



Article

Ship Detection in Low-Quality SAR Images via an Unsupervised Domain Adaption Method

Xinyang Pu ¹, Hecheng Jia ¹ , Yu Xin ², Feng Wang ¹ and Haipeng Wang ^{1,*}

¹ Key Laboratory for Information Science of Electromagnetic Waves (MoE), Fudan University, Shanghai 200433, China; xypu21@m.fudan.edu.cn (X.P.); hcjia19@fudan.edu.cn (H.J.); fengwang@fudan.edu.cn (F.W.)

² Beijing Institute of Remote Sensing Information, Beijing 100192, China; autoxinyu0902@163.com

* Correspondence: hpwang@fudan.edu.cn

Abstract: Ship detection in low-quality Synthetic Aperture Radar (SAR) images poses a persistent challenge. Noise signals in complex environments disrupt imaging conditions, hindering SAR systems from acquiring precise target information, thereby significantly compromising the performance of detectors. Some methods mitigate interference via denoising techniques, while others introduce noise using classical methods to learn target features in the presence of noise. This conundrum is regarded as a cross-domain problem in this paper; a ship detection method in low-quality images is proposed to learn features of targets and shrink serious deterioration of detection performance by utilizing Generative Adversarial Networks (GANs). First, an image-to-image translation task is implemented using CycleGAN to generate low-quality SAR images with complex interference from the source domain to the target domain. Second, with the annotation inheritance, these generated SAR images participate in a training process to improve the detection accuracy and model robustness. Multiple experiments indicate that the proposed method conspicuously improves the detection performance and efficaciously reduces the missed detection rate in the SAR ship detection task. This cross-domain approach achieved outstanding improvements in the form of 11.0% mAP and 3.22% mAP on the GaoFen-3 ship dataset and SRSSD-V1.0, respectively. In addition, the characteristics and potentials of near-shore and off-shore SAR image reconstruction with style transfer based on Generative Adversarial Networks were explored and analyzed in this work.



Citation: Pu, X.; Jia, H.; Xin, Y.; Wang, F.; Wang, H. Ship Detection in Low-Quality SAR Images via an Unsupervised Domain Adaption Method. *Remote Sens.* **2023**, *15*, 3326. <https://doi.org/10.3390/rs15133326>

Academic Editor: Mohammad Awrangjeb

Received: 4 May 2023

Revised: 13 June 2023

Accepted: 23 June 2023

Published: 29 June 2023



Copyright: © 2023 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/>).

Keywords: deep learning; domain adaption; synthetic aperture radar; oriented object detection; Generative Adversarial Networks

1. Introduction

Synthetic Aperture Radar (SAR) has an all-weather, all-time, cloud-free imaging mechanism and has been explored prosperously in both academic and application affairs. The development of high-resolution satellite technology stimulates considerable research on ship target detection in SAR images. Some traditional methods focus on ship target recognition via image processing techniques. The rapid development of artificial intelligence has triggered a great amount of research on deep-learning algorithms for ship detection tasks using SAR images. Furthermore, objects in aerial images have the characteristic of arbitrary orientations, which stimulate the progression of academic research in oriented object detection. Therefore, an increasing amount of SAR ship detection research has gradually employed oriented-object-detection algorithms based on deep learning because of their credible and accurate detection performance.

Even though the number of studies of object detection in SAR images has grown extensively in recent years, some obvious drawbacks still exist compared with the detection performance of optical aerial images on account of the SAR special imaging mechanism and image characteristics. In strong interference and high sea environments of a specific space–time range, there are a large number of high-frequency, strong-power, multiple, and

complex noise interferences from radio equipment, which lead to difficult communication conditions, signal blocking, and poor SAR image quality. Objects in low-quality SAR images are fuzzy and hidden from view by multiple interferences and are thus laborious to recognize using deep-learning models and human perception. Several examples of comparisons between high-quality SAR images and low-quality SAR images are presented in Figure 1, including random background noise, strip noise, and azimuth defocus.

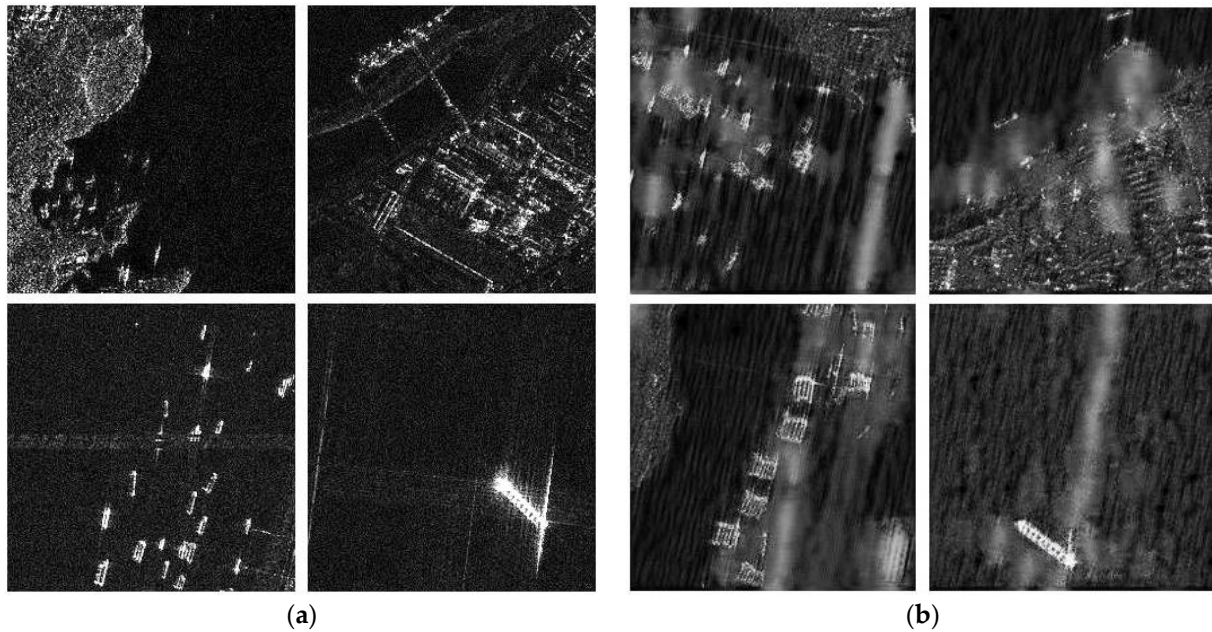


Figure 1. Examples of high-quality SAR images and low-quality SAR images. (a) High-quality SAR images; (b) Low-quality SAR images with various interferences.

Considering the relatively rare quantity of low-quality SAR images and the difficulties of labeling objects manually, training an object detection model of a fully supervised deep neural network using only low-quality SAR images is inefficient and inauthentic in practical applications. Therefore, most SAR ship detection tasks are implemented on public SAR ship datasets that contain adequate high-quality images with annotations when adopting deep neural networks, a method that has achieved tremendous success. Nevertheless, coming up with complicated imaging conditions and environments, the detectors will struggle with the noteworthy disparity of feature distributions between normal SAR images and low-quality SAR images in strong interference environments, given that the objects in low-quality SAR images with low contrast and random noise are indistinct and blur at the boundary. Training with high-quality SAR images, deep-learning models will have difficulty identifying these ambiguous targets in low-quality SAR images, and a high proportion of objects will be omitted by detectors and hurt the model performance.

To tackle the dilemma of strong interference during the SAR object detection task, two opposite solutions are deeply explored and confirmed to be reasonable via extensive research. One approach is SAR image denoising and despeckling based on both traditional methods and deep-learning methods as data-preprocessing techniques. As traditional methods, SAR denoising and despeckling focus on frequency filtering and geometric transformation, such as Lee filtering [1], Kuan Filtering [2], PPB [3], and SAR-BM3D [4] algorithms. In addition, numerous strategies based on deep learning have achieved excellent performance in image denoising. Jain [5] utilizes a CNN as well as SDA [6] and a Sparse de-noising self-encoder [7] to realize natural image denoising. Much research [8–10] has achieved SAR image despeckling via the CNN. However, these preprocessing methods are still risky for the downstream task since they change the object's appearance, and they are also time-consuming during inference processes. On the contrary, an alternative

approach is to generate artificial SAR images with a corresponding noise style as a data augmentation strategy to improve the model's robustness and performance. In analyzing the imaging mechanism and environment characteristics, diverse traditional noise simulation and image-processing methods are involved, including image saliency enhancement [11], saturation adjustment [12], and gamma noise simulation [13]. However, these mentioned methods are only suitable for single and specific interference and tend to be ineffective in the typical low-quality SAR object detection task considering the complicated interferences and ineffective background.

Training with high-quality SAR images with annotations and encountering low-quality SAR images with strong interferences in the inference process can be regarded as a specific domain shift problem, so exploring the domain adaption method is a reasonable and promising solution. Therefore, an unsupervised domain adaption method is proposed in this paper to surmount the detection performance deterioration during the low-quality SAR ship detection task. First, an image-to-image translation algorithm based on Generative Adversarial Networks (GANs) is explored to convert high-quality SAR images from the source domain to low-quality SAR images in the target domain. The cycle-consistency loss in CycleGAN [14] keeps the localization and appearance of ship targets invariant, which makes the annotation inheritance possible for the downstream task when attempting object detection. Second, the generated artificial low-quality SAR images with original labels from high-quality images are added to the training dataset, enriching the samples in the model training process and enhancing the capacity of detectors. Treating the proposed method as an efficacious data augmentation strategy, artificial large-scale nearshore SAR images are generated rigorously and accurately as training images for object detection models, encouraging significant progress in detecting recall and recognition precision. The main contributions of this paper can be summarized as follows:

1. A low-quality SAR image generation module is designed for generating SAR images with typical interference by utilizing unpaired image-to-image translation via CycleGAN, which is employed as a data augmentation strategy for the ship target detection task to enrich the training samples of detectors;
2. A framework for ship detection using SAR images is proposed to effectively learn target features in low-quality SAR images, thereby enhancing the detection performance of ship targets under the condition of strong interference;
3. Extensive experiments on public SAR datasets show that our method can generate effective SAR images for ship detection tasks while improving the recall and precision of ship targets in low-quality SAR images under strong interference and decreasing classification errors and missing detection.

The subsequent description of this paper is organized in the following order: in Section 2.1, we present some previous works related to the proposed method for object detection and the image generation task; in Section 2.2, the implementation details of our method are demonstrated via theorization; Section 3 introduces comprehensive experiments to evaluate the validity of this unsupervised domain adaption approach for ship target detection in SAR images; in Section 4, the discussion of this work and its future research are presented; the final section, Section 5, presents the conclusion of the whole paper.

2. Materials and Methods

2.1. Related Works

2.1.1. SAR Ship Detection

Ship target detection has been a fundamental and momentous task in SAR perceptual interpretation for many years. Situated at the initial stage, most traditional detection methods mainly consist of three parts: preprocessing, detection, and discrimination. The CFAR [15] detector has triggered a considerable amount of research on the SAR target detection task. The constant false alarm rate detection algorithm leverages the adaptive threshold according to the given false alarm rate and background data distribution to determine

the target region at the pixel level. Furthermore, multiple optimizing strategies [16–19] improve the statistical model and threshold determination based on the original CFAR algorithm for the ship target detection task.

On the other hand, with the rapid development of satellite technologies and artificial intelligence, readily available aerial image resources stimulate the wide utilization and applications of deep-learning algorithms for remote sensing as well as SAR ship detection. Girshick proposes several productive and landmark methods, such as the R-CNN [20], Fast R-CNN [21], and Faster R-CNN [22], all focusing on the object detection task. Feature extraction, region proposal strategies, classification, and regression benchmarks establish the foundations of detection network architectures. The Region Proposal Network (RPN) provides flexible bounding boxes with multiple aspect ratios. The Feature Pyramid Network (FPN) [23] maintains multi-scale, high-level features to enhance the feature extraction efficiency. The You Only Look One (YOLO) [24] series has developed into the most influential set of one-stage detectors in recent years. YOLOv3 [25] adopts the Darknet53 backbone and other higher-version networks to introduce more effective modules to improve the model performance, including Cross-Stage Partial connections (CSP), SPP-Block, and PANet.

Turning now to interdisciplinary approaches using the Synthetic Aperture Radar, typical feature extraction and a fusion module are proposed to optimize the model performance for SAR images. An attention mechanism has been deeply explored in SAR target detection via elaborate networks. The Convolutional Block Attention Module (CBAM) in the feature map has been proven to be beneficial by Cui while using the Dense Attention Pyramid Network (DAPN) [26]. Fu [27] balances the feature pyramid under the guidance of attention to detect small ship targets better in complicated backgrounds. Several data augmentation strategies are proposed to improve the model performance and achieve a significantly reduced annotation cost for classification [28] and semantic segmentation [29] tasks of SAR images. Since the ship object always appears with large aspect ratios and arbitrary orientations, traditional object detection methods, which are only capable of providing horizontal bounding boxes, still exhibit obvious shortcomings. Oriented-object-detection algorithms flourish under the condition of sufficient aerial images and annotations with oriented bounding boxes in remote sensing.

2.1.2. Oriented Object Detection

As a popular topic in computer vision, oriented object detection has been comprehensively researched for both algorithms and applications. Considering the characteristic of dense distribution and arbitrary orientations of objects in remote sensing, an increasing number of algorithms are explored to locate the rotated bounding boxes that strictly embrace the edges of objects on the public datasets of aerial images from satellites, including DOTA [30], HRSC2016 [31], DIOR-R [32], ICDAR2015 [33], and so on. The RoI Transformer [34] equips the RRoI learner and Rotated Position-Sensitive RoI Align, with the former learning the transformation from Horizontal RoIs to Rotated RoIs, and the latter being capable of extracting the rotation-invariant feature from RRoI to fulfill the oriented-object-detection task. Subsequently, a Rotation-equivariant Detector (ReDet) [35] is able to incorporate rotation-equivariant networks into the backbone and use Rotation-invariant RoI Align (RiRoI Align) to extract rotation-invariant features which facilitate the outstanding improvement of detection performance for aerial images with oriented bounding boxes. A gliding vertex [36] presents an effective representation of oriented bounding boxes to alleviate detection error and confusion problems. The Oriented R-CNN [37] leverages an oriented Region Proposal Network (oriented RPN) and Oriented R-CNN head to achieve state-of-the-art detection accuracy on two public datasets with the potential to obtain high-quality oriented proposals and refine the oriented Regions of Interest. Attempting to tackle three challenges of object detection in remote sensing, small objects, cluttered arrangement, and arbitrary orientations, the SCRNet [38] offers a feature fusion framework, a supervised multi-dimensional attention method, and an improved smooth L1 loss.

An efficient and fast single-stage detector named R3Det [39] first overcomes the feature misalignment problem for large aspect ratio object detection and realizes the SkewIoU loss to estimate object orientations more accurately. Circular Smooth Label (CSL) [40] regards the angle estimation of oriented bounding boxes as a classification task instead of regression. Densely Coded Labels (DCLs) [41] utilize a novel coding mechanism to speed up the training process of CSL-based methods and also bring up Angle Distance and Aspect Ratio-Sensitive Weighting (ADARSW) in an effort to enhance square-like object detection. GWD [42] tackles boundary discontinuity and inconsistency using an innovative regression loss derived from the Gaussian Wasserstein distance that converts the rotated bounding boxes to a 2D Gaussian distribution. Similarly, KLD [43] approximates the SkewIoU loss with the specified distribution distance based on Gaussian modeling, and KFIOU [44] is an easy-to-implement Gaussian-based loss helped by the Kalman filter to realize its full differentiability when mimicking the SkewIoU's computing mechanism. In our work, the Oriented R-CNN was explored as the baseline for its distinguished detection performance based on a public SAR dataset, as reported, and our experiences.

2.1.3. Unsupervised Domain Adaptation Based on GANs

Multiple studies have investigated the potential of Generative Adversarial Networks (GANs) when it comes to domain adaption problems. GANs have strong capability for modeling and approximating data distributions of images and manufacturing artificial image samples for specific data distribution. The image-to-image translation task is widely employed in domain transfer for generating images of the target domain. Requiring strictly paired data, Pix2Pix [45] illustrates a conditional generative adversarial network for mapping input to output images at the pixel level and plays a significant role in many classic tasks, such as style transfer. CycleGAN [14] releases the restrictions of expensive cost-paired data and expands the applications on image-to-image translation dramatically by introducing a cycle consistency loss and two mapping frameworks and realizing an image-to-image translation task between two domains with unpaired samples. Consequently, CycleGAN stimulates multitudinous derivation in extensive research areas. Arruda [46] adopts CycleGAN for the cross-domain car detection task by transferring images from light to dark. Pasqualino [47] explores an unsupervised domain adaptation scheme, including CycleGAN, for cultural artwork recognition. A data augmentation method based on CycleGAN is involved in multi-organ detection in CT images by Hammami [48]. Liu [49] also introduces CycleGAN to transfer simulated samples for SAR target classification. More in-depth works upgrade the architecture of CycleGAN according to the characteristics of their downstream tasks. FteGanOd [50] contains a feature translate-enhancement module based on CycleGAN and achieves multi-scale feature fusion for night-time vehicle detection. DE-CycleGAN [51] enhances the color information and sharpness of whole satellite images and also enhances features of the target at the object level when involving weak vehicle detection in aerial images. A physical-model-based CycleGAN [52] combines the physical model of a signal from the satellite sensor with CycleGAN and removes the thin cloud on remote sensing datasets by synthesizing a cloud-free image, thin cloud thickness map, and its thickness coefficient. In our work, CycleGAN is responsible for learning the mapping of typical interference in SAR images and generating SAR images as training samples for downstream tasks.

2.2. The Proposed Method

This section illustrates the proposed method of a data augmentation strategy in an object detection task based on unsupervised domain adaption in the following order: unpaired image-to-image translation, low-quality SAR image generation with annotations, and oriented object detection. At the initial stage, the background characteristic of low-quality SAR images is imitated using generators of CycleGAN via unpaired image-to-image translation. The cycle-consistency loss is adopted to achieve cycle mappings between two data distributions of two domains as follows: the high-quality SAR as the source domain

and the low-quality SAR with strong interferences as the target domain. Therefore, during the second step, the translator keeps the geographic and semantic information of scenes invariant and only simulates the appearances of strong interferences when converting the high-quality SAR images to low-quality SAR images so that the inheritance of original annotations (rotated bounding boxes) from the source domain to target domain is feasible and sensible. During the last process, a state-of-the-art oriented-object-detection algorithm based on deep learning, named the Oriented R-CNN, is explored for low-quality SAR ship detection with multiple traditional data augmentation strategies and the proposed method of generating artificial low-quality SAR images as training samples, which is verified to be beneficial and meaningful for detector performance improvement. Figure 2 shows the whole structure of the proposed method.

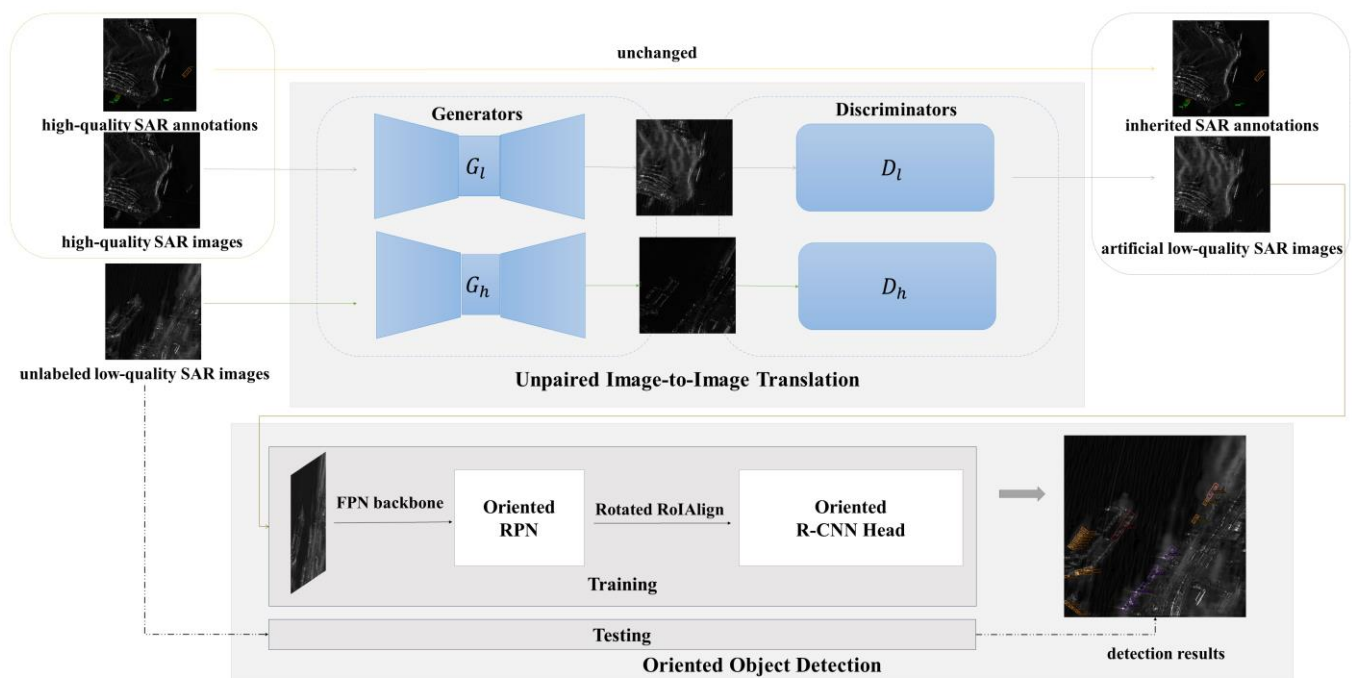


Figure 2. The whole architecture of the proposed method.

2.2.1. Unpaired Image-to-Image translation

The unpaired image-to-image translation realized via CycleGAN with a cycle-consistent loss based on Generative Adversarial Networks expands the potential of image generation for both academic and industrial applications since it releases the restriction of a paired data requirement. In particular, low-quality SAR images in strong interference environments are relatively unavailable and sparse, not to mention the fact that access to the paired high-quality SAR and low-quality SAR images at the identical geography scene is almost impossible to achieve. Relying on the capability of cycle-consistency mappings between two data distributions of CycleGAN, it becomes feasible to achieve the interference emulation of low-quality SAR images and artificial low-quality SAR training samples generation with their corresponding annotations for downstream tasks.

The proposed method attempts to simulate the complicated noises in low-quality SAR images and inject them into high-quality SAR images. Consequently, these SAR images are divided into two domains according to the imaging conditions. We treat high-quality SAR images as source domain X , and low-quality SAR images as target domain Y . A data distribution mapping $G : X \rightarrow Y$ is learned by CycleGAN, with $G(X) = Y$. Furthermore, iterating for the final objective, which combines with a cycle-consistency loss, the generator of CycleGAN enforces the inverse mapping $F : Y \rightarrow X$ to $F(G(X)) \approx X$ so that this image-to-image translation process is capable of maintaining the semantic information of SAR images at the object level.

The CycleGAN framework, depicted in Figure 2, receives two unpaired datasets from separate domains: low-quality SAR and high-quality SAR. Two generators, G_l and G_h , and two discriminators, D_l and D_h , are involved in the training process simultaneously. The generator G_l produces artificial low-quality SAR images with strong interference from high-quality SAR images, and the discriminator D_l distinguishes whether this image is a real sample or a fake one. By the same token, the generator G_h and discriminator D_h are responsible for filtering complex noise in low-quality SAR images and determining the facticity of input samples.

The whole training objective consists of three parts: adversarial losses of mapping $G : X \rightarrow Y$ and its discriminator D_l , adversarial losses of inverse mapping $F : Y \rightarrow X$ and its discriminator D_h , and a cycle-consistency loss between these two mappings.

$$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{GAN}(G, D_Y, X, Y) + \mathcal{L}_{GAN}(F, D_X, Y, X) + \lambda \mathcal{L}_{cyc}(G, F) \quad (1)$$

The adversarial losses of the low-quality SAR generation process can be described as follows:

$$\mathcal{L}_{GAN}(G, D_Y, X, Y) = \mathbb{E}_{y \sim p_{data}(y)} [\log D_Y(y)] + \mathbb{E}_{x \sim p_{data}(x)} [\log(1 - D_Y(G(x)))] \quad (2)$$

When G tries to minimize this objective against an adversary, D prefers to maximize it as follows: $\min_G \max_{D_Y} \mathcal{L}_{GAN}(G, D_Y, X, Y)$. Similarly, the adversarial losses of the inverse mapping follow the same formulation.

Except for the adversarial loss from two pairs of generators and discriminators, as mentioned before, a cycle-consistency loss plays an extremely essential role in the generation framework to guarantee the learned mappings of the domain shift are cycle-consistent. More specifically, the cycle-consistency condition contains forward cycle consistency and backward cycle consistency, which means that these two image translation architectures should ensure bringing every image x or y from its own domain X or Y back to the original image from the intermediate artificial image in the opposite domain Y or X , as $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$ and $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$ demonstrate. The cycle-consistency loss is expressed as follows:

$$\mathcal{L}_{cyc}(G, F) = \mathbb{E}_{x \sim p_{data}(x)} [\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim p_{data}(y)} [\|G(F(y)) - y\|_1] \quad (3)$$

The property of cycle consistency enables CycleGAN to realize the style transfer for images so that some important semantic details of objects remain unchanged, such as the object positions, categories, and terrestrial background. For SAR image generation, the trained CycleGAN distinguishes the main difference between two domains and learns to imitate the strong interferences for high-quality SAR images. The invariant geographical scenes and the ship targets make the bounding boxes inheritance from the high-quality SAR to the low-quality SAR practicable.

2.2.2. Oriented Object Detection

Oriented-object-detection algorithms have been extensively explored in SAR object detection tasks, especially for SAR ship target detection, considering the arbitrary orientations and extreme aspect ratios of densely distributed ship targets in SAR images. The Oriented R-CNN [37], chosen as the baseline of our proposed method, is a state-of-the-art two-stage rotated object detection algorithm that is widely implemented in aerial images for its promising accuracy and efficiency. Equipped with the oriented Region Proposal Network (oriented RPN), the first stage of the Oriented R-CNN produces accurately oriented proposals simply and efficiently. Architected via lightweight fully convolutional networks, the oriented RPN extracts multi-scale features using the five-stage FPN. Three types of horizontal anchors with various aspect ratios are devoted to delivering oriented proposals with a novel representation scheme of oriented bounding boxes $O = (x, y, w, h, \Delta\alpha, \Delta\beta)$ for objects named as a midpoint offset (representation shown in the top of Figure 3) so that the

coordinates $v = (v_1, v_2, v_3, v_4)$ of the four vertices of the oriented bounding boxes can be calculated as follows:

$$\begin{cases} v_1 = (x, y - h/2) + (\Delta\alpha, 0) \\ v_2 = (x + w/2, y) + (0, \Delta\beta) \\ v_3 = (x, y + h/2) + (-\Delta\alpha, 0) \\ v_4 = (x - w/2, y) + (0, -\Delta\beta) \end{cases} \quad (4)$$

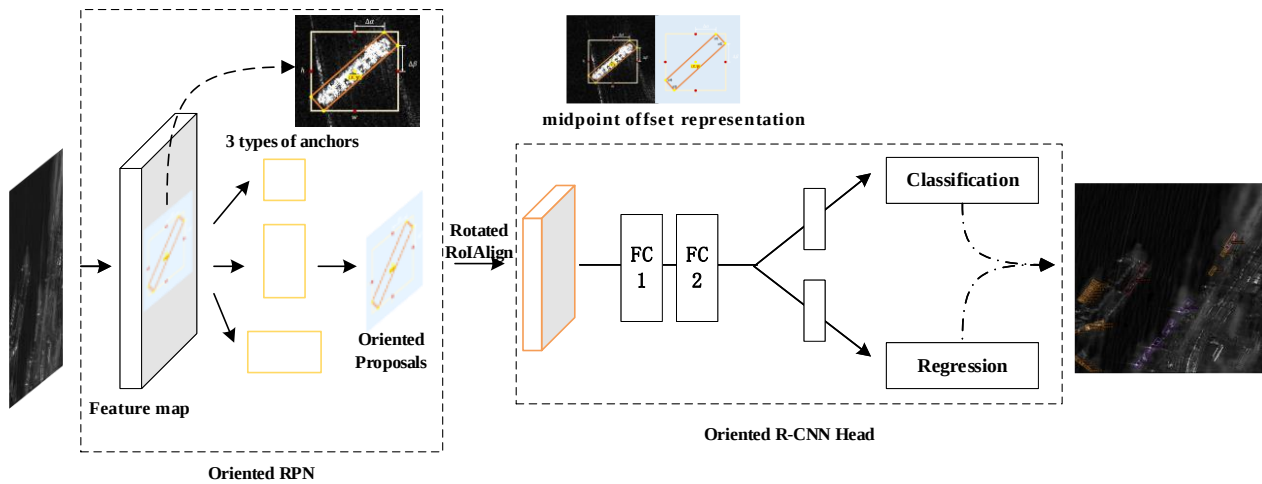


Figure 3. The pipeline of Oriented R-CNN.

In the second stage, an oriented R-CNN head is responsible for refinement and recognition of oriented Regions of Interest (oriented RoI). Fixed-size feature vectors are derived using rotated RoI alignment and sent into two fully connected layers, followed by two parallel fully connected layers for classification and regression tasks, respectively. Figure 3 illustrates the pipeline of Oriented R-CNN.

For SAR ship detection, the phenomena of dense distribution and near-shore scenes increase the difficulties in detecting objects precisely. Although the Oriented R-CNN achieves promising results in oriented object detection because of its capability of extracting image features and sketching the contours of objects with oriented bounding boxes, when encountering a domain shift problem as distinguished data distributions of training and testing images, it will come up against a brick wall, just as with other standard oriented-object-detection algorithms. For instance, missing detection and category confusion problems will occur frequently and damage model performance. Manifold data augmentation strategies have been proposed to improve the accuracy of detectors, but such domain shift problems continue to exist if only adopting simple data augmentation methods such as random crop, random flip, and some unsophisticated noise simulations, leading to limited performance improvements in our experiments for low-quality SAR ship detection. As a result, the proposed method is designed elaborately as a typical data augmentation strategy for the object detection task to improve the robustness of detectors in strong SAR interference environments based on CycleGAN by simulating global interference and imaging characteristics for generated low-quality SAR images and expanding training samples for detection models. Exhaustive experiments demonstrate that the proposed methods are efficient and credible, considering the detecting performance significantly increases and the missing rate decreases simultaneously in the SAR ship detection task.

3. Results

This section illustrates the methodology and datasets involved in the experiments of our proposed method. Comprehensive experiments are included in this section to analyze the model's efficiency and potential. Both the generation model and object detection algorithm are evaluated using standard metrics [53,54], such as Inception Score (IS), Frechet Inception Distance (FID), recall, precision, and mean Average Precision (mAP). Various

data conditions for image-to-image translation model training and multiple data augmentation strategies are investigated in our experiments and demonstrate the necessity and significance of the proposed method.

3.1. Datasets

3.1.1. GaoFen-3 Ship Dataset

Thirty scenes of estuaries and rivers from the GaoFen-3 (GF-3) satellite are involved in this dataset, demonstrating an azimuth resolution of 1.124 m and slant range resolution of 1.700–1.754 m. The original SAR images are preprocessed to uniform slices with a size of 1024×1024 , a size which includes ship targets. The training set contains a total of 2555 images and 7747 ship targets from 20 scenes, and the remaining 10 scenes are sliced to form the test dataset of 1037 images with 2680 ship targets. Both near-shore and off-shore scenes are explored in the experiments of SAR image generation and oriented object detection. In addition, the low-quality SAR images in the test dataset appear with strong interference, while the training dataset is derived from a normal high-quality imaging environment.

3.1.2. SRSDD-V1.0

SRSDD-V1.0 [55] is a high-resolution SAR rotation ship detection dataset. All data in this dataset are from the GF-3 Spotlight (SL) mode with a 1 m resolution, with each image having 1024×1024 pixels. The data of inshore scenes occupy a proportion of 63.1%, with complex backgrounds and much interference, making detection more challenging. The dataset contains multiple categories: a total of six categories of 2884 ships that are all annotated by rotatable bounding boxes. It contains a total of 666 images, including 532 training images and 134 test images. Similarly, we introduce typical noise and interference into the testing samples to explore the efficiency of the proposed method when it comes to the domain adaption problem shown in Figure 4.

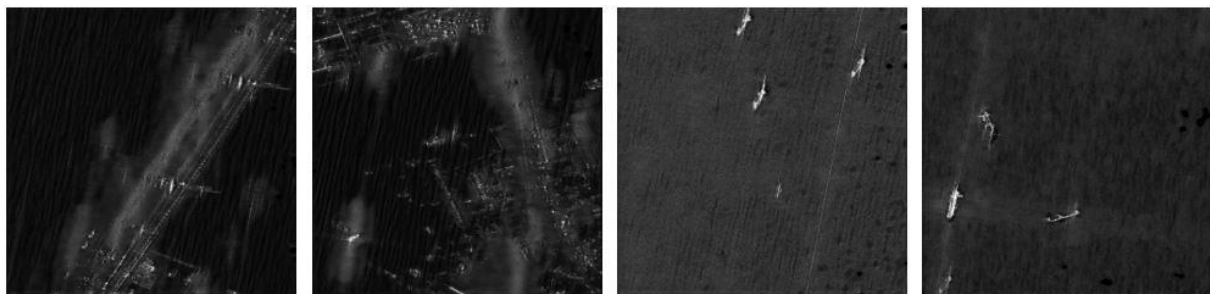


Figure 4. Examples of low-quality SAR images in the SRSDD-V1.0 test dataset.

3.2. Evaluation Criteria

3.2.1. Evaluation Metrics for GANs

- Inception Score (IS)

An Inception model is employed to evaluate the generated images' conditional label distribution $p(y|x)$, and a higher Inception Score (IS) indicates the better quality and diversity of the samples. Assuming images with reasonable semantic information bring about a conditional label distribution $p(y|x)$ with low entropy and rich diversity is accompanied by high entropy on the data distribution marginal, then it is expressed as follows:

$$\int p(y|x = G(z))dz \quad (5)$$

The IS for estimating the generation quality and diversity is set as follows:

$$\exp(E_x \text{KL}(p(y|x) \parallel p(y))) \quad (6)$$

However, the IS demonstrates some disadvantages when meeting model collapse, overfitting, and the specific generation task. For instance, in our method of generating low-quality SAR images with strong interference, the ideal samples tend to have a lower IS with a complicated imaging background. As a result, another standard evaluation metric, the Frechet Inception Distance (FID), is adopted as the main evaluation criterion for GANs in this paper.

- Frechet Inception Distance (FID)

The FID represents the similarity between the real data and generated data distributions using the Frechet distance between the two Gaussian distributions and is computed by the following equation:

$$\text{FID}(p_{data}, p_g) = \| \mu_r - \mu_g \| + \text{tr} \left(C_r + C_g - 2(C_r C_g)^{\frac{1}{2}} \right) \quad (7)$$

Here, φ is the Inception network's convolutional feature functions, and $\varphi(p_{data})$ and $\varphi(p_g)$ are the Gaussian random variables with empirical means μ_r ; μ_g and empirical covariance C_r ; C_g .

3.2.2. Evaluation Metrics for Oriented Object Detection

- Precision P and Recall R

The precision P and recall R are defined by the following equations:

$$P = \frac{TP}{TP + FP} \quad (8)$$

$$R = \frac{TP}{TP + FN} \quad (9)$$

These are the basic metrics in the object detection model, with TP , FP , and FN referring to True Positive (TP), False Positive (FP), and False Negative (FN), respectively, and they measure the accuracy of the detectors. As a supplement, the oriented Intersection Over Union (IOU) measures the overlapping area between the predicted bounding box B_p and the ground-truth bounding box B_{gt} divided by the area of union between them, as shown by the following equation:

$$\text{IOU} = \frac{B_p \cap B_{gt}}{B_p \cup B_{gt}} \quad (10)$$

When the IOU is larger than the given threshold t , the predicted bounding box is treated as a correct detection result.

- Average precision AP and mean AP (mAP)

As the most common evaluation metric in object detection, the measurement method of the average precision AP and the mean AP (mAP) is provided by the PASCAL VOC challenge over all object classes. AP represents the comprehensive performance of the detector since a trade-off exists between precision and recall under different IOU thresholds and can be visualized via the precision-recall curve $P(R)$ and is calculated as follows:

$$AP = \int_0^1 P(R) dR \quad (11)$$

Hence, mAP is the overall average AP of total N object classes, as shown in the following equation:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (12)$$

3.3. Image Generation Experiments

In the first stage of our experiments, the image-to-image translation task aims to imitate the characteristics of SAR interferences and only trains the generators with images from two different domains so that no annotations are required. Selecting from the whole original dataset, we build two domain datasets for CycleGAN: the high-quality SAR as the source domain and the low-quality SAR with strong interference as the target domain. For more elaborate analyses, near-shore SAR images that contain lands and off-shore SAR images with only ship targets in the ocean are separated for subsequent experiments to determine the data distribution modeling ability of CycleGAN on SAR images under different scenes.

Due to the high computing resource consumption of Generative Adversarial Networks, the SAR images are resized to 800×800 when training, while the size is maintained at 1024×1024 in the test phases in order to retain the original semantic information as much as possible for the downstream object detection task.

Inference on the architecture of CycleGAN and artificial low-quality SAR images are produced and inherit the bounding boxes annotations, respectively, from the high-quality SAR dataset. The quality of the image generation model is evaluated using GAN metrics such as FID and IS. From observations and trials, SAR images containing both land and ocean have more rich geographic semantic information, and the SAR images only displaying several ship objects with only ocean background information lack enough interference information for CycleGAN networks. Training with both near-shore and off-shore SAR images tends to be more effortless in developing a cycle-consistent generation model of CycleGAN, which leads to competitive generation performance. As a supplement, the unpaired SAR datasets consist of approximately 342, 522, and 864 SAR images for the near-shore, off-shore, and both scenes, for each domain, respectively.

As the copy and crop strategies are discarded in order to study the global features of the SAR images, CycleGAN is trained for a total of 80,000 iterations with batch size 1 on an NVIDIA GeForce RTX 3090 GPU. Adam optimizers are set for both generations and discriminators, and the learning rate is 0.0002. The evaluation of the FID and IS metrics takes place every 2000th iteration, with the best IS (highest) being 3.4309 at the 8000th iteration for near-shore scenes, and the best FID (lowest) being 49.7250 at the 40,000th iteration when training with both near-shore and off-shore SAR images. Table 1 shows all the quantified evaluation results of our experiments in separate subdatasets. Consequently, the generation model with the best FID is chosen to be responsible for producing low-quality SAR images as extended training samples from the high-quality SAR dataset, considering the IS metric is comparatively inappropriate for evaluating the quality of a generated low-quality SAR image in this task. Table 1 shows the evaluation results of the SAR image generation task.

Table 1. Evaluation results of the SAR image generation task. ‘↓’ means the lower the result value, the better; ‘↑’ means the higher the result value, the better. Bold results represent the optimal values.

Dataset	FID ↓	IS ↑
Off-shore	63.9560	2.4928
Near-shore	57.0183	3.4309
All	49.7250	3.1451

The training process of all near-shore and off-shore SAR image generation tasks is visualized in Figure 5 every 12,000 iterations, from the 4000th iteration to the 40,000th iteration, and the samples at the end of the iterations for visualization are random.

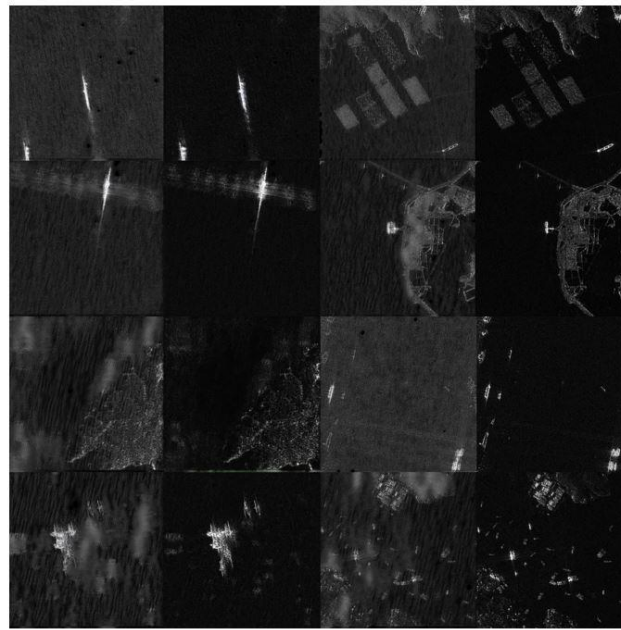


Figure 5. The visualization of the training process for the SAR image generation task based on CycleGAN. The rows from top to bottom: every 12,000th iteration ranging between 4000 and 40,000 iterations. The columns from left to right: original low-quality SAR, generated high-quality SAR from low-quality SAR, generated low-quality SAR from original high-quality SAR, and its original SAR image.

3.4. Oriented-Object-Detection Experiments

All these experiments were conducted on two NVIDIA GeForce RTX 3090 GPUs for a total of 36 epochs on the GaoFen-3 ship dataset and 120 epochs on the SRSDD-V1.0 dataset with a batch size of 4. The learning rate was 0.005, and the momentum was 0.9 with a weight decay of 0.0001 on an SGD optimizer. The Oriented R-CNN worked as the baseline for our experiments on SAR ship detection with rotated bounding boxes. In addition, random flipping and random cropping were regarded as standard data augmentation strategies. More training epochs were applied on the SRSDD-V1.0 dataset for enough iterations with fewer training samples when keeping the batch size constant.

3.4.1. Experiments on GaoFen-3 Ship Dataset

In training on the original SAR dataset, the high-quality SAR images dominate the direction of convergence in the model training process, and most of the ship targets in low-quality SAR images will be ignored by detectors in the inference phase. To determine the efficiency and significance of our proposed method, multiple data augmentation strategies and noise simulation methods based on both normal distribution and GANs were explored in comprehensive experiments.

On the one hand, several traditional data augmentation methods, including random flipping, random rotation, and Hue Saturation Value (HSV) augmentation, were utilized to improve the detection performance and robustness of the model. The results reveal that common data augmentation strategies are capable of improving detection accuracy to some extent but tend to be ineffective when it comes to the domain shift problem. However, after adding the artificial low-quality SAR image samples obtained from CycleGAN into the detector training process, the detection performance improves significantly due to the shrinking gap between the two domains. On the other hand, a common noise simulation method of gamma distribution noise was compared with the proposed method, and the evaluation results demonstrate that our method has distinguished performance considering the comprehensive interference generation abilities relying on CycleGAN. Furthermore, multiple-oriented-object-detection algorithms were investigated in the ex-

panded experiments to confirm the practicality and robustness of the proposed method as an unsupervised domain adaption method for the ship target detection task in SAR images. The Oriented R-CNN was selected as the final detection algorithm in this downstream task, establishing a trade-off between the model performance, time and memory consumption of the detector training process, and model stability. The evaluation results using recall, precision, and mAP metrics illustrate that other general augmentation strategies achieve feeble improvements in low-quality SAR ship detection, while our proposed method emerges as an effective approach that is specially designed for such typical imaging characteristics. Figure 6 presents some cases of the reductions in missing detection, and Table 2 shows the detection results of our experiments.

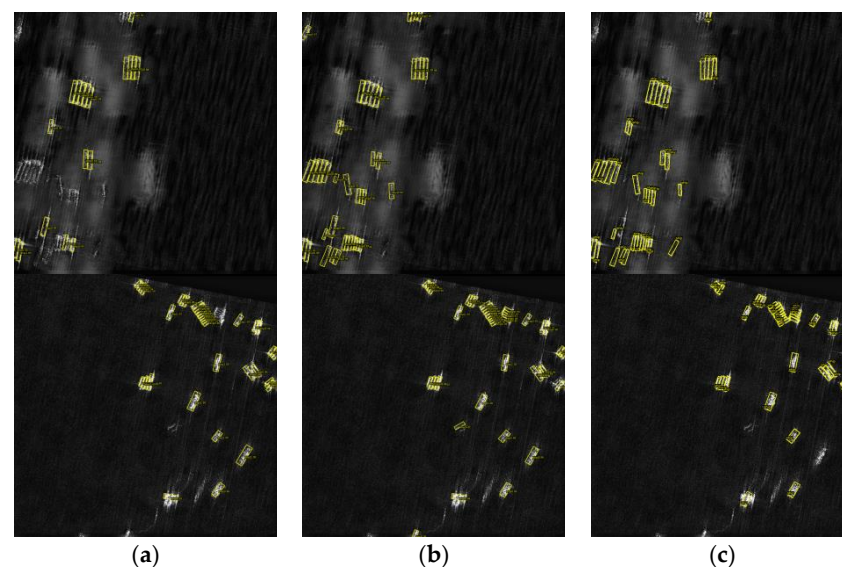


Figure 6. Reduction in missing detection in GaoFen-3 ship dataset at the (a) baseline, (b) proposed method, and (c) ground truth.

Considering there is a trade-off between precision and recall, especially for detection tasks with only one category, the quantitative evaluation results in Table 2 are not reliable enough to evaluate the detection performance of the proposed method, as a higher recall may bring lower precision spontaneously. Therefore, the visualized curve of precision and recall is displayed in Figure 7, which reveals that our methods ($R + F + H + C$ and $R + F + H + G + C$) outperform baseline and other data augmentation strategies.

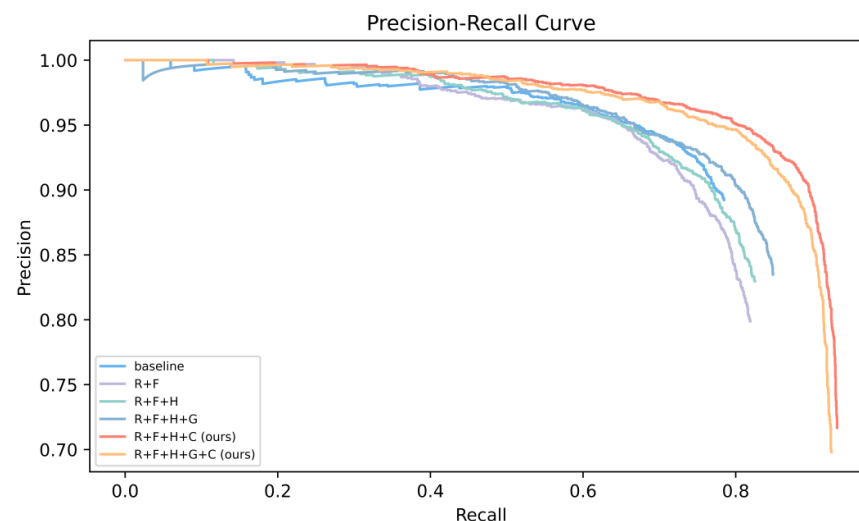


Figure 7. The visualized precision–recall curves of different methods based on the Oriented R-CNN.

Table 2. Evaluation results of recall, precision, and mAP on GaoFen-3 ship dataset. ‘R’, ‘F’, ‘H’, ‘G’, and ‘C’ indicate random rotation, random flipping, HSV augmentation, gamma noise simulation, and the proposed method based on CycleGAN, respectively. Bold results represent the optimal values in each method.

Method	Data Aug	Recall	Precision	mAP
Retinanet [56]	-	0.804	0.698	0.782
	R + F	0.851	0.568	0.808
	R + F + H	0.855	0.53	0.809
	R + F + H + G	0.874	0.596	0.838
	R + F + H + C (ours)	0.905	0.419	0.869
	R + F + H + G + C (ours)	0.896	0.388	0.850
Roi transformer	-	0.741	0.932	0.728
	R + F	0.762	0.845	0.739
	R + F + H	0.790	0.846	0.765
	R + F + H + G	0.819	0.873	0.802
	R + F + H + C (ours)	0.896	0.732	0.874
	R + F + H + G + C (ours)	0.896	0.703	0.871
Gliding vertex	-	0.737	0.922	0.721
	R + F	0.744	0.668	0.716
	R + F + H	0.784	0.698	0.758
	R + F + H + G	0.787	0.855	0.768
	R + F + H + C (ours)	0.876	0.727	0.851
	R + F + H + G + C (ours)	0.886	0.713	0.864
R3det	-	0.822	0.754	0.784
	R + F	0.868	0.499	0.821
	R + F + H	0.879	0.460	0.834
	R + F + H + G	0.866	0.571	0.827
	R + F + H + C (ours)	0.906	0.445	0.866
	R + F + H + G + C (ours)	0.906	0.424	0.864
Oriented reppoints [57]	-	0.868	0.422	0.805
	R + F	0.906	0.196	0.795
	R + F + H	0.919	0.268	0.829
	R + F + H + G	0.921	0.351	0.837
	R + F + H + C (ours)	0.956	0.282	0.895
	R + F + H + G + C (ours)	0.955	0.223	0.892
Redet	-	0.840	0.846	0.819
	R + F	0.881	0.80	0.859
	R + F + H	0.87	0.82	0.852
	R + F + H + G	0.854	0.873	0.837
	R + F + H + C (ours)	0.922	0.797	0.907
	R + F + H + G + C (ours)	0.933	0.753	0.915
Oriented R-CNN	-	0.785	0.893	0.766
	R + F	0.818	0.798	0.791
	R + F + H	0.825	0.829	0.799
	R + F + H + G	0.849	0.835	0.825
	R + F + H + C (ours)	0.932	0.716	0.909
	R + F + H + G + C (ours)	0.925	0.698	0.900

To ascertain the extensive applicability of our proposed approach, we perform supplementary experiments utilizing horizontal object detection algorithms to assess the efficacy of our approach as an efficacious data augmentation strategy, including YOLO series as one-stage networks and Mask R-CNN as two-stage networks. The details of the evaluation are displayed in Table 3.

Table 3. Evaluation results of horizontal object detection algorithms on the GaoFen-3 ship dataset. ‘R’, ‘F’, ‘H’, and ‘C’ indicate random rotation, random flipping, HSV augmentation, and the proposed method based on CycleGAN, respectively. Bold results represent the optimal values in each method.

Horizontal Object Detection Algorithms	Data Aug	Recall	Precision	mAP
Mask R-CNN [58]	F	0.726	0.881	0.695
	F + C (ours)	0.847	0.842	0.817
Mask R-CNN + ASPP + Attention	F	0.732	0.889	0.704
	F + C (ours)	0.854	0.863	0.826
Yolov5 [59]	R + F + H	0.792	0.393	0.697
	R + F + H + C (ours)	0.899	0.405	0.830
Yolov8 [60]	R + F + H	0.775	0.408	0.662
	R + F + H + C (ours)	0.903	0.431	0.836

Additionally, to avoid any possible bias caused by the individuality of the training and test dataset, the cross-validation experiments were conducted on the GaoFen-3 ship dataset, adopting the Oriented R-CNN baseline. The evaluation results shown in Table 4 verified that the proposed method improves the model performance in all data divisions, and also outperforms the traditional data augmentation strategies, driving us to reach conclusions that are consistent with those made in Table 2.

Table 4. Evaluation results of the cross-validation experiments with the Oriented R-CNN baseline on the GaoFen-3 ship dataset. ‘R’, ‘F’, ‘H’, and ‘C’ indicate random rotation, random flipping, HSV augmentation, and the proposed method based on CycleGAN, respectively. Division 1 adopts the same training and test dataset as Table 2. Bold results represent the optimal values in each division.

Cross-Validation Data Division	Data Aug	Recall	Precision	mAP
Division 1	-	0.785	0.893	0.766
	R + F + H	0.825	0.829	0.799
	R + F + H + C (ours)	0.932	0.716	0.909
Division 2	-	0.681	0.900	0.663
	R + F + H	0.754	0.844	0.729
	R + F + H + C (ours)	0.875	0.750	0.841
Division 3	-	0.774	0.843	0.747
	R + F + H	0.805	0.797	0.766
	R + F + H + C (ours)	0.913	0.749	0.875

3.4.2. Experiments on SRSDD-V1.0 Dataset

To authenticate the validity of our proposed method for a cross-domain problem encountered with SAR images, we generated strong interference for a public SAR dataset named SRSDD-V1.0, which contains six categories. It was assumed that the well-trained detector trained on normal SAR images would significantly struggle with the testing images with strong interference, and artificial low-quality SAR training samples are beneficial for mitigating such detector deterioration and improving the model’s robustness. The evaluation results for each category based on mAP, recall, and precision are shown in Table 5 and exhaustively prove our speculations. During multiple experiments, random flip and random rotate data augmentation strategies were abandoned due to their negative effects on the SRSDDV1.0 dataset. In addition, we present some visual examples of the classification and detection improvements with the proposed method in Figure 8.

Table 5. The evaluation results of AP, recall, and precision on the SRSDD-V1.0 dataset. ‘R’, ‘F’, ‘H’, ‘G’, and ‘C’ indicate random rotation, random flipping, HSV augmentation, gamma noise simulation, and the proposed method based on CycleGAN, respectively. Bold results represent the optimal values in each sub-table. (a) AP for each category and mean average AP (mAP) on the SRSDD-V1.0 dataset; (b) recall for each category and total recall on the SRSDD-V1.0 dataset; (c) precision for each category and total precision on the SRSDD-V1.0 dataset.

(a)								
Method	Data Aug	Ore-Oil	Cell-Container	Fishing	LawEnforce	Dredger	Container	mAP
Oriented R-CNN	-	0.315	0.507	0.238	0.273	0.668	0.364	0.394
	R + F	0.265	0.453	0.149	0	0.707	0.247	0.304
	H	0.445	0.697	0.334	0.273	0.737	0.474	0.493
	H + G	0.480	0.680	0.241	0.136	0.760	0.533	0.472
	H + C (ours)	0.516	0.725	0.315	0.273	0.772	0.552	0.525
	H + G + C (ours)	0.498	0.698	0.34	0.273	0.717	0.578	0.517
(b)								
Method	Data Aug	Ore-Oil	Cell-Container	Fishing	LawEnforce	Dredger	Container	Recall
Oriented R-CNN	-	0.412	0.773	0.268	0.200	0.783	0.464	0.466
	R + F	0.441	0.818	0.183	0	0.826	0.4	0.416
	H	0.559	0.818	0.329	0.200	0.804	0.596	0.579
	H + G	0.618	0.818	0.268	0.200	0.826	0.636	0.604
	H + C (ours)	0.647	0.818	0.341	0.200	0.804	0.667	0.635
	H + G + C (ours)	0.588	0.727	0.317	0.200	0.783	0.66	0.620
(c)								
Method	Data Aug	Ore-Oil	Cell-Container	Fishing	LawEnforce	Dredger	Container	Precision
Oriented R-CNN	-	0.304	0.175	0.5	0.333	0.581	0.495	0.442
	R + F	0.174	0.136	0.221	0	0.319	0.193	0.395
	H	0.358	0.231	0.659	0.200	0.435	0.589	0.516
	H + G	0.420	0.198	0.629	0.250	0.521	0.567	0.510
	H + C (ours)	0.379	0.346	0.571	0.143	0.493	0.542	0.511
	H + G + C (ours)	0.513	0.41	0.542	0.167	0.493	0.657	0.603

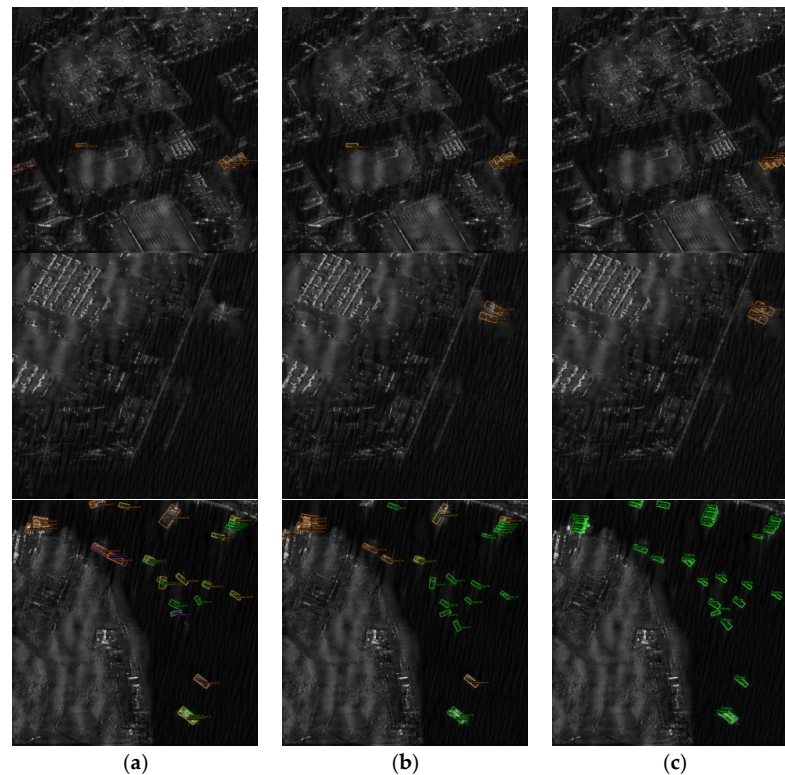


Figure 8. Visualizations of the classification and detection results at the (a) baseline, (b) the proposed method, and (c) ground truth.

4. Discussion

Ship target detection in low-quality SAR images is continuously challenging due to their complicated imaging conditions with strong interference and fewer quantities. To improve the detection performance for this typical condition, it is regarded as a domain shift problem, and unsupervised domain adaption is adopted to shrink the gap between different data domains. Generative Adversarial Networks are capable of simulating the image style of low-quality SAR images using deep-learning networks and generating artificial SAR images with special characteristics for expanding the training samples of detection models. Extensive experiments demonstrate that our proposed method is an efficient data augmentation strategy to improve the model performance and robustness of ship target detection in SAR images. We will further discuss how our work compares with other research in the field of computer vision and explore the influence of the unique characteristics of SAR images on our research.

- Compared with Traditional Methods

Different from the traditional methods of injecting the noise of simple statistical distribution into SAR images, our approach demonstrated superior performance in experiments for several reasons. Firstly, in traditional approaches to image processing, a limited variety of noises will be utilized without rigorous consideration of the statistical distributions of the noises in the real world and specific applications, so these data augmentation strategies relying on the traditional noise simulation make feeble improvements upon our task. However, the proposed method, leveraging style transferring on image generation, can more authentically simulate the sophisticated characteristics of interferences on low-quality SAR images. This not only provides our model with greater robustness when faced with low-quality SAR images, but it also bestows it with a more potent generalization ability. Secondly, compared to other studies that attempt to improve model performance using popular data augmentation methods in computer vision, our method delves deeper into the characteristics of SAR images. SAR possesses a unique imaging mechanism, resulting in typical noise and low quality. Our method not only takes these characteristics into account but also strives to enhance model performance through them. For instance, we harness the features of complex noise in low-quality SAR images and generate low-quality SAR images via CycleGAN, which has a strong ability to learn latent data distribution comprehensively. This allows the model to adapt to special conditions during training, improving its applicability to actual scenarios.

- Generalization Capability

In our discussion, we would also like to follow through with the analysis of the generalization ability of our model. The factors influencing a model's generalization ability include the volume of data and target morphology, among other factors.

On the one hand, the proposed method uses image translation techniques to create a greater volume of samples, improving the model's generalization ability, particularly within low-quality SAR images with complicated interference. The introduction of more data and samples typically implies an increase in the volume of data annotation, while the proposed method achieves image style transfer, keeping the target position consistent, enabling instance-level unsupervised training to generate augmented samples without necessitating additional annotation costs. This factor is vital as annotation, especially for SAR images, can be a tedious and time-consuming task, often requiring specialized knowledge about the target and environment. By adopting our method, more training data are generated without incurring the costs of increased manual labor.

On the other hand, our approach also takes into consideration the various target morphologies that may exist within SAR images. The use of an unsupervised domain adaption method based on image-to-image translation helps to simulate different levels of quality and noise within the images, providing a more comprehensive set of training data that covers a broader range of target appearances. This further enhances the robustness of our model, enabling it to better generalize and detect targets under varying circumstances.

In summary, our approach considers both the generalization ability of the model and the label cost in the practical application process, providing an efficient solution for target detection and identification in low-quality SAR images.

- Future Work

In the first stage of this work, CycleGAN is responsible for modeling a cycle map between two SAR image dataset domains. Both near-shore SAR images and off-shore images are involved in the experiments, and we discover that rich semantic information in various scenes in terms of both near-shore and off-shore data are beneficial for improving SAR image generation performance and training stabilities. Afterward, multiple training processes for the low-quality SAR image generation task were analyzed and explored, and the shortcomings of the Generative Adversarial Networks emerged, such as the model training stability, overfitting, and mode collapse. Especially for the SAR image generation task, the special physical imaging mechanism, diverse target characteristics, and multiple interferences increase the difficulty of the generator training based on the GANs. It is assumed that some well-designed modules based on the combinations of physical imaging theories and deep learning, which are not deeply involved and investigated in our works, will achieve outstanding improvements in the SAR image generation task under designated scenes.

According to the final stage of the object detection modules, our method was treated as a data augmentation strategy to simulate a typical image style and interference characteristics and enrich the training samples for narrowing the domain gap. Comprehensive experiments verify that the proposed method outperforms other traditional data augmentation strategies and works effectively in multiple-oriented-object-detection algorithms. On the other hand, an SAR image generation model based on CycleGAN is also available to convert a low-quality SAR image to a high-quality SAR image with denoising and despeckling abilities, so it is possible for it to be implemented as another data preprocessing method for detection tasks in the future. Considering the inference time consumption of the detection processes, we decided not to undergo the expansion of research and experiments for such preprocessing tricks in this study.

5. Conclusions

The proposed method improves the robustness and detection performance of the SAR ship detection model by implementing an unsupervised domain adaption image-to-image translation task based on Generative Adversarial Networks with cycle-consistency loss regarding the ship target detection task in low-quality SAR images with strong interference as a domain shift problem. Artificial low-quality SAR training samples that have a compatible imaging style with a strongly interfering environment are produced via CycleGAN, which is capable of modeling the data distributions on different domains. The mAP performance of the oriented-object-detection models significantly improved for the GaoFen-3 ship dataset, from 79.9% to 90.9% and from 49.3% to 52.5% for the SRSDD-V1.0 dataset; other evaluation metrics, such as recall and precision, also demonstrate the efficiency of the proposed method via the results of comprehensive experiments with multiple-oriented and horizontal-object-detection algorithms using one-stage and two-stage detectors. On the other hand, the obvious problems of a high missing rate and classification error for the domain shift problem were dramatically ameliorated, given the artificial training samples provide meaningful semantic information for the detection algorithm. In addition, this work investigated the capacity and potential of SAR image generation and style transferring tasks based on Generative Adversarial Networks for near-shore and off-shore scenes.

Author Contributions: Conceptualization, X.P. and H.W.; methodology, X.P.; software, X.P.; validation, X.P.; formal analysis, X.P.; investigation, X.P.; resources, X.P. and H.J.; data curation, X.P., H.J. and Y.X.; writing—original draft preparation, X.P.; writing—review and editing, H.J. and H.W.; visualization, X.P.; supervision, H.W. and F.W.; funding acquisition, H.W. and F.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China (Grant No. 62271153 and 61901122) and the Natural Science Foundation of Shanghai (Grant No. 22ZR1406700 and 20ZR1406300).

Data Availability Statement: The publicly available dataset SRSDD-V1.0 can be accessed at <https://github.com/HeuristicLU/SRSDD-V1.0>. For other experimental data, please contact Haipeng Wang (hpwang@fudan.edu.cn) for access.

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

References

1. Lee, J.-S. Speckle Analysis and Smoothing of Synthetic Aperture Radar Images. *Comput. Graph. Image Process.* **1981**, *17*, 24–32. [CrossRef]
2. Kuan, D.T.; Sawchuk, A.A.; Strand, T.C.; Chavel, P. Adaptive Noise Smoothing Filter for Images with Signal-Dependent Noise. *IEEE Trans. Pattern Anal. Mach. Intell.* **1985**, *PAMI-7*, 165–177. [CrossRef]
3. Deledalle, C.-A.; Denis, L.; Tupin, F. Iterative Weighted Maximum Likelihood Denoising With Probabilistic Patch-Based Weights. *IEEE Trans. Image Process.* **2009**, *18*, 2661–2672. [CrossRef]
4. Parrilli, S.; Poderico, M.; Angelino, C.V.; Verdoliva, L. A Nonlocal SAR Image Denoising Algorithm Based on LLMMSE Wavelet Shrinkage. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 606–616. [CrossRef]
5. Jain, V.; Seung, H.S. Natural Image Denoising with Convolutional Networks. In Proceedings of the 21st International Conference on Neural Information Processing Systems, Red Hook, NY, USA, 8–10 December 2008; pp. 769–776.
6. Vincent, P.; Larochelle, H.; Lajoie, I.; Bengio, Y.; Manzagol, P.-A. Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. *J. Mach. Learn. Res.* **2010**, *11*, 3371–3408.
7. Zhang, K.; Zuo, W.; Chen, Y.; Meng, D.; Zhang, L. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Trans. Image Process.* **2017**, *26*, 3142–3155. [CrossRef]
8. Chierchia, G.; Cozzolino, D.; Poggi, G.; Verdoliva, L. SAR Image Despeckling through Convolutional Neural Networks. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 5438–5441.
9. Wang, P.; Zhang, H.; Patel, V.M. SAR Image Despeckling Using a Convolutional Neural Network. *IEEE Signal Process. Lett.* **2017**, *24*, 1763–1767. [CrossRef]
10. Zhang, Q.; Yuan, Q.; Li, J.; Yang, Z.; Ma, X. Learning a Dilated Residual Network for SAR Image Despeckling. *Remote Sens.* **2018**, *10*, 196. [CrossRef]
11. Zhai, Y.; Shah, M. Visual Attention Detection in Video Sequences Using Spatiotemporal Cues. In Proceedings of the 14th ACM International Conference on Multimedia, Santa Barbara, CA, USA, 23–27 October 2006; pp. 815–824.
12. Yolov5. Available online: <https://github.com/ultralytics/yolov5> (accessed on 25 April 2023).
13. Bishnoi, S.; Sarkar, P.S.; Thomas, R.G.; Patel, T.; Pal, M.; Adhikari, P.S.; Sinha, A.; Saxena, A.; Gadkari, S.C. Preliminary Experimentation of Fast Neutron Radiography with D-T Neutron Generator at BARC. *J. Nondestruct. Eval.* **2019**, *38*, 13. [CrossRef]
14. Zhu, J.-Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2242–2251.
15. Gao, G.; Shi, G. CFAR Ship Detection in Nonhomogeneous Sea Clutter Using Polarimetric SAR Data Based on the Notch Filter. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4811–4824. [CrossRef]
16. Zhai, L.; Li, Y.; Su, Y. Inshore Ship Detection via Saliency and Context Information in High-Resolution SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1870–1874. [CrossRef]
17. Pappas, O.; Achim, A.; Bull, D. Superpixel-Level CFAR Detectors for Ship Detection in SAR Imagery. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1397–1401. [CrossRef]
18. Li, T.; Liu, Z.; Xie, R.; Ran, L. An Improved Superpixel-Level CFAR Detection Method for Ship Targets in High-Resolution SAR Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 184–194. [CrossRef]
19. Li, M.-D.; Cui, X.-C.; Chen, S.-W. Adaptive Superpixel-Level CFAR Detector for SAR Inshore Dense Ship Detection. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 4010405. [CrossRef]
20. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
21. Girshick, R. Fast R-CNN. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
22. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [CrossRef]
23. Lin, T.-Y.; Dollar, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature Pyramid Networks for Object Detection. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 936–944.

24. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
25. Redmon, J.; Farhadi, A. YoloV3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767.
26. Cui, Z.; Li, Q.; Cao, Z.; Liu, N. Dense Attention Pyramid Networks for Multi-Scale Ship Detection in SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 8983–8997. [[CrossRef](#)]
27. Fu, J.; Sun, X.; Wang, Z.; Fu, K. An Anchor-Free Method Based on Feature Balancing and Refinement Network for Multiscale Ship Detection in SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 1331–1344. [[CrossRef](#)]
28. Bi, H.; Xu, F.; Wei, Z.; Xue, Y.; Xu, Z. An Active Deep Learning Approach for Minimally Supervised PolSAR Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9378–9395. [[CrossRef](#)]
29. Bi, H.; Xu, L.; Cao, X.; Xue, Y.; Xu, Z. Polarimetric SAR Image Semantic Segmentation with 3D Discrete Wavelet Transform and Markov Random Field. *IEEE Trans. Image Process.* **2020**, *29*, 6601–6614. [[CrossRef](#)]
30. Xia, G.-S.; Bai, X.; Ding, J.; Zhu, Z.; Belongie, S.; Luo, J.; Datcu, M.; Pelillo, M.; Zhang, L. DOTA: A Large-Scale Dataset for Object Detection in Aerial Images. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 3974–3983.
31. Liu, Z.; Yuan, L.; Weng, L.; Yang, Y. A High Resolution Optical Satellite Image Dataset for Ship Recognition and Some New Baselines. In Proceedings of the International Conference on Pattern Recognition Applications and Methods, Porto, Portugal, 24–26 February 2017.
32. Cheng, G.; Wang, J.; Li, K.; Xie, X.; Lang, C.; Yao, Y.; Han, J. Anchor-Free Oriented Proposal Generator for Object Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5625411. [[CrossRef](#)]
33. Karatzas, D.; Gomez-Bigorda, L.; Nicolaou, A.; Ghosh, S.; Bagdanov, A.; Iwamura, M.; Matas, J.; Neumann, L.; Chandrasekhar, V.R.; Lu, S.; et al. ICDAR 2015 Competition on Robust Reading. In Proceedings of the 2015 13th International Conference on Document Analysis and Recognition (ICDAR), Tunis, Tunisia, 23–26 August 2015; pp. 1156–1160.
34. Ding, J.; Xue, N.; Long, Y.; Xia, G.-S.; Lu, Q. Learning RoI Transformer for Oriented Object Detection in Aerial Images. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 2844–2853.
35. Han, J.; Ding, J.; Xue, N.; Xia, G.-S. ReDet: A Rotation-Equivariant Detector for Aerial Object Detection. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 2785–2794.
36. Xu, Y.; Fu, M.; Wang, Q.; Wang, Y.; Chen, K.; Xia, G.-S.; Bai, X. Gliding Vertex on the Horizontal Bounding Box for Multi-Oriented Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 1452–1459. [[CrossRef](#)]
37. Xie, X.; Cheng, G.; Wang, J.; Yao, X.; Han, J. Oriented R-CNN for Object Detection. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 11–17 October 2021; pp. 3500–3509.
38. Yang, X.; Yang, J.; Yan, J.; Zhang, Y.; Zhang, T.; Guo, Z.; Sun, X.; Fu, K. SCRDet: Towards More Robust Detection for Small, Cluttered and Rotated Objects. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 8231–8240.
39. Yang, X.; Yan, J.; Feng, Z.; He, T. R3Det: Refined Single-Stage Detector with Feature Refinement for Rotating Object. *AAAI* **2021**, *35*, 3163–3171. [[CrossRef](#)]
40. Yang, X.; Yan, J. Arbitrary-Oriented Object Detection with Circular Smooth Label. In *Computer Vision—ECCV 2020*; Vedaldi, A., Bischof, H., Brox, T., Frahm, J.-M., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 677–694.
41. Yang, X.; Hou, L.; Zhou, Y.; Wang, W.; Yan, J. Dense Label Encoding for Boundary Discontinuity Free Rotation Detection. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 15814–15824.
42. Yang, X.; Yan, J.; Ming, Q.; Wang, W.; Zhang, X.; Tian, Q. Rethinking Rotated Object Detection with Gaussian Wasserstein Distance Loss. In Proceedings of the 38th International Conference on Machine Learning; Meila, M., Zhang, T., Eds.; PMLR: New York, NY, USA, 2021; Volume 139, pp. 11830–11841.
43. Yang, X.; Yang, X.; Yang, J.; Ming, Q.; Wang, W.; Tian, Q.; Yan, J. Learning High-Precision Bounding Box for Rotated Object Detection via Kullback-Leibler Divergence. In *Advances in Neural Information Processing Systems*; Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P.S., Vaughan, J.W., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2021; Volume 34, pp. 18381–18394.
44. Yang, X.; Zhou, Y.; Zhang, G.; Yang, J.; Wang, W.; Yan, J.; Zhang, X.; Tian, Q. The KFIOU Loss for Rotated Object Detection. *arXiv* **2022**, arXiv:2201.12558.
45. Isola, P.; Zhu, J.-Y.; Zhou, T.; Efros, A.A. Image-to-Image Translation with Conditional Adversarial Networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 5967–5976.
46. Arruda, V.F.; Paixao, T.M.; Berriel, R.F.; De Souza, A.F.; Badue, C.; Sebe, N.; Oliveira-Santos, T. Cross-Domain Car Detection Using Unsupervised Image-to-Image Translation: From Day to Night. In Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–8.
47. Pasqualino, G.; Furnari, A.; Signorello, G.; Farinella, G.M. An Unsupervised Domain Adaptation Scheme for Single-Stage Artwork Recognition in Cultural Sites. *Image Vis. Comput.* **2021**, *107*, 104098. [[CrossRef](#)]

48. Hammami, M.; Friboulet, D.; Kechichian, R. Cycle GAN-Based Data Augmentation For Multi-Organ Detection In CT Images Via Yolo. In Proceedings of the 2020 IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, 25–28 October 2020; pp. 390–393.
49. Liu, L.; Pan, Z.; Qiu, X.; Peng, L. SAR Target Classification with CycleGAN Transferred Simulated Samples. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 4411–4414.
50. Shao, X.; Wei, C.; Shen, Y.; Wang, Z. Feature Enhancement Based on CycleGAN for Nighttime Vehicle Detection. *IEEE Access* **2021**, *9*, 849–859. [\[CrossRef\]](#)
51. Gao, P.; Tian, T.; Li, L.; Ma, J.; Tian, J. DE-CycleGAN: An Object Enhancement Network for Weak Vehicle Detection in Satellite Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 3403–3414. [\[CrossRef\]](#)
52. Zi, Y.; Xie, F.; Song, X.; Jiang, Z.; Zhang, H. Thin Cloud Removal for Remote Sensing Images Using a Physical-Model-Based CycleGAN With Unpaired Data. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1004605. [\[CrossRef\]](#)
53. Gui, J.; Sun, Z.; Wen, Y.; Tao, D.; Ye, J. A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications. *IEEE Trans. Knowl. Data Eng.* **2023**, *35*, 3313–3332. [\[CrossRef\]](#)
54. Padilla, R.; Netto, S.L.; Da Silva, E.A.B. A Survey on Performance Metrics for Object-Detection Algorithms. In Proceedings of the 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), Niterói, Brazil, 1–3 July 2020; pp. 237–242.
55. Lei, S.; Lu, D.; Qiu, X.; Ding, C. SRSDD-v1.0: A High-Resolution SAR Rotation Ship Detection Dataset. *Remote Sens.* **2021**, *13*, 5104. [\[CrossRef\]](#)
56. Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollar, P. Focal Loss for Dense Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 318–327. [\[CrossRef\]](#)
57. Li, W.; Chen, Y.; Hu, K.; Zhu, J. Oriented RepPoints for Aerial Object Detection. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 1819–1828.
58. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the 2017 IEEE/CVF International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2961–2969.
59. Zhu, X.; Lyu, S.; Wang, X.; Zhao, Q. TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021.
60. Jocher, G.; Chaurasia, A.; Stoken, A.; Borovec, J.; NanoCode012; Kwon, Y.; Michael, K.; TaoXie; Fang, J.; Imyhxy; et al. *Ultralytics/Yolov5: V7.0—YOLOv5 SOTA Realtime Instance Segmentation*; Zenodo: Geneva, Switzerland, 2022. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.