



Article

SAR Target Recognition Using cGAN-Based SAR-to-Optical Image Translation

Yuchuang Sun ^{1,2} , Wen Jiang ^{1,2}, Jiyao Yang ^{1,2} and Wangzhe Li ^{1,2,*}

¹ National Key Lab of Microwave Imaging Technology, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China; sunyuchuang19@mails.ucas.ac.cn (Y.S.); jiangwen16@mails.ucas.ac.cn (W.J.); yangjiyao16@mails.ucas.ac.cn (J.Y.)

² The School of Electronics, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: wzli@mail.ie.ac.cn

Abstract: Target recognition in synthetic aperture radar (SAR) imagery suffers from speckle noise and geometric distortion brought by the range-based coherent imaging mechanism. A new SAR target recognition system is proposed, using a SAR-to-optical translation network as pre-processing to enhance both automatic and manual target recognition. In the system, SAR images of targets are translated into optical by a modified conditional generative adversarial network (cGAN) whose generator with a symmetric architecture and inhomogeneous convolution kernels is designed to reduce the background clutter and edge blur of the output. After the translation, a typical convolutional neural network (CNN) classifier is exploited to recognize the target types in translated optical images automatically. For training and testing the system, a new multi-view SAR-optical dataset of aircraft targets is created. Evaluations of the translation results based on human vision and image quality assessment (IQA) methods verify the improvement of image interpretability and quality, and translated images obtain higher average accuracy than original SAR data in manual and CNN classification experiments. The good expansibility and robustness of the system shown in extending experiments indicate the promising potential for practical applications of SAR target recognition.

Keywords: synthetic aperture radar (SAR); target recognition; SAR-to-optical image translation; deep learning; conditional generative adversarial network (cGAN)



Citation: Sun, Y.; Jiang, W.; Yang, J.; Li, W. SAR Target Recognition Using cGAN-Based SAR-to-Optical Image Translation. *Remote Sens.* **2022**, *14*, 1793. <https://doi.org/10.3390/rs14081793>

Academic Editors: Mi Wang, Hanwen Yu, Jianlai Chen and Ying Zhu

Received: 14 March 2022

Accepted: 4 April 2022

Published: 8 April 2022

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



Copyright: © 2022 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

Target recognition in synthetic aperture radar (SAR) imagery is widely used in civil and military scenarios because SAR can observe ground targets independent of weather and sunlight illumination. Targets in SAR imagery can be recognized by trained experts or automatic image recognition algorithms. However, it is universally viewed as a challenging task to recognize targets in SAR imagery accurately. Current computer vision methods based on typical optical images do not apply very well to SAR images [1] because SAR images have many effects distinct from optical images. First of all, the method of active coherent detection brings unavoidable speckle noise that is generated by the constructive and destructive interference of the coherent microwaves reflected from many microscopic surfaces in the same resolution cell [2–4]. Secondly, SAR images reflect the physical properties of the scenes in the range-azimuth domain with geometrical distortion and structural loss in the perspective of human vision, which vary significantly from different observation views, target orientations, and wavebands [5–7]. Furthermore, there are random salt-and-pepper noise and Gaussian noise additions to SAR images during the signal and image processing in the digital domain. The aforementioned drawbacks make SAR target recognition methods, both automatic and manual, difficult to obtain good effects, which undoubtedly limits the promotion and application of SAR technologies.

Compared with SAR target recognition, target recognition in optical remote sensing images has received extensive attention and made significant progress benefiting from the higher quality data with more accurate focus, less noise, and more texture [8–10]. Accordingly, the all-day and all-weather characteristics of SAR imaging and the good expression ability of optical imaging can be fully utilized by converting SAR images into optical images with the corresponding structural features preserved. Meanwhile, the mature processing algorithms of optical remote sensing images can be exploited for more accurate recognition and interpretation. On this point, SAR-to-optical image translation is explored to translate SAR images into their optical counterparts by taking advantage of multi-source sensor data. State-of-the-art SAR-to-optical methods are based on conditional generative adversarial networks (cGAN) [11], which is a deep learning method making exciting progress in image-to-image translation [12]. In 2018, M. Schmitt et al. [5] published an open-source SEN1-2 dataset comprising 282,384 pairs of matched SAR and optical image patches, collected from Sentinel-1 and Sentinel-2, respectively. From then on, considerable research has been conducted to improve the architectures of cGAN for more effective SAR-to-optical translation. For instance, [13] synthesized missing or corrupted multi-spectral optical images from the SAR data of the same scene using a classical cGAN. A later work [6] exploited CycleGAN to achieve the translation from different resolution SAR images to optical gray images. In atrous-cGAN [14] architecture, atrous convolutions and atrous spatial pyramid poolings were added to cGAN to enhance fine details of output optical images by exploiting spatial context at multiple scales. With the continuous development of SAR-to-optical technology, encouraging progress has been achieved in cloud removal [13,15], scene recognition [13,16,17], road extraction [6], semantic segmentation [14], and other applications. Similarly, in the target recognition domain, mature optical image-based processing and recognition algorithms can be used to improve recognition accuracy significantly if targets in SAR images are translated into optical expression. However, few studies have applied SAR-to-optical in target recognition, mainly for three reasons: Firstly, the translation effects of existing algorithms are better in mountains, rivers, forests, farmland, and other natural scenes, while man-made scenes like buildings and vehicles are hard to restore [6]. Furthermore, as the SAR data for training are mostly from Sentinel-1 platforms with a resolution of 5 m, existing applications of SAR-to-optical methods have been limited to the translation of large-scale scenes. Thus the new dataset including high-resolution SAR target images is necessary for SAR-to-optical target translation. Nevertheless, available SAR target datasets, such as the Moving and Stationary Target Acquisition Recognition (MSTAR) [18], lack matching optical images. This is because common passive optical imaging results like multi-spectral imagery are too different from SAR imagery in radiometric appearance, especially when reflecting the structure of targets. The former depends on the diffuse reflection of sunlight, and the latter depends on the specular and angular reflection of electromagnetic waves emitted actively [19]. Suitable active optical imaging data can reduce the difficulty of training a SAR-to-optical network greatly.

In consideration of the above points, a novel SAR target recognition system is proposed. By using a SAR-to-optical translation network based on cGAN, target recognition in complicated SAR images is converted to that in the optical images. The combination of SAR-to-optical translation and target recognition is achieved benefiting from the creation of a new multi-view SAR-optical dataset and a modified translation network better suited to SAR target images. The creation of the dataset is shown in Figure 1a: The new dataset, including multi-view matched SAR images and optical images of four types of small planes and helicopters, is named SPH4. An unmanned aerial vehicle (UAV) SAR is used to image ground aircraft targets in HH, HV, and VV polarization modes, and a computer simulation based on a ray-tracing algorithm is used to produce optical imagery with a similar radiometric appearance to SAR imagery. Then, pairs of SAR and optical target images are matched and used to form the SPH4 dataset. The training of the translation and the recognition network is shown in Figure 1b: With SPH4, the SAR-to-optical translation network is trained to generate translated images from SAR images. In the cGAN-based translation net-

work, a modified symmetric U-Net architecture with inhomogeneous convolution kernels is adopted as the generator of cGAN to reduce the background clutter and edge blur of output translated images. With regard to the discriminator, a PatchGAN classifier similar to pix2pix [12] is used to judge the authenticity of the image patches. After the translation, the translated images are labeled and used to train the recognition network, in which a typical convolutional neural network (CNN) classifier referring to LeNet [20] is exploited to recognize the target types. Experimental results verify the effectiveness of the proposed system. The translated images can be more easily recognized by the human eye than SAR images, and the evaluation result based on image quality assessment (IQA) methods show that the translated images have the higher peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) than those in original SAR images. The results of the CNN classifier and the manual recognition demonstrate that the translated images gained higher average accuracy than SAR images in both automatic and manual target recognition. Furthermore, extending robustness experiments show that the system can maintain stability in the case of noise addition and aircraft type extension. The main contributions of this paper can be summarized as the following four aspects:

1. A novel SAR target recognition system is proposed and developed, using SAR-to-optical translation to enhance target recognition for improving the accuracy of recognition.
2. A new approach to creating the matched SAR-optical dataset is presented by simulating optical images corresponding to SAR target images for SAR-to-optical translation, SAR target recognition, and other following research.
3. A modified cGAN network with a new generator architecture is explored, which can be competent for the SAR-to-optical translation of aircraft targets.
4. Experiments of noise addition and aircraft type extension are designed and implemented to demonstrate good robustness and extensibility of the proposed recognition system.

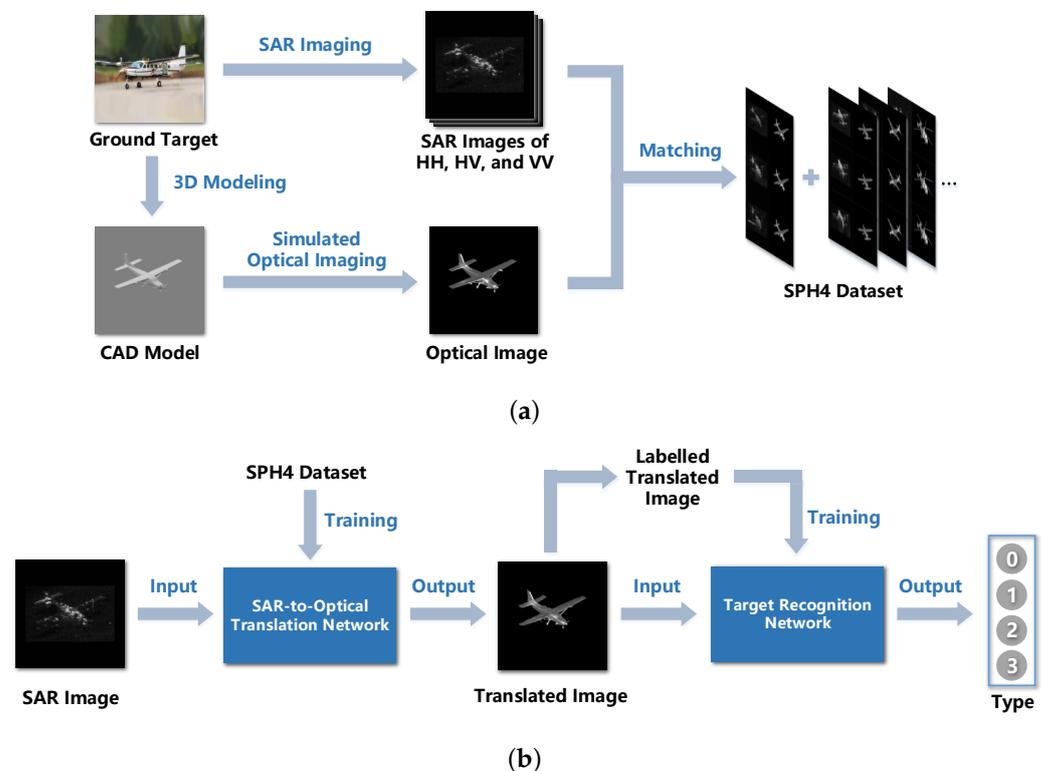


Figure 1. Entire recognition system. (a) Creation of the SPH4 dataset; (b) Training of the translation network and recognition network.

2. Related Works

In this section, we introduce some previous works of SAR automatic target recognition (ATR). ATR and discuss their problems first. Next, the theory of cGAN-based image-to-image translation is given in detail. Lastly, the model-based data generation used in the creation of the dataset is introduced in terms of applications.

2.1. SAR ATR

There is a great difference between SAR imagery characteristics and human visual habits, which brings many difficulties to manual target interpretation. Accordingly, the development of ATR algorithms for SAR targets is necessary [21]. The process of SAR ATR can be summarized into three steps: preprocessing, feature extraction, and feature classification. Preprocessing can make image features clear by reducing noise, improving the resolution, and so on. For instance, Novak et al. [22] enhanced the resolution of SAR images in MSTAR through a super-resolution algorithm and gained higher recognition accuracy than SAR ATR using conventional preprocessing. In [23], the target contours in MSTAR obtained by smooth denoising, semantic segmentation, and edge extraction were easier to recognize for their simple and clear data format. The local gradient ratio pattern histogram of SAR images is extracted in [24], which is proven to reduce the effects of local gradient variation and speckle noise. A recent study [25] processed the target pixel grayscale declines in MSTAR into graph representation and classified them with a graph CNN. Feature extraction collects local target information that is the basis to judge the type of targets. Early studies used intelligent algorithms to extract features, such as principal component analysis (PCA) [26], support vector machine (SVM) [27], and genetic algorithm (GA) [28]. Nowadays, deep learning plays an irreplaceable role in ATR, among which CNN has achieved excellent results in feature extraction. As a supervised learning algorithm, CNN extracts the local features of the image through convolution windows sliding on the whole image. Chen [21] adopted a new all-convolutional network (A-ConvNet) for SAR ATR and avoided the overfitting successfully with small training datasets. To tackle the same task, AlexNet was used in [29], which is considered one of the best CNN architectures, to obtain robustness on MSTAR data of extended operating conditions. Feature classification can evaluate and reduce the extracted features to determine the final type of targets. Wagner et al. [30] adopted SVM to replace the fully connected neural network in CNN, and the modified architecture obtains a higher accuracy in MSTAR. Most representative SAR ATR systems are based on MSTAR [18] containing SAR images of ten military vehicles, which is launched by the Defense Advanced Research Projects Agency (DARPA) and Air Force Research Laboratory (AFRL).

Although achieving an increasing accuracy, existing SAR ATR systems still have some problems: Firstly, preprocessing algorithms generally have no adaptive ability and can only improve image quality limitedly. Besides, most systems using the modified feature extractor or feature classifier improve the recognition accuracy by increasing algorithm complexity (the numbers of network layers or channels in deep learning methods), which increases the cost of calculation and brings the risk of overfitting to deep learning methods. There is one more point, since the acquisition of SAR images is costly and time-consuming, the vast majority of existing studies used the available part of MSTAR, while the scenarios in the practical target recognition are changing all the time [10]. The recognition of targets other than military vehicles, such as aircraft and ships, is rarely discussed, caused by a lack of data suitable for training. On the basis of these problems, our research mainly innovates in the stage of pre-processing. SAR images are translated into their optical expression through a SAR-to-optical translation network, which significantly reduces the difficulty of recognition. Feature extraction and classification are achieved through a CNN network and a fully connected network, respectively. In addition, SPH4, a new multi-view SAR dataset of aircraft targets, is created as an attempt to expand the application of SAR ATR technology.

2.2. Image-to-Image Translation Based on cGAN

In order to satisfy the problem of predicting a specific target image from a given similar image in computer vision, Isola et al. [12] presented the concept of image-to-image translation and introduced pix2pix, a common framework for this problem. This famous network is based on cGAN [11], in which two adversarial neural networks, a generator (G) and a discriminator (D), are trained simultaneously. G tries to synthesize realistic data similar enough to the reference data x , while D tries to distinguish synthesized data from x . The input of G includes a random noise z and an extra condition y , the output of G can be represented as $G(z|y)$. y can be an image class, an object property, or even a picture, which is the original SAR images in SAR-to-optical translation. The aim of training D is to make its output $D(x|y) = 1$ and $D(G(z|y)) = 0$. The value function $V(G, D)$ of cGAN is optimized by G and D during training:

$$\min_G \max_D V(G, D) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x | y)] + \mathbb{E}_{z \sim p_z(z)} [1 - D(G(z | y))] \quad (1)$$

On the one hand, the generator in pix2pix is based on a U-Net architecture, which has skip connections between corresponding encoder and decoder layers to share low-level features rather than just passing them through the bottleneck. On the other hand, instead of determining the authenticity of a whole output image, pix2pix uses a PatchGAN discriminator to score each part of the image, and takes the average score as the judgment result of the whole image. The discriminator of cGAN can be regarded as an adaptive loss function between the generator output and the reference image. Compared with the traditional fixed loss function, such as L2 loss, the loss of cGAN can adjust its weight to further optimize the output with the increase in training epochs. Additionally, the L1 loss is combined with the loss of PatchGAN in pix2pix through a hyper-parameter λ to take the whole and the local translation effects into account. The final loss function can be represented as:

$$\mathcal{L}_{\text{pix2pix}}(G, D) = \mathcal{L}_{\text{PatchGAN}}(G, D) + \lambda \mathcal{L}_{L1}(G) \quad (2)$$

Due to the excellent translation performance, pix2pix and its improved versions have numerous applications in image dehazing [31], image classification [32], image colorization [12,33], image super-resolution [34], semantic segmentation [12,35], and so on. Based on pix2pix, we modify the architecture of the U-Net generator and achieve better results in the SAR-to-optical translation of targets in this research.

2.3. Model-Based Data Generation

Deep learning is one of the most advanced target recognition methods, but the training of deep learning algorithms is seriously dependent on a good deal of data. The actual applications are often faced with the problem of data shortage. As an easily accessible material, 3D computer-aided design (CAD) models have received much attention for data augmenting, because images of all kinds of targets in any position and perspective can be generated with their 3D CAD models theoretically. For instance, Joerg [36] presented the method of multi-view object class detection with CAD model renderings assisting. In a later work [37], virtual images rendered from 3D models were used to replace real labeled images for training recognition methods based on whitened histogram of gradients (HOG) features and linear SVM. Similarly, the availability of utilizing CAD models to synthesize training datasets for deep convolutional neural networks was demonstrated in [38].

In terms of SAR target recognition, the model-based data generation is usually adopted to augment existing datasets for the data acquisition is more time-consuming and costly. Malmgren [39] trained a CNN model with a generated SAR dataset based on SAR simulation of 3D models and transferred the trained network to real SAR target recognition in MSTAR. Another case of model-based SAR data generation is described in the Synthetic and Measured Paired Labeled Experiment (SAMPLE) dataset [40] that consists of SAR images from the MSTAR and well-matched simulated data. Simulation in the SAMPLE used

elaborate CAD models with the same configurations and sensor parameters as the SAR imaging process during the MSTAR collection. Beyond traditional simulation algorithms, researchers have started to investigate the use of GAN to make simulated SAR images more realistic [41]. Our research draws on the idea of model-based data generation by using 3D models of targets to generate optical images as a supplement to SAR images.

3. Methods

In this section, architectures of the modified cGAN network for translation and the CNN network for recognition are provided. Lastly, optical image simulation for creating the SPH4 dataset is introduced.

3.1. Translation Network

In order to achieve the SAR-to-optical translation of targets, we present a new image-to-image translation network referring to pix2pix because the traditional pix2pix is not adequate for this translation task. Using SAR images and optical images as original data and reference data, respectively, the traditional pix2pix network is trained and has unacceptable grids and edge blur in the outputs. The generator in pix2pix is a simple repetition of the encoding units and decoding units of the same structure, using a large number of layers and parameters to obtain a higher fitting ability. As a result, pix2pix does not necessarily work well on a particular task. To solve the problem, a more powerful generator architecture is equipped to the new translation network to handle complicated SAR images affected by speckle noise and geometric distortion.

The translation network consists of a generator and a discriminator. In the generator, the traditional U-Net network is improved to an adapted version, with a symmetric architecture to handle the translation between pictures of the same size. Meanwhile, inhomogeneous convolution kernels are adopted instead of adding padding during convolution to keep the image size unchanged, so as to prevent information that does not belong to SAR images from being added to the edge. The adapted U-Net network can work very well in dealing with the target translation and its architecture is shown in Figure 2. Each blue cube represents a multi-channel feature map. The numbers at the top and the bottom left represent the channel number and the map size, respectively. The generator has an encoder-decoder network. The left part of the network is the encoder, which conducts dimension compression and channel expansion for feature maps through a series of convolution modules and max-pooling layers. The convolution module follows the typical structure of convolution + BatchNorm + ReLU [42], in which the use of convolution kernels of different sizes avoids adding padding. The max-pooling layer is adopted with a 2×2 sampling size and stride 2. The right part of the network is the decoder that restores the feature maps to their original size through the opposite operation through deconvolution modules and up-sampling layers. The deconvolution module compresses channels and expands dimensions with a structure of deconvolution + BatchNorm + ReLU. Deconvolutions with 2×2 and 3×3 kernel sizes are used for up-sampling. In addition, jump connections are added between the encoder and the decoder, so that some useful low-level information of input can be directly shared with the output. The overall generator with total of 34,176,065 parameters has the input and the output of 256×256 grayscale.

The discriminator has the same architecture as the pix2pix. After a 4-layer full convolutional neural network and a sigmoid function, a 256×256 grayscale is transformed to a 30×30 matrix. Each value in the matrix represents the correctness of the corresponding local patch. The mean of the matrix is used as the score indicating the image authenticity. The final loss function in pix2pix consists of L1 loss and PatchGAN as (2). L1 loss is concerned about the low-frequency features of overall outputs, while PatchGAN concentrates on the details in local patches of outputs. After a series of experiments, we set the hyper-parameter to 100, which means L1 loss occupies the dominant position, and PatchGAN, as an adaptive loss function, fine-tunes the overall loss to make the output closer to the reference data.

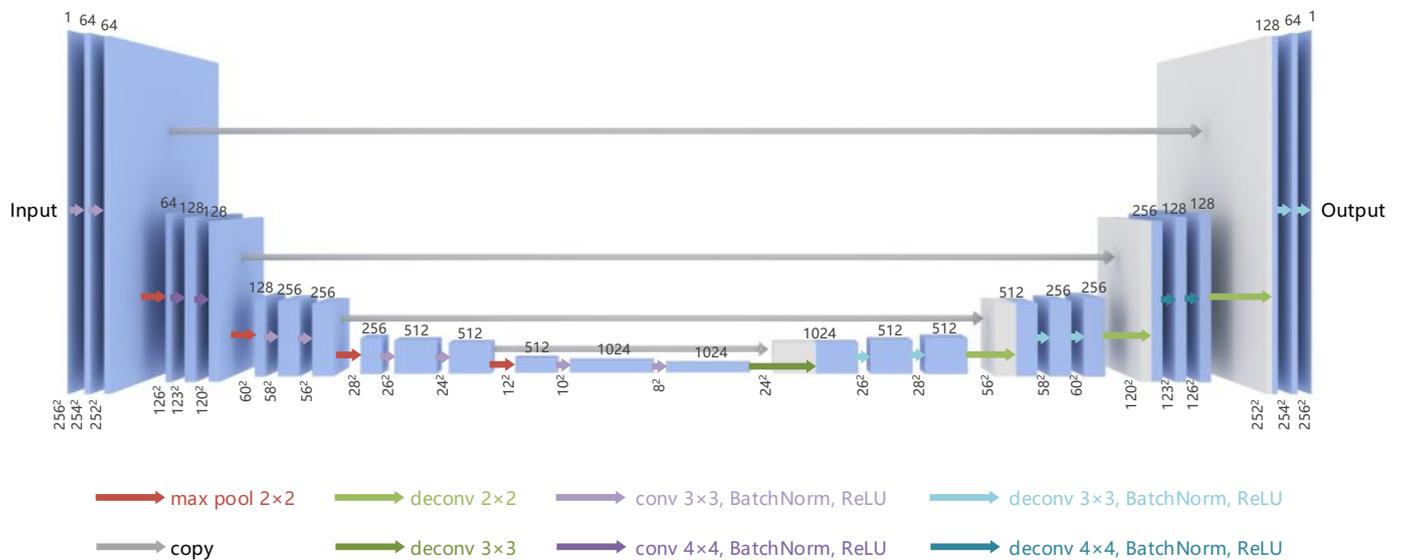


Figure 2. Architecture of the generator. Each blue cube represents a feature map, whose thickness (number on the **top**) and size (number at **lower left**) represent the channel number and the map size, respectively. The gray cube represents the copied feature map. The arrows in different colors represent specific actions to be performed on the feature map.

3.2. Recognition Network

To recognize the translated optical images, we adopt a LeNet [20] architecture, which is one of the first neural networks making a breakthrough in image recognition. The focus of this study is to validate if image translation can enhance the target recognition to achieve higher accuracy, rather than finding the most suitable recognition network to improve the accuracy close to the upper limit. Accordingly, LeNet is chosen for its simple architecture and typicality. In LeNet, a CNN consisting of convolution layers and pooling layers is used for feature extraction, and a fully connected network works as a classifier. We adjust the architecture of LeNet to adapt it to the image size in this recognition problem. The features of an input image are extracted through five convolution layers, each followed by an average pooling layer. Then feature map is flattened into a vector and with which the type of input is determined by four full connection layers.

3.3. Optical Imaging Simulation

There are significant differences in radiometric appearance between common photographic imagery and SAR imagery, which puts the data fusion and the image translation into a dilemma. Assuming an ideal image that has a similar radiometric appearance as the SAR image and no speckle characteristics, it must be an active incoherent imaging result. Additionally, the imaging band should be close to the visible band of electromagnetic waves, so as to bring comfortable image expression and sufficient structural fine granularity. Such an ideal imaging process has a similar imaging mechanism and result with active infrared imaging, so we simulate the optical images based on a ray-tracing algorithm referring to active infrared imaging. The UAV SAR imaging of the real scene and the simulated active infrared imaging of the CAD scene are illustrated in Figure 3. The CAD model of the scene is established based on prior knowledge of the real scene, with elaborate models of four types of aircraft placed according to the relative position in the real scene. It is worth mentioning that the surfaces of aircraft models are programmed to be smooth with strong specular reflection, according to the scattering of high-frequency electromagnetic waves from metal surfaces. Thereafter, we obtain the aircraft position during SAR imaging and set a camera at the corresponding position in the CAD scene, as the UAV works at a fixed altitude $H = 150$ m and the multi-view routes are known. Analogous to the electromagnetic plane wave in the far field, the light source is set to

parallel infrared rays with the same $\theta = 45^\circ$ as the viewpoint of SAR imaging in the CAD scene. Finally, the ray-tracing algorithm tracks the incoming ray backwards at each pixel of the image received by the camera and calculates the reflection and refraction of the ray by the target. While multiple scattering no more than four times is taken into account, the received ray intensity at each pixel is calculated according to the contribution of the light source, and the simulated optical image is produced. These optical images are consistent with corresponding SAR images in semantic information and radiometric appearance, for the characteristics of geometrical optics scattering used in the ray-tracing algorithm are similar to high-frequency microwave scattering. These highly matched optical images complement the SAR target data, which facilitates the SAR-to-optical translation of targets.

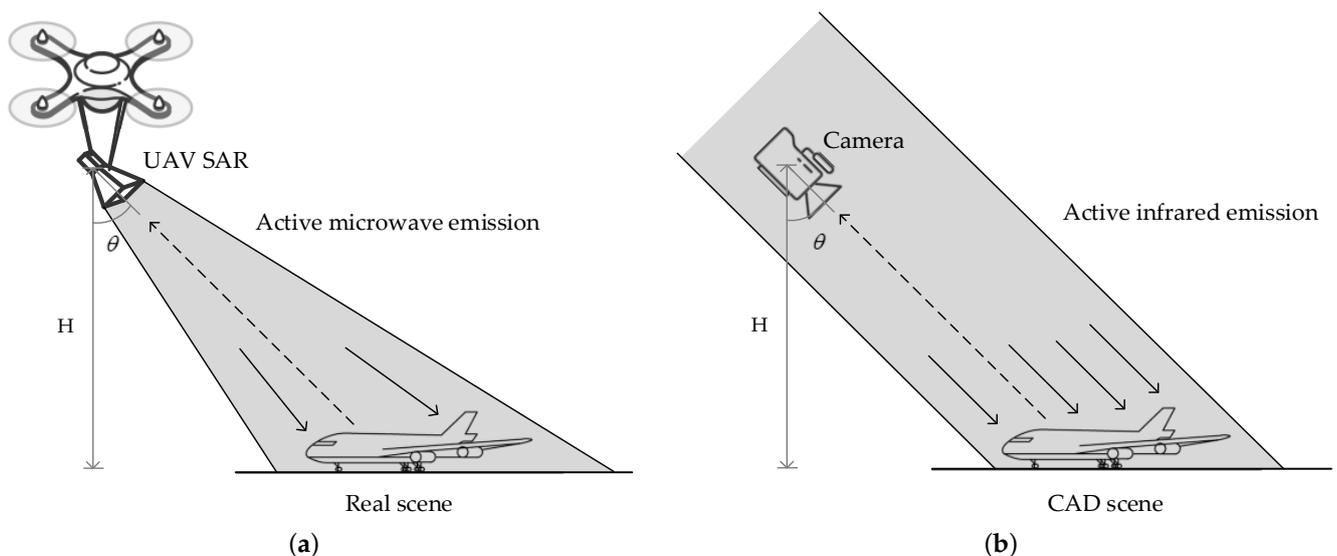


Figure 3. Comparison of SAR imaging and simulated optical imaging. (a) Schematic of SAR imaging; (b) Schematic of simulated optical imaging referring active infrared imaging.

4. Experiments and Results

In this section, details of the SPH4 dataset are introduced firstly. Next, training parameters and hardware configuration in experiments are provided. Lastly, results of the trained recognition system are shown.

4.1. SPH4 Dataset

A new dataset SPH4 is created for this research, which includes pairs of multi-view SAR-optical images of aircraft targets. Two types of small fixed-wing planes (Quest Kodiak 100 Series II and Cessna 208B) and two types of helicopters (Ka-32 and AS350) are selected as targets because of their typical characteristics of fixed-wing aircraft and helicopters. These characteristics can endow the feature extraction network with the ability to adapt to other types of aircraft. The SAR images have a resolution of $0.3 \text{ m} \times 0.3 \text{ m}$, including HH, HV, and VV polarized modes, derived from a Ku-band UAV SAR with a center frequency of 14.6 GHz, a bandwidth of 600 MHz, and a flight altitude of 150 m. Due to the range-based image mechanism and the low-altitude of UAV SAR, targets in the original SAR images are inverted with foreshortening effects, which is not consistent with the perspective of human visual habits. Therefore, the original SAR images are flipped upside down, which means the upper end of the image is proximal and the lower end is distal. Aircraft targets in the original SAR data are sliced, classified, and labeled to annotated 8-bit grayscales with a size of 256×256 pixels. Corresponding optical images of the same size are generated by exploiting the ray-tracing algorithm referring to the active infrared imaging. Images of each aircraft type under each viewpoint are grouped into a category that covers three SAR images in HH, HV, and VV, respectively, and one optical image. With SAR images in different polarization modes sharing one optical image in one category, the SPH4 contains

107 categories, a total of 321 pairs of SAR-optical images. As with the examples of SPH4 and photos of the targets shown in Figure 4, there is high consistency between the corresponding real target, the SAR image, and the simulated optical image.

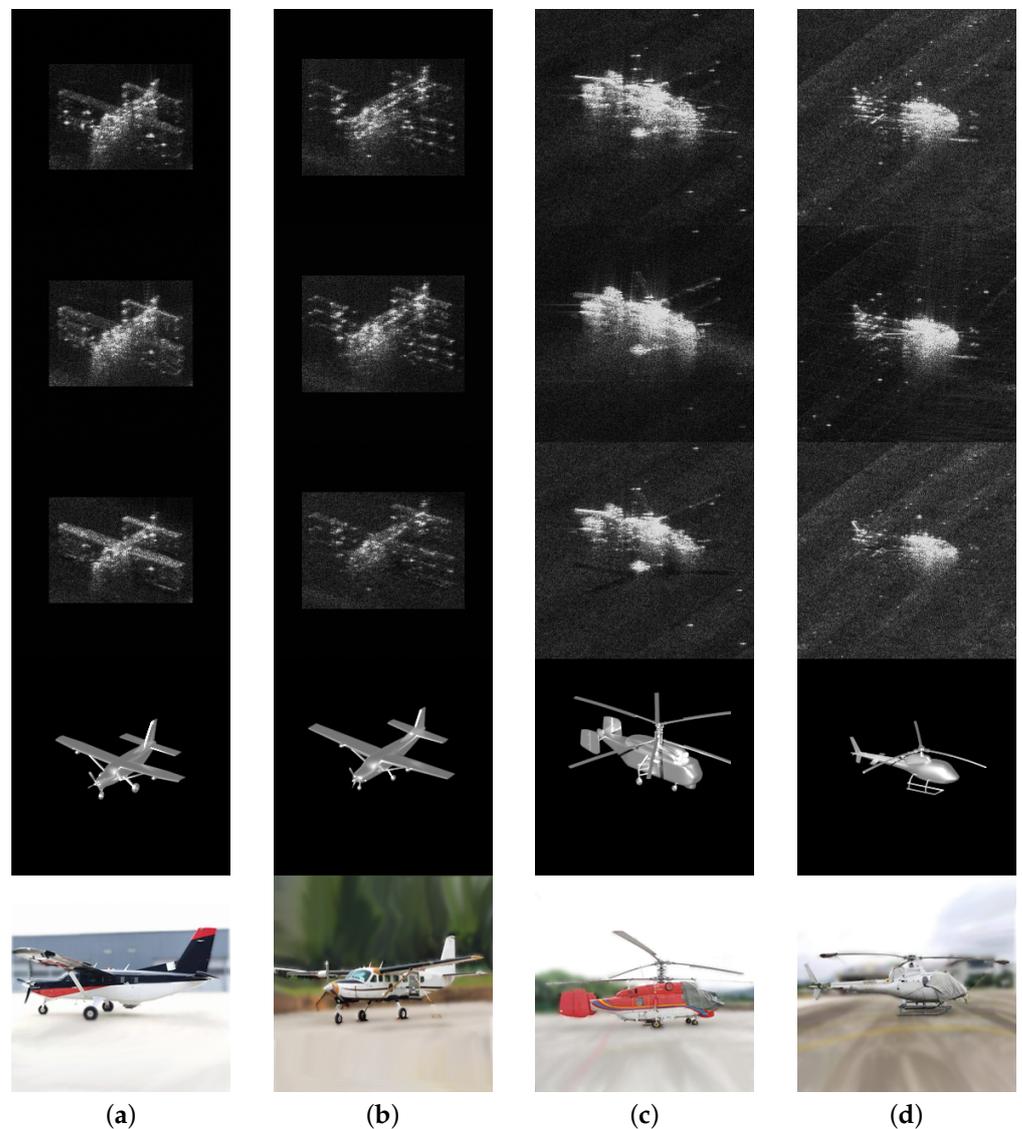


Figure 4. Examples of SPH4 and photos of the targets. Columns (a–d) show one of the categories of Quest Kodiak 100 Series II, Cessna 208B, Ka-32, and AS350, respectively. Images of each row from top to bottom belong to SAR images in HH, SAR images in HV, SAR images in VV, optical images, and photos of the targets, respectively.

4.2. Implement Details

In order to enable the translation network compatible with SAR images of different polarization modes, three SAR images of each category are used as separate inputs. Due to images of the same category having a certain similarity, the segmentation of training and test sets is based on categories to avoid data contamination and over-optimal results. According to aforesaid rules, the data are divided into five approximately equal parts randomly. Each part is used as the test set in turn while the rest are used as the training set. After five experiments, all SAR images in SPH4 are translated into optical.

Data augmentation can improve the robustness of the network and reduce overfitting when available data for training are insufficient. In this study, the training set is augmented by random horizontal flipping and center rotation within $\pm 5^\circ$, for the aircraft targets have

symmetry and the visual effect of optical images is not sensitive to fine-tuning of the viewpoint, respectively.

When training the translation network, random Gaussian noise of $\sigma = 15$ is added to the 256×256 input images and the iteration lasted 2500 epochs. The learning rate is reciprocated between 0.0001 and 0.0003 by the cosine annealing algorithm [43], which helps the model get rid of local minima by circularly changing learning rates. Additionally, we use Adam [44] as the optimizer of training.

The hardware we use in the experiments are Intel(R) Xeon(R) Gold 5218R CPU at 2.10 GHz and NVIDIA GeForce RTX 3090 GPU with a dedicated GPU memory of 24.0 GB. All the codes are written in Python 3.8.5, using the deep learning tools of Pytorch 1.7.1 package in Anaconda.

4.3. Results

In this subsection, we first introduce and evaluate the outputs of the SAR-to-optical translation network, following by automatic and manual SAR target recognition results.

4.3.1. Translation Results

Due to the huge differences in image features among different applications of image-to-image translation, there is no unified evaluation method suitable for all applications. The effect evaluation of translated images is a well-known task. Combining practical application and mathematical analysis, visual evaluation and IQA methods are used to evaluate the quality of translated images.

Examples of the SAR-to-optical translation results are shown in Figure 5: First of all, through local image reconstruction, separated speckles are translated into continuous areas and the background is effectively purified in translated images, which significantly improves the image quality and makes them more friendly to human eyes. Some clutter caused by the SAR imaging process, such as bright stripes in the background of HH SAR images in (f) and (g), are effectively judged as noise and eliminated. Secondly, not only are the main bodies of the aircraft targets, like fuselages, wings, and tail fins (including the horizontal stabilizer fin and the vertical fin), successfully restored to their optical counterpart, but the missing and distorted prior details of the aircraft structure like undercarriages are also recovered. Successful reconstruction of the aircraft structure can better enhance the recognition of aircraft types and orientations. For example, in (I), the structure and the orientation of the KA-32 in SAR images have become difficult to recognized due to defocus, while the same recognition problem in translated images is very easy referring to the features. These positive results verify the effectiveness of the SAR-to-optical translation network of targets.

It is necessary to numerically calculate the difference between translated images and optical images. Traditional IQA methods such as *PSNR* and *SSIM* are widely used in the evaluation of image-to-image translation. *PSNR* is simply defined by mean square error (*MSE*) to calculate the difference between the maximum signal and the background noise by comparing the corresponding pixels in the images point by point:

$$PSNR = 10 \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (3)$$

MAX represents the maximum value of the image pixels. In this study, all images are 8-bit grayscales, and *MAX* is 255. The higher the *PSNR* value, the less distortion it represents. *SSIM* evaluates the image from the perspective of human visual perception,

comprehensively considering luminance, contrast, and structure of the image, which is defined as:

$$SSIM(X, Y) = l(X, Y) \cdot c(X, Y) \cdot s(X, Y)$$

$$l(X, Y) = \frac{2\mu_X\mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1}, c(X, Y) = \frac{2\sigma_X\sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2}, s(X, Y) = \frac{2\sigma_{XY} + C_3}{\sigma_X\sigma_Y + C_3} \quad (4)$$

where μ_X and μ_Y represent the mean values of images X and Y , respectively, σ_X and σ_Y represent the variances of images X and Y , respectively, σ_{XY} represents the covariance of images X and Y , and C_1 , C_2 , and C_3 are constants to avoid the denominator being 0. SSIM ranges from 0 to 1. The larger the value is, the smaller the image distortion is. We assume that optical images are ideal images, and SAR images and translated images are the results of noise addition. As with the average PSNR and SSIM of SAR images and translated images shown in Table 1, the translation network successfully improved image quality and visual effects.

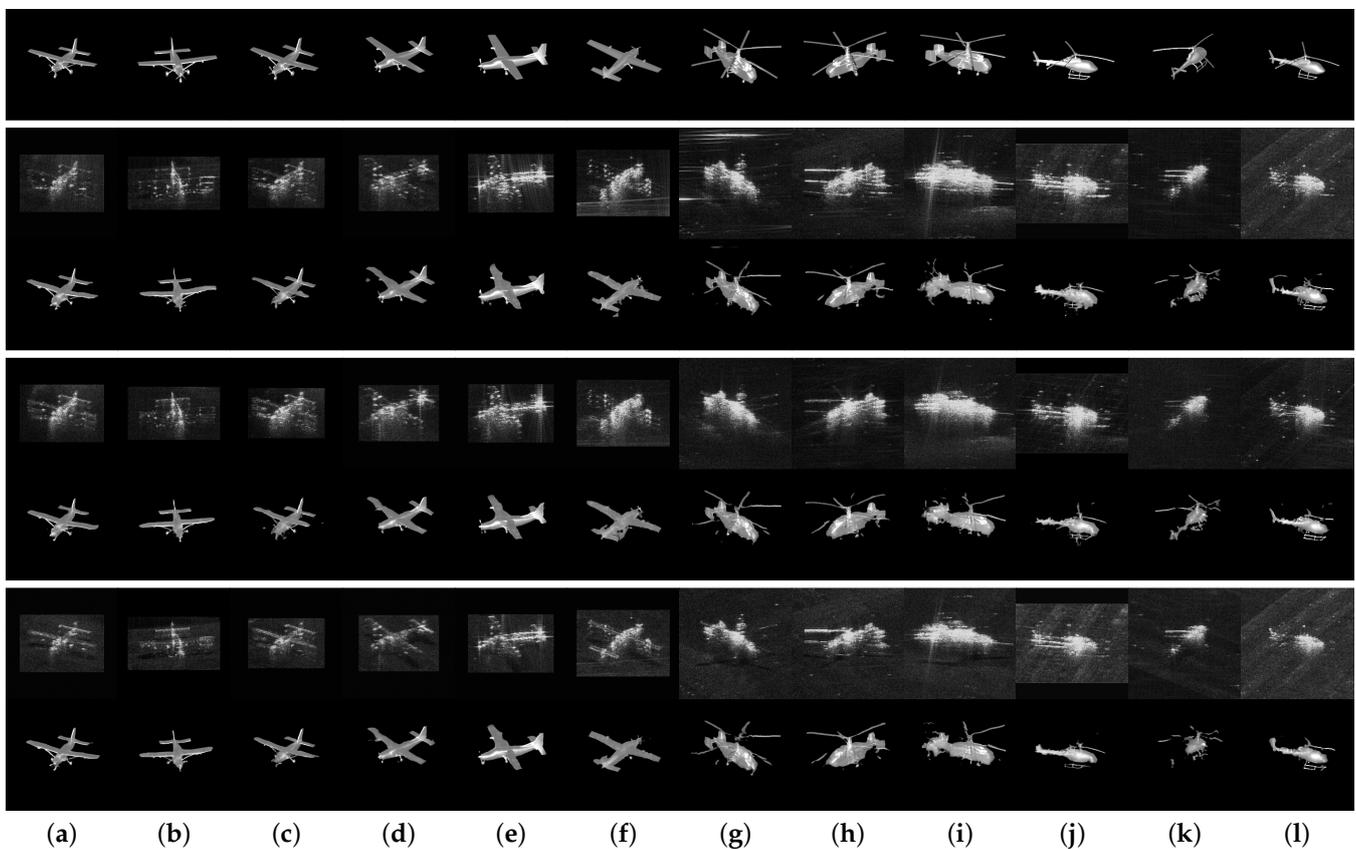


Figure 5. Examples of translation. Columns (a–l) show the results of different categories, respectively, among which (a–c) belong to Quest Kodiak 100 Series II, (d–f) Cessna 208B, (g,h) Ka-32, and (j–l) AS350. The first row shows optical images, then SAR images in HH, SAR images in HV, and SAR images in VV, each followed by corresponding translated images.

Table 1. PSNR and SSIM of different images.

Image Type	PSNR	SSIM
SAR	18.7729	0.4685
Translated	21.4738	0.7420

4.3.2. Recognition Results

A CNN classification network based on LeNet architecture is designed to recognize aircraft types. In the first recognition experiment, the labeled translated image dataset is randomly divided into two equal sets for training and testing the recognition network with no overlap in categories. Meanwhile, SAR images and optical images are tested in the same configuration as controls. Theoretically, with ideal image quality, optical images can get the upper limit of experimental results in classification. Translated images should achieve a higher accuracy rate than SAR images in the classification experiment for the translation reduces the noise and increases the structural details. The experiments are carried out ten times with random grouping, and the average accuracy is shown in Figure 6a. Average accuracy results of optical images < translated images < SAR images is consistent with the theory and shows that translated images generated by the SAR-to-optical translation are more suitable for the CNN classification algorithm.

In another experiment, the CNN network is trained with optical images to recognize translated images and SAR images. The division of training set and test set is consistent with that in the previous experiment. As can be seen in Figure 6b, the accuracy of translated images is less affected, while the accuracy of the SAR images drops significantly. This result makes it possible to recognize aircraft that do not belong to the dataset, because it is theoretically possible to generate as many SAR images of aircraft as we need for training the recognition network through simulation algorithms, as long as the CAD models are available. Both experiments verified the feasibility of enhancing SAR ATR by using SAR-to-optical translation. It can be believed that the optimization of the recognition network architecture and training hyper-parameters can further improve the accuracy, but this is beyond the scope of this research.

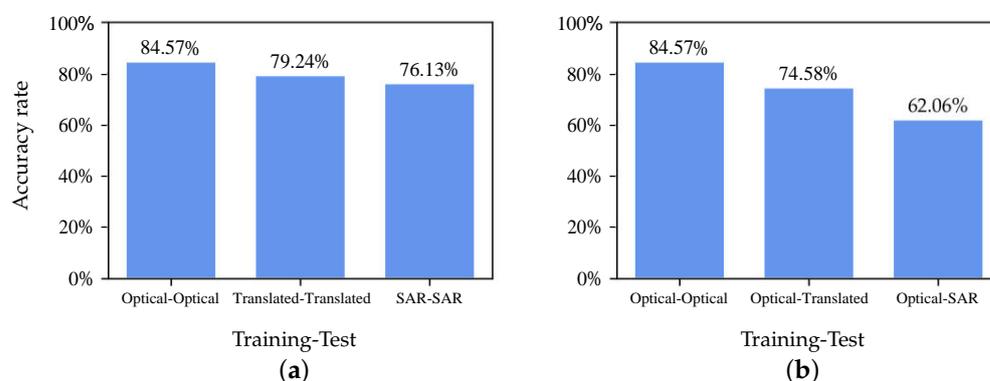


Figure 6. Results of recognition. The ordinates indicate the recognition accuracy. In the horizontal coordinate, the former is the type of training data, and the latter is the type of test data. (a) represents the results of training and testing using the same kind of data. (b) represents the results of training with optical images, testing with optical images, translated images, and SAR images.

Furthermore, the manual target recognition experiment (details of the experiment can be found in the Appendix A Figure A1) is implemented with six SAR professionals and the result is shown in Table 2. In terms of types and orientations, the average accuracies of translated image classification are 77.97% and 95.64%, respectively, which are significantly higher than 70.30% and 92.37% of SAR images classification.

Table 2. Results of manual target recognition experiment. Average results of six SAR professionals are in bold.

Name	Type Accuracy		Orientation Accuracy	
	SAR	Translated	SAR	Translated
K.F.	0.6854	0.7508	0.9221	0.9533

Table 2. Cont.

Name	Type Accuracy		Orientation Accuracy	
	SAR	Translated	SAR	Translated
J.H	0.6106	0.7445	0.8847	0.9533
J.L.	0.6791	0.6916	0.9159	0.9408
R.W.	0.7975	0.8723	0.9470	0.9626
Q.X.	0.6667	0.7165	0.9252	0.9626
J.Y.	0.7788	0.9003	0.9470	0.9657
Average	0.7030	0.7797	0.9237	0.9564

5. Discussion

In this section, extended experiments are implemented to test the anti-noise performance and adaptability of the trained translation network. In the end, some failed cases are discussed.

5.1. Expanding Experiments

In the previous section, we verify the feasibility of SAR-to-optical translation in enhancing SAR target recognition. However, in practical applications, SAR images not only need to face the noise generated during the signal and image processing but also often contain new aircraft that do not belong to the existing dataset. Therefore, the translation network needs the robustness to resist noise and the extensibility to adapt to new targets.

5.1.1. Noise Resistance

Gaussian noise and salt-and-pepper noise are two kinds of common noise in imaging processing [45]. Gaussian noise is usually caused by poor working conditions of imaging sensors and inadequate light sources, whose intensity can be described by the standard deviation σ of the Gaussian distribution. Salt-and-pepper noise, as a kind of pulse noise, randomly changes some pixel values of images, which is the black and white bright spot noise generated in transmission channels, decoding processing, and so on. The intensity of salt-and-pepper noise is determined by the change probability of pixels.

For testing the anti-noise performance of the translated network, a series of Gaussian noise or salt-and-pepper noise with different intensities are added to input SAR images. Examples of output results are shown in Figure 7. As can be seen from (a) and (b), with the increase in noise intensity, the main bodies of the aircraft in the SAR images are gradually submerged, which will bring great difficulties to automatic and manual target recognition. While the background noise is removed effectively, the main structural features of the aircraft are retained in the translated images, even in the face of strong noise. These results verify the robustness of the network in the case of high-intensity image noise.

5.1.2. Type Extension

Using the SAR images of new type aircraft to test the translation network can verify the performance of feature extraction. The ideal network can realize SAR-to-optical translation by recognizing and matching the local features of aircraft in SAR images and optical images. Because aircraft are designed based on aerodynamics and have a certain similarity in local features, the well-trained network should be able to recognize and translate extended types of aircraft. The SAR images of PC-12, Beech King Air 350, AT-504 fixed-wing aircraft, and AW 139 helicopter obtained under the same operating condition as SPH4 are used for testing in the type extension experiment, and the results are illustrated in Figure 8. It can be seen that the aircraft in the translated images have better visual effects. The fuselage, wing, and tail dimensions of fixed-wing aircraft are effectively restored, which makes it easier to identify aircraft types by aspect ratio. Additionally, the high horizontal tail of PC-12, the two propeller engines of the Beech King Air 350, and the tapered front end of the AW 139 can also be identified in the translated images. These features are different from the four

types of aircraft in SPH4, but the network still has the ability to restore them. Such a result proves that the network has successfully learned the core features of the aircraft and has strong extensibility.

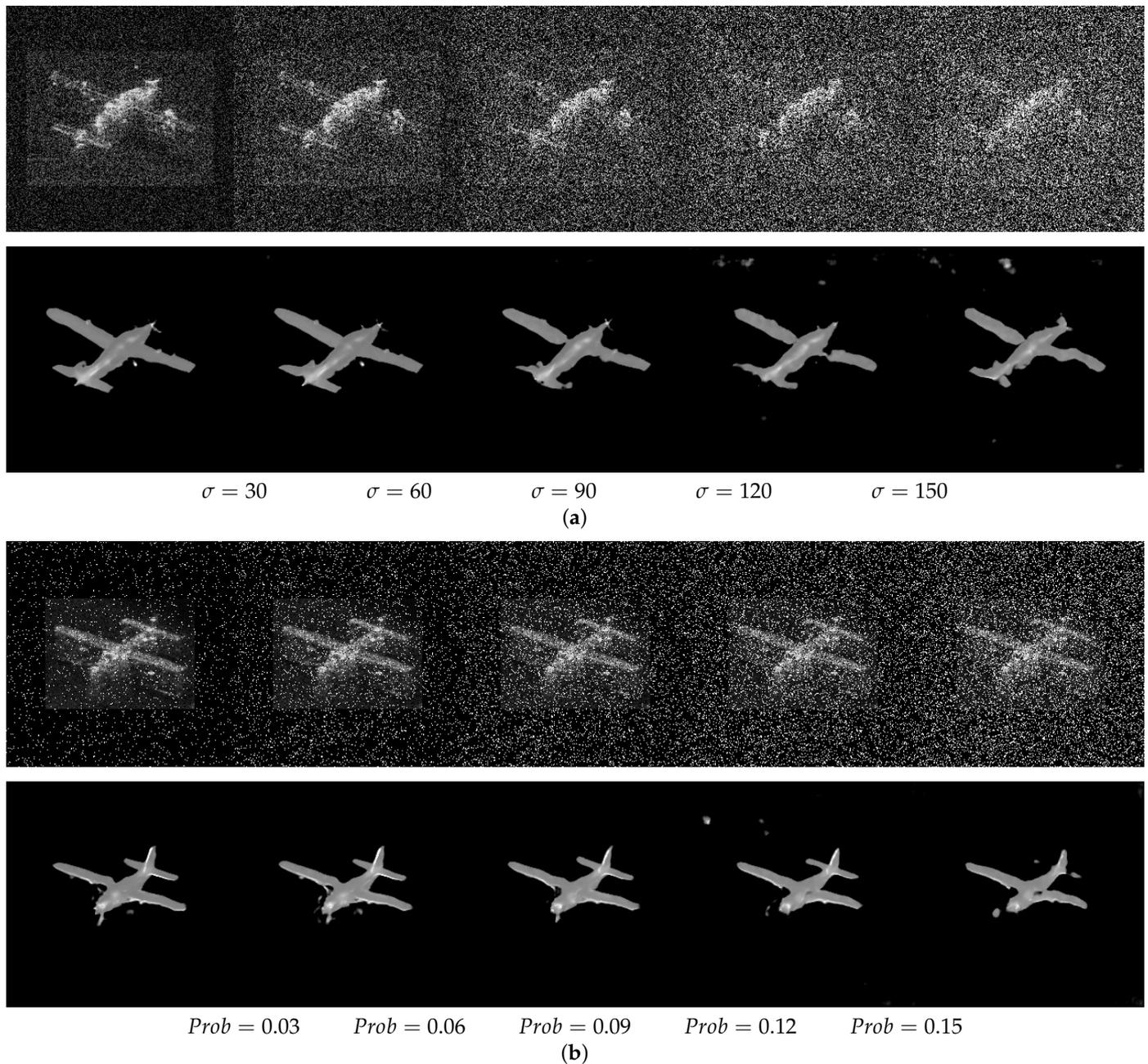


Figure 7. Translation results of noisy SAR images. (a) SAR images with Gaussian noise of different σ added (the top row) and corresponding translated images (the following row); (b) SAR images with salt-and-pepper noise of different probabilities added (the top row) and corresponding translated images (the following row).

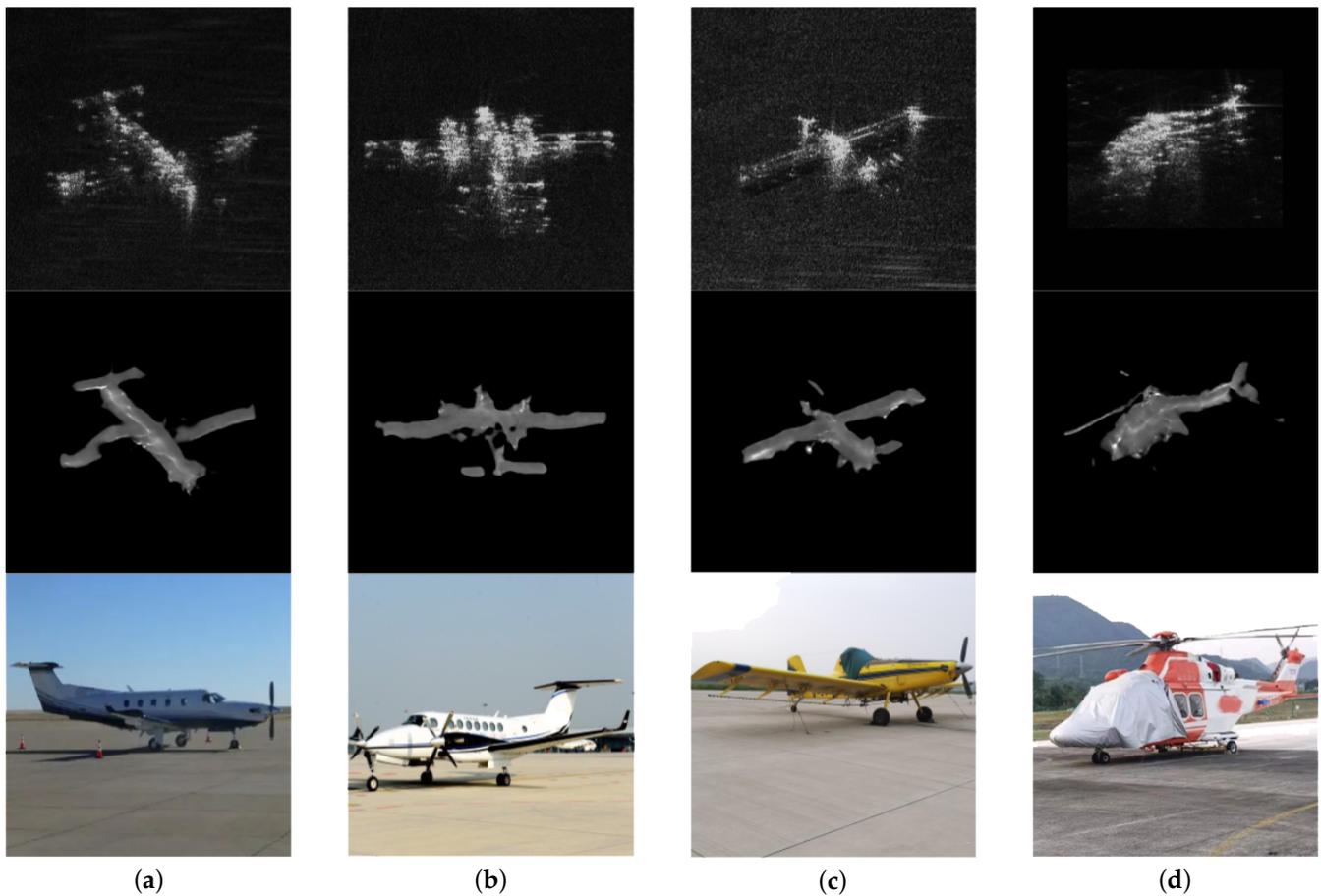


Figure 8. Translation results of extended aircraft types. Panels (a–d) show the results of PC-12, Beech King Air 350, AT-504, and AW 139, respectively. Images of each row from top to bottom belong to SAR images, translated images, and photos of the targets respectively.

5.2. Failed Cases

There are background noise and defocus in some SAR images during imaging processing. The SAR-to-optical translation in this paper can cope well with slight clutter, such as (a) and (b) in Figure 9. However, some failed cases appear because of excessive clutter, as shown in Figure 9: The VV SAR image of (c) has a wide range of background noise, which can be comparable with the aircraft target in radiation and causes the ghosts in the translated image. Defocus in the SAR images of (d) and (e) makes the original characteristics of the wings difficult to be recognized and translated by the network, which leads to the partial absence of the wings in the translated images. In (f), the tails of the helicopter in the SAR images are prolonged due to defocus, and are identified as tails of the fixed-wing aircraft, which causes wrong shape features of the horizontal stabilizer fin in the translated images. The fuselages of the helicopter in the SAR images of (g) are elongated due to defocus, which leads to the long fuselages in the translated images erroneously. The aircraft in the SAR images of (h) is at the edge of the images, resulting in the tail missing, while the same tail missing appears in the translated images. Most of these failures can be attributed to errors in SAR image imaging processing. The problematic SAR images make automatic and manual target recognition even more difficult, while it appears that the translation network can dumb this problem down. It is worth mentioning that these results of the problematic SAR images identify that the SAR-to-optical translation is achieved through local feature recognition and mapping, which gives full respect to SAR images and effectively avoids overfitting.

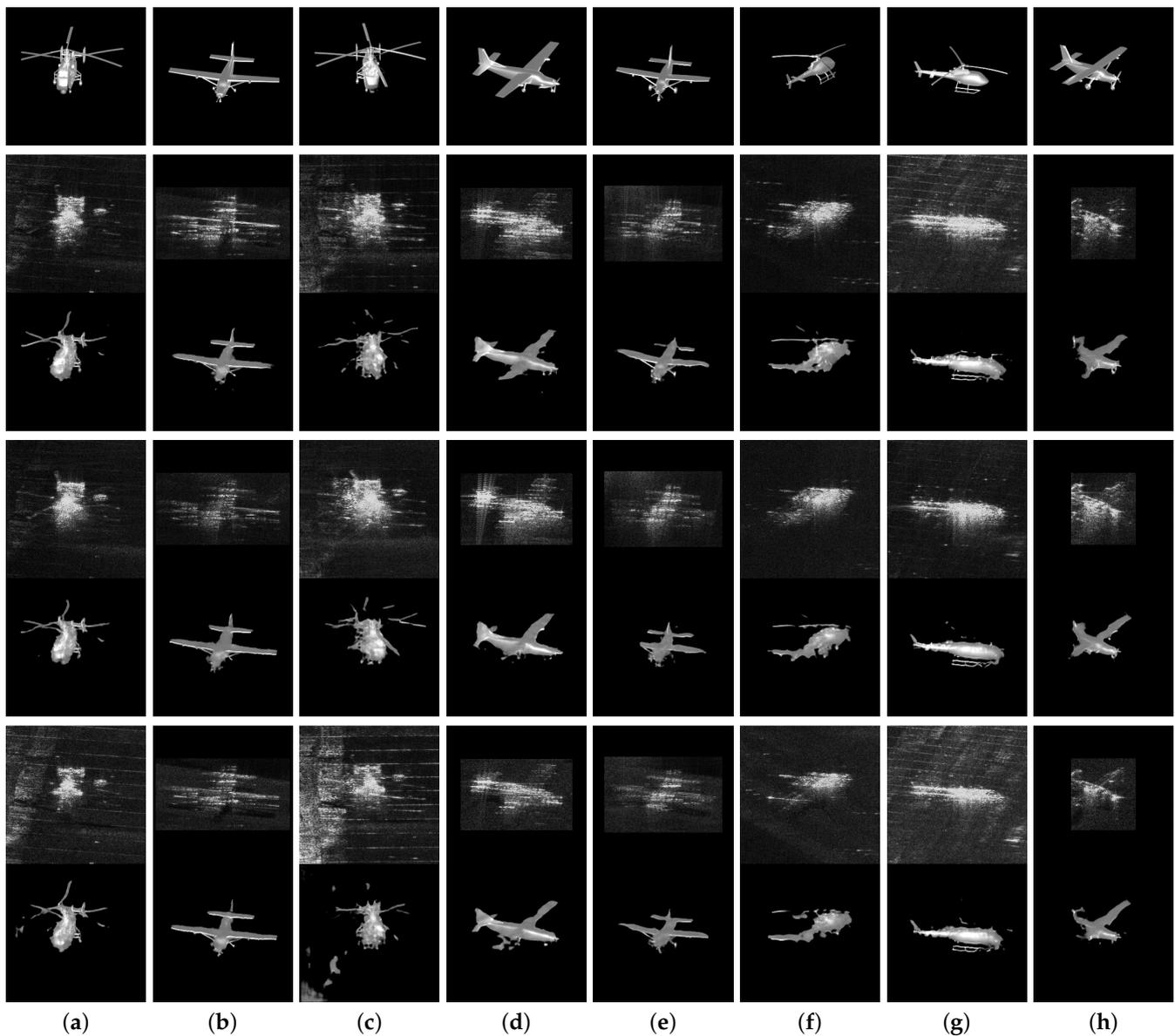


Figure 9. Examples of the failed cases. Columns (a–h) show the results of different categories, respectively. The first row shows optical images, then SAR images in HH, SAR images in HV, and SAR images in VV, each followed by corresponding translated images.

6. Conclusions and Outlook

The experimental results in this study demonstrate the great potential of SAR-to-optical translation in enhancing SAR target recognition, both automatic and manual. Firstly, a novel method is proposed to generate optical images well-matched with SAR images. This method can produce optical images that are highly consistent with SAR images in semantic information and radiometric appearance through the model-based computer simulation, which breaks the limitation of multi-sensor load and provides a new way for solving the shortage of SAR-to-optical datasets. Benefitting from this method, a new dataset SPH4 containing multi-view SAR-optical images of aircraft targets is created, which can be used in SAR-to-optical translation, target recognition, and other subsequent studies. Such a dataset opens the door to research on the SAR-to-optical translation of targets. Secondly, a cGAN-based translation network with a symmetric U-Net generator and a PatchGAN discriminator is proposed and its excellent performance on the SAR-to-optical translation of targets is verified through experiments. The evaluation of experimental results based on human vision and mathematical analysis shows that the translation

network can successfully translate SAR images into optical expression through noise reduction and local structural feature translation without overfitting. As a preprocessing step in the target recognition, SAR-to-optical translation offers a promising new route for improving the interpretability and quality of SAR target images, which can enhance manual target recognition and CNN-based SAR ATR to achieve a higher accuracy. In addition, with the SAR-to-optical translation network, ATR methods can use simulated optical images for training to recognize translated SAR targets, which breaks the limit of target types in training SAR data and improves the performance of SAR ATR. Finally, experiments of noise addition and aircraft type expansion verify the stability and adaptability of the translation network. All these results confirm the promising potential of this system towards practical applications ranging from airport monitoring to all-weather reconnaissance.

There are only a few types of aircraft used in training. For example, the fixed-wing aircraft are all single-wing, which makes it difficult to accurately restore the shape and structure of some special aircraft. In subsequent research, more aircraft types and imaging bands will be involved in the dataset to enable the SAR-to-optical translation network to gain greater compatibility with complex application scenarios.

Author Contributions: Conceptualization and methodology, Y.S. and W.J.; validation, Y.S. and J.Y.; software, data processing, and writing—original draft preparation, Y.S.; writing—review and editing, supervision, and project administration, W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China grant number (2018YFA0701900, 2018YFA0701901).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: If need be, please email to wzli@mail.ie.ac.cn to access the SPH4 dataset.

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

Appendix A. Manual Target Recognition Experiment

The manual target recognition experiment is implemented based on the criteria shown in Figure A1a,b with six SAR professionals. Every experimenter recognizes the type and the orientation of all the 521 SAR images and the 521 translated images in the dataset. There are five options for aircraft type and orientation recognition: 0, 1, 2, 3, or unable to tell. The accuracy only increases when the correct option is selected. Some examples of the manual recognition are shown in Figure A1c.

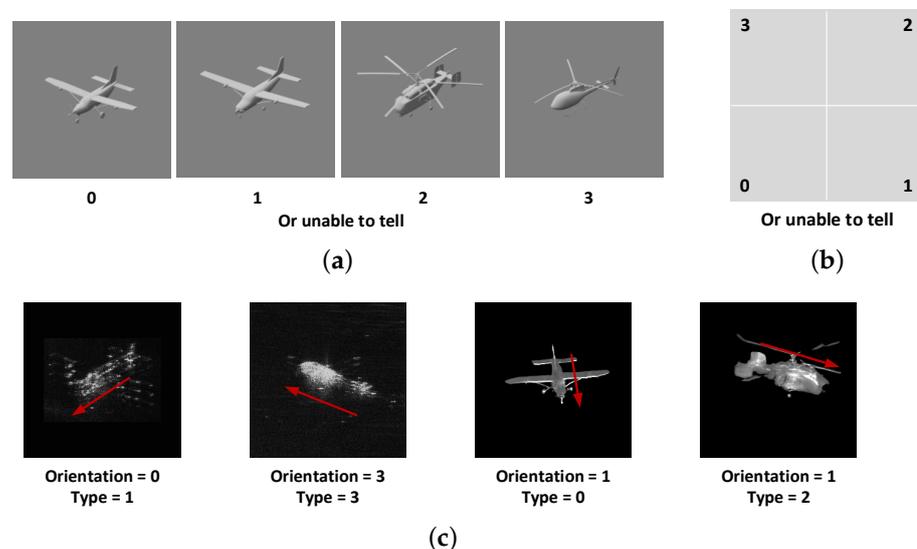


Figure A1. Details of manual target recognition experiments. (a,b) are the judgment criteria of types and orientations, respectively. (c) shows examples of recognition in SAR images and translated images.

References

1. Huang, Z.; Pan, Z.; Lei, B. What, Where, and How to Transfer in SAR Target Recognition Based on Deep CNNs. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 2324–2336. [[CrossRef](#)]
2. Goodman, J.W. Some fundamental properties of speckle. *JOSA* **1976**, *66*, 1145–1150. [[CrossRef](#)]
3. Wang, P.; Zhang, H.; Patel, V.M. SAR Image Despeckling Using a Convolutional Neural Network. *IEEE Signal Process. Lett.* **2017**, *24*, 1763–1767. [[CrossRef](#)]
4. 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. [[CrossRef](#)]
5. Schmitt, M.; Hughes, L.H.; Zhu, X.X. The SEN1-2 Dataset for Deep Learning in SAR-Optical Data Fusion. *arXiv* **2018**, arxiv:1807.01569.
6. Fuentes Reyes, M.; Auer, S.; Merkle, N.; Henry, C.; Schmitt, M. SAR-to-Optical Image Translation Based on Conditional Generative Adversarial Networks—Optimization, Opportunities and Limits. *Remote Sens.* **2019**, *11*, 2067. [[CrossRef](#)]
7. Hughes, L.H.; Schmitt, M.; Mou, L.; Wang, Y.; Zhu, X.X. Identifying Corresponding Patches in SAR and Optical Images with a Pseudo-Siamese CNN. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 784–788. [[CrossRef](#)]
8. Cheng, G.; Han, J. A survey on object detection in optical remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **2016**, *117*, 11–28. [[CrossRef](#)]
9. Dong, C.; Liu, J.; Xu, F. Ship Detection in Optical Remote Sensing Images Based on Saliency and a Rotation-Invariant Descriptor. *Remote Sens.* **2018**, *10*, 400. [[CrossRef](#)]
10. Ren, Y.; Zhu, C.; Xiao, S. Small Object Detection in Optical Remote Sensing Images via Modified Faster R-CNN. *Appl. Sci.* **2018**, *8*, 813. [[CrossRef](#)]
11. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784.
12. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134.
13. Bermudez, J.D.; Happ, P.N.; Feitosa, R.Q.; Oliveira, D.A.B. Synthesis of Multispectral Optical Images From SAR/Optical Multitemporal Data Using Conditional Generative Adversarial Networks. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1220–1224. [[CrossRef](#)]
14. Turnes, J.N.; Castro, J.D.B.; Torres, D.L.; Vega, P.J.S.; Feitosa, R.Q.; Happ, P.N. Atrous cGAN for SAR to Optical Image Translation. *IEEE Geosci. Remote Sens. Lett.* **2020**, *19*, 1–5. [[CrossRef](#)]
15. Darbaghshahi, F.N.; Mohammadi, M.R.; Soryani, M. Cloud removal in remote sensing images using generative adversarial networks and SAR-to-optical image translation. *arXiv* **2020**, arXiv:2012.12180.
16. Zhang, J.; Zhou, J.; Li, M.; Zhou, H.; Yu, T. Quality Assessment of SAR-to-Optical Image Translation. *Remote Sens.* **2020**, *12*, 3472. [[CrossRef](#)]
17. Wang, L.; Xu, X.; Yu, Y.; Yang, R.; Gui, R.; Xu, Z.; Pu, F. SAR-to-Optical Image Translation Using Supervised Cycle-Consistent Adversarial Networks. *IEEE Access* **2019**, *7*, 129136–129149. [[CrossRef](#)]
18. Keydel, E.R.; Lee, S.W.; Moore, J.T. MSTAR extended operating conditions: A tutorial. In *Algorithms for Synthetic Aperture Radar Imagery III*; SPIE: Bellingham, WA, USA, 1996; Volume 2757, pp. 228–242.
19. Pohl, C.; Van Genderen, J. *Remote Sensing Image Fusion*; CRC Press: Boca Raton, FL, USA, 2016. [[CrossRef](#)]
20. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [[CrossRef](#)]
21. Chen, S.; Wang, H.; Xu, F.; Jin, Y.Q. Target Classification Using the Deep Convolutional Networks for SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4806–4817. [[CrossRef](#)]
22. Novak, L.M.; Owirka, G.J.; Weaver, A.L. Automatic target recognition using enhanced resolution SAR data. *IEEE Trans. Aerosp. Electron. Syst.* **1999**, *35*, 157–175. [[CrossRef](#)]
23. Sun, Y.; Liu, Z.; Todorovic, S.; Li, J. Adaptive boosting for SAR automatic target recognition. *IEEE Trans. Aerosp. Electron. Syst.* **2007**, *1*, 112–125. [[CrossRef](#)]
24. Yuan, X.; Tang, T.; Xiang, D.; Li, Y.; Su, Y. Target recognition in SAR imagery based on local gradient ratio pattern. *Int. J. Remote Sens.* **2014**, *35*, 857–870. [[CrossRef](#)]
25. Zhu, H.; Lin, N.; Leung, H.; Leung, R.; Theodidis, S. Target Classification from SAR Imagery Based on the Pixel Grayscale Decline by Graph Convolutional Neural Network. *IEEE Sens. Lett.* **2020**, *4*, 1–4. [[CrossRef](#)]
26. Mishra, A.K.; Motaung, T. Application of linear and nonlinear PCA to SAR ATR. In Proceedings of the 2015 25th International Conference Radioelektronika (RADIOELEKTRONIKA), Pardubice, Czech Republic, 21–22 April 2015; pp. 349–354.
27. Zhao, Q.; Principe, J.C. Support vector machines for SAR automatic target recognition. *IEEE Trans. Aerosp. Electron. Syst.* **2001**, *37*, 643–654. [[CrossRef](#)]
28. Bhanu, B.; Lin, Y. Genetic algorithm based feature selection for target detection in SAR images. *Image Vis. Comput.* **2003**, *21*, 591–608. [[CrossRef](#)]
29. Majumder, U.; Christiansen, E.; Wu, Q.; Inkawhich, N.; Blasch, E.; Nehrbass, J. High-performance computing for automatic target recognition in synthetic aperture radar imagery. In *Cyber Sensing 2017*; International Society for Optics and Photonics: Bellingham, WA, USA, 2017; Volume 10185, p. 1018508.

30. Wagner, S.A. SAR ATR by a combination of convolutional neural network and support vector machines. *IEEE Trans. Aerosp. Electron. Syst.* **2016**, *52*, 2861–2872. [[CrossRef](#)]
31. Qu, Y.; Chen, Y.; Huang, J.; Xie, Y. Enhanced pix2pix dehazing network. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 8160–8168.
32. Wang, X.; Yan, H.; Huo, C.; Yu, J.; Pant, C. Enhancing Pix2Pix for remote sensing image classification. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 2332–2336.
33. Li, Y.; Fu, R.; Meng, X.; Jin, W.; Shao, F. A SAR-to-Optical Image Translation Method Based on Conditional Generation Adversarial Network (cGAN). *IEEE Access* **2020**, *8*, 60338–60343. [[CrossRef](#)]
34. Yuan, Y.; Liu, S.; Zhang, J.; Zhang, Y.; Dong, C.; Lin, L. Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 701–710.
35. Sato, M.; Hotta, K.; Imanishi, A.; Matsuda, M.; Terai, K. Segmentation of Cell Membrane and Nucleus by Improving Pix2pix. In Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2018), Funchal, Portugal, 19–21 January 2018; pp. 216–220.
36. Liebelt, J.; Schmid, C. Multi-view object class detection with a 3d geometric model. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 1688–1695.
37. Sun, B.; Saenko, K. From Virtual to Reality: Fast Adaptation of Virtual Object Detectors to Real Domains. In Proceedings of the BMVC 2014, Nottingham, UK, 1–5 September 2014; Volume 1, p. 3.
38. Peng, X.; Sun, B.; Ali, K.; Saenko, K. Learning deep object detectors from 3d models. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1278–1286.
39. 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](#)]
40. Lewis, B.; Scarnati, T.; Sudkamp, E.; Nehrbass, J.; Rosencrantz, S.; Zelnio, E. A SAR dataset for ATR development: The Synthetic and Measured Paired Labeled Experiment (SAMPLE). In *Algorithms for Synthetic Aperture Radar Imagery XXVI*; International Society for Optics and Photonics: Bellingham, WA, USA, 2019; Volume 10987, p. 109870H.
41. Lewis, B.; Liu, J.; Wong, A. Generative adversarial networks for SAR image realism. In *Algorithms for Synthetic Aperture Radar Imagery XXV*; International Society for Optics and Photonics: Bellingham, WA, USA, 2018; Volume 10647, p. 1064709.
42. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 448–456.
43. Loshchilov, I.; Hutter, F. SGDR: Stochastic Gradient Descent with Warm Restarts. *arXiv* **2017**, arXiv:1608.03983.
44. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
45. Boyat, A.K.; Joshi, B.K. A review paper: Noise models in digital image processing. *arXiv* **2015**, arXiv:1505.03489.