

# Super-Resolution for MIMO Array SAR 3-D Imaging Based on Compressive Sensing and Deep Neural Network

Chunxiao Wu<sup>1b</sup>, Zenghui Zhang<sup>1b</sup>, *Member, IEEE*, Longyong Chen<sup>1b</sup>, and Wenxian Yu<sup>1b</sup>

**Abstract**—Multiple-input multiple-output (MIMO) array synthetic aperture radar (SAR) can straightly obtain the 3-D imagery of the illuminated scene with the single-pass flight. Generally, the Rayleigh resolution of the elevation direction is unacceptable due to the length limitation of linear array. The super-resolution imaging algorithms within the compressive sensing (CS) framework have been extensively studied because of the essential spatial sparsity in the elevation direction. However, the super-resolution performance of the existing sparse reconstruction algorithms will deteriorate dramatically in the case of lower signal-to-noise ratio (SNR) level or a few antenna elements. To overcome this problem, a new super-resolution imaging structure based on CS and deep neural network (DNN) for MIMO array SAR is proposed in this article. In this new algorithm, the spatial filtering based on CS is first proposed to reserve the signals only impinging from the prespecified space subregions. Thereafter, a group of parallel end-to-end DNN regression models are designed for mapping the potential sparse recovery mathematical model and further locating the true scatterers in the elevation direction. Finally, extensive simulations and airborne MIMO array SAR experiments are investigated to validate that the proposed method can realize the state-of-the-art super-resolution imaging against other existing related methods.

**Index Terms**—Compressive sensing (CS), deep neural network (DNN), multiple-input multiple-output (MIMO), super-resolution, synthetic aperture radar (SAR), 3-D imaging.

## I. INTRODUCTION

3-D synthetic aperture radar (3-D SAR) imaging has been extensively studied in recent years because of its attractive and distinct advantages in many application fields, e.g., 3-D reconstruction and deformation monitoring of man-made structures [1], [2], forest biomass estimation [3], [4], and glacier ablation analysis [5]. In practice, the representative SAR 3-D imaging systems can be categorized into three types; that is, SAR

tomography (TomoSAR) [6], [7], circular SAR (CSAR) [8]–[11], and linear array SAR [12], [13]. Specifically, TomoSAR synthesizes the aperture in the elevation direction through repeat-pass baselines, and hence, it is applicable to resolve the stacked elevation scatterers corresponding to per azimuth-range cell of the 2-D SAR imagery. Nonetheless, it takes a long period to obtain the data of multiple baselines so that it cannot meet the time-sensitive application. CSAR can theoretically obtain the 3-D SAR image of the observation region by single track, but its height resolution is relatively low. Multiple flights of CSAR can improve the height resolution while at the cost of a lot of time consumption. Multiple-input multiple-output (MIMO) array SAR [14]–[16], as one of the linear array SAR systems, can straightly realize the goal of SAR 3-D imaging with single-pass flight by the use of multiple transmitters and receivers. Meanwhile, it can achieve the high resolution in the elevation direction. Hence, it is already playing an increasingly important role in the SAR 3-D imaging scenario.

Typically, the high-resolution 2-D SAR image on the azimuth-range plane can be achieved by introducing larger bandwidth and synthetic aperture in the range and azimuth directions, respectively. Whereas, for the elevation direction of MIMO array SAR, the Rayleigh resolution, restricted by the length limitation of antenna array, is relatively low. For this reason, the super-resolution algorithm for the elevation direction is desired. In practice, only a few (typically one to three) effective scatterers exist in the elevation direction related to per azimuth-range resolution cell for the 3-D imaging of urban buildings [17]. Therefore, it satisfies the sparse requirements of the compressive sensing (CS) theory regarding the strong spatial sparsity [18]–[21]. Based on this fact, different kinds of sparse reconstruction algorithms, such as discrete grids based  $\ell_q$ -norm ( $0 < q \leq 1$ ) minimization methods [17], [22], maximum *a posteriori* based Bayesian CS [23], and continuous domain based atomic norm minimization approach [24], have been extensively studied for the elevation super-resolution imaging. These algorithms have shown the superior performance involving the super-resolution capability and robustness against the noise interference. Nonetheless, the performance of the abovementioned algorithms regarding the super-resolution ability will deteriorate dramatically in the case of lower signal-to-noise ratio (SNR) or few antenna elements [24], [25]. To solve this problem, the moderate super-resolution imaging algorithm is expected.

Manuscript received February 18, 2020; revised April 12, 2020 and May 13, 2020; accepted June 3, 2020. Date of publication June 9, 2020; date of current version June 18, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant U1830103 and in part by National Pre-Research Foundation of China under Grant 61406190101. (Corresponding author: Zenghui Zhang.)

Chunxiao Wu, Zenghui Zhang, and Wenxian Yu are with the Shanghai Key Laboratory of Intelligent Sensing and Recognition School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: chunxiao.wu@sjtu.edu.cn; zenghui.zhang@sjtu.edu.cn; wxyu@sjtu.edu.cn).

Longyong Chen is with the Science and Technology on Microwave Imaging Laboratory, Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China (e-mail: lychen@mail.ie.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2020.3000760

In recent years, deep learning has been extensively studied in the fields of image processing [26]–[28], speech recognition [29], natural language processing [30], direction-of-arrival estimation [31], [32], etc. The deep neural network (DNN), as one of the most popular neural network models in the deep learning field, has attracted the interest of many researchers due to its outstanding learning and expression ability. It is well-known that DNN with at least one hidden layer can approximate any continuous function on a closed and bounded subset of  $\mathbb{R}^n$  depending on the universal approximation theorem [33], [34]. That is, the DNN is capable of learning and representing many convex or nonconvex and nonlinear mathematical models including the CS problems. Motivated by this fact, deep learning-based signal recovery methods have been extensively studied in recent years. In 2015, the stacked denoising autoencoder, which is similar to multilayer perceptron with multiple fully connected layers, is first proposed to recover structured signals from their corresponding undersampled measurements [35]. Then, with the development of deep learning, some improved deep learning networks are introduced to learn the inverse transformation in order to replace the conventional convex or greedy recovery algorithms. The representative methods include ReconNet [36], DeepInverse [37], deep residual reconstruction network [38], generative models based compressed sensing [39], etc. The abovementioned networks are all data-driven deep learning methods. Recently, in order to improve the interpretability of deep learning networks, the neural networks, which are only designed to replace some specific resulting procedures of the CS recovery algorithms, have been investigated by the researchers [40]–[42]. For example, Chang *et al.* [41] proposed a general framework that learns a proximal operator using the DNN to solve linear inverse problems. In this article, considering that the sparse recovery in the elevation direction of the MIMO array SAR system is usually solved within the framework of CS, it can be regarded as the convex and nonlinear issue when using the common  $\ell_1$ -norm minimization in practical applications. This implies that the DNN is applicable to the super-resolution imaging scenario of the elevation direction depending on the universal approximation theorem. More recently, in 2019, Budillon *et al.* [43] first attempted to utilize the deep learning method to improve the performance of TomoSAR imaging. Specifically, it aims to use a pretrained TomoSAR neural network, which consists of multiple connected layers, to detect a single scatterer in each azimuth-range resolution cell and predict its elevation. What the authors do not discuss is that it cannot be adapted to the more general and complicated cases, such as multiple scatterers and outliers discrimination. Therefore, the more advanced DNN framework is desired.

In this article, a new super-resolution imaging framework based on CS and DNN for MIMO array SAR is proposed. This approach mainly consists of the following three parts; that is, preliminary recovery with CS, spatial filtering, and a group of parallel DNN regression models that achieve the goal of super-resolution reconstruction. To be specific, first, provided that the scatterers are distributed over a series of uniform but rough grid points in the elevation direction, the scattering intensity can be preliminarily estimated through the  $\ell_1$ -norm minimization

algorithm. Second, the potential scope of the elevation direction is decomposed into  $P$  spatial subregions. In general, for the elevation reconstruction of urban buildings, there mainly exist three kinds of layover areas; that is, the ground, the facade, and the roof. Hence, the parameter  $P$  is set to be 3 in this article. Moreover, for each spatial subregion, the spatial filtering is conducted to guarantee that only the signal impinging from this subregion is reserved based on the previous CS recovery results. Finally, a group of parallel end-to-end DNN regression models are designed for the super-resolution reconstruction of scatterers under each spatial subregion. The main contributions of this article are summarized as follows.

- 1) The CS-based spatial filtering method is first proposed in this article. Compared with the conventional matrix spatial filtering method, besides the outputs can reserve the element-space data property, the advantage of the proposed method is that the phase information of the steering vectors related to the positions of scatterers is well retained even in the case of few antenna elements. In addition, this procedure can help to reduce the distribution divergences of scatterers in the elevation direction, which is beneficial to the DNN model with less training dataset and neurons.
- 2) The end-to-end DNN regression model is first introduced for the super-resolution imaging in the elevation direction of the MIMO array SAR system. The theoretical analysis indicates the fact that the super-resolution performance of the DNN regression model is better than other conventional sparse recovery methods. Also, the experimental results demonstrate that the proposed method can realize the state-of-the-art super-resolution imaging against other existing related methods even in the case of lower SNR level or few antenna elements.
- 3) The new imaging signal model is established in this article by eliminating the variable interference of the range factor  $R_0$  in the conventional imaging signal model. The advantage is that the generated training dataset based on the new echo signal model is adapted to the elevation reconstruction of the entire azimuth-range plane.

The remainder of this article is organized as follows. The geometrical and signal models of the MIMO array SAR system are introduced in Section II. Section III presents a new framework involving CS-based spatial filtering and DNN-driven super-resolution reconstruction for MIMO array SAR 3-D imaging. In Section IV, extensive simulations and airborne MIMO array SAR real experimental results are investigated to demonstrate the predominance of the proposed method. Finally, this article concludes with a brief summary in Section V.

## II. MIMO ARRAY SAR SYSTEM

In this section, we will briefly review the geometrical and signal models of the MIMO array SAR system.

### A. Geometrical Model

The imaging geometrical model of MIMO array SAR is depicted in Fig. 1, where the  $x$ -,  $y$ -, and  $z$ -axis stand for the azimuth

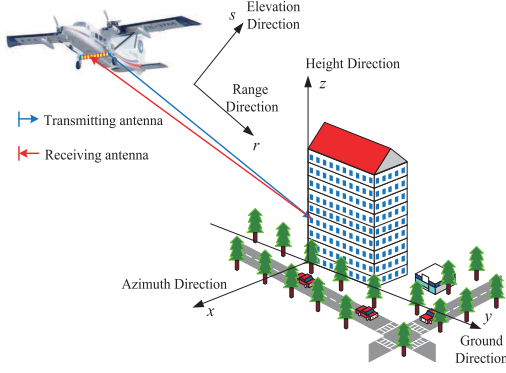


Fig. 1. MIMO array SAR imaging geometrical model.

direction, the ground range direction, and the height direction, respectively. The plane flies along the azimuth direction with a velocity  $v$  at the height  $H$ . The array antenna composed of  $N_t$  transmitters and  $N_r$  receivers is mounted under the belly of the plane as shown in Fig. 1. According to the principle of equivalent phase center [44], [45], the MIMO array antenna can be equivalent to a uniformly distributed virtual linear array composed of  $N = N_t \times N_r$  virtual elements with the constant interval. In practical applications, the phase compensation caused by the prior far-field hypothesis and the time division working mode should be implemented in advance before the imaging process [23], [46]. As a result, besides the number of real antenna elements can be reduced significantly, the echoes of the MIMO array are rearranged and equivalent to the uniformly distributed linear array with the self-transmitting and receiving working model.

### B. Signal Model

In general, the linear frequency modulation signal emitted by the  $n$ th ( $n = 0, 1, \dots, N-1$ ) virtual antenna is given by

$$s(t) = \text{rect}\left(\frac{t}{T_p}\right) \exp\left(j2\pi f_c t + j\pi k_r t^2\right) \quad (1)$$

where  $t$ ,  $T_p$ ,  $f_c$ , and  $k_r$  denote the fast time, the pulse width, the carrier frequency, and the chirp rate, respectively. Based on the point scattering model, the corresponding echo signal after removing the carrier frequency can be explicitly written as

$$\hat{s}(x_m, t, y_n) = \sum_{q=1}^Q \gamma_q \text{rect}\left(\frac{t - 2R_q/c}{T_p}\right) \cdot \exp\left[j\pi k_r \left(t - \frac{2R_q}{c}\right)^2 - j\frac{4\pi}{\lambda} R_q\right] \quad (2)$$

where  $\gamma_q$ ,  $\lambda$ , and  $R_q$  stand for the scattering intensity of the point scatterer  $q$  ( $q = 1, 2, \dots, Q$ ), the wavelength, and the instantaneous range between the  $n$ th antenna with position  $(x_m, y_n, H)$  and the point scatterer  $q$ , respectively. Generally, with the preprocessing operations of the range compression and the azimuth compression, the high-resolution 2-D SAR imagery can be obtained. Furthermore, with the preliminary deramping process in the elevation direction, the echo signal model in (2)

can be simplified and rewritten more compactly as [22]

$$\hat{s}(y_n) = \sum_{q=1}^Q \gamma_q \exp\left(j\frac{4\pi}{\lambda R_0} y_n y_q\right) \quad (3)$$

where  $R_0$  and  $y_q$  represent the projection of the range on the zero-Doppler plane and the position of the point scatterer  $q$  along the elevation direction, respectively. Now, suppose that the potential scatterers along the elevation direction are assumed to lie on a set of discrete grid points with the location  $y_k$  ( $k = 1, 2, \dots, K$ ). Then, in the presence of noise, the signal model in (3) can be rewritten in a matrix form of

$$\mathbf{s} = \mathbf{A}\boldsymbol{\gamma} + \mathbf{n} \quad (4)$$

where  $\mathbf{s} \in \mathbb{C}^{N \times 1}$  denotes the received echo signal vector along the elevation direction corresponding to the same azimuth-range resolution cell,  $\boldsymbol{\gamma} \in \mathbb{C}^{K \times 1}$  stands for the potential scattering information vector to be recovered, and  $\mathbf{A} \in \mathbb{C}^{N \times K}$  is the measurement matrix with the entry  $a_{nk} = \exp(j4\pi y_n y_k / \lambda R_0)$ . In practice, for per fixed azimuth-range resolution cell, there only exist a few point-like scatterers along the elevation direction. That is, the vector  $\boldsymbol{\gamma}$  presents strong spatial sparsity in which it satisfies the requirement of the CS theory very well [18]–[21]. Hence, the signal model in (4) can be solved within the framework of CS and the sparse signal vector  $\boldsymbol{\gamma}$  can be recovered with high probability. Typically, according to the  $\ell_1$ -norm minimization principle, the resulting mathematical model can be formulated as

$$\hat{\boldsymbol{\gamma}} = \arg \min_{\boldsymbol{\gamma}} \left\{ \frac{1}{2} \|\mathbf{s} - \mathbf{A}\boldsymbol{\gamma}\|_2^2 + \lambda \|\boldsymbol{\gamma}\|_1 \right\} \quad (5)$$

where  $\lambda$  is the regularization parameter that controls the balance between the model error and the signal sparsity. Subsequently, some postprocessing operations, e.g., model selection [25] and the iterative reweighted technique [47], are implemented to suppress the undesired outliers due to the noise interference and locate the true scatterers. Until now, the complete sparse signal recovery flow for the elevation direction based on the CS framework has been presented.

Generally speaking, compared with the conventional spectral estimation strategies including adaptive beamforming (BF) (CAPON) [48] and multiple signal classification [49], the sparse recovery algorithms within the framework of CS can break through the limitation of Rayleigh resolution and have presented the prominent advantage regarding the super-resolution capability. Nonetheless, the sparse recovery performance of the CS algorithms is affected by the restricted isometry property (RIP) of the measurement matrix. The mutual coherence of the measurement matrix as an alternative of the RIP is usually investigated in practical applications. Considering the MIMO array SAR system, the measurement matrix is predetermined by the given antenna positions and the grid points of scatterers in the elevation direction. Based on this fact, when the distance between the two adjacent scatterers is much smaller than the Rayleigh resolution, the mutual coherence of the corresponding two steering vectors of the measurement matrix will be greatly improved. That is, the sparse recovery performance cannot be

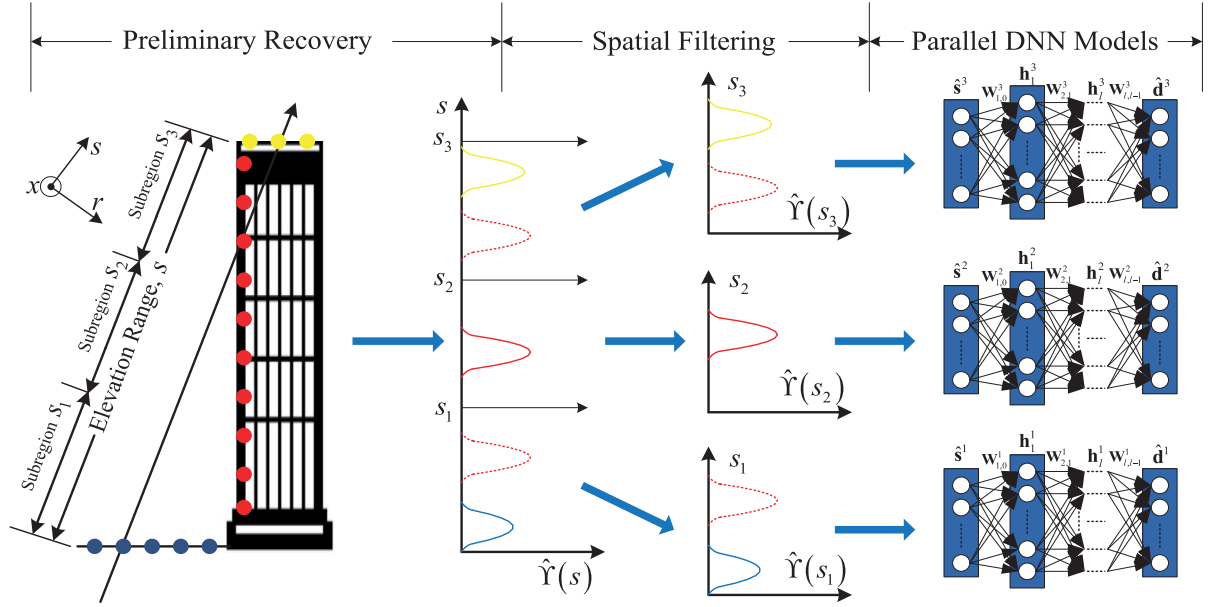


Fig. 2. Proposed a complete super-resolution imaging structure for the MIMO array SAR system.

guaranteed any more. Hence, it is expected that the recovery performance can get rid of the influence of the measurement matrix. In addition, the super-resolution performance of the abovementioned algorithms will also deteriorate dramatically in the case of lower SNR or a few antenna elements. For this reason, the following section is devoted to presenting a robust and advanced super-resolution imaging algorithm for the MIMO array SAR system.

### III. DNN-DRIVEN SUPER-RESOLUTION IMAGING ALGORITHM FOR MIMO ARRAY SAR

In this section, an advanced super-resolution imaging framework based on CS and DNN for MIMO array SAR is proposed. The main idea is that the pretrained DNN model is capable of distinguishing two scatterers with smaller space than the Rayleigh resolution. To specify this, first, the  $\ell_1$ -norm minimization method is utilized to retrieve the scattering information of scatterers distributed over a series of prespecified grid points. Second, the recovery range of the elevation direction is decomposed into three subregions and the spatial filtering is implemented in order to retain the signals only from the specific subspace. Finally, a group of parallel pretrained DNN regression models are utilized to realize the goal of robust super-resolution imaging in the elevation direction.

#### A. Super-Resolution Imaging Framework for MIMO Array SAR

The super-resolution imaging structure for the MIMO array SAR system in the elevation direction mainly consists of three parts; that is, preliminary recovery with CS, spatial filtering, and a group of parallel DNN regression models for the super-resolution recovery. A sketch of the complete super-resolution imaging structure is shown in Fig. 2.

For the TomoSAR imaging within the framework of deep learning, Budillon *et al.* have proposed a multilayer classifier to detect the scatterers in the elevation direction [43]. The authors have demonstrated that the proposed method is effective if there is only one scatterer in the elevation direction corresponding to each azimuth-range resolution cell. However, the proposed DNN classifier model cannot be adapted to the more complicated cases such as multiple scatterers. The reason is that the number of neurons of the DNN classifier will increase dramatically. To be specific, considering the case of a single scatterer and the number of discretized grid points in the elevation direction is set to be  $K$ , the number of neurons in the output layer of the DNN classifier is also  $K$ . However, if there are at most three scatterers along the elevation direction distributed in the ground, the facade, and the roof, respectively, the number of neurons in the output layer will be up to  $C_K^1 + C_K^2 + C_K^3$ . Thus, the network structure of the DNN classifier model will become very large and need to be improved. To overcome this problem, in this article, the DNN regression model is introduced to locate the point-like strong scatterers in the elevation direction. The advantage is that the number of neurons in the output layer is always equal to  $K$  regardless of the number of scatterers. And, there can be multiple effective output neurons depending on the number of potential scatterers in each azimuth-range unit. Despite this, when the SNR level is fixed, the number of types of training data is still equal to  $C_K^1 + C_K^2 + C_K^3$  in the case of at most three scatterers, which will bring large memory consumption and data storage involving the training dataset generation. Motivated by this, the feasible reconstruction range of the elevation direction is first decomposed into three subregions in this article as depicted in Fig. 2. The subregion  $s_1$  mainly contains scattering points from the ground. In addition, it may also have a few scatterers corresponding to the facade of the building. In subregion  $s_2$ , there are scatterers mainly coming from the facade. Considering



the subregion  $s_3$ , besides the scatterers coming from the roof, the scattering points belonging to the facade may also exist. That is, the imaging geometrical model guarantees that there are at most two scatterers in per subregion along the elevation direction for each azimuth-range resolution cell. Then, the spatial filtering is conducted to reserve the signals only impinging from the specific subspace.

For the filtered signal  $\hat{s}_i$  ( $i = 1, 2, 3$ ), a group of parallel DNN regression models are utilized to detect the potential scatterers in each subregion and predict their positions. Importantly, it can be noted that the number of neurons in the output layer for each DNN regression model can be reduced to  $K/3$ . This implies that the number of neurons of the DNN model can be greatly reduced. In addition, the number of types of the training dataset at the same SNR level will decrease to  $C_{K/3}^1 + C_{K/3}^2$  with at most two scatterers per space subregion. It means that the dataset storage and memory consumption for training the DNN model will be improved significantly. The following of this section is devoted to describing the processing steps of the proposed super-resolution imaging structure in detail.

### B. Preliminary Recovery With $\ell_1$ -Norm Minimization

For each azimuth-range resolution cell, the potential scatterers along the elevation direction are assumed to be distributed over a series of prespecified grid points. Importantly, it should be noted that the discrete grid here is relatively rough. Then, based on the established measurement matrix  $\mathbf{A}$ , the sparse signal vector  $\hat{\gamma}$  can be obtained by solving the aforementioned  $\ell_1$ -norm minimization model in (5). In this article, an off-the-shelf solver involving the CVX tool package is utilized to get the solution of the model [50]. As a result, the scattering complex value corresponding to each discrete grid is uniquely determined. That is, the preliminary recovery result along the elevation direction is identified.

### C. CS-Based Spatial Filtering

It is well-known that the purpose of spatial filtering is to retain some signals only impinging from the specific subregion while other signals corresponding to the remaining subspaces are absent. In general, the spatial filter processor mainly consists of two types; that is, BF [51] and the matrix spatial filter [52], [53]. In practical applications, the matrix spatial filter is often used as the prefiltering processor. The reason is that the matrix spatial filter can ensure the outputs reserve the element-space data property, which is significant in some scenarios that need the element-space data as the input including the SAR 3-D imaging within the CS framework. This is a unique advantage not available in the BF method. For the imaging signal model in (4), the output of matrix spatial filter can be formulated as

$$\begin{aligned} \mathbf{y} &= \mathbf{G}^H \mathbf{s} = \mathbf{G}^H \mathbf{A} \boldsymbol{\gamma} + \mathbf{G}^H \mathbf{n} \\ &= \mathbf{C} \boldsymbol{\gamma} + \mathbf{n}_c \end{aligned} \quad (6)$$

where  $\mathbf{G} \in \mathbb{C}^{N \times N}$ ,  $\mathbf{C} \in \mathbb{C}^{N \times K}$ , and  $\mathbf{n}_c \in \mathbb{C}^{N \times 1}$  stand for the matrix filter, the filtered measurement matrix, and the filtered noise term, respectively. Generally, the matrix filter  $\mathbf{G}$  should

satisfy the following constraint given by:

$$\mathbf{G}^H \mathbf{A}(y_k) = \begin{cases} \mathbf{A}(y_k), & y_k \in \mathbf{Y}_p \\ \mathbf{0}, & y_k \in \mathbf{Y}_s \end{cases} \quad (7)$$

where  $\mathbf{Y}_p$  denotes the subspace of interest containing the desired signals, which can be passed by the matrix filter  $\mathbf{G}$  without distortion. In contrast, for other signals located in other sectors  $\mathbf{Y}_s$ , they will no longer exist in the filtered signal  $\mathbf{y}$ . Many different kinds of algorithms have been developed to design the optimal matrix filter including the rank-deficient least-squares method and the second-order cone programming (SOCP) based matrix spatial filter design [54], [55]. Generally, the performance of the aforementioned spatial matrix filter design methods is reliable in the case of sufficient sampling with a large value of  $N$ . In practice, considering the payload of the plane and the cost of hardware equipment, the number of antenna elements is relatively small. Therefore, the performance of the conventional spatial matrix filter design methods cannot be guaranteed any more.

In this article, the CS-based spatial filtering method is first proposed. The output of the proposed algorithm is still the element-space data for the purpose of postprocessing. To be specific, the preliminary recovery result  $\hat{\gamma}$  is first decomposed into three parts; that is,  $\hat{\gamma} = [\hat{\gamma}_{s_1}^T, \hat{\gamma}_{s_2}^T, \hat{\gamma}_{s_3}^T]^T$ . In addition, the measurement matrix  $\mathbf{A}$  is also partitioned as  $\mathbf{A} = [\mathbf{A}_{s_1}, \mathbf{A}_{s_2}, \mathbf{A}_{s_3}]$ , where  $\mathbf{A}_{s_p}$  ( $p = 1, 2, 3$ ) denotes the measurement submatrix composed of steering vectors corresponding to the discrete grid points distributed in the subregion  $s_p$ . According to the theory of CS, the desired signal can be reconstructed with high probability if it is sparse in *a priori* known basis [19]. Hence, it can be believed that the initial recovery signal  $\hat{\gamma}$  is reliable to some extent. Based on this fact, for the  $p$ th ( $p = 1, 2, 3$ ) subregion, the output of spatial filter can be expressed as

$$\mathbf{s}_p = \mathbf{s} - \mathbf{A}_s \hat{\gamma}_s \quad (8)$$

where  $\mathbf{s}_p$ ,  $\mathbf{A}_s$ , and  $\hat{\gamma}_s$  denote the filtered observation vector in which it mainly contains the signal of interest arriving from the passband sector, the measurement matrix of the stopband sector, and the preliminarily recovered signal of the stopband sector based on the  $\ell_1$ -norm minimization, respectively. As mentioned above, the passband sector also consists of three types corresponding to the three subregions of the elevation direction. When the passband sector is determined, the remaining subspaces are regarded as the stopband sector. For example, if the subregion  $s_2$  is regarded as the passband sector, the remaining subregions  $s_1$  and  $s_3$  will be seen as a whole designated as the stopband sector. Furthermore, the parameters  $\mathbf{A}_s$  and  $\hat{\gamma}_s$  can be obtained by  $\mathbf{A}_s = [\mathbf{A}_{s_1}, \mathbf{A}_{s_3}]$  and  $\hat{\gamma}_s = [\hat{\gamma}_{s_1}^T, \hat{\gamma}_{s_3}^T]^T$ , respectively. Finally, the filtered echo signal vector  $\mathbf{s}_p$  only containing the scatterers of the subregion  $s_2$  can be determined by conducting the established model in (8). In summary, based on the preliminary recovery result of the  $\ell_1$ -norm minimization, the echo signal of scatterers located in the passband sector can be accurately determined by subtracting the echo signal corresponding to the stopband sector from the original observation signal. Compared with the spatial matrix filter method,

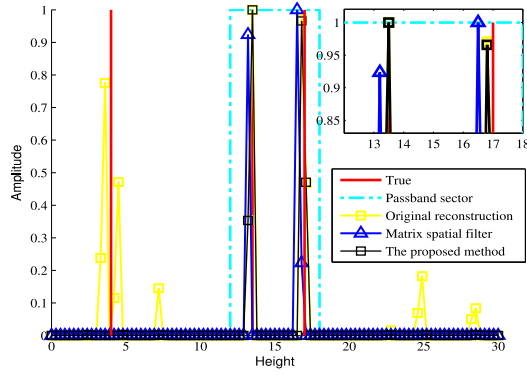


Fig. 3. Spatial filtering based sparse recovery results of the proposed method compared with other related methods.

the disadvantage of the former is that the phase information of the steering vectors belonging to the passband sector can be changed when the dimension of the matrix filter  $\mathbf{G}$  is small depending on the constraint in (7). It is known that the phase information of the steering vectors is related to the positions of scatterers distributed in the elevation direction. That is, the performance of location accuracy of the conventional spatial matrix filter method will degrade drastically in the case of few antenna elements. As for the proposed CS-based spatial filtering method, the phase information related to the positions of scatterers is well retained by avoiding the direct manipulation of the steering vectors. Thus, it can be believed that the location accuracy performance of our method is better than the spatial matrix filter approach when the number of antenna elements is small.

In order to validate the performance of spatial filtering of the proposed method compared with the spatial matrix filter method, the numerical simulation is conducted in this section and the experimental results are depicted in Fig. 3. The array antenna comprises 8 antenna elements and the SNR is set to be 10 dB. For the prefiltering processing, the SOCP-based peak stopband constrained least-squares passband criterion is introduced to implement the matrix spatial filtering [56]. In addition, the  $\ell_1$ -norm minimization is utilized to complete the subsequent sparse recovery of the passband sector using a solver such as the CVX package. The passband range is specified by the cyan dotted line and the other regions are the spatial sector not of interest. One can observe from this figure that there is a large deviation between the predicted locations of the matrix spatial filter based recovered signal and the true locations. That is, the performance of location accuracy will degrade to some extent compared with the straightforward sparse signal reconstruction by the  $\ell_1$ -norm minimization without spatial prefiltering. The main reason is that the dimension of the matrix filter  $\mathbf{G}$  is small such that the constraint property in (7) can no longer be strictly maintained. As we can see, the reconstruction result of the proposed method is approximately in agreement with the true value in the passband sector. Also, for the stopband sector, there is no any interference signal. This implies that the CS-based spatial prefiltering method is reliable even in the case of few antenna elements. Moreover, the output with the element-space

data property is also convenient for the following process, which is presented in the following section.

#### D. DNN-Driven Super-Resolution Imaging

DNN, as one of the representative neural network models, has been extensively explored in many practical scenarios. In accordance with the universal approximation theorem [33], it is well-acknowledged that DNN is able to deal with the convex or nonconvex and nonlinear mathematical models. Additionally, the sparse signal recovery for each subregion can also be solved within the framework of sparse reconstruction such as the  $\ell_1$ -norm minimization model in (5), which is a convex and nonlinear issue. That is, the resulting mathematical model involving the signal reconstruction of each subregion can be regarded as a black box and, hence, an end-to-end learning with DNN can be utilized to achieve the purpose of signal recovery. Specifically, the end-to-end fully connected DNN regression model is introduced in this article mainly due to three reasons. First, considering the received signal  $\mathbf{S} \in \mathbb{C}^{N \times 1}$ , every element, which is corresponding to the signal received by a certain antenna, in the vector  $\mathbf{S}$  is different and valuable. That is, the signal  $\mathbf{S}$  does not have local correlation. Thus, the input vector  $\mathbf{S}$  is fully connected with the hidden layers in the DNN model. Second, the imaging in the elevation direction is an underdetermined problem, which aims to use few measurements  $\mathbf{S} \in \mathbb{C}^{N \times 1}$  to recover signal  $\gamma \in \mathbb{C}^{K \times 1}$  ( $N \ll K$ ). Generally, the dimensionality increase of signal recovery can only be accomplished by the fully connected hidden layers. Third, the interpretability of the end-to-end DNN regression model can be well investigated by analyzing the relationship between the  $\ell_1$ -norm minimization and the proposed DNN structure, which will be presented later in this section.

Generally speaking, for the generation of training dataset, different number of scatterers distributed at different locations along the elevation direction is mainly considered. In addition, for the given MIMO array SAR system, the number of antenna elements is fixed. Thus, considering the echo signal model in (3), the echo signals can be massively simulated with different levels of SNR and random amplitude values of the scatterers. However, it should be noted that the echo signal model in (3) is applicable to the reconstruction of the scatterers corresponding to the resolution cell with the same range  $R_0$ . That is, the generated training dataset is only adapted to the elevation reconstruction of scatterers with the same slant range  $R_0$ . It is unacceptable that the training dataset is generated repeatedly with the change of the slant range  $R_0$  corresponding to different pixels along the range direction. In other words, the variable interference of the slant range factor  $R_0$  in (3) must be eliminated. Motivated by this fact, the new echo signal model and training dataset generation policy are proposed in this article. To specify this, the discrete grid space along the elevation direction is set as

$$\rho_d = \alpha \rho_e = \alpha \frac{\lambda R_0}{2A}, \quad 0 < \alpha < 1 \quad (9)$$

where  $\rho_e = \lambda R_0 / 2A$  denotes the Rayleigh resolution in the elevation direction and  $\alpha$  represents the discrete factor that controls the discrete degree of the grid. Now, for the  $k$ th scatterer,

its position can be given by

$$y_k = m\rho_d, \quad m = 0, 1, \dots, M-1 \quad (10)$$

where  $M$  denotes the total number of grid points that is determined by

$$M = \frac{\rho_{\text{euna}}}{\rho_d} = \frac{\lambda R_0/2\Delta A}{\rho_d} = \frac{A}{\alpha\Delta A} \quad (11)$$

where  $\rho_{\text{euna}} = \lambda R_0/2\Delta A$  stands for the unambiguous range of the elevation direction and  $\Delta A$  is the space of two adjacent antenna elements. By substituting (10) into (3), the new echo signal model can be rewritten as

$$\hat{s}(y_n) = \sum_{m=1}^M \gamma_m \exp\left(\frac{2\pi\alpha}{A} y_n m\right) + \varepsilon. \quad (12)$$

It can be seen that the reconstructed echo signal model involves the variables including the amplitude  $\gamma_m$ , the additive noise  $\varepsilon$ , and the relative position  $m$ . That is, the generated training dataset of the new echo signal model applies to the overall elevation reconstruction of the scatterers with different slant range. Until now, the training dataset generation policy based on the reconstructed echo signal model has been established.

In recent years, with the development of neural network, there exist two kinds of neural network models; that is, real-valued DNN and complex-valued DNN [57]. For the sake of simplicity, the real-valued DNN structure is introduced in this article. Generally, for the spatial filtering vector  $\mathbf{s}_p$  of the  $p$ th subregion, the complex value of each entry is decomposed into the real and imaginary parts given by

$$\mathbf{S}_p = [\text{Re}(\mathbf{s}_p)^T, \text{Im}(\mathbf{s}_p)^T]^T \quad (13)$$

where  $\text{Re}(\cdot)$  and  $\text{Im}(\cdot)$  represent the real and imaginary parts of a complex vector separately. Thereafter, the normalized vector is designed as the input of the DNN

$$\hat{\mathbf{S}}_p = \frac{\mathbf{S}_p}{\|\mathbf{S}_p\|}. \quad (14)$$

For the DNN structure of signal recovery, there are three parallel fully connected neural network models designed for mapping and expressing the essential mathematical model. Considering the DNN model of the  $p$ th subregion, there are totally three hidden layers designed for controlling the tradeoff between the expressive capability and the overfitting risk of the whole network. To be specific, the computation and update of the feed-forward neural network is given by

$$\mathbf{h}_l^p = f_a(\mathbf{W}_{l,l-1}^p \mathbf{h}_{l-1}^p + \mathbf{b}_l^p), \quad l = 1, 2, 3 \quad (15)$$

where  $\mathbf{h}_l^p$  denotes the output of the  $l$ th hidden layer and  $\mathbf{h}_0^p = \hat{\mathbf{S}}_p$ ;  $\mathbf{W}_{l,l-1}$  is the weight matrix between the  $(l-1)$ th layer and the  $l$ th layer;  $\mathbf{b}_l^p$  is the bias vector of the  $l$ th layer; and  $f_a$  stands for the rectified linear unit (ReLU) activation function given by [58]

$$f_a(x) = \text{ReLU}(x) = \max(0, x) = \begin{cases} x, & \text{for } x \geq 0 \\ 0, & \text{otherwise.} \end{cases} \quad (16)$$

The adoption of the ReLU nonlinear activation function can speed up the training process and enhance expressivity of the

DNN model for mapping the potential nonlinear mathematical model. Then, in view of the output layer, the associated truth value is set as

$$d_k^p = \begin{cases} 1, & \text{when the scatterer exists in the position } y_k \\ 0, & \text{otherwise} \end{cases} \quad (17)$$

where 1 and 0 denote the probability that a scatterer exists in the corresponding discrete grid point  $y_k$ . In addition, the identity function is introduced as the activation function for the output layer. So far, the DNN structure has been completely established.

*Relationship to the  $\ell_1$ -norm minimization:* In this article, we attempt to analyze the relationship between the proposed DNN structure and the  $\ell_1$ -norm minimization for the purpose of the interpretability of the DNN model. Generally, the resulting  $\ell_1$ -norm minimization model in (5) can be solved by the iterative soft thresholding (IST) algorithm in which it is composed of the following procedures:

$$\hat{\gamma}_{l+} = \hat{\gamma}_{l-1} - \mathbf{A}^H(\mathbf{A}\hat{\gamma}_{l-1} - \mathbf{s}) \quad (18)$$

$$\hat{\gamma}_l = \text{soft}_\lambda(\hat{\gamma}_{l+}) = \text{sign}(\hat{\gamma}_{l+})\max(|\hat{\gamma}_{l+}| - \lambda, 0) \quad (19)$$

where  $\hat{\gamma}_l$  denotes the recovery result of the  $l$ th iteration and  $\text{soft}_\lambda(\cdot)$  denotes the soft thresholding operation. Furthermore, let  $\mathbf{U}_l = \mathbf{I} - \mathbf{A}^H \mathbf{A}$  and  $\mathbf{b}_l = \mathbf{A}^H \mathbf{s}$ , the iterative resulting expressions in (18) and (19) can be simplified as

$$\hat{\gamma}_l = \text{soft}_\lambda(\mathbf{U}_l \hat{\gamma}_{l-1} + \mathbf{b}_l). \quad (20)$$

Apparently, the simplified resulting model is very similar to the output computation of the  $l$ th layer of the DNN model given in (15). The difference is that the ReLU activation function is introduced for the DNN model and the soft thresholding function is adopted in the IST algorithm. The same thing is that both functions are nonlinear functions. Thus, it can be considered that the  $l$ th layer update of the DNN model is similar to the  $l$ th iteration of the IST algorithm to some extent. Furthermore, the multiple iterations of the IST algorithm are replaced by the multiple connected layers of the DNN model. Importantly, it should be noted that the number of neurons in each hidden layer of the DNN model is usually large. In other words, the dimension of the weight matrix  $\mathbf{W}_{l,l-1}$  of the DNN model is much bigger than the dimension of the weight matrix  $\mathbf{U}_l$  of the IST algorithm, which can enhance the expressive capability of the neural network. Additionally, for the IST algorithm, the weight matrix  $\mathbf{U}_l$  and the bias vector  $\mathbf{b}_l$  are fixed and unchanged during each iteration. In contrast, the weight matrix  $\mathbf{W}_{l,l-1}$  and the bias vector  $\mathbf{b}_l$  are different in each hidden layer of the DNN model. Also, they can be better optimized depending on the training data and further get the best value. Hence, it can be believed that the recovery performance of the DNN model is better than other CS-based methods including the super-resolution ability, which will be demonstrated in the following experimental analysis section.

In the training stage, in order to evaluate the performance of the DNN regression model, the mean square error (MSE) criterion is introduced as the loss function to judge the position prediction accuracy

$$l = \frac{1}{K} \|\mathbf{d}^p - \hat{\mathbf{d}}^p\|_2^2 \quad (21)$$



where  $\hat{\mathbf{d}}^p$  represents the output vector in the output layer of the DNN regression model. It is conducive to the training of the DNN model. The reason is that the gradient of the MSE-based loss function can automatically change with different prediction errors, which is a unique advantage not available for the mean absolute error based loss function. Then, the back-propagation strategy is utilized to optimize the representations of the DNN structure. Specifically, the weight matrices and bias vectors are updated and optimized based on the gradients of the loss function with respect to these parameters as follows:

$$\xi_{\text{new}} = \xi_{\text{old}} - \eta \frac{\partial l}{\partial \xi} \quad (22)$$

where  $\xi$  denotes one of the parameters in the DNN structure;  $\eta$  is the learning rate automatically determined by the adaptive moment estimation (Adam) method [59]. Thus, the DNN regression models can be optimized successfully and are capable of dealing with the signal recovery problem of each subregion.

Finally, after obtaining all the prediction values in the output layers of the three parallel DNN regression models with the input vector  $\hat{\mathbf{S}}_p$ , the output of the whole DNN structure can be determined by concatenating the output of each subregion in order, that is

$$\hat{\mathbf{d}} = [(\hat{\mathbf{d}}^1)^T, (\hat{\mathbf{d}}^2)^T, (\hat{\mathbf{d}}^3)^T]^T. \quad (23)$$

Furthermore, the relative positions of the potential scatterers can be identified corresponding to the grid nodes with positive values approximately equating to 1 in  $\hat{\mathbf{d}}$ . Once the relative position  $m$  is determined, the absolute position of the scatterer is given by

$$y_q = \rho_d m_q, \quad q = 1, 2, \dots, Q. \quad (24)$$

Based on this fact, a much slimmer measurement matrix  $\hat{\mathbf{A}} \in \mathbb{C}^{N \times Q}$  is reconstructed and the final complex-valued reflectivity vector  $\hat{\gamma}$  can be determined as follows:

$$\hat{\gamma} = (\hat{\mathbf{A}}^H \hat{\mathbf{A}})^{-1} \hat{\mathbf{A}}^H \mathbf{s}. \quad (25)$$

As a consequence, the accurate positions and complex-valued scattering information of all the point-like strong scatterers are determined based on the proposed advanced imaging method with the DNN structure.

The abovementioned recapitulation illuminates the fact that by decomposing the reconstruction range of the elevation direction into three subregions and exploiting the spatial prefiltering based on the CS preliminarily recovered results, a group of parallel and feasible DNN regression models are established for the purpose of the sparse signal reconstruction. For the super-resolution performance and computational complexity analysis of the proposed method compared with other related methods, it will be detailedly illustrated in the following section.

#### IV. EXPERIMENT AND RESULT ANALYSIS

In this section, extensive simulations for the MIMO array SAR 3-D imaging are first presented to evaluate the performance of the proposed method. Then, the airborne MIMO array SAR

TABLE I  
PREDICTION ACCURACY FOR DIFFERENT DNN STRUCTURES

Structure	3 layers	4 layers	5 layers	6 layers	7 layers	10 layers
Accuracy	0.3415	0.7623	0.9935	0.9938	0.9938	0.9940

experiment is also conducted to demonstrate the advantages of our approach.

##### A. Experiment Setup

According to the reconstructed echo signal model in (12), the observed signal is related to three parameters; that is, the scattering information  $\gamma$ , the additive noise  $\varepsilon$ , and the relative position  $m$ . Thus, for the DNN-driven sparse reconstruction, the training dataset can be obtained with the change of these three parameters. First, for each subregion in the elevation direction, different number of scatterers distributed in different discrete grid points are considered. The space of the discrete grid points is set to be  $0.1\rho_e$  ( $\rho_e$  denotes the Rayleigh resolution in the elevation direction). Then, once the locations of scatterers are determined, the echo signals can be massively simulated with randomly generated amplitude values between  $[-1, 1]$  with 250 samples and different levels of SNR uniformly distributed between  $[0 \text{ dB}, 20 \text{ dB}]$  with 21 samples. In this article, a total of 5 922 000 echo signal vectors are obtained to generate the training dataset for each subregion, which is divided into two parts with 80% for training and 20% for validation. In the training phase, we have gradually increased the number of layers of the DNN model and the corresponding prediction accuracy on testing dataset is presented in Table I. As we can see, with the increase of the number of layers from three to five, the position prediction accuracy can be significantly improved. Then, when the number of network layers continues to increase, the prediction accuracy is slightly improved, which is negligible to some extent. Simultaneously, more layers of the DNN structure will usually bring more time consumption for training the DNN model. Thus, the DNN structure composed of five layers is designed in this article, which is enough to guarantee that the positions of scatterers in the elevation direction can be accurately predicted in practical applications. Specifically, since the number of antenna elements is 8 and the received signal is decomposed into real and imaginary parts with normalization for input, the dimensionality of the input layer is 16. There are a total of three hidden layers, and the hidden nodes are 256, 512, and 256, respectively. The number of neurons in the output layer depends on the number of discrete grid points. In this article, for the performance verification on real data, the elevation range is uniformly decomposed into three subregions and the number of neurons in the output layer for each DNN regression model is 47. In addition, the Adam optimizer is used and the training process is based on Keras [60]. Finally, we can obtain the pretrained DNN model.

##### B. Performance Analysis of the Resolving Ability

Generally speaking, the high-resolution 2-D SAR image on the azimuth-range plane can be satisfied by introducing larger



TABLE II  
MIMO ARRAY SAR 3-D IMAGING SYSTEM PARAMETERS

Symbol	Parameter	Value
$f_c$	Carrier Frequency	15 GHz
$B$	Transmitting Signal Bandwidth	500 MHz
$V$	Platform Fly Velocity	70 m/s
$PRF$	Pulse Repeat Frequency	1 KHz
$H$	Platform Height	1070 m
$\theta_{abw}$	Azimuth Beamwidth	2°
$\theta_{rbw}$	Range Beamwidth	27°
$\theta_{r,c}$	Center Incidence Angle	25°

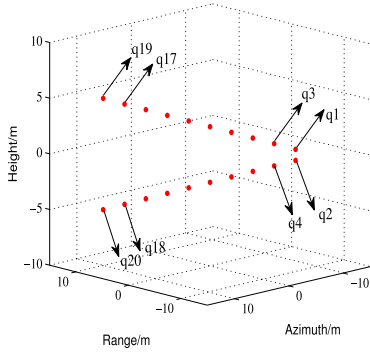


Fig. 4. 3-D distributed scene.

signal bandwidth and synthetic aperture. However, the resolving ability in the elevation direction is usually unacceptable due to the essential length limitation of linear array. Thus, the performance analysis of the super-resolution capability with different methods along the elevation direction is mainly considered here. Specifically, it is based on the BF method, the  $\ell_1$ -norm minimization method, and the proposed DNN-driven imaging algorithm, respectively. The airborne MIMO array SAR system parameters are listed in Table II. It is composed of two transmitting antenna elements and eight receiving antenna elements, which is equivalent to a fully distributed uniform linear array with the number of 16 antenna elements working in the self-transmitter and receiver model. In addition, the baseline length is 1.4 m. The 3-D distributed scene of point scatterers is shown in Fig. 4. In order to focus on the resolving ability in the elevation direction, ten pairs of scatterers with different elevation space changing from  $0.2\rho_e$  to  $2\rho_e$  are mainly investigated here. The detailed coordinates of these scatterers are listed in Table III. White Gaussian noise is added to the SAR 3-D imaging scene and the SNR is set to be 10 dB.

Depending on the common BF method, the MIMO array SAR 3-D imaging results are shown in Fig. 5. As can be seen, the two scatterers with the elevation space smaller than the Rayleigh resolution cannot be distinguished successfully. It means that the BF method has no super-resolution ability. Also, it always suffers from the high sidelobe interference as illustrated in the corresponding azimuth-elevation and range-elevation projection planes of Fig. 5. The SAR 3-D imaging results based on the  $\ell_1$ -norm minimization approach are also shown in Fig. 5. One can observe from this figure that the

TABLE III  
LOCATIONS OF ALL SCATTERERS

$(x_r, y_a, z_h)$		$(x_r, y_a, z_h)$	
$q_1$	$(-8, -8, 0.5)$	$q_2$	$(-8, -8, -0.5)$
$q_3$	$(-6, -6, 1.0)$	$q_4$	$(-6, -6, -1.0)$
$q_5$	$(-4, -4, 1.5)$	$q_6$	$(-4, -4, -1.5)$
$q_7$	$(-2, -2, 2.0)$	$q_8$	$(-2, -2, -2.0)$
$q_9$	$(0, 0, 2.5)$	$q_{10}$	$(0, 0, -2.5)$
$q_{11}$	$(2, 2, 3.0)$	$q_{12}$	$(2, 2, -3.0)$
$q_{13}$	$(4, 4, 3.5)$	$q_{14}$	$(4, 4, -3.5)$
$q_{15}$	$(6, 6, 4.0)$	$q_{16}$	$(6, 6, -4.0)$
$q_{17}$	$(8, 8, 4.5)$	$q_{18}$	$(8, 8, -4.5)$
$q_{19}$	$(10, 10, 5.0)$	$q_{20}$	$(10, 10, -5.0)$

$\ell_1$ -norm minimization method can separate the two scatterers with the interval slightly smaller than the Rayleigh resolution. When the elevation space of  $q_3$  and  $q_4$  closes to  $0.4\rho_e$ , there is only one false scatterer existing in the reconstruction result. This implies that the super-resolution performance of the  $\ell_1$ -norm minimization cannot be guaranteed any more in the case of the distribution space of scatterers far smaller than the Rayleigh resolution. Finally, the 3-D SAR imagery based on the proposed DNN-driven super-resolution imaging algorithm is depicted in the bottom of Fig. 5. The reconstruction result is in agreement with the original 3-D distributed scene of scatterers. It can be seen that the proposed algorithm is capable of distinguishing two scatterers and locating their positions successfully even when the interval is equivalent to one-fifth of the Rayleigh resolution. That is, it presents the state-of-the-art super-resolution performance compared with other related methods.

Additionally, for a more explicit demonstration of the resolving ability regarding the aforementioned reconstruction algorithms, the azimuth-elevation plane 2-D imaging results and the 1-D slices in the elevation direction corresponding to the specific four scatterers  $q_1$ ,  $q_2$ ,  $q_5$ , and  $q_6$  are shown in Fig. 6. It can be seen that the paired scatterers are superimposed as a false scatterer for the BF reconstruction. Considering the  $\ell_1$ -norm minimization algorithm, it can separate the two scatterers  $q_5$  and  $q_6$  with  $0.6\rho_e$  interval. However, for the reconstruction of  $q_1$  and  $q_2$  with  $0.2\rho_e$  space, there is only one scatterer existing in the recovered result. The main reason is that the interval of two scatterers is too small such that the mutual coherence of the corresponding steering vectors for the measurement matrix is greatly increased. This implies that the sparse recovery result is no longer reliable according to the CS theory. In contrast, a quick sanity shows that the proposed approach can separate the paired scatterers successfully and predict their locations accurately. As a consequence, it can be preliminarily concluded that the super-resolution ability of the proposed approach can be improved significantly. This is attributed to the powerful learning and mapping ability of the end-to-end DNN regression model.

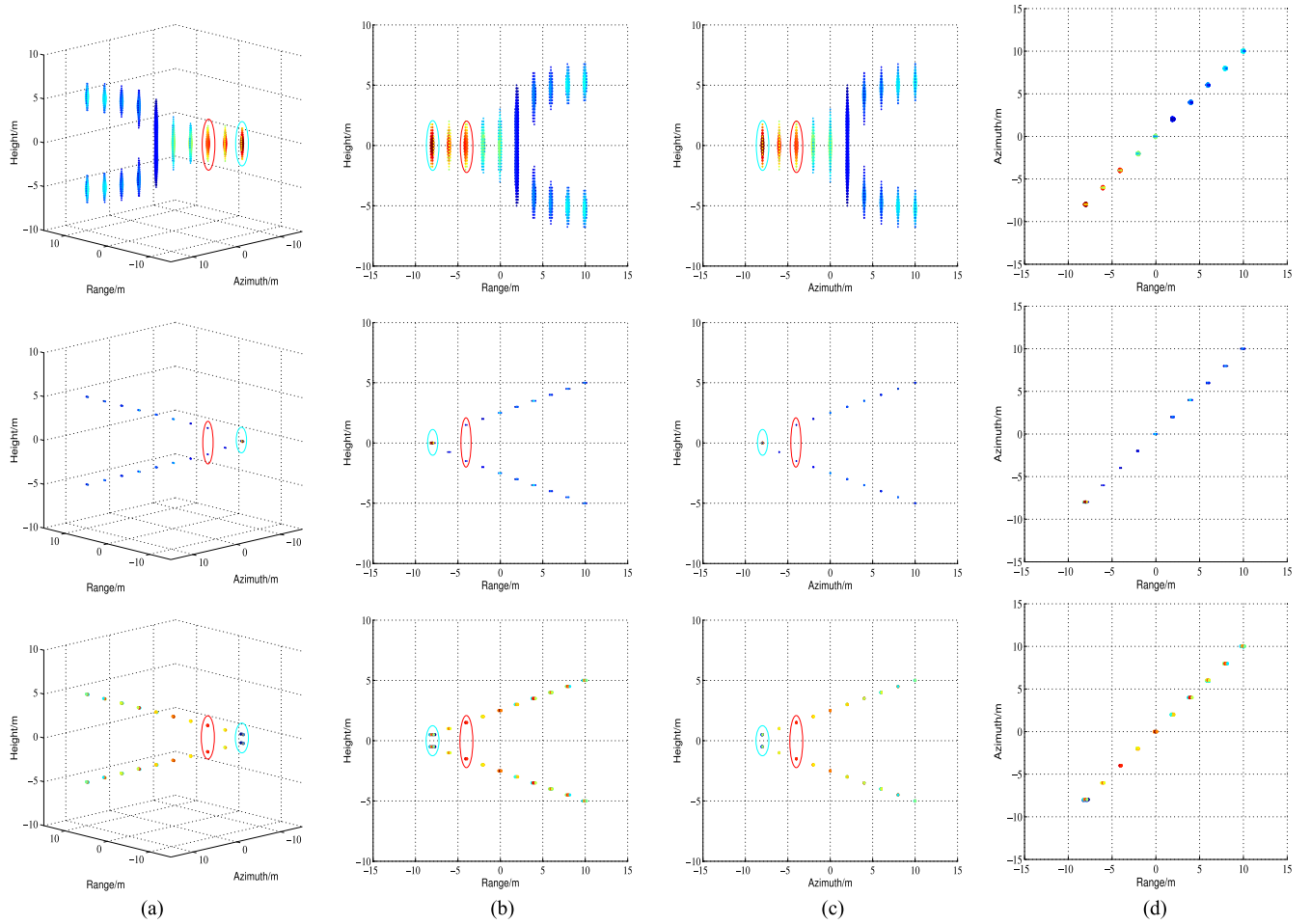


Fig. 5. 3-D imaging results by the BF, the  $\ell_1$ -norm minimization, and the proposed DNN-driven imaging algorithm from top to bottom, respectively. (a) 3-D imaging view. (b) Range-height plane. (c) Azimuth-height plane. (d) Range-azimuth plane.

In order to further verify the super-resolution performance of the proposed DNN-driven imaging algorithm, the separation success rates of the specific scatterers with changing elevation space versus different levels of SNR and sampling rate are observed here. In this article, two scatterers with the same range and azimuth positions distributed along the elevation direction are mainly considered. Specifically, the elevation location of the first one is fixed, and the second one is distributed from  $0.1\rho_e$  to  $2\rho_e$  referred to the first scatterer. The reconstruction is considered successful if the location error is less than the threshold, which is set to be  $0.1\rho_e$ . Fig. 7 presents the separation success rate of the related methods with different number of antenna elements as a function of the normalized distance. And, the SNR level is set to be 10 dB and the number of Monte Carlo trials is 1000. An examination of this figure shows that the reconstruction probability of the  $\ell_1$ -norm minimization degrades dramatically when the corresponding interval of two scatterers is less than half of the Rayleigh resolution. In contrast, the separation success rate of the proposed method is improved significantly and slightly affected by the normalized distance. It can be seen that the recovery probability can even reach 98% with the specific  $0.1\rho_e$  distance of two scatterers. Additionally,

the impact of different number of antenna elements on the separation success rate regarding the proposed method is relatively slight. This is attributed to the adequate training dataset and the powerful pretrained DNN regression model. The separation success rate under different levels of SNR is depicted in Fig. 8. In this circumstance, eight antenna elements are introduced with 1000 Monte Carlo trials. As we can see, the super-resolution performance of the  $\ell_1$ -norm minimization method cannot be guaranteed any more especially in the case of lower SNR. For the proposed DNN-driven imaging algorithm, it can be noted that the separation success rate can slightly affected by different SNR levels. Moreover, the recovery probability can also exceed 98% even when the interval of two scatterers is equivalent to  $0.1\rho_e$  in the case of SNR = 3 dB and 8 antenna elements. Finally, considering other deep learning-based sparse signal recovery methods, as a representative, the modified deep convolutional network (DCN) model designated as DeepInverse [37], which first utilizes a fully connected layer to boost the dimensionality of the input from  $\mathbb{R}^N$  to  $\mathbb{R}^K$  followed by convolutional layers without pooling, is also introduced here to investigate the super-resolution performance. In the training stage, we gradually increase the number of hidden layers and found that the best

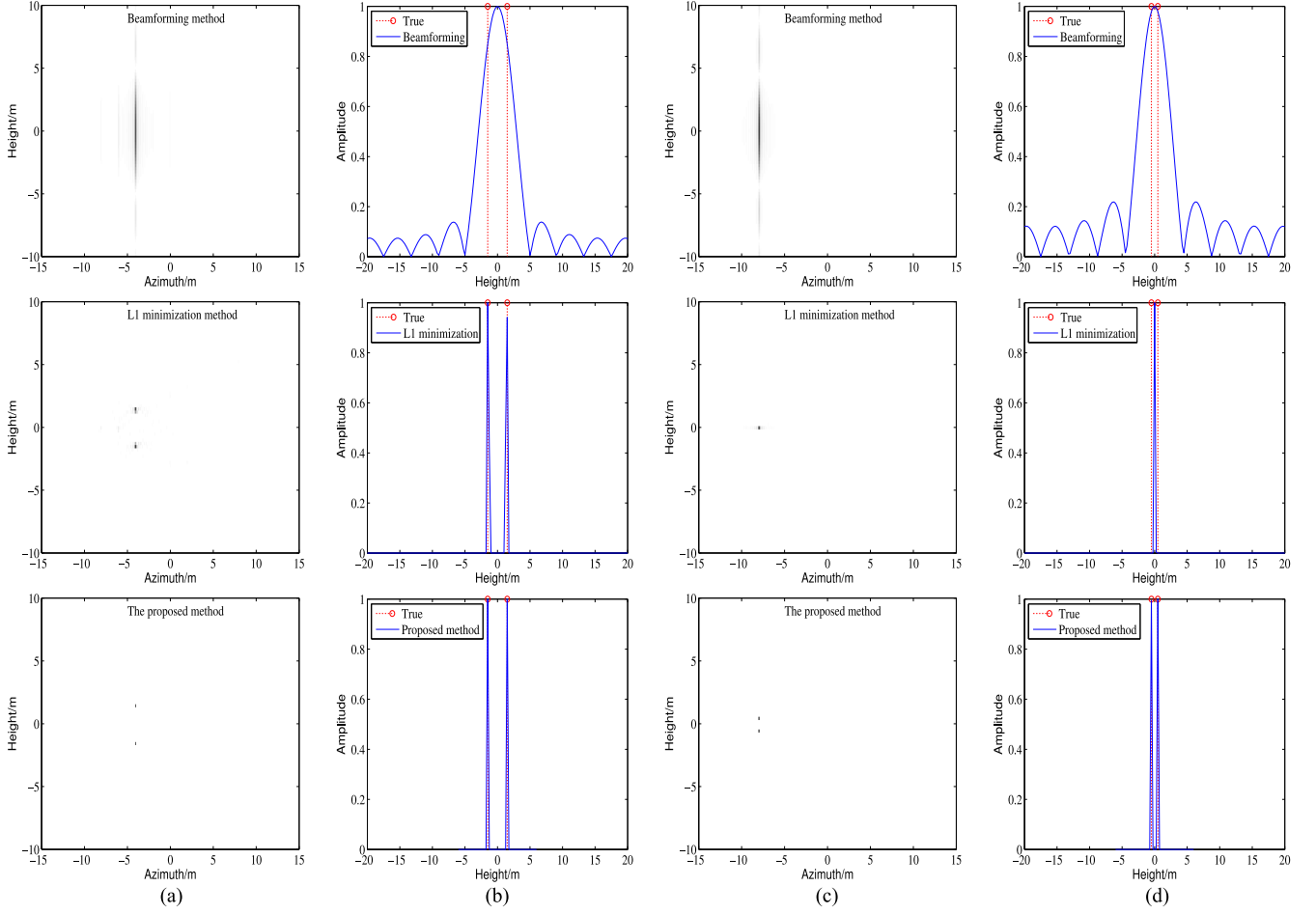


Fig. 6. Imaging results by the BF, the  $\ell_1$ -norm minimization, and the proposed DNN-driven imaging algorithm from top to bottom, respectively. (a) 2-D imaging results of  $q_5$  and  $q_6$ . (b) 1-D slices of  $q_5$  and  $q_6$ . (c) 2-D imaging results of  $q_1$  and  $q_2$ . (d) 1-D slices of  $q_1$  and  $q_2$ .

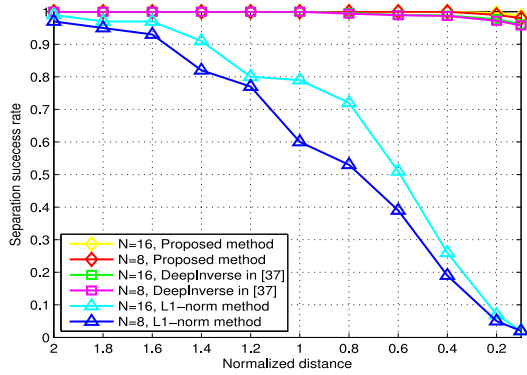


Fig. 7. Separation success rate as a function of normalized distance at  $\text{SNR} = 10$  dB under 1000 Monte Carlo trials.

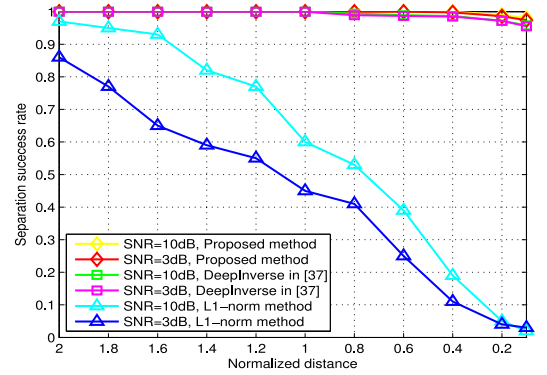


Fig. 8. Separation success rate as a function of normalized distance at  $N = 8$  under 1000 Monte Carlo trials.

performance can be achieved with the following specifications. The input layer is fully connected with the first hidden layer to boost the dimensionality from  $\mathbb{R}^{16}$  to  $\mathbb{R}^{47}$ , which is equivalent to the dimension of the recovered signal  $\gamma \in \mathbb{C}^{K \times 1}$  ( $K = 47$ ). For the following convolutional layers without pooling, there are 4 layers in total and the number of filters is 256, 512, 256, and 64, respectively. The sizes of these filters are  $47 \times 1$ . The output

layer has 1 filter with the size of  $47 \times 1$ . Also, we found that the super-resolution performance will degrade if the size of the filter is reduced. The reason is that the input signal does not have local correlation, and hence, the conventional local connections and shared weights in the convolutional layers will weaken the expressive power of the neural network to some extent. Therefore, it is necessary to design the filter with a complete



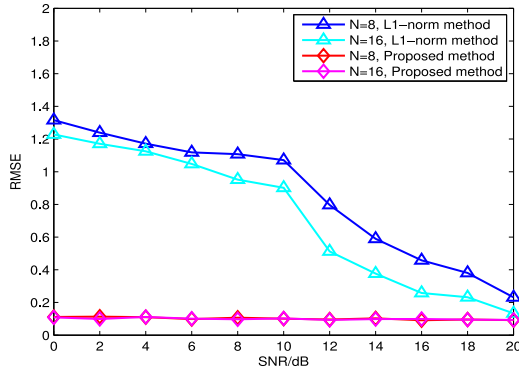


Fig. 9. RMSE of different algorithms with different SNR levels and antenna elements.

receptive field for the input signal. Finally, the super-resolution performance of the DeepInverse method versus different antenna elements and SNR levels is also presented in Figs. 7 and 8, respectively. As we can see, both of the proposed method and DeepInverse approach can improve the super-resolution performance significantly compared with the conventional sparse reconstruction algorithms. In addition, it can be observed from this figure that the super-resolution performance of the proposed method is better than the DeepInverse approach when the spacing of two scatterers is less than  $0.8\rho_e$ , and the performance gap reaches the maximum in the case of the specific  $0.1\rho_e$  distance between two scatterers. As mentioned above, the reason is that the local connections and shared weights in the convolutional layers weaken the expressive capability of the neural network to a certain extent when compared with the dense layers. Thus, in extreme cases, for example, the interval of two scatterers is  $0.1\rho_e$ , the super-resolution performance of the proposed method is better than that of the modified DCN model including DeepInverse. In summary, it can be concluded that the proposed DNN-driven imaging algorithm in the elevation direction can reach the state-of-art super-resolution performance than other existing sparse recovery methods within the CS framework even in the case of lower sampling rate and SNR levels.

### C. Performance Analysis of the Location Accuracy

Considering the performance of location accuracy, the root mean square error (RMSE) of the scatterers' positions is investigated in Fig. 9. It is based on the proposed method and the  $\ell_1$ -norm minimization with different number of antenna elements and different levels of SNR. For the  $\ell_1$ -norm minimization, the location estimation performance can be improved significantly with higher SNR level and more antenna elements. However, in the case of lower SNR level and few antenna elements, the performance of location accuracy will degrade drastically. In contrast, one can observe from this figure that the location accuracy performance of the proposed DNN-driven imaging method is much better than that of the  $\ell_1$ -norm minimization. The accurate location estimation performance can still be well maintained with 8 antenna elements at SNR = 0 dB. In



Fig. 10. Optical image of the test site.

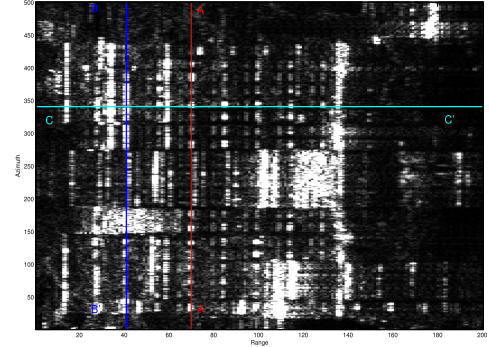


Fig. 11. 2-D SAR image.

brief, the location accuracy performance of the proposed method can be improved significantly compared with the conventional sparse recovery algorithms especially in the case of lower SNR level or few antenna elements.

### D. Performance Verification on MIMO Array SAR Real Data

In order to better verify the performance of the proposed DNN-driven imaging algorithm, the MIMO array SAR system, developed by the Institute of Electronics of the Chinese Academy of Sciences, is utilized to collect the raw data of the illuminated scene to validate our method. Meanwhile, as a representative, the  $\ell_1$ -norm minimization approach within the framework of CS is also introduced for the latter performance comparison. The optical image of the test site is shown in Fig. 10. Correspondingly, Fig. 11 presents the 2-D SAR image of the specific manmade structure as depicted by the red rectangle in Fig. 10. In real scenario, some preliminary processing, e.g., multiple SAR images registration and phase calibration, are first conducted on the 2-D SAR dataset. Additionally, the pixels of the 2-D SAR image with stronger scattering amplitude, which can well reflect the building structure, are assumed to be the 3-D reconstruction candidates. For the 3-D imaging results, there are always some outliers existing in the elevation direction due to the noise interference. Typically, the iterative reweighted technique and model selection method can have the capability to suppress the undesired outliers and locate the true scatterers. In this article, unless otherwise stated, the model selection method based on the Bayesian information criterion is considered the default choice [61].

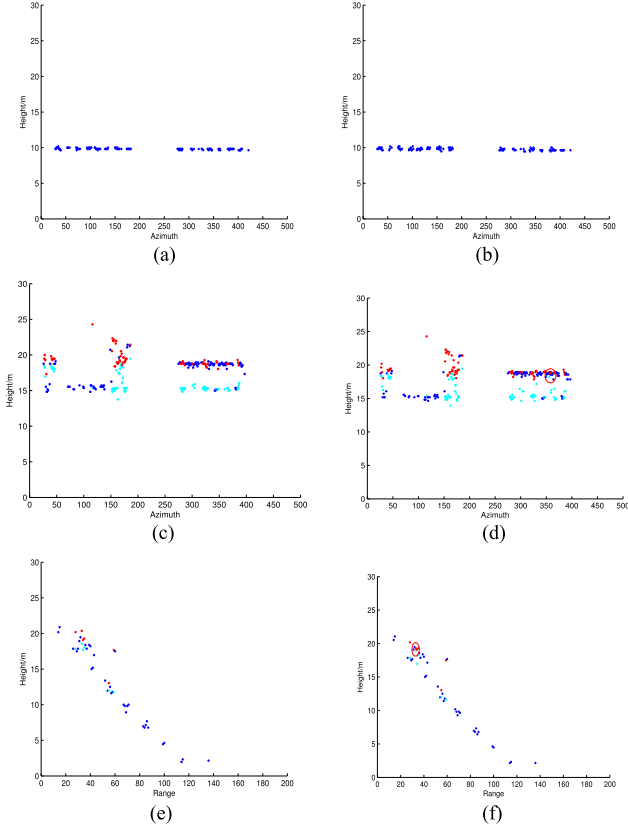


Fig. 12. (a) and (b) Profiles of slice AA' by  $\ell_1$  and the proposed method, respectively. (c) and (d) Profiles of slice BB' by  $\ell_1$  and the proposed method, respectively. (e) and (f) profiles of slice CC' by  $\ell_1$  and the proposed method, respectively.

First, for the purpose of validating the performance of the proposed method, the slices AA', BB', and CC' (see Fig. 11) are preliminarily reconstructed and depicted in Fig. 12, respectively. It can be seen that the color of the scatterers is divided into three types; that is, blue, cyan, and red. Specifically, single scatterer located in the azimuth-range resolution cell belongs to the blue point. In regard to the overlaid double scatterers, they are represented by the paired cyan and red points. For the retrieved profile of slice AA', one can observe from the figure that the reconstructed scatterers are well gathered at a specific height, which denotes one floor of the building. In view of slice BB', besides the single scatterers distributed in a floor of the building, an examination of this figure shows that there are also paired double scatterers existed in the reconstruction result. The reason is that the two scatterers located in the same range cell are distributed in the facade and roof of the building separately. With respect to the profile of slice CC', a quick sanity shows that the scatterers belonging to different floors of the building are accurately identified. In addition, it can be noted that some paired double scatterers with much smaller interval marked by the red ellipse [see Fig. 12(d) and (f)] can be successfully separated by the proposed DNN-driven imaging method. In contrast, these paired scatterers cannot be detected by the  $\ell_1$ -norm minimization method [see Fig. 12(c) and (e)].

Thus, it can be concluded that the super-resolution performance of the proposed method is better than other sparse reconstruction algorithms.

The complete 3-D imaging results of the test site by the proposed approach and the  $\ell_1$ -norm minimization are shown in Fig. 13, respectively. One can observe from this figure that the potential scatterers distributed in different floors of the building are clearly identified and well gathered at a certain height. A quick sanity shows that most detected scatterers are single scatterers, which mainly come from the facade of the building and can well represent the outline and details of the structure. In regard to the overlaid double scatterers, they are mainly distributed in the top and bottom of the building, respectively. The reason is that the echoes mainly come from the roof and facade of the building for the higher parts. In view of the lower parts, the paired scatterers are generated due to the staircase and entrance existing in the front of the building. To be specific, for the  $\ell_1$ -norm minimization method, 9616 single scatterers and 1405 double scatterers are detected as depicted in the first row of Fig. 13. In contrast, with respect to the proposed DNN-driven imaging algorithm, 9635 single scatterers and 1536 double scatterers are identified. It can be seen that the number of double scatterers detected by the proposed method is bigger than that of the conventional CS algorithm. This is mainly due to the super-resolution capability of the proposed DNN-driven imaging algorithm, which can separate two adjacent scatterers successfully with space more smaller than the Rayleigh resolution. In addition, some specific double scatterers marked by the red ellipse in Fig. 13 are investigated to validate the super-resolution performance. After quantitative analysis, it can be found that the spacing of these paired scatterers is exactly equivalent to the discrete grid spacing. Considering that the grid discrete factor  $\alpha$  is 0.1, this implies that the interval of the overlaid scatterers is equal to one tenth of the Rayleigh resolution. This fact leads to the super-resolution advantage of the proposed method. Thus, it can be concluded that the super-resolution performance of the proposed DNN-driven imaging algorithm can be improved significantly.

As for the computational complexity, the computational time of the proposed method against the  $\ell_1$ -norm minimization approach is mainly considered here. Considering the imaging processing steps of  $\ell_1$ -norm minimization, the computational consumption mainly focuses on the sparse recovery and model selection. In regard to the proposed DNN-driven imaging algorithm, the processing time is mainly spent on preliminary sparse recovery with  $\ell_1$ -norm minimization, DNN training and prediction. Thus, the  $\ell_1$ -norm minimization based SAR 3-D imaging takes about 6 h. In contrast, the proposed DNN-driven imaging algorithm only costs about 2.7 h. The improvement mainly comes from two aspects. One reason is that the preliminary recovery with the discrete factor  $\alpha = 0.5$  by the  $\ell_1$ -norm minimization can save a lot of computational time in comparison with the straightforward sparse recovery by  $\ell_1$  with fine discrete grid points ( $\alpha = 0.1$ ). The other reason is that the proposed parallel DNN regression models with less training dataset and neurons in the multiple connected

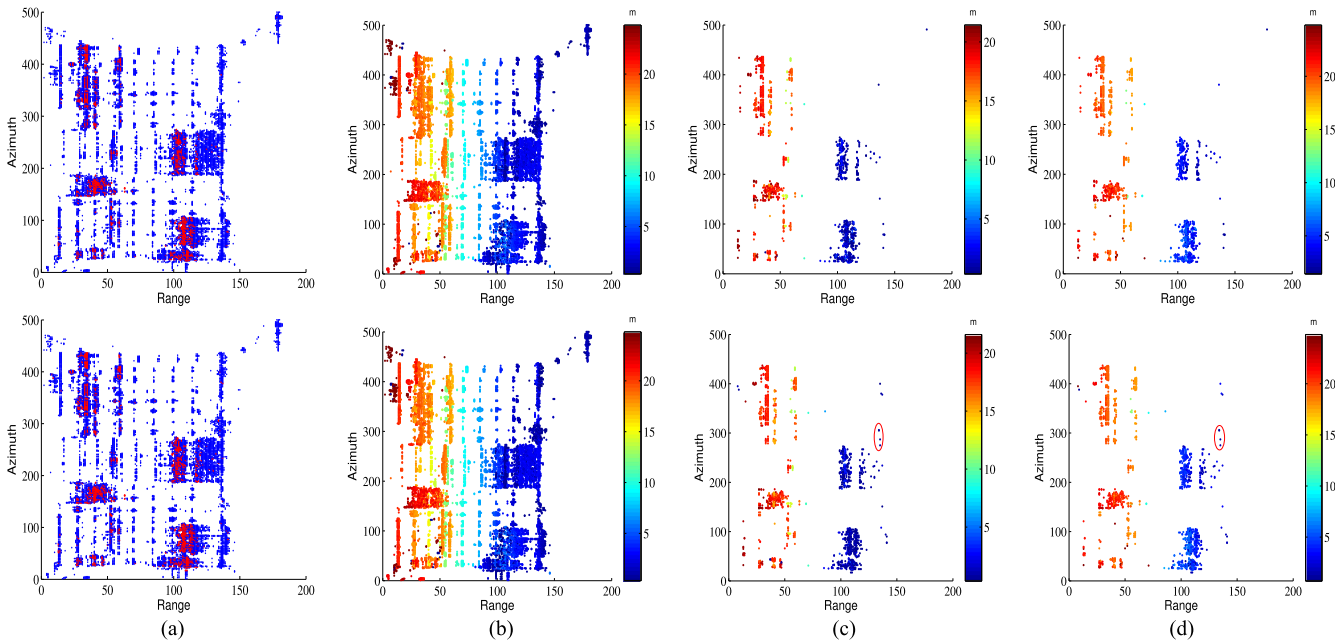


Fig. 13. 3-D imaging results by the  $\ell_1$ -norm minimization and the proposed DNN-driven imaging algorithm from top to bottom, respectively. (a) Number of scatterers in each pixel. Blue for single scatterers and red for double scatterers. (b) Heights for single scatterers. (c) Heights for lower paired scatterers. (d) Heights for higher paired scatterers.

layers can greatly reduce the computational consumption. In summary, the performance of our proposed algorithm regarding the super-resolution capability and computational complexity can be improved significantly compared with other existing related methods.

## V. CONCLUSION

In this article, an advanced super-resolution imaging framework based on CS and DNN for MIMO array SAR is proposed. First, in accordance with the preliminary recovery results by the  $\ell_1$ -norm minimization method, the spatial filtering is introduced to retain scattering signals only impinging from the prespecified spatial subregions. Based on this fact, a group of parallel end-to-end fully connected DNN regression models are designed for mapping the potential sparse recovery mathematical model and further locating the true scatterers in the elevation direction. Finally, numerical simulations and airborne MIMO array SAR experiments are shown to demonstrate the state-of-the-art super-resolution performance against other existing related methods. As a comparison, the proposed algorithm is capable of separating two scatterers with an interval of one-tenth of the Rayleigh resolution. In addition, the performance of the proposed method regarding the location accuracy and computational efficiency is also better than other sparse recovery methods within the CS framework.

## ACKNOWLEDGMENT

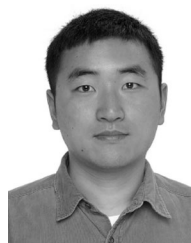
The authors would like to thank the Associate Editor and the anonymous reviewers for their valuable comments and suggestions that significantly improved this article.

## REFERENCES

- [1] G. Fornaro, F. Lombardini, and F. Serafino, "Three-dimensional multipass SAR focusing: Experiments with long-term spaceborne data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 702–714, Apr. 2005.
- [2] X. X. Zhu and R. Bamler, "Very high resolution spaceborne SAR tomography in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4296–4308, Dec. 2010.
- [3] O. Frey and E. Meier, "Analyzing tomographic SAR data of a forest with respect to frequency, polarization, and focusing technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3648–3659, Oct. 2011.
- [4] S. Tebaldini and F. Rocca, "Multibaseline polarimetric SAR tomography of a boreal forest at P- and L-bands," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 232–246, Jan. 2012.
- [5] S. Tebaldini, T. Nagler, H. Rott, and A. Heilig, "L-band 3D imaging of an alpine glacier: Results from the AlpTomoSAR campaign," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2015, pp. 5212–5215.
- [6] K. K. Knaell and G. P. Cardillo, "Radar tomography for the generation of three-dimensional images," *IEEE Proc.—Radar, Sonar Navigat.*, vol. 142, no. 2, pp. 54–60, Apr. 1995.
- [7] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.
- [8] A. Ishimaru, T.-K. Chan, and Y. Kuga, "An imaging technique using confocal circular synthetic aperture radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 5, pp. 1524–1530, Sep. 1998.
- [9] T.-K. Chan, Y. Kuga, and A. Ishimaru, "Experimental studies on circular SAR imaging in clutter using angular correlation function technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2192–2197, Sep. 1999.
- [10] O. Ponce *et al.*, "Fully polarimetric high-resolution 3-D imaging with circular SAR at L-band," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3074–3090, Jun. 2014.
- [11] O. Ponce, P. Prats-Iraola, R. Scheiber, A. Reigber, and A. Moreira, "First airborne demonstration of holographic SAR tomography with fully polarimetric multicircular acquisitions at L-band," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6170–6196, Oct. 2016.
- [12] M. Weib and J. H. G. Ender, "A 3D imaging radar for small unmanned airplanes—ARTINO," in *Proc. Eur. Radar Conf.*, Oct. 2005, pp. 209–212.
- [13] J. Nouvel, H. Jeuland, G. Bonin, S. Roques, O. D. Plessis, and J. Peyret, "A Ka band imaging radar: DRIVE on board ONERA motorglider," in *Proc. IEEE Int. Symp. Geosci. Remote Sens.*, Jul. 2006, pp. 134–136.



- [14] J. Li, P. Stoica, and X. Zheng, "Signal synthesis and receiver design for MIMO radar imaging," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3959–3968, Jul. 2008.
- [15] W. Roberts, P. Stoica, J. Li, T. Yardibi, and F. A. Sadjadi, "Iterative adaptive approaches to MIMO radar imaging," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 1, pp. 5–20, Feb. 2010.
- [16] W. Q. Wang, "MIMO SAR imaging: Potential and challenges," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 28, no. 8, pp. 18–23, Aug. 2013.
- [17] X. X. Zhu and R. Bamler, "Tomographic SAR inversion by  $L_1$ -norm regularization/the compressive sensing approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3839–3846, Oct. 2010.
- [18] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- [19] E. J. Candès *et al.*, "Compressive sampling," in *Proc. Int. Congr. Math.*, Madrid, Spain, Aug. 2006, vol. 3, pp. 1433–1452.
- [20] D. L. Donoho *et al.*, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [21] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [22] C. Wu, Z. Zhang, X. Liang, L. Chen, W. Yu, and T.-K. Truong, "Fast 3-D imaging algorithm based on unitary transformation and real-valued sparse representation for MIMO array SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7033–7047, Sep. 2019.
- [23] X. Ren, L. Chen, and J. Yang, "3D imaging algorithm for down-looking MIMO array SAR based on Bayesian compressive sensing," *Int. J. Antennas Propag.*, vol. 2014, no. 7, pp. 1–9, May 2014.
- [24] Q. Bao, X. Peng, Z. Wang, Y. Lin, and W. Hong, "DLSLA 3-D SAR imaging based on reweighted gridless sparse recovery method," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 6, pp. 841–845, Jun. 2016.
- [25] X. X. Zhu and R. Bamler, "Super-resolution power and robustness of compressive sensing for spectral estimation with application to spaceborne tomographic SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 247–258, Jan. 2011.
- [26] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.
- [27] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.
- [28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [29] G. Hinton *et al.*, "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.
- [30] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," in *Proc. 25th Int. Conf. Mach. Learn.*, 2008, pp. 160–167.
- [31] Z.-M. Liu, C. Zhang, and S. Y. Philip, "Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 7315–7327, Dec. 2018.
- [32] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549–8560, Sep. 2018.
- [33] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989.
- [34] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [35] A. Mousavi, A. B. Patel, and R. G. Baraniuk, "A deep learning approach to structured signal recovery," in *Proc. Annu. Allerton Conf. Commun., Control, Comput.*, Oct. 2015, pp. 1336–1343.
- [36] 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.*, Jun. 2016, pp. 449–458.
- [37] A. Mousavi and R. G. Baraniuk, "Learning to invert: Signal recovery via deep convolutional networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2017, pp. 2272–2276.
- [38] H. Yao *et al.*, "DR<sup>2</sup>-Net: Deep residual reconstruction network for image compressive sensing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2017, pp. 451–462.
- [39] A. Bora, A. Jalal, E. Price, and A. Dimakis, "Compressed sensing using generative models," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 537–546.
- [40] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. Int. Conf. Mach. Learn.*, 2010, pp. 399–406.
- [41] J. H. R. Chang, C. Li, B. Pczos, B. V. K. V. Kumar, and A. C. Sankaranarayanan, "One network to solve them all—Solving linear inverse problems using deep projection models," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 5889–5898.
- [42] M. Borgerd and P. Schniter, "Onsager-corrected deep learning for sparse linear inverse problems," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Dec. 2016, pp. 227–231.
- [43] A. Budillon, A. Johnsy, G. Schirrinzi, and S. Vitale, "SAR tomography based on deep learning," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 3625–3628.
- [44] D. Bliss and K. Forsythe, "Multiple-input multiple-output (MIMO) radar and imaging: degrees of freedom and resolution," in *Proc. IEEE 37th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2003, vol. 1, pp. 54–59.
- [45] G. Krieger, N. Gebert, and A. Moreira, "Unambiguous SAR signal reconstruction from nonuniform displaced phase center sampling," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 260–264, Oct. 2004.
- [46] Y. Hou, D. Li, and W. Hong, "The thinned array time division multiple phase center aperture synthesis and application," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2008, vol. 5, pp. 25–28.
- [47] P. Ma, H. Lin, H. Lan, and F. Chen, "On the performance of reweighted  $L_1$  minimization for tomographic SAR imaging," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 4, pp. 895–899, Apr. 2014.
- [48] F. Gini, F. Lombardini, and M. Montanari, "Layover solution in multibaseline SAR interferometry," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 4, pp. 1344–1356, Oct. 2002.
- [49] C. Chen and Z. Xiaoling, "A new super-resolution 3D-SAR imaging method based on MUSIC algorithm," in *Proc. IEEE RadarCon*, Kansas, MO, USA, May 2011, pp. 525–529.
- [50] M. Grant and S. Boyd, *CVX: Matlab Software for Disciplined Convex Programming*, 2008, [Online]. Available: <http://cvxr.com/cvx/>
- [51] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.
- [52] R. J. Vaccaro and B. F. Harrison, "Optimal matrix-filter design," *IEEE Trans. Signal Process.*, vol. 44, no. 3, pp. 705–709, Mar. 1996.
- [53] Z. Zhu, S. Wang, H. Leung, and Z. Ding, "Matrix filter design using semi-infinite programming with application to DOA estimation," *IEEE Trans. Signal Process.*, vol. 48, no. 1, pp. 267–271, Jan. 2000.
- [54] C. S. MacInnes, "Source localization using subspace estimation and spatial filtering," *IEEE J. Ocean. Eng.*, vol. 29, no. 2, pp. 488–497, Jul. 2004.
- [55] M. Pesavento, A. B. Gershman, and Z.-Q. Luo, "Robust array interpolation using second-order cone programming," *IEEE Signal Process. Lett.*, vol. 9, no. 1, pp. 8–11, Aug. 2002.
- [56] S. Yan, C. Hou, and X. Ma, "Direction-of-arrival estimation using matrix spatial prefiltering approach," *Chin. J. Acoust.*, vol. 27, no. 1, pp. 727–739, 2008.
- [57] C. Trabelsi *et al.*, "Deep complex networks," in *Proc. Int. Conf. Learn. Representations*, 2018, [Online]. Available: <https://openreview.net/forum?id=HIT2hmZAb>
- [58] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. 14th Int. Conf. Artif. Intell. Statist.*, 2010, vol. 15, pp. 315–323.
- [59] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," Jan. 2017, *arXiv:1412.6980v9*.
- [60] F. Chollet, *Keras*, 2015, [Online]. Available: <http://keras.io/>
- [61] K. P. Burnham and D. R. Anderson, "Multimodel inference: understanding AIC and BIC in model selection," *Sociol. Methods Res.*, vol. 33, no. 2, pp. 261–304, 2004.



**Chunxiao Wu** received the B.Sc. degree in automation engineering from University of Electronic Science and Technology of China, Chengdu, China, in 2015. He is currently working toward the Ph.D. degree with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China.

His research interests include radar signal processing, SAR 3-D imaging, compressive sensing, and deep learning.



**Zenghui Zhang** (Member, IEEE) received the B.Sc. degree in applied mathematics, the M.Sc. degree in computational mathematics, and the Ph.D. degree in information and communication engineering from the National University of Defense Technology (NUDT), Changsha, China, in 2001, 2003, and 2008, respectively.

From 2008 to 2012, he was a Lecturer with the Department of Mathematics and System Science, NUDT. Since 2013, he has been an Associate Professor with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China. His main research interests include MIMO SAR imaging methods, SAR image detection and classification, joint radar and communication system, and clutter mitigation and target detection.



**Wenxian Yu** was born in Shanghai, China, in 1964. He received the B.Sc. degree in radio measurement and control and data transmission, the M.Sc. degree in communication and electronic system, and the Ph.D. degree in communication and information system from the National University of Defense Technology (NUDT), Changsha, China, in 1985, 1988, and 1993, respectively.

From 1996 to 2008, he was a Professor with the College of Electronic Science and Engineering, NUDT, where he was the Deputy Head and the Assistant Director of the National Key Laboratory of Automatic Target Recognition. He was the Executive Dean of Shanghai Jiao Tong University, Shanghai, China, from 2009 to 2011. He is currently with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, where he is a Yangtze River Scholar Distinguished Professor and the Head of Research. His current research interests include radar target recognition, remote sensing information processing, multisensor data fusion, and integrated navigation system. In these areas, he has authored or coauthored more than 200 research papers.



**Longyong Chen** received the B.Sc. degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2003 and the M.Sc. and Ph.D. degrees in signal and information processing from the University of Chinese Academy of Sciences, Beijing, China, in 2006 and 2009, respectively.

He is currently a Professor with the Science and Technology on Microwave Imaging Laboratory, Institute of Electronics, Chinese Academy of Sciences, Beijing, China. He has product responsibility for many important Chinese national programs, including high-resolution synthetic aperture radar (SAR) systems. His research interests include real-time radar signal processing, coherent polarimetric, and interferometric SAR systems.