



Article

Nonsparse SAR Scene Imaging Network Based on Sparse Representation and Approximate Observations

Hongwei Zhang ¹, Jiacheng Ni ^{1,*}, Kaiming Li ¹, Ying Luo ¹ and Qun Zhang ^{1,2}

¹ School of Information and Navigation, Air Force Engineering University, Xi'an 710077, China; zhanghong1941@mail.nwpu.edu.cn (H.Z.); likaiming1982@163.com (K.L.); luoying2002521@163.com (Y.L.); afeuzq@163.com (Q.Z.)

² Key Laboratory for Information Science of Electromagnetic Waves (Ministry of Education), Fudan University, Shanghai 200433, China

* Correspondence: littlenjc@sina.com

Abstract: Sparse-representation-based synthetic aperture radar (SAR) imaging technology has shown superior potential in the reconstruction of nonsparse scenes. However, many existing compressed sensing (CS) methods with sparse representation cannot obtain an optimal sparse basis and only apply to the sensing matrix obtained by exact observation, resulting in a low image quality occupying more storage space. To reduce the computational cost and improve the imaging performance of nonsparse scenes, we formulate a deep learning SAR imaging method based on sparse representation and approximated observation deduced from the chirp-scaling algorithm (CSA). First, we incorporate the CSA-derived approximated observation model and a nonlinear transform function within a sparse reconstruction framework. Second, an iterative shrinkage threshold algorithm is adopted to solve this framework, and the solving process is unfolded as a deep SAR imaging network. Third, a dual-path convolutional neural network (CNN) block is designed in the network to achieve the nonlinear transform, dramatically improving the sparse representation capability over conventional transform-domain-based CS methods. Last, we improve the CNN block to develop an enhanced version of the deep SAR imaging network, in which all the parameters are layer-varied and trained by supervised learning. The experiments demonstrate that our proposed two imaging networks outperform conventional CS-driven and deep-learning-based methods in terms of computing efficiency and reconstruction performance of nonsparse scenes.

Keywords: synthetic aperture radar (SAR) imaging; compressive sensing (CS); deep learning (DL); approximated observation; sparse representation



Citation: Zhang, H.; Ni, J.; Li, K.; Luo, Y.; Zhang, Q. Nonsparse SAR Scene Imaging Network Based on Sparse Representation and Approximate Observations. *Remote Sens.* **2023**, *15*, 4126. <https://doi.org/10.3390/rs15174126>

Academic Editors: Hui Bi, Daiyin Zhu and Jingjing Zhang

Received: 17 July 2023

Revised: 8 August 2023

Accepted: 16 August 2023

Published: 22 August 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) can achieve high-resolution images by wideband signal pulse compression and virtual aperture synthesis, which is not affected by time and weather [1]. As an active microwave imaging system, SAR transmits a series of pulses and receives echoes, which are generally processed by matched filter (MF)-based SAR imaging methods, such as the range-Doppler algorithm (RDA) [2], chirp-scaling algorithm (CSA) [3], and back-projection algorithm (BPA) [4]. With the development of these traditional MF-based focusing algorithms, more measurements of the echo data are required to satisfy the demand for increased resolution, which makes the storage and processing of SAR data difficult for the system hardware. Over the past decade, as compressed sensing (CS) and related deep learning (DL) methods have emerged, SAR imaging technology has progressed to high-dimensional data and plays an increasingly significant role in military and civilian fields [5].

1.1. Previous Work

Different from the Nyquist sampling theorem, CS provides the possibility of accurately reconstructing sparse signals using a small quantity of observation data [6,7]. CS-driven SAR imaging, also known as sparse SAR imaging, has recently become a controversial issue and has received extensive attention [8]. These studies primarily focus on the potential application of SAR imaging technology innovated by CS theory. They have indicated that if the SAR imaging result is a sparse signal in the spatial domain or transform domain, it can be reconstructed with high precision from an incomplete SAR echo by constructing an optimization problem [9,10]. Sparse SAR imaging can be considered an inverse optimization problem, which can be directly solved using iterative algorithms when the true scattered field is sparse in the spatial domain, e.g., the iterative shrinkage threshold algorithm (ISTA) [11], approximate message passing (AMP) algorithm [12], and alternating direction method of multipliers (ADMM) [13]. However, the scattered field of the SAR scene is always nonsparse or weakly sparse, and thus, it needs to be sparsely represented. This limitation restricts the application of sparse SAR imaging.

To solve the above-mentioned limitation, reconstruction methods that combine iterative algorithms with sparse representation, which formulate the reconstruction framework with sparsity constraints, have been developed. Rilling et al. [14] investigated a transform-domain-based reconstruction technique in which only a specific transform basis, i.e., a wavelet basis, was employed to transform the nonsparse areas. In practice, the features of SAR images are complicated, so it is difficult to obtain a common transform basis or sparse representation. As a result, joint sparse transform-based approaches, including combined dictionaries, structured sparse representation, and mixed sparse representation [15–18], have been successively proposed to enhance the sparsity exploration ability of the SAR reconstruction framework. In [15], an image formation technique with sparse representation, which relied on combined dictionaries to obtain multiple feature-enhanced reconstructed images, was developed. By incorporating an adaptive sparse representation space into the SAR reconstruction framework, Shen et al. [16] proposed a structural-sparse-representation-assisted SAR imaging algorithm in which SAR image reconstruction and adaptive sparse space updating were constructed as a joint optimization problem. However, the above studies adopted the exact observation model to build an optimization problem for reconstructing nonsparse SAR images. Because of the vectorization of 2D scattering coefficient matrices, such exact observation models require much more computational resources, thus rendering them inefficient for large-scale matrices. Fortunately, a sparse reconstruction framework based on approximated observations was proposed in [19] to solve this problem. On this basis, Li et al. [18] proposed the mixed sparse representation for approximated observation SAR imaging, dramatically improving the computational efficiency of large-scale matrices while guaranteeing the sparsity of related signals. Although this kind of imaging algorithm can eliminate the difficulties in storing and processing data, it remains limited by two general shortcomings of CS-driven SAR imaging: the difficulty of determining the optimal parameter values and the complexity of the reconstruction mechanism.

The recent development of DL technology has triggered extensive research on SAR image restoration, providing new research methods to overcome the shortcomings of conventional CS-driven SAR imaging. Convolutional neural networks (CNNs), as powerful networks in DL technology, have been employed in the field of SAR imaging to advance state-of-the-art technology for image reconstruction and target focusing. A representative study on embedding CNN into reconstruction was published in [20], in which a complex-valued CNN structure was introduced to enhance the imaging functionality of the inverse SAR (ISAR) signal processing framework. In [21], Mu et al. proposed a deep CNN imaging method named DeepImaging to refocus SAR moving targets and eliminate residual clutter. Similarly, U-Net, a CNN-based network, was improved in [22] to achieve a high-resolution focused image by addressing the defocused data generated by the RDA. In [23], a deep convolutional encoder, also based on a CNN, was combined with a residual network to approximate the SAR imaging process of the traditional RDA. In this case, the encoder

mapped the echo data to the SAR image, and the residual network subsequently enhanced the image resolution. For all the above methods, the CNN is utilized as a “black box” to learn the mapping from sampled data or defocused images to high-quality images. Consequently, these works lack interpretable characteristics since they do not explicitly investigate imaging models or explore image sparsity.

In contrast to traditional neural networks, deep unfolding methods combine model-driven iterative algorithms and data-driven deep networks. This approach solves the inverse imaging problem by converting the traditional CS-driven SAR reconstruction algorithm to a deep network, which improves imaging performance and interpretability. Thus, deep unfolding is more appropriate for SAR image formation than CNN. Inspired by this concept, several deep unfolding imaging methods based on iterative algorithms have emerged in the field of SAR imaging, but they remain in the preliminary stage. In [24], Yonel et al. unfolded the ISTA on a recurrent neural network architecture to establish an inverse problem solver for passive SAR reconstruction. Then, the solver was further transformed into an autoencoder (named ISTA-Net) by adding a forward model as the decoder so that it could be trained by unsupervised learning. However, due to its vectorized SAR imaging model, ISTA-Net wastes storage space, which limits the scope of imaging spaces. An extension network named range-migration kernel-based ISTA-Net (RMIST-Net), which was presented in [25], was designed to directly process 2D complex-valued echo data instead of vectorization. Similarly, a target-oriented SAR imaging network, obtained by unfolding an MF-based ADMM iterative solution, was proposed in [26] to enhance the signal-to-clutter ratio of the desired target. For SAR imaging of moving targets, deep-unfolding-based focusing methods have been proposed to reconstruct sparse ship targets [27,28]. In addition, deep unfolding technology can also be applied to the sparse imaging of ISAR, such as ADMM-Net [29] and AF-AMPNet [30]. Sufficient experiments verified that deep-unfolding-based imaging methods outperformed conventional CS-driven reconstruction algorithms. Note that the above microwave imaging networks for stationary and moving targets can only be applied to sparse scenes. To overcome this limitation, an RDA-based deep unfolding imaging network named RDA-Net is proposed in [31]. Although RDA-Net can reconstruct nonsparse SAR scenes by learning the corresponding compensation matrices of RDA, setting these matrices as learnable parameters still requires a substantial amount of storage space. Furthermore, due to the lack of sparse representation modules, the robustness and accuracy of RDA-Net with unsupervised learning are low in nonsparse scenes. In our previous work [5], we proposed a 1D SAR imaging network with sparse representation capability, but it required a much higher computational cost due to the exact observation model. As a result, it is difficult for current deep unfolding technology to effectively solve the SAR image reconstruction problem for complex scenes while saving storage space.

1.2. Motivation

After analyzing the previous works on SAR imaging methods, we came to the following conclusions. (1) Traditional CS methods require the integration of multiple algorithms to solve the nonsparse SAR scene imaging problem. The lack of optimal sparse basis results in unsatisfactory imaging efficiency. (2) The current DL-based methods cannot achieve satisfactory reconstruction results for complex and large-scale scenes.

In this paper, we investigate a deep SAR learning imaging method, which was motivated by the desire to enhance the working efficiency and image quality of nonsparse SAR scene imaging. Therefore, we propose a novel deep unfolding strategy for SAR imaging of complex scenes to reduce the computational burden and improve the reconstruction performance.

1.3. Main Work and Contributions

The proposed strategy utilizes the significant representation performance of a CNN to accurately reconstruct nonsparse scenes, rather than learning the large-scale MF compen-

sation matrices in SAR imaging networks such as RDA-Net. First, we establish a sparse reconstruction framework for complex scenes by integrating a CSA-derived approximated observation model and a nonlinear transform function. Second, the overall architecture of an end-to-end deep SAR imaging network is determined by unfolding the iterative process of solving this framework with the ISTA. To improve the imaging accuracy of nonsparse SAR scenes, we design a learnable dual-path CNN block that outperforms conventional nonlinear transform functions in terms of sparse transform. In addition, we develop an enhanced version of the deep SAR imaging network by strengthening the sparse representation capability of the CNN block. Experimental results demonstrate that our proposed imaging networks (abbreviated as nets) achieve impressive reconstruction results of nonsparse SAR scenes while reducing the computational cost. The main contributions of this article are summarized as follows:

- (1) Compared to the 1D exact observation model, the advantages and feasibility of the 2D approximated observation model are analyzed. Then, by decoupling the CSA operation process, we develop a specific 2D SAR sparse reconstruction framework and introduce its imaging mechanism with the ISTA as a solution example.
- (2) Combining the advantages of deep unfolding and a CNN, a deep SAR imaging net and its improved versions (SR-CSA-Net and SR-CSA-Net-plus, respectively) are developed for nonsparse SAR reconstruction, with internal modules designed as dual-path structures to address complex-valued SAR echoes as input to imaging nets.
- (3) A CNN-based nonlinearity module is designed in SR-CSA-Net to automatically learn the optimal sparse transform. Furthermore, we redesign the nonlinearity module with skip connections in SR-CSA-Net-plus to fully exploit the high-frequency component of the scattering coefficient and improve the reconstruction performance.
- (4) The proposed imaging nets are trained end-to-end via supervised learning, and the parameters are learned using cleverly designed loss functions. Specifically, these well-designed loss functions simultaneously take into account both structure symmetry constraints and reconstruction discrepancies.
- (5) Extensive experiments verify that our two proposed imaging nets outperform conventional CS-driven sparse reconstruction algorithms and several existing deep unfolding imaging methods by large margins.

This paper is organized as follows: Section 2 discusses the CSA-derived SAR approximated observation imaging model. Section 3 illustrates the details of sparse-representation-based deep SAR imaging nets. In Section 4, experimental results and performance analyses are presented. Section 5 concludes this paper.

Notations: A vector is denoted by the bold lowercase letter $\mathbf{x}$, a matrix is denoted by the bold capital letter $\mathbf{X}$, and a variable or scalar is denoted by the lowercase letter x . The real field domain and complex field domain are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. In addition, $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote the transpose, conjugate, and Hermitian transpose, respectively, of matrices or operators.

2. SAR Imaging Model Based on Approximated Observation

In Section 2.1, the elementary theory of sparse SAR imaging based on the exact observation model is summarized. To reduce the computational cost, we investigate the general formalization of the 2D approximated observation model and verify its feasibility in Section 2.2. In Section 2.3, we propose a concrete example derived from the CSA procedure to explicitly construct the 2D approximated observation model. The CSA-derived model is generalized to a sparse-representation-based SAR imaging model in Section 2.4.

2.1. One-Dimensional Exact Observation Model for SAR Imaging

Vertical side-looking, strip-mode SAR was adopted for sparse imaging in this paper, and the geometric structure of the imaging system is shown in Figure 1. The observation area could be regarded as a series of scattering centers located on the $P \times Q$ grids, in which the grid coordinates along the range direction (x -axis) and azimuth (y -axis) direction are

$p = 1, 2, \dots, P$ and $q = 1, 2, \dots, Q$, respectively. Therefore, the scattering coefficient of the imaging scene is expressed as the following scattering coefficient matrix:

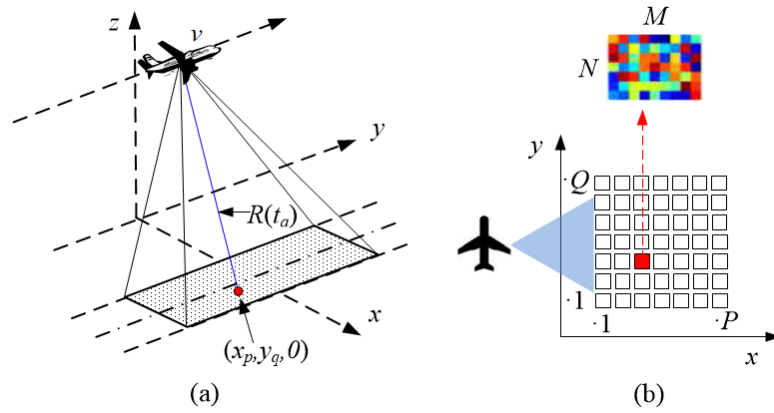


Figure 1. Geometric structure of the SAR imaging system: (a) 3D model; (b) 2D model.

$$\Xi = \begin{bmatrix} \sigma(x_1, y_1) & \cdots & \sigma(x_1, y_Q) \\ \vdots & \sigma(x_p, y_q) & \vdots \\ \sigma(x_P, y_1) & \cdots & \sigma(x_P, y_Q) \end{bmatrix}_{P \times Q} \quad (1)$$

where $\sigma(x_p, y_q)$ denotes the scattering coefficient at coordinate (x_p, y_q) . As shown in Figure 1b, it is assumed that the range and azimuth sampling points on each grid are $m = 1, 2, \dots, M$ and $n = 1, 2, \dots, N$, respectively.

The 1D exact observation model of sparse SAR imaging represents the relationship between SAR echo signals and the scattering distribution of imaging scenes. The down-sampling observation model of the echo signal is expressed as

$$\mathbf{s}_d = \mathbf{\Psi} \cdot \mathbf{s} + \mathbf{n}_0 = \mathbf{\Psi} \mathbf{\Phi} \boldsymbol{\sigma} + \mathbf{n}_0 \quad (2)$$

where $\mathbf{s} \in \mathbb{C}^{MN \times 1}$ is the vectorized original echo signal; $\mathbf{s}_d \in \mathbb{C}^{M'N' \times 1}$ is the complex-valued echo vector after downsampling; $\mathbf{n}_0$ is the additional white Gaussian noise (AWGN); $\mathbf{\Psi} \in \mathbb{R}^{M'N' \times MN}$ is the downsampling matrix, which generally adopts a partial identity matrix; $\mathbf{\Phi} \in \mathbb{C}^{MN \times PQ}$ is the measurement matrix composed of phase terms; and $\boldsymbol{\sigma} = \text{vec}(\Xi) \in \mathbb{C}^{PQ \times 1}$ is the scattering coefficient vector, where $\text{vec}(\cdot)$ is the vector operator. If the imaging scene is sparse enough and the measurement matrix $\mathbf{\Phi}$ meets the conditions of the restricted isometry property (RIP) [32], the unknown SAR image $\boldsymbol{\sigma}$ can be reconstructed from the known echo vector $\mathbf{s}$ by solving the following optimization problem based on ℓ_1 decoupling [24]

$$\hat{\boldsymbol{\sigma}} = \arg \min_{\boldsymbol{\sigma}} \{ \|\mathbf{s}_d - \mathbf{\Psi} \mathbf{\Phi} \boldsymbol{\sigma}\|_2^2 + \lambda \|\boldsymbol{\sigma}\|_1 \} \quad (3)$$

where $\hat{\boldsymbol{\sigma}}$ approximates the true scattered field $\boldsymbol{\sigma}$, $\lambda \|\cdot\|_1$ denotes the ℓ_1 norm-based regularization constraint term, and λ is the predefined regularization parameter.

2.2. Two-Dimensional Approximated Observation Model for SAR Imaging

The 1D exact observation model vectorizes the 2D echo signal and scattering coefficient matrix. In this case, the size of the measurement matrix $\mathbf{\Phi}$ will rapidly increase with the expansion of the echo signal and imaging scene scale. The model introduces a substantial computational and storage burden to the sparse imaging processing of the 1D exact observation model and simultaneously limits its application in sparse imaging. Therefore, a 2D approximated observation model is designed in this subsection to replace the conventional 1D exact observation model, reducing computational and storage costs by a large margin.

It is known that the MF algorithm focuses the echo signal s in the 2D frequency domain. It is assumed that $\mathbf{M}$ represents the MF operation procedure, with which the scattering coefficient vector σ is approximately reconstructed by $\hat{\sigma} = \mathbf{M}s$, where $\mathbf{M}$ satisfies $\mathbf{M}\Phi \approx \mathbf{I}$. It is difficult to establish a 2D approximated observation model by accurately decoupling the measurement matrix Φ [31]. Conversely, the MF operation procedure $\mathbf{M}$ can be decoupled into a series of operators processed in the range or azimuth dimension, and the inverse operation of $\mathbf{M}$ denoted as $\mathbf{M}^{-1}$ is also able-decoupled. Hence, the inverse operation $\mathbf{M}^{-1}$, as an acceptable alternative to Φ , is integrated with the CS method to design the 2D approximated observation model. We further establish a general conclusion as

$$\mathbf{G} = \mathbf{M}^{-1} \approx \Phi \quad (4)$$

where $\mathbf{M}$ represents any high-resolution MF algorithm, such as RDA or CSA, and $\mathbf{G}$ is the generalized inverse operation of $\mathbf{M}$, which is referred to as the approximated observation matrix and can be used to implicitly approximate Φ . By substituting $\mathbf{G}$ for Φ , the observation model in Equation (2) is rewritten as an approximated observation model

$$s_d = \Psi \Phi \sigma + n_0 = \Psi \mathbf{G} \hat{\sigma} + n_0. \quad (5)$$

To further establish the 2D approximated observation model, the Kronecker product decomposition, a typical method to reduce the size of the matrix, is introduced in this paper. Assuming that the approximated observation matrix $\mathbf{G} \in \mathbb{C}^{MN \times PQ}$ is expressed as $\mathbf{G} = \mathbf{G}_a^T \otimes \mathbf{G}_r$ via the Kronecker product decomposition, where $\otimes$ represents the Kronecker product, and $\mathbf{G}_r \in \mathbb{C}^{M \times P}$ and $\mathbf{G}_a \in \mathbb{C}^{Q \times N}$ denote the observation matrices of the range dimension and azimuth dimension, respectively. Note that the coupling characteristic of measurement matrix Φ makes its range measurement matrix Φ_r and azimuth measurement matrix Φ_a unable to be wholly separated. Thus, we utilize $\mathbf{G}$ to approximate the observation matrix Φ . On this basis, the 2D approximated observation model is derived as follows:

$$\begin{aligned} s_d &= \Psi \mathbf{G} \hat{\sigma} + n_0 \\ &= \Psi (\mathbf{G}_a^T \otimes \mathbf{G}_r) \hat{\sigma} + n_0 \\ &= \Psi (\mathbf{G}_a^T \otimes \mathbf{G}_r) \text{vec}(\hat{\Sigma}) + \text{vec}(\mathbf{N}_0) \\ &= \text{vec}(\Psi_r \mathbf{G}_r \hat{\Sigma} \mathbf{G}_a \Psi_a + \mathbf{N}_0) \\ &= \text{vec}(\Psi_r \mathbf{S} \Psi_a + \mathbf{N}_0) \\ &= \text{vec}(\mathbf{S}_d) \end{aligned} \quad (6)$$

where $\mathbf{S} \in \mathbb{C}^{M \times N}$ is the original 2D echo matrix; $\mathbf{S}_d \in \mathbb{C}^{M' \times N'}$ denotes the 2D echo matrix after downsampling; $\Psi_r \in \mathbb{R}^{M' \times M}$ formed by M' rows of an M -by- M identity matrix and $\Psi_a \in \mathbb{R}^{N \times N'}$ formed by N' columns of an N -by- N identity matrix are the downsampling matrices in the range dimension and azimuth dimension, respectively; and $\hat{\Sigma} \in \mathbb{C}^{P \times Q}$ is the reconstructed 2D scattering coefficient matrix. The 2D approximated observation model and its downsampling form with AWGN then become

$$\begin{cases} \mathbf{S} = \mathbf{G}_r \hat{\Sigma} \mathbf{G}_a \\ \mathbf{S}_d = \Psi_r \mathbf{S} \Psi_a + \mathbf{N}_0 \end{cases} \quad (7)$$

In summary, due to the challenge of decomposing Φ with the Kronecker product decomposition rule, we adopt the approximated observation matrix $\mathbf{G}$ derived from the MF algorithm to construct the 2D approximated observation model rather than searching for a direct application to the observation matrix Φ .

2.3. CSA-Derived Approximated Observation Model

Since the MF algorithm is a linear process including a phase multiplication and fast Fourier transform (FFT) in the range and azimuth dimensions, the approximated

observation matrices $\mathbf{G}_r$ and $\mathbf{G}_a$ are implicitly represented by a series of suboperations obtained from the inversion of the classical MF focusing methods (e.g., RDA, CSA, and BPA). In this subsection, an admissible 2D approximated observation model derived from CSA is examined to show how to achieve the analytic representation of Equation (7).

CSA is a dual-domain hybrid processing imaging algorithm that includes three steps of phase multiplication: (1) differential range cell migration correction (RCMC) (range-Doppler domain), (2) range compression and bulk RCMC (2D frequency domain), and (3) azimuth compression (range-Doppler domain). Then, the imaging operator denoted as $\mathbf{M}(\cdot)$ and the approximated observation operator denoted as $\mathbf{G}(\cdot)$ are described as

$$\begin{aligned}\hat{\mathbf{E}} &= \mathbf{M}(\mathbf{S}) \\ &= \left\{ \mathbf{F}_r^H \langle \mathbf{F}_r [(\mathbf{S}\mathbf{F}_a) \circ \mathbf{H}_1] \circ \mathbf{H}_2 \rangle \circ \mathbf{H}_3 \right\} \mathbf{F}_a^H\end{aligned}\quad (8)$$

$$\begin{aligned}\mathbf{S} &= \mathbf{G}(\hat{\mathbf{E}}) \\ &= \left\{ \mathbf{F}_r^H \langle \mathbf{F}_r [(\hat{\mathbf{E}}\mathbf{F}_a) \circ \mathbf{H}_3^*] \circ \mathbf{H}_2^* \rangle \circ \mathbf{H}_1^* \right\} \mathbf{F}_a^H\end{aligned}\quad (9)$$

where $\circ$ denotes the Hadamard product; $\mathbf{F}_r$ and $\mathbf{F}_a$ are the FFT matrices along the range direction and azimuth direction, respectively; $\mathbf{F}^H$ represents the inverse FFT (IFFT) matrix; and $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_3$ are the quadratic phase function for chirp scaling operation, the phase function for range compression and bulk RCMC, and the phase function for azimuth compression and residual phase compensation, respectively. Since the phase multiplication operation is a unitary transformation, $\mathbf{G}(\cdot)$ is obtained by deriving the inverse operation of $\mathbf{M}(\cdot)$. Furthermore, $\mathbf{G}(\cdot)$ is also referred to as the CSA operator, which can be regarded as an implicit operation procedure of decomposing $\mathbf{G}$ with the Kronecker product decomposition rule, avoiding the vectorization of echoes and scattering coefficients.

The above operators have a prominent property that both $\mathbf{M}(\cdot)$ and $\mathbf{G}(\cdot)$ are linear processes whose relationship, similar to Equation (4), is described as $\mathbf{G}(\cdot) = \mathbf{M}(\cdot)^{-1} = \mathbf{M}(\cdot)^H$. This conclusion can be further generalized to $\mathbf{G} = \mathbf{M}^H$.

Theorem 1. *The imaging operator $\mathbf{M}(\cdot)$ and approximated observation operator $\mathbf{G}(\cdot)$ are linear operators, and the relationship satisfies $\mathbf{G}(\cdot) = \mathbf{M}(\cdot)^H$ and $\mathbf{G} = \mathbf{M}^H$.*

Proof. Based on Equations (8) and (9) above, the vector form of $\hat{\mathbf{E}}$ and $\mathbf{S}$ is expressed as $\text{vec}(\hat{\mathbf{E}}) = \text{vec}(\mathbf{M}(\mathbf{S}))$ and $\text{vec}(\mathbf{S}) = \text{vec}(\mathbf{G}(\hat{\mathbf{E}}))$. According to the definition of the CSA, the operators $\mathbf{M}(\cdot)$ and $\mathbf{G}(\cdot)$ are rewritten as the multiplication of a series of suboperation matrices. Then, we deduce

$$\hat{\sigma} = \text{vec}(\mathbf{M}(\mathbf{S})) = \mathbf{M}\mathbf{s} = \hat{\mathbf{F}}_a^H \hat{\mathbf{H}}_3 \hat{\mathbf{F}}_r^H \hat{\mathbf{H}}_2 \hat{\mathbf{F}}_r \hat{\mathbf{H}}_1 \hat{\mathbf{F}}_a \mathbf{s} \quad (10)$$

$$\mathbf{s} = \text{vec}(\mathbf{G}(\hat{\mathbf{E}})) = \mathbf{G}\hat{\sigma} = \hat{\mathbf{F}}_a^H \hat{\mathbf{H}}_1 \hat{\mathbf{F}}_r^H \hat{\mathbf{H}}_2^* \hat{\mathbf{F}}_r \hat{\mathbf{H}}_3^* \hat{\mathbf{F}}_a \hat{\sigma} \quad (11)$$

$$\hat{\mathbf{H}}_x = \text{diag}(\text{vec}(\mathbf{H}_x)) \in \mathbb{C}^{MN \times MN}, x = 1, 2, 3 \quad (12)$$

$$\begin{cases} \hat{\mathbf{F}}_r = \mathbf{I}_a \otimes \mathbf{F}_r \in \mathbb{C}^{MN \times MN} \\ \hat{\mathbf{F}}_a = \mathbf{F}_a^T \otimes \mathbf{I}_r \in \mathbb{C}^{MN \times MN} \end{cases} \quad (13)$$

where $\hat{\mathbf{F}}_r^H = \mathbf{I}_a \otimes \mathbf{F}_r^H$, $\hat{\mathbf{F}}_a^H = (\mathbf{F}_a^H)^T \otimes \mathbf{I}_r$, $\mathbf{I}_r \in \mathbb{R}^{M \times M}$, and $\mathbf{I}_a \in \mathbb{R}^{N \times N}$ are the identity matrices, and $\hat{\mathbf{H}}_x$ is the diagonal matrix that satisfies $\hat{\mathbf{H}}_x^* = \hat{\mathbf{H}}_x^H$. Consequently, it can be concluded from Equations (10)–(13) that $\mathbf{G} = \mathbf{M}^H$ and $\mathbf{G}(\cdot) = \mathbf{M}(\cdot)^H$. $\square$

Theorem 1 proves that the inverse operation of $\mathbf{M}$, as an approximated observation, can be applied to design the 2D approximated observation model by replacing the measurement matrix Φ . Restricted by the form of operator $\mathbf{G}(\cdot)$, the downsampling observation

model $\mathbf{S}_d$ in Equation (7) is rewritten as $\mathbf{S}_d = \mathbf{\Psi}_r \mathbf{G}(\hat{\Xi}) \mathbf{\Psi}_a + \mathbf{N}_0$. Similar to the 1D exact observation model, the scattering coefficient matrix Ξ is reconstructed by transforming the downsampling observation model into an optimization problem, which is expressed as

$$\begin{aligned}\hat{\Xi} &= \arg \min_{\Xi} \left\{ \|\mathbf{S}_d - \mathbf{\Psi}_r \hat{\mathbf{S}} \mathbf{\Psi}_a\|_F^2 + \lambda \|\Xi\|_1 \right\} \\ &= \arg \min_{\Xi} \left\{ \|\mathbf{S}_d - \mathbf{\Psi}_r \mathbf{G}(\Xi) \mathbf{\Psi}_a\|_F^2 + \lambda \|\Xi\|_1 \right\}\end{aligned}\quad (14)$$

where $\|\cdot\|_F$ denotes the Frobenius norm, and $\hat{\mathbf{S}} = \mathbf{G}(\Xi)$ is the approximate echo data calculated by introducing the true scattered field Ξ into the CSA operator $\mathbf{G}(\cdot)$.

The above derivation method can be extended to yield more approximate observation-based CS algorithms from the inverse operation of other MF methods, such as the RDA and BPA. Note that the decoupled nature of MF methods enables their inverse operation by definition. This paper does not present other possible extensions due to space restrictions.

2.4. Sparse Representation of the SAR Imaging Model

The matrix Ξ is considered sparse if the dominant parts of the scattering coefficients are either zeros or very close to zeros. The scattered field of complex SAR scenes is generally nonsparse, making it challenging to obtain high-quality SAR images by directly adopting the optimization model in Equation (14). Assuming that the SAR scene is sparse after adopting traditional transform domain methods (such as the discrete wavelet transform (DWT) [33] or discrete cosine transform (DCT) [34]), the optimization model is rewritten as a sparse-representation-based reconstruction model $\hat{\Xi} = \arg \min_{\Xi} \left\{ \|\mathbf{S}_d - \mathbf{\Psi}_r \hat{\mathbf{S}} \mathbf{\Psi}_a\|_F^2 + \lambda \|\Xi_{SR}\|_1 \right\}$, where $\Xi_{SR} = \mathbf{Y}\Xi$ denotes the sparse representation of Ξ with respect to the transform matrix $\mathbf{Y}$, and the constraint term $\lambda \|\cdot\|_1$ enforces the sparsity of matrix Ξ_{SR} . Then, we further generalize the ℓ_1 norm-based regularizer $\|\cdot\|_1$ to a more general regularization function $\rho(\cdot)$, which represents the piecewise linear function (PLF) [35] or the constraint function based on the ℓ_q norm, where $q \in [0, 1]$. The ℓ_q norm-based regularizers with $q = 0$ and $q = 1$ correspond to the hard threshold functions and soft threshold functions, respectively.

Assuming that the transform matrix $\mathbf{Y}$ is an orthonormal sparse basis matrix, the scattering coefficient matrix is defined as $\Xi = \mathbf{Y}^+ \Xi_{SR}$, where $\mathbf{Y}\mathbf{Y}^+ = \mathbf{I}$ and $\mathbf{I}$ is the identity operator. In this case, the optimization problem of nonsparse scene reconstruction has been efficiently addressed due to the orthogonality of $\mathbf{Y}$. However, it remains nontrivial and laborious to obtain a high-quality SAR reconstructed image if the sparse basis is nonorthogonal or even nonlinear in a more complex imaging scene. Then, the sparse transform matrix is further generalized to a general nonlinear transform function $F(\cdot)$ to relax its reversibility and orthogonality requirements. We substitute $\rho[F(\Xi)]$ for $\|\mathbf{Y}\Xi\|_1$ to develop a generalized reconstruction model based on sparse representation.

$$\hat{\Xi} = \arg \min_{\Xi} \left\{ \|\mathbf{S}_d - \mathbf{\Psi}_r \hat{\mathbf{S}} \mathbf{\Psi}_a\|_F^2 + \lambda \rho[F(\Xi)] \right\}. \quad (15)$$

Many iterative optimization algorithms have been developed to solve this optimization problem, including the iterative threshold algorithm (ITA), ADMM, AMP, etc. Nevertheless, these algorithms universally demand many iterations and extensive calculations to achieve satisfactory reconstruction results. In this paper, the nonlinear transform function and other iterative parameters are set as undetermined variables in network learning to improve the imaging speed and reconstruction performance.

3. SAR Imaging Nets Based on Sparse Representation and Approximated Observation

In Section 3.1, first, we elaborate on the design of the CSA- and ISTA-based deep SAR imaging net architecture. Second, Sections 3.2 and 3.3 present SR-CSA-Net and its enhanced version, which apply to SAR imaging of complex scenes. Last, the network structure and training strategy analysis are demonstrated in Section 3.4.

3.1. CSA-Derived 2D SAR Imaging Net Architecture

As one of the ITAs, the ISTA with the ℓ_1 norm-based regularizer is well known for solving the regularization optimization problem and is widely utilized in sparse SAR imaging. In this subsection, the detailed construction process from the ISTA to the deep-unfolding-based SAR imaging net architecture is introduced step by step. Specifically, the ISTA solving the optimization problem in Equation (3) is divided into two stages—operator update and nonlinear transformation—which are expressed as

$$\mathbf{r}_l = \hat{\sigma}_{l-1} + \mu \cdot (\mathbf{\Psi}\mathbf{\Phi})^H (\mathbf{s}_d - \mathbf{\Psi}\mathbf{\Phi}\hat{\sigma}_{l-1}) \quad (16)$$

$$\hat{\sigma}_l = \text{soft}(\mathbf{r}_l; T) = \text{sign}(\mathbf{r}_l) \max[|\mathbf{r}_l| - T; 0] \quad (17)$$

where l denotes the iteration index, μ represents the step size, which affects the convergence of the ISTA, $\mathbf{r}_l$ is the operator that processes the residuals, $\text{soft}(\cdot; \cdot)$ denotes the nonlinear reconstruction function corresponding to the ℓ_1 norm-based regularizer ($q = 1$), λ and μ are merged into a threshold parameter $T = \lambda\mu$, and $\text{sign}(\cdot)$ denotes the sign function. Similarly, the ISTA can also effectively solve the 2D optimization problem in Equation (14) due to the linear characteristic of the CSA operator $\mathbf{G}(\cdot)$. The 2D scattering coefficient matrix $\hat{\Xi}$ is reconstructed by iterating between the operator update and the nonlinear transformation, which can be rewritten as

$$\mathbf{R}_l = \hat{\Xi}_{l-1} + \mu_l \cdot \mathbf{M} \left\{ \mathbf{\Psi}_r^T [\mathbf{S}_d - \mathbf{\Psi}_r \mathbf{G}(\hat{\Xi}_{l-1}) \mathbf{\Psi}_a] \mathbf{\Psi}_a^T \right\} \quad (18)$$

$$\hat{\Xi}_l = \text{soft}(\mathbf{R}_l; T) = \text{sign}(\mathbf{R}_l) \max[|\mathbf{R}_l| - T; 0]. \quad (19)$$

In CS-driven SAR imaging methods, the ISTA is frequently required to undertake numerous iterations to acquire acceptable reconstruction results, accompanied by the challenge of selecting the most appropriate hand-crafted parameters. To overcome the disadvantages of 1D exact observation models and online sparse imaging techniques, we developed a CSA-derived deep SAR imaging net by unfolding the above-mentioned 2D iterative algorithm. This net architecture for 2D SAR imaging consists of a fixed number of L layers, each corresponding to one ISTA iteration. The two stages represented by Equations (18) and (19) are mapped to the monolayer topology, which comprises two update modules: the linearity module $\mathbf{R}$ and the nonlinearity module $\mathbf{N}$. Figure 2 depicts the topology architecture of the deep-unfolding-based 2D SAR imaging net.

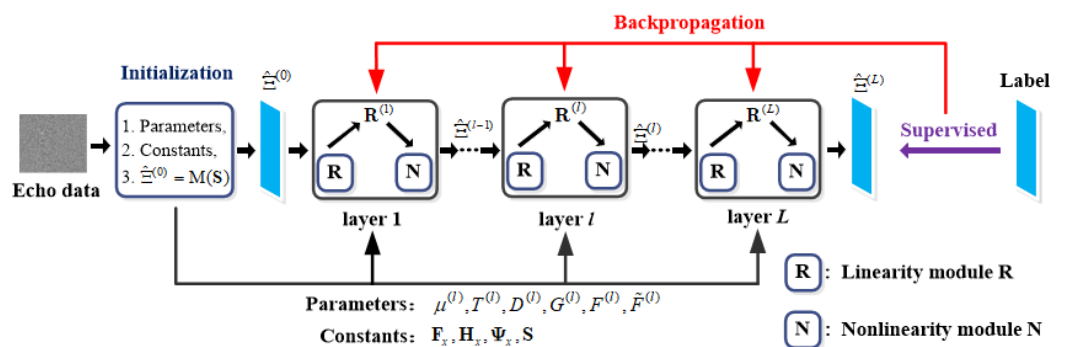


Figure 2. Topology architecture of 2D SAR imaging net.

Then, we describe the internal structure of network modules. In terms of Equation (18), the linearity module $\mathbf{R}$ is expressed as

$$\mathbf{R} : \mathbf{R}^{(l)} = \hat{\Xi}^{(l-1)} + \mu^{(l)} \cdot \mathbf{M} \left\{ \mathbf{\Psi}_r^T [\mathbf{S}_d - \mathbf{\Psi}_r \mathbf{G}(\hat{\Xi}^{(l-1)}) \mathbf{\Psi}_a] \mathbf{\Psi}_a^T \right\}. \quad (20)$$

where $\mathbf{R}^{(l)}$ denotes the 2D linear reconstruction result, and $\mu^{(l)}$ and $T^{(l)}$ represent the learnable parameters in the l th layer, while they are manually tuned in the ISTA. Parameter learnability is conducive to resolving the challenge of adjusting a priori. Since the radar signal is complex-valued data, we address the matrices in Equation (20) by separating them into real and imaginary parts, i.e., $\mathbf{R}^{(l)} = \mathbf{Re}(\mathbf{R}^{(l)}) + j\mathbf{Im}(\mathbf{R}^{(l)})$, where $\mathbf{Re}(\cdot)$ denotes the real part of the matrix, $\mathbf{Im}(\cdot)$ denotes the imaginary part of the matrix, and j denotes the imaginary unit. Hence, the operators $G(\cdot)$ and $M(\cdot)$ in module $\mathbf{R}$ are designed as a dual-path structure, with which the real and imaginary parts of the radar signal are separately processed.

We use $G(\cdot)$ as an example to introduce the dual-path structure in detail. The calculation process of $G(\cdot)$ is divided into three steps to simplify the calculation expression, and the real and imaginary parts of each step are derived according to the complex-valued operation rules. The first step $\mathbf{F}_r[(\hat{\mathbf{z}}^{(l-1)}\mathbf{F}_a) \circ \mathbf{H}_3^*]$ with respect to $\hat{\mathbf{z}}^{(l-1)}$, denoted as a suboperator $\mathbf{F}_r[\cdot]$, is given in Equation (21), where $\text{Con}(\mathbf{A}, \mathbf{B}) = [\mathbf{A}^T \mathbf{B}^T]^T$ represents the consolidation operator, with which the matrices $\mathbf{A}$ and $\mathbf{B}$ are consolidated to form an array $[\mathbf{A}^T \mathbf{B}^T]^T$, and $d\text{Con}_U([\mathbf{A}^T \mathbf{B}^T]^T) = \mathbf{A}$ and $d\text{Con}_L([\mathbf{A}^T \mathbf{B}^T]^T) = \mathbf{B}$ represent the inverse operation of $\text{Con}(\cdot, \cdot)$, with which the upper part and lower part, respectively, of the array are extracted.

The second step $\mathbf{F}_r^H(\cdot) = \mathbf{F}_r^H(\mathbf{F}_r[\cdot] \circ \mathbf{H}_2^*)$ based on suboperator $\mathbf{F}_r[\cdot]$ is expressed as Equation (22), and the third step $\{\cdot\}^H = \{\mathbf{F}_r^H(\cdot) \circ \mathbf{H}_1^*\}^H$, namely, $G(\cdot)$, is given in Equation (23). Since $M(\cdot)$ is the inverse operation of $G(\cdot)$, $M(\cdot)$ has the same structure as $G(\cdot)$ and contains different phase functions. The structure of $M(\cdot)$ can be developed by swapping the input and output of $G(\cdot)$, and the topology architecture of module $\mathbf{R}$ in the l th layer is illustrated in Figure 3 based on the above analysis. Note that all the matrix operations related to the FFT and downsampling in Figure 3 are left multiplications.

$$\begin{bmatrix} \mathbf{Re}(\mathbf{F}_r[\cdot]) \\ \mathbf{Im}(\mathbf{F}_r[\cdot]) \end{bmatrix} = \begin{bmatrix} \mathbf{Re}(\mathbf{F}_r) & -\mathbf{Im}(\mathbf{F}_r) \\ \mathbf{Im}(\mathbf{F}_r) & \mathbf{Re}(\mathbf{F}_r) \end{bmatrix} \cdot \text{Con} \left(\begin{bmatrix} \mathbf{Re}(\hat{\mathbf{z}}^{(l-1)}\mathbf{F}_a) \circ \mathbf{Re}(\mathbf{H}_3^*) - \mathbf{Im}(\hat{\mathbf{z}}^{(l-1)}\mathbf{F}_a) \circ \mathbf{Im}(\mathbf{H}_3^*), \\ \mathbf{Im}(\hat{\mathbf{z}}^{(l-1)}\mathbf{F}_a) \circ \mathbf{Re}(\mathbf{H}_3^*) + \mathbf{Re}(\hat{\mathbf{z}}^{(l-1)}\mathbf{F}_a) \circ \mathbf{Im}(\mathbf{H}_3^*) \end{bmatrix} \right) \\ = \begin{bmatrix} \mathbf{Re}(\mathbf{F}_r) & -\mathbf{Im}(\mathbf{F}_r) \\ \mathbf{Im}(\mathbf{F}_r) & \mathbf{Re}(\mathbf{F}_r) \end{bmatrix} \cdot \text{Con} \left(\begin{array}{l} d\text{Con}_U \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_a) & \mathbf{Im}(\mathbf{F}_a) \\ -\mathbf{Im}(\mathbf{F}_a) & \mathbf{Re}(\mathbf{F}_a) \end{bmatrix} \begin{bmatrix} \mathbf{Re}(\hat{\mathbf{z}}^{(l-1)})^T \\ \mathbf{Im}(\hat{\mathbf{z}}^{(l-1)})^T \end{bmatrix} \right)^T \circ \mathbf{Re}(\mathbf{H}_3^*) \\ -d\text{Con}_L \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_a) & \mathbf{Im}(\mathbf{F}_a) \\ -\mathbf{Im}(\mathbf{F}_a) & \mathbf{Re}(\mathbf{F}_a) \end{bmatrix} \begin{bmatrix} \mathbf{Re}(\hat{\mathbf{z}}^{(l-1)})^T \\ \mathbf{Im}(\hat{\mathbf{z}}^{(l-1)})^T \end{bmatrix} \right)^T \circ \mathbf{Im}(\mathbf{H}_3^*), \\ d\text{Con}_L \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_a) & \mathbf{Im}(\mathbf{F}_a) \\ -\mathbf{Im}(\mathbf{F}_a) & \mathbf{Re}(\mathbf{F}_a) \end{bmatrix} \begin{bmatrix} \mathbf{Re}(\hat{\mathbf{z}}^{(l-1)})^T \\ \mathbf{Im}(\hat{\mathbf{z}}^{(l-1)})^T \end{bmatrix} \right)^T \circ \mathbf{Re}(\mathbf{H}_3^*) \\ +d\text{Con}_U \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_a) & \mathbf{Im}(\mathbf{F}_a) \\ -\mathbf{Im}(\mathbf{F}_a) & \mathbf{Re}(\mathbf{F}_a) \end{bmatrix} \begin{bmatrix} \mathbf{Re}(\hat{\mathbf{z}}^{(l-1)})^T \\ \mathbf{Im}(\hat{\mathbf{z}}^{(l-1)})^T \end{bmatrix} \right)^T \circ \mathbf{Im}(\mathbf{H}_3^*) \end{array} \right) \quad (21)$$

$$\begin{bmatrix} \mathbf{Re}(\mathbf{F}_r^H(\cdot)) \\ \mathbf{Im}(\mathbf{F}_r^H(\cdot)) \end{bmatrix} = \begin{bmatrix} \mathbf{Re}(\mathbf{F}_r^H) & -\mathbf{Im}(\mathbf{F}_r^H) \\ \mathbf{Im}(\mathbf{F}_r^H) & \mathbf{Re}(\mathbf{F}_r^H) \end{bmatrix} \cdot \text{Con} \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_r[\cdot]) \circ \mathbf{Re}(\mathbf{H}_2^*) - \mathbf{Im}(\mathbf{F}_r[\cdot]) \circ \mathbf{Im}(\mathbf{H}_2^*), \\ \mathbf{Im}(\mathbf{F}_r[\cdot]) \circ \mathbf{Re}(\mathbf{H}_2^*) + \mathbf{Re}(\mathbf{F}_r[\cdot]) \circ \mathbf{Im}(\mathbf{H}_2^*) \end{bmatrix} \right) \\ = \begin{bmatrix} \mathbf{Re}(\mathbf{F}_r^H) & -\mathbf{Im}(\mathbf{F}_r^H) \\ \mathbf{Im}(\mathbf{F}_r^H) & \mathbf{Re}(\mathbf{F}_r^H) \end{bmatrix} \cdot \text{Con} \left(\begin{array}{l} d\text{Con}_U \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_r[\cdot]) \\ \mathbf{Im}(\mathbf{F}_r[\cdot]) \end{bmatrix} \right) \circ \mathbf{Re}(\mathbf{H}_2^*) - d\text{Con}_L \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_r[\cdot]) \\ \mathbf{Im}(\mathbf{F}_r[\cdot]) \end{bmatrix} \right) \circ \mathbf{Im}(\mathbf{H}_2^*), \\ d\text{Con}_L \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_r[\cdot]) \\ \mathbf{Im}(\mathbf{F}_r[\cdot]) \end{bmatrix} \right) \circ \mathbf{Re}(\mathbf{H}_2^*) + d\text{Con}_U \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_r[\cdot]) \\ \mathbf{Im}(\mathbf{F}_r[\cdot]) \end{bmatrix} \right) \circ \mathbf{Im}(\mathbf{H}_2^*) \end{array} \right) \quad (22)$$

$$\begin{bmatrix} \mathbf{Re}(G(\cdot))^T \\ \mathbf{Im}(G(\cdot))^T \end{bmatrix} = \begin{bmatrix} \mathbf{Re}(\{\cdot\}^H) \\ \mathbf{Im}(\{\cdot\}^H) \end{bmatrix}^T \\ = \begin{bmatrix} \mathbf{Re}(\mathbf{F}_a^H) & \mathbf{Im}(\mathbf{F}_a^H) \\ -\mathbf{Im}(\mathbf{F}_a^H) & \mathbf{Re}(\mathbf{F}_a^H) \end{bmatrix}^T \cdot \text{Con} \left(\begin{array}{l} \{\mathbf{Re}(\mathbf{F}_r^H(\cdot)) \circ \mathbf{Re}(\mathbf{H}_1^*) - \mathbf{Im}(\mathbf{F}_r^H(\cdot)) \circ \mathbf{Im}(\mathbf{H}_1^*)\}^T, \\ \{\mathbf{Im}(\mathbf{F}_r^H(\cdot)) \circ \mathbf{Re}(\mathbf{H}_1^*) + \mathbf{Re}(\mathbf{F}_r^H(\cdot)) \circ \mathbf{Im}(\mathbf{H}_1^*)\}^T \end{array} \right) \\ = \begin{bmatrix} \mathbf{Re}(\mathbf{F}_a^H) & \mathbf{Im}(\mathbf{F}_a^H) \\ -\mathbf{Im}(\mathbf{F}_a^H) & \mathbf{Re}(\mathbf{F}_a^H) \end{bmatrix}^T \cdot \text{Con} \left(\begin{array}{l} \left\{ d\text{Con}_U \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_r^H(\cdot)) \\ \mathbf{Im}(\mathbf{F}_r^H(\cdot)) \end{bmatrix} \right) \circ \mathbf{Re}(\mathbf{H}_1^*) - d\text{Con}_L \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_r^H(\cdot)) \\ \mathbf{Im}(\mathbf{F}_r^H(\cdot)) \end{bmatrix} \right) \circ \mathbf{Im}(\mathbf{H}_1^*) \right\}^T, \\ \left\{ d\text{Con}_L \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_r^H(\cdot)) \\ \mathbf{Im}(\mathbf{F}_r^H(\cdot)) \end{bmatrix} \right) \circ \mathbf{Re}(\mathbf{H}_1^*) + d\text{Con}_U \left(\begin{bmatrix} \mathbf{Re}(\mathbf{F}_r^H(\cdot)) \\ \mathbf{Im}(\mathbf{F}_r^H(\cdot)) \end{bmatrix} \right) \circ \mathbf{Im}(\mathbf{H}_1^*) \right\}^T \end{array} \right) \quad (23)$$

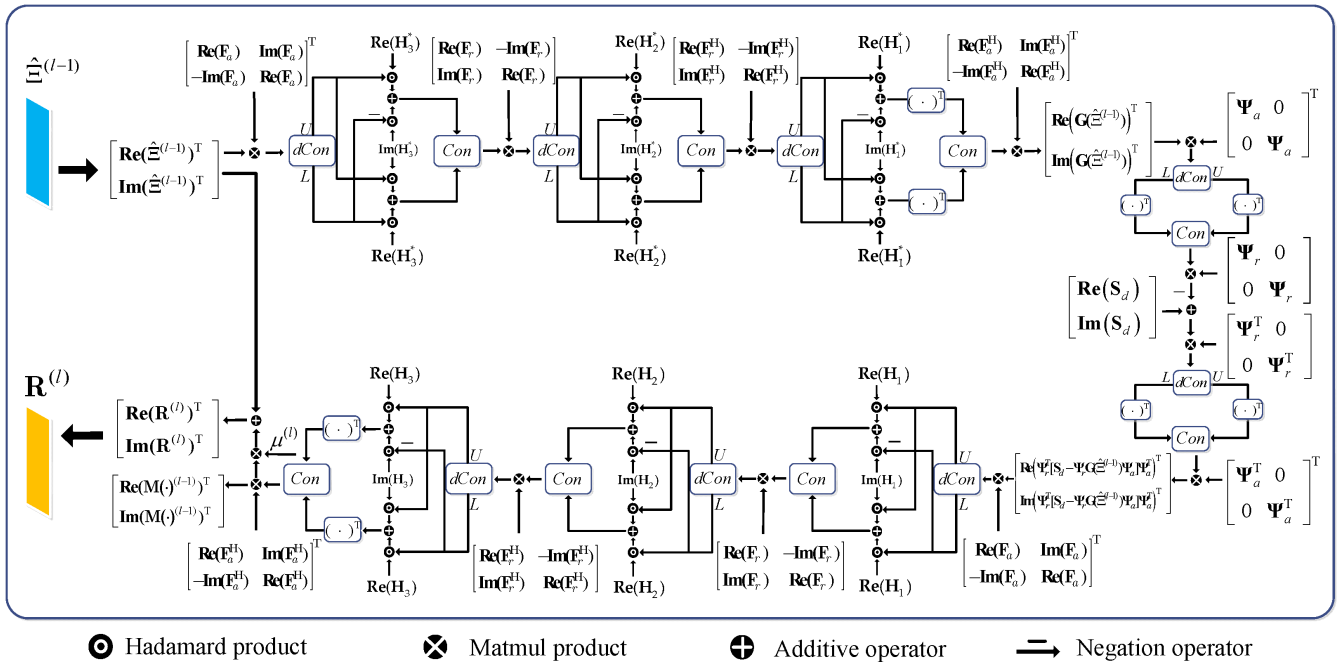


Figure 3. Topology architecture of module R.

Furthermore, if the imaging scene is sparse in the spatial domain, module N in Equation (19) can be combined with module R to establish a CSA-based 2D SAR imaging net (CSA-Net) applied to solve Equation (14). Conversely, if the imaging scene is nonsparse or weakly sparse, module N should be redesigned to sparsely represent the corresponding scene. Aimed at the difficulty of SAR imaging in complex scenes, a novel 2D SAR imaging strategy that can introduce sparse representation and transform sparsity into the network architecture is introduced in Section 3.3.

3.2. Deep SAR Imaging Net Based on Sparse Representation: SR-CSA-Net

In this subsection, a CNN-based module N with sparse representation is designed for solving the complex-scene, SAR imaging problem in Equation (15); then, it is incorporated with the CSA-based module R to construct a dual-path deep SAR imaging net that is referred to as SR-CSA-Net.

If we use traditional transform-domain methods to achieve sparse representation, the output of layer l can be calculated by the nonlinear function $\text{soft}(\cdot; \cdot)$ with the input of the linear reconstruction result $\mathbf{R}^{(l)}$ and threshold parameter $T^{(l)}$, which are expressed as follows:

$$\begin{cases} \hat{\Xi}_{SR}^{(l)} = \text{soft}(\mathbf{Y}\mathbf{R}^{(l)}; T^{(l)}) \\ \hat{\Xi}^{(l)} = \mathbf{Y}^+ \hat{\Xi}_{SR}^{(l)} \end{cases} \quad (24)$$

where the transform matrix $\mathbf{Y}$ and its inverse form $\mathbf{Y}^+$ are substitutable and nonoptimal. To sparsify the scattering coefficient matrix $\hat{\Xi}$ and optimize the reconstruction performance, we develop a CNN-based nonlinear transform function $F(\cdot)$ that can be utilized to learn an optimal sparse basis. Then, this well-designed transform function is embodied as a CNN block and adopted in SR-CSA-Net instead of using the handcrafted transform matrix $\mathbf{Y}$.

Considering the powerful fitting ability of a CNN, the convolution operators $C_i(\cdot)$, activation function $\text{ReLU}(\cdot)$, and batch normalization (BN) operator $B(\cdot)$ are combined to design the nonlinear transform function $F(\cdot)$, in which the two convolution operators $C_1(\cdot)$ and $C_2(\cdot)$ are successively separated by $B(\cdot)$ and $\text{ReLU}(\cdot)$. The BN operator preceding the

activation function effectively prevents problems such as vanishing gradient, overfitting, and slow convergence [36], which are expressed as

$$B(\mathbf{X}) = \frac{\mathbf{X} - E[\mathbf{X}]}{\sqrt{\text{Var}[\mathbf{X}] + \epsilon}} \cdot \gamma + \beta \quad (25)$$

where $E[\cdot]$ denotes the mean function, $\text{Var}[\cdot]$ represents the variance function, ϵ is a very small number that prevents the instability of numerical calculation, and γ and β are learnable affine parameters. On this basis, the linear reconstruction result $\mathbf{R}^{(l)}$ in the l th layer can be sparsely represented by using the well-designed transform function $F(\cdot)$, which is formulated as

$$F(\mathbf{R}^{(l)}) = C_2 \left\langle \text{ReLU} \left\{ B[C_1(\mathbf{R}^{(l)})] \right\} \right\rangle \quad (26)$$

where the first convolution operator $C_1(\cdot)$ corresponds to a set of N_f filters with a size of $\omega_f \times \omega_f \times 1$, whose purpose is to perform a channelwise conversion from one channel input to N_f channel output, the second convolution operator corresponds to another set of N_f filters with a size of $\omega_f \times \omega_f \times N_f$, and the rectified linear unit (ReLU) function is denoted as $\text{ReLU}(\mathbf{X}) = \max(\mathbf{X}, 0)$. In summary, the CNN block $F(\cdot)$ is important in module **N** and has rich sparse representation capability due to its nonlinearity and learnability.

It has been indicated in [37] that $\|\hat{\mathbf{E}}^{(l)} - F(\mathbf{R}^{(l)})\|_F^2$ and $\|\hat{\mathbf{E}}^{(l)} - \mathbf{R}^{(l)}\|_F^2$ satisfy the linear relationship under a reasonable assumption. Thus, $\hat{\mathbf{E}}_{SR}^{(l)}$, the sparse representation of $\hat{\mathbf{E}}^{(l)}$, is derived by $\hat{\mathbf{E}}_{SR}^{(l)} = F(\hat{\mathbf{E}}^{(l)}) = \text{soft}[F(\mathbf{R}^{(l)}); T^{(l)}]$. Furthermore, motivated by the structure symmetric constraint, a mirror-symmetrical structure with different weights, denoted by $\tilde{F}(\cdot)$ such that $\tilde{F}(\cdot) \times F(\cdot) = I$, is designed to extract the unknown scattering coefficient matrix from the obtained result of sparse representation. Therefore, the following closed-form expression efficiently computes the unknown scattering coefficient matrix $\hat{\mathbf{E}}^{(l)}$ in layer l :

$$\begin{aligned} \hat{\mathbf{E}}^{(l)} &= \tilde{F}(\hat{\mathbf{E}}_{SR}^{(l)}) \\ &= \tilde{F} \left\{ \text{soft} [F(\mathbf{R}^{(l)}); T^{(l)}] \right\} \end{aligned} \quad (27)$$

where the symmetric CNN block $\tilde{F}(\hat{\mathbf{E}}_{SR}^{(l)})$ is modeled as $\tilde{F}(\hat{\mathbf{E}}_{SR}^{(l)}) = \tilde{C}_2 \left\langle \text{ReLU} \left\{ \tilde{B}[\tilde{C}_1(\hat{\mathbf{E}}_{SR}^{(l)})] \right\} \right\rangle$. The nonlinearity module **N** is also divided into two paths to connect the real and imaginary parts output by the linearity module **R**. Moreover, due to the learnability of $\tilde{F}(\cdot)$ and $F(\cdot)$, each layer of SR-CSA-Net has its own CNN block, and thus, module **N** in the l th layer is redesigned as

$$\mathbf{N} : \begin{cases} \mathbf{Re}(\hat{\mathbf{E}}^{(l)}) = \tilde{F}^{(l)} \left\{ \text{soft} [F^{(l)}(\mathbf{Re}(\mathbf{R}^{(l)})); T^{(l)}] \right\} \\ \mathbf{Im}(\hat{\mathbf{E}}^{(l)}) = \tilde{F}^{(l)} \left\{ \text{soft} [F^{(l)}(\mathbf{Im}(\mathbf{R}^{(l)})); T^{(l)}] \right\} \end{cases} \quad (28)$$

The nonlinear reconstruction function $\text{soft}(\cdot; \cdot)$ is determined by Equation (19). For the nonzero elements of complex x , the sign function is calculated by $\text{sign}(x) = x./|x|$. Therefore, when processing complex-valued signals, the function $\text{soft}(\cdot; \cdot)$ in module **N** is redefined as a nonlinear function including the complex normalization operation. Selecting the real and imaginary parts of $\mathbf{R}^{(l)}$ as examples, the function $\text{soft}(\cdot; \cdot)$ is expressed as

$$\begin{aligned} \text{soft}[\mathbf{Re}(\mathbf{R}^{(l)}); T^{(l)}] &= \text{sign}(\mathbf{Re}(\mathbf{R}^{(l)})) \max[(|\mathbf{R}^{(l)}| - T^{(l)}); 0] \\ &= (\mathbf{Re}(\mathbf{R}^{(l)})./|\mathbf{R}^{(l)}|) \max[(|\mathbf{R}^{(l)}| - T^{(l)}); 0] \end{aligned} \quad (29)$$

$$\begin{aligned} \text{soft}[\mathbf{Im}(\mathbf{R}^{(l)}); T^{(l)}] &= \text{sign}(\mathbf{Im}(\mathbf{R}^{(l)})) \max[(|\mathbf{R}^{(l)}| - T^{(l)}); 0] \\ &= (\mathbf{Im}(\mathbf{R}^{(l)})./|\mathbf{R}^{(l)}|) \max[(|\mathbf{R}^{(l)}| - T^{(l)}); 0] \end{aligned} \quad (30)$$

where $\mathbf{Re}(\mathbf{R}^{(l)})./|\mathbf{R}^{(l)}|$ and $\mathbf{Im}(\mathbf{R}^{(l)})./|\mathbf{R}^{(l)}|$ represent the normalization operations of the real part and imaginary part, respectively, and $|\mathbf{R}^{(l)}| = \text{abs}(\mathbf{Re}(\mathbf{R}^{(l)}), \mathbf{Im}(\mathbf{R}^{(l)}))$ is the

elementwise absolute operation. Figure 4a illustrates the dual-path topology architecture of module **N** in the l th layer, and the monolayer topology of the proposed SR-CSA-Net is composed of modules **R** and **N**.

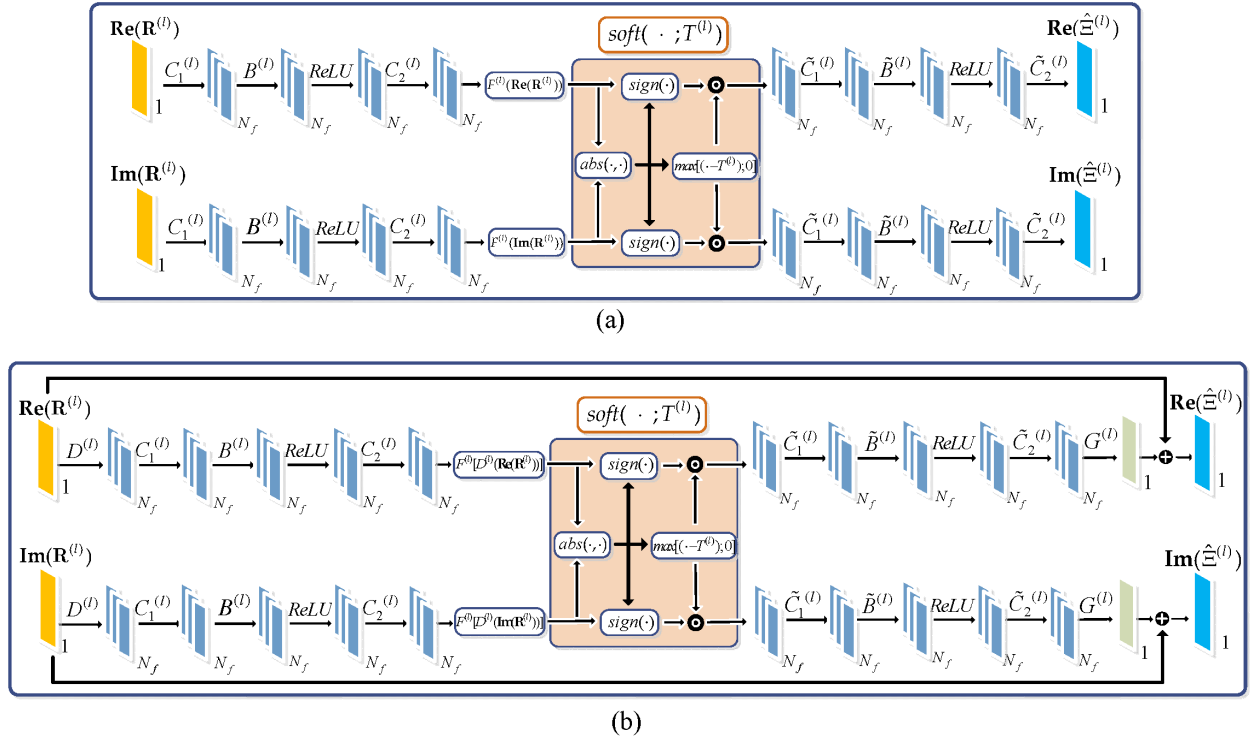


Figure 4. Topology architecture of the nonlinearity module. (a) Module **N**; (b) module **N_{plus}**.

3.3. Enhanced Version: SR-CSA-Net-Plus

To accelerate the convergence speed and expect further performance improvement from the sparse reconstruction, we designed an enhanced network named SR-CSA-Net-plus by improving module **N** from SR-CSA-Net. Suppose that $\hat{\mathbf{z}}^{(l)} = \mathbf{R}^{(l)} + \mathbf{W}^{(l)} + \mathbf{N}_0^{(l)}$, where $\mathbf{W}^{(l)}$ denotes the high-frequency component missing in the linear reconstruction result $\mathbf{R}^{(l)}$, which can be recovered by a linear operator $\Pi(\cdot)$ from the closed-form solution in Equation (28), i.e., $\mathbf{W}^{(l)} = \Pi(\hat{\mathbf{z}}^{(l)}) = \Pi\langle \tilde{F}^{(l)}\{\text{soft}[F^{(l)}(\mathbf{R}^{(l)}); T^{(l)}]\} \rangle$.

Specifically, the linear operator $\Pi(\cdot)$ consists of two convolution operators, represented by $\Pi(\cdot) = G(\cdot) \times D(\cdot)$, where $D(\cdot)$ corresponds to a set of N_f filters with a size of $\omega_f \times \omega_f \times 1$, and $G(\cdot)$ corresponds to one filter with a size of $\omega_f \times \omega_f \times N_f$. In this case, we redefine the size of convolution operators in both CNN blocks to address the dimensionality mismatch, and both $C_i(\cdot)$ and $\tilde{C}_i(\cdot)$ are defined as one set of N_f filters with a size of $\omega_f \times \omega_f \times N_f$. On this basis, we introduce skip connections to develop an improved module **N**, named module **N_{plus}**, which is given by Equation (31), and its topology architecture is illustrated in Figure 4b.

The nonlinear reconstruction result of SR-CSA-Net-plus is updated by module **N_{plus}**. Similar to module **N** in Equation (28), $\Pi(\cdot)$, $F(\cdot)$, and $\tilde{F}(\cdot)$ are also learnable in module **N_{plus}**, and they are not constrained to be the same on each layer, namely, they are layer-varied.

$$\mathbf{N}_{plus} : \begin{cases} \mathbf{Re}(\hat{\mathbf{z}}^{(l)}) = \mathbf{Re}(\mathbf{R}^{(l)}) + \Pi^{(l)}\langle \tilde{F}^{(l)}\{\text{soft}[F^{(l)}(\mathbf{Re}(\mathbf{R}^{(l)})); T^{(l)}]\} \rangle \\ \quad = \mathbf{Re}(\mathbf{R}^{(l)}) + G^{(l)}\langle \tilde{F}^{(l)}\{\text{soft}[F^{(l)}(D^{(l)}(\mathbf{Re}(\mathbf{R}^{(l)}))]; T^{(l)}]\} \rangle \\ \mathbf{Im}(\hat{\mathbf{z}}^{(l)}) = \mathbf{Im}(\mathbf{R}^{(l)}) + \Pi^{(l)}\langle \tilde{F}^{(l)}\{\text{soft}[F^{(l)}(\mathbf{Im}(\mathbf{R}^{(l)})); T^{(l)}]\} \rangle \\ \quad = \mathbf{Im}(\mathbf{R}^{(l)}) + G^{(l)}\langle \tilde{F}^{(l)}\{\text{soft}[F^{(l)}(D^{(l)}(\mathbf{Im}(\mathbf{R}^{(l)}))]; T^{(l)}]\} \rangle \end{cases} \quad (31)$$

Note that in SR-CSA-Net-plus, $F(\cdot) \times D(\cdot)$ and $G(\cdot) \times \tilde{F}(\cdot)$ are adopted to replace $F(\cdot)$ and $\tilde{F}(\cdot)$, respectively, in SR-CSA-Net, which is conducive to further exploring the sparsity of SAR images. The monolayer topology of the proposed SR-CSA-Net-plus comprises modules $\mathbf{R}$ and $\mathbf{N}_{plus}$.

3.4. Network Analysis

Structural analysis: SR-CSA-Net and SR-CSA-Net-plus mainly consist of a linearity module $\mathbf{R}$ and a nonlinearity module $\mathbf{N}$ or $\mathbf{N}_{plus}$. The nonlinearity module of each layer is composed of convolution operators, BN operators, and nonlinear activation functions. Therefore, the framework of our proposed imaging nets is a mixture of linear reconstruction, convolution, batch normalization, and nonlinear activation. Compared with conventional CNN-based or deep-unfolding-based methods, our proposed imaging nets are impressive in network structure and operation mechanism. These two nets embed the approximated observation model into a linear reconstruction architecture and convert the sparse transform function into a CNN block. Moreover, the distinct superiority of SR-CSA-Net-plus lies in its more prominent sparse representation capability and the addition of skip connections. The skip connection structure facilitates network convergence and has been widely applied to ResNet [38] and DenseNet [39].

Learnable parameters: In our proposed imaging nets, the parameters of different layers are set to be nonshared. Γ_{net} is the parameter set of SR-CSA-Net, which has three types of parameters: the step size $\mu^{(l)}$, the soft threshold $T^{(l)}$, and the CNN blocks $F^{(l)}$ and $\tilde{F}^{(l)}$. Hence, Γ_{net} is expressed as $\Gamma_{net} = \{\mu^{(l)}, T^{(l)}, F^{(l)}, \tilde{F}^{(l)}\}_{l=1}^L$, where L is the total number of network layers, and $F^{(l)}$ and $\tilde{F}^{(l)}$ include parameters γ, β, N_f , and ω_f . Similar to SR-CSA-Net, each layer of SR-CSA-Net-plus also has its trainable network parameters, and its parameter set is denoted by $\Gamma_{net}^+ = \{\mu^{(l)}, T^{(l)}, D^{(l)}, G^{(l)}, F^{(l)}, \tilde{F}^{(l)}\}_{l=1}^L$. In this paper, we adopt supervised training to learn all parameters and use the minibatch gradient descent (MBGD) algorithm [36] in combination with the backpropagation algorithm [40] to update the gradient.

Complexity analysis: All of the above parameters in Γ_{net} and Γ_{net}^+ are real-valued. The dimensionality of $\mu^{(l)}$ and $T^{(l)}$ is 1, and the dimensionality of $D^{(l)}, G^{(l)}, F^{(l)}$, and $\tilde{F}^{(l)}$ is calculated by N_f and ω_f . For SR-CSA-Net, the number of network parameters in $F^{(l)}$ and $\tilde{F}^{(l)}$ is the same and equal to $1 \times \omega_f \times \omega_f \times N_f + N_f \times (2 + \omega_f \times \omega_f) \times N_f$. Then, the total number of network parameters in SR-CSA-Net is denoted as $O_{\Gamma_{net}} = L \times \{2 + 2 \times [N_f \times \omega_f^2 + N_f^2 \times (2 + \omega_f^2)]\}$. Similarly, for SR-CSA-Net-plus, the numbers of parameters in $D^{(l)}, G^{(l)}, F^{(l)}$, and $\tilde{F}^{(l)}$ are $N_f \times \omega_f \times \omega_f \times 1, 1 \times \omega_f \times \omega_f \times N_f, N_f \times (2 + \omega_f \times \omega_f + \omega_f \times \omega_f) \times N_f$, and $N_f \times (2 + \omega_f \times \omega_f + \omega_f \times \omega_f) \times N_f$, respectively. Therefore, the total number of network parameters in SR-CSA-Net-plus is $O_{\Gamma_{net}^+} = L \times \{2 + 2 \times [N_f \times \omega_f^2 + N_f^2 \times (2 + 2 \times \omega_f^2)]\}$.

For recent and state-of-the-art SAR imaging nets, namely, ISTA-Net [24], RMIST-Net [25], TISTA [41], C-TISTA [42], LAMP [43], AF-AMPnet [30], and RDA-Net [31], the number of parameters for each of these imaging nets is $L(PQ^2 + MN \times PQ + 1)$, $2L$, $L + 2$, $4L$, $L(MN \times PQ + 2)$, $2L + 2$, and $L(2MN + M^2 + N^2 + 2)$, respectively. Since the values of N_f and ω_f are generally much smaller than those of MN and PQ , the number of parameters in our proposed two imaging nets is acceptable and moderate. According to the above-mentioned analysis, SR-CSA-Net and SR-CSA-Net-plus achieve a tradeoff between reconstruction quality and computational complexity compared with the existing SAR imaging nets.

Loss function: The training procedure of the proposed imaging nets can be viewed as updating the learnable parameter set Γ_{net} or Γ_{net}^+ by minimizing the loss function. Instead of using a general loss function such as the averaged normalized mean square error (NMSE) [44], we adopt a cleverly designed loss function in this paper, in which both the reconstruction error and symmetry constraint of CNN blocks are considered. More specifically, the structure symmetry constraint $\tilde{F}(\cdot) \times F(\cdot) = I$ is satisfied in the

well-designed loss function while reducing the discrepancy between the label and the final reconstruction result obtained in layer L .

The total loss functions for SR-CSA-Net and SR-CSA-Net-plus are denoted as $\mathcal{L}_{\text{total}}$ and $\mathcal{L}_{\text{total}}^+$, respectively, which are defined as follows:

$$\mathcal{L}_{\text{total}}(\mathbf{I}_{\text{net}}) = \alpha_1 \mathcal{L}_1 + \alpha_2 \mathcal{L}_2 \quad (32a)$$

$$\mathcal{L}_{\text{total}}^+(\mathbf{I}_{\text{net}}^+) = \alpha_1 \mathcal{L}_1 + \alpha_3 \mathcal{L}_3 \quad (32b)$$

$$\mathcal{L}_1 = \frac{1}{2J} \sum_{j=1}^J \|\hat{\mathbf{e}}_j^{(L)} - \mathbf{D}_j\|_F^2 \quad (32c)$$

$$\begin{aligned} \mathcal{L}_2 = & \frac{1}{2J} \sum_{j=1}^J \sum_{l=1}^L \left\| \tilde{F}^{(l)} \left[F^{(l)} (\mathbf{Re}(\mathbf{R}_j^{(l)})) \right] - \mathbf{Re}(\mathbf{R}_j^{(l)}) \right\|_F^2 \\ & + \frac{1}{2J} \sum_{j=1}^J \sum_{l=1}^L \left\| \tilde{F}^{(l)} \left[F^{(l)} (\mathbf{Im}(\mathbf{R}_j^{(l)})) \right] - \mathbf{Im}(\mathbf{R}_j^{(l)}) \right\|_F^2 \end{aligned} \quad (32d)$$

$$\begin{aligned} \mathcal{L}_3 = & \frac{1}{2J} \sum_{j=1}^J \sum_{l=1}^L \left\| \tilde{F}^{(l)} \left\{ F^{(l)} \left[D^{(l)} (\mathbf{Re}(\mathbf{R}_j^{(l)})) \right] \right\} - D^{(l)} (\mathbf{Re}(\mathbf{R}_j^{(l)})) \right\|_F^2 \\ & + \frac{1}{2J} \sum_{j=1}^J \sum_{l=1}^L \left\| \tilde{F}^{(l)} \left\{ F^{(l)} \left[D^{(l)} (\mathbf{Im}(\mathbf{R}_j^{(l)})) \right] \right\} - D^{(l)} (\mathbf{Im}(\mathbf{R}_j^{(l)})) \right\|_F^2 \end{aligned} \quad (32e)$$

where $\mathcal{L}_1$ is the average Euclidean distance loss function, $\mathcal{L}_2$ and $\mathcal{L}_3$ are the symmetry constrained loss functions, J is the total number of training samples, $\hat{\mathbf{e}}_j^{(L)}$ represents the network output corresponding to the j th training sample, and $\mathbf{D}_j$ indicates the label corresponding to the j th training sample. In addition, α_1 , α_2 , and α_3 are adjustable parameters by which the compromise between $\mathcal{L}_1$ and $\mathcal{L}_2$ or between $\mathcal{L}_1$ and $\mathcal{L}_3$ can be controlled. In this paper, they were set to 1, 0.1, and 0.1, respectively.

4. Experimental Results and Analyses

In this section, the superiority of the proposed SR-CSA-Net and SR-CSA-Net-plus is demonstrated by simulated data and measured data from the RADARSAT-1 satellite. Table 1 lists the main parameters of the simulation and RADARSAT-1. These system parameters determine the phase functions $\mathbf{H}_x$ and FFT/IFFT matrices $\mathbf{F}_x$.

Table 1. Main parameters of the simulation and RADARSAT-1 satellite.

Parameters	Simulation	RADARSAT-1
Range FM rate	62.50 MHz/ μ s	0.72 MHz/ μ s
Azimuth FM rate	66.67 Hz/s	1733 Hz/s
Center frequency	10 GHz	5.3 GHz
Pulse duration	1.2 μ s	41.74 μ s
Pulse repetition frequency	100 Hz	1257 Hz
Effective radar velocity	100 m/s	7062 m/s

In the simulated experiments, we adopted the nonspare images in [37] as labels $\mathbf{D}_j|_{j=1}^J$ with a size of 256×256 . A sufficient number of simulated echoes $\mathbf{S}_j|_{j=1}^J$ was generated using labels with known geometry, 8000 of which were employed for training and 1000 were utilized for testing. With this strategy, the training set contained 8000 pairs of $\{\mathbf{D}_j, \mathbf{S}_j\}_{j=1}^{J=8000}$ with 20 dB AWGN added to each $\mathbf{S}_j$. Some nonspare labels in the training set are shown in Figure 5. In the measured experiments, a portion of the original RADARSAT-1 data containing nonspare scenes was applied to further investigate the proposed imaging strategy. The training set was formulated by utilizing 1000 slices of the original echo data, where the size of each slice was 512×512 , and a 20 dB AWGN was added. In particular, we adopted the traditional CSA to generate the labels of measured data, in which the sidelobe

was further suppressed by feature enhancement. The following three cases were taken into account to assess the effectiveness of our proposed imaging nets.

- (1) *Case I*: Simulated experiments of nonsparse scenes.
- (2) *Case II*: Simulated experiments of real scenes.
- (3) *Case III*: Measured experiments.

The fixed mutual parameters were set as follows: the convolution size was $N_f = 32$, the filter size was $\omega_f \times \omega_f = 3 \times 3$, the learning rate was 1×10^{-4} , the default number of layers was $L = 9$, and the default number of epochs was 101. All our experiments were implemented in the PyTorch framework with the Adam optimizer and accelerated by an NVIDIA Tesla V100 GPU.

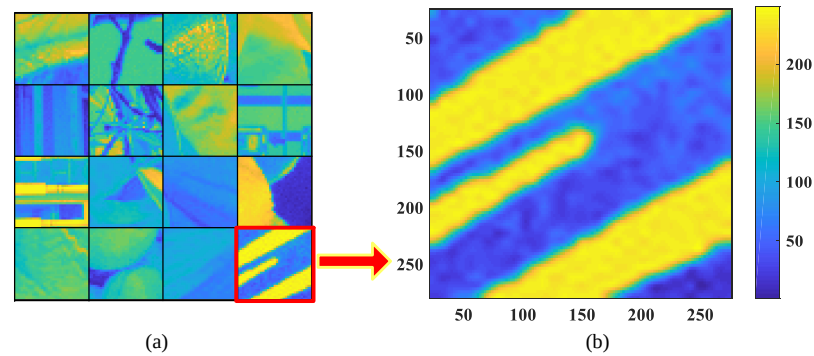


Figure 5. (a) Partial nonsparse labels in the training set; (b) one example of nonsparse labels.

4.1. Simulated Experiments of Nonsparse Scenes

Here, we focused on a nonsparse imaging scene in which the discretized grids $P \times Q$ were fixed to 256×256 . In this section, five algorithms, including MF (CSA), traditional CS without sparse representation, DCT-based CS (DCT-CS) [17], mixed sparse representation-based CS (MSR-CS) [18], and CSA-Net, were adopted as comparative experiments. CSA-Net is a simplified SR-CSA-Net without CNN blocks, which can be regarded as a 2D version of ISTA-Net [24] improved by the CSA operator $G(\cdot)$. In addition, four evaluation indices, including the NMSE, peak signal-to-noise ratio (PSNR) [45], structural similarity index for measuring [45], and mean computing time, were applied to evaluate the reconstruction performance of different algorithms. The NMSE is defined as

$$\text{NMSE} = \frac{\|\hat{\mathbf{X}} - \mathbf{X}_{\text{label}}\|_F^2}{\|\mathbf{X}_{\text{label}}\|_F^2} \quad (33)$$

where $\hat{\mathbf{X}}$ represents the reconstruction result and $\mathbf{X}_{\text{label}}$ denotes the label image. The definition of PSNR is

$$\begin{cases} \text{PSNR} = 10 \cdot \log_{10} \left(\frac{K_{\hat{\mathbf{X}}}^2}{\text{MSE}} \right) \\ \text{MSE} = \frac{1}{PQ} \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} [\hat{\mathbf{X}}(p, q) - \mathbf{X}_{\text{label}}(p, q)]^2 \end{cases} \quad (34)$$

where MSE is the mean square error between $\hat{\mathbf{X}}$ and $\mathbf{X}_{\text{label}}$, and $K_{\hat{\mathbf{X}}}$ denotes the pixel value range of $\hat{\mathbf{X}}$. In addition, the SSIM is defined as

$$\text{SSIM} = \frac{(2E[\hat{\mathbf{X}}]E[\mathbf{X}_{\text{label}}] + \alpha_1)(2\text{Cov}[\hat{\mathbf{X}}, \mathbf{X}_{\text{label}}] + \alpha_2)}{(E[\hat{\mathbf{X}}]^2 + E[\mathbf{X}_{\text{label}}]^2 + \alpha_1)(\text{Var}[\hat{\mathbf{X}}] + \text{Var}[\mathbf{X}_{\text{label}}] + \alpha_2)} \quad (35)$$

where $\text{Cov}[\cdot, \cdot]$ represents the cross-covariance function, and α_1 and α_2 were set to $(0.01 \cdot K_{\hat{\mathbf{X}}})^2$ and $(0.03 \cdot K_{\hat{\mathbf{X}}})^2$, respectively.

According to the definitions in Equations (33)–(35), it can be concluded that the NMSE describes the accuracy of the reconstruction results, the PSNR reflects the distortion of images after reconstruction, and the SSIM indicates the similarity of the reconstruction result and ground truth, where a larger SSIM value means better performance.

(1) *Results with complete sampling*: As illustrated in Figure 6, the reconstruction results of the proposed SR-CSA-Net and SR-CSA-Net-plus were compared with those of the above five algorithms with complete sampling. Figure 6 verifies that the reconstruction results of SR-CSA-Net and SR-CSA-Net-plus have more details and sharper edges than other methods, showing the superiority of the proposed imaging nets.

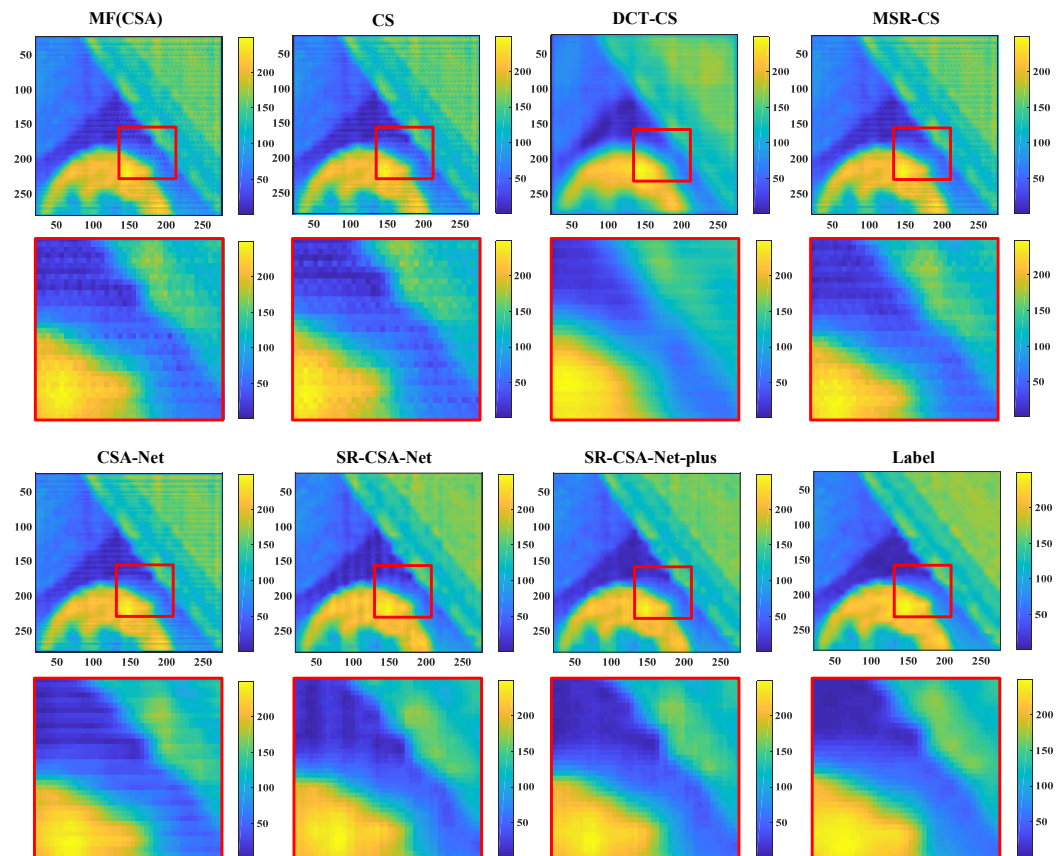


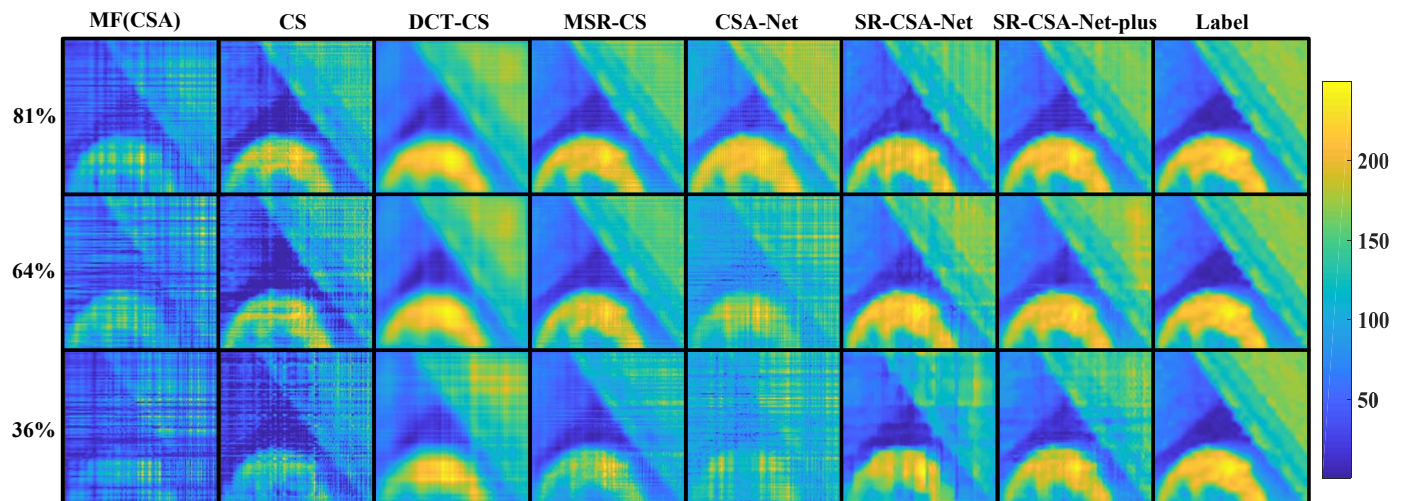
Figure 6. Comparison of seven SAR imaging methods with complete sampling (including the proposed SR-CSA-Net and SR-CSA-Net-plus), where more details of the imaging results are shown below the corresponding image.

Table 2 lists the three evaluation indices of the imaging results in Figure 6, obtained from the average of 100 Monte Carlo experiments. In addition, the optimal evaluation values are in bold font. Although traditional MF, CS-driven, and CSA-Net without CNN blocks obtain acceptable imaging results, the proposed SR-CSA-Net and SR-CSA-Net-plus outperform all the above methods by a large margin. As expected, SR-CSA-Net-plus has better performance and is superior to SR-CSA-Net based on the evaluation values.

(2) *Results with different sampling rates*: Then, we examined the influence of the sampling rate denoted as $\eta = M'N'/MN$ on the reconstruction performance. We adopted the same training set to ensure a fair comparison, and different sampling rates, including $\eta = 81\%$, 64% , and 36% , were applied to the above seven SAR imaging methods to more deeply investigate the superiority of our SR-CSA-Net and SR-CSA-Net-plus. The imaging results by MF, CS-driven methods, and deep unfolding-based methods under the premise of downsampling are shown in Figure 7 from left to right.

Table 2. Evaluation values with complete sampling.

Method	NMSE	PSNR (dB)	SSIM
MF(CSA)	0.0113	24.72	0.7685
CS	0.0173	22.88	0.7587
DCT-CS [17]	0.0108	24.91	0.7713
MSR-CS [18]	0.0096	25.45	0.7858
CSA-Net	0.0092	25.62	0.8106
SR-CSA-Net	0.0024	31.74	0.9069
SR-CSA-Net-plus	0.0018	32.54	0.9496

**Figure 7.** Comparison of seven SAR imaging methods with different sampling rates (including $\eta = 81\%$, 64% , and 36%).

In Figure 7, the ambiguity phenomenon appears in the reconstruction results of CSA and CSA-Net because both the conventional MF and the deep unfolding method without sparse representation cannot address the downsampled echo of the nonsparse scene. In contrast, the ambiguity phenomenon is eliminated in the reconstruction results of CS-driven methods and our proposed imaging nets because those methods are effective for the downsampled echo of the nonsparse scene. While CS without sparse representation may alleviate noise and sidelobe interference to some extent, its ambiguity phenomenon is also severe and grows in severity as sampling rates decrease. The reconstruction results of sparse-representation-based CS methods, including DCT-CS and MSR-CS, are acceptable when $\eta = 81\%$ and $\eta = 64\%$ but seriously deteriorate when $\eta = 36\%$. As a result, the proposed SAR imaging nets achieve satisfactory reconstruction results with complete sampling data to 36% downsampling data.

The PSNR, NMSE, and SSIM values of various η and the mean computing time of various algorithms are listed in Table 3. It shows that the imaging quality declines for all the methods with a decrease in sampling rate. The evaluation values of CSA, CS, and CSA-Net significantly decrease under the premise of downsampling. For DCT-CS and MSR-CS, the evaluation indices do not decline significantly when $\eta \geq 64\%$, but the evaluation indices decline significantly when $\eta = 36\%$. Although the evaluation indices of our proposed imaging nets also decline with a reduction in the sampling rate, the proposed SR-CSA-Net-plus obtains the best NMSE, PSNR, and SSIM values in each case, followed by SR-CSA-Net.

Table 3. Evaluation values with different sampling rates.

Method	$\eta = 81\%$			$\eta = 64\%$			$\eta = 36\%$			Runtime (s)
	NMSE	PSNR (dB)	SSIM	NMSE	PSNR (dB)	SSIM	NMSE	PSNR (dB)	SSIM	
MF(CSA)	0.1930	12.40	0.2981	0.1991	11.89	0.2110	0.3450	9.88	0.1409	1.37/-
CS	0.0740	16.57	0.4603	0.1262	14.25	0.3043	0.2654	11.02	0.2062	7.86/-
DCT-CS [17]	0.0144	23.67	0.7331	0.0176	23.13	0.7113	0.0507	18.16	0.5726	60.94/-
MSR-CS [18]	0.0118	24.55	0.7499	0.0162	23.16	0.7421	0.0739	16.45	0.5526	123.15/-
CSA-Net	0.0556	17.81	0.6287	0.0693	16.85	0.3489	0.1219	14.40	0.2076	-/0.036
SR-CSA-Net	0.0074	26.55	0.8153	0.0092	25.16	0.7995	0.0479	19.16	0.6069	-/0.064
SR-CSA-Net-plus	0.0044	29.27	0.9014	0.0066	27.09	0.8566	0.0207	21.77	0.7054	-/0.068

The superior reconstruction performance of the proposed SR-CSA-Net and SR-CSA-Net-plus are demonstrated by the PSNR and NMSE values. Furthermore, the SSIM values and the mean computing time indicate that the proposed imaging nets have outstanding target enhancement performance and reconstruction efficiency. For different sampling rates, the echo matrix is filled to a fixed size, so there is almost no difference in imaging time. The three deep-unfolding-based imaging nets have comparable imaging times, and CSA-Net takes the shortest imaging time among the three methods. Due to the existence of module N and module N_{plus} , the parameter sets in SR-CSA-Net and SR-CSA-Net-plus include learnable parameters related to sparse representation, which results in a slightly larger imaging time of the proposed imaging nets than that of CSA-Net.

(3) *Ablation study*: To further evaluate the superiority of our proposed strategy, an ablation study of SR-CSA-Net-plus with different layers, epochs, and network modules was performed and is presented in this part. The comparison results of PSNR with respect to different imaging methods are presented in Figure 8 and Table 4, where SR-CSA-Net-plus without skip connections (SCs) and $II(\cdot)$ stands for SR-CSA-Net, and SR-CSA-Net-plus without SCs, $II(\cdot)$, $F(\cdot)$, and $\tilde{F}(\cdot)$ refers to CSA-Net.

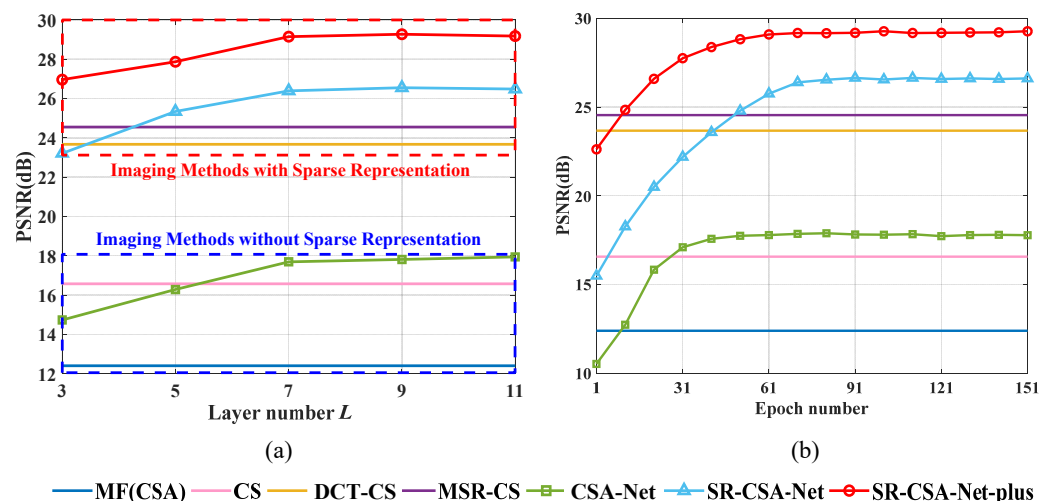


Figure 8. Performance investigation with respect to the SC and sparse representation with different layers and epochs ($\eta = 81\%$). (a) Layer; (b) epoch.

The PSNR curves of the proposed imaging nets increase with the number of layers or epochs and converge at a certain point, as shown in Figure 8. We also observe that SR-CSA-Net-plus significantly outperforms SR-CSA-Net, CSA-Net, and other conventional methods in terms of reconstruction performance. When we remove SCs and $II(\cdot)$, the results are still acceptable. However, after removing the sparse representation blocks $F(\cdot)$ and $\tilde{F}(\cdot)$, the system suffers a significant performance degradation. The PSNR curves converge when $L \geq 7$. Table 4 indicates that under this condition, SR-CSA-Net-plus achieves approximately 3 dB and 12 dB gains over SR-CSA-Net and CSA-Net, respectively. Furthermore, CSA-Net achieves the fastest training convergence as the number of epochs

increases, while SR-CSA-Net-plus registers as the second-fastest and achieves the best reconstruction performance among the three deep-unfolding-based imaging nets.

Table 4. PSNR Value of ablation study ($\eta = 81\%$).

Method	SC	$\Pi(\cdot)$	$F(\cdot)$	$\tilde{F}(\cdot)$	PSNR (dB) (epoch = 101)		
					$L = 7$	9	11
SR-CSA-Net-plus	✓	✓	✓	✓	29.15	29.27	29.18
SR-CSA-Net	×	×	✓	✓	26.39	26.55	26.48
CSA-Net	×	×	×	×	17.69	17.81	17.94

Method	SC	$\Pi(\cdot)$	$F(\cdot)$	$\tilde{F}(\cdot)$	PSNR (dB) ($L = 9$)		
					epoch = 71	101	131
SR-CSA-Net-plus	✓	✓	✓	✓	29.17	29.27	29.20
SR-CSA-Net	×	×	✓	✓	26.37	26.55	26.61
CSA-Net	×	×	×	×	17.86	17.81	17.79

The BN operator was designed in our proposed imaging nets to accelerate convergence, which is also critical to the improvement in network performance. To directly reflect the influence of the BN operator, the proposed two imaging nets were compared with those without BN operators. Figure 9 and Table 5 show the NMSE comparison results of four different imaging nets with various numbers of layers and epochs, where w/o is the abbreviation of “without”.

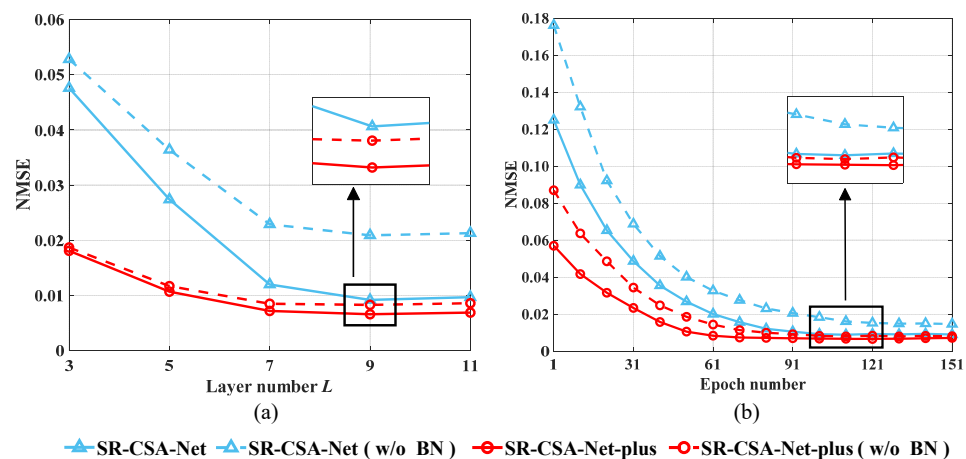


Figure 9. Performance investigation with respect to the BN operator with different layers and epochs ($\eta = 64\%$). (a) Layer; (b) epoch.

As illustrated in Figure 9 and Table 5, the NMSE values of SR-CSA-Net and SR-CSA-Net-plus are lower than those without BN operators at each layer or epoch. The performance of SR-CSA-Net-plus is not appreciably affected by the removal of BN operators, whereas the advantages of BN operators are more apparent for SR-CSA-Net. Figure 9a verifies that the NMSE curves of SR-CSA-Net-plus (two red curves) converge and tend to be flat when $L \geq 7$, while the curves of SR-CSA-Net (two blue curves) converge when $L \geq 9$. Figure 9b shows that SR-CSA-Net-plus with BN operators performs best in terms of convergence speed and achieves a stable performance over sixty epochs. Hence, the convergence speed and reconstruction accuracy of imaging nets can be improved to some extent by introducing the BN operator. Nevertheless, Table 5 suggests that SR-CSA-Net-plus without BN operators still achieves better reconstruction performance than SR-CSA-Net with BN operators.

Table 5. NMSE Value of ablation study ($\eta = 64\%$).

Method	SC	$\Pi(\cdot)$	BN	NMSE (epoch = 101)		
				$L = 7$	9	11
SR-CSA-Net-plus	✓	✓	✓	0.0072	0.0066	0.0069
SR-CSA-Net-plus (w/o BN)	✓	✓	×	0.0085	0.0083	0.0086
SR-CSA-Net	×	×	✓	0.0120	0.0092	0.0097
SR-CSA-Net (w/o BN)	×	×	×	0.0229	0.0206	0.0213
Method	SC	$\Pi(\cdot)$	BN	NMSE ($L = 9$)		
				epoch = 71	101	131
SR-CSA-Net-plus	✓	✓	✓	0.0074	0.0066	0.0068
SR-CSA-Net-plus (w/o BN)	✓	✓	×	0.0114	0.0083	0.0081
SR-CSA-Net	×	×	✓	0.0157	0.0092	0.0091
SR-CSA-Net (w/o BN)	×	×	×	0.0276	0.0206	0.0168

In conclusion, $F(\cdot)$ and $\tilde{F}(\cdot)$ are notably beneficial for the high-quality reconstruction of nonsparse SAR scenes. The superiority of SR-CSA-Net-plus over SR-CSA-Net can be attributed to the high-frequency component recovery operator $\Pi(\cdot)$ and SCs introduced by module N_{plus} . These two designs in SR-CSA-Net-plus improve the sparse representation ability and reconstruction performance. For both SR-CSA-Net and SR-CSA-Net-plus, the BN operator introduced by module N and module N_{plus} facilitates the convergence of the network training while improving the reconstruction accuracy.

4.2. Simulated Experiments of Real Scenes

The simulated experiments in Section 4.1 proved the effectiveness and feasibility of SR-CSA-Net and SR-CSA-Net-plus for nonsparse SAR reconstruction by comparing them with traditional MF and CS-driven methods. In this subsection, we further compare them with state-of-the-art deep unfolding methods, including RMIST-Net [25] and RDA-Net [31], which also reduce the storage burden and are designed by unfolding ISTA. To demonstrate the universality of the proposed imaging nets, we chose three real scenes with weak sparsity in an open SAR ships detection dataset (SSDD) [46] to generate the SAR echoes, which were directly input into the trained imaging nets in Section 4.1, avoiding unnecessary training and calculation. The three real scenes with the indices 231, 1080, and 1088 in the SSDD were cut into images of size of 256×256 , and their sparsity degrees were scene 3 > scene 1 > scene 2. The radar system parameters were the same as those in the previous section, and the fixed mutual parameters of all networks were set to be the same for a fair comparison. Figure 10 shows the imaging results of three real scenes obtained by RMIST-Net, RDA-Net, SR-CSA-Net, and SR-CSA-Net-plus when downsampling with $\eta = 81\%$ and $\eta = 64\%$. In addition, the corresponding evaluation indices are listed in Table 6, where the best values are marked in boldface.

Figure 10 and Table 6 reflect the superiority and robustness of our proposed two imaging nets for a real SAR scene reconstruction when downsampling. Although the reconstruction quality of all the above imaging nets weakens with the decreasing scene sparsity and sampling rate, SR-CSA-Net-plus has the best advantage in all three scenes and both downsampling schemes, followed by SR-CSA-Net. For the other two comparative imaging nets, the imaging results of RDA-Net and RMIST-Net are acceptable when $\eta = 81\%$, but those of RMIST-Net are streaky when $\eta = 64\%$. Furthermore, RDA-Net performs better than RMIST-Net because the iterative parameters and compensation matrices in RDA-Net are learned, while RMIST-Net only learns the ISTA parameters, and its RM kernel is predefined. This issue leads to RDA-Net having the largest number of learnable parameters but RMIST-Net having the least number of learnable parameters. In contrast, the number of parameters in our methods is moderate. Therefore, our proposed imaging nets achieve a compromise between reconstruction performance and computing speed compared with RDA-Net and RMIST-Net.

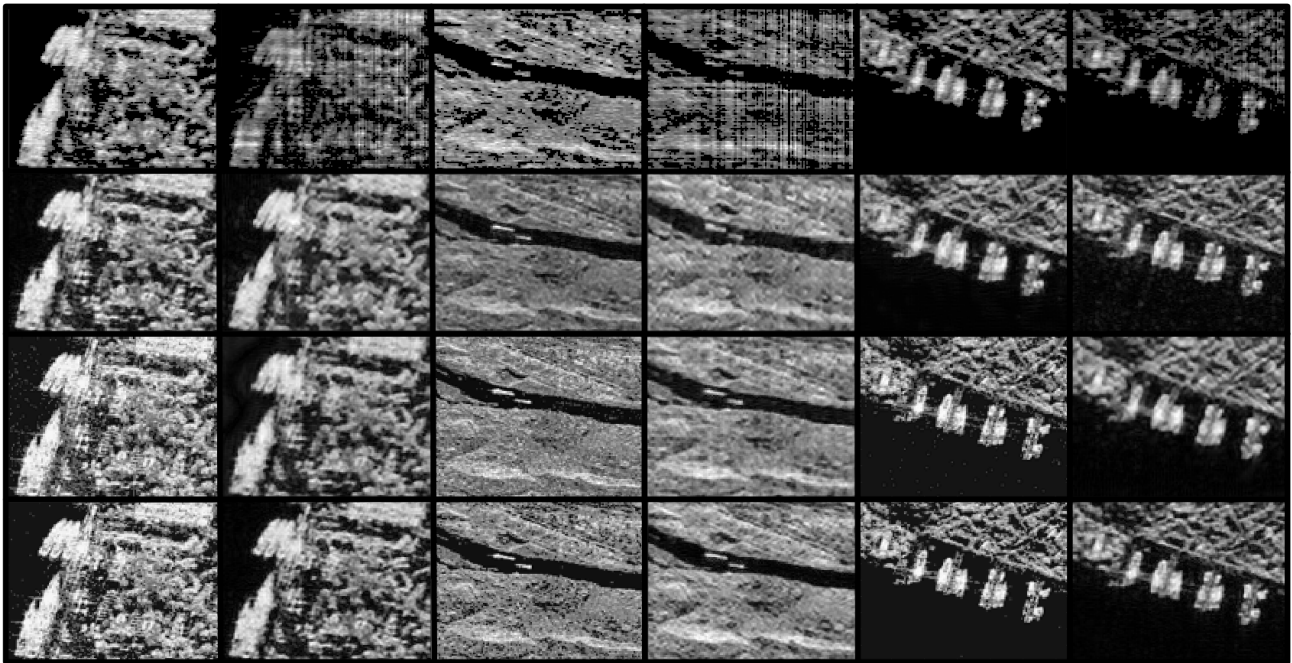


Figure 10. Imaging results of three real scenes by RMIST-Net, RDA-Net, SR-CSA-Net, and SR-CSA-Net-plus from top to bottom. (Columns 1 and 2: scene 1. Columns 3 and 4: scene 2. Columns 5 and 6: scene 3. Columns 1, 3, and 5: downsampling with $\eta = 81\%$. Columns 2, 4, and 6: downsampling with $\eta = 64\%$.)

Table 6. Performance comparison of different imaging nets in nonsparse SAR scenes.

Method	NMSE/PSNR (dB)/SSIM, $\eta = 81\%$			NMSE/PSNR (dB)/SSIM, $\eta = 64\%$			Runtime (s)
	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3	GPU
RMIST-Net [25]	0.0938/16.08/0.5866	0.1487/15.28/0.4947	0.1791/17.16/0.3930	0.1984/12.82/0.4241	0.1685/14.74/0.3871	0.2413/15.89/0.3352	0.042
RDA-Net [31]	0.0449/19.27/0.7533	0.0811/17.92/0.6314	0.1042/18.81/0.6073	0.0677/17.49/0.6564	0.1482/15.30/0.5495	0.1609/17.41/0.4627	0.075
SR-CSA-Net	0.0259/21.66/0.8575	0.0490/20.11/0.8603	0.0527/22.47/0.8279	0.0542/18.45/0.7117	0.1281/15.93/0.6167	0.1414/18.19/0.4521	0.064
SR-CSA-Net-plus	0.0123/24.88/0.9033	0.0283/22.49/0.8931	0.0237/25.94/0.9164	0.0439/19.37/0.7669	0.1187/16.26/0.6974	0.1066/19.42/0.6326	0.068

4.3. Measured Experiments

The previous simulated experiments validated that SR-CSA-Net and SR-CSA-Net-plus were superior to conventional CS-driven and deep-unfolding-based algorithms in both reconstruction quality and efficiency. In this subsection, some experimental results are given based on the measured data of RADARSAT-1 to investigate whether the proposed imaging strategy still performs well in the measured experiments. Since SR-CSA-Net-plus performed better, we chose it as an example to further verify our proposed strategy and retrain it in the way introduced at the beginning of Section 4. SR-CSA-Net-plus was tested and compared with the four above-mentioned methods, including CSA, MSR-CS, RMIST-Net, and RDA-Net. In addition, an AMP-unfolded 2D SAR imaging method inspired by [30], called AMP-Net, was introduced in this experiment for comparison with other SAR imaging methods unfolded by ISTA. The corresponding imaging results with $\eta = 81\%$ and $\eta = 64\%$ of a sparse scene and a nonsparse scene, namely, a harbor and seashore, are shown in Figures 11 and 12, respectively. The image entropy (ENT) and the previously mentioned evaluation values are listed in Table 7.

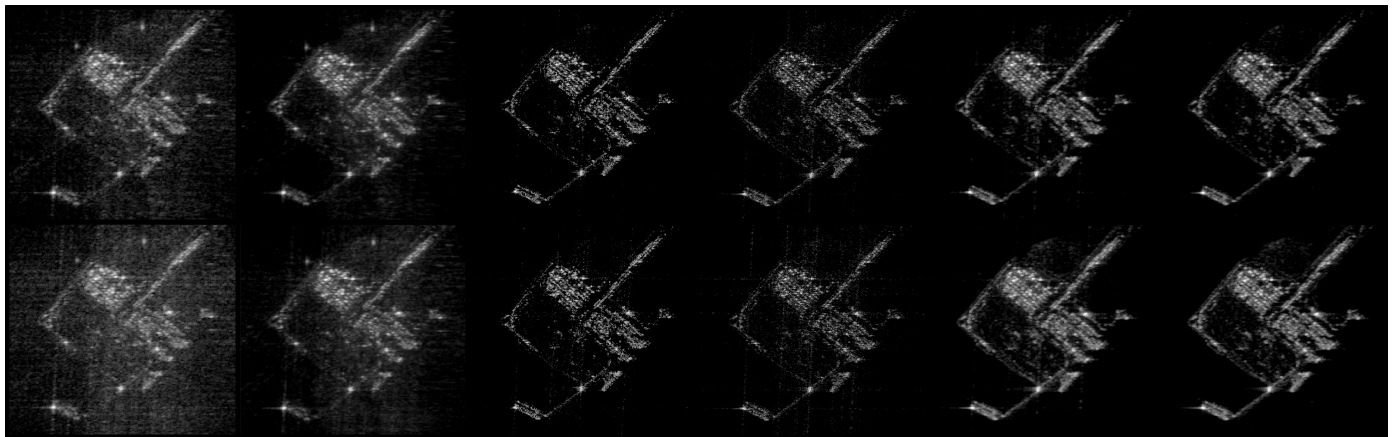


Figure 11. Imaging results of measured data for a harbor by CSA, MSR-CS, AMP-Net, RMIST-Net, RDA-Net, and SR-CSA-Net-plus from left to right. (Rows 1: downsampling with $\eta = 81\%$. Rows 2: downsampling with $\eta = 64\%$).

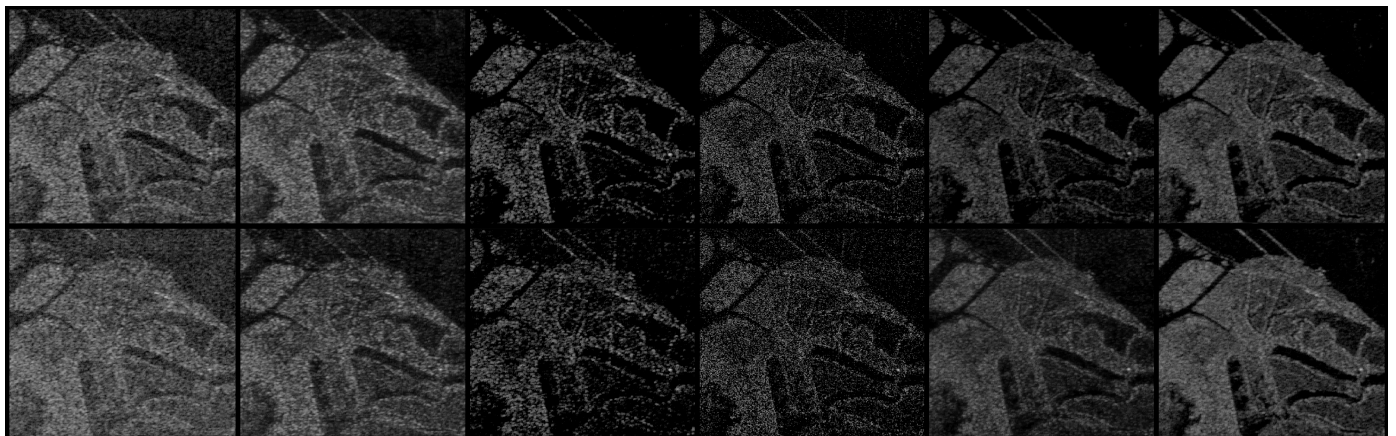


Figure 12. Imaging results of measured data for a seashore by CSA, MSR-CS, AMP-Net, RMIST-Net, RDA-Net, and SR-CSA-Net-plus from left to right. (Rows 1: downsampling with $\eta = 81\%$. Rows 2: downsampling with $\eta = 64\%$).

Table 7. Performance comparison of different imaging nets on measured data.

Method	NMSE/PSNR (dB)/SSIM/ENT, $\eta = 81\%$		NMSE/PSNR (dB)/SSIM/ENT, $\eta = 64\%$	
	Harbor	Seashore	Harbor	Seashore
CSA	0.0509/18.07/0.4988/3.7511	0.0488/18.41/0.6472/4.0985	0.0587/17.18/0.4511/3.9849	0.0536/17.96/0.6206/4.1333
MSR-CS [18]	0.0110/23.86/0.6246/3.1717	0.0266/19.90/0.6310/4.0857	0.0127/23.24/0.5823/3.5203	0.0302/19.35/0.5978/3.9964
AMP-Net [30]	0.0065/27.14/0.8858/ 1.5613	0.0143/23.30/0.7062/ 2.6594	0.0066/26.10/0.8651/ 1.5712	0.0122/23.27/0.6721/ 2.6949
RMIST-Net [25]	0.0063/27.08/0.8632/1.7454	0.0278/21.80/0.7284/2.7329	0.0065/25.08/0.8376/1.9739	0.0334/19.87/0.6349/2.7519
RDA-Net [31]	0.0042/27.93/0.8946/2.1576	0.0055/26.49/0.8682/3.4257	0.0055/25.85/0.8632/2.4056	0.0071/24.61/0.7405/3.8262
SR-CSA-Net-plus	0.0037/28.13/0.9083/2.1361	0.0026/29.67/0.9145/3.2951	0.0049/26.77/0.8878/2.2255	0.0043/27.86/0.8680/3.6207

From Figures 11 and 12 and Table 7, the four networks achieve comparable reconstruction results in the sparse scene, but the proposed strategy has apparent advantages in the nonsparse scene. Specifically, SR-CSA-Net-plus achieves approximately 3~8 dB PSNR gains over RDA-Net, AMP-Net, and RMIST-Net, and reconstructs more details and sharper edges. Furthermore, the NMSE, PSNR, and SSIM values of SR-CSA-Net-plus are optimal among the six algorithms, which is consistent with the simulated experiments, while the ENT of SR-CSA-Net-plus ranks third. However, for nonsparse scenes, ENT cannot be used solely as an indicator to evaluate reconstruction performance. Although AMP-Net and RMIST-Net obtain smaller ENT values for both the harbor and seashore, they lose much information in the reconstruction process due to the lack of sparse representation ability, such as the structure, edges, and smooth components of scattered fields. This loss of information is especially pronounced in nonsparse scenes. In addition, since AMP is superior to ISTA, the reconstruction performance of AMP-Net is slightly better than that of RMIST-Net and has comparable performance with RDA-Net for sparse scenes. Thus, we are interested in combining AMP-Net with a sparse representation structure, which is expected to achieve better reconstruction performance than SR-CSA-Net-plus in our future work. According to the complexity analysis in Section 3.4, the number of parameters in RMIST-Net, AMP-Net, RDA-Net, and SR-CSA-Net-plus is 18, 20, 9.437×10^6 , and 3.738×10^5 , respectively, when the echo size is 512×512 . Therefore, RMIST-Net spends the least runtime among the four networks under the premise of the same fixed mutual parameters and 2D imaging mechanism, followed by AMP-Net and SR-CSA-Net-plus. In conclusion, the proposed strategy combines the merits of the CS and DL methods and achieves consistently high-quality reconstruction results in sparse and weakly sparse scenes while remaining computationally competent.

5. Conclusions

In this article, we presented a novel deep unfolding strategy for SAR imaging of nonsparse scenes. There are two limitations in conventional nonsparse SAR scene reconstruction: the vectorized CS optimization model requires a high computational cost to process the large-scale matrix, and the transform-domain-based methods cannot obtain an optimal sparse basis. Thus, the approximated observation model and CNN blocks were introduced in place of the conventional exact observation model and sparse transform function. This strategy combined deep unfolding and traditional neural network architectures, unlike model-driven CS reconstruction algorithms and data-driven DL imaging methods. In practical terms, we incorporated the CSA operator and ISTA within an efficient deep SAR imaging net, SR-CSA-Net, and developed its enhanced version, SR-CSA-Net-plus. Theoretically, these two imaging nets pertained to the model-driven deep hierarchical architecture. Each layer of our proposed imaging nets was designed as a combination of the linearity and nonlinearity modules, corresponding to the iteration steps of ISTA optimizations. Moreover, we embedded CNN blocks in the nonlinearity module for the sparse representation instead of adopting the existing sparse transform methods. The experimental results indicated that our proposed imaging strategy achieved significant performance improvements for the reconstruction quality of nonsparse scenes, showing superiority in terms of computational efficiency and complexity compared with conventional CS-driven and deep unfolding methods.

As a general imaging strategy, we envisage that such model-driven deep unfolding methods with CNN structures will have significant potential in various SAR imaging applications. However, our proposed nonsparse SAR scene imaging nets only consider the side-looking mode; they suffer from imaging quality degradation in squint mode. In future work, we are interested in integrating the processing of range cell migration and Doppler center shift (resulting in geometry distortion) into the proposed strategy to improve its adaptability and reconstruction ability for squint imaging.

Author Contributions: Conceptualization and methodology, H.Z. and J.N.; software, H.Z. and K.L.; validation, H.Z., K.L. and J.N.; resources, Q.Z. and Y.L.; writing—original draft preparation, H.Z.; writing—review and editing, H.Z., J.N. and Q.Z.; funding acquisition, Q.Z. and J.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under grant 62131020, grant 62001508, and grant 61971434.

Data Availability Statement: Not applicable.

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

References

1. Lan, L.; Marino, A.; Aubry, A.; De Maio, A.; Liao, G.; Xu, J.; Zhang, Y. GLRT-Based Adaptive Target Detection in FDA-MIMO Radar. *IEEE Trans. Aerosp. Electron. Syst.* **2021**, *57*, 597–613. [\[CrossRef\]](#)
2. Bamler, R. A comparison of range-Doppler and wavenumber domain SAR focusing algorithms. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 706–713. [\[CrossRef\]](#)
3. Raney, R.; Runge, H.; Bamler, R.; Cumming, I.; Wong, F. Precision SAR processing using chirp scaling. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 786–799. [\[CrossRef\]](#)
4. Ulander, L.; Hellsten, H.; Stenstrom, G. Synthetic-aperture radar processing using fast factorized back-projection. *IEEE Trans. Aerosp. Electron. Syst.* **2003**, *39*, 760–776. [\[CrossRef\]](#)
5. 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. [\[CrossRef\]](#)
6. Donoho, D. Compressed sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306. [\[CrossRef\]](#)
7. Eldar, Y.C.; Kutyniok, G. *Compressed Sensing: Theory and Applications*; Cambridge University Press: Cambridge, UK, 2012.
8. Bi, H.; Zhu, D.; Bi, G.; Zhang, B.; Hong, W.; Wu, Y. FMCW SAR Sparse Imaging Based on Approximated Observation: An Overview on Current Technologies. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 4825–4835. [\[CrossRef\]](#)
9. Bi, H.; Lu, X.; Yin, Y.; Yang, W.; Zhu, D. Sparse SAR Imaging Based on Periodic Block Sampling Data. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–12. [\[CrossRef\]](#)
10. Kelly, S.; Yaghoobi, M.; Davies, M. Sparsity-based autofocus for undersampled synthetic aperture radar. *IEEE Trans. Aerosp. Electron. Syst.* **2014**, *50*, 972–986. [\[CrossRef\]](#)
11. Beck, A.; Teboulle, M. A fast Iterative Shrinkage-Thresholding Algorithm with application to wavelet-based image deblurring. In Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 19–24 April 2009; pp. 693–696. [\[CrossRef\]](#)
12. Rangan, S. Generalized approximate message passing for estimation with random linear mixing. In Proceedings of the 2011 IEEE International Symposium on Information Theory Proceedings, St. Petersburg, Russia, 31 July–5 August 2011; pp. 2168–2172. [\[CrossRef\]](#)
13. Shi, W.; Ling, Q.; Yuan, K.; Wu, G.; Yin, W. On the Linear Convergence of the ADMM in Decentralized Consensus Optimization. *IEEE Trans. Signal Process.* **2014**, *62*, 1750–1761. [\[CrossRef\]](#)
14. Rilling, G.; Davies, M.; Mulgrew, B. Compressed sensing based compression of SAR raw data. In Proceedings of the SPARS'09—Signal Processing with Adaptive Sparse Structured Representations, Saint Malo, France, 6–9 April 2009; Gribonval, R., Ed.; Inria Rennes-Bretagne Atlantique: Rennes, France, 2009.
15. Samadi, S.; Çetin, M.; Masnadi-Shirazi, M.A. Multiple Feature-Enhanced SAR Imaging Using Sparsity in Combined Dictionaries. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 821–825. [\[CrossRef\]](#)
16. Shen, F.; Zhao, G.; Liu, Z.; Shi, G.; Lin, J. SAR Imaging With Structural Sparse Representation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 3902–3910. [\[CrossRef\]](#)
17. Ni, J.C.; Zhang, Q.; Luo, Y.; Sun, L. Compressed Sensing SAR Imaging Based on Centralized Sparse Representation. *IEEE Sens. J.* **2018**, *18*, 4920–4932. [\[CrossRef\]](#)
18. Bo, L.; Liu, F.; Zhou, C.; Zheng, W.; Hao, H. Mixed sparse representation for approximated observation-based compressed sensing radar imaging. *J. Appl. Remote Sens.* **2018**, *12*, 035015.
19. 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. [\[CrossRef\]](#)
20. Hu, C.; Wang, L.; Li, Z.; Zhu, D. Inverse Synthetic Aperture Radar Imaging Using a Fully Convolutional Neural Network. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1203–1207. [\[CrossRef\]](#)
21. Mu, H.; Zhang, Y.; Ding, C.; Jiang, Y.; Er, M.H.; Kot, A.C. DeepImaging: A Ground Moving Target Imaging Based on CNN for SAR-GMTI System. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 117–121. [\[CrossRef\]](#)
22. Lu, Z.J.; Qin, Q.; Shi, H.Y.; Huang, H. SAR moving target imaging based on convolutional neural network. *Digit. Signal Process.* **2020**, *106*, 102832. [\[CrossRef\]](#)
23. Rittenbach, A.; Walters, J.P. RDAnet: A Deep Learning Based Approach for Synthetic Aperture Radar Image Formation. *arXiv* **2020**, arXiv:2001.08202.

24. Yonel, B.; Mason, E.; Yazıcı, B. Deep Learning for Passive Synthetic Aperture Radar. *IEEE J. Sel. Top. Signal Process.* **2018**, *12*, 90–103. [\[CrossRef\]](#)
25. Wang, M.; Wei, S.; Liang, J.; Zeng, X.; Wang, C.; Shi, J.; Zhang, X. RMIST-Net: Joint Range Migration and Sparse Reconstruction Network for 3-D mmW Imaging. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5205117. [\[CrossRef\]](#)
26. Li, M.; Wu, J.; Huo, W.; Jiang, R.; Li, Z.; Yang, J.; Li, H. Target-Oriented SAR Imaging for SCR Improvement via Deep MF-ADMM-Net. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5223314. [\[CrossRef\]](#)
27. Zhang, H.; Ni, J.; Xiong, S.; Luo, Y.; Zhang, Q. Omega-KA-Net: A SAR Ground Moving Target Imaging Network Based on Trainable Omega-K Algorithm and Sparse Optimization. *Remote Sens.* **2022**, *14*, 1664. [\[CrossRef\]](#)
28. Chen, L.; Ni, J.; Luo, Y.; He, Q.; Lu, X. Sparse SAR Imaging Method for Ground Moving Target via GMTSI-Net. *Remote Sens.* **2022**, *14*, 4404. [\[CrossRef\]](#)
29. Li, R.; Zhang, S.; Zhang, C.; Liu, Y.; Li, X. Deep Learning Approach for Sparse Aperture ISAR Imaging and Autofocusing Based on Complex-Valued ADMM-Net. *IEEE Sens. J.* **2021**, *21*, 3437–3451. [\[CrossRef\]](#)
30. Wei, S.; Liang, J.; Wang, M.; Shi, J.; Zhang, X.; Ran, J. AF-AMPNet: A Deep Learning Approach for Sparse Aperture ISAR Imaging and Autofocusing. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5206514. [\[CrossRef\]](#)
31. Kang, L.; Sun, T.; Luo, Y.; Ni, J.; Zhang, Q. SAR Imaging Based on Deep Unfolded Network With Approximated Observation. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5228514. [\[CrossRef\]](#)
32. Song, C.B.; Xia, S.T. Sparse Signal Recovery by ℓ_q Minimization Under Restricted Isometry Property. *IEEE Signal Process. Lett.* **2014**, *21*, 1154–1158. [\[CrossRef\]](#)
33. Shensa, M.J. The discrete wavelet transform: Wedding the a trous and Mallat algorithms. *IEEE Trans. Signal Process.* **1992**, *40*, 2464–2482. [\[CrossRef\]](#)
34. Ahmed, N.; Natarajan, T.; Rao, K. Discrete Cosine Transform. *IEEE Trans. Comput.* **1974**, *100*, 90–93. [\[CrossRef\]](#)
35. Yang, Y.; Sun, J.; Li, H.; Xu, Z. ADMM-CSNet: A deep learning approach for image compressive sensing. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *42*, 521–538. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Cui, Y.; Wu, D.; Huang, J. Optimize TSK Fuzzy Systems for Classification Problems: Minibatch Gradient Descent With Uniform Regularization and Batch Normalization. *IEEE Trans. Fuzzy Syst.* **2020**, *28*, 3065–3075. [\[CrossRef\]](#)
37. Zhang, J.; Ghanem, B. ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Online, 1 July 2018; pp. 1828–1837.
38. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Online, 5 July 2016; pp. 770–778.
39. Iandola, F.; Moskewicz, M.; Karayev, S.; Girshick, R.; Darrell, T.; Keutzer, K. Densenet: Implementing efficient convnet descriptor pyramids. *arXiv* **2014**, arXiv:1404.1869.
40. Yu, X.; Efe, M.O.; Kaynak, O. A general backpropagation algorithm for feedforward neural networks learning. *IEEE Trans. Neural Netw.* **2002**, *13*, 251–254.
41. Ito, D.; Takabe, S.; Wadayama, T. Trainable ISTA for sparse signal recovery. *IEEE Trans. Signal Process.* **2019**, *67*, 3113–3125. [\[CrossRef\]](#)
42. Takabe, S.; Wadayama, T.; Eldar, Y.C. Complex trainable ista for linear and nonlinear inverse problems. In Proceedings of the ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 5020–5024.
43. Borgerding, M.; Schniter, P.; Rangan, S. AMP-Inspired Deep Networks for Sparse Linear Inverse Problems. *IEEE Trans. Signal Process.* **2017**, *65*, 4293–4308. [\[CrossRef\]](#)
44. Xiong, T.; Xing, M.; Wang, Y.; Wang, S.; Sheng, J.; Guo, L. Minimum-entropy-based autofocus algorithm for SAR data using chebyshev approximation and method of series reversion, and its implementation in a data processor. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 1719–1728. [\[CrossRef\]](#)
45. Horé, A.; Ziou, D. Image Quality Metrics: PSNR vs. SSIM. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 2366–2369. [\[CrossRef\]](#)
46. Li, J.; Qu, C.; Shao, J. Ship detection in SAR images based on an improved faster R-CNN. In Proceedings of the 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSAR DATA), Beijing, China, 13–14 November 2017; pp. 1–6. [\[CrossRef\]](#)

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