



Article

Sparse SAR Imaging Based on Non-Local Asymmetric Pixel-Shuffle Blind Spot Network

Yao Zhao ¹, Decheng Xiao ¹, Zhouhao Pan ², Bingo Wing-Kuen Ling ¹ , Ye Tian ³ and Zhe Zhang ^{4,5,6,7,8,*}

¹ Guangdong University of Technology, Guangzhou 510006, China; yaozhao@gdut.edu.cn (Y.Z.); 2112203071@mail2.gdut.edu.cn (D.X.); yongquanling@gdut.edu.cn (B.W.-K.L.)

² China Academy of Electronics and Information Technology, Beijing 100041, China; panzhouhao@cetc.com.cn

³ China Telecom Satellite Application Technology Research Institute, Beijing 100035, China; tianye6@chinatelecom.cn

⁴ Suzhou Key Laboratory of Microwave Imaging, Processing and Application Technology, Suzhou 215000, China

⁵ Suzhou Aerospace Information Research Institute, Suzhou 215000, China

⁶ National Key Laboratory of Microwave Imaging Technology, Beijing 100190, China

⁷ Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China

⁸ School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100190, China

* Correspondence: zhangzhe01@aircas.ac.cn

Abstract: The integration of Synthetic Aperture Radar (SAR) imaging technology with deep neural networks has experienced significant advancements in recent years. Yet, the scarcity of high-quality samples and the difficulty of extracting prior information from SAR data have experienced limited progress in this domain. This study introduces an innovative sparse SAR imaging approach using a self-supervised non-local asymmetric pixel-shuffle blind spot network. This strategy enables the network to be trained without labeled samples, thus solving the problem of the scarcity of high-quality samples. Through asymmetric pixel-shuffle downsampling (AP) operation, the spatial correlation between pixels is broken so that the blind spot network can adapt to the actual scene. The network also incorporates a non-local module (NLM) into its blind spot architecture, enhancing its capability to analyze a broader range of information and extract more comprehensive prior knowledge from SAR data. Subsequently, Plug and Play (PnP) technology is used to integrate the trained network into the sparse SAR imaging model to solve the regularization term problem. The optimization of the inverse problem is achieved through the Alternating Direction Method of Multipliers (ADMM) algorithm. The experimental results of the unlabeled samples demonstrate that our method significantly outperforms traditional techniques in reconstructing images across various regions.

Keywords: sparse SAR imaging; blind spot network; ADMM



Citation: Zhao, Y.; Xiao, D.; Pan, Z.; Ling, B.W.-K.; Tian, Y.; Zhang, Z. Sparse SAR Imaging Based on Non-Local Asymmetric Pixel-Shuffle Blind Spot Network. *Remote Sens.* **2024**, *16*, 2367. <https://doi.org/10.3390/rs16132367>

Academic Editor: Jean-Marc Le Caillec

Received: 30 April 2024

Revised: 15 June 2024

Accepted: 26 June 2024

Published: 28 June 2024



Copyright: © 2024 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 a radar system that produces high-resolution images of the ground [1]. Its working principle is to generate high-resolution images of a target area by using radar antennas to emit and receive reflected signals along a certain path and synthesizing these signals for processing. SAR is characterized by its all-weather, all-time capabilities, making it crucial in fields such as military reconnaissance, earth observation, agricultural monitoring, forestry, and geological exploration [2]. Sparse SAR imaging refers to a synthetic aperture radar technique based on compressed sensing, a signal processing technology that allows for signal reconstruction at sampling ratios significantly below those required by the Nyquist sampling theorem, assuming the signal exhibits sparsity [3–5]. Sparse SAR imaging offers several advantages, including a reduced system sampling rate, enhanced imaging performance, and decreased system complexity.

In the field of synthetic aperture radar (SAR) imaging, various methodologies provide effective solutions for data processing. The Chirp scaling algorithm (CSA) [6] performs well in processing large-scale SAR datasets, but its performance significantly deteriorates in sparse scenarios. The L1 + TV regularization method [7] enhances image resolution by promoting sparsity and smoothness, making it well-suited for processing sparse signals. However, this method is highly sensitive to parameter selection. Improperly chosen parameters can lead to reduced imaging quality or low computational efficiency, thus requiring careful parameter tuning. The probabilistic patch-based (PPB) [8] method utilizes probabilistic information from analogous regions within images to elevate quality and effectively mitigate noise. Nevertheless, its performance critically depends on the homogeneity of image regions and can falter in complex scenes that lack repetitive textures.

Deep learning has demonstrated significant advantages in addressing sparse inverse problems and has been applied to sparse SAR imaging, achieving a range of benefits. These include strong generalization capabilities, increased imaging speed, and enhanced reconstruction quality [9–13]. Sparse SAR imaging techniques based on deep learning rely on a large and diverse set of data samples to learn the mapping relationship from input to output. The quality, quantity, and diversity of samples directly impact the training effectiveness and overall performance of deep learning models. Yet, the acquisition of high-quality samples often encounters various difficulties in sparse SAR imaging, mainly reflected in the time-consuming and costly annotation of data. This process particularly requires professional knowledge for accurate annotation, and the consistency between different annotators can vary greatly, which may affect the quality of the imaging. SAE-Net [14] introduces a deep neural network that employs a self-supervised strategy for SAR imaging, effectively addressing the challenges associated with sparse SAR data. This approach not only underscores the research value of self-supervised strategies in SAR imaging but also highlights their potential to enhance data utilization and improve imaging quality under limited data conditions.

The Blind-Spot Network (BSN) is a deep learning architecture that is designed to address self-supervised learning challenges within image processing tasks, particularly in image restoration [15]. Its core concept involves intentionally “blinding” or obscuring the central pixel or area during the network’s prediction process, compelling the network to predict the content of that central pixel or area based on the surrounding contextual information. This method can be viewed as a specialized self-supervised learning strategy that learns not directly from raw input data but by predicting parts of the input data to learn useful data representations. Consequently, the blind-spot network can be trained without extensive labeled data, making it exceptionally suited for applications in sparse SAR imaging. It has been widely applied in fields such as MRI [16–18]. Building on this, further studies have introduced asymmetric pixel-shuffle downsampling (AP) [19] using different stride factors during training and inference to better break the spatial correlation and reduce aliasing, enabling BSN to adapt to the actual scene. Others have also introduced non-local modules [20], utilizing the internal non-local similarities to expand the perceptual range for analyzing information from a broader scope.

In this paper, we introduce a novel sparse SAR imaging method based on non-local asymmetric pixel-shuffle blind spot network (NLAPBSN), an improvement of the asymmetric pixel-shuffle blind spot network by incorporating a non-local module applicable into the blind spot architecture. Initially, we improve the self-supervised asymmetric pixel-shuffle downsampling BSN model by innovatively adding non-local modules adapted to the blind-spot architecture. This modification allows the NLAPBSN model to utilize a broader range of pixel information during training and inference, significantly enhancing image processing quality. The pre-trained model is obtained using SAR image data during the training process, ensuring that the model acquires a rich prior of SAR imagery. We then incorporate the plug and play (PnP) framework [21] that allows for the integration of the pretrained NLAPBSN model with specific imaging algorithms. Finally, we employ the Alternating Direction Method of Multipliers (ADMM) [22] to iteratively optimize the

imaging algorithm and solve the imaging problem. Through a series of simulations and empirical data experiments on SAR imagery, we validate the effectiveness of our method. These experimental results not only demonstrate the practicality of our approach but also highlight its potential in real-world applications. The main contributions of this paper are as follows:

- A novel method that integrates NLAPBSN with sparse SAR imaging techniques is proposed. This approach utilizes NLAPBSN to learn richer prior information from SAR data, which is then incorporated as a regularization term in sparse SAR imaging. By introducing this effective constraint mechanism, the SAR imaging performance is significantly enhanced.
- The non-local module suitable for blind-spot networks was first embedded into the blind spot network model with asymmetric pixel-shuffle downsampling, enhancing the model's understanding of the overall structure and content of SAR images. This improvement has consequently boosted both the performance and the generalization capability of the model.
- A self-supervised learning strategy was introduced in the task of sparse SAR imaging, effectively addressing the challenge of obtaining high-quality samples in the field of SAR imaging.
- Experimental validation demonstrates that the sparse SAR imaging algorithm based on the self-supervised blind-spot network proposed in this paper significantly outperforms other comparative methods in sparse scenarios.

The remainder of this paper is organized as follows. Section 2 describes the signal model and the sparse SAR imaging method based on the blind-spot network. Section 3 presents the results of simulation experiments and real data processing. Section 4 discusses our methods and experimental outcomes. Section 5 provides the conclusions.

2. Materials and Methods

2.1. Signal Models and Imaging Principles

SAR imaging is an inverse problem that reconstructs the scene's reflectivity through echo data. The fundamental principle involves the use of radar sensors to emit electromagnetic waves and receive signals reflected back from the targets. By analyzing the time delay and frequency changes of these echoes, the image of the surface can be constructed. In SAR imaging, the two-dimensional echo data collected by the SAR system is represented as $\mathbf{Y} \in \mathbb{C}^{M_\mu \times M_\tau}$, where M_μ and M_τ denote the range and azimuth dimensions of the data, respectively. The backscattering coefficient matrix, derived from the reconstruction of the echo data, is denoted as $\mathbf{X} \in \mathbb{C}^{N_p \times N_q}$, where N_p and N_q represent the range and azimuth dimensions, respectively. The formulation of sparse SAR imaging can be articulated as follows:

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{n} \quad (1)$$

In this context, $\mathbf{y} = \text{vec}(\mathbf{Y}) \in \mathbb{C}^{M \times 1}$ represents the column vector formed by the echo signals sampled over time, where $M = M_\mu \times M_\tau$. $\mathbf{x} = \text{vec}(\mathbf{X}) \in \mathbb{C}^{N \times 1}$ denotes the column vector of the scene's backscatter coefficients after spatial discretization, where $N = N_p \times N_q$. The operation $\text{vec}(\cdot)$ is used to denote stacking of the columns of a matrix in sequence to form a column vector. Φ is referred to as the system's observation matrix, and $\mathbf{n}$ represents the additive noise within the echo.

Notably, SAR data exhibits coupling between the azimuth and range dimensions, necessitating the rearrangement of the two-dimensional echo matrix into a one-dimensional vector. In large-scale observation scenarios, this approach can consume substantial memory. To overcome these issues and accelerate the imaging process, we have employed the Chirp Scaling algorithm (CSA) to construct an approximate observation model. By utilizing an echo simulation operator based on chirp scaling, we are able to efficiently process imaging:

$$\mathbf{K}_{cs}\{\mathbf{Y}\} = \mathbf{F}_a^{-1} \left\{ \mathbf{H}_3 \odot \left\{ \mathbf{H}_2 \odot [\mathbf{H}_1 \odot (\mathbf{F}_a \mathbf{Y} \mathbf{F}_r) \mathbf{F}_r^{-1}] \right\} \right\} \quad (2)$$

In the context of this formula, F_a and F_r denote the Fourier transforms in the azimuth and range dimensions, respectively. Conversely, F_a^{-1} and F_r^{-1} indicate the inverse Fourier transforms for the azimuth and range dimensions. The symbol $\odot$ is used to represent the Hadamard product. The terms H_1 , H_2 , and H_3 refer to phase functions that are integral to the imaging process. The imaging operator $K_{cs}\{\cdot\}$ involves working with matrices or operators, and the interactions between these operators consist exclusively of matrix multiplication and Hadamard multiplication. Importantly, all the operators are linear [23], allowing for the decomposition of $K_{cs}\{\cdot\}$ into a composition of three distinct linear operators.

$$K_{cs}\{\cdot\} = K_{ac}\{K_{rc}\{K_{sc}\{\cdot\}\}\} \quad (3)$$

The operator K_{sc} represents the chirp-scaling decoupling operator, K_{rc} represents the range compression and consistent range cell migration correction operators, and K_{ac} represents the azimuth compression and phase correction operators. The imaging of observation echoes $\mathbf{y}$ with chirp scaling imaging operators $K_{cs}\{\cdot\}$ can be expressed as follows:

$$\hat{\mathbf{x}} = K_{cs}\{\mathbf{y}\} \quad (4)$$

The reversal of the imaging operator built on the Chirp Scaling algorithm is called the approximate observation operator $J\{\cdot\}$.

$$J\{\cdot\} = K_{sc}^H\{K_{rc}^H\{K_{ac}^H\{\cdot\}\}\} \quad (5)$$

Thus, we can obtain an approximation of the observation matrix.

$$\mathbf{y} = J\{\hat{\mathbf{x}}\} \quad (6)$$

which means we can use it to relate to Formula (1).

For sparse SAR imaging, we can then represent SAR imaging in regularized form as follows:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - J\{\mathbf{x}\}\|_2^2 + \lambda R(\mathbf{x}) \quad (7)$$

The formula can be considered as composed of the data fidelity term and the regularization term. The former is primarily designed to align the predictive model $J\{\mathbf{x}\}$ as closely as possible with the actual echo data $\mathbf{y}$. Meanwhile, the regularization term $R(\mathbf{x})$ introduces prior information and constraints on the expected solution $\mathbf{x}$, modulated by the parameter λ . The balance between these two elements ensures both the accuracy of the predictive model and the reasonableness of the expected solution.

2.2. Non-Local Asymmetric Pixel-Shuffle Blind Spot Network

2.2.1. Asymmetric Pixel-Shuffle Blind Spot Network

The blind spot network is a variant of convolutional neural networks that is designed to reconstruct a clean image pixel from its noisy version by exploiting the surrounding pixel values and intentionally not using the target pixel itself during the inference [15]. It fundamentally differs from traditional neural networks in how they predict pixel values. Traditional networks typically predict the value of a specific pixel by considering both the pixel itself and its neighboring pixels. This approach, when used during training with a noisy image as both input and target, often leads the network to learn a simplistic direct mapping from input to output. In contrast, the blind spot network eliminates the direct dependence on input pixels through an innovative approach. The network utilizes only the information surrounding the target pixel when predicting pixel values, excluding data from the target pixel itself. This forces the network to learn to infer a pixel's properties based on its environment; the information surrounding it. This approach enables blind spot networks to achieve a deeper understanding of the image content by focusing on the interrelations and dependencies among the pixels in the vicinity, leading to potentially more robust and accurate image processing capabilities. Figure 1 reveals the basic difference between blind spot networks and traditional neural networks for image data processing.

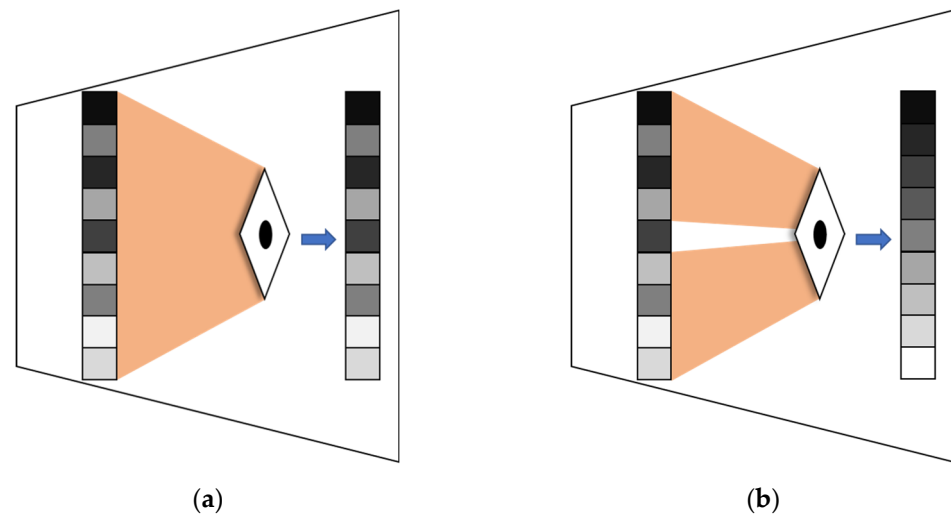


Figure 1. Differences between blind spot networks and conventional networks. (a) Conventional networks, when processing a single pixel, rely on the information provided by its surrounding pixels, including the pixel itself. (b) Blind spot networks, when processing a single pixel, rely on the information provided by its surrounding pixels and exclude the pixel itself.

Pixel-shuffle downsampling [24] is a technique that is primarily used in image processing that rearranges pixels in a specific way to reduce the spatial correlation between adjacent pixels in an image. It involves subsampling the image with a stride factor s , transforming the image into multiple sub-images with reduced spatial correlation between pixels. Since the use of BSN requires the noise in the image to be pixel-wise independent, using these PD-processed sub-images as inputs for BSN meets this requirement. However, a stride factor s that is too small cannot sufficiently reduce the spatial correlation between pixels, and too large a stride factor can result in aliasing artifacts. To address the limitations of applying PD with BSN, the asymmetric pixel-shuffle [19] introduces different stride factors for training and inference. This asymmetry helps balance the trade-off between breaking pixel spatial correlations and preserving image details. Figure 2 shows the simplified structure of the asymmetric pixel-shuffle blind spot network.

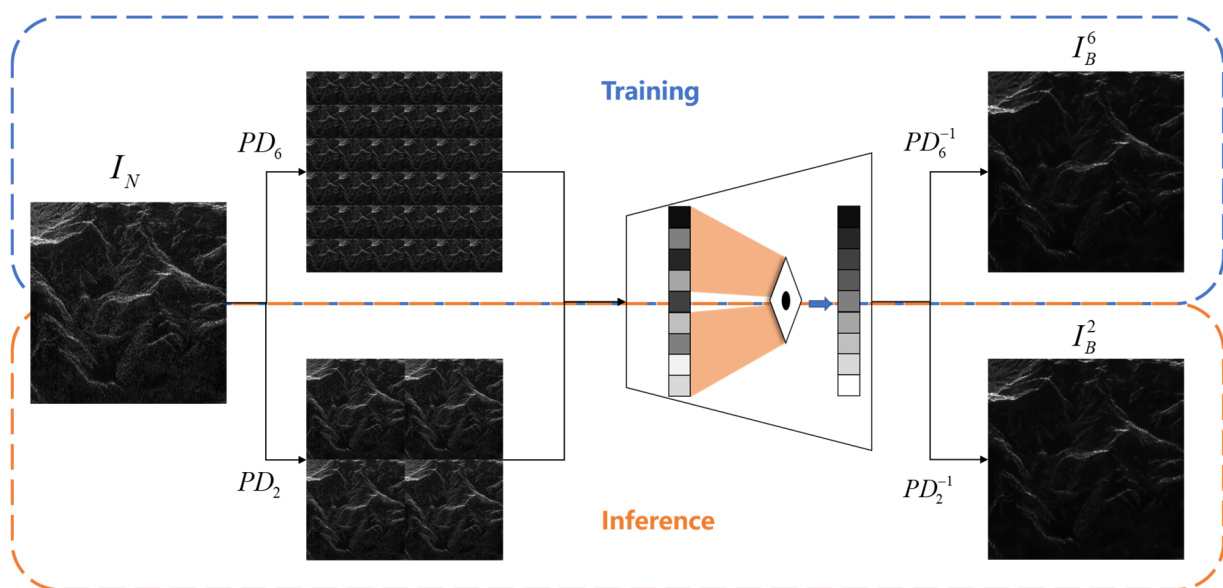


Figure 2. Simplified architecture of asymmetric pixel-shuffle blind spot network.

2.2.2. Non-Local Module

The non-local module (NLM) is a widely employed network component in the fields of deep learning and computer vision. It captures long-range feature dependencies by considering the global information of input feature maps. This module significantly enhances the network's understanding of the overall image structure, particularly proving effective in handling images with complex scenes and structures. It plays a crucial role in improving model accuracy and robustness. Additionally, as a standalone layer, the non-local module can be seamlessly integrated into various existing neural network architectures, endowing the model with an enhanced capability in processing global information.

A recent study introduced a novel non-local module utilizing a soft patch matching approach, which incorporates linear embedding and Gaussian-weighted Euclidean distance to measure similarity [25]. Although initially designed for traditional convolutional neural networks, to adapt it to the blind spot network architecture, we propose adjustments incorporating masking techniques to ensure compatibility. Specifically, we redefine the linear embedding as follows:

$$\begin{aligned} \delta(U_{ij}) &= \varphi(U_{ij}, U_{pij}) = \exp\{\Theta(U_{ij})\Xi(U_{pij})^T\}, \forall i, j, \\ \Theta(U_{ij}) &= U_{ij}V_{\Theta}, \Xi(U_{pij}) = U_{pij}V_{\Xi}, G(U_{ij}) = U_{pij}V_g, \forall i, j. \end{aligned} \quad (8)$$

where $\Theta(U_{ij})$ and $\Xi(U_{pij})$ denote the embedding functions corresponding to the feature vectors at locations i, j and to each feature vector in the neighboring block p , respectively. $\delta(U_{ij})$ represents a specific linear embedding function for encoding nonlocal correlations, which primarily serves to quantify the nonlocal associations between the feature vector U_{ij} at position i, j and each of its feature vectors in the neighborhood patch U_{pij} . V_{Θ} , V_{Ξ} , and V_g are used as transform weights to compute these embeddings. Their sizes are $w \times k$, $w \times k$, and $w \times w$, respectively. Then the nonlocal operation is reformulated as follows:

$$W_{ij} = \frac{1}{\gamma_{ij}(U_{ij})} (S \odot \exp\{U_{ij}V_{\Theta}V_{\Xi}^T U_{pij}^T\}) U_{pij}V_g, \forall i, j, \quad (9)$$

$$\gamma_{ij}(U_{ij}) = \sum_{pij} S \odot \varphi(U_{ij}, U_{pij}) \quad (10)$$

where $\gamma_{ij}(U_{ij})$ is the normalization factor, W_{ij} represents the obtained output feature vector, and S represents the mask operation. With this improvement, the nonlocal operation is better adapted to the structure of the blind spot architectures, allowing for the exclusion of vector-specific effects during the computation of the feature vectors. By integrating this non-local mechanism, the network is able to analyze the information from a wider scope, exploiting the non-local similarities within the image to compensate for information loss caused by occlusions. This approach improves the performance of the network in reconstructing details and interpreting complex image structures.

2.2.3. Architecture of the NLAPBSN

The NLAPBSN is an improvement upon the asymmetric pixel-shuffle blind spot network by incorporating the non-local module that is applicable to the blind spot network. In NLAPBSN, the initial layer is the center mask convolutional layer, which serves to blur the input pixel information. It performs blind spot operations on different branches using center mask convolution kernels of different sizes, such as 3×3 and 5×5 , respectively. Such an architectural design promotes the network's dependence on neighboring pixels rather than a single central pixel when processing information, thereby enhancing its ability to extract prior knowledge. Each branch includes inflated convolutional modules with different expansion rates, and a total of nine such modules are embedded. Non-local modules are added after the inflated convolutional modules to further enhance the network's ability to perceive the global structure of the image. This configuration allows the network to capture a wider range of spatial information. Figure 3 illustrates the architecture of our network.

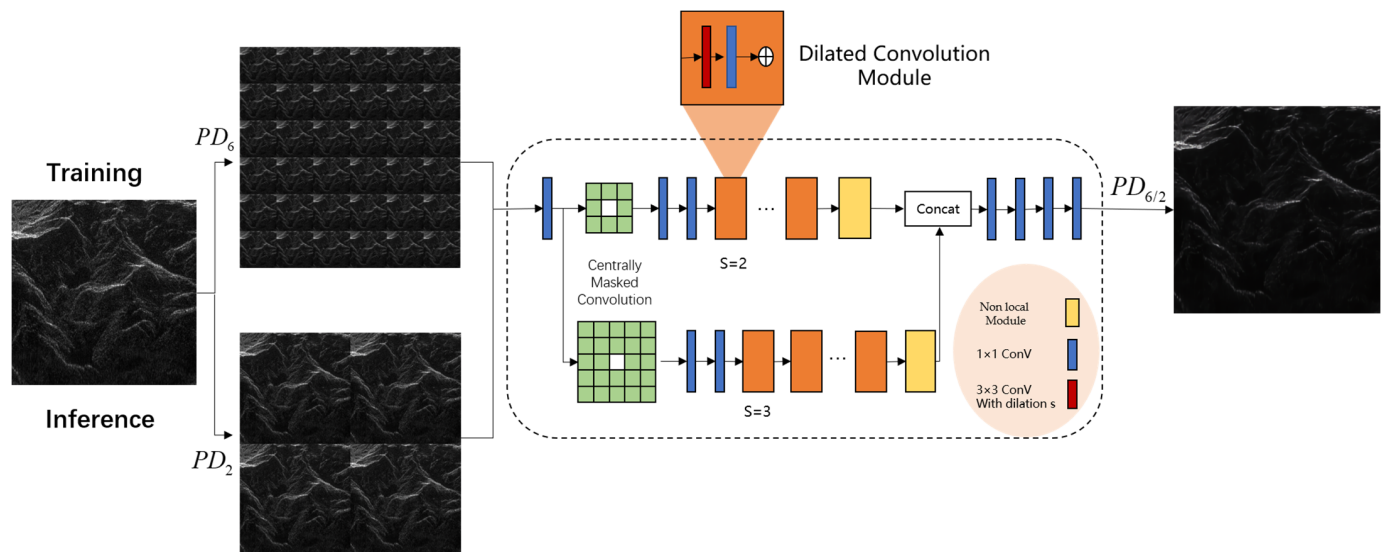


Figure 3. Illustration of NLAPBSN architecture.

In this network, the upper path employs a $2\times$ dilation rate, which expands the field of view and allows the network to recognize large patterns and textures. This is effective for detecting broad noise trends and large-scale image features. The lower path utilizes a $3\times$ expansion rate, effectively improving the capture of global information. To further extend the network's perceptual capabilities, we add non-local modules to the two-layer paths to capture contextual information more comprehensively. Through asymmetric pixel-shuffle downsampling, we use different step factors during the training and inference processes to further improve the ability of the network to extract the prior information from the image.

The objective of this study is to advance the application of our proposed network for SAR images using a self-supervised approach. Consequently, we have chosen the following loss function to train the NLAPBSN:

$$L_B = \left\| PD_s^{-1}(B(PD_s(I_N))) - I_N \right\|_1 = \left\| I_B^s - I_N \right\|_1 \quad (11)$$

where PD_s and PD_s^{-1} denote the pixel-shuffle downsampling and its inverse operation, respectively. I_B^s and I_N denote the output and input of the NLAPBSN. Here, $B(\cdot)$ represents the BSN enhanced with a non-local module and the L1 norm is used for better generalization.

2.3. Sparse SAR Imaging Model Based on NLAPBSN

In this study, we propose a method to address the sparse SAR imaging inverse problem by employing the NLAPBSN pre-trained model as a regularization prior within the plug-and-play (PnP) framework. This approach leverages advanced image restoration techniques by utilizing a pretrained model that incorporates prior information, replacing traditional regularization methods. However, when employing this strategy, models pretrained based on images from other domains may not provide sufficient a priori information for SAR images. While supervised learning for SAR images can provide rich a priori information, this usually relies on a large number of labeled samples, which are often difficult to obtain. The blind spot network we employ not only enables deeper mining of a priori knowledge in SAR images but also overcomes the challenge of obtaining high-quality labeled samples.

We use the PnP framework to solve (7). First, the original problem of (7) is reformulated by separating the data fidelity term and the regularization term using the auxiliary variable h . The problem is then formulated as follows:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - J\{\mathbf{x}\}\|_2^2 + \lambda R(\mathbf{h}) \quad \text{s.t. } \mathbf{x} = \mathbf{h} \quad (12)$$

The Lagrange multiplier $\mathbf{u}$ is then used to construct the Lagrange function, which includes an augmentation term for dealing with the equation constraints:

$$L_\beta = \frac{1}{2} \|\mathbf{y} - J\{\mathbf{x}\}\|_2^2 + \lambda R(\mathbf{h}) + \mathbf{u}^T(\mathbf{x} - \mathbf{h}) + \frac{\beta}{2} \|\mathbf{x} - \mathbf{h}\|_2^2 \quad (13)$$

where β is the ADMM penalty parameter. Iterative updates are performed using the framework of ADMM, and each iteration is divided into three steps:

The first step is to update $\mathbf{x}$ by minimizing the part of the Lagrangian function L_β with respect to $\mathbf{x}$.

$$\mathbf{x}^{k+1} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{y} - J\{\mathbf{x}\}\|_2^2 + \frac{\beta}{2} \|\mathbf{x} - \mathbf{h}^k + \mathbf{u}^k\|_2^2 \quad (14)$$

The second step is to update the auxiliary variable $\mathbf{h}$ by using the pre-trained model of NLAPBSN that extracts a priori information about the SAR data as the regularization term. In other words, we update $\mathbf{h}$ by NLAPBSN pretrained model $\mathbf{D}$ as follows:

$$\mathbf{h}^{k+1} = \mathbf{D}(\mathbf{x}^{k+1} + \mathbf{u}^k) \quad (15)$$

In the third step, the Lagrange multiplier $\mathbf{u}$ is then updated based on the difference between $\mathbf{x}$ and $\mathbf{h}$:

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{x}^{k+1} - \mathbf{h}^{k+1} \quad (16)$$

The solution to this optimization problem can be found through continuous iterative updating. The key advantage of this method is that the pretrained model acts as a prior in the regularization step, which helps to guide the updating step and improves the quality of the reconstruction. This mathematically maintains the decomposition of the optimization problem, which allows for an efficient solution at each step. In addition, the convergence of the ADMM ensures that a stable solution to the problem is found as the algorithm iterates.

2.4. Evaluation Indices

2.4.1. Peak Signal-to-Noise Ratio (PSNR)

PSNR is a measure of image quality that evaluates the fidelity of an image by quantifying the difference between the processed image and the original reference image. It calculates the variance in pixel values between two images based on the logarithmic scale Mean Square Error (MSE), which is determined by calculating the mean squared error between the reference image and the processed image. A higher PSNR value usually implies that the image is less distorted, indicating that the processed image performs better in maintaining the similarity with the original image, where $\max(\mathbf{X}_1)$ denotes the maximum pixel value in the reconstructed image; $\mathbf{X}_1$ and $\mathbf{X}_2$ represent the reconstructed image and the reference image, respectively; and $N_{\mathbf{X}_2}$ represents the total number of pixels in the image $\mathbf{X}_2$.

$$\text{PSNR} = 10 \lg \left(\frac{\max(\mathbf{X}_1)^2}{\text{MSE}} \right) \quad (17)$$

$$\text{MSE} = \frac{1}{N_{\mathbf{X}_2}} \|\mathbf{X}_1 - \mathbf{X}_2\|_F^2 \quad (18)$$

2.4.2. Structural Similarity Index Measure (SSIM)

SSIM is a metric that is commonly used to assess the degree of image similarity by considering the key features of brightness, contrast, and structural features of an image. The parameters in equation include the mean $\mu_{\mathbf{X}_1}$, $\mu_{\mathbf{X}_2}$, variance $\sigma_{\mathbf{X}_1}$, $\sigma_{\mathbf{X}_2}$, covariance $\sigma_{\mathbf{X}_1\mathbf{X}_2}$, and constants c_1 , c_2 .

$$\text{SSIM} = \frac{(2\mu_{\mathbf{X}_1}\mu_{\mathbf{X}_2} + C_1)(2\sigma_{\mathbf{X}_1\mathbf{X}_2} + C_2)}{(\mu_{\mathbf{X}_1}^2 + \mu_{\mathbf{X}_2}^2 + C_1)(\sigma_{\mathbf{X}_1}^2 + \sigma_{\mathbf{X}_2}^2 + C_2)} \quad (19)$$

2.4.3. Equivalent Number of Looks (ENL)

ENL is an indicator used to assess the level of noise and smoothness in SAR images. It represents the equivalent number of independent views. The higher the ENL, the lower the noise level in the image and the better the image quality.

$$\text{ENL} = \left(\frac{\mu_O}{\sigma_O} \right)^2 \quad (20)$$

2.4.4. Edge Preservation Index (EPI)

The edge preservation index (EPI) is a metric that is used to evaluate the ability of an image processing algorithm to reduce noise or smooth an image while preserving the image's detail and edge sharpness. $\mathbf{X}_d$ and $\mathbf{X}_s$ represent images without and with noise, respectively.

$$\text{EPI} = \frac{\sum_{i,j} \sqrt{([\mathbf{X}_d(i,j) - \mathbf{X}_d(i+1,j)]^2 + [\mathbf{X}_d(i,j) - \mathbf{X}_d(i,j+1)]^2)}}{\sum_{i,j} \sqrt{([\mathbf{X}_s(i,j) - \mathbf{X}_s(i+1,j)]^2 + [\mathbf{X}_s(i,j) - \mathbf{X}_s(i,j+1)]^2)}} \quad (21)$$

2.4.5. Target Background Ratio (TBR)

TBR is a metric that is used to assess the quality of SAR imaging. It is calculated by comparing the brightness or signal intensity of the target area with that of its surrounding background. A high TBR value usually indicates a clear contrast between the target and the background and a clearer target. r_T and r_B represent the target and background areas, respectively.

$$\text{TBR} = 10 \lg \left(\frac{\sum_{\mathbf{X}_{i,j} \in r_T} |\mathbf{X}_{i,j}|^2}{\sum_{\mathbf{X}_{i,j} \in r_B} |\mathbf{X}_{i,j}|^2} \right) \quad (22)$$

3. Results

In this study, we demonstrate a novel sparse SAR learning imaging technique that is based on NLAPBSN. We compare this technique with several conventional imaging methods, including the chirp scaling algorithm (CSA) [6], the L1+TV [7] regularization technique, PPB [8], and SAE-Net [14]. We comparatively analyze the performance of the various techniques by performing experiments on both simulated and real datasets. In these experiments, the methods are based not only on quantitative evaluation metrics but also include qualitative visual effect comparisons.

In our simulations, we assessed the imaging performance of various methods by adjusting the sampling rates, and we evaluated each algorithm using PSNR and SSIM metrics. The simulated data were derived from the AIR-SARShip dataset. In the experiments involving real data, we selected distinct regions characterized by unique features, including plains, a ferry terminal, and ships, to conduct a thorough analysis. For these varied regions, we computed the corresponding evaluation metrics to gauge the imaging capabilities of each method. The real data utilized in these experiments were sourced from the RADARSAT dataset.

3.1. Simulations

In this simulation, we selected some sparse scene images from the AIR-SARShip dataset as training samples for preprocessing. These samples consist of a total of 1000 images of 256×256 pixels, of which 100 are selected as the validation set. These images are also used to generate simulated echo data. In the training phase, we set 200 epochs and used a learning rate of 0.0001 and a batch size of 32 with the goal of minimizing the loss

function to optimize the network performance. The system parameters for the simulation are shown in Table 1.

Table 1. Parameters of the experiment based on the simulation data.

Parameters	Value
Pulse duration time	45 μ s
Radar center frequency	5.4 GHz
Bandwidth	60.5 MHz
Range FM rate	1.3 MHz/ μ s
Pulse repetition frequency	1200 Hz

3.1.1. The Selection of Stride Factors

In order to balance the aliasing noise in the training and inference phases and maximize the training effect, we need to choose the appropriate stride factors for each phase. By experimenting with different pacing factors, we can observe how they affect the model's performance and identify the most suitable configuration for a particular dataset. Through this systematic tuning and comparison, we are able to ensure that the chosen stride factors really help to improve the model performance. The results presented in Figure 4 demonstrate that the SNR is highest when the spanning factors are set to $a = 2$ and $b = 6$. This finding suggests that the model achieves optimal performance at these parameter settings.

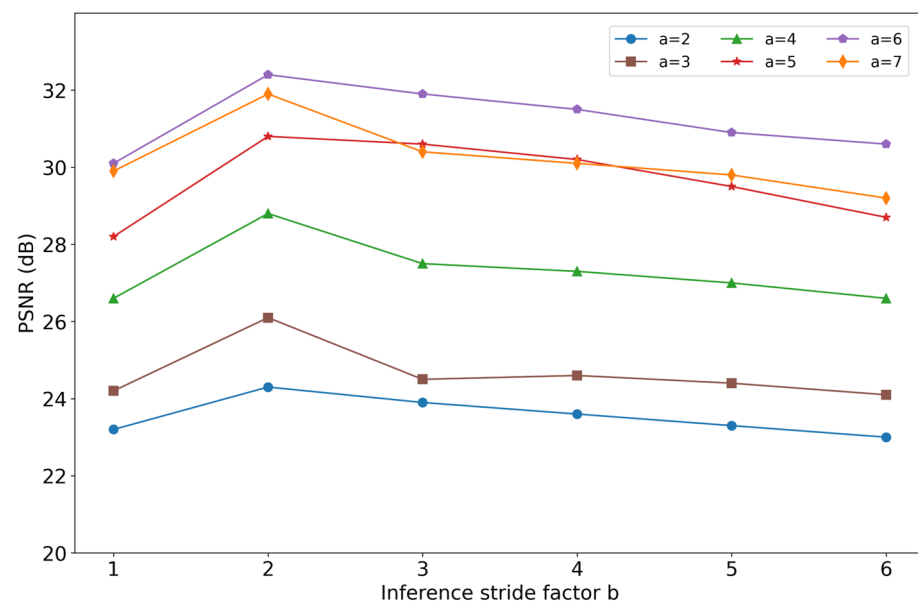


Figure 4. Effects of stride factors a and b.

3.1.2. SAR Imaging Results Based on Simulated Data

In this section, we designed comparative experiments across various sampling ratios to evaluate the image reconstruction capabilities of our proposed method against four alternative approaches. In the initial set of experiments, scenes containing ships were reconstructed using different techniques at an 80% sampling ratio. Figure 5 illustrates the SAR imaging outcomes at this sampling rate. Images reconstructed using the chirp scaling algorithm (CSA) showed noticeable distortions. In contrast, the image quality was notably enhanced using the L1 + TV and PPB methods. Compared to the aforementioned methods, both SAE-Net and our proposed approach demonstrated superior imaging quality, with more distinct edge features.

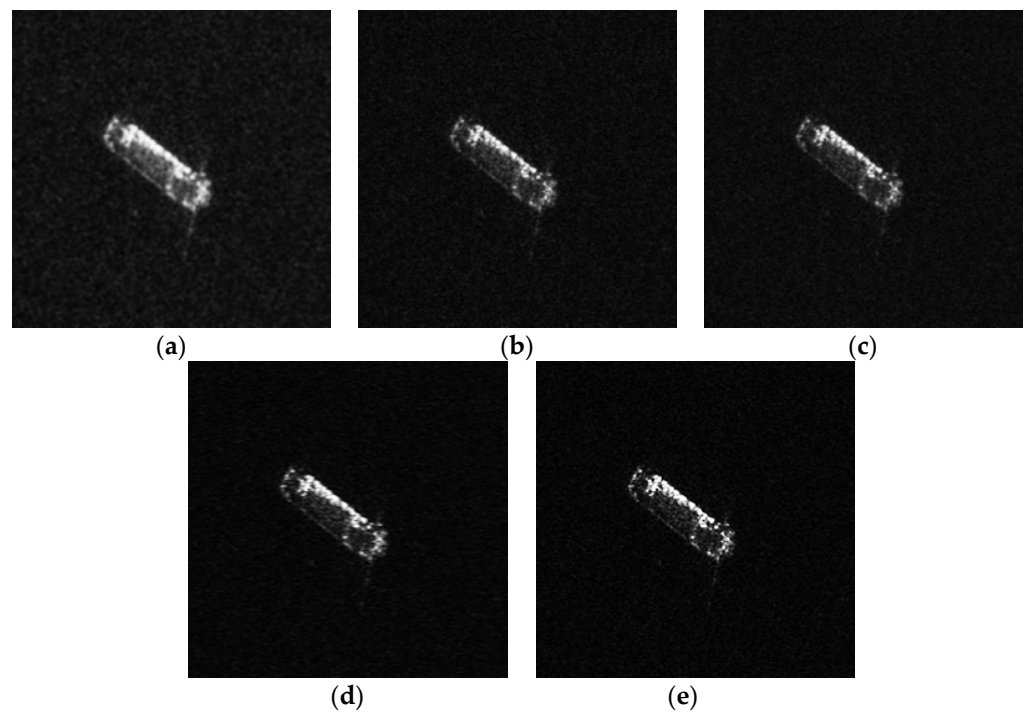


Figure 5. Imaging results for simulated data at an 80% undersampling ratio. (a) CSA; (b) L1 + TV; (c) PPB; (d) SAE-Net; (e) proposed method.

In subsequent experiments, we reduced the sampling ratio to 60% and employed various methods for image reconstruction to explore their performance under reduced sampling conditions. Figure 6 presents the imaging results from these methods. At a reduced sampling ratio of 60%, the reconstruction quality of all the methods deteriorated significantly, as indicated by the increased noise levels. Yet, our proposed method continued to excel in enhancing imaging quality, particularly in noise suppression.

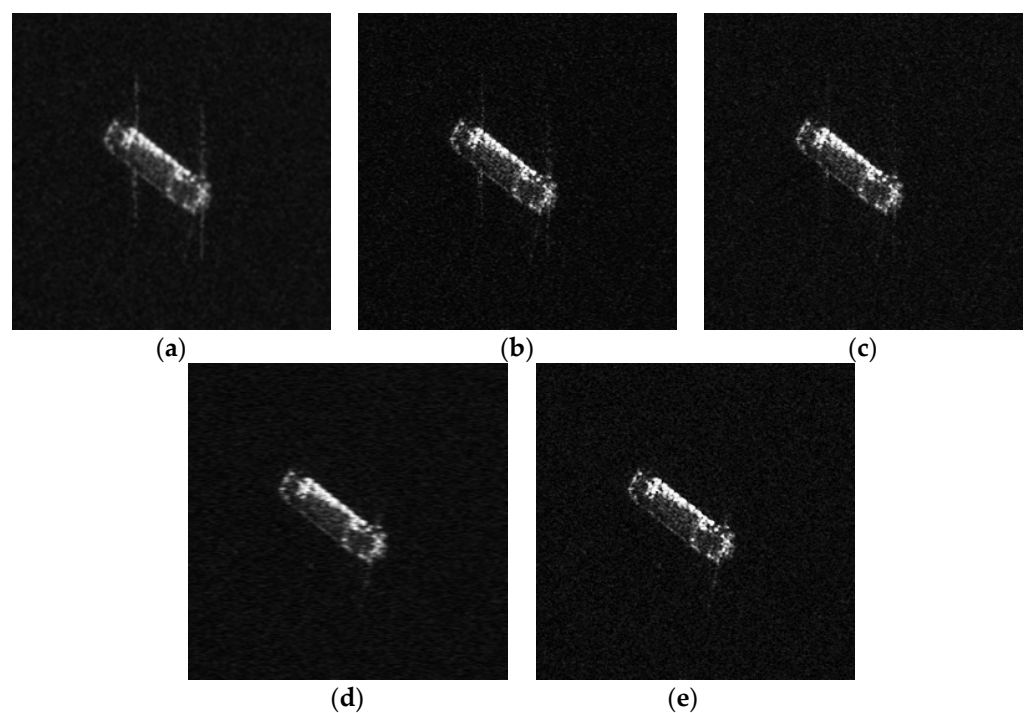


Figure 6. Imaging results for simulated data at 60% undersampling ratio. (a) CSA; (b) L1 + TV; (c) PPB; (d) SAE-Net; (e) proposed method.

Table 2 provides an analysis of the reconstruction outcomes of the various methods at different sampling ratios. The data reveal that the proposed method consistently outperforms the others in terms of PSNR and SSIM values across all sampling ratios. Relative to traditional approaches such as CSA, L1 + TV, and PPB, both SAE-Net and our proposed method demonstrate superior performance metrics, with our method slightly surpassing SAE-Net. These findings further corroborate the robustness of our proposed method, particularly under conditions of reduced sampling ratios.

Table 2. PSNR and SSIM of different methods of imaging in simulation experiments.

Method	PSNR/SSIM			
	80% Sampling	60% Sampling	40% Sampling	20% Sampling
CSA	20.65/0.8245	17.25/0.6529	14.89/0.5398	9.75/0.4837
L1 + TV	27.43/0.9283	24.64/0.8624	21.23/0.7156	16.93/0.6386
PPB	28.97/0.9365	24.57/0.8535	21.03/0.7078	17.32/0.6145
SAE-Net	31.86/0.9527	29.14/0.8987	22.98/0.7305	19.43/0.6517
Proposed	32.13/0.9534	29.68/0.9074	23.62/0.7356	19.86/0.6631

Additionally, we analyzed the imaging time for the five algorithms discussed. Table 3 presents a comparison of imaging speeds for different methods under the same experimental scenario. The chirp scaling algorithm (CSA) had the shortest actual runtime, while our proposed method exhibits moderate performance in terms of imaging speed. This finding indicates that embedding the pre-trained model into the iterative part of the ADMM algorithm somewhat increases the computation time. However, our proposed method still surpasses the imaging speed of the PPB method and SAE-Net. Considering both imaging speed and image quality, the method proposed in this paper is reasonable and effective.

In our simulation, we assessed the effectiveness of the proposed method in extracting distributed targets with distinct edge features, utilizing rectangles, triangles, and circles as test objects. Figure 7 presents the outcomes of applying our method to these various shapes. Specifically, Figure 7a displays the original reference images, Figure 7b shows images with substantial noise, and Figure 7c depicts the results after applying our proposed method. The findings demonstrate that the method effectively suppresses noise while preserving the clarity of edges across all three distinct target types. This outcome confirms the efficacy of the proposed method in accurately extracting prior information, underscoring its potential utility for applications requiring precise edge detection in noisy environments.

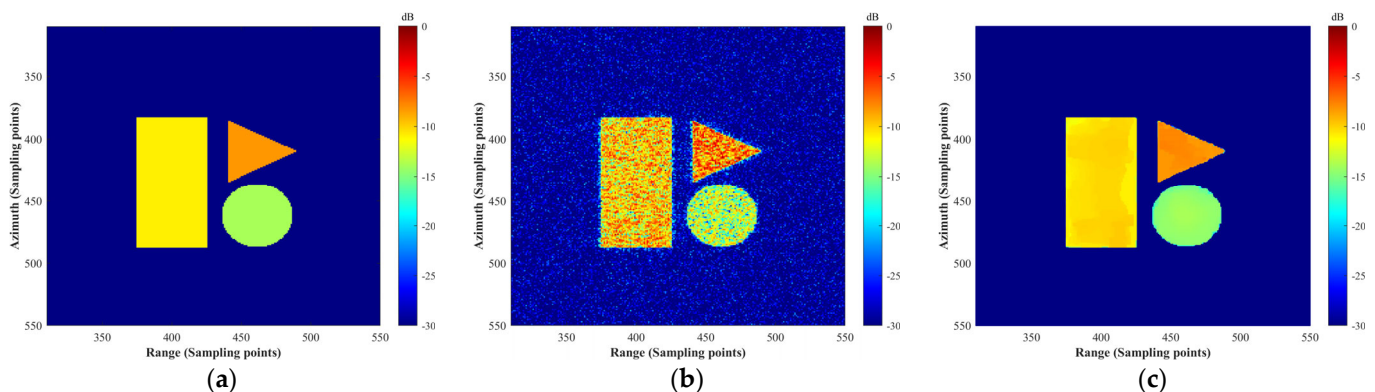


Figure 7. Results for surface targets of different shapes. (a) Original reference images; (b) images with substantial noise; (c) images processed using the proposed method.

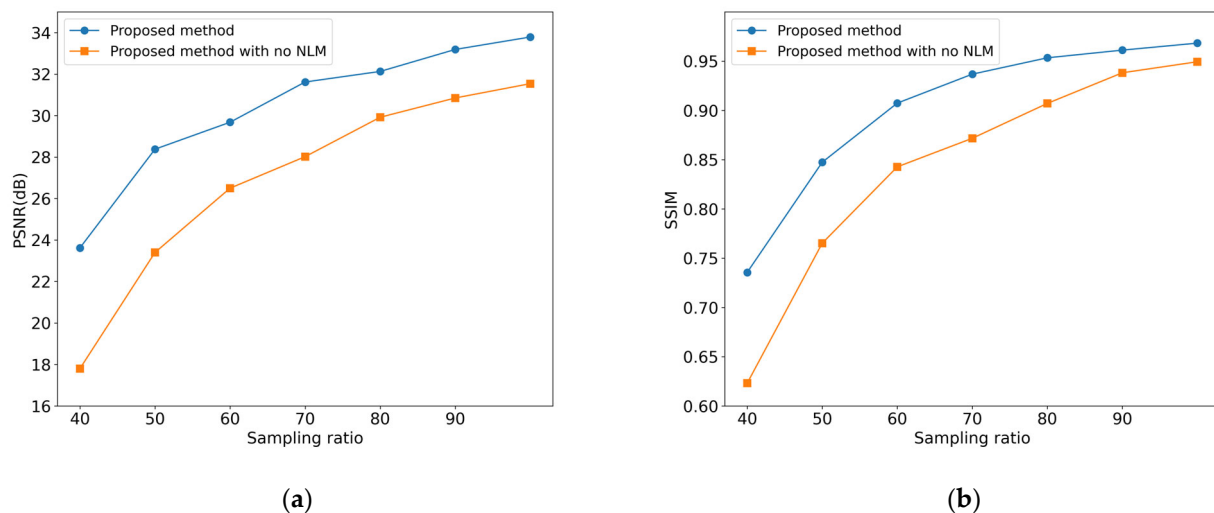
Table 3. Performance comparison of imaging speed.

Method	Time(s)
CSA	0.36
L1 + TV	13.5
PPB	116.1
SAE-Net	93.4
Proposed method	71.8

3.1.3. Ablation Experiment

In order to verify the improvement effect of the proposed non-local module on the image reconstruction performance, we meticulously analyze and experiment with the method before and after the addition of the non-local module. Under the condition that the other parameters remain consistent, we compare the image reconstruction effects before and after the introduction of the nonlocal module. We exhaustively analyze the specific impact of this change on performance using PSNR and SSIM as evaluation metrics. This approach allows us to accurately assess the effectiveness of non-local modules in image reconstruction.

Figure 8 clearly illustrates the comparison of the image reconstruction results before and after applying the non-local module at different undersampling ratios. This vividly highlights the changes in overall image quality. Compared to the method without the non-local module, our approach demonstrates a significant advantage in maintaining the structural similarity and fidelity of the image. Furthermore, our method retains its performance advantage even when the sampling ratio is reduced. These experimental results validate the significant contribution of our added non-local operation, effectively improving the reconstruction quality of the SAR images.

**Figure 8.** Performance analysis of the non-local module adapted to the blind-spot network. (a) PSNR, (b) SSIM.

We further investigated the impact of asymmetric pixel-shuffling downsampling operation on model performance through ablation studies that examined their effects on image reconstruction capabilities. Holding other parameters constant, we assessed the image reconstruction performance both before and after implementing the asymmetric pixel-shuffling operation. We employed PSNR and SSIM as evaluation metrics to determine the influence of this operation across various sampling ratios.

Figure 9 illustrates the image reconstruction outcomes prior to and following the integration of asymmetric pixel-shuffling downsampling operation, with the sampling ratio serving as a variable factor. The inclusion of this technique resulted in varying degrees

of improvement in the reconstruction metrics, indicating an enhancement in the overall model's reconstruction performance. These experimental findings confirm the efficacy of the asymmetric pixel-shuffling downsampling operation in boosting the model's reconstruction capabilities, thereby supporting its potential utility in image processing applications.

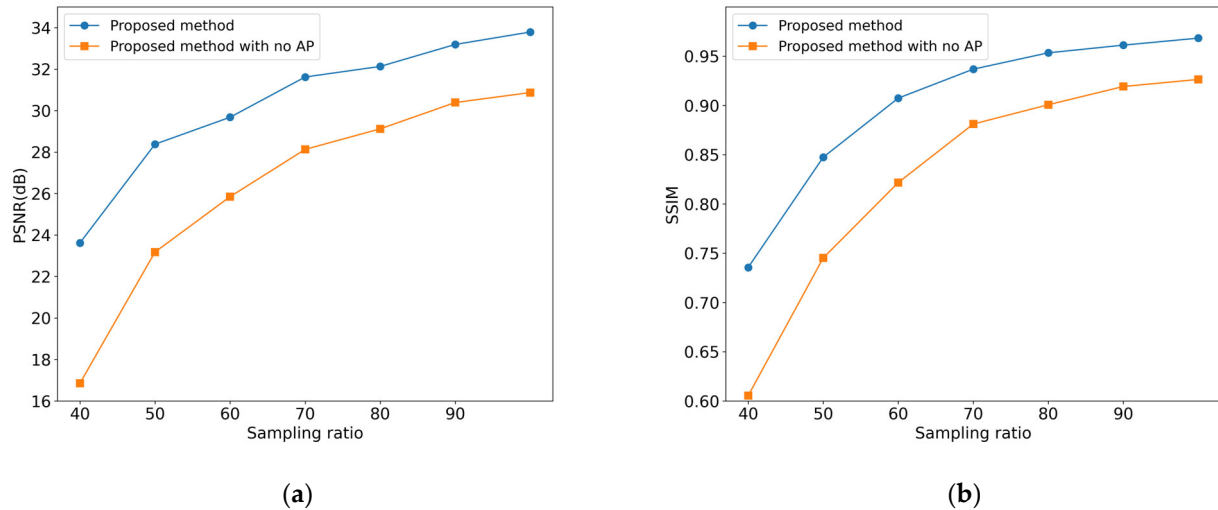


Figure 9. Performance analysis of the asymmetric pixel-shuffling downsampling operation. (a) PSNR, (b) SSIM.

3.2. Experiments Based on Real Data

In this section, we conduct experiments using real data from RADARSAT, which are processed using the range doppler algorithm to produce images with scattered spots. Next, these images are preprocessed and organized into a training set containing 1000 slices that are 256×256 pixels in size. After the network training was completed, we used the real data for validation and selected the CSA, L1 + TV, and PPB methods for comparison. We first conduct experiments in a plain region and evaluate the reconstruction performance of each method by calculating the ENL of its reconstruction results. Next, the ferry terminal area is selected for testing, and its reconstruction ability for texture and edge features is evaluated by analyzing the EPI of the images obtained using the different methods. Finally, for a scene containing ships, reconstruction is performed at different undersampling ratios, and the reconstruction performance is measured by calculating the TBR obtained using each method. The system parameters for the experiments based on real data are shown in Table 4.

Table 4. Parameters of the experiments based on real data.

Parameters	Value
Pulse duration time	41.7 μ s
Radar center frequency	5.3 GHz
Bandwidth	30.1 MHz
Range FM rate	721.5 MHz/ μ s
Pulse repetition frequency	1257 Hzwe

Figure 10 displays the imaging results obtained using various methods in the plain region. It is apparent from the figure that the imaging results from the CSA method exhibit significantly more pronounced noise, leading to a more distorted and less smooth appearance. In contrast, the other methods are more effective at suppressing noise, particularly our proposed method.

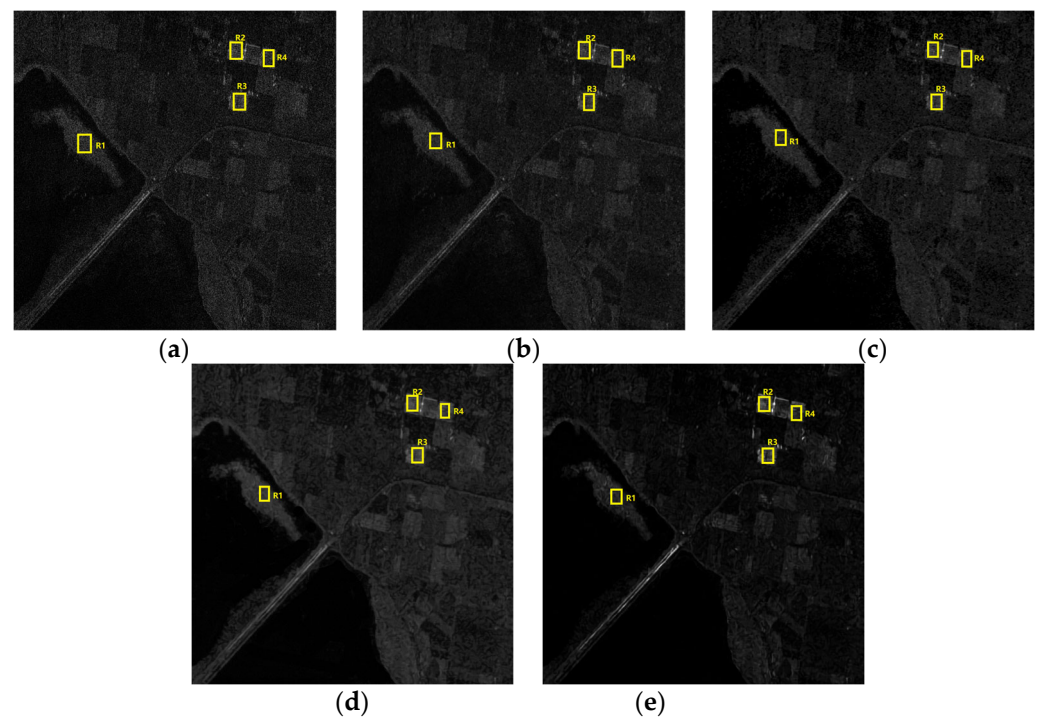


Figure 10. Imaging results of plain region with different methods. (a) CSA; (b) L1 + TV; (c) PPB; (d) SAE-Net; (e) proposed method.

To accurately assess the reconstruction accuracy, we selected four specific ground regions as distributed targets to compute the ENL from the imaging results shown in Figure 10. These regions are marked by yellow rectangles in Figure 10. Table 5 details the ENL values for these specific ground regions in the reconstructed images at various sampling ratios. Compared to the traditional CSA method, all other methods demonstrated improvements in ENL values, with the SAE-Net imaging method showing notably greater enhancements. Significantly, our proposed method achieved the highest ENL values among all the evaluated methods.

Table 5. ENL in the plain regions using different methods.

Region	Sample Ratio	ENL				
		CSA	L1 + TV	PPB	SAE-Net	Proposed
R1	80%	8.3661	26.3815	26.1593	30.7385	31.6412
	40%	5.1284	18.9374	19.3086	26.5819	26.7158
	20%	2.9834	10.8506	10.4376	17.4527	17.6397
R2	80%	10.6766	27.8015	28.1306	33.2653	34.3871
	40%	7.6893	22.7108	22.6751	29.5748	29.7635
	20%	4.8379	14.3796	14.8926	19.9734	20.8659
R3	80%	12.6434	28.6274	27.3973	34.1055	34.2976
	40%	8.1546	24.0141	23.8174	28.1974	28.6849
	20%	5.3985	15.7537	15.3719	18.8735	19.7593
R4	80%	7.9263	16.3578	17.1923	23.1883	23.7843
	40%	3.7850	10.1836	10.4783	16.2587	16.3705
	20%	1.8463	6.8962	6.9631	9.0754	9.8972

To evaluate the performance of our method more thoroughly, we expanded our analysis by selecting various regions for the reconstruction and ENL calculation. Figure 11 displays the imaging results obtained using the different methods. The yellow rectangular boxes in the figure highlight the regions where the ENL was calculated. The ENL values for

these regions, determined at different sampling rates using various methods, are presented in Table 6. Notably, our proposed method achieved the highest ENL values compared to all the other methods evaluated.

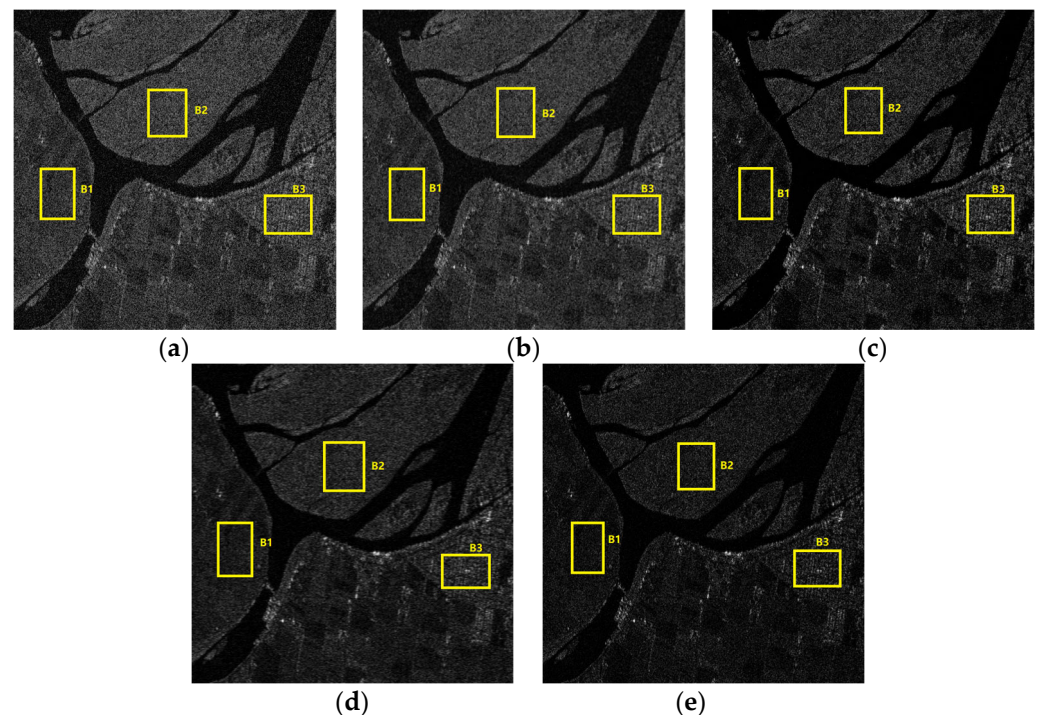


Figure 11. Imaging results using different methods. (a) CSA; (b) L1 + TV; (c) PPB; (d) SAE-Net; (e) proposed method.

Table 6. ENL of the target regions using different methods.

Region	Sample Ratio	ENL				
		CSA	L1 + TV	PPB	SAE-Net	Proposed
B1	80%	7.6391	23.5837	25.1948	29.7495	30.3051
	40%	5.2830	13.6036	15.2763	23.3084	23.9803
	20%	1.1739	7.3097	8.3685	15.8936	16.2569
B2	80%	8.2639	25.2583	26.6273	31.6682	32.4761
	40%	5.7386	20.1634	20.2123	24.7403	25.3647
	20%	2.2687	15.0726	15.2395	17.2079	18.4952
B3	80%	4.3824	18.3874	18.4193	22.6483	23.2208
	40%	2.4967	11.0638	10.8364	15.7438	16.1337
	20%	1.0261	7.3759	7.3872	11.2573	12.4519

Figure 12 demonstrates the imaging results of the various methods in the selected ferry terminal area at a sampling ratio of 60%. As can be seen from the figure, the imaging results of the CSA, L1 + TV, and PPB methods show energy diffusion and poor noise suppression under downsampling, leading to distortion of texture features. In contrast, SAE-Net and our proposed method effectively scatter noise in the downsampling scenario and is not affected by energy diffusion, though it does introduce slight detail distortion. Overall, our proposed method still slightly outperforms the traditional methods in terms of reconstruction performance, effectively suppressing noise even at lower sampling ratios, resulting in better preservation of texture details.

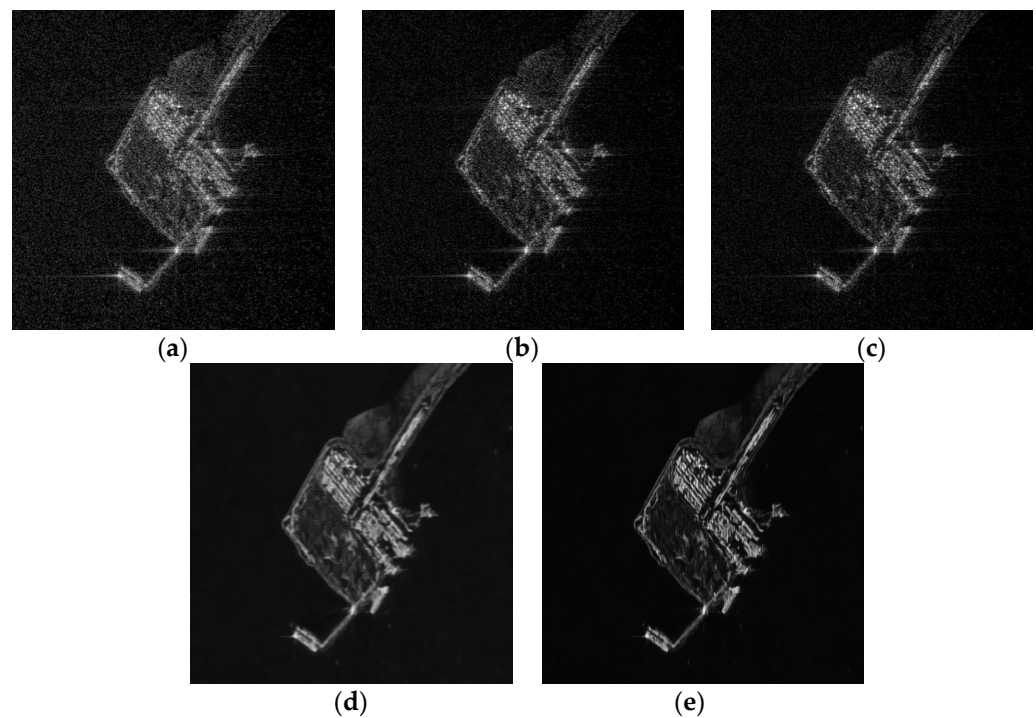


Figure 12. Imaging results of the ferry terminal area using different methods. (a) CSA; (b) L1 + TV; (c) PPB; (d) SAE-Net; (e) proposed method.

To more objectively demonstrate the advantages of our method, we utilized the EPI to evaluate the capability of preserving image edge details after removing scatter noise. According to the data in Table 7, the EPI values for our method surpassed those of L1 + TV, PPB, and SAE-Net in terms of edge detail retention. These results clearly indicate that our method exhibits superior performance compared to the other methods in retaining image edge details.

Table 7. EPI in the ferry terminal area for different methods.

Method	EPI
CSA	0.6023
L1 + TV	0.5986
PPB	0.6269
SAE-Net	0.6769
Proposed method	0.6875

Figure 13 illustrates the imaging results of the various methods when different undersampling ratios are used for the ships. At full sampling, the noise is more obvious in the imaging results of CSA compared to the other methods. When the sampling ratio is reduced, the noise in the CSA-reconstructed image becomes more pronounced, and both CSA and L1 + TV images exhibit more prominent sidelobes. In comparison, the other methods show good results in suppressing noise, while our proposed method demonstrates better results.

In order to more effectively demonstrate that our method outperforms other methods in terms of reconstruction quality, we objectively assessed the reconstruction quality of ships based on the TBR. Table 8 presents the TBR values for the imaging results obtained using various methods under both full sampling and undersampling conditions. The results demonstrate that our proposed method consistently outperforms the other methods in terms of TBR values across both conditions. This evidence further supports the superior reconstruction performance of our method in ship reconstructions.

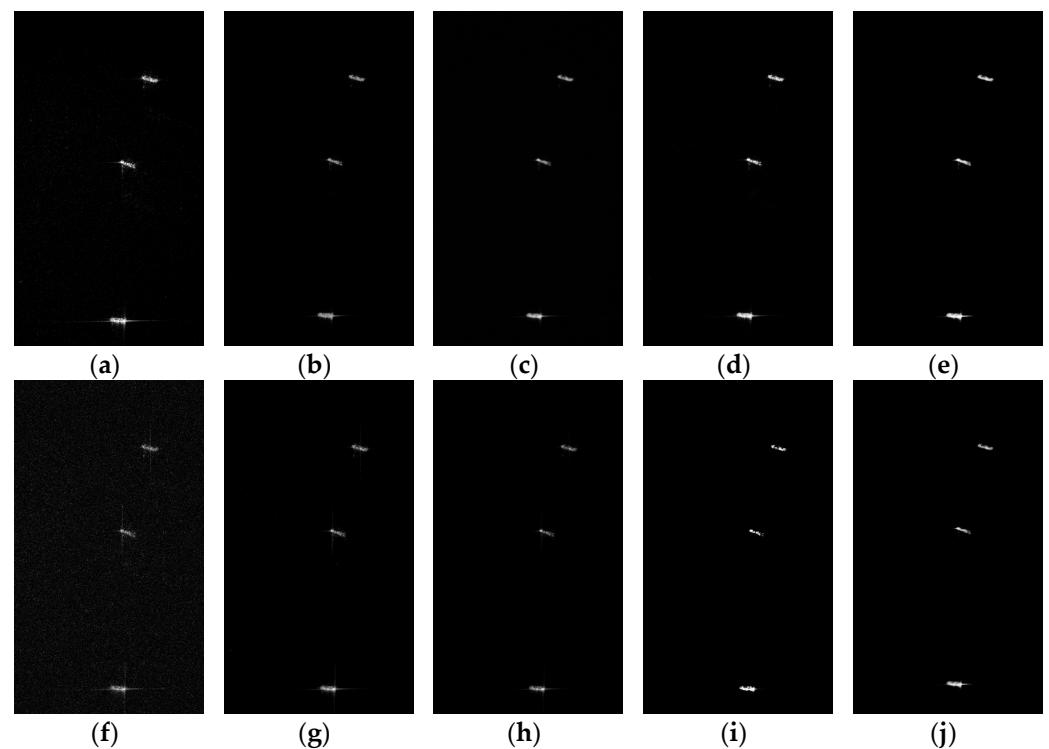


Figure 13. Imaging results of ships comparison across different methods and sampling ratios. The first row (a–e): Full Sampling Results (CSA, L1+TV, PPB, SAE-Net, Proposed method); the second row (f–j) 80% Sampling Results (CSA, L1+TV, PPB, SAE-Net, Proposed method).

Table 8. TBR for different methods at various sampling ratios.

Method	TBR			
	Full Sampling	80% Sampling	40% Sampling	20% Sampling
CSA	12.9651	5.7164	4.3759	2.5384
L1 + TV	43.7182	35.8763	23.7395	14.6936
PPB	45.8916	36.0846	23.8794	14.9527
SAE-Net	47.6825	40.8953	29.6731	20.8945
Proposed	47.7274	41.2981	30.5638	21.6473

The experiments conducted on real data demonstrate that our method surpasses the other techniques in reconstructing plain regions, ferry terminal areas, and ships. Both in preserving image details and in reducing noise, our method shows excellent performance for different scenes and conditions. This not only demonstrates the superiority of our method but also proves its generalization ability. These results emphasize the technical advantages of our method and demonstrate its wide potential in practical applications, especially in areas where accurate reconstruction and higher-quality images are required.

4. Discussion

In this study, we present an innovative self-supervised sparse SAR imaging method based on blind spot networks. Compared to other SAR imaging techniques, our method employs an echo simulation operator based on the chirp scaling algorithm (CSA) to construct an approximate observation model, and it trains an improved blind spot network to obtain a priori information about SAR images. This a priori information effectively addresses the issue of selecting regularization terms. Through experiments with both simulated and real data, we have comprehensively assessed the reconstruction capabilities of the algorithm at various sampling ratios. Our method performs superiorly in suppressing noise and clear targets, and it maintains clear imaging of the targets under downsampling

conditions. The effectiveness of our method compared to multiple conventional techniques is confirmed by comparative experiments performed on simulations and real data. These findings indicate that self-supervised learning holds substantial potential for SAR imaging. Moreover, integrating a priori knowledge through advanced deep learning techniques is essential for driving technological advancements in SAR imaging.

5. Conclusions

In this paper, we explore the integration of synthetic aperture radar (SAR) imaging technology with advanced deep learning techniques, particularly focusing on the innovative application of self-supervised non-local asymmetric pixel-shuffle blind spot networks. This self-supervised sparse SAR imaging method demonstrates significant potential in dealing with the scarcity of high-quality samples and offers a new direction for the development of SAR imaging technology.

Future research and improvements may pursue the following avenues: (1) Investigating how unsupervised data augmentation techniques can improve self-supervised learning to enhance network training outcomes while reducing reliance on labeled data. This approach could help us to better leverage the intrinsic data properties and inherent redundancies in SAR images. (2) Exploring the optimization of non-local module integration to increase its efficiency and effectiveness in capturing long-range dependencies within SAR images. Enhancing this aspect of the model could significantly improve its ability to process and analyze complex spatial relationships in SAR data. (3) Customizing models to better suit different imaging scenarios is crucial for future research. Tailoring deep learning architectures to specific applications ranging from military surveillance to environmental monitoring can ensure that the models effectively meet the diverse and specific needs of various SAR applications.

Author Contributions: Conceptualization, Z.Z. and Z.P.; methodology, Y.Z. and Y.T.; software, Y.T. and D.X.; writing—original draft preparation, Y.Z. and D.X.; writing—review and editing, B.W.-K.L. and Y.Z.; supervision, B.W.-K.L. and Z.Z.; funding acquisition, Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Guangdong Province under Grant 2021A1515012009.

Data Availability Statement: The data utilized in the simulated experiments of this study were sourced from the AIR-SARShip, which can be found here: (<https://radars.ac.cn/web/data/getData?newsColumnId=d25c94d7-8fe8-415f-a897-cb88657a8141&pageType=en>, accessed on 29 April 2024). The data utilized in the experiments based on real data in this study were sourced from the RADARSAT dataset, which can be found here: (http://us.artechhouse.com/Assets/downloads/Cumming_058-3.zip, accessed on 29 April 2024).

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

References

1. Wei, S.; Su, H.; Ming, J.; Wang, C.; Yan, M.; Kumar, D.; Shi, J.; Zhang, X. Precise and Robust Ship Detection for High-Resolution SAR Imagery Based on HR-SDNet. *Remote Sens.* **2020**, *12*, 167. [CrossRef]
2. Zhou, F.; Zhao, B.; Tao, M.; Bai, X.; Chen, B.; Sun, G. A Large Scene Deceptive Jamming Method for Space-Borne SAR. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 4486–4495. [CrossRef]
3. Donoho, D.L. Compressed Sensing. *IEEE Trans. Inform. Theory* **2006**, *52*, 1289–1306. [CrossRef]
4. Xu, Z.; Zhang, B.; Zhou, G.; Zhong, L.; Wu, Y. Sparse SAR Imaging and Quantitative Evaluation Based on Nonconvex and TV Regularization. *Remote Sens.* **2021**, *13*, 1643. [CrossRef]
5. Jiang, C.L.; Zhang, B.C.; Fang, J.; Zhe, Z.; Hong, W.; Wu, Y.R.; Xu, Z.B. Efficient ℓ_q Regularisation Algorithm with Range–Azimuth Decoupled for SAR Imaging. *Electron. Lett.* **2014**, *50*, 204–205. [CrossRef]
6. Raney, R.K.; Runge, H.; Bamler, R.; Cumming, I.G.; Wong, F.H. Precision SAR Processing Using Chirp Scaling. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 786–799. [CrossRef]
7. Güven, H.E.; Güngör, A.; Çetin, M. An Augmented Lagrangian Method for Complex-Valued Compressed SAR Imaging. *IEEE Trans. Comput. Imaging* **2016**, *2*, 235–250. [CrossRef]

8. Deledalle, C.-A.; Denis, L.; Tupin, F. Iterative Weighted Maximum Likelihood Denoising with Probabilistic Patch-Based Weights. *IEEE Trans. Image Process.* **2009**, *18*, 2661–2672. [[CrossRef](#)] [[PubMed](#)]
9. Yonel, B.; Mason, E.; Yazici, B. Deep Learning for Waveform Estimation and Imaging in Passive Radar. *IET Radar Sonar Navig.* **2019**, *13*, 915–926. [[CrossRef](#)]
10. Zhao, S.; Ni, J.; Liang, J.; Xiong, S.; Luo, Y. End-to-End SAR Deep Learning Imaging Method Based on Sparse Optimization. *Remote Sens.* **2021**, *13*, 4429. [[CrossRef](#)]
11. Pu, W. Deep SAR Imaging and Motion Compensation. *IEEE Trans. Image Process.* **2021**, *30*, 2232–2247. [[CrossRef](#)] [[PubMed](#)]
12. Gao, J.; Deng, B.; Qin, Y.; Wang, H.; Li, X. Enhanced Radar Imaging Using a Complex-Valued Convolutional Neural Network. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 35–39. [[CrossRef](#)]
13. Zhang, H.; Ni, J.; Xiong, S.; Luo, Y.; Zhang, Q. SR-ISTA-Net: Sparse Representation-Based Deep Learning Approach for SAR Imaging. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 4513205. [[CrossRef](#)]
14. Pu, W. SAE-Net: A Deep Neural Network for SAR Autofocus. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5220714. [[CrossRef](#)]
15. Krull, A.; Buchholz, T.-O.; Jug, F. Noise2Void—Learning Denoising From Single Noisy Images. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 2124–2132.
16. Wu, Q.; Ji, X.; Gu, Y.; Xiang, J.; Quan, G.; Li, B.; Zhu, J.; Coatrieux, G.; Coatrieux, J.-L.; Chen, Y. Unsharp Structure Guided Filtering for Self-Supervised Low-Dose CT Imaging. *IEEE Trans. Med. Imaging* **2023**, *42*, 3283–3294. [[CrossRef](#)] [[PubMed](#)]
17. de Negreiros, A.C.S.V.; Giraldo, G.; Werner, H.; Santos, Í.M.F. Self-Supervised Image Denoising Methods: An Application in Fetal MRI. In *Workshop de Visão Computacional (WVC)*; SBC: São Bernardo do Campo, Brazil, 2023; pp. 137–141.
18. Huang, C.; Hong, D.; Yang, C.; Cai, C.; Tao, S.; Clawson, K.; Peng, Y. A New Unsupervised Pseudo-Siamese Network with Two Filling Strategies for Image Denoising and Quality Enhancement. *Neural Comput. Appl.* **2023**, *35*, 22855–22863. [[CrossRef](#)]
19. Lee, W.; Son, S.; Lee, K.M. AP-BSN: Self-Supervised Denoising for Real-World Images via Asymmetric PD and Blind-Spot Network. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 17704–17713.
20. Molini, A.B.; Valsesia, D.; Fracastoro, G.; Magli, E. Speckle2Void: Deep Self-Supervised SAR Despeckling with Blind-Spot Convolutional Neural Networks. *IEEE Trans. Geosci. Remote Sens.* **2020**, *60*, 5204017. [[CrossRef](#)]
21. Alver, M.B.; Saleem, A.; Cetin, M. Plug-and-Play Synthetic Aperture Radar Image Formation Using Deep Priors. *IEEE Trans. Comput. Imaging* **2021**, *7*, 43–57. [[CrossRef](#)]
22. Boyd, S. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *FNT Mach. Learn.* **2010**, *3*, 1–122. [[CrossRef](#)]
23. Fang, J.; Xu, Z.; Zhang, B.; Hong, W.; Wu, Y. Fast Compressed Sensing SAR Imaging Based on Approximated Observation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 352–363. [[CrossRef](#)]
24. Zhou, Y.; Jiao, J.; Huang, H.; Wang, Y.; Wang, J.; Shi, H.; Huang, T. When AWGN-Based Denoiser Meets Real Noises. *Proc. AAAI Conf. Artif. Intell.* **2020**, *34*, 13074–13081. [[CrossRef](#)]
25. Liu, D.; Wen, B.; Fan, Y.; Loy, C.C.; Huang, T.S. Non-Local Recurrent Network for Image Restoration. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: New York, NY, USA, 2018; Volume 31.

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.