Communication

Deep Learning Electromagnetic Inversion Solver Based on a Two-Step Framework for High-Contrast and Heterogeneous Scatterers

He Ming Yao, Michael Ng, and Lijun Jiang

***Abstract*— This communication proposes a novel electromagnetic (EM) inversion solver based on a two-step deep learning (DL) framework. The framework consists of the deep convolutional asymmetric encoder–decoder structure (DCAEDS) followed by the deep residual convolutional neural network (DRCNN). In the first step, DCAEDS utilizes EM scattered field data from a single-frequency one-time measurement to coarsely retrieve the initial contrasts (permittivities) of target scatterers. In the second step, DRCNN employs a mixed input scheme, comprising the initially reconstructed permittivities from the first step and the original EM scattered field data, to significantly improve the retrieved contrasts (permittivities) and refine the reconstruction of targets. Consequently, the proposed EM inversion solver achieves excellent accuracy and efficiency, even for high-contrast targets. The proposed solver is flexible as it is required only for a single-frequency one-time measurement on the EM scattered field. Moreover, the proposed two-step DL-based solver overcomes the limitations of conventional methods, such as high computational costs and ill-posedness. Numerical benchmarks based on various dielectric objects demonstrate the feasibility of the proposed EM inversion solver, highlighting its potential as a candidate for real-time quantitative EM inversion for high-contrast targets.**

***Index Terms*— Convolutional neural network, electromagnetic (EM) inverse scattering, high contrast, residual learning, two-step process.**

I. Introduction

Real-time electromagnetic (EM) inversion [1], [2], [3] has been a significant challenge in research fields such as subsurface sensing [4], microwave remote sensing [5], and biomedical imaging [6]. Over the past few decades, various important methods have been reported to solve the EM inverse scattering (EMIS) problem, including contrast source inversion methods [7], [8], Born iterative method [9], contrast-source extended Born [10], Gauss–Newton methods [2], and subspace optimization method [11]. However, these conventional methods typically require iterative computation to optimize the objective function and need to compute complex Green's functions, leading to longer computational time and higher memory costs. Moreover, conventional methods usually require tedious measurement operations and a considerable amount of measurement data to achieve better reconstruction performance [1], [2], [3], [4], [5], [6], [12]. Normally, obtaining better reconstruction results for the unknown scatterers requires many incident waves, which imposes heavy burdens on measurement and raises the reconstruction computation complexity [1], [2], [3], [12]. Because of these difficulties, the conventional methods for EMIS hardly realize teal-time EM inversion.

The rapid development of artificial intelligence (AI) techniques, including machine learning (ML) [13] and deep learning (DL) [14], has significantly contributed to computational EMs (CEM) research, including EM computation [15], [16], [17], hybrid field-circuit simulations [18], [19], and remote sensing [20], [21]. Indeed, DL strategies have been employed to enhance traditional methodologies in this field and provide promising performances for solving EMIS problems, as presented in [22], [23], [24], and [25]. The current DL-based methods for realizing EM inversion can be summarized into two major categories: 1) the "black-box"-based methods [22], [23], [24], [25], [26], where the EM fields or initially reconstructed contrasts are the input and the contrasts of targeted scatterers are selected as the output; 2) physically inspired methods [27], [28], where physics-based prior information can be integrated into the model during the training process. Under normal circumstances, these DL-based methods can reduce computation costs and add more prior information for realizing EM inversion. However, nearly all these DL-inspired methods rely on multi-incident waves and multiple measurements (multiple incident angles and even multiple frequencies [22], [23], [24], [25], [26], [27], [28]) to conduct the measurement/simulation for both training and testing, resulting in large quantities of measurement data. The tedious use of multiple measurements undoubtedly increases the complexity of preparing training data for DL models [22], [23], [24], [25], [26], as well as the size of input for DL models, leading to unnecessary computation costs during their application [22], [23], [24], [25], [26]. Moreover, these DL-based methods only utilize single-step DL models, most of which cannot operate independently of conventional methods. Typically, they either substitute segments of the process in conventional methods or necessitate initial inputs derived from conventional methods, such as the backpropagation (BP) method [24]. Although some so-called physically inspired methods, including supervised descent method (SDM) [27], [28], attempt to avoid using DL models, they still need to create training data based on large quantities of measurements.

AI techniques offer the potential to break down complex tasks into several independent modules, each with its own objective (e.g., image recognition [29]). These modules can then be combined to effectively complete the complex tasks [13], [14]. While most existing DL techniques for solving EMIS problems rely on a single DL model to reconstruct contrast (permittivity), EMIS problems can potentially be divided into several subtasks. In fact, conventional optimization methods for solving EMIS problems can also be viewed as iteratively optimizing a series of subtasks [13], [14]. In this work, we propose a novel EM inversion solver based on a two-step DL framework, which differs from the structures presented in [22], [23], [24],

Manuscript received 14 August 2022; revised 7 August 2023; accepted 11 September 2023. Date of publication 12 March 2024; date of current version 7 June 2024. This work was supported in part by Hong Kong Research Grant Council General Research Fund (GRF) under Grant 12300218, Grant 12300519, Grant 17201020, Grant 17300021, Grant C1013-21GF, and Grant C7004-21GF; in part by the Joint Natural Science Foundation of China—Research Grants Council (NSFC-RGC) under Grant N-HKU76921; and in part by the Research Grants Council of Hong Kong Special Administrative Region, China, under Grant HKU PDFS2122-7S05. *(Corresponding author: Michael Ng.)*

He Ming Yao is with the Department of Materials, Imperial College London, SW7 2BX London, U.K. (e-mail: yaohmhk@connect.hku.hk).

Michael Ng is with the Department of Mathematics, Hong Kong Baptist University, Hong Kong, China (e-mail: michael-ng@hkbu.edu.hk).

Lijun Jiang is with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, China (e-mail: jianglj@hku.hk).

Color versions of one or more figures in this communication are available at https://doi.org/10.1109/TAP.2024.3372772.

Digital Object Identifier 10.1109/TAP.2024.3372772

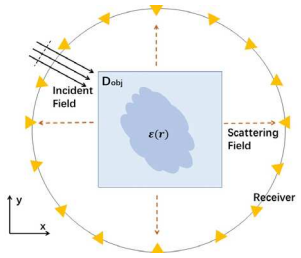

Fig. 1. Schematic of TM_z wave scattering from D_{obj}.

and [26]. Our approach consists of a deep convolutional asymmetric encoder–decoder structure (DCAEDS) followed by a deep residual convolutional neural network (DRCNN). In the first step, DCAEDS uses EM scattered field data from a single measurement to roughly retrieve initial contrasts of target scatterers. In the second step, DRCNN employs a mixed input scheme, which includes the initially reconstructed contrasts from the first step and the original EM scattered field data, ultimately improving the target reconstruction as the final output.

Compared with existing conventional methods and DL-based methods [3], [26], [27], the proposed two-step DL framework provides several advantages for a fast EM inversion solver as follows.

1) *Simplicity:* The application of our DL solver requires only a single far-field simulation/measurement at a single frequency, which significantly reduces the size of input data and simplifies the measurement operation. This enables real-time EM inversion. In addition, training for both the proposed models can utilize identical dataset, created using a simple synthetic dataset, thus avoiding extra costs for creating training data.
2) *Effectiveness:* Compared to conventional methods, our DL EM inversion solver can effectively handle heterogeneous scatterers with extremely high contrasts.
3) *Flexibility:* We can incorporate a broader scope of prior knowledge in a more adaptable way. This is achievable by constructing the training dataset, including acquired prior during the offline phase.
4) *Accuracy:* The methodology we introduce surpasses traditional techniques in accuracy when addressing EMIS challenges, especially in scenarios involving scatterers with high-contrast and heterogeneous properties.

II. Problem and Formulation

A. Problem Formulation

Fig. 1 illustrates a typical use case for EMIS issue, showcasing a 2-D transverse magnetic (TM) wave, denoted as E^{in}, impinging upon the objective domain D_{obj}. Surrounding D_{obj}, M receivers are uniformly positioned to record the scattered wave E^s. D_{obj} is methodically divided into a grid of $N \times N$ uniform segments. The EMIS phenomenon is encapsulated by two principal, (1) and (2), commonly referred to as the Lippmann–Schwinger equations [1], [2], [3], [4], [5], [6], [30]. These equations encapsulate the intricate interplay between the scattered field E^s, the incident field E^{in}, and the cumulative EM field E^t. Specifically, (1) captures the dynamic interactions occurring within the scatterer elements constituting D_{obj}

$$E^t(\boldsymbol{r}) = E^{in}(\boldsymbol{r}) + k_0^2 \int_{D_{obj}} G(\boldsymbol{r}, \boldsymbol{r'}) \chi(\boldsymbol{r'}) E^t(\boldsymbol{r'}) d\boldsymbol{r'} \quad (1)$$

where $G(\boldsymbol{r}, \boldsymbol{r'})$ is the Green's function. For the TM_z wave, it can be denoted as $G(\boldsymbol{r}, \boldsymbol{r'}) = -(j/4)H_0^{(2)}(k_0|\mathbf{r} - \mathbf{r}'|)$. $H_0^{(2)}$ is the Hankel function of the zeroth order of the second kind, while k_0 is the wavenumbers in the free space. $\boldsymbol{r} = (x, y)$ and $\boldsymbol{r'} = (x', y')$ are, respectively, the field and source points in D_{obj}. The contrast function is defined as $\chi(\mathbf{r}') = \varepsilon_r(\boldsymbol{r}') - 1$. The following equation describes the relationship between E^s and E^t:

$$E^s(\boldsymbol{r}) = k_0^2 \int_{D_{obj}} G(\boldsymbol{r}, \boldsymbol{r'}) \chi(\mathbf{r}') E^t(\mathbf{r}') d\boldsymbol{r'} \quad (2)$$

where $\boldsymbol{r} = (x_R, y_R)$ denotes the receiver locations and $\boldsymbol{r'} = (x', y')$ represents the coordinates for segments within D_{obj}. The target of EM inversion is to reconstruct χ of scatterers by E^s. Traditional approaches are typically characterized by an optimization routine, as formulated in (3), which seeks to deduce unknown parameters by reducing discrepancy between empirical data and predictions of the refined model. Within this framework, $f(\chi)$ is defined to facilitate the optimization endeavor. Regrettably, the extraction of χ based on (3) often manifests as nonlinear operation, especially in scenarios involving high contrast [1], [2], [3]

$$\min: f(\chi) = \sum_{i=1}^{N_i} \| E_i^s - E_i^s(\chi)' \| + \alpha L(\chi) \quad (3)$$

where the received E_i^s induced by N_i different incident field E^{in} is approached by the optimized scattered fields $E_i^s(\chi)'$ by iterative calculation in (3). In this context, α represents the fixed coefficient for regularization, with L denoting the regularization term itself [1].

B. Two-Step DL Framework

To address the challenges posed by conventional methods in achieving EM inversion, we propose a two-step DL framework that enables high-precision EM inversion even for high-contrast scatterers. In this novel framework, the EM scattered field, obtained from a single-frequency measurement using one transmitter, is first transformed into an initial rough reconstruction. Building upon this rough reconstruction, the second step further refines the reconstruction quality and provides the final prediction of this EM inversion solver. Given that the requirement for many training samples is difficult to meet through real-world experiments, we employ simulation data for training and testing the proposed DL-based solver [22], [23], [24], [25]. The newly developed DL framework for EM inversion can be summarized in the following two steps.

Step 1 (DCAEDS for Retrieving the Initial Contrast): The first step converts the measured EM scattered field to a preliminary contrast prediction of target scatterers. It can be concluded as a process that transforms the received EM scattered field into the output initially retrieved results of targets. In contrast to traditional iterative approaches for addressing EMIS problem [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], the proposed first step can be seen as the initial value selection process, where the initial contrast χ_0 is calculated for the subsequent iterative optimization process in (3). In fact, it is a challenge for the conventional methods to provide valid initial contrast χ_0 for the iterative optimization process, particularly in high-contrast scatterer cases. Generally, the reconstruction from the conventional methods [1], [2], [3] hardly provide meaningful information if only employing one-time measurement on the target.

In the proposed first step, we employ DCAEDS, which uses the ground-truthed contrast χ of the scatterer as the output, while the original scattered fields E^s are selected as its input. Due to the complexity of this physics-based problem, the size of the measured data often differs from that of the target [1], [2], [3]. In some extreme cases, researchers can only obtain limited measured data with a completely different size from the target [4], [5], [6]. Most existing DL models for EM inversion usually select an input size that is larger or approximated to the output size, in order to enrich the information utilized by the models. This approach, however, leads to tedious measurements to acquire large amounts of EM scattered

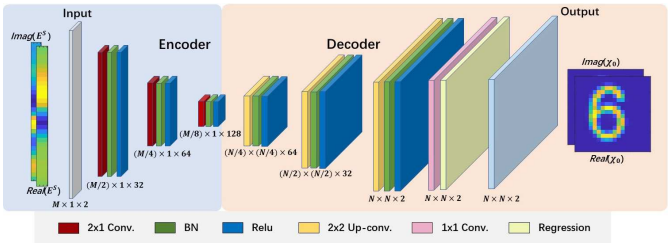

Fig. 2. Architecture of DCAEDS for the first step. The complete spellings of the abbreviations are: Conv.: convolution, BN: batch normalization, ReLU: rectified linear unit, and Up-conv.: up-convolution.

field data and increases the complexity of the DL model [22], [23]. Unlike these models, the application of DCAEDS allows for a smaller input size compared to the output size, providing more flexibility in terms of input and output dimensions. This greatly simplifies the measurement process and reduces the complexity of the model.

The proposed DCAEDS is revised based on SegNet [31], which is extensively applied in the image segmentation field. Fig. 2 presents the internal structure of the proposed DCAEDS for the first step, which can be divided into two modules in general: encoder and decoder. Different from the conventional SegNet, DCAEDS narrows down its encoder and makes its encoder and decoder not symmetric anymore, which brings much convenience to the application [32], [33]. Sequential layers within the encoder module process the 1-D EM scattered field, denoted as E^s, compressing it into information-rich segments. The inputs have the size of $M \times 1 \times 2$, where the row M represents M receivers, one column stands for only one incident EM testing wave, and two tubes are the real and imaginary parts of E^s. Within the decoder module, the previously isolated informational segments are reconstructed to formulate the scatterer's contrast χ_0. Its outputs are the $N \times N \times 2$ matrix composed of the real and imaginary parts of the contrast (permittivities) of targets. In detail, the encoder module repeatedly conducts convolution (Conv), rectified linear unit (ReLU), and batch normalization (BN). The decoding module utilizes a series of up-convolution (Upconv) processes, BN, and ReLU activations in sequence. Within the encoding module, various layers are designed to extract and encode salient features from the input. Specific details concerning the number of convolutional layers, and the quantity and dimensions of filters (kernels) are delineated in Fig. 2. Concluding the decoding sequence, the architecture incorporates a 1×1 convolution layer followed by a regression layer, culminating in the predicted contrast profile, denoted as χ_0. Here, we need to emphasize that the size of input (i.e., $M \times 1$) from one-time measurement is much smaller than that of the output (i.e., $N \times N$), because of the application of DCAEDS. The mean square error (mse) is the loss function [22], [23], [24], [25], [26].

Step 2 (DRCNN for Refining the Initial Contrast): The second step is actually the process to enhance the originally reconstructed contrast from the first step to the accurate final result. Thus, this step perfectly replaces the iterative optimization process utilized for the conventional methods for EMIS problems [2], [7], [8], [9], [10], [11], [12]. In the second step, DRCNN utilizes the mixed input scheme as the input, consisting of the initially reconstructed permittivities (contrasts) from the first step and the originally measured EM scattered field data. Based on this mixed input, DRCNN finally realizes the reconstruction refinement and provides the final prediction of the target scatterers. Thus, the input of DRCNN with the size of $N \times N \times 4$ is the coarsely reconstructed contrast χ_0 from the previous step and the EM scattered field data E^s, while its corresponding ground-truthed contrast χ is chosen as the output with the size of $N \times N \times 2$. To keep the size uniformity for four tubes of the input, the real part of E^s with the size as $M \times 1$ is interpolated to the dimension as $N \times N$, where the $M \times 1$ vector is first interpolated into $N \times 1$ vector, and then, the interpolated $N \times 1$ vector is copied N times to form the $N \times N$ "scattering image" for the real part of E^s. The same operation has also been done to the imaginary part of E^s to obtain the "scattering image" with the size of $N \times N$ for its imaginary part. In this way, we could easily and efficiently provide input to DRCNN.

The proposed DRCNN is based on the so-called U-Net structure [34], consisting of three parts: encoding, bridging, and decoding. The encoding part encodes the input contrast images into compact representations, while the corresponding decoding part recovers the representations. The middle bridging part acts as the bridge connecting the other two mentioned paths. In addition, unlike the conventional U-net, only the skip connection is exerted for residual learning between the other two mentioned parts. The encoding architecture is designed with multiple iterations of Conv, BN, and ReLU in sequential applications. On the counterpart, the decoding part is equipped with the repetitive applications of Upconv, BN, and ReLU operation, shown in Fig. 3.

Therefore, the second step functions as the iterative refining process in conventional iterative optimization approaches [2], [7], which refine the contrast by optimization process in (3). For the second process, the relationship between E^s, the coarsely reconstructed contrast χ_0, and the final refined contrast χ can be described as (4), where $\mathcal{F}$ represents the nonlinear operation in DRCNN

$$\chi = \mathcal{F}\left(\left[\chi_0, E^s\right]\right). \tag{4}$$

Several special issues about the proposed two models for EM inversion should be emphasized.

1) Mixed Input Scheme for DRCNN: The input of the proposed DRCNN is mixed by two kinds of data: the first part is the initial reconstruction computed from DCAEDS and the second part is based on the received EM scattered field data. In this way, both the initial estimation from DCAEDS and the "raw" EM scattered field data can be involved together to be utilized as the input by DCAEDS. Due to the application of DCAEDS, only one-transmitter-measured EM scattered field data (much smaller than the size of expected reconstruction image) is required for both DCAEDS and DRCNN. In addition, EM inversion challenges are inherently complex-valued issues, where the real and imaginary parts can be adjusted independently across different channels. This distinction enhances the versatility and adaptability of two DL models for practical application requirements.

2) Computational Complexity: Within DCAEDS and DRCNN, the computational burden is primarily governed by convolutional processes, BN, and the ReLU [22], [23], [24], [25]. Owing to the minimal size of the filter operations, the predominant computational expense is attributed to convolutional activities [35], [37]. The input dimension for DRCNN is established at $N \times N \times 4$, and it outputs a dimension of $N \times N \times 2$. Each layer employs R convolutional filters, each with dimensions $K \times K$, and the total count of layers is denoted by f. Thus, the computational complexity for the DRCNN can be expressed as $O(N^2K^2R^2f)$ [35], [37]. On the counterpart, the input and the output sizes of DCAEDS are $M \times 1 \times 2$ and $N \times N \times 2$ respectively. For the decoder in DCAEDS, the architecture of each layer incorporates R convolutional filters, each measuring $K \times K$, with the entire network comprising f layers in total. As a result, the decoder's computational complexity can be quantified as $O(N^2K^2R^2f)$ [35], [37]. Because the size of the required EM scattered field data is much smaller than that of the expected reconstruction (i.e., $M \ll N^2$), the computational complexity of DCAEDS can be approximated as $O(N^2K^2R^2f)$. Moreover, the configuration of two DL models, namely, the DCAEDS and DRCNN, is characterized by similar parameters. Consequently, the computational complexity of both models is approximately equivalent. Thus, the computation complexity of the entire two-step framework can be

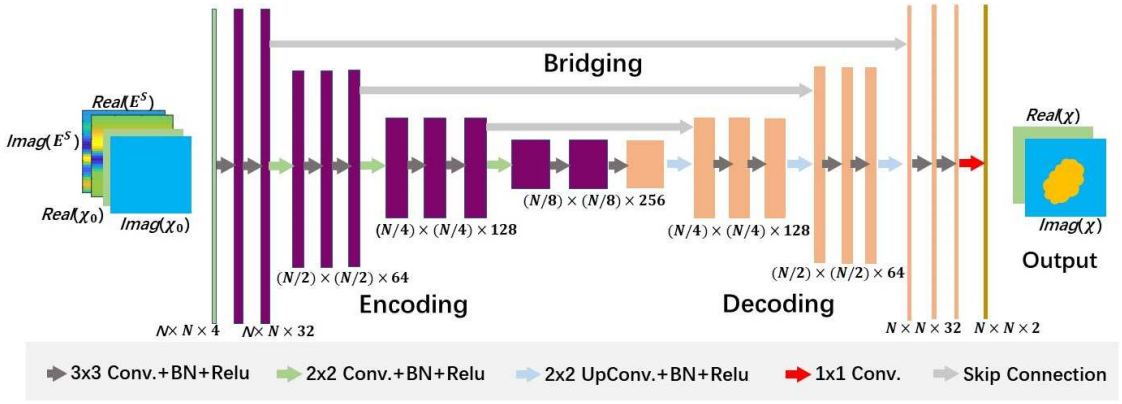

Fig. 3. Architecture of DRCNN for the second step. The complete spellings of the abbreviations are: Conv.: convolution, BN: batch normalization, ReLU: rectified linear unit, and Up-conv.: up-convolution.

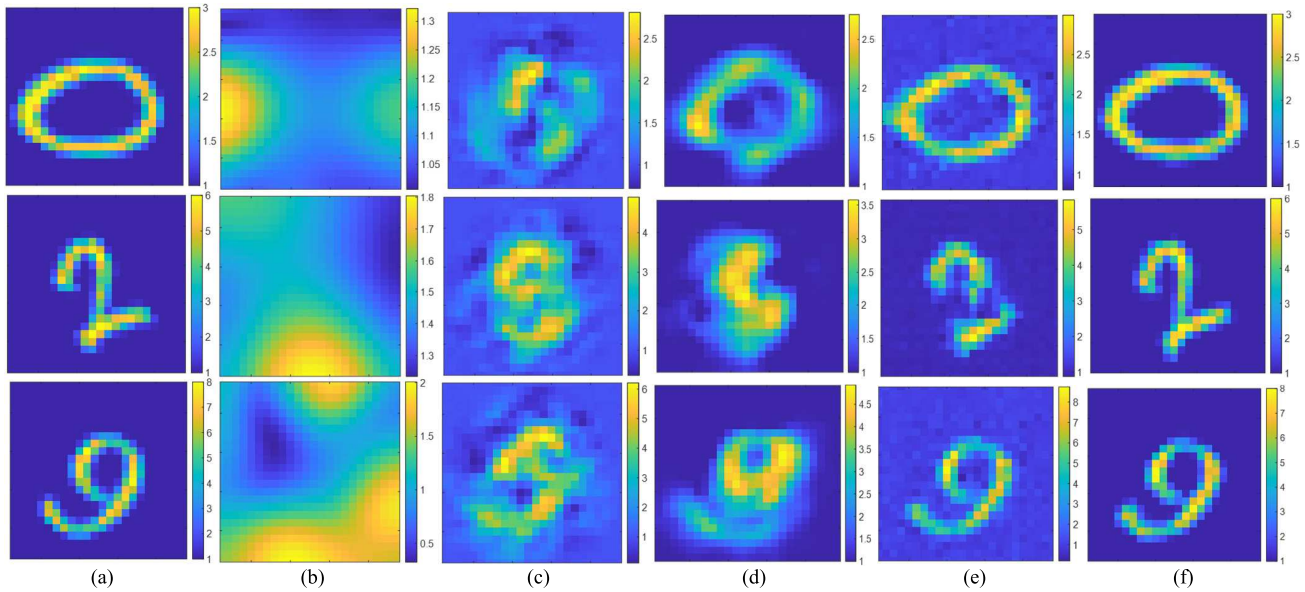

Fig. 4. Comparison of reconstruction of number-shaped targets (the maximum of randomly selected permittivities is: 3 for first row, 6 for second row, and 8 for third row). (a) Ground truths. (b) Reconstruction from Gauss–Newton method. (c) Reconstruction from MCNN. (d) Reconstruction from SDM. (e) Reconstruction from DCAEDS in the first step. (f) Reconstruction from DRCNN in the second step.

$O(N^2K^2R^2f)$. In addition, the demand for memory complexity in both DL models is represented as $O(N^2K^2R^2f)$ because it primarily influenced by the dimensions of the filters and biases [37].

III. Numerical Results

A. Numerical Setup and Off-Line Training

The training data for the proposed DCAEDS and DRCNN are created based on only one group of simple synthetic dataset, i.e., MNIST [22], [23], [24], [25], [26], [27], [28], [29]. As shown in Fig. 1, the size of each sample from MNIST is set to $\lambda \times \lambda$ (the wavelength $\lambda = 1$ m in free space) with the nonhomogeneous relative permittivity ε_r. Each sample is meshed into 24×24 uniform pieces (i.e., $N = 24$). The number of receivers uniformly set around D_{obj} is $M = 24$, of which the distance is 30λ. Only one TM_z incident plane wave illuminates D_{obj} with the incident angle equal to 0° (i.e., $N_i = 1$). In this section, both training and testing datasets are created by full-wave EM simulations [38]. Unlike the scatterers with homogeneous contrast [22], the relative permittivity ε_r of the used samples is heterogeneous rather than being homogeneous (constant). These heterogeneous digit-shaped scatterers from MNIST have the relative permittivity ε_r ranging from 1 to 8, which is very challenging for the conventional methods [2], [4], [5], [6], [7], [8], [9], [10], [11], [12]. Based on MNIST, 5000 samples are randomly chosen to create the training data. Two quantitative indicators, i.e., normalized mean-square error (NMSE) and structural similarity index (SSIM), are employed to evaluate the reconstructions [22], [23], [24], [25], [26], [27], [28], [29]. In addition, for comparative purposes, the Gauss–Newton method [1], [2], [3], [4] has been applied to identical test samples illustrated in Sections III-B and C. The innovative two-step DL framework integrates DRCNN with DCAEDS, effectively achieving EM inversion through two distinct yet cohesive modules. The proposed DL models are implemented into DL Toolbox in MATLAB 2021a [39], while adaptive moment estimation (Adam) optimizer is used to optimize the mentioned loss functions [40].

B. Performance on Number-Shaped Scatterers

Thousand new samples in the MNIST dataset are randomly chosen in this section to test the trained two-step DL framework, while 5% white noise is added into the produced EM scattered field. Fig. 4 displays a comparative analysis showcasing the actual ground truth, the reconstruction outcomes achieved through Gauss–Newton, and the reconstruction using our two-step solver, where the maximum of nonhomogeneous permittivities for the samples is randomly selected as 3, 6, and 8. While the initial reconstruction from DCAEDS can provide meaningful information about target, the results from the following DRCNN agree very well with the ground truth. However, the Gauss–Newton method fails to realize reconstruction of the mentioned scatterers, which is led by its dysfunction on high-contrast

TABLE I
PERFORMANCE COMPARISON OF TWO-STEP FRAMEWORK AND OTHER METHODS

Reconstruction	MNIST	Letter	Fresnel
DCAEDS	0.01033s	0.01015s	0.01022s
DRCNN	0.01237s	0.01276s	0.01235s
SDM	0.1325s	0.1307s	0.1342s
MCNN	0.005823s	0.005744s	0.005936s
Gauss-Newton (10 times iteration)	2.2739s	2.3058s	2.2982s

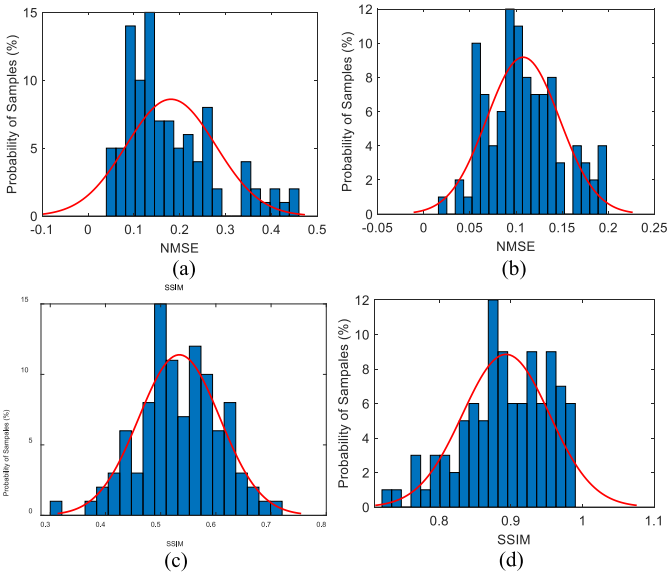

Fig. 5. NMSE and SSIM statistical histograms of the reconstruction quality. (a) NMSE obtained from DCAEDS. (b) NMSE obtained from DRCNN. (c) SSIM obtained from DCAEDS. (d) SSIM obtained from DRCNN.

targets and by the limited measured data (i.e., one-time-measured single-frequency scattered field). Unlike the results from DCAEDS, these wrong predictions from Gauss–Newton cannot provide any meaningful information for reconstruction. Furthermore, other popular DL-based inversion methods (i.e., the multiple-layer convolutional neural network (MCNN) [16] and the SDM [27], [28]) have also been adopted as a comparison for demonstrating the capability of our two-step DL framework. As presented in Fig. 4, MCNN can provide bad reconstruction results with little "meaningful" information, while SDM can generally provide some "meaningful" prediction and roughly describe the shape of the scatterers. Compared with the excellent reconstruction from DRCNN, they suffer from heavy distortion and blurring outline of reconstruction. Moreover, our developed DL approach, described as this two-step solver, significantly accelerates the reconstruction process compared to the traditional Gauss–Newton, as evidenced in Table I, which typically requires approximately 2.3 s to perform ten iterative optimization calculations for a single scatterer's reconstruction, while its final reconstruction nearly cannot provide any meaningful information, as illustrated in Fig. 4(b). Besides, MCNN and SDM utilize about 0.006 and 0.13 s, respectively, to complete reconstruction for one sample. On the contrary, the proposed two-step solver utilizes less than 0.022 s (0.01 s for DCAEDS and 0.012 s for DRCNN) to successfully complete the reconstruction computation of one scatterer, which makes it much more suitable for the real-time application.

For our two-step DL solver, Fig. 5(a)–(d) illustrates the statistical analyses for the testing results: 1) NMSE: the average of NMSE is about 0.2 for DCAEDS, while the average improved to 0.1 by DRCNN, and 2) SSIM: the SSIM average is around 0.55 for the reconstruction of DCAEDS, while this average can be larger than

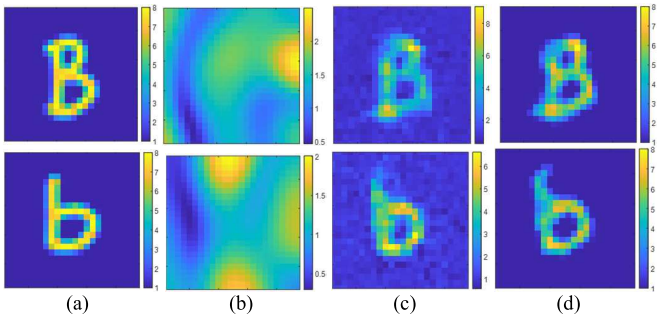

Fig. 6. Comparison of reconstruction of letter-shaped targets. (a) Ground truths. (b) Gauss–Newton method. (c) DCAEDS in the first step. (d) DRCNN in the second step.

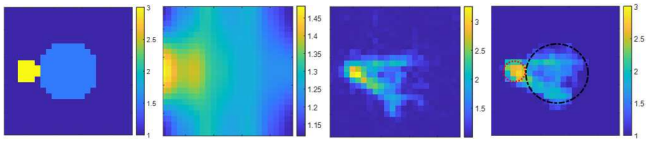

Fig. 7. Comparison of reconstruction for FoamDielExt. (a) Ground truths. (b) Gauss–Newton method. (c) DCAEDS in the first step. (d) DRCNN in the second step.

0.9 by DRCNN. Therefore, the proposed approach can realize EM inversion with excellent performances.

C. Study on Performance of Two-Step DL Framework

In this section, various numerical examples are utilized to challenge the boundary and limitation of the proposed two-step DL framework.

1) Performance on Letter-Shaped Scatterers: A new synthetic dataset is first employed, called "Letter," where the heterogeneous letter-shaped scatterers are assumed to be randomly located in D_{obj}. The size of each sample in "Letter" is set to $\lambda \times \lambda$, while their relative permittivities ε_r are randomly set into the range (1, 8]. Thousand new scatterers from "Letter" are randomly selected to test the trained DL solver for EM inversion, while 5% white noise is added to the received scattered field data for input. Fig. 6 illustrates the comparison between the ground truth, the reconstruction from the Gauss–Newton method, and the reconstruction from DCAEDS and DRCNN. Obviously, while the reconstruction from DCAEDS has demonstrated much meaningful information, the final reconstruction from DRCNN in the second step agrees well with the ground truth. The traditional approach (Gauss–Newton method) falls short in yielding acceptable outcomes. In contrast, our designed DL solver demonstrates a markedly more efficient reconstruction process, completing the task in approximately 0.022 s (0.01 s for DCAEDS and 0.012 s for DRCNN) for a single sample. This is a significant reduction in time compared to the Gauss–Newton method, which takes about 2.3 s, as shown in Table I. Taking into account that our DL model is exclusively trained on MNIST dataset, the resulting reconstructions yield gratifying outcomes to solve the EM inversion problem.

2) Performance on Experimental Data: To further illustrate the capability of our two-step DL method, the widely used experimental data measured at Institute Fresnel are employed [23], [24], [27], [28]. In this experiment (i.e., "FoamDielExt"), the scatterers consist of two cylinders: while the bigger one has a diameter of 0.08 m with the relative permittivity 1.45($\pm$0.15), the smaller one has a diameter of 0.031 m with the relative permittivity 3 ($\pm$0.3). To adapt the experimental data from Fresnel into the proposed DL method, we select received scattered field only from 24 receivers (only about 1.2% experimental measurement data) resulting from one transmitter. In response to the constraints of the experiment,

we retrain DL models. We emphasize that only the MNIST dataset is still used to recreate simulation-based training data for our DL method. The performance of two retrained models has been presented in Fig. 7. The Gauss–Newton method cannot provide satisfactory reconstruction again. However, our DL method provides an acceptable reconstruction on the experimental data.

IV. Conclusion

This communication presents a novel two-step DL solver for realizing EM inversion, which offers significant advantages over existing methods. Only one-time measurement is needed to provide the input for the DCAEDS in the first step. While the first step realizes the coarse reconstruction on the target, the second step utilizes DRCNN to further refine the reconstruction to the final prediction by adopting the mixed input scheme. As a result, EM inversion can be successfully realized with much higher accuracy and efficiency even for high-contrast targets. Numerical examples demonstrate the capability and feasibility of the proposed two-step DL solver with the clear accuracy and efficiency improvement.

References

[1] W. C. Chew et al., "Nonlinear diffraction tomography: The use of inverse scattering for imaging," *Int. J. Imag. Syst. Technol.*, vol. 7, no. 1, pp. 16–24, 1996.

[2] A. Abubakar, T. M. Habashy, G. Pan, and M.-K. Li, "Application of the multiplicative regularized Gauss–Newton algorithm for three-dimensional microwave imaging," *IEEE Trans. Antennas Propag.*, vol. 60, no. 5, pp. 2431–2441, May 2012.

[3] M. Pastorino, *Microwave Imaging*. Hoboken, NJ, USA: Wiley, 2010.

[4] P. M. Meaney, M. W. Fanning, D. Li, S. P. Poplack, and K. D. Paulsen, "A clinical prototype for active microwave imaging of the breast," *IEEE Trans. Microw. Theory Techn.*, vol. 48, no. 11, pp. 1841–1853, Nov. 2000.

[5] M. Li and A. Abubakar, "Electromagnetic inverse problems [guest editorial]," *IEEE Antennas Propag. Mag.*, vol. 59, no. 5, pp. 9–115, Oct. 2017.

[6] J. Behura, K. Wapenaar, and R. Snieder, "Autofocus imaging: Image reconstruction based on inverse scattering theory," *Geophysics*, vol. 79, no. 3, pp. A19–A26, May 2014.

[7] A. Abubakar et al., "The contrast source inversion method for location and shape reconstructions," *Inverse Problems*, vol. 18, no. 2, pp. 495–510, 2002.

[8] S. Sun, B. J. Kooij, T. Jin, and A. G. Yarovoy, "Cross-correlated contrast source inversion," *IEEE Trans. Antennas Propag.*, vol. 65, no. 5, pp. 2592–2603, May 2017.

[9] M. Moghaddam and W. C. Chew, "Study of some practical issues in inversion with the born iterative method using time-domain data," *IEEE Trans. Antennas Propag.*, vol. 41, no. 2, pp. 177–184, Feb. 1993.

[10] L. Crocco et al., "Testing the contrast source extended Born inversion method against real data: The TM case," *Inverse Problems*, vol. 21, no. 6, p. S33, 2005.

[11] Y. Liu, Z. Zhao, X. Zhu, Z. Nie, and Q. H. Liu, "A CSEB subspace-based optimization method for reconstruction of uniaxial anisotropic objects," in *Proc. Prog. Electromagn. Res. Symp.-Fall (PIERS-FALL)*, 2017, pp. 93–98.

[12] K. I. Hopcraft and P. R. Smith, *An Introduction to Electromagnetic Inverse Scattering*. Boston, MA, USA: Kluwer Academic, 1992.

[13] M. Usama et al., "Unsupervised machine learning for networking: Techniques, applications and research challenges," *IEEE Access*, vol. 7, pp. 65579–65615, 2019.

[14] S. Dargan, M. Kumar, M. R. Ayyagari, and G. Kumar, "A survey of deep learning and its applications: A new paradigm to machine learning," *Arch. Comput. Methods Eng.*, vol. 27, pp. 1071–1092, Jun. 2019.

[15] H. M. Yao and L. Jiang, "Machine-learning-based PML for the FDTD method," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, pp. 192–196, 2019.

[16] H. H. Zhang, H. M. Yao, L. Jiang, and M. Ng, "Deep long short-term memory networks-based solving method for the FDTD method: 2-D case," *IEEE Microw. Wireless Technol. Lett.*, vol. 33, no. 5, pp. 499–502, May 2023.

[17] H. M. Yao, L. Jiang, and M. Ng, "Implementing the fast full-wave electromagnetic forward solver using the deep convolutional encoder–decoder architecture," *IEEE Trans. Antennas Propag.*, vol. 71, no. 1, pp. 1152–1157, Jan. 2023.

[18] H. H. Zhang, L. J. Jiang, and H. M. Yao, "Embedding the behavior macromodel into TDIE for transient field-circuit simulations," *IEEE Trans. Antennas Propag.*, vol. 64, no. 7, pp. 3233–3238, Jul. 2016.

[19] H. H. Zhang, L. J. Jiang, H. M. Yao, and Y. Zhang, "Transient heterogeneous electromagnetic simulation with DGTD and behavioral macromodel," *IEEE Trans. Electromagn. Compat.*, vol. 59, no. 4, pp. 1152–1160, Aug. 2017.

[20] R. Guo et al., "Joint inversion of audio-magnetotelluric and seismic travel time data with deep learning constraint," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 9, pp. 7982–7995, Sep. 2021.

[21] S. Chen, H. Wang, F. Xu, and Y.-Q. Jin, "Target classification using the deep convolutional networks for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4806–4817, Aug. 2016.

[22] H. M. Yao, L. Jiang, and W. E. I. Sha, "Enhanced deep learning approach based on the deep convolutional encoder–decoder architecture for electromagnetic inverse scattering problems," *IEEE Antennas Wireless Propag. Lett.*, vol. 19, pp. 1211–1215, 2020.

[23] H. H. Zhang, H. M. Yao, L. Jiang, and M. Ng, "Solving electromagnetic inverse scattering problems in inhomogeneous media by deep convolutional encoder–decoder structure," *IEEE Trans. Antennas Propag.*, vol. 71, no. 3, pp. 2867–2872, Mar. 2023.

[24] L. Li, L. G. Wang, F. L. Teixeira, C. Liu, A. Nehorai, and T. J. Cui, "DeepNIS: Deep neural network for nonlinear electromagnetic inverse scattering," *IEEE Trans. Antennas Propag.*, vol. 67, no. 3, pp. 1819–1825, Mar. 2019.

[25] M. Salucci, M. Arrebola, T. Shan, and M. Li, "Artificial intelligence: New frontiers in real-time inverse scattering and electromagnetic imaging," *IEEE Trans. Antennas Propag.*, vol. 70, no. 8, pp. 6349–6364, Aug. 2022.

[26] L. Li, L. G. Wang, and F. L. Teixeira, "Performance analysis and dynamic evolution of deep convolutional neural network for electromagnetic inverse scattering," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, pp. 2259–2263, 2019.

[27] H. M. Yao, R. Guo, M. Li, L. Jiang, and M. K. P. Ng, "Enhanced supervised descent learning technique for electromagnetic inverse scattering problems by the deep convolutional neural networks," *IEEE Trans. Antennas Propag.*, vol. 70, no. 8, pp. 6195–6206, Aug. 2022.

[28] R. Guo et al., "Pixel- and model-based microwave inversion with supervised descent method for dielectric targets," *IEEE Trans. Antennas Propag.*, vol. 68, no. 12, pp. 8114–8126, Dec. 2020.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[30] M. A. Fiddy and R. S. Ritter, *Introduction to Imaging From Scattered Fields*. Boca Raton, FL, USA: CRC Press, 2014.

[31] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[32] F. Wuttke, H. Lyu, A. S. Sattari, and Z. H. Rizvi, "Wave based damage detection in solid structures using spatially asymmetric encoder–decoder network," *Sci. Rep.*, vol. 11, no. 1, p. 20968, Oct. 2021.

[33] I. A. Kazerouni, G. Dooly, and D. Toal, "Ghost-UNet: An asymmetric encoder–decoder architecture for semantic segmentation from scratch," *IEEE Access*, vol. 9, pp. 97457–97465, 2021.

[34] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2015, pp. 234–241.

[35] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.

[36] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA, USA: Academic, 1999.

[37] J. Cong and B. Xiao, "Minimizing computation in convolutional neural networks," in *Proc. Int. Conf. Artif. Neural Netw.*, 2014, pp. 281–290.

[38] M. F. Catedra et al. *The CG-FFT Method: Application of Signal Processing Techniques to Electromagnetics*. Norwood, MA, USA: Artech House, 1995.

[39] P. Kim, *MATLAB Deep Learning*, New York, NY, USA: Apress, 2017.

[40] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2015, pp. 1–41.