International Journal of Microwave and Wireless Technologies

cambridge.org/mrf

Physical model-driven deep networks for through-the-wall radar imaging

Yuhao Wang, Yue Zhang, Mingcheng Xiao, Huilin Zhou, Qiegen Liu and Jianfei Gao

School of Information Engineering, Nanchang University, 999 Xuefu Avenue, Honggutan New District, Nanchang, Jiangxi, China

Research Paper

Cite this article: Wang Y, Zhang Y, Xiao M, Zhou H, Liu Q, Gao J (2023). Physical model-driven deep networks for through-the-wall radar imaging. *International Journal of Microwave and Wireless Technologies* **15**, 82–89. https://doi.org/10.1017/S1759078722000071

Received: 12 September 2021
Revised: 4 January 2022
Accepted: 6 January 2022
First published online: 3 February 2022

Key words:
Compressed sensing; deep learning; image reconstruction; TWRI; unrolling network

Author for correspondence:
Yuhao Wang,
E-mail: wangyuhao@ncu.edu.cn

Abstract

In order to merge the advantages of the traditional compressed sensing (CS) methodology and the data-driven deep network scheme, this paper proposes a physical model-driven deep network, termed CS-Net, for solving target image reconstruction problems in through-the-wall radar imaging. The proposed method consists of two consequent steps. First, a learned convolutional neural network prior is introduced to replace the regularization term in the traditional iterative CS-based method to capture the redundancy of the radar echo signal. Moreover, the physical model of the radar signal is used in the data consistency layer to encourage consistency with the measurements. Second, the iterative CS optimization is unrolled to yield a deep learning network, where the weight, regularization parameter, and the other parameters are learnable. A quantity of training data enables the network to extract high-dimensional characteristics of the radar echo signal to reconstruct the spatial target image. Simulation results demonstrated that the proposed method can achieve accurate target image reconstruction and was superior to the traditional CS method, in terms of mean squared error and the target texture details.

Introduction

The through-the-wall radar imaging (TWRI) technique has been widely used in a variety of applications such as counter-terrorism, life detection in calamity rescuing, and security check for its non-destructive and non-contactable characteristics [1–3]. These applications must be guaranteed with high-resolution and high-sensitivity detection properties. However, satisfying these properties is considerably challenging due to the factors such as strong wall clutter reflection, radar echo signal sparseness characteristics, and prolonged data collection [4]. To suppress the wall clutter reflection, numerous research studies have been exploited such as low-rank and sparse representation (LRSR), spatial filtering, and subspace projection [5–7]. To reconstruct the target image, back projection (BP), and compressed sensing (CS) can be implemented. Compared to BP, CS can recover a sparse signal from a small number of data sets, whose sampling rate is far lower than what the Nyquist sampling theory claims, by exploiting signal sparsity [8]. Hence, a CS-based method is possible to reconstruct target image with fewer data collection, which alleviates the burden of prolonged data collection and computational cost [9, 10]. In general, CS-based microwave imaging methods can be regarded as linear imaging under the born approximation. It can provide an accurate reconstruction scheme when the target dielectric constant is small and the target size is appropriate.

Since the TWRI is an ill-posed and inverse problem, it can be solved by using a minimum iterative optimization method enforced by regularization term. There are various regularization terms that can fit different scenarios. The ℓ_2 (or Tikhonov) regularization yields linear reconstruction algorithms so that it can be applied to smooth scenarios [11]. Total-variation (TV) regularization preserves the rapid transitions and discontinuities in the signal to impose stability to signal recovery so that it is more suitable for piecewise smooth scenarios [12]. Low-rank and joint sparsity regularization in CS methodology implement wall clutter estimation with discrete prolate spheroidal sequence basis before estimating the target image so that it is suitable for handling the nonlinear inverse problem [4]. An autofocus CS algorithm is proposed to construct the imagery under the condition of unknown wall parameters. In the autofocus CS algorithm, the nonlinear conjugate gradient method is used to estimate the hyperparameters during the iterative framework [13]. Unfortunately, the traditional regularization term is not always optimal since they are not omni-knowledge to all types of scenarios.

In order to tackle these issues, deep learning has attracted much attention in handling the signal recovery problem in recent years [14–17, 19]. In [14], a model-based deep learning network was proposed to recover the target image from noisy and sparse measurements based on an alternating recursive algorithm with the physical model. In [15], a framework based on deep learning is proposed for solving the electromagnetic (EM)inverse scattering

© The Author(s), 2022. Published by Cambridge University Press in association with the European Microwave Association

CAMBRIDGE UNIVERSITY PRESS

problem, which constructs and expands the functions of existing physics-based inversion algorithms. In [16], there are also related physics-based deep learning methods for solving the inverse scattering problem. In [17], a convolutional neural network (CNN)-based human posture reconstruction method for TWRI is proposed to reconstruct accurate human posture behind walls. The training process follows a supervision-prediction learning pipeline inspired by the cross-modal learning technique. In [18], an approach based on generative adversarial nets is proposed to suppress multi-path ghosts, side/grating lobe artifacts, and wall penetration effect suppression in MIMO TWRI. In [19], the authors proposed a fully convolutional network (FCN) to extract multi-scale features to implement target imaging in TWRI. However, the lack of physical model in the pure FCN in [19] causes the network to depend heavily on a large amount of training data and is hard to converge.

Motivated by the analysis mentioned above, a TWRI scheme is proposed based on a physical model-driven deep learning network prior, termed CS-Net, which is unrolled from the iterative CS optimization. Compared with the pure data-driven deep learning method in TWRI, CS-Net can achieve faster convergence speed with less training data. Our main contributions are summarized as follows:

(i) A physical model of the radar signal is considered in our proposed deep network to reduce the number of network parameters and avoid the network to be a black-box solver.
(ii) We propose a deep network unrolled from the iterative CS optimization. Compared with the traditional CS method, it greatly improves the performance of texture detail reconstruction and the reconstructed image has a higher resolution.
(iii) The performance of the target image reconstruction in TWRI using CS-Net is evaluated and compared with those traditional CS, BP, and LRSR schemes. Simulation results demonstrate that the performance of the proposed method is better than the traditional methods in terms of minimum square error (MSE) and target texture details.

The rest of this paper is organized as follows. Section "Signal model and compressed sensing" briefly introduces the radar signal physical model and reviews the CS model for target image reconstruction in TWRI. Section "CS-NET architecture" introduces the architecture of the model-driven deep learning network and loss function. Section "Simulation results" presents the experimental results and analysis. The final section concludes.

Notations: Boldface capital letters denote matrices and boldface lower case letters denotevectors. $(\cdot)^{\mathrm{H}}$ denotes the conjugate transpose operator. $\|\cdot\|_2$ denotes the Euclidean norm of a vector, respectively. $\mathbb{C}$ denotes the set of all the complex number, and $\mathbb{R}$ denotes the set of all the real number. $*$ denotes the convolution operation. $\mathbf{0}$ denotes the zero vector.

Signal model and compressed sensing

System scenario and signal model

The TWRI system in this study is considered in a two-dimensional (2-D) domain consisting of a pair of signal transmitter and receiver, a front wall, and a target. As depicted in Fig. 1, the signal transmitter and the receiver move against the wall with each step forms an antenna pair, which is equivalent to a multi-input-multi-output scanning array. In order to better illustrate the TWRI problem, the polarization of the antenna of the signal transmitter and receiver pair is along the z-axis. Also, a 2-D EM wave with the longitude direction along the z-axis is considered. Moreover, to investigate the performance limitation of our proposed method, the TWRI scenario is assumed to be surrounded by a perfect match area (PMA), which absorbs all the EM waves without reflection to form a free-space setting.

Hence, the radar echo signal $s_r(t)$ is given by the convolution of the transmitted signal $s_t(t)$ and the system function $h(t)$, which describes the pulse response of the imaging area:

$$s_r(t) = s_t(t) * h(t). \tag{1}$$

Assuming the material of the wall and the target is non-dispersive and homogeneous, the signal channel can only influence the magnitude and the phase of the signal [20]. $h(t)$ can be written as:

$$h(t) = \sum_{i=0}^{L} \sum_{j=0}^{W} \alpha_{i,j} \sigma_{i,j} \delta\left(t - \tau_{ij}\right), \tag{2}$$

where $\alpha_{i,j}$ and $\sigma_{i,j}$, respectively, denote the complex reflectivity coefficient and the path loss factor of a single point located at $\mathbf{x}_{i,j} = (x_i, y_j)$. $\delta(t - \tau_{ij})$ denotes the impulse response of a single point located at $\mathbf{x}_{i,j}$ with a propagation delay from transmitter to the single point, and back to the receiver. L and W denote the number of pixels along the length and width of the imaging area, respectively.

Here, we use matrix $\mathbf{X}$ to represent the complex reflectivity coefficient distribution of the imaging area:

$$\mathbf{X} = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,W} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,W} \\ \vdots & \vdots & \alpha_{i,j} & \vdots \\ \alpha_{L,1} & \alpha_{L,2} & \cdots & \alpha_{L,W} \end{pmatrix}. \tag{3}$$

Since the radar echo signal is affected by the complex reflectivity coefficient distribution of the imaging area, matrix $\mathbf{X}$ can be calculated by solving the inverse problem using the CS methodology based on the physical model of the signal in the next subsection.

Compressed sensing and physical model

The TWRI system can be thought of as an operator $\mathbf{A}$ that acts on an imaging domain vector $\mathbf{x}$ to yield a vector of measurement $\mathbf{b}$ [7]. The goal of the CS imaging scheme is to recover vector $\mathbf{x}$ from the measurement vector $\mathbf{b}$, which is given by:

$$\mathbf{b} = \mathbf{MPx} = \mathbf{Ax}, \tag{4}$$

where $\mathbf{A} = \mathbf{MP}$ denotes the sensing matrix. $\mathbf{M} \in \mathbb{R}^{Q \times (K \times N_a)}$ and $\mathbf{P} \in \mathbb{C}^{(K \times N_a) \times J}$ denote the Gaussian random measurement matrix and the dictionary matrix, respectively [21]. Q is the number of rows of matrix $\mathbf{M}$, K is the number of frequencies point, N_a represents the number of antenna pairs and $J = L \times W$ represents the image pixels. $\mathbf{x} \in \mathbb{C}^{J \times 1}$ denotes the image of complex reflectivity coefficient vector, which is obtained by resizing $\mathbf{X}$ into a column vector. In the TWRI system, the dictionary matrix $\mathbf{P}$ is given by:

https://doi.org/10.1017/S1759078722000071 Published online by Cambridge University Press

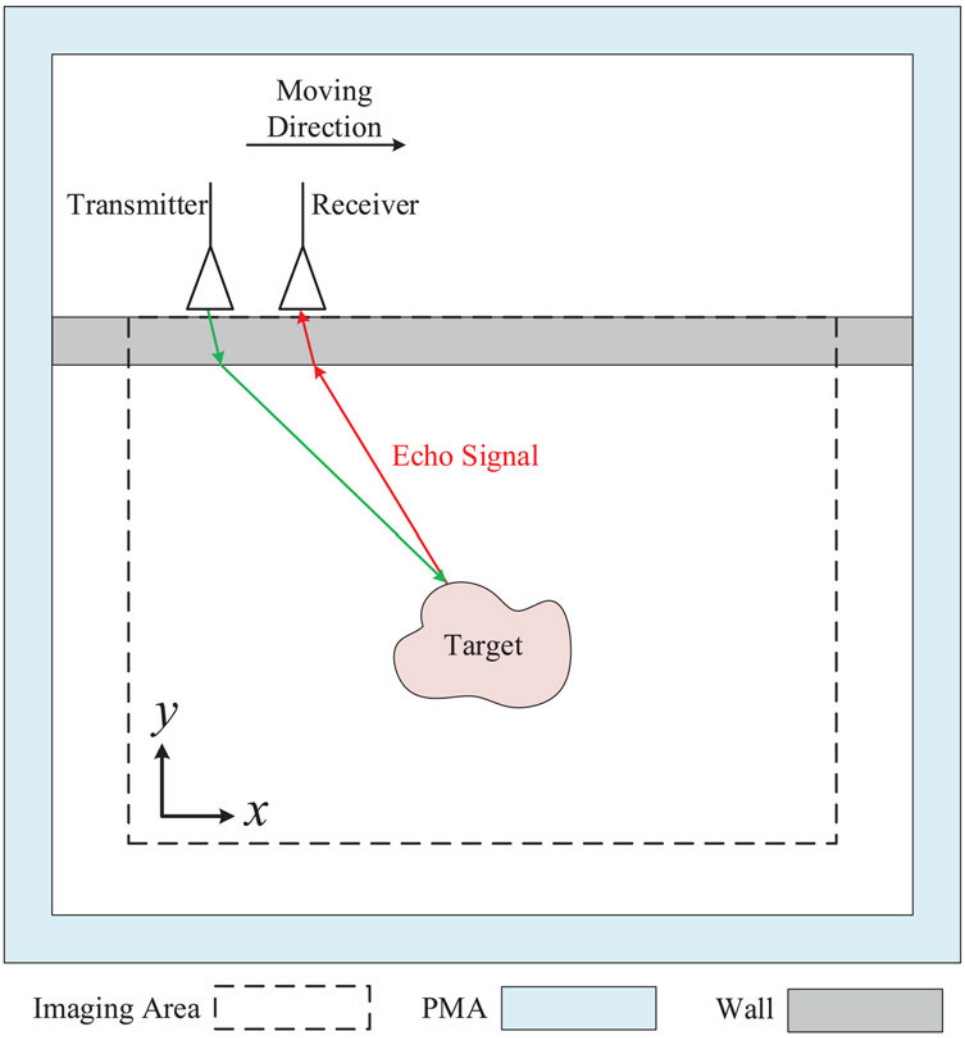


Fig. 1. TWRI system scenario consists of a pair of transmitter and receiver, a front wall, a target, and a PMA. The transmitter and receiver pair move against the wall to form a B-scan.

$$\mathbf{P} = \left(\mathbf{P}_1^{\mathrm{H}}, \mathbf{P}_2^{\mathrm{H}}, \ldots, \mathbf{P}_i^{\mathrm{H}}, \ldots, \mathbf{P}_{N_a}^{\mathrm{H}}\right)^{\mathrm{H}}, \tag{5a}$$

$$[\mathbf{P}_i]_{jk} = \exp(-j2\pi \mathbf{f}_k \mathbf{t}_{ij}), \tag{5b}$$

where $\mathbf{f} = (f_1, f_2, \ldots, f_k, \ldots, f_K)^{\mathrm{H}}$ denotes the frequencies point of the echo signal from the receiver with number of frequencies point K, where f_k denotes the k-th frequency point. $\mathbf{t}_{ij}$ denotes the round-trip delay between the i-th antenna position and the j-th image pixel. Since $\mathbf{A}$ represents the mapping from the imaging field to measurement field, measurement vector $\mathbf{b} \in \mathbb{C}^{Q\times 1}$ can be obtained by down-sampling radar echo signal matrix $\mathbf{Y}$ in frequency domain and vector $\mathbf{b}$ is given by:

$$\mathbf{b} = \mathbf{M}vec(\mathbf{Y}), \tag{6a}$$

$$\mathbf{Y} = \mathbf{FR}, \tag{6b}$$

where $\mathbf{R} \in \mathbb{C}^{T_i \times N_a}$ denotes the echo signal matrix from the receiver in time domain and $\mathbf{F} \in \mathbb{C}^{K\times T_i}$ denotes the 2-D discrete Fourier transform base matrix that acting on $\mathbf{R}$. Here, T_i represents the index of traveling time in time domain and $vec(\mathbf{Y})$ represents resize matrix $\mathbf{Y}$ into a column vector. Then, a high-resolution image vector $\mathbf{x}_{\mathrm{rec}}$ can be recovered from the under determined equations $\mathbf{b} = \mathbf{Ax}$ by solving the problem of the following minimization:

$$\mathbf{x}_{\mathrm{rec}} = \arg\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{Ax}\|_2^2 + \lambda R(\mathbf{x}), \tag{7}$$

where λ is a regularization parameter that adjusts the weights of data consistency (DC) term and the regularization term. $R(\mathbf{x})$ is an regularization prior. $R(\mathbf{x})$ is a small scalar when $\mathbf{x}$ is a noise-free image, while its value is high for noisy and artifact image [14]. Traditional choices include Tikhonov [11], TV [12], as well as their combinations. Nevertheless, the method mentioned above is not always optimum due to the following reasons:

(i) The reconstruction result can be affected by the initiative value $\mathbf{x}_0$, but there is no theoretical instruction on how to choose a suitable initiative value.
(ii) The λ plays a vital role in the reconstruction. But it is challenging to choose the most suitable λ value.

https://doi.org/10.1017/S1759078722000071 Published online by Cambridge University Press

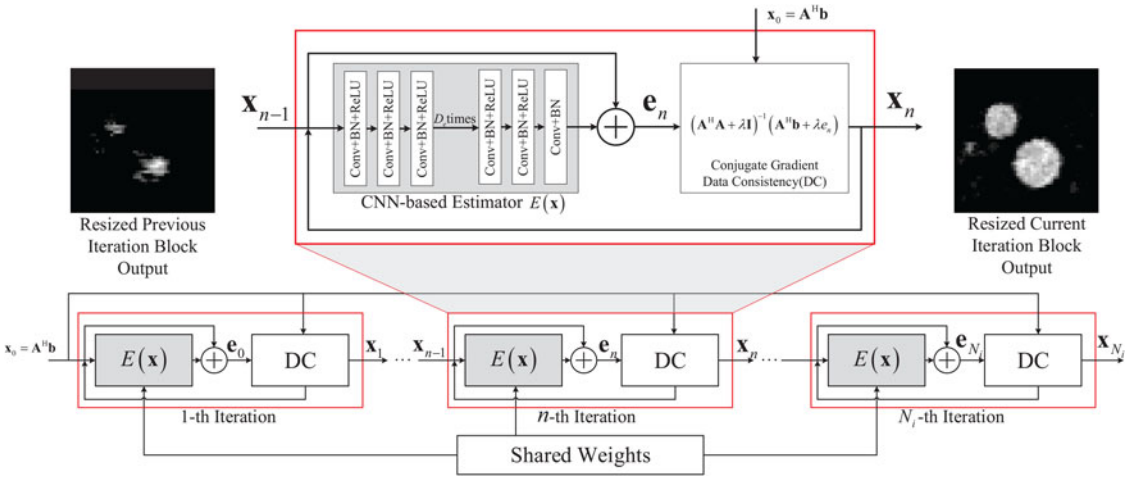


Fig. 2. Flowchart of the CS-Net architecture. The network consists of a cascade of iteration block, with each block consisting of a CNN-based estimator $E(\mathbf{x})$ and a DC sub-network DC. Each layer in $E(\mathbf{x})$ is composed of three types of operations: convolution (Conv), batch normalizing (BN), and rectified linear units (ReLU).

Table 1. Basic simulation parameters

Description	Notation	Value
Size of the domain	$l_d \times w_d$	2.5 m × 2 m
Size of imaging area	$l_i \times w_i$	1 m × 1 m
Thickness of the wall	d_w	$0.1\tilde{}$m ≤ d_w ≤ $0.6\tilde{}$m
The starting position of Tx	Tx	(0.5 m, 1.5 m)
The starting position of Rx	Rx	(0.64 m, 1.5 m)
Distance between Tx and Rx	d_a	0.14 m
Number of antenna pairs	N_a	43
Frequencies point	K	64
Central frequency of the transmit signal	f_t	0.5 GHz
Training data set	N_p	200
Training time	t_{train}	51 h

In order to tackle these issues, we will introduce a deep learning network to yield a physical model-driven deep network (CS-Net) in the next subsection.

CS-Net architecture

To overcome the limitations mentioned above, we propose an efficient and flexible iterative CNN-based framework CS-Net based on the physical model to reconstruct the target image. The structure of the CS-Net is shown in Fig. 2. The iterative network alternates between a CNN-based estimator sub-network $E(\mathbf{x})$ and a DC sub-network based on the conjugate gradient (CG) method. In order to enhance the network performance and effectively speed up the convergence rate, each layer in $E(\mathbf{x})$ is composed of three types of operations: convolution (Conv), batch normalizing (BN), and rectified linear units (ReLU). The first and the middle layers "CNN + BN + ReLU" use 64 filters with size 3 × 3 × 64. The last layer "CNN + BN" adopts 64 filters with size 3 × 3 × 64 to reconstruct the output image.

In Section "Signal model and compressed sensing," we have introduced that solving the problem of TWRI is to recover high-resolution image vector $\mathbf{x}_{\text{rec}}$ by (7). Here, we will solve the problem efficiently by introducing a CNN-based regularization prior, and the problem (7) can be reformulated as:

$$\mathbf{x}_{\text{rec}} = \arg\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|N_w(\mathbf{x})\|^2, \tag{8a}$$

$$N_w(\mathbf{x}) = \mathbf{x} - E(\mathbf{x}), \tag{8b}$$

where N_w is a learned CNN estimator of noise and alias patterns, which depends on the learned parameters w. λ is a trainable regularization parameter. $E(\mathbf{x})$ is a CNN-based estimator sub-network of CS-Net. It can be regarded as a denoising part of $\mathbf{x}$ to learn the redundancies radar echo signals and remove ghosts and noises.

Then, by introducing an auxiliary intermediate variable e_n, we can obtain an alternating iterative formulation to unroll the recursive network:

$$\mathbf{x}_{n+1} = \arg\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x} - e_n\|^2, \tag{9a}$$

$$e_n = E(\mathbf{x}_n). \tag{9b}$$

By calculating the gradient of problem (9a) and setting it to be zero, then we can obtain the following equation:

$$\mathbf{A}^{\text{H}}(\mathbf{A}\mathbf{x} - \mathbf{b}) + \lambda(\mathbf{x} - e_n) = 0. \tag{10}$$

Then problem (9a) can be solved by:

$$\mathbf{x}_{n+1} = \left(\mathbf{A}^{\text{H}}\mathbf{A} + \lambda \mathbf{I}\right)^{-1}\left(\mathbf{A}^{\text{H}}\mathbf{b} + \lambda e_n\right). \tag{11}$$

Problem (11) can be solved by using CG method. Furthermore, by regarding each iteration in CG scheme as a layer, it forms the DC sub-network in the whole network.

Taking each iteration as one layer, the above update rule can be regarded as an unrolled deep CNN called CS-Net. After initializing with $\mathbf{x}_0 = \mathbf{A}^{\text{H}}\mathbf{b}$, it alternatively updates e_n and x_{n+1} by CNN-based estimator step (9b) and CG step (11). The sub-network $E(\mathbf{x})$ learns the redundancies of the radar echo signal and remove ghosts and noises. Then, the DC sub-network, which encourages consistency with the measurements, fetches

https://doi.org/10.1017/S1759078722000071 Published online by Cambridge University Press

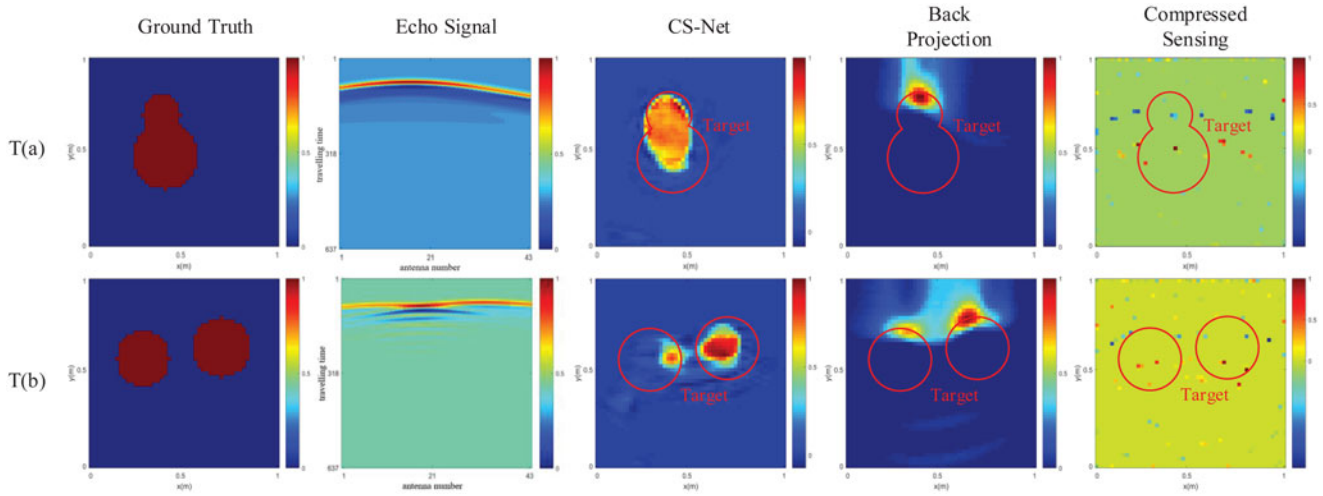


Fig. 3. Target texture detail recovery performance of the reconstructed target images achieved by CS-Net with WSE and CS, BP schemes. The first column shows the ground truth images for two representative scenarios and the other columns are the echo signal and the reconstruction results, respectively.

the input of $E(\mathbf{x})$ and the output of $E(\mathbf{x})$ to implement CG algorithm. We can obtain the radar targets reconstruction image after several iterations.

In the general compressive sensing problem model, matrix $\mathbf{A}$ is usually a measurement matrix without physical meaning. In our physical model driven CS-Net, matrix $\mathbf{A}$ represents the mapping from the imaging field to the measurement field in the radar detection scene. Since the choice of λ is affected by the spectrum of $\mathbf{A}$, in this study, λ is seen as one of the training parameters in the CS-Net. CS-Net learns the spectrum of $\mathbf{A}$ and update the λ and e_n to a suitable value during the training process. Hence, we set the initial value of $e_n = \mathbf{0}$ and $\lambda = 0.5$. The input of CS-Net is initialized to be $\mathbf{x}_0 = \mathbf{A}^{\mathrm{H}}\mathbf{b}$, and vector $\mathbf{b}$ represents the radar echo signal data after down-sampling in frequency domain. Once the network depth D_e and iteration number N_i is fixed, the update step in (9) and (11) and be viewed as an unrolled deep network in Fig. 2. Since each iteration uses the same solving structure, we use weight sharing mechanism at different iterations to reduce the number of network parameters and speed up the training time. Moreover, we use the same regularization parameter λ in each iteration to ensure the consistency of each iteration.

In order to balance the network performance and computational cost, similar to the parameter settings in [14], the depth of the network is set as 5 and the number of iterations is set as 6. After that, the network is trained by minimizing the loss function, which is specified as the MSE between $\mathbf{x}$ and the training label $\mathbf{l}$:

$$loss = \sum_{i=1}^{N_p} \|\mathbf{x}(i) - \mathbf{l}(i)\|_2^2, \tag{12}$$

where N_p is the number of training data sets. $\mathbf{x}(i)$ and $\mathbf{l}(i)$ denote the i-th output and the i-th label, respectively.

Simulation results

This section mainly shows the experimental details of CS-Net, the training process evaluation, and the comparison with the reconstruction results of other methods.

In the simulation experiments, we set the imaging area as a rectangle and the target as two cylinders. In order to obtain diversified sample data, the position is randomly distributed. Some specific simulation parameters about radar detection are given in Table 1. The radar echo signal data $\mathbf{R}$ is obtained by simulating the EM wave propagation in the set scene, and the ground truth of the target is obtained by simulating the distribution of the detected target as the label data of the CS-Net. After the transformation of radar echo data $\mathbf{R}$ in (6), the down-sampled frequency domain data $\mathbf{b}$ is obtained as the data set needed by the network. The network training data set is 200, and the test data set is 100. After the network training is completed, the down-sampled frequency domain data $\mathbf{b}$ in the test data is input into the network model to obtain the target reconstruction image in the corresponding scene.

In order to explore the performance limitation of the CS-Net, we implement wall signal elimination (WSE) to remove the echo signal from the wall by using time-gate method. Figure 3 depicts the reconstructed target images by the CS-Net with WSE and two traditional representative schemes. As shown in Fig. 3, CS-Net can recover most of the target texture. Reconstructed image with the BP method in T(a) can present the target position. However, the larger target in the reconstructed image with the BP method in T(b) is immersed by the echo signal from the smaller one. The reconstructed image with the CS method can reveal the target position. Nevertheless, similar to the BP method, the larger target in T(b) is immersed in the echo signal from the smaller one.

In order to reflect the anti-interference performance to the front wall of the CS-Net, we test the network with the echo signal without implementing WSE and compare the reconstructed result with those obtained by three traditional representative schemes. Based on the same setting in Fig. 3, the experimental data in Fig. 4 is the echo signal that do not implement WSE by using the time-gate method, and is marked as without WSE. Specifically, under the same scenario and parameter settings, T(a) in Fig. 3 includes the ground truth, echo signal, and the reconstructed target images of different methods after implementing WSE, while T(b) in Fig. 4 do not implement WSE, and obtained the corresponding results. The reconstructed images with the CS-Net still can present the target position and some of the related texture details with the wall eliminated. Meanwhile, the results with the LRSR method show the target

https://doi.org/10.1017/S1759078722000071 Published online by Cambridge University Press

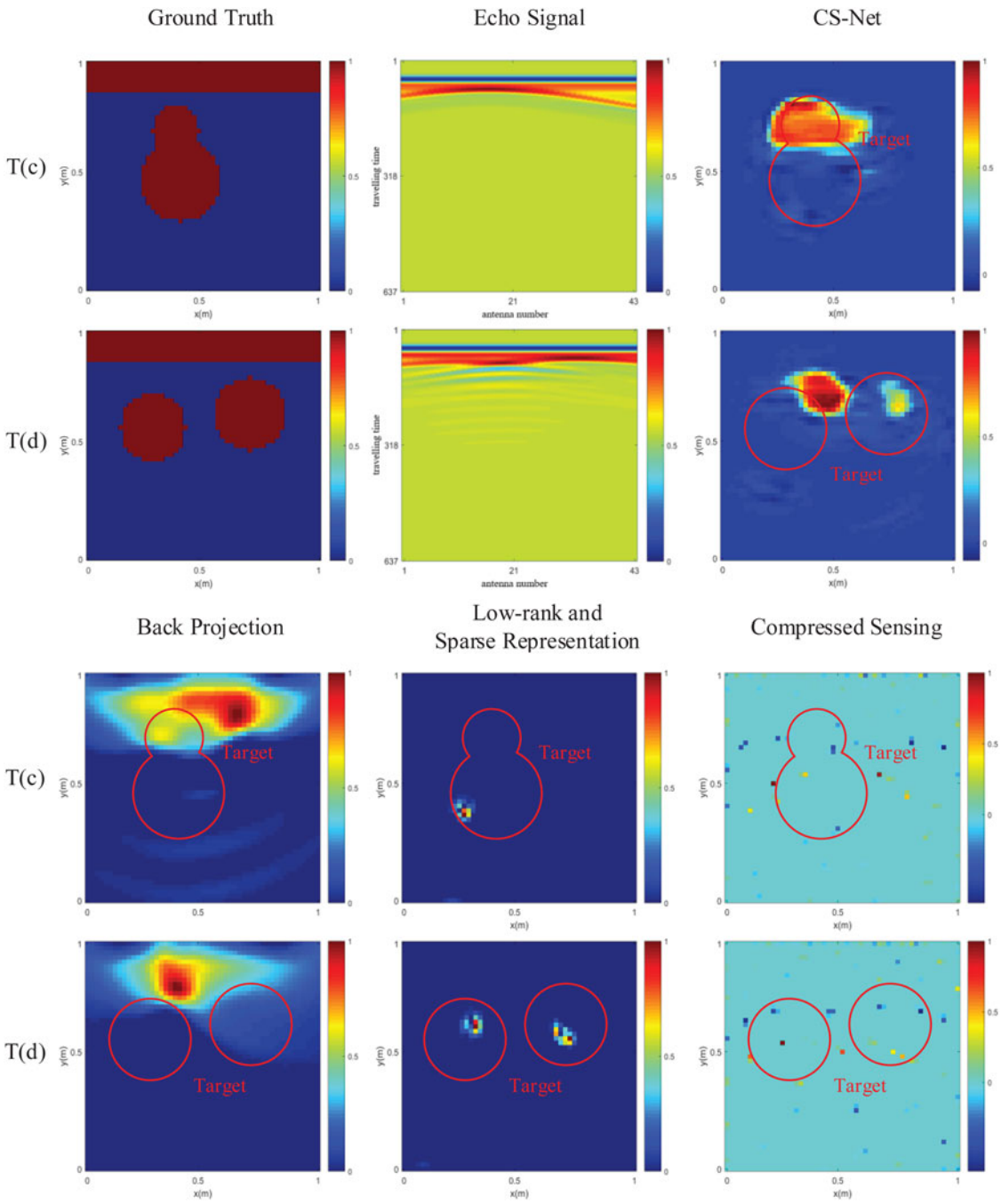


Fig. 4. Target texture detail recovery performance of the reconstructed target images achieved by CS-Net without WSE and CS, BP, LRSR schemes. The first column shows the ground truth images for two representative scenarios and the other columns are the echo signal and the reconstruction results, respectively.

position and number without the target texture details, since it reconstructs the target image based on the echo signal from the front wall and targets. Comparatively, the results with the BP and CS methods even cannot reflect the target position due to the interference from the front wall.

To better reflect the performance of the reconstructed image with CS-Net in the TWRI system, without generality, we use MSE and program running time (PRT) as an evaluation index. MSE is used to evaluate the error between the ground truth and the reconstruction result, and PRT is used to reflect imaging time. Since the output vector from each method is in different value scales, we normalize the i-th value x_i in the output vector $\mathbf{x} = (x_1, x_2, \ldots, x_i, \ldots x_N)^{\mathrm{T}}$ to $x_i \in [0, 1]$ as shown below:

$$x_i = \frac{x_i - \mathbf{x}_{\min}}{\mathbf{x}_{\max} - \mathbf{x}_{\min}}, \tag{13}$$

where $\mathbf{x}_{\max}$ and $\mathbf{x}_{\min}$ denote the maximum and minimum value of the vector $\mathbf{x}$, respectively.

Table 2 compares the performance achieved by CS-Net with those obtained by the traditional method based on CS, BP, and LRSR schemes. Table 2 shows that the MSE of the CS-Net is smaller than those with CS, BP, and LRSR schemes, both in with WSE and without WSE configurations. It validates that the quality of

https://doi.org/10.1017/S1759078722000071 Published online by Cambridge University Press

Table 2. MSE performance of target image reconstruction achieved by CS-Net and those obtained by CS, BP, and LRSR schemes

Method	T(a)	T(b)	T(c)	T(d)
CS-Net	0.054	0.088	0.103	0.131
CS	0.173	0.206	0.141	0.145
BP	0.121	0.149	0.146	0.191
LRSR	–	–	0.119	0.137

Table 3. PRT target image reconstruction achieved by CS-Net and those obtained by CS, BP, and LRSR schemes

Method	T(a)	T(b)	T(c)	T(d)
CS-Net	1.523	1.528	1.586	1.574
CS	45.447	45.826	48.641	49.275
BP	2.084	2.143	2.492	2.427
LRSR	–	–	18.924	19.468

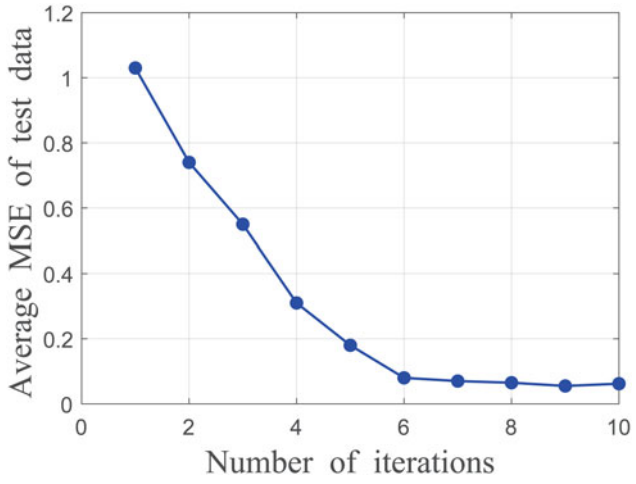


Fig. 5. MSE as a function of number of iterations. As the number of iterations of the network increases during training, the average MSE of 100 sets of test data changes. It shows that MSE stabilizes around six iterations.

the reconstructed image with the proposed CS-Net is better than those with the traditional scheme. Moreover, the MSE of those with WSE configuration in T(a) and T(b) is smaller than those without WSE in T(c) and T(d). Thus, the echo signal from the wall can largely affect the quality of the reconstructed image and it is necessary to eliminate the echo signal from the wall before implementing the reconstruction method.

Table 3 compares the imaging time achieved by CS-Net with those obtained by the traditional method based on CS, BP, and LRSR schemes. Table 3 shows that the PRT of the CS-Net is shorter than those with CS, BP, and LRSR schemes, both in with WSE and without WSE configurations. It validates that CS-Net can reconstruct the target image faster than the traditional algorithm, and has great advantages in imaging time.

Figures 5 and 6 show the evaluation in the network training process. Figure 5 displays the influence of the number of alternating iterations of $E(\mathbf{x})$ and DC on MSE. It can be observed that as the increase of network iterations, MSE gradually decreases to around six iterations to reach a stable value. Figure 6 indicates

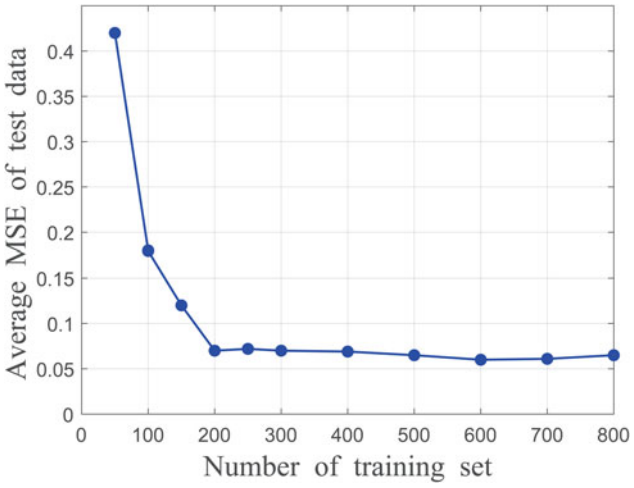


Fig. 6. MSE as a function of number of training set. As the number of training set increases during training, the average MSE of test data changes. It shows that MSE stabilizes around 200 sets of test data.

the impact of the number of network training sets on MSE. It can be observed that MSE tends to be stable when the training data set is around 200. This shows that CS-Net can use less training data to achieve faster network convergence speed.

Conclusion

In this paper, a physical model-driven deep network, termed CS-Net, was proposed to estimate the target image of a TWRI system. The proposed method introduced a regularization prior based on CNN and alternately iterates with the DC layer to form a physical model-driven network solution model, without manually adjusting algorithm parameters. Simulation results have shown the superiority of our proposed CS-Net estimation scheme in comparison with the existing traditional schemes. Compared with the traditional iterative solution method, the deep learning method proposed in this study can significantly improve the speed of reconstructing the target image and accurately reconstruct the target image and show excellent computational performance in TWRI.

References

1. **Muqaibel AH** (2015) Improved compressive sensing with antenna directivity for TWRI. in Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting, 402–403.
2. **Sun HB** (2019) Through-the-wall human motion sensing based on forward scattering. in 2019 IEEE Radar Conference, 1–5.
3. **Laviada J, Arboleya A, Lopez-Gayarre F and Las-Heras F** (2021) Broadband synthetic aperture scanning system for three-dimensional through-the-wall inspection. *IEEE Geoscience and Remote Sensing Letters*, **13**, 97–101.
4. **Tang VH, Bouzerdoum A and Phung SL** (2020) Compressive radar imaging of stationary indoor targets with low-rank plus jointly sparse and total variation regularizations. *IEEE Transactions on Image Processing*, **29**, 4598–4613.
5. **Tivive FHC, Bouzerdoum A and Amin MG** (2015) A subspace projection approach for wall clutter mitigation in through-the-wall radar imaging. *IEEE Transactions on Geoscience and Remote Sensing*, **53**, 2108–2122.
6. **Yoon YS and Amin MG** (2009) Spatial filtering for wall-clutter mitigation in through-the-wall radar imaging. *IEEE Transactions on Geoscience and Remote Sensing*, **47**, 3192–3208.

https://doi.org/10.1017/S1759078722000071 Published online by Cambridge University Press

7. **Zhang Y and Xia T** (2016) In-wall clutter suppression based on low-rank and sparse representation for through-the-wall radar. *IEEE Geoscience and Remote Sensing Letters*, **13**, 671–675.
8. **Zhang Q, Chen YJ, Chen YG, Chi L and Wu Y** (2015) A cognitive signals reconstruction algorithm based on compressed sensing. in 2015 IEEE 5th Asia-Pacific Conference on Synthetic Aperture Radar, 724–727.
9. **Lei L, Huang JJ and Sun Y** (2016) Compressed sensing MIMO radar waveform optimization without signal recovery. 2016 CIE International Conference on Radar, 1–4.
10. **Yan HC, Xu J and Zhang XD** (2015) Compressed sensing radar imaging of off-grid sparse targets. 2015 IEEE Radar Conference, 0690–0693.
11. **Unser M, Fageot J and Gupta H** (2016) Representer theorems for sparsity-promoting ℓ_1 regularization. *IEEE Transactions on Information Theory*, **62**, 5167–5180.
12. **Hosseini MS and Plataniotis KN** (2014) High-accuracy total variation with application to compressed video sensing. *IEEE Transactions on Image Processing*, **23**, 3869–3884.
13. **Jin T** (2017) Autofocus compressed sensing imaging based on nonlinear conjugate gradient. in 2017 XXXIInd General Assembly and Scientific Symposium of the International Union of Radio Science, 1–4.
14. **Aggarwal HK, Mani MP and Jacob M** (2019) MoDL: model-based deep learning architecture for inverse problems. *IEEE Transactions on Medical Imaging*, **38**, 394–405.
15. **Sanghvi Y, Kalepu Y and Khankhoje UK** (2020) Embedding deep learning in inverse scattering problems. *IEEE Transactions on Computational Imaging*, **6**, 46–56.
16. **Chen XD, Wei Z, Li MK and Rocca P** (2020) A review of deep learning approaches for inverse scattering problems. *Progress in Electromagnetics Research*, **167**, 67–81.
17. **Zheng ZJ, Pan J, Ni ZK, Shi C, Ye SB and Fang GY** (2021) Human posture reconstruction for through-the-wall radar imaging using convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, **19**.
18. **Zhang HY, Song RY, Chen SY, Wang G, Jia Y and Yan C** (2019) Target Imaging Based on Generative Adversarial Nets in Through-wall Radar Imaging. in International Conference on Control, Automation and Information Sciences (ICCAIS), 1–6.
19. **Li HQ, Cui GL, Guo SS, Kong LJ and Yang XB** (2021) Human target detection based on FCN for through-the-wall radar imaging. *IEEE Geoscience and Remote Sensing Letters*, **18**, 1565–1569.
20. **Tivive FHC and Bouzerdoum A** (2021) Clutter removal in through-the-wall radar imaging using sparse autoencoder with low-rank projection. *IEEE Transactions on Geoscience and Remote Sensing*, **59**, 1118–1129.
21. **Li MC, Xi XL, Zhang XH and Liu GH** (2021) Joint compressed sensing and spread spectrum through-the-wall radar imaging. *IEEE Access*, **9**, 6259–6267.



Yuhao Wang received his Ph.D. degree from Wuhan University, Wuhan, China, in 2006. He is currently a Professor with the Cognition Sensor Network Laboratory, School of Information Engineering, Nanchang University (NCU), Nanchang, China. He is the Dean of the Artificial Intelligence Industry Institute, NCU, and also is the Head of Jiangxi Embedded Systems Engineering Research Center. His current research interests include wideband wireless communication and radar sensing fusion systems, channel measurement and modeling, nonlinear signal processing, smart sensors, image and video processing, and machine learning, and visible light communication.



Yue Zhang received her bachelor's degree from the East China University of Technology, Nanchang, China. She is currently pursuing her master's degree with the School of Information Engineering, Nanchang University, Nanchang, China. Her research interests include radar sensor signal processing, through-the-wall radar imaging methods, and deep learning.



Mingcheng Xiao received his bachelor's degree and his master's degree from the School of Information Engineering, Nanchang University, Nanchang, China. His research interests include through-the-wall radar, MIMO radar imaging, and cognitive radio sensor networks.



Huilin Zhou received his Ph.D. degree in information engineering from Wuhan University, Wuhan, China. He is currently a Professor with the School of Information Engineering, Nanchang University, Nanchang, China. His research interests include radar systems, radar signal processing, and radar imaging.



Qiegen Liu received his B.S. degree in applied mathematics from Gannan Normal University, Ganzhou, China, his M.S. degree in computation mathematics, and his Ph.D. degree in biomedical engineering from Shanghai Jiaotong University, Shanghai, China. Since 2012, he has been with the School of Information Engineering, Nanchang University, Nanchang, China, where he is currently a Professor. During 2015–2017, he was also a Postdoc with UIUC and University of Calgary, Calgary, AB, Canada. His current research interests include sparse representations, deep learning and their applications in image processing, computer vision, and MRI reconstruction.



Jianfei Gao received his bachelor's degree from the College of Applied Science, Jiangxi University of Technology, Ganzhou, China. He is currently pursuing his master's degree with the School of Information Engineering, Nanchang University, Nanchang, China. His research interests include radar imaging, clutter suppression, and deep learning.

https://doi.org/10.1017/S1759078722000071 Published online by Cambridge University Press