, Canada, 8–11 December 2014; pp. 2672–2680.
33. Radford, A.; Metz, L.; Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv* **2015**, arXiv:1511.06434.
34. Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein generative adversarial networks. In Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; pp. 214–223.
35. Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; Courville, A.C. Improved training of wasserstein gans. In Proceedings of the Advances in Neural Information Processing Systems 30 (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; pp. 5767–5777.
36. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784.
37. Cao, C.; Cao, Z.; Cui, Z. LDGAN: A Synthetic Aperture Radar Image Generation Method for Automatic Target Recognition. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 3495–3508. [\[CrossRef\]](#)
38. Blum, A.; Mitchell, T. Combining labeled and unlabeled data with co-training. In Proceedings of the COLT98: The 11th Annual Conference on Computational Learning Theory, Madison, WI, USA, 5–8 July 1995; pp. 92–100.
39. Balcan, M.F.; Blum, A.; Yang, K. Co-training and expansion: Towards bridging theory and practice. In Proceedings of the Advances in Neural Information Processing Systems 28 (NIPS 2015), Montreal, QC, Canada, 7–12 December 2015; pp. 89–96.
40. Wang, W.; Zhou, Z.H. Analyzing co-training style algorithms. In *European Conference on Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 454–465.
41. Wang, W.; Zhou, Z.H. Co-training with insufficient views. In Proceedings of the Asian conference on machine learning, Canberra, Australia, 13–15 November 2013; pp. 467–482.
42. Keydel, E.R.; Lee, S.W.; Moore, J.T. MSTAR extended operating conditions: A tutorial. In Proceedings of the Aerospace/Defense Sensing and Controls, Orlando, FL, USA, 10 June 1996; pp. 228–242.
43. Huang, L.; Liu, B.; Li, B.; Guo, W.; Yu, W.; Zhang, Z.; Yu, W. OpenSARShip: A dataset dedicated to Sentinel-1 ship interpretation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *11*, 195–208. [\[CrossRef\]](#)