



## Article

# Blind Super-Resolution for SAR Images with Speckle Noise Based on Deep Learning Probabilistic Degradation Model and SAR Priors

Chongqi Zhang, Ziwen Zhang, Yao Deng, Yueyi Zhang, Mingzhe Chong, Yunhua Tan \* and Pukun Liu

School of Electronics, Peking University, Beijing 100871, China

\* Correspondence: tanggeric@pku.edu.cn; Tel.: +86-159-1101-2611

**Abstract:** As an active microwave coherent imaging technology, synthetic aperture radar (SAR) images suffer from severe speckle noise and low-resolution problems due to the limitations of the imaging system, which cause difficulties in image interpretation and target detection. However, the existing SAR super-resolution (SR) methods usually reconstruct the images by a determined degradation model and hardly consider multiplicative speckle noise, meanwhile, most SR models are trained with synthetic datasets in which the low-resolution (LR) images are down-sampled from their high-resolution (HR) counterparts. These constraints cause a serious domain gap between the synthetic and real SAR images. To solve the above problems, this paper proposes an unsupervised blind SR method for SAR images by introducing SAR priors in a cycle-GAN framework. First, a learnable probabilistic degradation model combined with SAR noise priors was presented to satisfy various SAR images produced from different platforms. Then, a degradation model and a SR model in a unified cycle-GAN framework were trained simultaneously to learn the intrinsic relationship between HR–LR domains. The model was trained with real LR and HR SAR images instead of synthetic paired images to conquer the domain gap. Finally, experimental results on both synthetic and real SAR images demonstrated the high performance of the proposed method in terms of image quality and visual perception. Additionally, we found the proposed SR method demonstrates the tremendous potential for target detection tasks by reducing missed detection and false alarms significantly.

**Keywords:** synthetic aperture radar (SAR); blind super-resolution (SR); cycle-GAN; generative adversarial networks (GAN); probabilistic degradation model



**Citation:** Zhang, C.; Zhang, Z.; Deng, Y.; Zhang, Y.; Chong, M.; Tan, Y.; Liu, P. Blind Super-Resolution for SAR Images with Speckle Noise Based on Deep Learning Probabilistic Degradation Model and SAR Priors. *Remote Sens.* **2023**, *15*, 330. <https://doi.org/10.3390/rs15020330>

Academic Editors: Igor Yanovsky and Jing Qin

Received: 15 November 2022

Revised: 29 December 2022

Accepted: 3 January 2023

Published: 5 January 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/>).

## 1. Introduction

Synthetic aperture radar (SAR) is an active microwave coherent imaging technology, which can produce high-resolution (HR) images regardless of adverse light and weather conditions. Hence, SAR plays an extremely important role in target detection and recognition, which is widely used for military and civilian purposes. Therefore, high-quality SAR images with more details and accurate information are required. However, the resolution is related to the signal bandwidth, the center frequency and imaging modes, and improving these system configurations costs a lot. Moreover, SAR images often suffer from the interference of speckle noise and lack high-frequency information due to their coherence imaging mechanism, which causes difficulty in scene interpretation and analysis. Therefore, the development of an SAR image super-resolution (SR) reconstruction algorithm that can provide more details without additional hardware cost remains an attractive but challenging problem.

In recent years, various methods of SR and noise despeckling have been extensively studied in the literature, respectively, and a comprehensive review of SR was provided in [1]. Nevertheless, noise must be taken into account, otherwise, the SR process will amplify the noise and will not give the image any more useful information due to the heavier noise.

However, little SR research about SAR images has been proposed. Additionally, most SAR SR methods merely consider classical additive Gaussian noise models rather than multiplicative and nonwhite speckle noise. Therefore, those methods are incompatible with the real image degradation model and only few addressed this problem.

In recent years, SR reconstruction algorithms for various images have been developing rapidly, and single-image SR (SISR) has gained more attention than multi-image super-resolution (MISR) due to its high efficiency (no need for extra images) [2]. It is hard to obtain SAR images with a series of corresponding images, so we mainly focus on SISR. Image SR methods can mainly be classified into traditional methods and deep-learning-based ones [3]. Traditional ones contain interpolation, pansharpening [4] and some other digital image processing methods [5–9]; Interpolation algorithms [10], such as bilinear and bicubic interpolation, have become the most popular methods to render SR images because of their low complexity and high computing speed, but interpolation does not allow for the obtaining of more details due to its simplicity. Meanwhile, pansharpening also requires LR-HR paired images. With the emergence of compressive sensing technology, sparse representation-based algorithms [11–15] have been used for SR reconstruction, but these algorithms remain complicated. Over the past decades, we have witnessed significant development in deep learning algorithms, and these methods have produced state-of-the-art results in image restoration [16,17] and image SR [18]. Deep-learning-based methods can be divided into two groups, one utilizes convolutional neural networks (CNN) and the other is based on generative adversarial networks [19] (GAN). SRCNN [18] is the first CNN-based method, which can learn an end-to-end LR-HR model for image SR. Following this work, various architectures with powerful techniques (residual blocks and recursive supervision) [20,21] based on SRCNN have been proposed, and more complex and deeper models are emerging. Enhanced deep SR [22] (EDSR) employs several residual blocks to extract image features, deconvolution layers are introduced in FSRCNN [23] and the feedback network employed in SRFBN [24] also performs well. Moreover, GAN shows great potential in SR tasks due to its powerful ability to generate indistinguishable images, which improves visual effects in terms of visual metrics. The SRGAN [25] algorithm is the first one to apply GAN to SISR; meanwhile, ESRGAN [26] improves this method by replacing the basic block with a residual-in-residual dense block (RRDB) and removing batch normalization (BN) layers. For blind SR, several methods show great performance in the real-world images by designing degradation models elaborately, such as real-ESRGAN [27], real-SR [28] and PDM-SR [29].

However, most deep-learning-based SR methods aimed at SAR images just migrate the methods used in optical images, which may not take full account of the SAR imaging characteristics (such as complicated background and multiplicative speckle noise) and cause undesirable effects. One of them [30] adopted a non-local means denoising combined BP network to achieve HR SAR images. SRGAN [31] proposed a GAN with a perceptual loss function, which made remarkable progress in both reconstruction accuracy and computational efficiency. SNGAN [32] followed the ESRGAN to reduce the computational requirement and model oscillation by canceling the BN layers. A novel model [33] with deconvolution and PReLU activation function was designed for PolSAR SR tasks. OGSRN [34] obtained comprehensive information from co-registered High-Resolution (HR) optical images to guide the SAR image reconstruction. Ref. [35] employed GAN and solved the domain gap between synthetic and real-world LR DEMs for InSAR HR DEMs estimation. Nevertheless, these SAR SR methods hardly take the multiplicative speckle noise into account.

SAR denoising and deblurring problem has been studied extensively due to their significance. SAR speckle noise has been analyzed in several models [36–38] with different statistical properties, and plenty of methods [39,40] have been designed to alleviate that. Adaptive spatial filters such as the Frost filter [41] and Gamma MAP filter [42], wavelet-domain methods such as block-matching 3D [43] (BM3D) algorithm and total variation [44,45] (TV) methods have been developed in the literature. Thus far, several

deep-learning-based denoising methods such as SAR-CNN [46] and ID-CNN [47] have been applied and achieve a great restoration. Deep-learning-based methods have the advantage of adapting to complicated data without estimating obscure model parameters, and incorporating SAR priors can accelerate training and enhance performance effectively.

Although the SAR SR and denoising algorithms have excellent performance in HR image restoration, they still have some shortcomings in three aspects. First, deep-learning-based SAR images SR and denoising methods are separated, and these SR methods hardly consider multiplicative speckle noise and blur since they may pay more attention to SR process instead of noise. SAR noise is so complicated that it is hard to introduce SAR noise model priors into deep learning architectures. Most SAR SR methods only consider the classical additive Gaussian noise model, which may cause bad results in real SAR images with speckle noise. N. Karimi's work [48] is one of the first works to combine an SAR image SR model and multiplicative speckle noise model, even so, it is still not a deep-learning-based method. Second, the degradation models are determined in most methods. Actually, it is not appropriate to regard SAR degradation as a specific model due to its complex and uncertain SAR imaging systems. Meanwhile, the degradation model and the estimation parameters vary from different data sources, which requires strong adaptability of the method. Third, the training data sources severely limit the performance of the method. Most existing SAR SR models are trained with synthetic datasets in which the LR images are bicubic down-sampled from their HR counterparts. Additionally, most existing learning-based methods adopt a simple degradation model when formulating the SR datasets. However, these models become less effective when applied to real-world scenarios due to the domain gap between the synthetic and real LR SAR images. Some methods even adopt the grayscale optical images and superimposed noise artificially to simulate the realistic scenes, which may only learn the degradation processes they defined instead of various practical relationships between LR and HR images with implicit information.

In this paper, we proposed a cycle-GAN-based blind SR method for SAR images with speckle noise by learning the degradation model and introducing SAR priors to solve the above problems. The contributions of this paper are listed as follows:

(1) SAR priors. To the best of our knowledge, our method was one of the first SAR image SR methods based on deep learning which introduced statistical properties of speckle noise such as Gamma distribution initialization.

(2) A learnable probabilistic degradation model. Inspired by PDM-SR [29], we introduced a learnable probabilistic degradation model instead of a determined degradation model for blind SAR SR tasks, and we modified the architecture in terms of SAR noise prior criterion, which is conducive to satisfy various SAR images.

(3) Unified cycle GAN framework. We trained the degradation model and SR model simultaneously in a unified cycle GAN framework to learn the intrinsic relationship between the HR-LR domains. Additionally, we trained the model with different resolution real SAR images completely, instead of using synthetic images, to conquer the domain gap.

(4) Experimental results. Results of both synthetic and realistic SAR images with various levels of speckle noise demonstrated the high performance of the proposed method in terms of both image quality and visual perception. The results also showed that SR and denoise tasks can be well realized simultaneously. We found tremendous potential for target detection tasks due to its powerful capabilities for exposing targets by generating the details and eliminating the noise, which can reduce missed detection and false alarms significantly. The proposed method can improve mAP and reduce the training epochs effectively.

The remainder of this paper is structured as follows: In Section 2, we introduce the proposed method. Section 3 presents the experimental results and evaluation. Section 4 presents the conclusion of the proposed method.

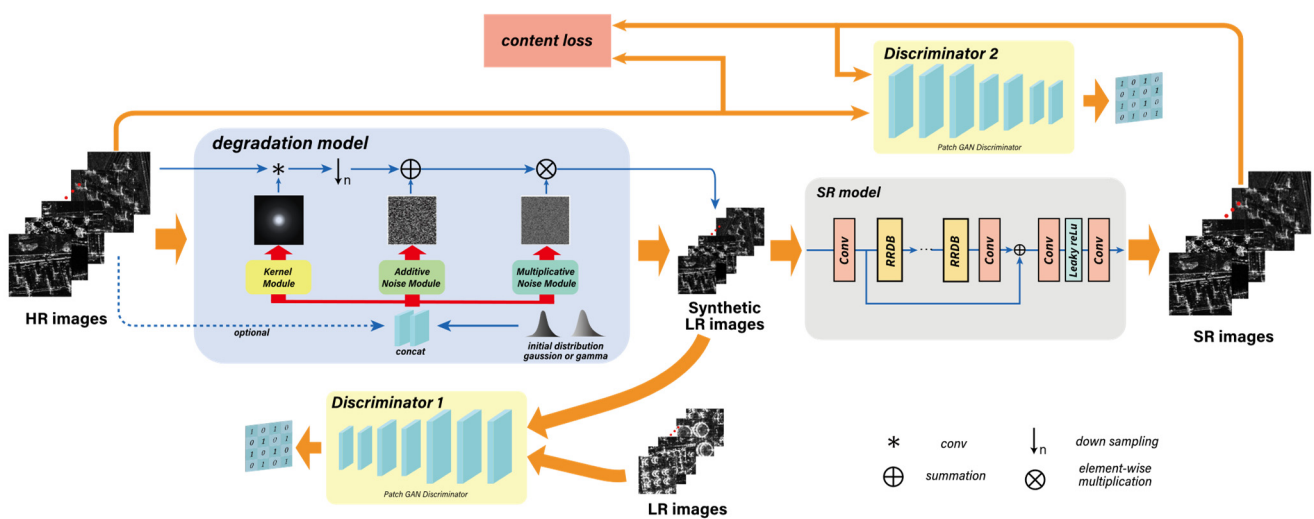
## 2. Methodology

In this section, we provide the details of the proposed method. The proposed method is a GAN-based blind SR method for SAR images with speckle by learning the degradation model and introducing SAR priors. The model consists of a learnable degradation model and an SR model in a cycle-GAN [49] framework to conquer the domain gap. A probabilistic degradation model is adopted to fit the diverse distributions of SAR images. A kernel module and a noise module are introduced to rebuild the degradation. The SAR noise model priors are integrated into the degradation processes. The SR module is designed as an RRDBnet. Finally, adversarial loss and content loss are combined to train the whole network simultaneously. Real SAR images that have different resolutions are adopted as HR and LR images to learn the intrinsic relationship between them. These parts are explained in detail below.

### 2.1. Model Framework

The model framework is a cycle-GAN-based model. GAN is a significant generative network architecture where a generator network and a discriminator network are trained simultaneously. The former is trained to gain samples that are similar to the target domain; meanwhile, the latter is trained to determine the authenticity of the generated samples. Cycle-GAN is an efficient way to resolve the unavailable paired data problem, which can learn better from a source domain to a target domain by introducing another GAN network with a cycle consistency loss. In this work, the HR images ( $I^H$ ) and the LR images ( $I^L$ ) are regarded as two different domains that have distinctive features, and our goal is to learn a degradation process  $Deg : I^H \rightarrow I^L$  and an SR process  $SR : I^L \rightarrow I^H$  simultaneously to ensure the good consistent performance of the whole cycle-GAN based SR model.

The model framework is shown in Figure 1, which consists of a degradation model (D) and an SR model (S), D focuses on learning the degradation process and S is to recover the HR images from the generated LR images. To make sure that the proposed method is able to conquer the domain gap between LR and HR caused by different imaging formations from different imaging platforms, we encourage the D to learn the intrinsic relationship between LR and HR to generate synthetic LR images similar to the real LR images, then, the synthetic LR images carried with real LR characteristic are restored to the initial real HR images, which proves that the SR model has the capacity to give more details than the real LR images.



**Figure 1.** Overall framework of the model. The whole model consists of a degradation model and an SR model in a cycle-GAN framework. Two discriminators are used to supervise the two generators.



We designed two discriminator networks to supervise the corresponding networks respectively. The objective function of the network can be formulized as min-max problems in (1) and (2).

$$\min_{SR} \max_{D_H} V(SR, D_H) = \mathbb{E}_{I^H \sim P_H(I^H)} [\log(D_H(I^H))] + \mathbb{E}_{I^L \sim P_L(I^L)} [\log(1 - D_H(SR(I^L)))] \tag{1}$$

$$\min_{Deg, D_L} V(Deg, D_L) = \mathbb{E}_{I^L \sim P_L(I^L)} [\log(D_L(I^L))] + \mathbb{E}_{I^H \sim P_H(I^H)} [\log(1 - D_L(Deg(I^H)))] \tag{2}$$

where  $D_H$  and  $D_L$  denote the discriminator 2 and the discriminator 1, as depicted in Figure 1.

Finally, we introduce a consistency loss to measure the distance between  $SR(Deg(I^H))$  and  $I^H$ , which ensures the reconstruction ability of the model, i.e.,  $I^H \rightarrow Deg(I^H) \rightarrow SR(Deg(I^H)) \approx I^H$ , therefore, we design a target function in (3)

$$\operatorname{argmin}_{Deg, SR} \left\| SR(Deg(I^H)) - I^H \right\| \tag{3}$$

The workflow of the model is shown in Figure 1, where the blue box and the gray box are the degradation model and the SR model, respectively, which can be regarded as stage 1 and stage 2. In the first stage, the degradation model extracts the degradation features adjust to the HR images by learning three modules. Then, synthetic LR images can be obtained after operation of the learned modules. In stage 2, the SR model is built to generate SR images from synthetic LR images. RRDB blocks and residual connections are introduced to strengthen the capacity of the model. Finally, the two models are trained simultaneously by adversarial training; the two yellow boxes are the discriminators. Discriminator 1 will distinguish the real LR images from synthetic LR images and discriminator 2 will distinguish the recovered SR images from the original real HR images. To ensure the similarity between HR and SR, we introduce content loss to restrain the whole process.

### 2.2. Probabilistic Degradation Model

Inspired by PDM-SR, we introduce a probabilistic degradation model with SAR characteristics into the learnable architecture. The partial input of the probabilistic model can be initialized as a specific distribution in each training epoch to model the random factors in the degradations, which can improve the generalization to adapt to the diversity of the SAR images. The probabilistic degradation model consists of a kernel module, an additive noise module and a multiplicative noise module, where we introduce the SAR noise and blur priors by reconstructing the architecture, and the whole degradation model is described in Figure 2. The inputs of these modules are initialized in different distributions according to the characteristics of each part, and the HR images are optional to be concatenated with the initial distribution as the input of each module to guarantee that the modules are fit to the input images. The forward processes are designed on the base of the real SAR image degradation model.

To make sure the SAR degradation priors can be introduced properly, the characteristics of the real SAR images are taken into account. We introduced the SAR speckle noise statistical property into the degradation based on the conventional degradation form to guide the degradation model to learn the gap between LR and SR so that the following SR model can recognize and remove the SAR noise in the process of super resolution.

Generally, the degradation process can be formulized as an equation of (4)

$$D(x) = (x \otimes k) \downarrow_s + n \tag{4}$$

where  $D(x)$  denotes the degradation function,  $x$  denotes the HR images,  $\otimes$  stands for convolution,  $k$  and  $n$  mean blur kernels and noise and  $\downarrow_s$  means down-sampling in  $s$  scale. The main problem is to estimate the  $k$  and  $n$ . We can regard the distribution of  $D$  as the joint distribution of  $k$  and  $n$ . Therefore, we can learn the mapping from a specific distribution of

$k$  and  $n$  to the target distribution of  $D$  by initializing the original distribution of  $k$  and  $n$  in a deep-learning method. However, this requires a precondition that the distributions of  $k$  and  $n$  should be independent so that we can design a model conformed to the real imaging mechanism. Usually, the classical additive Gaussian noise and the blurring process are independent, but this theorem is not followed in SAR images due to the coherent imaging mechanism and the multiplicative speckle noise. The SAR noise model has been discussed for decades and the noise has an accepted formula to describe. Assuming that  $\mathbf{Y} \in \mathbb{R}^{W \times H}$  is the intensity of the observed image,  $\mathbf{X} \in \mathbb{R}^{W \times H}$  is the noise-free counterpart and  $\mathbf{F} \in \mathbb{R}^{W \times H}$  is the speckle noise. The noise model can be formulized in (5)

$$\mathbf{Y} = \mathbf{F} \bullet \mathbf{X} \tag{5}$$

where  $\bullet$  denotes the elementwise multiplication.  $F$  follows a Gamma distribution with unit mean and variance  $1/L$  and its probability density function is shown in (6)

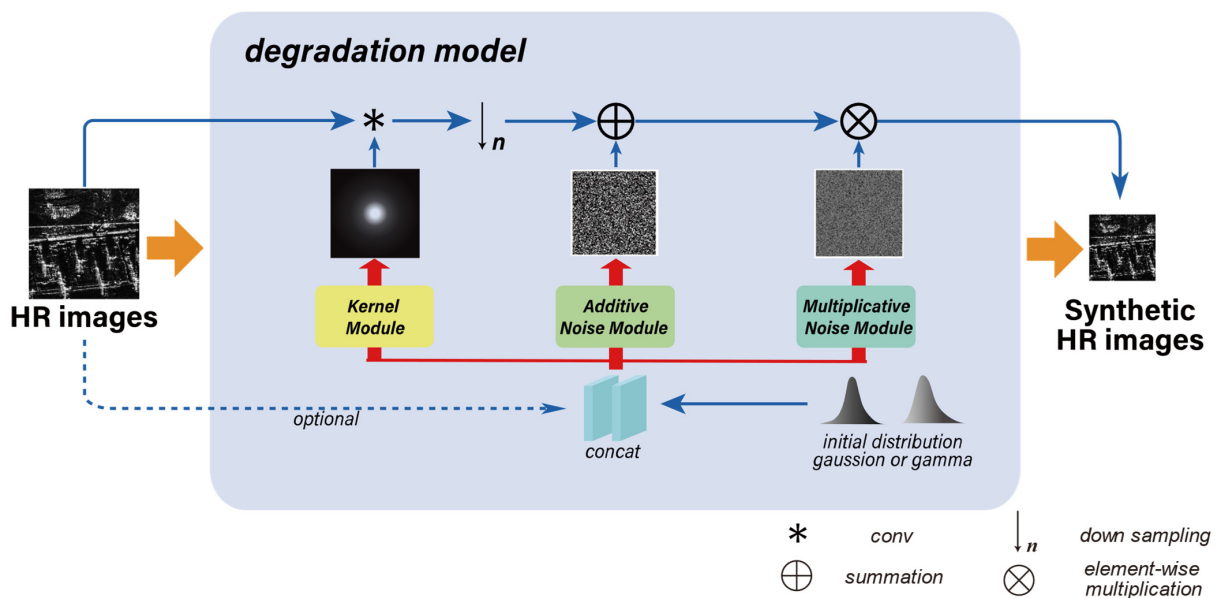
$$p(F) = \frac{1}{\Gamma(L)} L^L F^{L-1} e^{-LF} \tag{6}$$

where  $\Gamma(\cdot)$  denotes the Gamma function and  $F \geq 0, L \geq 1$  [50], and  $L$  represents the SAR images.

The ideal model mentioned above may encounter some trouble in the real SAR scenes. The signal is not only influenced by the signal-dependent multiplicative speckle noise caused by coherent imaging mechanism but also affected by the signal-independent additive fluctuation noise caused by the SAR system circuit or natural environment. Thus, we improve the degradation model for SAR images, as shown in (7),

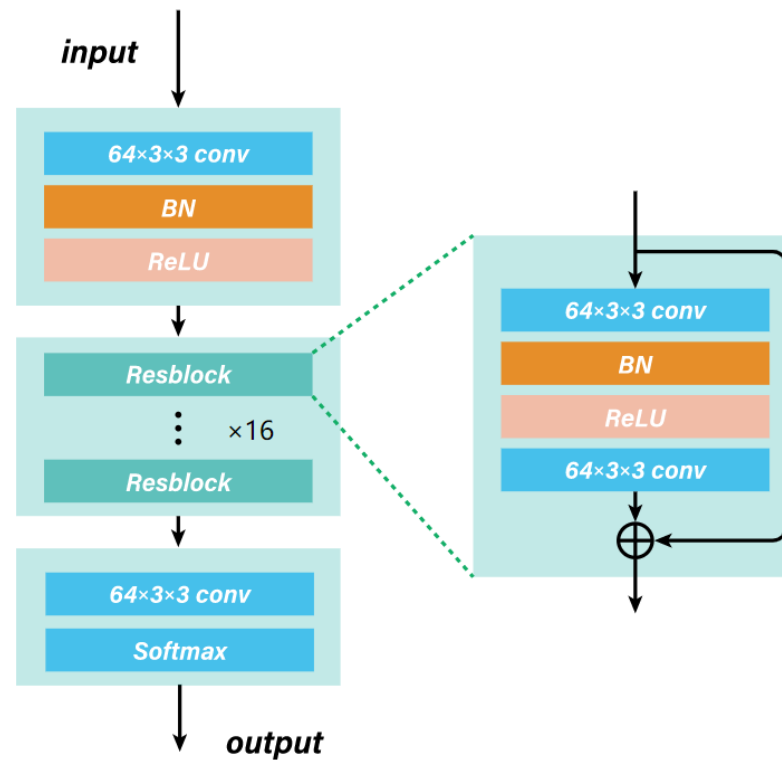
$$D(x) = [(x \otimes k) \downarrow_s + n] * n' \tag{7}$$

where  $*$  means elementwise multiplication and  $n'$  denotes multiplicative noise, which has different distribution from the additive noise. Therefore, the network can be designed under the above principle, and the model is shown in Figure 2. The whole degradation is trained in an adversarial framework, and the distribution of  $D$  can adapt to the target domain automatically. In this way, the SAR priors can be introduced in the model.



**Figure 2.** The architecture of the probabilistic degradation model. The PDM model consists of a kernel module, an additive noise module and a multiplicative noise module. These modules are organized in a specific degradation function.

The details of the modules are depicted in Figure 3. The blur kernel module and the noise module consist of head, body and tail. In the head block, a single convolution layer with  $1 \times 1$  or  $3 \times 3$  kernel, 1 stride and 64 channels make the first layer. The following is a batch normalization layer that may guarantee the stability of the model. Additionally, a ReLU activation layer is used to increase the nonlinearity. The body block is designed as 16 residual blocks with 64 channels, which are shown in Figure 3. The tail adopts a  $1 \times 1$  or  $3 \times 3$  convolution layer and a softmax layer, which is used to ensure that all the output elements sum to one.



**Figure 3.** Details of the kernel module and noise module. The module consists of head, body and tail parts.  $\oplus$  means element-wise addition.

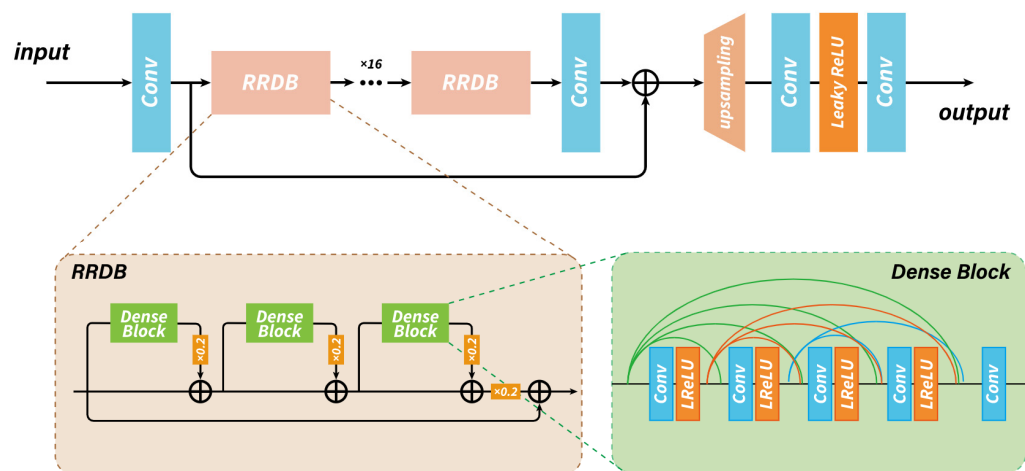
The differences between the blur kernel module and the noise module are the initial distribution of the input and the kernel sizes of the body block. We define a standardized normal distribution as the input of the blur kernel module and the additive noise module, while we define a Gamma distribution as the input of the multiplicative noise module. Particularly, the kernel size of the convolution layer is related to whether the processes are correlated spatially. The SAR signal after the range-Doppler algorithm [51] can be modeled as (8)

$$S_{ac}(\tau, t_a) = A_0 \operatorname{sinc}\left(\tau - \frac{2R_s}{c}\right) \operatorname{sinc}(t_a) \exp\left(-j4\pi \frac{R_s}{\lambda}\right) \exp(j2\pi f_{dc} t_a) \quad (8)$$

where  $A_0$  denotes the amplitude of the target signal,  $R_s$  denotes the range of the target away from the radar,  $c$  denotes the speed of light,  $\tau$  and  $t_a$  denote the fast-time and the slow-time and  $\lambda$  and  $f_{dc}$  denote the carrier wavelength and Doppler center frequency. As the equation shows, the SAR images after demodulation are made up of the sinc function at the scattering center, which is related to the adjacent pixels. Additionally, the speckle noise is spatially coherent. Therefore, the kernel size of additive noise module is  $1 \times 1$  while the blur kernel noise module and the multiplicative noise module are  $3 \times 3$ .

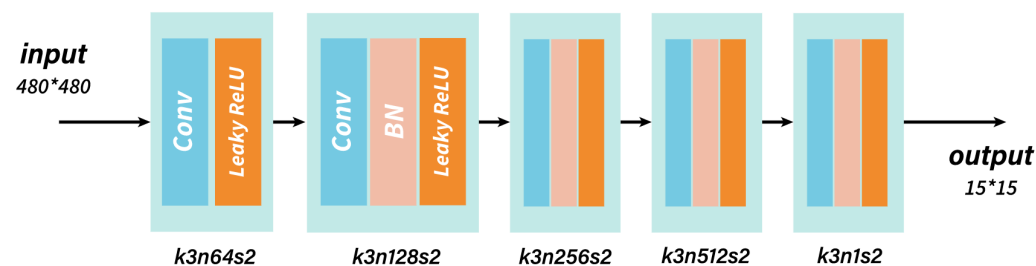
### 2.3. SR Model

Inspired by ESRGAN, the SR model is mainly composed of several residual-in-residual dense blocks (RRDBs) without BN layers, which can improve the capacity and reduce computational complexity. As depicted in Figure 4, the RRDBnet mainly consists of 16 RRDBs and each RRDB consists of three dense blocks which adopt complex connections in five convolution layers with leakyReLU activation layers, and a long skip connection is used to combine different depth features.



**Figure 4.** The architecture of RRDBnet. RRDBnet contains several RRDB; the yellow box and the green box are the details of RRDB and Dense block, respectively.  $\oplus$  means element-wise addition.

The SR model is also trained in an adversarial framework. The discriminator network employs a patchGAN discriminator, which is used in cycleGAN. The discriminator network is shown in Figure 5, and the kernel size, feature channel and the stride are shown in it. The input images are divided into several blocks that represent the discrimination result of each patch, and each picture is discriminated by the whole matrix. PatchGAN can pay attention to more areas and discriminate more accurately.



**Figure 5.** Details of the PatchGAN discriminator. k means the kernel size, n represents the number of channels and s denotes the stride.

### 2.4. Loss Function

The overall training loss is designed as the weighted combination of two adversarial losses and content loss. The two adversarial losses are designed to supervise the two generator models, i.e., the degradation model and the SR model. To guarantee the consistency between the original real HR images and the SR images after a series of processes, we introduce the content loss to add the constraints preventing from learning arbitrarily.

- Adversarial loss:

According to the min-max problem shown in (1) and (2), the adversarial loss is defined as

$$L_{GAN1}(G_1, D_Y, X, Y) = \mathbb{E}_{x \sim p_{data}(x)} [\log(1 - D_Y(G_1(x)))] + \mathbb{E}_{y \sim p_{data}(y)} [\log(D_Y(y))] \quad (9)$$

$$L_{GAN2}(G_2, D_X, Y, X) = \mathbb{E}_{y \sim p_{data}(y)} [\log(1 - D_X(G_2(y)))] + \mathbb{E}_{x \sim p_{data}(x)} [\log(D_X(x))] \quad (10)$$

where  $X$  indicates the HR images and  $Y$  indicates the LR images. In  $L_{GAN1}$ ,  $G_1$  aims to generate fake data that looks like  $Y$ , while  $D_Y$  aims to discriminate between the fake samples  $G_1(x)$  and real samples  $y$ .  $G$  tends to minimize the objective against the adversary  $D$  that tries to maximize it. The  $L_{GAN2}$  works the same.

- Content loss.

The content loss consists of  $L_1$  loss and perception loss with a regularization parameter. The content loss aims to prevent the whole model from learning arbitrarily.  $L_1$  loss measures the mean absolute error between the ground true  $y$  and generator output  $G(x)$  at the pixel level, which is formularized as

$$L_1 = \frac{1}{WH} \sum_{i,j} |G(x_{i,j}) - y_{i,j}| \quad (11)$$

where  $x$  indicates the synthetic LR images and  $y$  indicates the ground true HR images.

However, the MAE lacks information on high-level features. Thus, we introduce perception loss [52], which is defined as the feature representation distance between the real high images  $x$  and the reconstructed image  $G(y)$ . We adopt a pretrained VGG19 [53] model as the feature extractor, and the perception loss takes the form as

$$L_{percep}^{SR} = \frac{1}{WH} \sum_{i,j} (\phi_{ij}(x) - \phi_{ij}(G(y))) \quad (12)$$

where  $W$  and  $H$  indicate the width and height of the image,  $i$ , and  $j$  are the index of the feature maps and  $\phi$  denotes the feature map obtained after the fourth convolution layer and before the fifth pooling layer in the VGG19 network.

Total loss is designed as a weighted summation of adversarial loss and content loss, which is formularized as

$$L(G, F, D_X, D_Y) = L_{GAN1} + \lambda L_{GAN2} + \eta L_1 + \mu L_{percep} \quad (13)$$

where  $\lambda$ ,  $\eta$  and  $\mu$  denote the adjustable parameters which balance the weights of different losses.

### 3. Results

#### 3.1. Dataset and Training Details

To solve the blind SAR SR problem, we built an unpaired SR dataset that consists of LR and SR SAR images from real SAR products. To make a fair comparison with methods requiring paired images, we also built a paired dataset that has the same SR images as the unpaired SR dataset. The real SAR images are obtained from Terra-SAR and Gaofen-3 satellites. The details of the datasets are listed in Table 1. The original data we accessed is a level-1A (L1A) single look complex (SLC) product, and the raw data is stored in 16 bits. We stretch all the raw SAR images into 8 bits as a preprocessing step for the following calculations. The complex data are processed into grayscale amplitude images. Therefore, we mainly focus on the SR problems on grayscale amplitude images.



**Table 1.** The details of unpaired datasets and paired datasets.

Datasets	Unpaired Dataset	Paired Dataset
Train data	17,926	17,926
Test data	205	801
resolution	1 m HR & 3 m LR	1 m HR & 4 m LR
LR sources	real images	down-sampling

All experiments in this paper are carried out on a NVIDIA Telsa V100 GPU (16 G). The optimizer is Adam [54], with an initial learning rate of  $2 \times 10^{-4}$ . The batch size is set as 8. We train the model for  $1 \times 10^5$  iterations. The upscale factor is set as 4. The codes are modified based on BasicSR [55].

### 3.2. Metrics

We introduce two kinds of evaluation metrics, which are classified by whether the reference images are required. We adopt the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity [56] (SSIM) as reference metrics to value the SR performance, and we adopt the Equivalent Number of Looks [57,58] (ENL) as the no-reference metric, which requires no ground true images to value the denoising performance.

Aimed at the traditional SR problems, the metrics of PSNR and SSIM are proper to evaluate the performance. Nevertheless, for blind SR problems, these metrics are inappropriate due to the inexistence of reference images. Therefore, we evaluate the unpaired dataset mainly in terms of the visual perception and the non-reference metric ENL. PSNR, SSIM and ENL are defined in Equations (14)–(16).

$$PSNR = 10 \times \log \left[ \frac{255^2}{\frac{1}{WH} \|I^{SR} - I^{HR}\|_2^2} \right] \quad (14)$$

$$SSIM = \frac{(2\mu_{I_{hr}}\mu_{I_{sr}} + c_1)(2\sigma_{I_{hr}I_{sr}} + c_2)}{(\mu_{I_{hr}}^2 + \mu_{I_{sr}}^2 + c_1)(\sigma_{I_{hr}}^2 + \sigma_{I_{sr}}^2 + c_2)} \quad (15)$$

$$ENL = \frac{1}{n_{patch}} \sum_{i=1}^{n_{patch}} \frac{\mu_{I_{SR}i}^2}{\sigma_{I_{SR}i}^2} \quad (16)$$

where  $I^{SR}$  and  $I^{HR}$  are SR images and ground true images,  $\mu_{I_{sr}}$  and  $\mu_{I_{hr}}$  are the averages of  $I^{SR}$  and  $I^{HR}$ ,  $\sigma_{I_{sr}}^2$  and  $\sigma_{I_{hr}}^2$  are the variances of  $I^{SR}$  and  $I^{HR}$ ,  $\sigma_{I_{hr}I_{sr}}$  is the covariance between  $I^{SR}$  and  $I^{HR}$ ,  $c_1$  and  $c_2$  are two constants to maintain the stability of the equation,  $n_{patch}$  is the number of patches in one image.

### 3.3. Experiment Results

#### • Results on SR Dataset.

The proposed method is compared with the bicubic interpolation, SNGAN, Real-ESRGAN and PDM-SR. SNGAN is a GAN-based method aiming at SAR images and requires paired datasets. Real-ESRGAN and PDM-SR are blind SR methods that do not need paired datasets. To compare these methods effectively, we trained the first method with the paired dataset and trained the last two methods with the unpaired dataset. Finally, we tested the above methods with synthetic images and real SAR images, respectively.

Tables 2 and 3 lists the average PSNR, SSIM and ENL of several SR methods at a scale factor of  $\times 4$  in synthetic and real SAR datasets. As Table 2 listed, the proposed method achieves the best performance on the metrics of ENL and shows the second-best result on PSNR and SSIM, except for the bicubic interpolation method. The bicubic interpolation method gets good results on PSNR and SSIM because the input LR images are bicubic down-sampled from the reference HR images, however, the better performance on metrics cannot prove that they would be better on visual perception, as shown in Figures 6 and 7.

SNGAN is trained with synthetic paired data and shows worse performance than the other methods when compared to SSIM, which reveals the domain gap between the real and synthetic SAR images. Real ESRGAN gets better results due to overcoming the domain gap by learning with real unpaired SAR images. Real ESRGAN pays more attention to the degradations such as blur, resize, noise and JPEG compression, and it adopts a high-order degradation model. However, if these degradation model designs are not suitable for the SAR mechanism, especially for the SAR noise mechanism, they may not obtain great results on real SAR images in terms of ENL. Thus, Real ESRGAN shows better performance on reconstruction metrics but worse on the noise intensity metric. The PDM SR adopts a probabilistic degradation model on noise module design, which pays more attention to estimating the noise in a probabilistic way, thus, it gains great performance on ENL second to the proposed method. The proposed method introduces several SAR noise priors into the degradation model in a probabilistic way and demonstrates the best performance on denoising in terms of ENL.

**Table 2.** Valuation results of synthetic SAR images dataset.

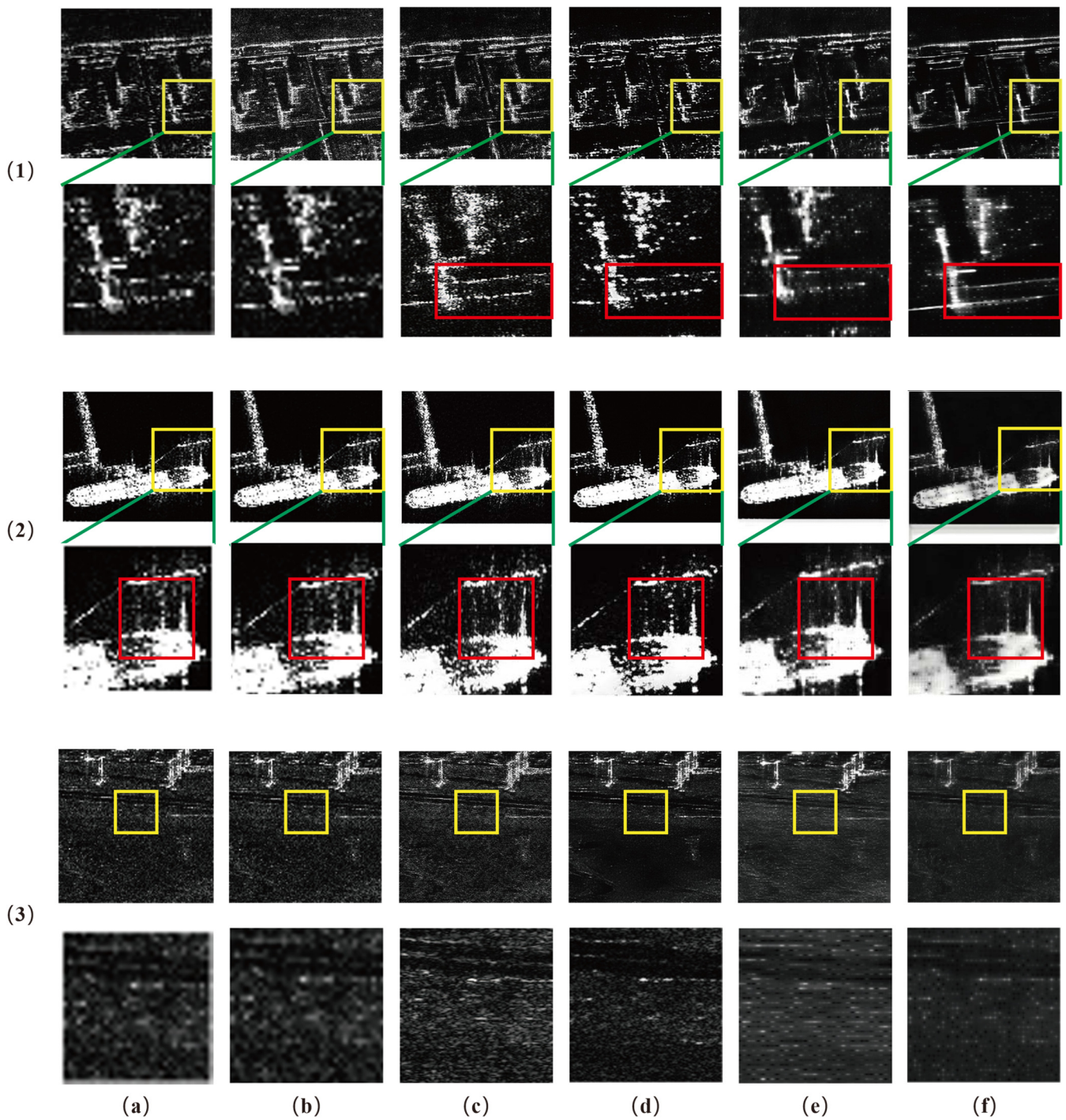
Metrics	Real	Up Bicubic	SNGAN	Real ESRGAN	PDM SR	Proposed
PSNR	-	20.805	18.773	19.544	16.126	19.254
SSIM	-	0.489	0.329	0.394	0.309	0.362
ENL	0.906	1.044	0.981	0.841	1.420	2.295

**Table 3.** Valuation results of real SAR images dataset.

Metrics	Real	Up Bicubic	SNGAN	Real ESRGAN	Proposed
ENL	1.332	1.401	1.245	0.920	1.907

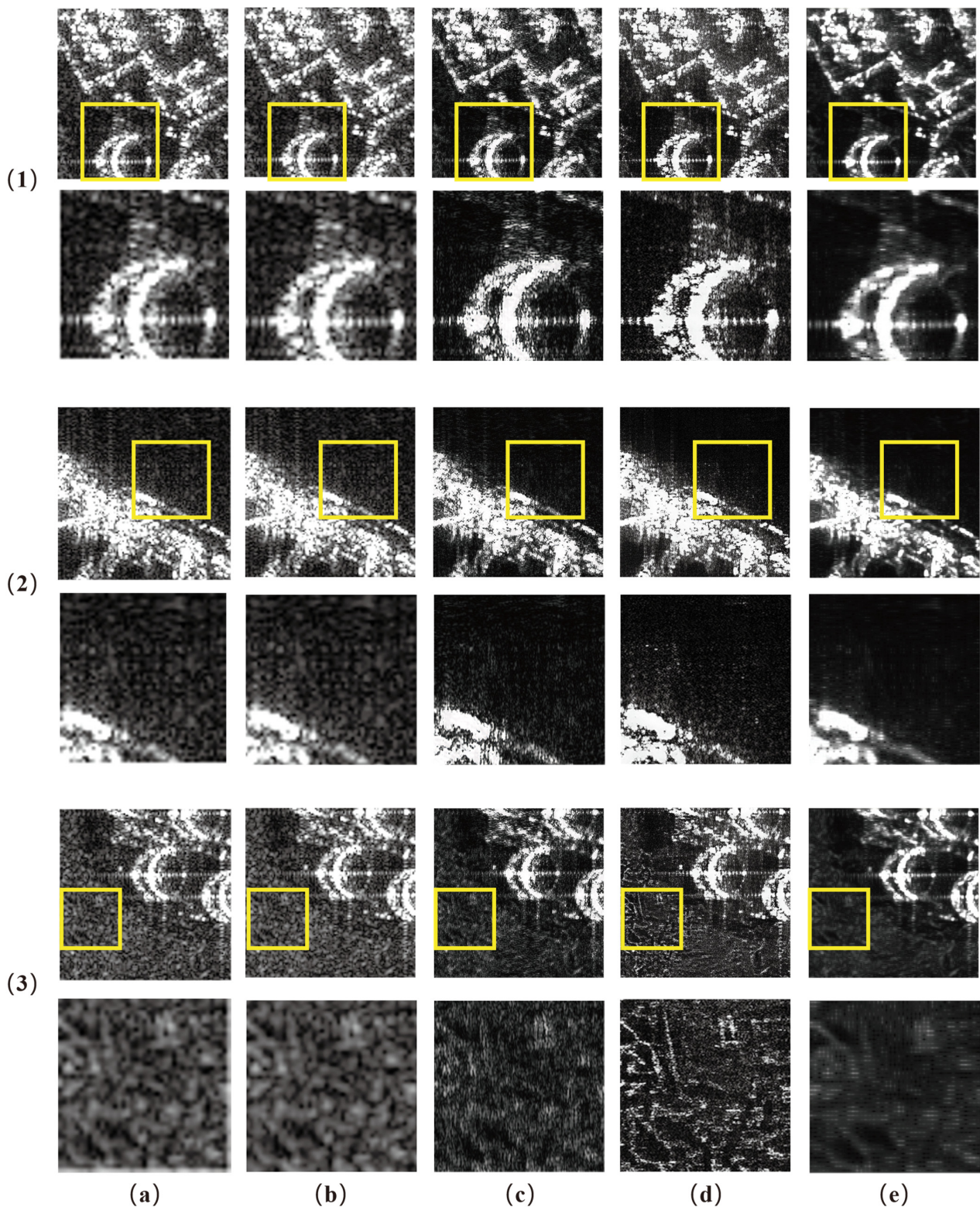
Our proposed method is a joint super-resolution reconstruction and despeckling co-processing, which inherited the advantages of both tasks. The metric of ENL represents the noise level and the other methods do not take full account of the noise problems, therefore, the proposed method with the denoising function gets the best score in ENL. The input images of the synthetic paired dataset are obtained from the corresponding real high-resolution images in a bicubic down-sampling way, and some of the methods we compared are trained with the synthetic data in a supervised training strategy, while the proposed method is aimed at the real scene SAR images and is trained with real SAR images all in an unsupervised way, which means that the former methods have introduced the bicubic deterministic degradation model in the training process, while the proposed method learns the probabilistic implicit degradation mode between real LR and HR SAR images that follows, definitely not the bicubic mode. The degradation modes that the two kinds of methods learned are different. The synthetic paired dataset fits the methods trained with bicubic mode data and they will get better scores on PSNR and SSIM because the forward process is just the reverse process of the bicubic down-sampling. However, it is not good for the blind SR problem because the degradation mode in the real scene is definitely not bicubic mode. Therefore, the proposed method may not get better performance in the synthetic dataset in which the LR images are bicubic down-sampled from the corresponding HR images.

However, the SR problem is ill-posed, and there is no sole criterion to evaluate the SR performance. Additionally, it is unfair to judge the blind SR method by the reference metrics because the other methods used the reference HR images in the training process, while the blind SR methods did not. Therefore, we evaluate the SR results comprehensively by combining visual perception and metrics. When compared with other methods using real SAR image datasets, the proposed method also shows the best performance on ENL.



**Figure 6.** Visual results of synthetic SAR image dataset. The second row is the enlarged view of the first row in each of three paired pictures. (a) Input image (bicubic down-sampling from the HR image). (b) Bicubic interpolation. (c) SNGAN. (d) Real ESRGAN. (e) PDM-SR. (f) The proposed method. (1)–(3) represent three pairs of examples to illustrate the performance. The second row of each pair of examples is the enlarged view of the corresponding yellow box in the first row. The red boxes show the outstanding performance of the proposed method in recovering texture and denoising.





**Figure 7.** Visual results of real SAR image dataset. The second row is the enlarged view of the first row in each of three paired pictures. (a) Input image (real LR image). (b) Bicubic interpolation. (c) SNGAN. (d) Real ESRGAN. (e) The proposed method. (1)–(3) represent three pairs of examples to illustrate the performance. The second row of each pair of examples is the enlarged view of the corresponding yellow box in the first row.

To further demonstrate the advantages of our method, we select several visualization results shown in Figures 6 and 7. Figure 6 shows the SR result of several typical targets such as buildings, flat surfaces and ships in the synthetic SAR dataset. In order to compare the details of SR performance, we provide enlarged views of the areas marked by yellow rectangles. From the perspective of image visual effects, it can be found that the proposed method shows powerful capability both in denoising and SR. In Figure 6, we present three sets of pictures processed by different methods, as shown in each column, and the areas in the yellow boxes are enlarged for more details visually, as shown in the second row. Compared to interpolation and SNGAN, which require paired datasets, the buildings processed by the proposed method in Figure 6 are exposed obviously due to the suppression of backing noise, which demonstrates that the proposed method tends to improve the denoising of SR images instead of the SR counterpart, similar to the real images. At the same time, the texture of buildings and roads can be restored well by the proposed method. Real ESRGAN also gains great denoising performance, but some of the results show artifacts and generate some weird patterns that are incompatible with the real scenes. In this aspect, the proposed method can restore continuous lines from corrupt images without artificial patterns, as shown in the red box in the second row, which is conducive to subsequent image interpretation. In the second example, the proposed method can suppress the noise near the strong scattering target, and it can retain the target scattering elements when suppressing cross bright spots, which may resolve the cross bright spots of SAR images caused by signal processing of pulse compression. In the third example, the original image has plenty of noise in flat areas, which can be suppressed after the proposed SR process. Its enlarged parts shown in the second row demonstrate that the proposed method can suppress the granularity noise and retain the texture.

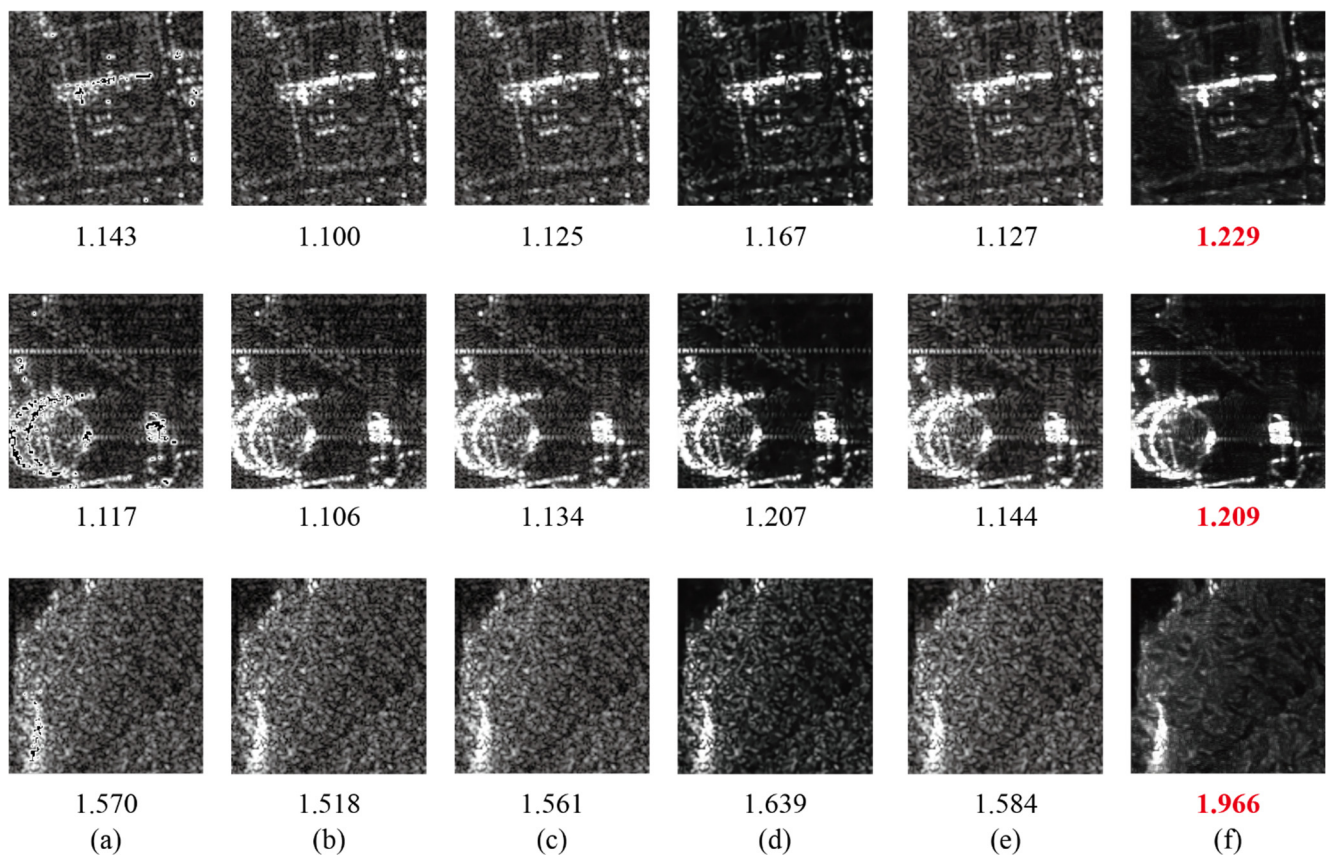
In Figure 7, we select several scenes from real SAR images to illustrate the generalization capability when processing real scene images. The proposed method can also reduce the impact of sinc bright spots caused by strong scattering targets. As shown in the first example, the oil tanks suffer from noise near the strong scattering area, like the two-dimensional sinc function, and the proposed method can suppress the sinc noise by suppressing the side lobes. In the second example, the proposed method also generates the image with the least noise in the areas of the target and the backing ground. Meanwhile, GAN-based methods usually produce artifacts and nonexistent texture due to the arbitrariness of the generator. As depicted in the third example, SNGAN and real-ESRGAN generate some weird texture near the oil tanks while the proposed method suppresses the speckle noise without artifacts. In brief, the proposed method can retain the texture of original images and suppress noise effectively.

- Despeckling Evaluation

To further illustrate the despeckling effect, we compare our method with different kinds of despeckling methods. The spatial domain filtering method used by Kuan [59] and Frost [41], the non-local mean method NLM [60] and SAR-BM3D [61], and the deep-learning-based method DnCNN [16] are implemented for comparison with the proposed method. For a fair comparison, the LR images are restored by a basic bicubic upsampling process. Figure 8 shows the best performance of the proposed method among the mainstream despeckling methods in terms of ENL.

To illustrate the stability for different intensities of noise, we select the real scene SAR images with noise of different intensities. As Figure 9 depicted, the SAR images with the noise of lower intensities (higher ENL scores) can get better despeckling performance, while the images with lower ENL scores may not get too much improvements in terms of ENL scores due to the inherent severe noise. From the perspective of visual effects, the processed images also get great despeckling performance according to the row b in Figure 9.





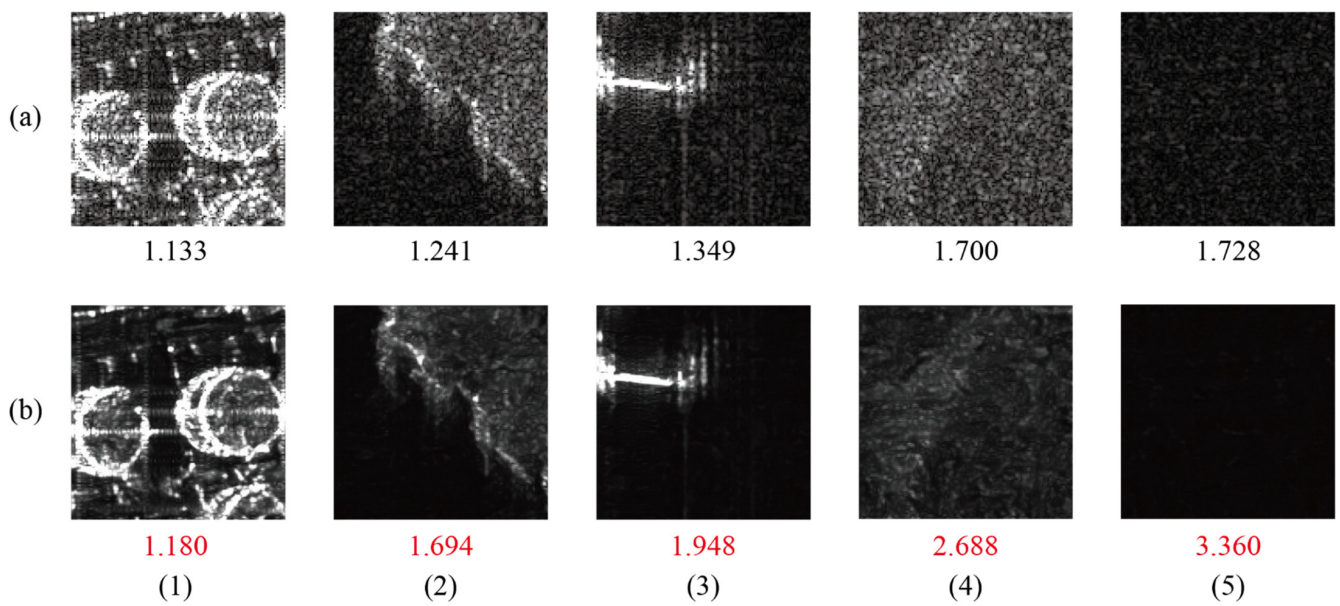
**Figure 8.** The despeckling results of different despeckling methods. (a) Kuan. (b) NLM (c) Frost. (d) DnCNN. (e) SAR-BM3D. (f) The proposed method. The scores below the images are ENLs. The red bold format represents the best scores.

- Target Detection Evaluation

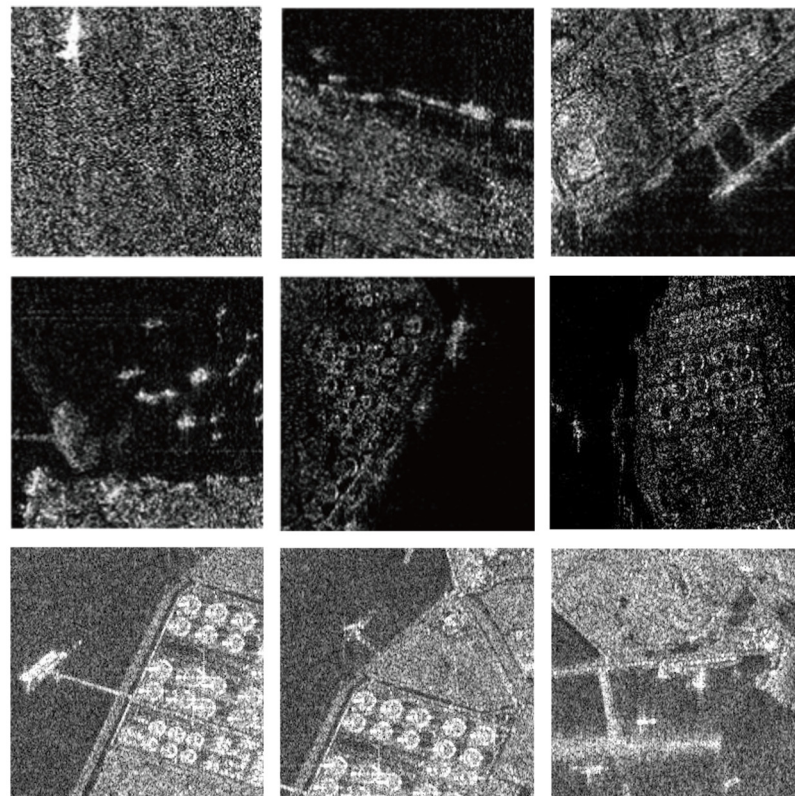
Additionally, we discover an enormous potential for detection due to the powerful capabilities for exposing targets by generating the details and eliminating the noise. The proposed method can improve mAP and reduce the training epochs effectively, which indicates that the SR methods can reduce missed detection and false alarms significantly. We train a target detection algorithm based on yolov5 with and without the SR preprocessing, and the other parts of the algorithm are set identically. We train the model with the MSAR [62] dataset, which contains 28,449 SAR images of ships, bridges, planes and oil tanks. Table 4 shows the detail of the MSAR dataset. MSAR contains more categories than other mainstream SAR target detection datasets, and it remains challenging on severe noise and unbalanced samples, as shown in Figure 10.

**Table 4.** Details of MSAR dataset.

Category	Indicator
Scenes	HISEA-1
Polarization	HH, VV, HV, VH
Size (pixel)	256 × 256, 2048 × 2048
Number of pictures	28,449
Number of ships	39,858
Number of oil tanks	12,319
Number of aircraft	6368
Number of bridges	1851



**Figure 9.** The despeckling results for noise of different intensities. (a) the input images with noise of different intensities. (b) images processed by the proposed method. (1) to (5) represent the examples with the noise of different intensities. The scores below the images are ENLs. The red format represents the best scores.



**Figure 10.** Illustration of the noise challenging of MSAR dataset.

The metrics to measure the detection effect mainly include accuracy, recall and average accuracy ( $mAP$ ), which refer to the rate of correctly recognized samples in all positively

detected samples, the rate of correctly identified samples in all ground truth samples, and comprehensive metric calculated by  $P$  and  $R$ , respectively. The definitions are as follows:

$$P = \frac{TP}{TP + FP} \tag{17}$$

$$R = \frac{TP}{TP + FN} \tag{18}$$

$$mAP = \int_0^1 P(P)dR \tag{19}$$

where  $TP$ ,  $FP$  and  $FN$  refer to positive samples predicted as positive, negative samples predicted as positive and negative samples predicted as negative, respectively.

We modify the yolov5 with the proposed SR processing, which is used as pre-processing before the whole detection mode. The SR processing brings a huge improvement to detection. Figure 11 shows the training curve of the above training method and Tables 5 and 6 show the comparison indicators of each target. From the training curve, the SR process improves mAP to 0.838 within less than 100 epochs, while the baseline can only reach 0.687 within 500 epochs, which indicates that the proposed SR process can improve the efficiency and accuracy of the target detection task significantly. In Tables 5 and 6, we can also see that the metrics of each category are improving to varying degrees, while the decline of mAP in the bridge category is mainly caused by unbalanced quantity. The substantial improvement is mainly caused by extra information from the SR process, which provides effective information based on LR images and eliminates the interference of noise. The powerful capabilities for exposing targets by generating the details and eliminating the noise show great potential for detection tasks by significantly reducing missed detection and false alarms.

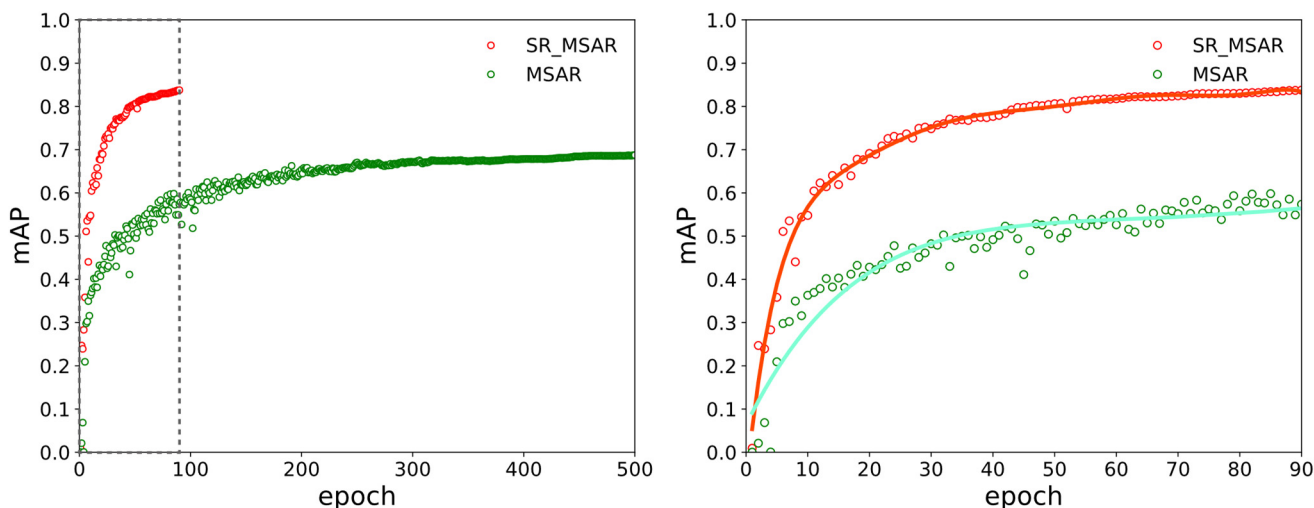


Figure 11. Comparison of yolov5 with and without proposed SR processing. The zoomed gray box is illustrated on the right diagram. The orange line and the cyan line are the fitted curves of the points.

Table 5. Results of yolov5 in MSAR dataset.

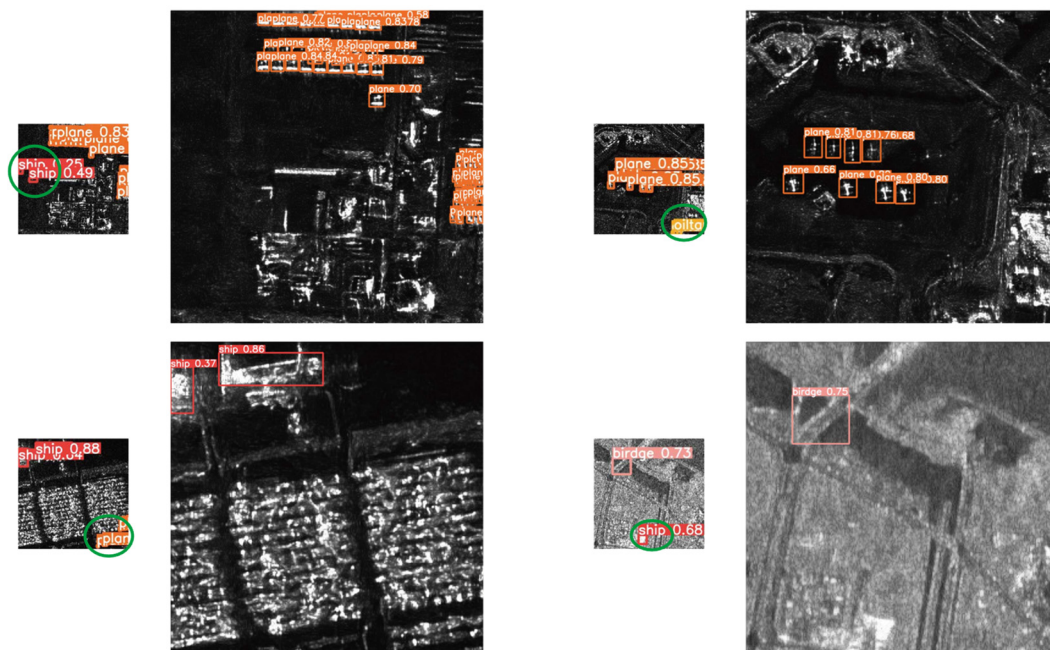
Class Images	P	R	mAP_0.5	mAP_0.5:0.95
All	0.848	0.654	0.687	0.408
Ship	0.844	0.891	0.916	0.562
Bridge	0.893	0.669	0.718	0.417
Plane	0.758	0.364	0.399	0.137
Oil tank	0.896	0.692	0.715	0.515



**Table 6.** Results of SR modified yolov5 in MSAR dataset.

Class Images	P	R	mAP_0.5	mAP_0.5:0.95
All	0.862	0.766	0.838	0.498
Ship	0.87	0.937	0.95	0.647
Bridge	0.884	0.254	0.566	0.232
Plane	0.784	0.895	0.855	0.348
Oil tank	0.908	0.979	0.98	0.765

To further illustrate how the SR process improves the performance of detection by reducing missed detection and false alarms, we select several examples of the detection results under the two methods, as shown in Figures 12 and 13. It is clear that the results with SR methods have less missed detection and false alarms. The LR images contain plenty of strong local noise, which can easily be falsely recognized as targets. As shown in Figure 12, several ships and planes are falsely detected on land. Additionally, the heavy noise causes the targets to be drowned out, as shown in Figure 11; an oil tank and several ships are missed due to the heavy noise. The SR images processed by the proposed method have lower noise and clearer targets, which is conducive to various interpretations.

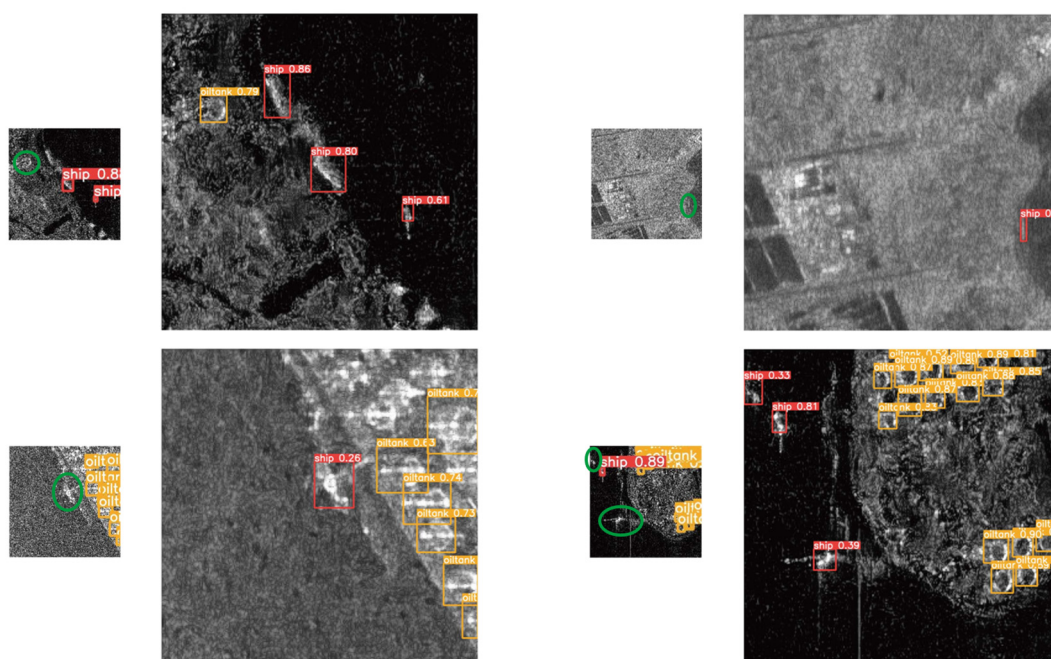


**Figure 12.** Illustration of false alarms. Green circles mark fault detection. The left is the detection results without SR process and the right is processed in SR at a factor of 4. Different colors means different kinds of targets.

To compare the proposed method with other detection methods in this challenging dataset, we choose P, R, mAP and time for each image as indicators. Table 7 shows the performance of these methods, the first three methods trained with MSAR dataset are carried out in [62]. The proposed method improves the recognition accuracy greatly with little increase in time consumption.

**Table 7.** Comparison with the latest target detection methods.

Class Images	mAP_0.5
RetinaNet [63]	0.562
FCOS [64]	0.577
Yolov5	0.687
Proposed	0.837



**Figure 13.** Illustration of missing detection. Green circles mark the position of missing detection. The left insets are the detection results without the SR process. Different colors means different kinds of targets.

### 3.4. Ablation Results

To illustrate the effectiveness of the proposed method, we design ablation experiments with and without each part of the proposed method. We mainly explore the effectiveness of training data (paired or unpaired), SR model (EDSR or RRDBnet), degradation model (original or improved PDM), kernel design (whether adjust the kernel size of convolution layer) and loss (with or without perception loss). The details of the ablation results are listed in Table 8, in which, the highest score is marked in bold red and the second highest score is marked in bold black.

Through five sets of comparisons, the proposed method does not achieve the highest accuracy in every index at the same time, but achieves the best results in balancing each index and visual perception, and gets the second-best score. For training data, the adoption of unpaired data will not improve evaluation indicators such as PSNR and SSIM that require reference images, but will tend to find internal connections, which is reflected in improving ENL. For the degradation model, the improved PDM can greatly improve the image quality due to the introduction of SAR prior information into the model, which can make SR better while maintaining the same content of the training data. For the SR model, RRDBnet is an improvement of the super-resolution network, which has a stronger learning ability; therefore, it can get better results. For the kernel design, when adopting the improved kernel module design according to the noise and image characteristics without adding the perception loss, although the ENL is greatly improved and the noise is greatly reduced, the quality of the image structure cannot be guaranteed, which is reflected in the blurring of the image and the thickening of the lines. The perception loss is to prevent the GAN model from learning arbitrarily, so that the visual effect of the images can be closer to the results that people accept most; therefore, the metrics show good improvement when compared to the baseline method.



**Table 8.** Influence of training data on experimental results.

SR Model	Deg Model	Kernel Design	Perception Loss	Training Data	PSNR	SSIM	ENL
EDSR	Original	×	×	Paired	17.440489	0.332912	1.241069
EDSR	Original	×	×	unpaired	16.125568	0.309423	1.419914
EDSR	Improved	×	×	unpaired	19.224419	<b>0.375096</b>	1.285013
RRDBnet	Improved	×	×	unpaired	<b>19.930641</b>	<b>0.416975</b>	1.701983
RRDBnet	Improved	✓	×	unpaired	17.426025	0.275015	<b>3.937221</b>
RRDBnet	Improved	✓	✓	unpaired	<b>19.254030</b>	0.361569	<b>2.295177</b>

#### 4. Conclusions

In this paper, we proposed a blind SR method for SAR images by introducing SAR priors in a cycle-GAN framework, which conquered the domain gap caused by severe speckle noise and low-resolution problems and provided great assistance for subsequent image interpretation and target detection. First, a learnable probabilistic degradation model combined with statistical properties of SAR priors noise was presented to satisfy various situations. Furthermore, we trained the degradation model and SR model simultaneously in a unified cycle GAN framework to learn the intrinsic relationship between HR-LR domains. Additionally, we trained the model with real SAR images instead of synthetic images to conquer the domain gap. Finally, experimental results on both synthetic and real SAR images demonstrated the high performance of the proposed method in terms of image quality and visual perception. Experimental results on both synthetic and realistic SAR images with various levels of speckle noise demonstrated the high performance of the proposed method in terms of both image quality and visual perception. The results also showed that SR and denoise tasks can be well realized simultaneously. Additionally, we found a tremendous potential for target detection tasks by significantly reducing missed detection and false alarms due to its powerful capabilities for exposing targets by generating target details and eliminating noise. The proposed method can effectively improve mAP and reduce the training epochs. In the future, we will try adding more SAR priors into deep-learning-based methods, including SAR statistical properties and the SAR imaging mechanism. In this work, we also found that the method based on GAN with the proposed multiplicative noise module may make the training unstable, resulting in difficult convergence after several training epochs. Therefore, we will try to add more constraints to the model to improve its robustness.

**Author Contributions:** Conceptualization, C.Z. and Y.T.; methodology, C.Z.; software, C.Z.; validation, C.Z. and Z.Z.; formal analysis, C.Z.; investigation, C.Z. and Y.D.; resources, Y.T.; data curation, C.Z.; writing—original draft preparation, C.Z.; writing—review and editing, Z.Z., Y.D., Y.Z., M.C. and Y.T.; visualization, Z.Z.; supervision, Z.Z. and P.L.; project administration, C.Z.; funding acquisition, Y.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Natural Science Foundation of China, grant number 61991423.

**Data Availability Statement:** The data presented in this study are available on request from the first author.

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

#### References

1. Yue, L.; Shen, H.; Li, J.; Yuan, Q.; Zhang, H.; Zhang, L. Image super-resolution: The techniques, applications, and future. *Signal Process.* **2016**, *128*, 389–408. [\[CrossRef\]](#)
2. Yang, C.-Y.; Ma, C.; Yang, M.-H. Single-image super-resolution: A benchmark. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 372–386.
3. Karwowska, K.; Wierzbicki, D. Using Super-Resolution Algorithms for Small Satellite Imagery: A Systematic Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 3292–3312. [\[CrossRef\]](#)

4. Marcello, J.; Ibarrola-Ulzurrun, E.; Gonzalo-Martin, C.; Chanussot, J.; Vivone, G. Assessment of hyperspectral sharpening methods for the monitoring of natural areas using multiplatform remote sensing imagery. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 8208–8222. [[CrossRef](#)]
5. Kanakaraj, S.; Nair, M.S.; Kalady, S. Adaptive importance sampling unscented Kalman filter with kernel regression for SAR image super-resolution. *IEEE Geosci. Remote Sens. Lett.* **2020**, *19*, 4004305. [[CrossRef](#)]
6. Shkvarko, Y.V.; Yañez, J.I.; Amao, J.A.; del Campo, G.D.M. Radar/SAR image resolution enhancement via unifying descriptive experiment design regularization and wavelet-domain processing. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 152–156. [[CrossRef](#)]
7. Kanakaraj, S.; Nair, M.S.; Kalady, S. Adaptive importance sampling unscented Kalman filter based SAR image super resolution. *Comput. Geosci.* **2019**, *133*, 104310. [[CrossRef](#)]
8. Biondi, F. Recovery of partially corrupted SAR images by super-resolution based on spectrum extrapolation. *IEEE Geosci. Remote Sens. Lett.* **2016**, *14*, 139–143. [[CrossRef](#)]
9. Wang, Z.-m.; Wang, W.-w. Fast and adaptive method for SAR superresolution imaging based on point scattering model and optimal basis selection. *IEEE Trans. Image Process.* **2009**, *18*, 1477–1486. [[CrossRef](#)]
10. Zhang, X.; Cao, K.; Jiao, L. A contourlet-based interpolation restoration method for super-resolution of SAR image. In Proceedings of the 2009 2nd Asian-Pacific Conference on Synthetic Aperture Radar, Xian, China, 26–30 October 2009; pp. 1068–1071.
11. Yang, J.; Wright, J.; Huang, T.S.; Ma, Y. Image super-resolution via sparse representation. *IEEE Trans. Image Process.* **2010**, *19*, 2861–2873. [[CrossRef](#)]
12. Dong, W.; Zhang, L.; Shi, G.; Wu, X. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. *IEEE Trans. Image Process.* **2011**, *20*, 1838–1857. [[CrossRef](#)]
13. He, C.; Liu, L.; Xu, L.; Liu, M.; Liao, M. Learning based compressed sensing for SAR image super-resolution. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 1272–1281. [[CrossRef](#)]
14. Kulkarni, N.; Nagesh, P.; Gowda, R.; Li, B. Understanding compressive sensing and sparse representation-based super-resolution. *IEEE Trans. Circuits Syst. Video Technol.* **2011**, *22*, 778–789. [[CrossRef](#)]
15. Karimi, N.; Taban, M.R. Nonparametric blind SAR image super resolution based on combination of the compressive sensing and sparse priors. *J. Vis. Commun. Image Represent.* **2018**, *55*, 853–865. [[CrossRef](#)]
16. 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](#)] [[PubMed](#)]
17. Fu, X.; Huang, J.; Ding, X.; Liao, Y.; Paisley, J. Clearing the skies: A deep network architecture for single-image rain removal. *IEEE Trans. Image Process.* **2017**, *26*, 2944–2956. [[CrossRef](#)] [[PubMed](#)]
18. Dong, C.; Loy, C.C.; He, K.; Tang, X. Learning a deep convolutional network for image super-resolution. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 184–199.
19. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144. [[CrossRef](#)]
20. Kim, J.; Lee, J.K.; Lee, K.M. Deeply-recursive convolutional network for image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1637–1645.
21. Kim, J.; Lee, J.K.; Lee, K.M. Accurate image super-resolution using very deep convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1646–1654.
22. Lim, B.; Son, S.; Kim, H.; Nah, S.; Mu Lee, K. Enhanced deep residual networks for single image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 136–144.
23. Dong, C.; Loy, C.C.; Tang, X. Accelerating the super-resolution convolutional neural network. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; pp. 391–407.
24. Li, Z.; Yang, J.; Liu, Z.; Yang, X.; Jeon, G.; Wu, W. Feedback network for image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 3867–3876.
25. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4681–4690.
26. Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Qiao, Y.; Change Loy, C. Esrgan: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018; pp. 1–16.
27. Wang, X.; Xie, L.; Dong, C.; Shan, Y. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 1905–1914.
28. Ji, X.; Cao, Y.; Tai, Y.; Wang, C.; Li, J.; Huang, F. Real-world super-resolution via kernel estimation and noise injection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 466–467.
29. Luo, Z.; Huang, Y.; Li, S.; Wang, L.; Tan, T. Learning the degradation distribution for blind image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 6063–6072.

30. Wu, Z.; Wang, H. Super-resolution reconstruction of SAR image based on non-local means denoising combined with BP neural network. *arXiv* **2016**, arXiv:1612.04755.
31. Wang, L.; Zheng, M.; Du, W.; Wei, M.; Li, L. Super-resolution SAR image reconstruction via generative adversarial network. In Proceedings of the 2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE), Hangzhou, China, 3–6 December 2018; pp. 1–4.
32. Zheng, C.; Jiang, X.; Zhang, Y.; Liu, X.; Yuan, B.; Li, Z. Self-normalizing generative adversarial network for super-resolution reconstruction of SAR images. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 1911–1914.
33. Shen, H.; Lin, L.; Li, J.; Yuan, Q.; Zhao, L. A residual convolutional neural network for polarimetric SAR image super-resolution. *ISPRS J. Photogramm. Remote Sens.* **2020**, *161*, 90–108. [[CrossRef](#)]
34. Yanshan, L.; Li, Z.; Fan, X.; Shifu, C. OGSRN: Optical-guided super-resolution network for SAR image. *Chin. J. Aeronaut.* **2022**, *35*, 204–219.
35. Wu, Z.; Zhao, Z.; Ma, P.; Huang, B. Real-World DEM Super-Resolution Based on Generative Adversarial Networks for Improving InSAR Topographic Phase Simulation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 8373–8385. [[CrossRef](#)]
36. Goodman, J.W. Some fundamental properties of speckle. *JOSA* **1976**, *66*, 1145–1150. [[CrossRef](#)]
37. López-Martínez, C.; Fabregas, X. Polarimetric SAR speckle noise model. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2232–2242. [[CrossRef](#)]
38. Xie, H.; Pierce, L.E.; Ulaby, F.T. Statistical properties of logarithmically transformed speckle. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 721–727. [[CrossRef](#)]
39. Simard, M.; DeGrandi, G.; Thomson, K.P.; Benie, G.B. Analysis of speckle noise contribution on wavelet decomposition of SAR images. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1953–1962. [[CrossRef](#)]
40. Lopez-Martinez, C.; Fabregas, X. Modeling and reduction of SAR interferometric phase noise in the wavelet domain. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2553–2566. [[CrossRef](#)]
41. Frost, V.S.; Stiles, J.A.; Shanmugan, K.S.; Holtzman, J.C. A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Trans. Pattern Anal. Mach. Intell.* **1982**, *PAMI-4*, 157–166. [[CrossRef](#)]
42. Lopes, A.; Nezry, E.; Touzi, R.; Laur, H. Structure detection and statistical adaptive speckle filtering in SAR images. *Int. J. Remote Sens.* **1993**, *14*, 1735–1758. [[CrossRef](#)]
43. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image denoising with block-matching and 3D filtering. In *Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning*; SPIE: Bellingham, WA, USA, 2006; pp. 354–365.
44. Ma, X.; Shen, H.; Zhao, X.; Zhang, L. SAR image despeckling by the use of variational methods with adaptive nonlocal functionals. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3421–3435. [[CrossRef](#)]
45. Bioucas-Dias, J.M.; Figueiredo, M.A. Multiplicative noise removal using variable splitting and constrained optimization. *IEEE Trans. Image Process.* **2010**, *19*, 1720–1730. [[CrossRef](#)] [[PubMed](#)]
46. 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.
47. Wang, P.; Zhang, H.; Patel, V.M. SAR image despeckling using a convolutional neural network. *IEEE Signal Process. Lett.* **2017**, *24*, 1763–1767. [[CrossRef](#)]
48. Karimi, N.; Taban, M.R. A convex variational method for super resolution of SAR image with speckle noise. *Signal Process. Image Commun.* **2021**, *90*, 116061. [[CrossRef](#)]
49. Zhu, J.-Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
50. Ulaby, F.; Dobson, M.C.; Álvarez-Pérez, J.L. *Handbook of Radar Scattering Statistics for Terrain*; Artech House: London, UK, 2019.
51. Bamler, R. A comparison of range-Doppler and wavenumber domain SAR focusing algorithms. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 706–713. [[CrossRef](#)]
52. Johnson, J.; Alahi, A.; Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; pp. 694–711.
53. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
54. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
55. Wang, X.; Yu, K.; Chan, K.C.; Dong, C.; Loy, C.C. *BasicSR: Open Source Image and Video Restoration Toolbox*; GitHub: San Francisco, CA, USA, 2018.
56. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]
57. Lee, J.-S. Speckle analysis and smoothing of synthetic aperture radar images. *Comput. Graph. Image Process.* **1981**, *17*, 24–32. [[CrossRef](#)]
58. Oliver, C.; Quegan, S. *Understanding Synthetic Aperture Radar Images*; SciTech Publishing: Raleigh, NC, USA, 2004.
59. 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](#)]

60. Buades, A.; Coll, B.; Morel, J.-M. A non-local algorithm for image denoising. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–26 June 2005; pp. 60–65.
61. 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.* **2011**, *50*, 606–616. [[CrossRef](#)]
62. Xia, R.; Chen, J.; Huang, Z.; Wan, H.; Wu, B.; Sun, L.; Yao, B.; Xiang, H.; Xing, M. CRTransSar: A Visual Transformer Based on Contextual Joint Representation Learning for SAR Ship Detection. *Remote Sens.* **2022**, *14*, 1488. [[CrossRef](#)]
63. Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2980–2988.
64. Tian, Z.; Shen, C.; Chen, H.; He, T. Fcos: Fully convolutional one-stage object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 9627–9636.

**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.