







CSR-Net: A Novel Complex-Valued Network for Fast and Precise 3-D Microwave Sparse Reconstruction

Mou Wang , *Student Member, IEEE*, Shunjun Wei , *Member, IEEE*, Jun Shi , *Member, IEEE*, Yue Wu , *Student Member, IEEE*, Qizhe Qu , *Student Member, IEEE*, Yuanyuan Zhou, Xiangfeng Zeng , and Bokun Tian

Abstract—Since the compressed sensing (CS) theory broke through the limitation of the traditional Nyquist sampling theory, it has attracted extensive attention in the field of microwave imaging. However, in 3-D microwave sparse reconstruction application, conventional CS-based algorithms always suffer from huge computational cost. In this article, a novel 3-D microwave sparse reconstruction method based on a complex-valued sparse reconstruction network (CSR-Net) is proposed, which converts complex number operations into matrix operations for real and imaginary parts. Using the unfolding + network approximate scheme, each iteration process of CS-based iterative threshold optimization is designed as a block of CSR-Net, and a modified shrinkage term is introduced to improve the convergence performance of the approach. In addition, CSR-Net adopts a convolutional neural network module to replace a nonlinear sparse representation process, which dramatically reduces computational complexity and improves reconstruction performance over conventional CS-based iterative threshold optimization algorithms. Then, we divide the 3-D scene into a series of 2-D slices, and a phase correction scheme is adopted to ensure that the whole 3-D scene can be reconstructed with measurement matrix of a slice. Moreover, an efficient position–amplitude–random training method without additional real-measured data is employed for the proposed network, which effectively train the CSR-Net without enough real-measured data. Extensive experiment results demonstrate that CSR-Net outperforms both conventional iterative threshold optimization methods and deep network-based ISTA-NET-plus large margins. Its speed and reconstruction accuracy in 3-D imaging can achieve a state-of-the-art level.

Index Terms—3-D microwave imaging, complex-valued network, compressed sensing (CS), convolutional neural network (CNN), deep learning (DL), iterative threshold optimization.

I. INTRODUCTION

3-D MICROWAVE imaging technology [1]–[4] and its application [5], [6] have become a hotspot. Up to now, various 3-D microwave imaging solutions have been proposed [7], [8].

Manuscript received January 11, 2020; revised July 9, 2020; accepted August 4, 2020. Date of publication August 6, 2020; date of current version August 20, 2020. This work was supported in part by the National Key R&D Program of China under Grant 2017-YFB0502700, in part by the National Natural Science Foundation of China under Grant 61671113, 61501098, and in part by the High-Resolution Earth Observation Youth Foundation under Grant GFZX04061502. (Corresponding author: Mou Wang.)

The authors are with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: 798171954@qq.com; weishunjun@uestc.edu.cn; shijun@uestc.edu.cn; yuewu@std.uestc.edu.cn; 201822011037@std.uestc.edu.cn; zhouyuanyuan@std.uestc.edu.cn; zxf@std.uestc.edu.cn; tianbokun@std.uestc.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2020.3014696

They play an important role in topographic mapping, urban inspection, and so on. Compared with the traditional 2-D imaging methods [9], 3-D solutions overcome some drawbacks and can obtain more essential information about the target. However, the problems existing in the conventional 3-D imaging algorithms based on the traditional Nyquist sampling theorem and the classical signal processing theory, such as difficulty in system implementation, too many antenna elements, and low imaging resolution, are still the main difficulties in high-precision 3-D fast imaging.

In recent years, an emerging theory of compression sensing has brought revolutionary breakthroughs for accurate reconstruction of sparse signals, and attracted high attention in different application fields, including but not limited to compression imaging [10]–[12], channel coding [13], [14], medical imaging [15], pattern recognition [16], wireless network [17], and radar technology [18]–[20]. CS theory demonstrates that as long as the original signal is sparse or compressible, it can be recovered with a sub-Nyquist sampling rate, and the stronger the signal sparsity, the fewer observation data required for sparse reconstruction. CS theory uses a nonadaptive linear projection to maintain the structural information of the original signal and then reconstructs the original signal by solving the optimal solution of the underdetermined equations. In [21], this process is a dubbed sparse linear inverse problem. In the past ten years, plenty of algorithms are proposed to solve the sparse linear inverse problem, such as the Iterative Shrinkage-Thresholding Algorithm (ISTA) [22], Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [23], the Approximate Message Passing algorithm (AMP) [24], and so on. In spite of the good performance in solving linear inverse problems, conventional CS reconstruction algorithms always perform a high time complexity in 3-D microwave imaging applications, because of their iterative properties and a large amount of echo data.

Recently deep learning (DL) has been widely studied, and it has dramatically advanced the state-of-the-art for many problems in most fields. For example, target recognition, image processing, natural language processing, and many others. A representative work to embed deep networks into CS reconstruction did by Kulkarni *et al.* in [25], they designed a novel convolutional neural network (CNN) architecture dubbed ReconNet, which takes in CS measurements as input and outputs an intermediate reconstruction. They showed that ReconNet provides high

quality reconstructions for a wide range of measurement rates in real time. In [26], Mason *et al.* introduced a DL framework for inverse problems in synthetic aperture radar (SAR) imaging, and they designed a recurrent neural network architecture as an inverse solver based on the iterations of proximal gradient descent optimization methods. But their method is only for passive radar imaging problems. As for general microwave imaging problems, essentially, it can be converted into a special type of sparse linear inverse problem. Considering the wide application of linear inverse problems, CS reconstruction algorithms based on deep networks have been proposed. Deep unfolding [27] is a widely used method to convert conventional iteration algorithms into deep networks. Inspired by the conception, Borgerding *et al.* proposed the AMP inspired deep networks to solve the sparse linear inverse problems [21]. Sufficient experiments show that their network has excellent performance. But its excellent performance is overly dependent on the number of iteration layers. This greatly limits its speed. Recently, Wu *et al.* proposed a new superresolution imaging structure based on CS and deep neural network (DNN) for multiple-input multiple-output (MIMO) array SAR [28]. They designed a spatial filtering method based on CS and a group of parallel end-to-end DNN regression models to realize superresolution in the elevation direction. And they validated the method can realize the state-of-the-art superresolution imaging with a higher speed than conventional methods, but the imaging process still costs several hours. Aiming at developing a fast yet accurate network, in [29], Zhang and Ghanem proposed a novel structured deep network dubbed ISTA-NET. They adopted a novel strategy to design their network, that is to combine the speed advantage of CNN and the accuracy advantage of deep unfolding. Adequate experimental results shown that their networks outperformed existing CS methods by large margins in the reconstruction of natural images.

In this article, we introduce a deep network termed CSR-Net adapted to 3-D microwave imaging which is inspired by the work in [29]. Taking into account the needs of real-time imaging in the security field, we use the same network construction strategy as ISTA-NET, but an iterative shrinkage term derived from a classical iterative threshold optimization algorithm FISTA is additionally introduced to get better convergence performance. Our network is a combination of deep unfolding structure and CNN modules. We unfold the iteration process into a blocks based network. In each block, a CNN module is designed to represent an optimal sparse basis, which is a key part of the sparse linear inverse problem. Aiming at performing a fast yet accurate 3-D imaging process with CSR-Net, we adopt a scheme of converting 3-D imaging space to slices and reconstruct them respectively, 3-D imaging results are obtained by a heap of reconstruction results of all slices with an additional approximate fixed phase correction process. Furthermore, we propose a novel training method dubbed PAR for the proposed CSR-Net and other networks based on CS theory. Adequate experimental results show that the training method can be effectively used for 3-D microwave imaging. A summary of the main contributions of our work are as follows.

- 1) A complex-valued sparse reconstruction network (CSR-Net) is constructed, which converts complex number operations into matrix operations for real and imaginary parts. Compared with ISTA-NET-plus, we introduce a modified shrinkage term from the FISTA algorithm, which gives our network better convergence performance.
- 2) CSR-Net adopts a CNN module to learn an optimal non-linear sparsifying transform, which dramatically reduces computational complexity and improves reconstruction performance.
- 3) A phase correction scheme is introduced to fix the phase difference of echo data of each slice, aiming at ensuring the generality of the measurement matrix to slices.
- 4) An efficient positionamplituderandom (PAR) training method is designed. Even if there is not enough measured data for training, this method can effectively deal with the training problem of microwave imaging network.
- 5) CSR-Net is applied to 3-D microwave sparse reconstruction, both simulation and real-measured experimental results demonstrate that our method outperforms both conventional iterative threshold optimization methods and deep network-based method ISTA-NET-plus large margins.

The rest of this article is organized as follows. In Section II, we introduce the theory of classical iterative threshold optimization methods for CS, and a simple convergence performance comparison is analyzed. In Section III, the CSR-Net network is introduced from both unfolding structure and CNN network module. In Section IV, a 3-D microwave imaging phase correction scheme is introduced. In order to adopt CSR-Net to microwave imaging application, a complex-valued CSR-Net is modified. At the end of the section, we design an efficient training method for CSR-Net which can be extended to all CS-based networks. In Section V, we analyze the performance of the CSR-Net network and its performance on 3-D microwave imaging application. Numerical results demonstrate that our method can achieve state-of-the-art performance. Section VI concludes this article.

II. ITERATIVE THRESHOLD OPTIMIZATION METHODS FOR COMPRESSED SENSING

Compressed sensing aims to recover a signal $x \in \mathbb{R}^N$ from a linear measurement $y \in \mathbb{R}^M$ of the form

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

where $\Phi \in \mathbb{R}^{M \times N}$ dubbed measurement matrix, CS ratio is defined as $\frac{M}{N}$, and $n \in \mathbb{R}^M$ is the measurement noise assumed to be Gaussian distributed. In most application scenarios, $M \ll N$, in other words, Φ is a flat matrix. In these assumptions, the problem can be converted into solving under-determined equations.

We assume that $s \in \mathbb{R}^N$ is a sparse representation of x approximately expressed as $x = \Psi s$, where $\Psi \in \mathbb{R}^{N \times N}$ is a suitable sparse basis, such as Fourier, Wavelet, and Discrete Cosine basis. Thus, we can rewrite (1) as follows:

$$y = \Phi \Psi s + e \triangleq A s + e \quad (2)$$

where $A = \Phi\Psi$ dubbed the sparse measurement matrix. As for measurement y , the CS theory demonstrates that the signal s can be reconstructed with a high probability when it conforms to the sparse structure in a transform domain as Ψ , on the premise that Φ satisfies the RIP constraint [30]. According to the reconstruction results \hat{s} , the estimate of signal x can be effectively recovered via $\hat{x} = \Psi\hat{s}$.

One of the best-known methods of solving the reconstruction problem is through solving the convex optimization problem

$$\hat{s} = C(s) = \arg \min_s \frac{1}{2} \|y - As\|_2^2 + \lambda \|s\|_1 \quad (3)$$

where λ is a tunable parameter that controls the tradeoff between sparsity and reconstruction error constraint.

A simple way to solve (3) is iterative threshold optimization, such as ISTA and FISTA. The major iterative steps of ISTA expressed as follows:

$$\begin{cases} v_k = y - As_k \\ r_k = s_k + \chi A^T v_k \\ s_{k+1} = \eta(r_k, \theta) \end{cases} \quad (4)$$

where θ is the soft threshold, χ is a step size, the reconstruction error after the t th iteration is v_k , r_k represents rough estimate of s_k , and $\eta(r, \theta) : \mathbb{R}^N \rightarrow \mathbb{R}^N$ is the shrinkage function of the form

$$\eta(r, \theta) \triangleq \text{sgn}(r) \max\{|r| - \theta, 0\} \quad (5)$$

where $\text{sgn}(\cdot)$ represents symbolic function.

In [23], Beck and Teboulle prove that ISTA has a worst-case complexity result of $O(1/k)$ and they introduce a new ISTA with an improved complexity result of $O(1/k^2)$ dubbed FISTA. The major iterative steps are similar to ISTA, but a shrinkage term associated with historical iteration values is added

$$\begin{cases} v_k = y - As_k \\ r_k = s_k + \chi A^T v_k + \left(\frac{t_k - 1}{t_k + 1}\right) (s_k - s_{k-1}) \\ s_{k+1} = \eta(r_k, \theta) \end{cases} \quad (6)$$

where t_k is the shrinkage pseudo coefficient, and it can be iteratively updated by the following process:

$$t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2}. \quad (7)$$

Plenty of researches shows that FISTA converges in order of magnitude fewer iterations than ISTA.

III. DESIGN OF CSR-NET

In order to develop a fast yet accurate method for 3-D microwave imaging, a novel CNN-based unfolding structured network dubbed CSR-Net is introduced. Unlike the work in [21], it simply uses an unfolding structure for iterative algorithms, while our method mainly consists of two modules, unfolding module and CNN module.

A. Unfolding Module

In general, the proposed CSR-Net adopts the idea of unfolding and constructs an iterative-based network model. Furthermore,

we unfolded every FISTA iteration phase into an update block. In order to improve the performance of the network and get the reconstruction result directly, transform basis Ψ module and its inverse form Ψ^{-1} added for each iteration block instead of simply attaching module Ψ to the end of the whole network. This work appears to be redundant now, but in the following sections, its elegance will be gradually revealed. The unfolding structure of CSR-Net is shown in Fig. 1, where Δx_k is the difference between k th and $(k-1)$ th iteration result. The proposed CSR-Net is composed of several iteration blocks, for each iteration block, four main update steps are performed. We take block k as an example.

1) *Shrinkage Coefficient Update*: This step can be divided into two more detailed steps. First, t_{k+1} can be updated as (7) according to historical shrinkage pseudo coefficient t_k , then the shrinkage coefficient c_k can be calculated of the form

$$c_k = \frac{t_k - 1}{t_{k+1}}. \quad (8)$$

2) *Shrinkage Terms Update*: In this step, iteration block need input consisting of x_{k-1} , x_{k-2} , and shrinkage coefficient c_k . Thus, shrinkage terms can be updated according to the following formula:

$$\omega_k = c_k \Delta x_{k-1} = c_k (x_{k-1} - x_{k-2}). \quad (9)$$

Essentially, the FISTA algorithm is a kind of gradient descent, and each iteration is a gradient drop of problem (3). As for formula (3), this process can be split into two separate steps as follows.

3) *MMSE Gradient Descent*: Considering the first item of (3), each gradient drop update result \hat{r}_k can be calculated as follows:

$$\begin{aligned} \hat{r}_k &= x_{k-1} - \chi \nabla C(x_{k-1}) \\ &= x_{k-1} - \chi \Phi^T (y - \Phi x_{k-1}). \end{aligned} \quad (10)$$

Then, the shrinkage factor r_k will be updated as

$$r_k = \hat{r}_k + \omega_k. \quad (11)$$

4) *Shrinkage for Sparsity Constraint*: Now we consider the second term of (3), sparsity constraint, it is also termed l_1 -regularization, aims to seek the sparsest solution of s . A simple way to solve this problem is also gradient descent. But we note that $\nabla \|\cdot\|_1 = \text{sgn}(\cdot)$, that is to say l_1 -regularization is not differentiable at 0. Recent years, some researchers found (5) is an effective method to solve the problem, and (5) is a special case of so-called proximal mapping. According to the method, result of block- k can be calculated as follows with the input of r_k and soft threshold θ

$$\begin{aligned} s_k &= \text{sgn}(r_k) \max\{|r_k| - \theta, 0\} \\ x_k &= \Psi s_k. \end{aligned} \quad (12)$$

It is worth noting that steps 1) and 2) are the main differences between ISTA and FISTA. Essentially, a block in CSR-Net represents these four steps above, and it also equals an iteration phase in the conventional FISTA algorithm.

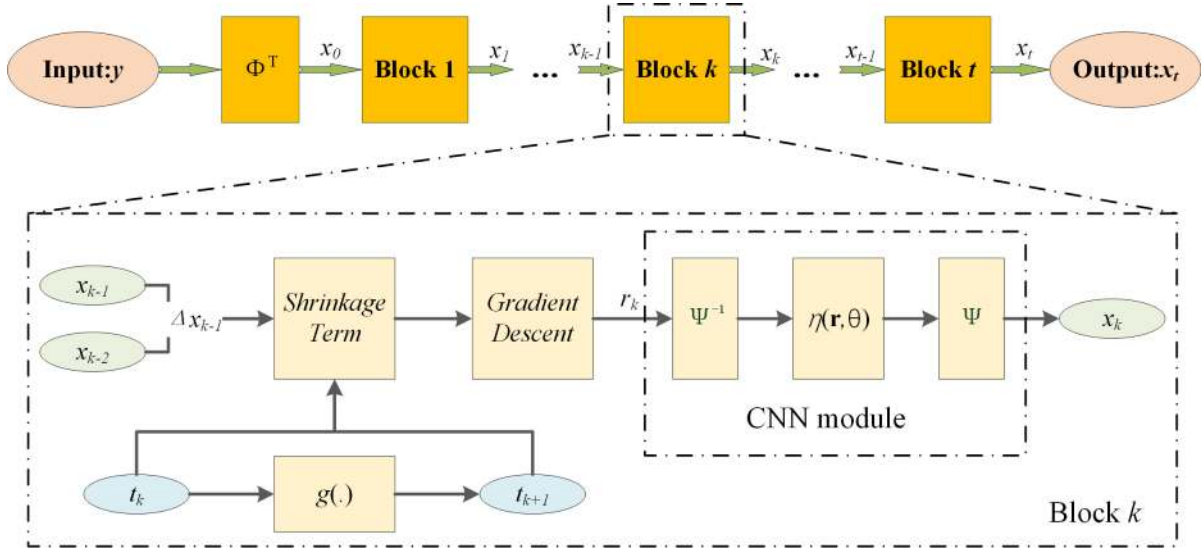


Fig. 1. Unfolded structure of CSR-Net.

B. CNN Module

In order to obtain a most sparse representation of x , and get a state of art reconstruction performance. A key point is to find an optimal sparse basis. Inspired by the work of ISTA-Net-plus [29], we adopt a more elegant way to improve this work for our application.

Considering the uncertainty and substitutability of Ψ , we choose to design a general nonlinear transform function $\Psi(\cdot)$ to sparsify rough estimation of reconstruction signal r_k in the k th block. $\Psi(\cdot)$ can be formulated in matrix form as

$$\Psi(x) = B_3 \times \text{ReLU}(B_2 B_1 \tilde{x}) \quad (13)$$

where $x \in \mathbb{R}^{1 \times n}$ represents the input signal, $\tilde{x} \in \mathbb{R}^{1 \times (n+s-1)}$ is the zero expansion of x , and $\text{ReLU}(x) = \max(0, x)$. Define learnable weights $w_i \in \mathbb{R}^{s \times 1}$, arrange them in subbasis matrices B_i as follows:

$$B_i = \begin{pmatrix} w_{i,1} & 0 & \cdots & 0 \\ w_{i,2} & w_{i,1} & \ddots & \vdots \\ \vdots & w_{i,2} & \ddots & 0 \\ w_{i,s} & \vdots & \ddots & w_{i,1} \\ 0 & w_{i,s} & & w_{i,2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{i,s} \end{pmatrix} \quad (14)$$

where $B_i \in \mathbb{R}^{(n+s-1) \times n}$, $w_{i,j}$ represents the j th element of w_i . Obviously, matrix multiplications of B_i can also be equivalently transformed into the single-channel convolution operations with w_i the kernels. If we redefine $w_i \in \mathbb{R}^{m \times t}$, and arrange them in B_i the same way, matrix multiplication can be transformed into convolution operations with m channels input and t channels output.

Thus, we can adopt a CNN network module to replace $\Psi(\cdot)$, it is a combination of convolution operators, ReLU and batch

normalization (BN). Compared with the work in [29], an extra BN layer is introduced to prevent overfitting and vanishing gradient problem. In more detail, every two convolution layers are separated by a ReLU unit, and each ReLU is preceded by a BN layer. The first convolution layer corresponds to 1 filter in order to adapt single-channel data. And the second convolution layer corresponds to N_{conv} filters of the size $N_f \times N_f$, the next two convolution layers are the same size as the second one. It is of note that $s = N_f \times N_f$ and $t = N_{\text{conv}}$. Furthermore, a symmetric network denoted by $\Psi^{-1}(\cdot)$ with a mirror-symmetrical structure and different weights is constructed as the representation of Ψ^{-1} .

Essentially, step 4) in Section II can be converted into an optimization problem of the form

$$x_k = \arg \min_x \frac{1}{2} \|x - r_k\|_2^2 + \lambda \|\Psi x\|_1. \quad (15)$$

As proved in [29], under a reasonable assumption that each term $x - r_k$ follows an independent normal distribution with common zero mean and variance δ^2 , the following equation is established:

$$\|\Psi^{-1}(x) - \Psi^{-1}(r_k)\|_2^2 \approx \kappa \|x - r_k\|_2^2 \quad (16)$$

where κ is a coefficient related to convolutional weights. Thus, with the proximal mapping theory in (5), (15) can be approximated to another form with the definition of $\Psi(\cdot)$ and $\Psi^{-1}(\cdot)$

$$x_k = \Psi(\eta(\Psi^{-1}(r_k), \theta)). \quad (17)$$

The CNN module in k th block of CSR-Net is shown in Fig. 2, it is inspired by (17), the network achieved a combination of simplicity and high performance. In another perspective, this is also why we choose to lead in extra Ψ and Ψ^{-1} modules in each iteration block.

As for loss function, we adopted the similar solution in [29] to consider the constraint of $\Psi(\cdot) \times \Psi^{-1}(\cdot) = I$. Where I is the identity operator. Both reconstruction error and sparsity

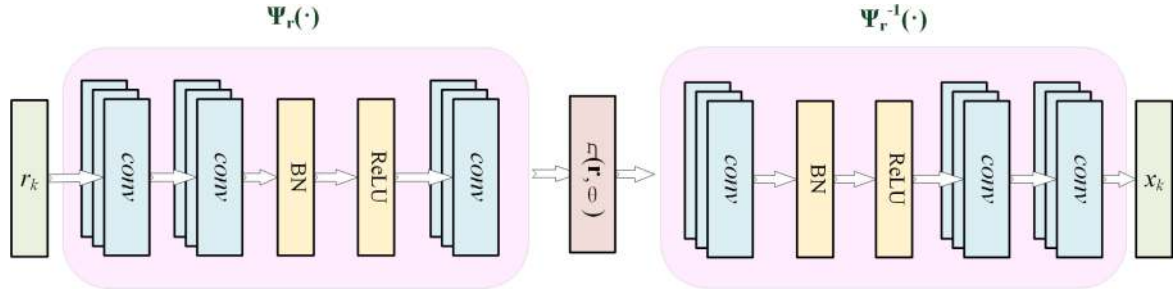


Fig. 2. CNN module in k th block.

constraint were considered, and adjustable parameters were adopted to control the tradeoff between two constrained terms. With the training data $\{x_i, y_i\}_{i=1}^{N_{\text{train}}}$ and learnable parameters $\Theta = \{\chi, \theta, \Psi, \Psi^{-1}\}$, the global loss function $l_g(\Theta)$ is defined as

$$l_g(\Theta) = \frac{\beta_1}{N_{\text{set}}} \sum_{i=1}^{N_{\text{set}}} \|\hat{x} - x_i\|_2^2 + \frac{\beta_2}{N_{\text{set}}} \sum_{i=1}^{N_{\text{set}}} \sum_{k=1}^t \|\Psi^{-1}(\Psi(r_k)) - r_k\|_2^2 \quad (18)$$

where t is the number of iteration blocks, N_{set} is the number of training samples. And β_1, β_2 are the weight parameters for the two different constrained factors, in our experiments they are set to 1 and 0.1, respectively.

C. Complex-Valued Network

Taking the basic operators for complex-valued data into consideration, especially for multiplication operators, it can be converted into a matrix operators based on the real and imaginary parts respectively, i.e., for complex multiplication $c = a \times b$, it can be expressed as

$$\begin{pmatrix} \text{Re}(c) \\ \text{Im}(c) \end{pmatrix} = \begin{pmatrix} \text{Re}(a) & -\text{Im}(a) \\ \text{Im}(a) & \text{Re}(a) \end{pmatrix} \begin{pmatrix} \text{Re}(b) \\ \text{Im}(b) \end{pmatrix} \quad (19)$$

here, $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ represent real and imaginary operators, respectively. Inspired by (19), we establish the following data structures for echo y and imaging results α , which is simply double the data dimensional size by separating the real and imaginary parts of original data

$$\alpha_k = \begin{pmatrix} \text{Re}(\alpha_k) \\ \text{Im}(\alpha_k) \end{pmatrix} y_k = \begin{pmatrix} \text{Re}(y_k) \\ \text{Im}(y_k) \end{pmatrix}. \quad (20)$$

As for measurement matrix A_{sys} , we adopt the modification as

$$a_{ij} = \begin{pmatrix} \text{Re}(a_{ij}) & -\text{Im}(a_{ij}) \\ \text{Im}(a_{ij}) & \text{Re}(a_{ij}) \end{pmatrix} \quad (21)$$

where a_{ij} is an element of A_{sys} . With the modifications above, we can effectively convert complex-valued problems into real-valued ones. Thus, the CSR-Net can be applied to microwave imaging problems.

IV. 3-D MICROWAVE IMAGING VIA CSR-NET

In most instances, the microwave imaging process can be equivalent to a special case of compressed sensing problem of recovering a complex signal $\alpha \in \mathbb{C}^N$ dubbed target scene scattering coefficients form a complex measurement $y \in \mathbb{C}^M$ dubbed echo. The echo data can be expressed as

$$y = A_{\text{sys}}\alpha + e \quad (22)$$

where $A_{\text{sys}} \in \mathbb{C}^{M \times N}$ is determined by the imaging system. To our best knowledge, CS-based imaging has been widely studied, but a majority of researches focused on 2-D imaging or conventional iteration algorithms. And in this section, an effective and efficient method for 3-D microwave imaging based on the proposed CSR-Net will be introduced.

A. Correction Scheme for 3-D Microwave Imaging

For 3-D microwave imaging via CSR-Net, the key point is to build an appropriate measurement matrix for the imaging system. As for a single imaging scene slice, considering an imaging scene with $N_s \times N_s$ resolution units, and we assume N_a antenna elements randomly distributed. We consider the round-trip phase difference $\Delta\varphi_{i,j}$ between the imaging resolution unit P_s^i and the antenna array element P_a^j , it can be calculated as

$$\Delta\varphi_{i,j} = 4\pi \frac{\|P_s^i - P_a^j\|}{\lambda} = \frac{4\pi f}{c} \|P_s^i - P_a^j\| \quad (23)$$

where f is the carrier frequency, c is the velocity of light, $\lambda = \frac{c}{f}$ is the wavelength, P_s^i represents the coordinate vector of imaging resolution unit i , and P_a^j represents the coordinate vector of antenna array element j . Thanks to the exponential form of imaging system echo, we can get a $\Delta\varphi_{i,j}$ based measurement matrix $A_{\text{sys}} \in \mathbb{C}^{N_a \times N_s^2}$ of the form

$$A_{\text{sys}} = \begin{pmatrix} e^{-j\Delta\varphi_{1,1}} & \dots & e^{-j\Delta\varphi_{1,N_s^2}} \\ \vdots & \ddots & \vdots \\ e^{-j\Delta\varphi_{N_a,1}} & \dots & e^{-j\Delta\varphi_{N_a,N_s^2}} \end{pmatrix}. \quad (24)$$

It is worth noting that A_{sys} is determined by the imaging system and geometric parameters. And in some other application scenarios A_{sys} can be constructed in a different way, i.e., FFT. We assume target scattering coefficient is $\alpha \in \mathbb{C}^{N_s^2 \times 1}$, and its element defined as α_i . Thus, echo $y \in \mathbb{C}^{N_a \times 1}$ can be expressed

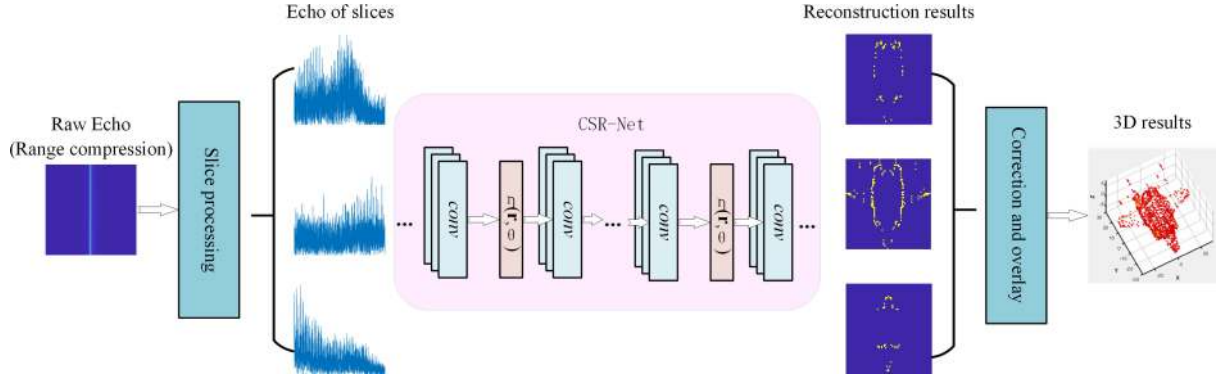


Fig. 3. Schematic of 3-D imaging via CSR-Net.

as

$$y = A_{\text{sys}}\alpha. \quad (25)$$

In pursuit of simplicity, 3-D imaging scenes were divided into several slices in our scheme, therefore, we can perform sparse reconstruction imaging via CSR-Net on each slice separately, and then correct the imaging results and superimpose them to obtain 3-D imaging results. The schematic is shown in Fig. 3. Based on the scheme above, for a 3-D imaging scene divided into N_{slice} slices, we assume that target scattering coefficient for slice k is $\alpha_k \in \mathbb{C}^{N_s^2 \times 1}$, and echo for slice k is $y_k \in \mathbb{C}^{N_a \times 1}$, where $k = 1, 2, \dots, N_{\text{slice}}$. Thus, in a whole 3-D scene, target scattering coefficient $\alpha_{3D} \in \mathbb{C}^{N_s^2 \times N_{\text{slice}}}$ and echo $y_{3D} \in \mathbb{C}^{N_a \times N_{\text{slice}}}$ can be extended to the following forms respectively:

$$\begin{aligned} \alpha_{3D} &= (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_{N_{\text{slice}}}) \\ y_{3D} &= (y_1 \ y_2 \ \cdots \ y_{N_{\text{slice}}}) \\ &= A_{\text{sys}} (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_{N_{\text{slice}}}) \\ &= A_{\text{sys}}\alpha_{3D}. \end{aligned} \quad (26)$$

For microwave imaging, the problem is to reconstruct α_{3D} from a known A_{sys} and y_{3D} , that can be performed by the proposed CSR-Net efficiently.

Noting (26), in practical 3-D imaging cases, A_{sys} usually vary from slice to slice, and network parameters vary with A_{sys} . So a constant measurement matrix A_{sys}^0 must be found as a reference, we assume that the measurement matrix for slice k is A_{sys}^k . Essentially, for the reference measurement matrix A_{sys}^0 , A_{sys}^k can be represented as $A_{\text{sys}}^k \approx \Theta_k A_{\text{sys}}^0$, where Θ_k is a scalar termed phase correction factor. We perform a calibration process on y_{3D} , and get a close form of (26) with only a constant measurement matrix A_{sys}^0

$$y_{3D}\Theta^{-1} = A_{\text{sys}}^0\alpha_{3D} \quad (27)$$

where Θ^{-1} is the inverse of correction matrix formed by correction factors Θ_k . We propose the following method to get a close form of Θ . As shown in Fig. 4. R_0 is the distance between the array plane and the reference plane, R_k is the distance between the array plane and the k th slice, and the size of resolution unit is ΔP_s , the antenna array element spacing is ΔP_a . Thus, in k th

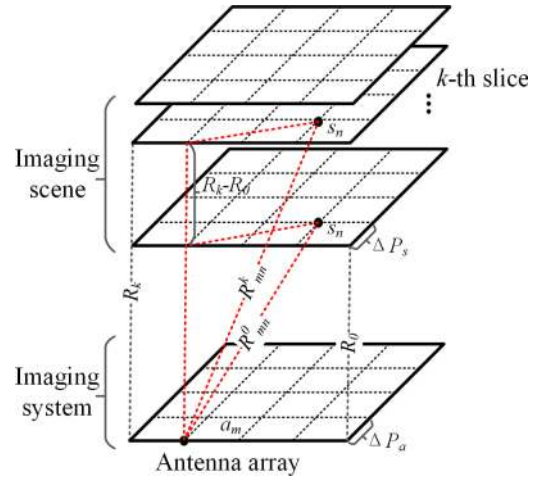


Fig. 4. Schematic of distance difference between slices.

slice, the distance between antenna element a_m and imaging resolution unit s_n can be expressed as

$$R_{mn}^k = \sqrt{R_k^2 + \left(\sqrt{(i_s^2 + j_s^2)}\Delta P_s - \sqrt{(i_a^2 + j_a^2)}\Delta P_a \right)^2} \quad (28)$$

where i_s and j_s are the indexes of s_n , i_a and j_a are the indexes of a_m . For simplicity, in the following description, we define $f_{ij}^a = \sqrt{(i_a^2 + j_a^2)}$ as the distance index of a_m , and the similar f_{ij}^s represents the distance index of s_n . We carry out Taylor expansion of (28) and ignore the higher-order terms, it can be expressed as

$$R_{mn}^k \approx R_k + \frac{(f_{ij}^s \Delta P_s - f_{ij}^a \Delta P_a)^2}{2R_k}. \quad (29)$$

We assume that $R_k \gg \Delta P_s$ and $R_k \gg \Delta P_a$ in microwave imaging problem, this relationship is always be met. (29) can simply converted into $R_{mn}^k \approx R_k$. Therefore, for each resolution unit in slice k , the phase differences $\Delta\varphi_k$ between reference slice

and slice k are of the same form

$$\Delta\varphi_k \approx \frac{4\pi f}{c} (R_k - R_0). \quad (30)$$

Thus, Θ_k can be expressed as

$$\Theta_k = e^{-j\Delta\varphi_k}. \quad (31)$$

And correction matrix Θ can be expressed as

$$\Theta = \text{diag}(\Theta_k) = \begin{pmatrix} \Theta_1 & & & 0 \\ & \Theta_2 & & \\ & & \ddots & \\ 0 & & & \Theta_{N_{\text{slice}}} \end{pmatrix}. \quad (32)$$

Using (27), we can perform a 3-D imaging process with only a constant measurement matrix A_{sys}^0 rather than plenty of them, which is a significant modification for the adoption of proposed CSR-Net in real 3-D imaging application. In this method, the 3-D imaging process can be written as

$$\hat{\alpha}_{3D} = f(y_{3D}\Theta^{-1}, A_{\text{sys}}^0) \quad (33)$$

where $\hat{\alpha}$ represents imaging result, and $f(\cdot)$ is imaging process.

B. Training Scheme for CSR-Net

Another novelty for our work is that we propose a simple but efficient method dubbed PAR to train CSR-Net networks for microwave imaging application scenarios. In seek for accurate results, conventional deep networks always should be trained with quite a large number of training samples that are usually hard to get, especially, in the microwave imaging application scenarios. But for the microwave imaging based on CS theory, we found that the imaging process is mainly affected by the pseudo inverse of A_{sys} . So we attempted to create a training pattern related to A_{sys} only, and it is not important whether there are enough real-measured data samples.

Aiming at generating a data set $\{x_i, y_i\}_{i=1}^{N_{\text{train}}}$ matching specific imaging tasks, a key problem is how to determine the range of label amplitude. In our work, the MF reconstruction results are taken as a reference, and the amplitudes of label vectors $x_i, i = 1, 2, \dots, N_{\text{train}}$ are in the same order of magnitude and randomly distributed. Moreover, label vectors are of the same size as imaging resolution plane $N_s \times N_s$, and none zero values of them are distributed in random places. According to the imaging system and geometric model parameters, the measurement matrix A_{sys} can be calculated by (24). Thus, the measurement for each constructed label vector can be generated as $y_i = A_{\text{sys}}x_i$. Under these assumptions, an arbitrarily sized training set can be constructed. And it is worth noting that the number of target N_{target}^i of training data $\{x_i, y_i\}$ can be calculated as

$$N_{\text{target}}^i = \|x_i\|_1. \quad (34)$$

According to the theory of compressed sensing [7], the following restrictions for N_{target}^i need to be met:

$$1 \leq N_{\text{target}}^i \leq \frac{N_a}{2}. \quad (35)$$

We construct a training data sample with a target number that meets the above constraint, moreover, the number of training

targets is evenly distributed in the restricted interval. Therefore, in the manner described above, we can generate any amount of data needed to train CSR-Net.

V. EXPERIMENTAL RESULTS

In this article, we analyze the performance of the proposed CSR-Net in a numerical test environment first. Then, the proposed 3-D imaging method is verified in terms of performance and speed. And in both cases, our method performs better over conventional and network-based methods.

In our experiments, we trained all networks for different test conditions with the following fixed mutual hyperparameters: convolution size of 32, filter size of 3×3 , initial learning rate of 0.0001, and batch size of 8. And we set χ and θ to 0.0001 and 0.001, respectively. It is of note that χ and θ are learnable. All of our training and testing experiments were implemented in the Tensorflow framework with the Adam optimizer and performed on a platform with 3.70 GHz Intel Core i7-8700 K CPU (64 G RAM) and RTX 2060 GPU (6 G Memory).

For the quantitative evaluation of reconstruction performance, we adopt two indexes. First, the normalized mean square error (NMSE) [31] value was used which is defined as

$$\text{NMSE} = \frac{\|\hat{\alpha} - \alpha_{\text{label}}\|_2^2}{\|\alpha_{\text{label}}\|_2^2} \quad (36)$$

where $\hat{\alpha}$ and α_{label} denote the reconstructed images and ground truth, respectively. We also use the target-to-background ratio (TBR) [32], which is defined by

$$\text{TBR} = 10 \lg \left[\frac{\sum_{(i,j) \in R_T} |\hat{\alpha}_{ij}|^2}{\sum_{(i,j) \in R_B} |\hat{\alpha}_{ij}|^2} \right] \quad (37)$$

where R_T and R_B are the set of target points and the set of background points, respectively. To effectively evaluate the imaging performance of measured data, we use the image entropy (ENT), which is defined as

$$\text{ENT} = - \sum_{g(i)} p_i \log p_i \quad (38)$$

where p_i represents the proportion of pixels with amplitude i in the image, $g(i)$ is total gray value of image histogram. It reflects the average amount of information in the image, the smaller the ENT value, the clearer the image.

A. Performance of CSR-Net

We generate $\{x_i, y_i\}_{i=1}^{N_{\text{train}}}$ for training, and the elements of matrices x_i are randomly distributed, we adopt 0.5-1 to represent the scattering coefficients with different strengths, and a zero value means no target. For learnable parameters $\Theta = \{\chi, \theta, \Psi, \Psi^{-1}\}$, in this article, we adopt untied training mode to both CSR-Net and its comparative experiment ISTA-NET-plus, in this mode, in which learnable parameters can be represented as $\Theta = \{\chi_k, \theta_k, \Psi_k, \Psi_k^{-1}\}_{k=1}^{t-1}$.

First, four different networks are designed to verify the impact of BN layers. We adopt the method described in Section IV-B to generate a training set with 10 000 samples, 20 dB additive white

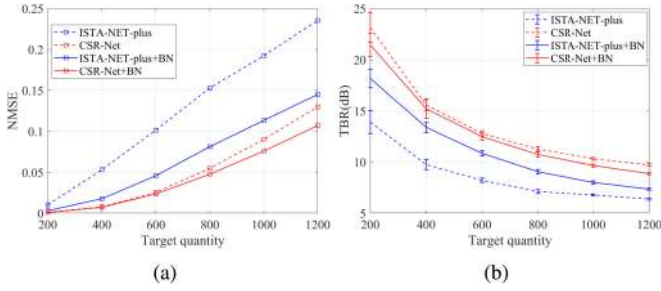


Fig. 5. Performance investigation on BN layers. (a) NMSE. (b) TBR.

Gaussian noise (AWGN) is added to generated clean measurements. For each sample in the set, the number of targets sets to be a random number, and it meets the constraint in (35), all networks use the same training set. Furthermore, as a default setting, all networks are fixed in 9 blocks, and we set the number of resolution units as 64×64 , the range of scenarios as $60 \text{ m} \times 60 \text{ m}$, the platform height R_0 equals 3000 m which is much larger than the size of the resolution unit. Thus, the measurement matrix $A_{\text{sys}} \in \mathbb{C}^{N_a \times N_s^2}$ can be calculated of the form in (24). Furthermore, we generate a testing set consist of 50 different samples in each number of targets, and the number of targets is in range of $\{200, 400, 600, 800, 1000, 1200\}$. If not specified, 20 dB AWGN noise is added to all test sets in this section by default. Fig. 5(a) shows networks with BN layers get lower NMSE values, and that is more obvious for ISTA-NET-plus. Fig. 5(b) shows CSR-Net with BN layers perform slightly worse than the original one in TBR, but ISTA-NET-plus improves quit a lot. It suggests that for CSR-Net, the BN layer can improve the reconstruction accuracy to a certain extent; but it reduces the TBR value. However, for ISTA-NET-plus, the introduction of the BN layer can greatly improve the reconstruction accuracy, and at the same time increase the TBR value. Nevertheless, after the introduction of the BN layer, CSR-Net still performs better than ISTA-NET-plus in two indicators. So in what follows, BN layers are introduced by default.

1) *Target Quantity*: In this section, we study the effect of target density on reconstruction performance. CS ratio is fixed to 50%. For a fair comparison, the same training set is used, We adopt testing sets with target density as sparse ($0.1 \times N_{\text{total}} = 135$), half ($0.5 \times N_{\text{total}} = 675$) and full ($N_{\text{total}} = 1351$), where $N_{\text{total}} = 1351$ is the total number of target in an imaging scenario. Results are shown in Fig. 6.

In order to investigate the performance of target density more detailedly, we compare the performance of CSR-Net with not only network-based ISTA-NET-plus but also conventional Matched Filtering (MF, baseline) ISTA, FISTA algorithms. To be more precise, the same test set described above adopted. We analyze NMSE and TBR performance of all methods mentioned above, Fig. 7 shows NMSE and TBR performance versus target quantity. It demonstrates that CSR-Net obtains compelling reconstruction results compared to the MF baseline, it also significantly outperforming other algorithms involved whether in a scene with sparse or dense targets.

2) *Number of Blocks*: Then, we analyze the effect of the maximum number of blocks k on reconstruction performance.

We train our networks with a maximum number of blocks in a range of $\{1, 3, 6, 9, 12, 15\}$, and the same training process is implemented in ISTA-NET-plus. A testing set consists of 50 samples applied, in which each sample is randomly distributed with 500 targets. For a fair comparison, the testing set is also applied to test conventional ISTA and FISTA algorithms with the same number of iterations. Taking MF as baseline, Fig. 8 shows curves of reconstruction performance under different blocks/iterations. Experimental results show that NMSE of reconstructed results is quite worse in a small number of blocks, at this time, we can get quite a considerable drop of NMSE with only a small increase in the number of blocks. Inversely, when the number of blocks is quite large, only little performance improvement can be made by adding extra blocks. However, TBR shows the opposite trend. In Fig. 8(a), the NMSE values of proposed CSR-Net are smaller than ISTA-NET-plus, it indicates that our method outperforms ISTA-NET-plus in reconstruction accuracy, and it also exceeds conventional methods quite large margins. Furthermore, Fig. 8(b) shows that TBR obtained by CSR-Net is higher than the other four methods. It suggests that our method has good reconstruction accuracy. In more detail, we find that both NMSE and TBR curves are tend to be stable when the number of blocks exceeds nine. So, in the following experiments, nine blocks are adopted by default.

3) *SNR*: To further investigate the robustness of CSR-Net, AWGN noise of different intensities is added to the testing set. SNRs of testing sets are in the range of $\{0, 5, 10, 15, 20, 25, 30\}$ dB. Both networks and conventional algorithms are composed of nine blocks/iterations, and networks are trained for 50 epochs with the same training set. Fig. 9 demonstrates performance under different SNRs. We note that Fig. 9 shows the proposed method outperforms ISTA-NET-plus and conventional algorithms in both NMSE and TBR. This indicates that no matter what noise conditions, CSR-Net has higher reconstruction accuracy. Furthermore, the trend of curves shows that all methods is fairly robust to noise except MF performs worse in high noise level.

B. Investigation on 3-D Imaging

We adopt the proposed CSR-Net in 3-D imaging simulation experiments using the method described in Section IV-A. The number of resolution units is fixed as 64×64 , and the platform height R_0 is 3000 m that much higher than the size of resolution units, more detailed parameters of the simulation system is shown in Table I. In what follows, we analyze the imaging performance with 9 blocks/iterations and adopt 32 slices in height by default.

1) *Superresolution Test*: In the simulation environment, we set pulse repetition rate (PRF) to 420 Hz, and platform velocity v_s is set to be 100 m/s. With parameters in Table I, the azimuth and range resolutions can be calculated as follows:

$$R_c = \frac{\lambda H}{2L_a} = 1 \text{ m}$$

$$R_a = \frac{\lambda H}{2W_a} = \frac{\lambda H}{2 \times (S_a - 1)v_s/\text{PRF}} = 1 \text{ m} \quad (39)$$

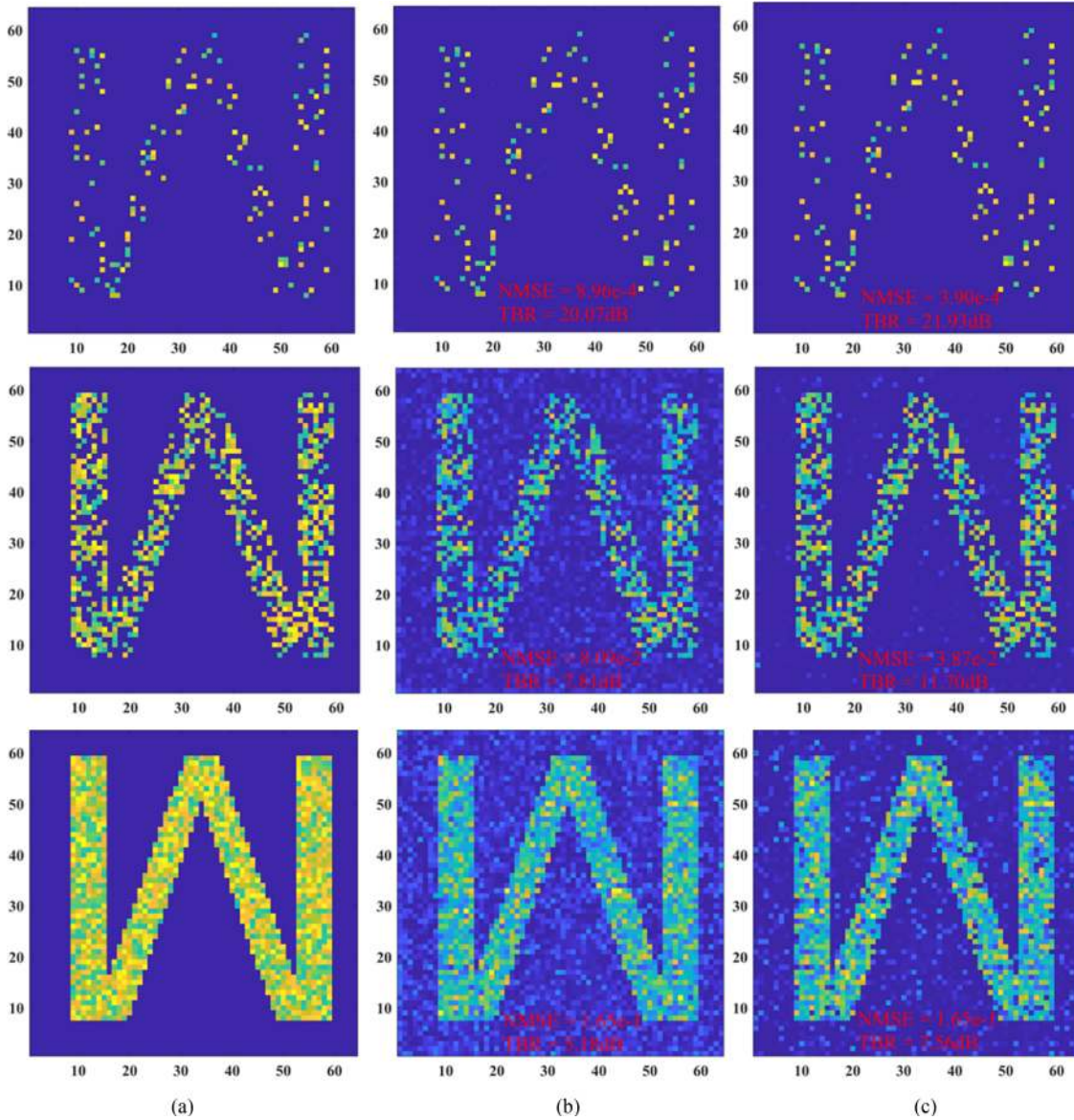


Fig. 6. Reconstruction results of CSR-Net and ISTA-NET-plus. Row 1 represents sparse ($0.1 \times N_{\text{total}}$) case. Row 2 are half ($0.5 \times N_{\text{total}}$). Row 3 full (N_{total}). (a) Ground truth. (b) ISTA-NET-plus. (c) CSR-Net.

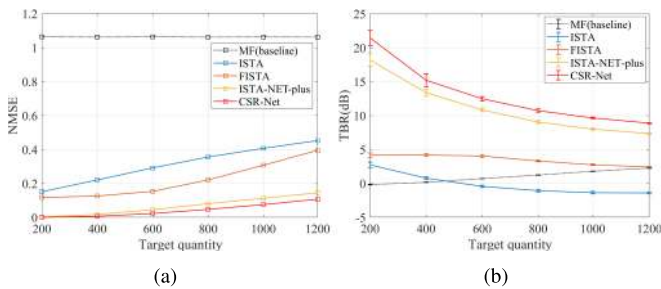


Fig. 7. Performance under different target quantity. (a) NMSE. (b) TBR.

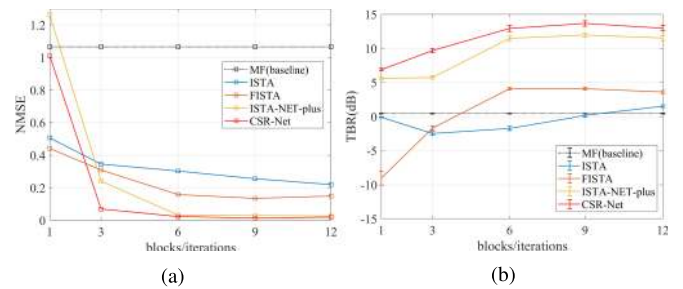


Fig. 8. Performance under different blocks/iterations. (a) NMSE. (b) TBR.

where $\lambda = \frac{c}{f_c}$ is the wavelength, c is speed of light, and W_a is Azimuth sampling length, it can be calculated as $W_a =$

$(S_a - 1) \frac{v_a}{\text{PRF}}$. In order to investigate whether CSR-Net can realize superresolution reconstruction, we set the reconstruction resolution in range of $0.3R_{\text{sys}} - 0.9R_{\text{sys}}$ by adjusting image

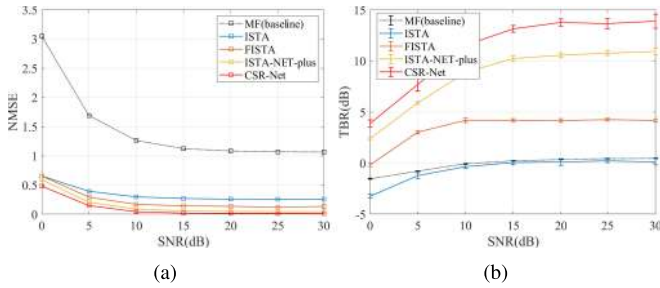


Fig. 9. Performance under different SNRs. (a) NMSE. (b) TBR.

TABLE I
PARAMETERS OF SIMULATION DATA

Parameters	Symbol	Value
Center frequency	f_c	30GHz
Bandwidth	B	150MHz
Platform height	H	3000m
Linear array elements	N_a	64
Azimuth sampling points	S_a	64
LAA length	L_a	15m

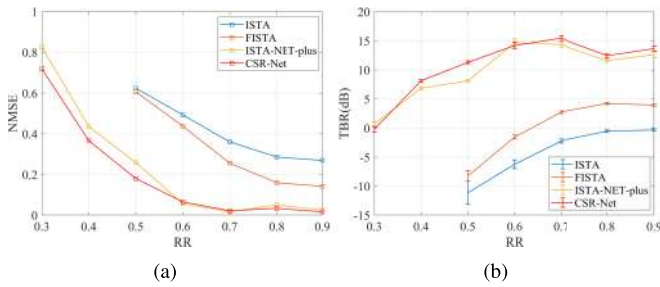


Fig. 10. Imaging performance under different RR. (a) NMSE. (b) TBR.

range and fixing the number of resolution units, here in $R_{\text{sys}} = R_c = R_a = 1$ m. And we fix the CS ratio as 50% which is related to the number of antenna elements activated and resolution units. It is worth noting that the measurement matrix varies with the number of resolution units and the size of the scene range. So we generate training data sets with 10 000 samples for each different reconstruction resolution, and 20 dB AWGN noise added to all training sets.

Fig. 10 shows imaging performance under different resolution ratio (RR). Fig. 10(a) shows that CSR-Net performs the lowest NMSE than other methods in all range of resolution interval. And it can be noted that networks based methods exceed conventional ones by large margins, especially in a low RR value. A similar phenomenon also presented in the curves of TBR versus RR. In addition, both network-based methods mentioned in this article still perform well when the resolution reaches twice the theoretical value of the system ($0.5R_{\text{sys}}$), which means both CSR-Net and ISTA-NET-plus can realize double superresolution reconstruction. It is of note that the performance of conventional ISTA and FISTA deteriorate significantly when RR in range of 0.3–0.5, which even leads to reconstruction failure. So the performance of them in this range is not reflected in Fig. 10. In addition, an

F35 fighter model is used to perform a 3-D imaging experiment, Fig. 11 shows some 3-D imaging results via different methods mentioned in this article when $RR = 0.5$.

2) *CS Ratio*: At last, we investigate the effect on the CS ratio of 3-D imaging performance. CS ratios are changed in range $\{0.1, 0.3, 0.5, 0.7, 0.9\}$. For the microwave imaging application scenario, it can be achieved by adjusting the number of antenna elements or sampling the received signal. In this article, we adopt the former method, and schematic diagrams of antenna array plane in different CS ratio are shown in Fig. 12. 3-D imaging results via CSR-Net and ISTA-NET-plus are shown in Fig. 13. It suggests that CSR-Net obtains better results in 3-D imaging applications regardless of CS ratio. And Fig. 14 show the performance under different CS ratio. Fig. 14(a) shows that network-based methods have a worse reconstruction accuracy than conventional algorithms in a low CS ratio, but all of them do not perform well. When the CS ratio exceeds 0.2, network-based methods outperform conventional ones, especially, the proposed CSR-Net shows a fairly low reconstruction error than other methods. In addition, all methods perform a considerable low reconstruction error with a CS ratio near 1. Fig. 14(b) shows the proposed CSR-Net exceeds other methods by quite large margins in TBR. That suggests it has a better target resolution.

3) *Speed Investigation*: In this section, the speed performance of 3-D imaging process via different methods are investigated. Noting that conventional CS algorithms often suffer from high data-dependence and low parallelism, which is not suitable for GPU acceleration, so they are only tested on CPU. While network-based methods are examined on both CPU and GPU platforms. Leveraging the fact that the data feeding process costs a considerable amount of time and this process can be easily optimized by adopting an asynchronous FIFO in actual application, thus, only reconstruction (forward inference) process is counted in imaging speed test for network-based methods. And reconstruction speed is evaluated in two different conditions, the same blocks/iterations and $NMSE < 0.001$. We generate a 3-D imaging testing set consist of 20 slices using the PAR method. Runtime(s) of all methods in two different conditions are shown in Tables II and III, respectively. As can be seen in Table II, networks with the same blocks show a similar speed, and both of them exceed conventional algorithms by more than two orders of magnitude. Moreover, Table III means that the proposed CSR-Net achieves a higher speed in the same MSE level.

C. Verification on Measured Data

Two experimental ground-based linear array SAR (GB-LASAR) systems are adopted to verify the performance of CSR-Net. A nine blocks CSR-Net network is trained for imaging, furthermore, MF, OMP, ISTA, and FISTA algorithms are used for performance comparison in these experiments.

1) *Experiment 1*: Imaging experiment is adopted using measured data obtained from ground-based equivalent LASAR experimental system 1. System parameters are listed in Table IV, and the experimental scenario and system are shown in Fig. 15 top row. As shown in Table IV, the number of virtual 2-D array

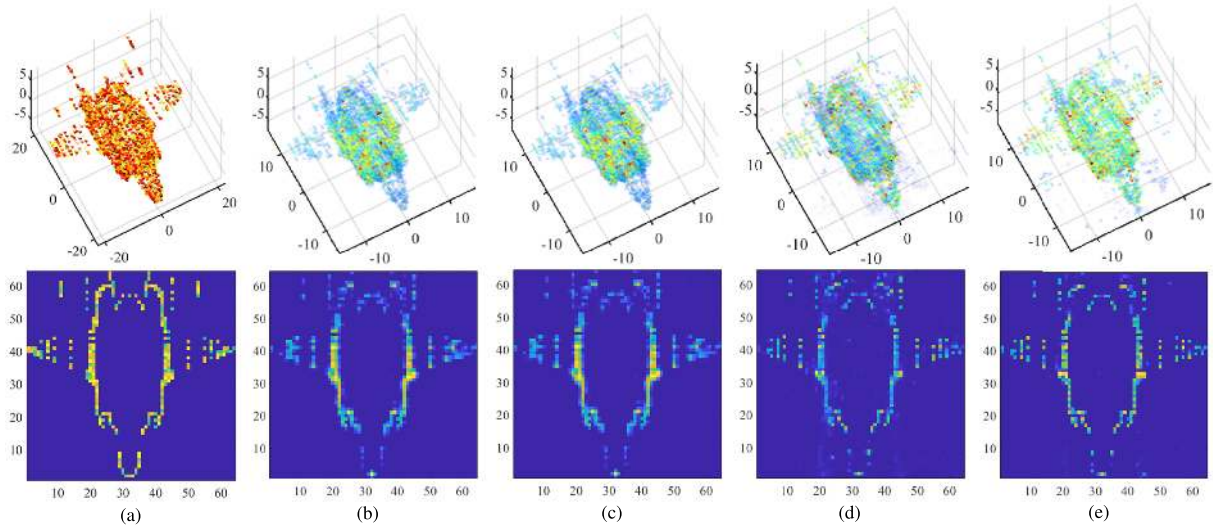


Fig. 11. 3-D imaging results via different methods when resolution ratio $RR = 0.5$, row 1 represent 3-D results, and row 2 are slices of them. (a) Ground truth. (b) ISTA. (c) FISTA. (d) ISTA-NET-plus. (e) CSR-Net.

TABLE II
RUNTIME (S) OF DIFFERENT METHODS WHEN BLOCKS/ITERATIONS ARE 9 (CPU/GPU)

	ISTA	FISTA	ISTA-NET-plus	CSR-Net
A slice	11.489/-	11.462/-	0.594/0.047	0.584/0.048
3-D result	229.780/-	229.240/-	11.880/0.940	11.680/0.960

TABLE III
RUNTIME (S) OF DIFFERENT METHODS WHEN $NMSE < 0.001$ (CPU/GPU)

	ISTA	FISTA	ISTA-NET-plus (15 blocks)	CSR-Net (12 blocks)
A slice	12.313/-	11.697/-	0.946/0.077	0.751/0.059
3-D result	246.260/-	233.940/-	18.920/1.540	15.020/1.180

TABLE IV
PARAMETERS OF EXPERIMENTAL SYSTEM

Parameters	System1	System2	mmW platform
Center frequency	9 GHz	9.62 GHz	77GHz
Bandwidth	2 GHz	80MHz	3.55GHz
APC(full)	8394	8928	41107
Distance	5 m	100m	280mm
LAA length	1.4 m	1.25m	-

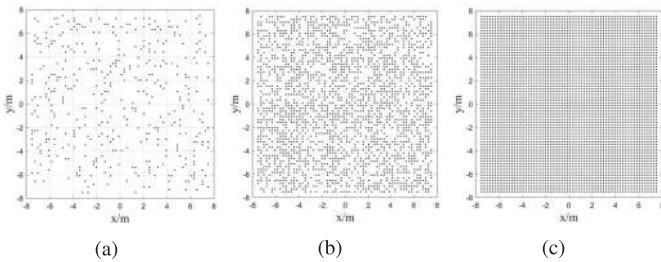


Fig. 12. Antenna array elements distribution diagram. (a) Spare, CS ratio = 0.1. (b) Half, CS ratio = 0.5. (c) Full, CS ratio = 1.

elements is 8394, from which we randomly sample 4000 and 2000 elements, that is to say, the CS ratio is fixed as 47.7% and 23.8%, respectively. During system movement, the size of the virtual antenna array formed is 1.5×1.3 m. As a full sampling baseline shown in Fig. 16, Fig. 17 shows imaging results obtained via MF, OMP, ISTA, FISTA, and CSR-Net with 4000 and 2000 elements.

2) *Experiment 2*: A different experiment system is adopted to verify the universality of CSR-Net. System parameters are listed in Table IV, and the experimental scenario and system are shown in Fig. 15 bottom row. In this system,

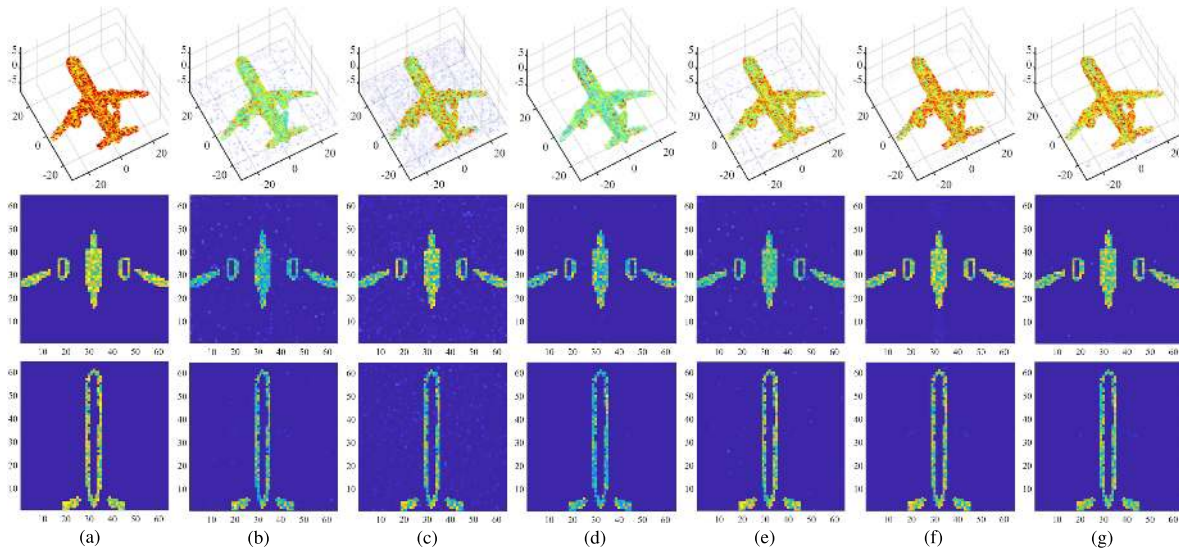


Fig. 13. 3-D imaging results via CSR-Net and ISTA-NET-plus in different CS ratio, row 1 represent 3-D results, and row 2 and 3 are slices of them. (a) Ground truth. (b)–(d) ISTA-NET-plus. (e)–(g) CSR-Net. (b,e) CS ratio 0.3. (c,f) CS ratio 0.5. (d,g) CS ratio 0.7.

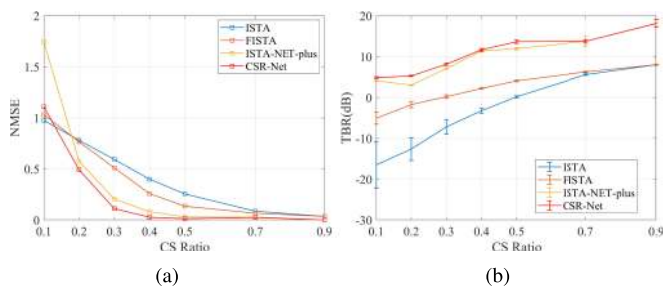


Fig. 14. Imaging performance under different CS ratio. (a) NMSE. (b) TBR.

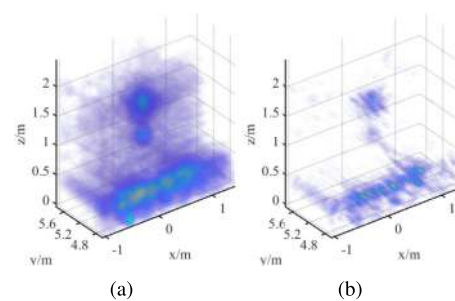


Fig. 16. Full sample (8394 elements) 3-D imaging via MF and OMP used for comparison. (a) MF. (b) OMP.



Fig. 15. Experimental systems and scenarios, the top and bottom rows correspond to experiment 1 and 2, respectively. (a) Experimental systems. (b) Experimental scenarios.

the number of virtual 2-D array elements is 8928, and we also randomly sample 4000 and 2000 elements, the CS ratio is fixed as 44.8% and 22.4%, respectively. As a baseline, Fig. 18 shows results via MF and OMP in the full sampling

case. Fig. 19 shows 3-D imaging results in 4000 and 2000 elements.

3) *Performance and Speed Analysis*: Compared with baseline Fig. 16, Fig. 17 shows our method still perform well at low CS ratios, and it can be seen in Fig. 17, results of CSR-Net get fewer false target points than that of ISTA-NET-plus and conventional algorithms in 4000 elements sampling, but networks based methods show a significant performance degradation at low CS ratio (2000 elements sampling), especially ISTA-NET-plus. Even so, CSR-Net can still achieve reconstruction accuracy comparable to traditional algorithms, this also conforms to the law presented in Fig. 14 in Section V-B. Table V numerically shows the ENT of the experiments, we note that CSR-Net gets lower ENT values, it suggests that CSR-Net achieves a better performance. Similarly, the conclusion can also be drawn from experiment 2.

In addition, Table VI shows reconstruction time of all methods. Since MF adopts formula $\hat{x} = A_{\text{sys}}^T y$ directly to obtain imaging results, its reconstruction speed is fairly fast by just simply using one matrix multiplication. As for other algorithms, the proposed CSR-Net is noticeably better than ISTA, FISTA,

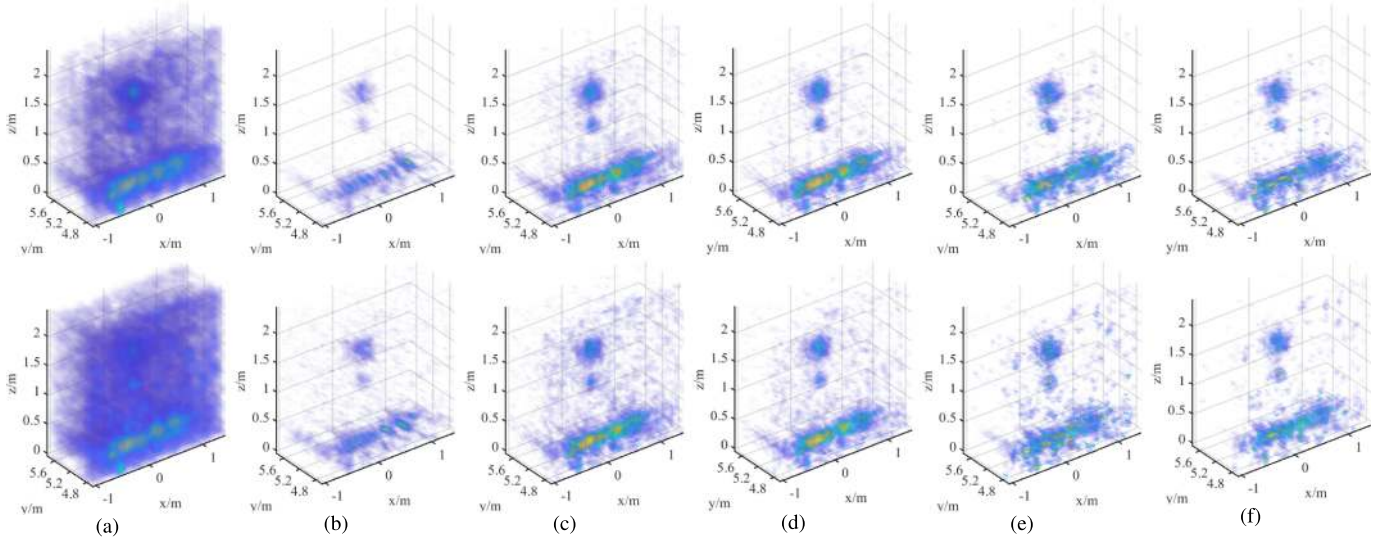


Fig. 17. Experiment 1, double balls 3-D imaging results, the top and bottom rows represent 4000 and 2000 elements, respectively. (a) MF. (b) OMP. (c) ISTA. (d) FISTA. (e) ISTA-NET-plus. (f) CSR-Net.

TABLE V
ENT VALUES OF DIFFERENT METHODS IN REAL 3-D IMAGING APPLICATIONS

	Sampling (elements)	MF	OMP	ISTA	FISTA	ISTA-NET-plus	CSR-Net
Exp.1 (16 slices)	2000	1.4210	0.6857	0.5434	0.4362	0.4586	0.3745
	4000	1.2700	0.5314	0.5411	0.3487	0.3892	0.3453
Exp.2 (21 slices)	2000	1.4444	0.5381	0.2941	0.2826	0.2872	0.2762
	4000	1.3367	0.3274	0.2909	0.2618	0.2595	0.1885

TABLE VI
RUNTIME (S) OF DIFFERENT METHODS IN REAL 3-D IMAGING APPLICATIONS(CPU/GPU)

	Sampling (elements)	MF	OMP	ISTA	FISTA	ISTA-NET-plus	CSR-Net
Exp.1 (16 slices)	2000	0.890/-	67.900/-	188.696/-	182.154/-	10.472/0.625	10.649/0.627
	4000	1.622/-	209.130/-	432.424/-	428.300/-	10.338/0.654	10.356/0.652
Exp.2 (21 slices)	2000	0.925/-	197.235/-	294.886/-	253.721/-	11.899/0.703	12.154/0.687
	4000	1.833/-	397.510/-	572.053/-	567.591/-	11.860/0.719	11.586/0.741

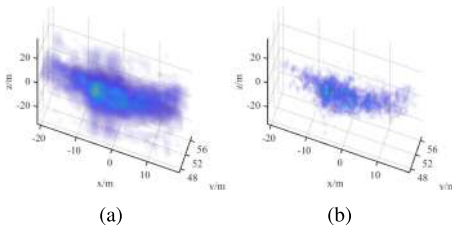


Fig. 18. Full sample (8928 elements) 3-D imaging via MF and OMP used for comparison. (a) MF. (b) OMP.

and OMP in reconstruction speed. Since the same block number adopted in both CSR-Net and ISTA-NET-plus, they have similar reconstruction time.

D. Discussion

1) *Generalization Ability*: Considering that the simulation and real-measured experiments of this article are based on the far-field LASAR imaging platforms. In order to demonstrate the generality of the method, this section applies the proposed method to the recently released “TWR1443” MIMO millimeter-wave (MMW) sensors real-measured data [33] for near-field 3-D reconstruction. The significant radar system parameters are listed in Table IV. The target scenarios are shown in Fig. 20. After the measurement matrix is constructed according to the imaging geometric parameters, the MF scattering amplitude of scenario 1 is estimated. Referring to the estimated amplitude, 10 000 training samples are generated by PAR method, CS ratio is fixed to 0.1216 (5000 elements of 41107), and the number of resolution units is 64×64 . Set network parameters to default values (described at the beginning of this article), we

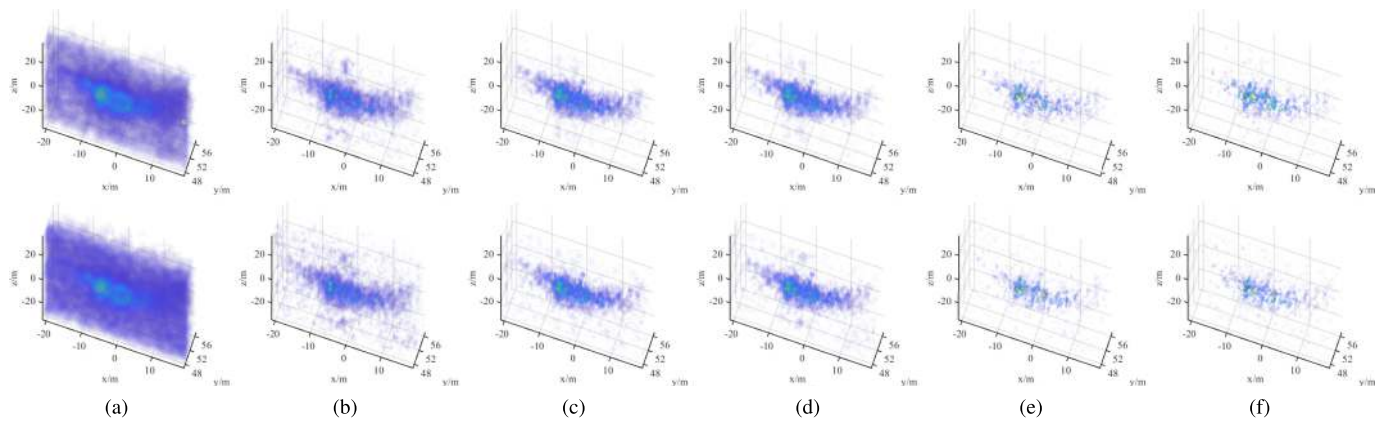


Fig. 19. Experiment 2, curved wall 3-D imaging results, the top and bottom rows represent 4000 and 2000 elements, respectively. (a) MF. (b) OMP. (c) ISTA. (d) FISTA. (e) ISTA-NET-plus. (f) CSR-Net.

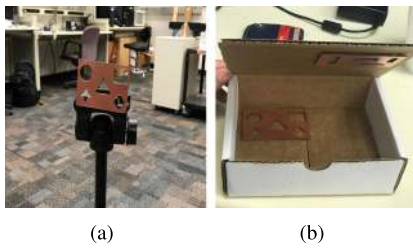


Fig. 20. MMW experimental scenarios. (a) Scenario 1, simple target. (b) Scenario 2, hidden targets.

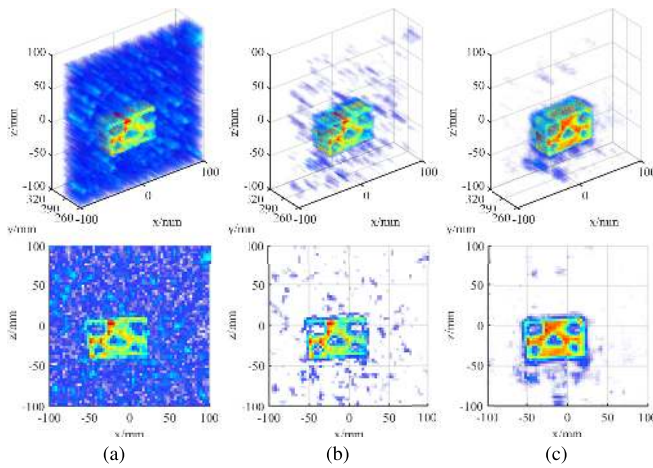


Fig. 21. 3-D MMW reconstruction results of scenario 1 via different methods, row 1 represent 3-D results, and row 2 are front views of them. (a) MF. (b) FISTA. (c) CSR-Net.

trained a CSR-net with nine blocks. For the sake of fairness, the parameters of the conventional FISTA algorithm are set to be the same with CSR-NET initialization parameters. Taking MF as the baseline, Fig. 21 shows 3-D imaging results of the proposed method and the conventional FISTA algorithm. In addition, to verify the adaptability of the proposed method to different target scenes, the well-trained network was directly applied to scenario 2. The imaging results are shown in Fig. 22.

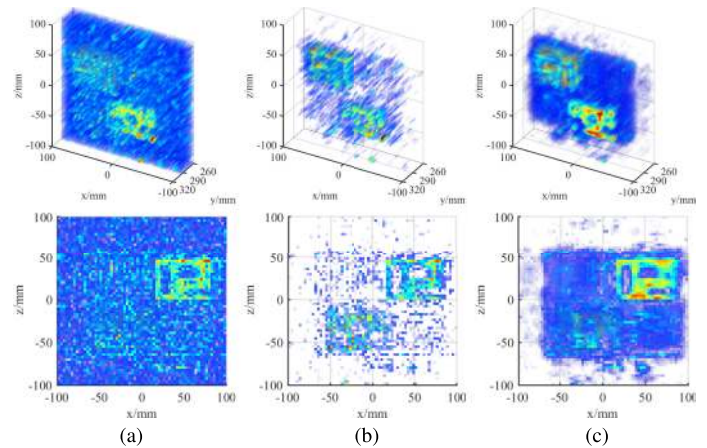


Fig. 22. 3-D MMW reconstruction results of scenario 2 via different methods, row 1 represent 3D results, and row 2 are front views of them. (a) MF. (b) FISTA. (c) CSR-Net.

In Fig. 21, it can be observed that the FISTA and MF will bring many artifacts and grating-lobes due to the lack of samples under 5000 sampling elements, but CSR-Net obtained compelling reconstruction results. Note that Scenario 2 is the imaging scene of the hidden object in the box, and the low-value region in the figure represents the weak scattering region dominated by the box. Similarly, imaging results in Fig. 22 demonstrate the MF contains a significant amount of artifacts, FISTA loses a large number of weak scattering targets, while CSR-NET achieved more accurate reconstruction results. Therefore, the proposed method is suitable not only for far-field LASAR 3-D imaging but also for near-field MIMO MMW 3-D reconstruction, the CSR-Net trained in Scenario 1 still achieves competitive reconstruction performance on Scenario 2, showing its good generalization ability.

2) *Characteristic Analysis:* Compared with the traditional algorithm, the parameters of CSR-NET are learnable. In the training process of the network, CSR-NET learns the optimal parameters automatically, while the traditional algorithm needs to manually debug these parameters, which requires a lot of time.

In addition, CSR-NET adopts CNN to learn an optimal nonlinear sparse representation process instead of the hand-crafted transform basis in conventional algorithms. Therefore, CSR-NET has a significant performance improvement and maintains an efficient reconstruction speed. And CSR-NET also has a better adaptability performance to weakly sparse scenes. Benefiting from the feedforward data-path of CSR-Net, we can accelerate it using GPU. But due to iterative nature of conventional algorithms (OMP, ISTA, FISTA), they often suffer from high data-dependence and low parallelism, which is not suitable for GPU acceleration. This gives CSR-Net a speed gain of about two orders of magnitude over conventional algorithms. However, because the proposed method converts complex operations into matrix operations and vectorizes 2-D scene slices, it requires more memory than traditional algorithms, thus the imaging task of large scenes is difficult to be realized directly. In future work, we will exploit the effective method to apply CSR-NET to large scene imaging tasks.

VI. CONCLUSION

In this article, we propose a novel 3-D microwave sparse reconstruction method based on a complex-valued sparse reconstruction network (CSR-Net). First, aiming at getting a better convergence performance, a shrinkage term derived from FISTA is introduced to CSR-Net. In addition, CSR-Net adopts a convolutional neural module to learn an optimal nonlinear sparsifying transform, which dramatically reduces computational complexity and improves reconstruction performance. Furthermore, in order to ensure the generality of the measurement matrix to slices, a 3-D microwave imaging correction scheme is introduced to fix the phase difference of measurement data of each slice. Considering the absence of measured data, the PAR method is designed to train our network effectively. Finally, experiments carried out on simulated and real-measured data demonstrate the advanced accuracy and efficiency of the proposed method, compared with the network-based ISTA-NET-plus, conventional MF, OMP, ISTA, and FISTA algorithms. It is of note that the proposed method is extremely fast in reconstruction, especially on GPU, we can envisage the proposed method to be implemented in a real-time microwave imaging application.

Note that the proposed method converts complex number operations to matrix operations, this will take up a too large amount of memory resources to achieve large scene reconstruction. A multichannel processing scheme may have great potential for this work and deserves further attention.

REFERENCES

- [1] X. Peng, W. Tan, W. Hong, C. Jiang, Q. Bao, and Y. Wang, "Airborne DL-SLA 3-D SAR image reconstruction by combination of polar formatting and l_1 regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 213–226, Jan. 2016.
- [2] W. Qiu, J. Zhou, H. Zhao, and Q. Fu, "Three-dimensional sparse turntable microwave imaging based on compressive sensing," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 4, pp. 826–830, Apr. 2014.
- [3] S. A. Rezaeieh, A. Zamani, and A. Abbosh, "3-D wideband antenna for head-imaging system with performance verification in brain tumor detection," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 910–914, 2014.
- [4] B. Tian *et al.*, "A fast sparse recovery algorithm via resolution approximation for laser 3D imaging," *IEEE Access*, vol. 7, pp. 178 710–178 725, 2019.
- [5] L. Li *et al.*, "3D SAR image background separation based on seeded region growing," *IEEE Access*, vol. 7, pp. 179842–179863, 2019.
- [6] L. Pu, X. Zhang, J. Shi, S. Wei, L. Li, and X. Tang, "Three dimensional image-based radar cross section extrapolation via planar projective transforms," *IEEE Access*, vol. 7, pp. 138 990–139 000, 2019.
- [7] N. Simonov, B.-R. Kim, K.-J. Lee, S.-I. Jeon, and S.-H. Son, "Advanced fast 3-D electromagnetic solver for microwave tomography imaging," *IEEE Trans. Med. Imag.*, vol. 36, no. 10, pp. 2160–2170, Oct. 2017.
- [8] B. Tian, S. Wei, L. Dang, M. Yan, and X. Zhang, "SAR 3D sparse imaging based on CLA," *J. Eng.*, vol. 2019, no. 19, pp. 5543–5547, 2019.
- [9] S. Vakalis and J. A. Nanzer, "Microwave imaging using noise signals," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 12, pp. 5842–5851, Dec. 2018.
- [10] A. Bourquard and M. Unser, "Binary compressed imaging," *IEEE Trans. Image Process.*, vol. 22, no. 3, pp. 1042–1055, Mar. 2012.
- [11] S.-G. Miaou, F.-S. Ke, and S.-C. Chen, "A lossless compression method for medical image sequences using JPEG-LS and interframe coding," *IEEE Trans. Inf. Technol. Biomedicine*, vol. 13, no. 5, pp. 818–821, Sep. 2009.
- [12] J. Li, Y. Fu, G. Li, and Z. Liu, "Remote sensing image compression in visible/near-infrared range using heterogeneous compressive sensing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 12, pp. 4932–4938, Dec. 2018.
- [13] A. Movahed, M. C. Reed, N. Aboutorab, and S. E. Tajbakhsh, "Exit chart analysis of turbo compressed sensing using message passing dequantization," *IEEE Trans. Signal Process.*, vol. 64, no. 24, pp. 6600–6612, Dec. 2016.
- [14] H.-C. Huang, F.-C. Chang, T.-K. Huang, and P.-L. Chen, "Error control for compressed sensing transmission with polar codes," in *Proc. IEEE Ist Global Conf. Life Sci. Technol.*, 2019, pp. 160–161.
- [15] T. Küstner *et al.*, "Mr image reconstruction using a combination of compressed sensing and partial Fourier acquisition: Espresso," *IEEE Trans. Med. Imag.*, vol. 35, no. 11, pp. 2447–2458, Nov. 2016.
- [16] S. Rousseau and D. Helbert, "Compressive color pattern detection using partial orthogonal circulant sensing matrix," *IEEE Trans. Image Process.*, vol. 29, pp. 670–678, 2019.
- [17] M. Leinonen, M. Codreanu, and M. Juntti, "Sequential compressed sensing with progressive signal reconstruction in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1622–1635, Mar. 2015.
- [18] Y. Álvarez, Y. Rodríguez-Vaqueiro, B. Gonzalez-Valdes, C. M. Rappaport, F. Las-Heras, and J. A. Martínez-Lorenzo, "Three-dimensional compressed sensing-based millimeter-wave imaging," *IEEE Trans. Antennas Propag.*, vol. 63, no. 12, pp. 5868–5873, Dec. 2015.
- [19] J. Yang, J. Thompson, X. Huang, T. Jin, and Z. Zhou, "Random-frequency SAR imaging based on compressed sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 983–994, Feb. 2013.
- [20] S. Zhang, W. Zhang, Z. Zong, Z. Tian, and T. S. Yeo, "High-resolution bistatic ISAR imaging based on two-dimensional compressed sensing," *IEEE Trans. Antennas Propag.*, vol. 63, no. 5, pp. 2098–2111, May 2015.
- [21] M. Borgerding, P. Schniter, and S. Rangan, "AMP-inspired deep networks for sparse linear inverse problems," *IEEE Trans. Signal Process.*, vol. 65, no. 16, pp. 4293–4308, Aug. 2017.
- [22] A. Chambolle, R. A. De Vore, N.-Y. Lee, and B. J. Lucier, "Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 319–335, Mar. 1998.
- [23] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
- [24] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18 914–18 919, 2009.
- [25] K. Kulkarni, S. Lohit, P. Turaga, R. Kerviche, and A. Ashok, "ReconNet: Non-iterative reconstruction of images from compressively sensed measurements," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 449–458.
- [26] E. Mason, B. Yonel, and B. Yazici, "Deep learning for radar," in *Proc. IEEE Radar Conf.*, 2017, pp. 1703–1708.
- [27] J. R. Hershey, J. L. Roux, and F. Weninger, "Deep unfolding: Model-based inspiration of novel deep architectures," 2014.
- [28] C. Wu, Z. Zhang, L. Chen, and W. Yu, "Super-resolution for MIMO array SAR 3-D imaging based on compressive sensing and deep neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3109–3124, 2020.

- [29] J. Zhang and B. Ghanem, "ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 1828–1837.
- [30] M. A. Davenport and M. B. Wakin, "Analysis of orthogonal matching pursuit using the restricted isometry property," *IEEE Trans. Inform. Theory*, vol. 56, no. 9, pp. 4395–4401, Sep. 2010.
- [31] Y. Han and J. C. Ye, "Framing U-Net via deep convolutional framelets: Application to sparse-view CT," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1418–1429, Jun. 2018.
- [32] L. Zhang, Z.-J. Qiao, M.-D. Xing, J.-L. Sheng, R. Guo, and Z. Bao, "High-resolution ISAR imaging by exploiting sparse apertures," *IEEE Trans. Antennas Propag.*, vol. 60, no. 2, pp. 997–1008, Feb. 2012.
- [33] M. E. Yanik and M. Torlak, "Near-field MIMO-SAR millimeter-wave imaging with sparsely sampled aperture data," *IEEE Access*, vol. 7, pp. 31 801–31 819, 2019.
- [34] S. J. Hamilton and A. Hauptmann, "Deep D-bar: Real-time electrical impedance tomography imaging with deep neural networks," *IEEE Trans. Med. Imag.*, vol. 37, no. 10, pp. 2367–2377, Oct. 2018.
- [35] J.-G. Park and S. Jo, "Bayesian weight decay on bounded approximation for deep convolutional neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2866–2875, Sep. 2019.
- [36] Z. Wei and X. Chen, "Deep-learning schemes for full-wave nonlinear inverse scattering problems," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1849–1860, Apr. 2019.
- [37] J. Fu, J. Dong, and F. Zhao, "A deep learning reconstruction framework for differential phase-contrast computed tomography with incomplete data," *IEEE Trans. Image Process.*, vol. 29, pp. 2190–2202, 2020.
- [38] J. Zhang, D. Zhao, and W. Gao, "Group-based sparse representation for image restoration," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3336–3351, Aug. 2014.
- [39] J. Sun *et al.*, "Deep ADMM-net for compressive sensing MRI," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 10–18.
- [40] Z. Wang, D. Liu, J. Yang, W. Han, and T. Huang, "Deep networks for image super-resolution with sparse prior," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 370–378.
- [41] M. Al-Shoukairi, P. Schniter, and B. D. Rao, "A gamp-based low complexity sparse Bayesian learning algorithm," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 294–308, Jan. 2018.
- [42] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.
- [43] M. T. McCann, M. Nilchian, M. Stampanoni, and M. Unser, "Fast 3D reconstruction method for differential phase contrast X-ray CT," *Opt. Express*, vol. 24, no. 13, pp. 14 564–14 581, 2016.
- [44] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Int. Conf. Med. Image Comput. and Comput.-Assisted Intervention*, Springer, 2015, pp. 234–241.
- [45] K. Kulkarni, S. Lohit, P. Turaga, R. Kerviche, and A. Ashok, "Reconnet: Non-iterative reconstruction of images from compressively sensed measurements," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 449–458.
- [46] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2015.
- [47] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, "k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI," *Magn. Reson. Medicine: An Official J. Int. Soc. Magn. Reson. Medicine*, vol. 61, no. 1, pp. 103–116, 2009.



Shunjun Wei (Member, IEEE) received the B.S., M.Sc., and Ph.D. degrees in electronic engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2006, 2009, and 2013, respectively.

In 2014, he joined UESTC, where he is currently an Associate Professor. His research interests include radar signal processing and SAR systems.



Jun Shi (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electronic engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2002, 2005, and 2009, respectively.

He is currently an Associate Professor and his research interests include radar signal processing and synthetic aperture radar systems.



Yue Wu (Student Member, IEEE) received the B.S. degree from the School of Physics and Engineering, Zhengzhou University, Zhengzhou, China, in 2018, and the M.S. degree from the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China.

Her research interests includes radar signal processing.



Qizhe Qu (Student Member, IEEE) received the B.S. degree from the Information Countermeasure Technology, North University of China, Taiyuan, China, in 2018. He is currently working toward the M.S. degree with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China.

His current research interests include radar signal recognition and machine learning.



Mou Wang (Student Member, IEEE) received the B.S. degree in the communication engineering from the Chongqing University of Posts and Telecommunications, Chongqing, China. He is currently working toward the Ph.D. degree with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China.

His current research interests include compressed sensing, SAR imaging, and machine learning.



Yuanyuan Zhou received the B.S. degree in electronic information science and technology from the Qingdao University of Science and Technology, Qingdao, China, in 2014. He is currently working toward the Ph.D. degree in signal and information processing with the University of Electronic Science and Technology of China, Chengdu, China.

His research interests include deep learning and SAR/remote sensing image processing and classification.



Xiangfeng Zeng received the B.S. degree in the information and communication engineering from the University of Electronic Science and Technology of China, Qingdao, China, in 2019. He is currently working toward the M.S. degree with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China.

His current research interests include SAR ship detection and machine learning.



Bokun Tian received the B.S. degree in electronic engineering, in 2015 from the University of Electronic Science and Technology of China, Chengdu, China, where he is currently working toward the Ph.D. degree in signal and information processing.

His research interests include radar signal processing, SAR systems, and compressed sensing.