Towards Intelligent Electromagnetic Inverse Scattering Using Deep Learning Techniques and Information Metasurfaces

CHE LIU [1,2] (Member, IEEE), **HONGRUI ZHANG**[3], **LIANLIN LI** [3] (Senior Member, IEEE),
AND TIE JUN CUI [1,2] (Fellow, IEEE)

(Invited Paper)

[1]Institute of Electromagnetic Space, Southeast University, Nanjing 211189, China
[2]State Key Laboratory of Millimeter Wave, Southeast University, Nanjing 211189, China
[3]School of Electronic, Peking University, Beijing 100871, China

CORRESPONDING AUTHORS: Tie Jun Cui; Lianlin Li (e-mail: tjcui@seu.edu.cn; lianlin.li@pku.edu.cn).

ABSTRACT Electromagnetic inverse scattering (EMIS) is uniquely positioned among many inversion methods because it enables to image the scene in a contactless, quantitative and super-resolution way. Although many EMIS approaches have been proposed to date, they usually suffer from two important challenges, i.e., time-consuming data acquisition and computationally -prohibitive data post processing, especially for large-scale objects with high and even moderate contrasts. To tackle the challenges, we here propose a framework of intelligent EMIS with the aid of deep learning techniques and information metasurfaces, enabling to the efficient data acquisition and instant data processing in a smart way. Towards this goal, as illustrative examples, we considerably extend the canonical contrast source inversion (CSI) algorithm, a canonical EMIS method by updating the contrast via the generative adversarial network (GAN), an unsupervised deep learning approach, leading to a novel physics-informed unsupervised deep learning method for EMIS, referred to as CSI-GAN in short. Compared with existing deep learning solutions for EMIS, our method relies on the supervision of physical law instead of the labeled training dataset, beating the bottleneck arising from the collection of labeled training datasets. Furthermore, we propose a scheme of adaptive data acquisition with the use of information metasurface in a cost-efficiency way, remarkably reducing the number of measurements and thus speeding up the data acquisition but maintaining the reconstruction's quality. Illustrative examples are provided to demonstrate the performance gain in terms of reconstruction quality, showing the promising potentials for providing the intelligent scheme for the EMIS problems.

INDEX TERMS MTT 70th Anniversary Special Issue, electromagnetic inverse scattering (EMIS), contrast source inversion, physics-informed neural network (PINN), unsupervised deep learning method, information metasurface.

I. INTRODUCTION

Electromagnetic inverse scattering (EMIS) has been shown a powerful contactless super-resolution examination technique, which enables us to "see" clearly the internal structure of scene [1], and has found its valuable applications in nondestructive evaluation [2], [3], [4], geological exploration [5], [6], [7], cancer detection [8], [9], [10], and security check [11], [12], to name a few. By now, various inverse scattering methods have been developed, such as contrast source inversion (CSI) [13], [14], [15], [16], distorted Born/Rytov iterative methods [17], [18], and stochastic methods [19], [20], [21], etc. However, they usually suffer from two formidable challenges, i.e., time-consuming data acquisition and

computationally prohibitive data post processing, especially for large-scale objects with high and even moderate contrasts. To deal with the challenges, here we propose a concept of intelligent electromagnetic inverse scattering by exploring the deep learning techniques for the adaptive data postprocessing on digital level and the information metaurface for the adaptive data acquisition on physical level.

Technically, the aforementioned EMIS methods are developed in the context of data postprocessing. Parallelly, the deep learning techniques are gained ever-increasingly attentions mainly in the digital world since the huge success of artificial intelligence (AI) board-game-Go program (AlphaGo) [22], such as in the areas of speech recognition [23], [24], [25], image recognition [26], [27], [28], automatic translation [29], [30], [31], image editing [32], [33], [34], [35], robot control [36], [37], [38], etc. Naturally, the deep learning techniques can be explored to solve the difficulties arising in the current EMIS algorithms. By now, various deep-learning-based EMIS methods have been proposed for pixel-based inverse scattering imaging [39], [40]. In [41], a U-Net structure [42] was used to learn the mapping from the electromagnetic input data to the pixelated contrast images, and three schemes to preprocess the input data were discussed: raw measured data, back propagation (BP), and dominant current schemes, where the latter two schemes were proved able to reconstruct satisfactory images. The effectiveness of the U-Net structure was also verified in the image reconstruction from multiple scattered light measurements [43], where it is proved that modified contrast is more suitable to be the input of the U-Net for full-wave EMIS problems [44]. In [45], three different input schemes to U-Net for solving EMIS with phaseless data (PD) were proposed and discussed, i.e., the raw phaseless measured field data, the dominant induced currents by the Levenberg-Marquardt method, and the contrast source inversion method, where except inputting raw PD, the other two input schemes both performed robust and accurate imaging results. In order to improve the imaging quality in EMIS when directly inputting the raw field data to the deep learning networks, data compression and preprocessing tricks were introduced by a pretrained autoencoder [46] or semiphysics-driven subnet [47] to reduce mapping workloads of the inverse deep learning networks.

Besides the U-Net, some generative adversarial networks (GANs), which are instinctively suitable for the imaging generation tasks, also prove to work well for electromagnetic pixelated imaging. In [48], GAN was trained to generate virtual dielectric anatomical breast phantoms that were similar to real human breasts, which realized data expansion for further researches. In the forward electromagnetic scattering problem, a framework of GAN, named pix2pix [49], was used to immediately predict the graphical induced currents when given permittivity contrast and incident field [50]. In [51], a complex-valued pix2pix was proposed for EMIS, in which a generator composed of multi-layer complex-valued convolutional neural network was trained to learn the mapping from

the BP result to the contrast images. Although the generator was complex-valued, the researchers in [51] admitted that their proposed method still lack interpretability, which is a common problem for the deep learning schemes. Further, in [52], an attention-assisted pix2pix with spatial attention mechanism was used to reconstruct scatterers including perfectly electric conductor (PEC) and dielectric objects, which is the first time that the learning-based approach was introduced to solve the EMIS with mixed boundary conditions.

The cost for lack of interpretability is that the training of deep learning networks is a black-box process and a great deal of training data are needed for a satisfactory result. In order to reduce the dependence on training data and increase the generalization ability, physical priori knowledge should be manually added into the design of deep learning schemes. Recently, some physics-informed neural networks driven by partial differential equations have been rapidly developed, which have shown great effectiveness in solving the classic physical problems including fluid mechanics and quantum mechanics [53]. For the inverse scattering, inspired by the similarities between the iterative optimization method of nonlinear inverse scattering and the ANN architecture, a cascade of end-to-end complex-valued residual convolutional neural network (CNN) modules, termed DeepNIS, was proposed in our previous work [54], in which the generalization ability was verified by testing with the experimental data. Another physics-informed work for EMIS was demonstrated in [55], where a cascaded end-to-end CNN was designed to hierarchically fit the induced current from scattered fields by setting multiple training labels in the intermediate layers. As a result, the nonlinearity of the mapping relation that the proposed CNN needed to fit was greatly reduced, which was verified by numerous numerical and experimental tests [55].

Besides these physics-informed end-to-end deep learning networks, the deep learning frameworks can also be iterative by integrated with iterative physical optimization operators, which may not be as fast as the end-to-end frameworks, but the inversion process can be relatively more controllable and interpretable. Such iterative deep learning frameworks mainly work by two sequential stages: offline training and online optimization. As a classical case presented in [56], in the offline training stage, Guo et al. trained two kinds of neural networks named "FWD SOLVER" and "subnetwork" in ISP SOLVER to surrogate the numerical forward solver [57] and cascaded inversion solver respectively. In the online optimization stage, the residual between the observed data and simulated data was input to the subnetwork to guide the iterative update of model reconstruction. Another iterative deep learning framework was introduced in [58], where a pretrained supervised deep learning network was embedded in a CSI method to iteratively map the signal subspace of the contrast source to an estimate of the true contrast source.

Although the prior physical knowledge was introduced to increase the generalization by these abovementioned attempts, it was not enough to get rid of paired training data. In other

words, these deep learning schemes are still supervised methods, which need to collect the scattered fields for thousands of training samples. Both simulation and experiment methods for scattered field collection are time consuming and expensive, especially for experiments.

In our daily research in iterative optimization methods [1] for EMIS of intricate scatterers, we found that if the contrast image was initialized as a shape topologically similar to the ground truth, the iterative methods could eventually converge to a satisfactory result. On the contrary, if the contrast image was not well initialized, such as initialized by a fuzzy BP result, the iterative methods were easy to generate blurry results. The main reason for this phenomenon is that the strong nonlinearity and ill-posedness of EMIS make the iterative methods trap in local optimum, which is hard to escape without outside assistances, such as revision with topological features. For solving the EMIS problems, we realized that respecting physical formulas is the inherent characteristic of the iterative optimization methods while adding the topological features is the advantages of deep learning methods especially GANs. Inspired by this idea, in this article, we propose a physics-informed unsupervised deep learning framework for solving EMIS, where a whole iterative optimization method is directly integrated in the framework to provide the physical priori knowledge, while an unsupervised GAN with physical loss function is trained to add the topological features into the optimization process of EMIS. In order to demonstrate the procedure and effectiveness of the proposed framework, we choose CSI as the example of iterative optimization backbones integrated in our deep learning framework, which is termed CSI-GAN in this article. Benefitting from the abundant physical mechanism brought by CSI, the measured scattered fields in training dataset are no longer needed and the contrast images alone are enough to train our CSI-GAN. Time-consuming and laborious full-wave simulations for the training-data preparation could be skipped, which lowers the threshold for the usage of the deep learning methods in the EMIS problems. Several simulation tests are conducted to fully vindicate the effectiveness of our proposed framework. Nevertheless, it must be noted that the unsupervised framework of CSI-GAN is accompanied by higher costs of computational resources and imaging time compared with the supervised deep learning frameworks. From the user's perspective, choosing the supervised or unsupervised framework depends on whether the paired training samples are easy to be collected or not, or the user prefers to save time on the data preparation process or the imaging process.

For the data acquisition of EMIS, there are three schemes of measurement configurations: the real aperture (RA), synthetic aperture (SA), and coding aperture (CA). The RA method uses a great number of elements in a large aperture, which is flexible in controlling the measurement modes, but suffers from its big size, heavy weight, high power, and expensive price of hardware. On the contrary, the SA method works with the mechanical movement of a single sensor to form virtually a large-size aperture in digital world. Apparently, the SA system

has relatively low hardware cost compared to the RA system; however, it is inefficient in data acquisition besides the relatively low signal-to-noise ratio (SNR) due to the mechanical movement of the single sensor. The CA method relies on the theory of compressive sensing [59], in which a sequence of random modulators is manipulated to achieve the *compressive* measurements. It is apparent that information metasurface can play the role of the dynamic modulators. The information metasurface was developed from programmable metasurface [60], [61], [62], [63], which consists of an array of controllable units (called as meta-atoms), where each meta-atom is connected to an external control unit so that its resonance can be damped by application of a bias voltage. Through applying different voltages to the control circuit, the selected subsets of the elements can be switched on to create the desirable radiation patterns. As such, the information metasurface is capable of producing arbitrary radiation patterns in a very flexible and dynamic manner. In this way, spatial information of an imaging domain can be encoded onto this set of radiation patterns, or the measurements, which can be processed to reconstruct the targets in the scene. The readers may refer to [63] for the comprehensive review on the information metasurfaces. Apparently, this kind of system has the advantages of simplified architecture and low hardware cost compared to the phased-array techniques. Here, we would like to point out that the alliance of the information metasurface with machine learning techniques opens a new exciting door for designing smarter systems with lower hardware costs.

We remark that a much shorter conference version of this paper will appear in [64]. Our initial conference paper did not address the detailed structure of the improved CSI-GAN nor the termination strategy of CSI-GAN iteration process. This paper will address these issues and add a comprehensive discussion of how information metasurface is used in solving EMIS.

II. FORMULATIONS OF THE EMIS PROBLEM

In this section, we detailly introduce the formulations describing the EMIS problem, while the formulations of the CSI algorithm that are needed to integrate with our deep learning framework are explained in APPENDIX in details. Some inverse imaging results of three CSI algorithms (CSI [13], ECSI [14] and MCSI [15]) are also illustrated.

A. FORWARD SCATTERING PROBLEM

For convenience, we consider a two-dimensional (2D) EMIS problem illuminated by plane waves with transverse-magnetic (TM) polarization, as illustrated in Fig. 1. A nonmagnetic scatterer with the relative permittivity $\varepsilon_r(\mathbf{r})$ is located in a domain of interest (DOI). Outside the scatterer is free space with permittivity ε_0 and permeability μ_0. For measurements, the DOI is illuminated successively by a total number of K_{in} transmitting antennas. For each illumination, the scattered electric fields are simultaneously recorded by a total number of K_s receiving antennas. The transmitters and receivers are alternately and uniformly arranged on a circle with radius
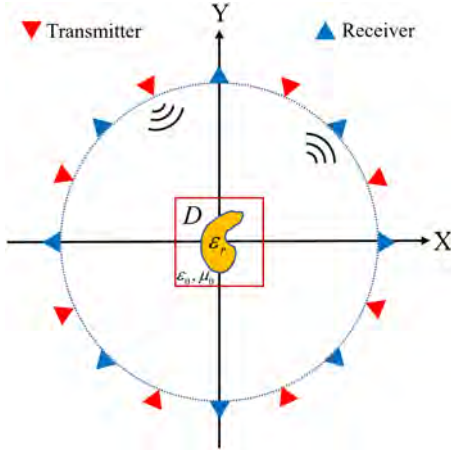
FIGURE 1. Measurement setup for 2D EMIS, where the unknown scatterers are located in a DOI (*D*). The transmitters and the receivers are alternately arranged around the DOI.

R sharing the same center with the DOI. We remark that *R* should be large enough so that the incident fields on the scatterer by the transmitting antennas are approximately plane waves.

Referring to the symbolic representation in [65], we use $E^t(\mathbf{r})$ and $E^i(\mathbf{r})$ to represent the total and incident electric fields at the field point $\mathbf{r}$, respectively. In domain *D* (DOI), the total electric field is calculated by:

$$E^t(\mathbf{r}) = E^i(\mathbf{r}) + i\omega\mu_0 \int_D g(\mathbf{r}, \mathbf{r}')[-i\omega\varepsilon_0(\varepsilon_r(\mathbf{r}') - 1)E^t(\mathbf{r}')]d\mathbf{r}', \mathbf{r} \in D \tag{1}$$

In free space *S* outside the domain *D*, the scattered field $E^s(\mathbf{r})$ radiated by scatterer can be written as:

$$E^s(\mathbf{r}) = i\omega\mu_0 \int_D g(\mathbf{r}, \mathbf{r}')[-i\omega\varepsilon_0(\varepsilon_r(\mathbf{r}') - 1)E^t(\mathbf{r}')]d\mathbf{r}', \mathbf{r} \in S \tag{2}$$

where $g(\mathbf{r}, \mathbf{r}')$ denotes the 2D Green's function from source point $\mathbf{r}'$ to field point $\mathbf{r}$ in free space:

$$g(\mathbf{r}, \mathbf{r}') = \frac{i}{4}H_0^{(1)}(k_0 |\mathbf{r} - \mathbf{r}'|) \tag{3}$$

where $H_0^{(1)}$ is the first-kind zeroth-order Hankel function. In order to simplify (1) and (2), a normalized contrast current density $J(\mathbf{r}')$ can be defined:

$$J(\mathbf{r}') = \chi(\mathbf{r}')E^t(\mathbf{r}') \tag{4}$$

where $\chi(\mathbf{r}')$ is the contrast defined by: $\chi(\mathbf{r}') = \varepsilon_r(\mathbf{r}') - 1$. Then (1) and (2) can be simplified as:

$$E^t(\mathbf{r}) = E^i(\mathbf{r}) + k_0^2 \int_D g(\mathbf{r}, \mathbf{r}')J(\mathbf{r}')d\mathbf{r}', \mathbf{r} \in D \tag{5}$$

$$E^s(\mathbf{r}) = k_0^2 \int_D g(\mathbf{r}, \mathbf{r}')J(\mathbf{r}')d\mathbf{r}', \mathbf{r} \in S \tag{6}$$

It is difficult to directly calculate the integral equations in (5) and (6). We use the method of moment (MOM) to transform these equations into discrete versions and refer to the symbolic representation in [55] to rewrite the formulations in their discrete versions. We assume that the numbers of transmitters and receivers are both *K*. For the convenience of centralized data processing, for each physical quantity, we gather the data measured at the *k*-th ($k = 1,2, \ldots,K$) radiation into one complex matrix and provide their computational formulas. Firstly, we discretize the domain *D* into $N \times N$ subunits, and the contrast values at the centers of these $N \times N$ subunits are represented by an $N \times N$ matrix, which is then flattened into a column vector $\overline{\chi}$. Then, we denote the incident field vector radiated by the *k*-th transmitter on $\overline{\chi}$ as $\overline{E}_k^i$, whose dimension is the same as $\overline{\chi}$. All of $\overline{E}_k^i(k = 1 2, \ldots,K)$ can be gathered into an $N^2 \times K$ complex matrix $\overline{\overline{\mathbf{E}}}^i$:

$$\overline{\overline{\mathbf{E}}}^i = [\overline{E}_1^i, \overline{E}_2^i, \ldots, \overline{E}_K^i] \tag{7}$$

Similarly, we define an $N^2 \times K$ complex matrix $\overline{\overline{\mathbf{E}}}^t$ to represent the total electrical field in domain *D*, whose definition is the same as $\overline{\overline{\mathbf{E}}}^i$ but replacing the incident field in $\overline{\overline{\mathbf{E}}}^i$ with the total electrical field. The normalized contrast current density in domain *D* under all *k*-th ($k = 1 2, \ldots,K$) radiations can then be represented by an $N^2 \times K$ complex matrix $\overline{\overline{\mathbf{J}}}$, which can be calculated by

$$\overline{\overline{\mathbf{J}}} = \overline{\overline{\boldsymbol{\xi}}} \cdot \overline{\overline{\mathbf{E}}}^t \tag{8}$$

to give the discrete form of (4), where $\overline{\overline{\boldsymbol{\xi}}}$ means the diagonalizable matrix of $\overline{\chi}$. The total electric field $\overline{\overline{\mathbf{E}}}^t$ can be calculated by

$$\overline{\overline{\mathbf{E}}}^t = \overline{\overline{\mathbf{E}}}^i + \overline{\overline{\mathbf{G}}}_D \cdot \overline{\overline{\mathbf{J}}} \tag{9}$$

to give the discrete form of (5), where $\overline{\overline{\mathbf{G}}}_D$ is an $N^2 \times N^2$ complex matrix whose definition is the same to [55]:

$$\overline{\overline{\mathbf{G}}}_D(n, n') = \frac{ik_0\pi a}{2}J_1(k_0a)H_0^{(1)}(k_0 |\mathbf{r}_n - \mathbf{r}_{n'}|), n \neq n' \tag{10}$$

$$\overline{\overline{\mathbf{G}}}_D(n, n') = \frac{ik_0\pi a}{2}H_1^{(1)}(k_0a) - 1, n = n' \tag{11}$$

Here, $a = (s/\pi)^{1/2}$, and *s* is the area of one subunit, $J_1(*)$ donates the Bessel function of the first order, $H_0^{(1)}$ and $H_1^{(1)}$ donate the zeroth and the first orders of the Hankel function of the first kind, respectively. Physically, $\overline{\overline{\mathbf{G}}}_D(n, n')$ means the propagation coefficient from the *n-th* subunit to the *n'-th* subunit of domain *D*. Similarly, (6) can also be represented in a discrete version:

$$\overline{\overline{\mathbf{E}}}^s = \overline{\overline{\mathbf{G}}}_S \cdot \overline{\overline{\mathbf{J}}} \tag{12}$$

where $\overline{\overline{\mathbf{E}}}^s$ is a $K \times K$ complex matrix, and each column represents the measured scattered fields under one radiation. $\overline{\overline{\mathbf{G}}}_S$ is a $K \times N^2$ complex matrix containing the propagation coefficient from the *n-th* subunit of domain *D* to the *m-th* receiver:

$$\overline{\overline{\mathbf{G}}}_S(m, n) = \frac{ik_0\pi a}{2}J_1(k_0a)H_0^{(1)}(k_0 |\mathbf{r}_n - \mathbf{r}_m|) \tag{13}$$
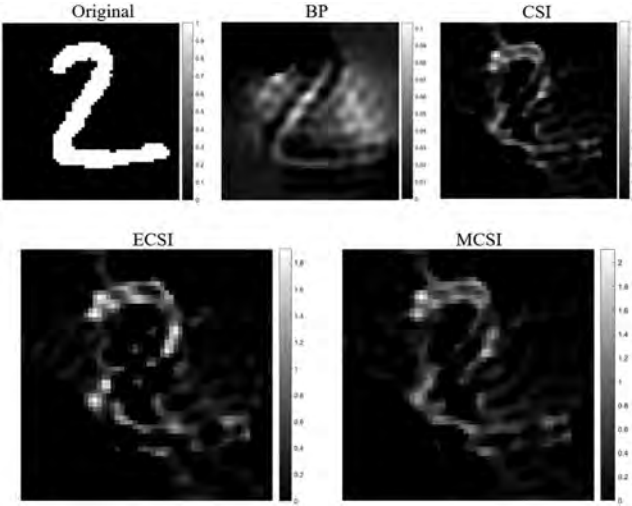
FIGURE 2. Testing results for four kinds of inverse scattering methods. The color bar indicates the value of the contrast $\chi(\mathbf{r}')$.

Lastly, combining (8) and (9), together with (12), we can get the two equations we need to define this EMIS, which are termed as data equation and state equation:

$$\text{Data equation} : \bar{\bar{\mathbf{J}}} = \bar{\bar{\boldsymbol{\xi}}} \cdot [\bar{\bar{\mathbf{E}}}^{\mathrm{i}} + \bar{\bar{\mathbf{G}}}_D \cdot \bar{\bar{\mathbf{J}}}]$$
$$\text{State equation} : \bar{\bar{\mathbf{E}}}^{\mathrm{s}} = \bar{\bar{\mathbf{G}}}_S \cdot \bar{\bar{\mathbf{J}}} \tag{14}$$

B. NUMERICAL SIMULATIONS OF CSI ALGORITHMS

In the numerical test, the domain of interest D with reference to Fig. 1 is set to be a $5.6\lambda_0 \times 5.6\lambda_0$ square ($\lambda_0 = 7.5$ cm is the working wavelength in vacuum), which is uniformly discretized into 64×64 sub-squares for the simulations. 36 linearly TM-polarized transmitters are located uniformly over the circle with a radius of $12\lambda_0$; and 36 co-polarized receivers alternately distributed with the transmitters are used to simultaneously measure the electrical field scattered from the probed scene. In the numerical forward problem calculations with MoM, the objects are set to be lossless dielectrics with a relative permittivity of $\varepsilon_r = 2$ and 2% noise has been added for all simulations throughout this article.

We randomly choose an image from the MNIST data set [66] and binarize it as the imaging object in the domain of interest D. Four kinds of inverse scattering methods (BP, CSI, ECSI and MCSI) are numerically tested to recover the permittivity-distribution of the imaging object. The testing results are demonstrated in Fig. 2. It can be found that all four inverse scattering methods fail to produce a legible image with sharp edge and complete topological structure.

III. THE CSI-GAN METHOD

In this section, we will introduce the methods of the proposed CSI-GAN in detail. First, the overall structure of CSI-GAN would be illustrated and detailly explained. Then, the design strategy for the loss function of GAN part would be interpreted. Last, the tricks of the training process of CSI-GAN

would be demonstrated for a stable convergence. The definition of all relevant symbols appeared in the CSI iterations, such as $\overline{\chi}_n$, $\bar{\bar{\boldsymbol{\xi}}}_n$, $\bar{\bar{\mathbf{J}}}_n$, $\bar{\bar{\mathbf{g}}}_n$ and $\bar{\bar{\mathbf{p}}}_n$, can be found in the Appendix section.

A. OVERALL STRUCTURE OF CSI-GAN

The schematic diagram of the proposed CSI-GAN in the n-th iteration is shown in Fig. 3. At the beginning ($n=1$), we execute the classic CSI method [13] for 100 iterations to obtain the initial values of the normalized contrast current density $\bar{\bar{\mathbf{J}}}_1$, the contrast $\bar{\bar{\boldsymbol{\xi}}}_1$, the gradient $\bar{\bar{\mathbf{g}}}_0$ and the Polak-Ribière conjugate gradient search direction $\bar{\bar{\mathbf{p}}}_0$. Because $\bar{\bar{\boldsymbol{\xi}}}_n$ is the diagonalizable matrix of the contrast vector $\overline{\chi}_n$, and the contrast image is the two-dimensional form of $\overline{\chi}_n$, in the following paragraphs we define a bold symbol " $\boldsymbol{\chi}_n$ " to represent the contrast image corresponding to $\overline{\chi}_n$ or $\bar{\bar{\boldsymbol{\xi}}}_n$.

The image of $\overline{\chi}_1$ (represented as $\boldsymbol{\chi}_1$) which is directedly initialized by CSI method is blurry in most cases because of the strong nonlinearity and ill-posedness of EMIS. In other words, the restriction imposed on $\boldsymbol{\chi}_1$ is not enough to let it converge to a high-resolution solution. Inspired by the fact that most of the images in real word own semantic information, in other words, they have regular shapes and could be recognized by people, thus we introduce a GAN to continually add topological constraint in the update process of the contrast image.

The schematic diagram of GAN is illustrated in Fig. 4. In the n-th iteration, the generator (Gen.) of GAN receives the image of $\overline{\chi}_{n+1}$ (represented as $\boldsymbol{\chi}_{n+1}$) from CSI, optimizes it towards the direction of adding more topological and semantic information, and outputs a modified image $\boldsymbol{\chi}_{n+1}^R$. Then $\boldsymbol{\chi}_{n+1}$ would be replaced by $\boldsymbol{\chi}_{n+1}^R$ in the next optimization iteration of CSI to help it jump out of the local optima and converge to the correct image.

The abstract structure diagram of CSI-GAN displayed in Fig. 3 is shown in Fig. 4. As we can see, after the contrast image is modified by GAN ($\boldsymbol{\chi}_n = \boldsymbol{\chi}_n^R$), the CSI iteration is only executed once to update the normalized contrast current density ($\bar{\bar{\mathbf{J}}}_n \rightarrow \bar{\bar{\mathbf{J}}}_{n+1}$), which is not enough to make $\bar{\bar{\mathbf{J}}}_{n+1}$ match with $\boldsymbol{\chi}_n$ from the angle of objective function $F(\bar{\bar{\mathbf{J}}}_{n+1}, \bar{\bar{\boldsymbol{\xi}}}_n)$ and will bring difficulty to the convergence of CSI-GAN . In order to make the convergence of CSI-GAN more stable, we improve the CSI-GAN workflow, as illustrated in Fig. 6. The improved workflow can be tackled in three steps. Firstly, the normalized contrast current density $\bar{\bar{\mathbf{J}}}_n$ is input into GAN, driving the training process of GAN before $\boldsymbol{\chi}_n$ is modified by GAN and replaced by $\boldsymbol{\chi}_n^R$.

Secondly, after $\boldsymbol{\chi}_n$ is replaced by $\boldsymbol{\chi}_n^R$, we keep its value unchanged while 50 iterations of the modified gradient method [14] are continuously executed to update $\bar{\bar{\mathbf{J}}}_n$ alone ($\bar{\bar{\mathbf{J}}}_n \rightarrow \bar{\bar{\mathbf{J}}}_n^R$) to make $\boldsymbol{\chi}_n$ and $\bar{\bar{\mathbf{J}}}_n^R$ match with each other:

$$\bar{\bar{\mathbf{J}}}_n^R = \underset{\bar{\bar{\mathbf{J}}}}{\arg\min} \, F(\bar{\bar{\mathbf{J}}}, \bar{\bar{\boldsymbol{\xi}}}_n) \tag{15}$$
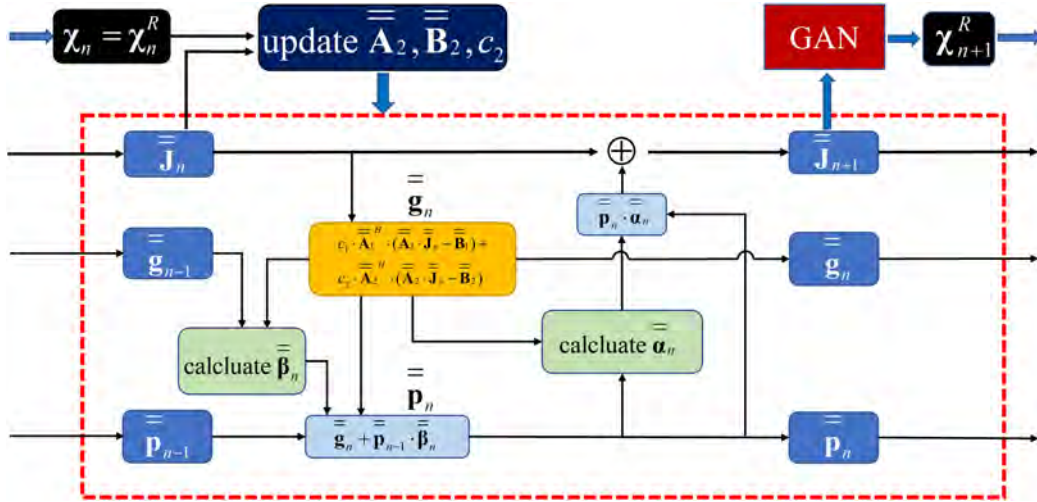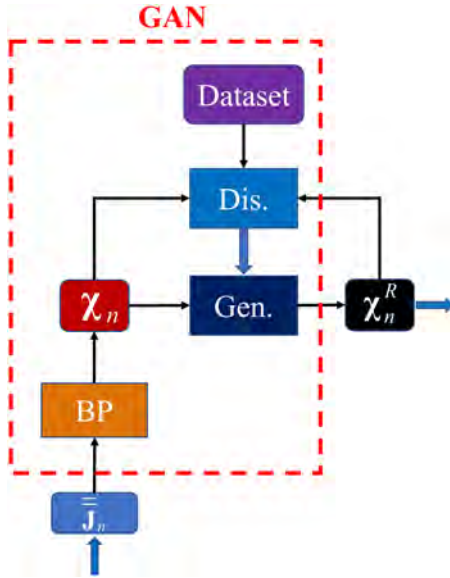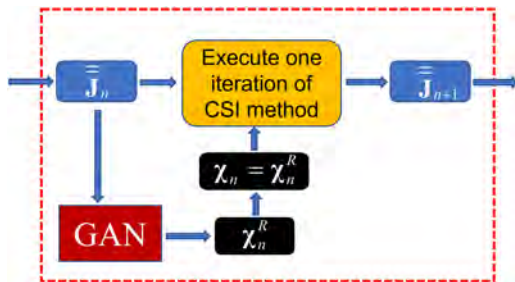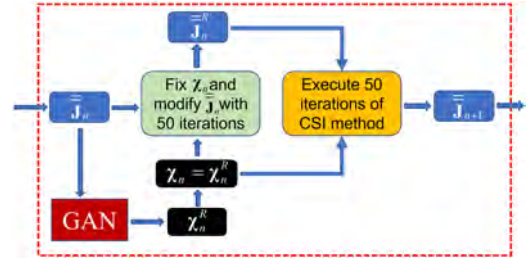
FIGURE 3. The structure diagram of CSI-GAN in the n-th iteration.

FIGURE 4. The abstract schematic diagram of GAN in CSI-GAN.

FIGURE 5. The abstract structure diagram of CSI-GAN in the n-th iteration.

Thirdly, a classic CSI procedure initialized by $\bar{\bar{\mathbf{J}}}_n^R$ and χ_n is executed by 50 iterations to update the normalized contrast current density ($\bar{\bar{\mathbf{J}}}_n \to \bar{\bar{\mathbf{J}}}_{n+1}$) before it is input to the GAN

FIGURE 6. The improved structure diagram of CSI-GAN in the n-th iteration.

module in the next iteration of CSI-GAN:

$$\bar{\bar{\mathbf{J}}}_{n+1} = \underset{\bar{\bar{\mathbf{J}}}, \bar{\bar{\xi}}}{\arg\min} F(\bar{\bar{\mathbf{J}}}, \bar{\bar{\xi}}) \tag{16}$$

Thus, the modified gradient iterations, CSI iterations and topological inpainting by GAN are alternately executed in the CSI-GAN pipeline. Meanwhile, all auxiliary variables can flow smoothly in the pipeline.

B. DERAILED TRUCTURE OF GAN

As we know, GANs [32] are able to transform Gaussian noise to target images, in other words, the output of GANs could be totally different from the input. As we just want to revise the contrast image (χ_n) using GAN, we wish that the output of GAN (χ_n^R) could share some common features with its input image, rather than outputting a brand-new image. So, the degree of freedom for GAN should be restricted to some extents. Inspired by this reason, we introduce a special GAN structure named as Cycle-GAN [67], which is widely used in the unpaired image-to-image translation to solve the mismatch problem between the input and output images, into the structure design of GAN in CSI-GAN.

The structure of the generator in the proposed CSI-GAN combined with Cycle-GAN is shown in Fig. 7, in which G_χ^R is the revised generator module responsible for the topological
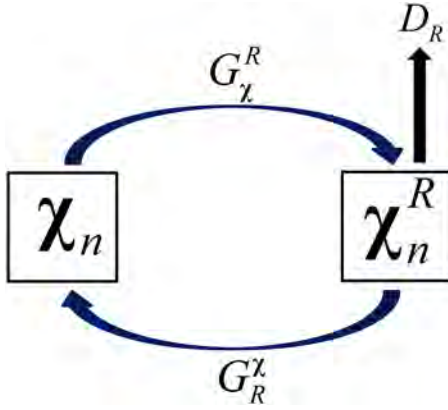
FIGURE 7. The structure of the generator in the proposed CSI-GAN combined Cycle-GAN.

inpainting, and G_R^χ is the inverse generator coupled with G_χ^R to restore χ_n^R to χ_n, insuring that χ_n^R and χ_n share some common topological features and restrict the freedom of G_χ^R. D_R is the discriminator to judge to which extent that χ_n^R shares the same topological features with the training dataset consisting of congeneric scatterers. The loss function of this Cycle-GAN structure will be provided in the next section.

C. LOSS FUNCTION OF GAN

For the design of CSI-GAN, we refer the error function of MSCI [15], which can be described as:

$$F_{MCSI}(\bar{\bar{\mathbf{J}}}, \bar{\bar{\boldsymbol{\xi}}}) = [F_S(\bar{\bar{\mathbf{J}}}) + F_D(\bar{\bar{\mathbf{J}}}, \bar{\bar{\boldsymbol{\xi}}})] \cdot F^R(\bar{\bar{\boldsymbol{\xi}}}) \quad (17)$$

where $F^R(\bar{\bar{\boldsymbol{\xi}}})$ is a topological regularization item indicating the flatness level of image edge. In the loss function of CSI-GAN, we replace $F^R(\bar{\bar{\boldsymbol{\xi}}})$ in (17) with a discriminator error and change the multiplicative regularization to the additive one. As a result, the loss functions of G_χ^R and G_R^χ in Fig. 7 can be written as:

$$Loss\_G_\chi^R = a_1 \cdot F_D\left(\bar{\bar{\mathbf{J}}}_n, G_\chi^R(\chi_n)\right)$$
$$+ a_2 \cdot \mathrm{CL}_1(\chi_n) + a_3 \cdot \mathrm{CL}_2(\chi_n^R)$$
$$- a_4 \cdot D_R\left(G_\chi^R(\chi_n)\right) \quad (18)$$

$$Loss\_G_R^\chi = a_2 \cdot \mathrm{CL}_1(\chi_n) + a_3 \cdot \mathrm{CL}_2(\chi_n^R) \quad (19)$$

where a_1 to a_4 are fixed parameters, while $\mathrm{CL}_1(\chi_n)$ and $\mathrm{CL}_2(\chi_n^R)$ are cycle consistency losses represented as:

$$\mathrm{CL}_1(\chi_n) = \left\| \chi_n - G_R^\chi\left(G_\chi^R(\chi_n)\right) \right\|_1 \quad (20)$$

$$\mathrm{CL}_2(\chi_n^R) = \left\| \chi_n^R - G_\chi^R\left(G_R^\chi(\chi_n^R)\right) \right\|_1 \quad (21)$$

in which $\|*\|_1$ represents the L1-norm.

For the loss function of discriminator D_R, we use the loss function of WGAN [68] to simulate the calculation of Wasserstein distance between the generated images and real images:

$$Loss\_D_R = \mathop{\mathbb{E}}_{\tilde{x} \sim P_g} [D(\tilde{x})] - \mathop{\mathbb{E}}_{x \sim P_t} [D(x)]$$
$$+ \lambda \mathop{\mathbb{E}}_{\hat{x} \sim P_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2] \quad (22)$$

where P_g and P_t represent the distributions of the generated images from G_χ^R (fake samples) and real scatterer images (real samples) in the training dataset, respectively.

D. TRAINING PROCESS OF GAN

Similar to [56], we train the GAN in the CSI-GAN by two sequential stages: offline training and online optimization. In the offline training process, we pre-train the generator alone using paired images with randomly smudged images by gaussian noise as input, and the corresponding raw images selected from the training dataset as the output to make it have the preliminary inpainting ability. In other words, the generator is pre-trained as a denoising network. However, a denoising mapping is now powerful enough to modify the contrast image in CSI iterations because it lacks the prior knowledge to guide the mapping from estimated contrast images to real ones in CSI iterations. Thus, in the online optimization process, we further train the revised generator G_χ^R with the back propagation gradients from the loss function (18) which has a physically-guided error term $F_D(*)$ to instruct G_χ^R in fitting the modified mapping we want. It must be noted that for different testing cases, the offline training process is executed ahead (no need to repeat) but the online optimization process needs to be restarted. Different testing cases do not share the network parameters of GAN in their online optimization processes. In other words, the GAN is a specialized image-corrector for each testing case and its network parameters are dynamically adapted in a testing case.

One of the most important things for GAN's training process in online optimization stage is to ensure that the generator is fully trained while preventing the discriminator from falling into overfitting. In the training process of discriminator, there may encounter a problem that fake samples generated from the revised generator G_χ^R are much less than the real samples from the training dataset, making the discriminator easily trap into overfitting and cause mode collapse. One way to relieve this problem is increasing the diversity of the fake samples. We use $G_\chi^R(\chi_n)$, $G_\chi^R(\chi_n^R)$ and their weighted means with random weighting coefficients as the whole batch of fake samples. Meanwhile, after the CSI-GAN procedure begins, in the n-th iteration, the GAN will be trained with 10 iterations before it is used to modify the contrast image. Another way is restricting the learning progress of discriminator. In one training iteration of GAN, the generator will be trained 3 times after the discriminator is trained once.

Some hyper-parameters need to be carefully designed for stable convergence. We choose the Resnet structure [69] with 34 layers and 4 layers as the generator and discriminator networks, respectively. For the fixed parameters in (18) and
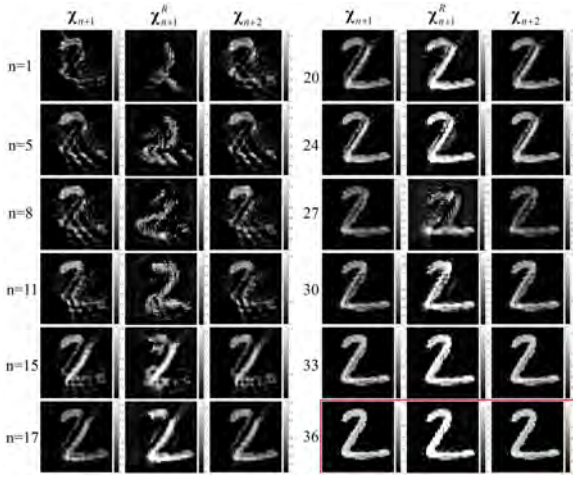
FIGURE 8. The imaging results of χ_{n+1}, χ_{n+1}^R and χ_{n+2} in the n-th optimization iteration of CSI-GAN for the target scatterer in Fig. 2.

(19), the values are taken as follows:

$$30 \leq a_1 \leq 100, a_2 = a_3 = 10, a_4 = 1 \tag{23}$$

The whole CSI-GAN pipeline is run on the Pytorch deep-learning platform including the modified gradient method iterations and CSI iterations to make full use of the parallel computing capability of the graphics processing unit (GPU). In other words, the modified gradient and CSI methods are integrated in the whole deep learning pipeline as a part of the deep neural network. The GPU device we use is NVIDIA RXT 3090. In our testing cases, the run time for one iteration of CSI-GAN including training the GAN is about 6s.

IV. TESTING RESULTS OF CSI-GAN

The settings of parameters for simulations are the same as those of Fig. 2. The image of the scatterer is chosen from the MNIST dataset of handwritten digits. We use 6000 images from MNIST as the training dataset for the GAN in CSI-GAN to learn the topological features of the potential target scatterer. It should be guaranteed that GAN will not see the exact image of the target scatterer. So, other 1000 images from MNIST are made as the testing dataset, from which a sample image is randomly chosen and then binarized to be used as the ground truth contrast image of the target scatterer. It must be noted that the contrast images in the training dataset need not to be numerically simulated to get their scattering fields.

When given a target scatterer, we get the corresponding scattered field $\bar{\bar{\mathbf{E}}}^s$ and put it into the CSI method to obtain an initial normalized contrast current density and other auxiliary variables to launch our CSI-GAN procedure. As the process goes on, the output of the generator would gradually converge to the correct image of the target scatterer. The imaging results of χ_{n+1}, χ_{n+1}^R and χ_{n+2} in the n-th optimization iterations of CSI-GAN for the target scatterer in Fig. 2 are shown in Fig. 8. As we can see, the GAN could help the integrated CSI method jump out of local optimum and repair the image with more topological features. The imaging result of the 36-th iteration
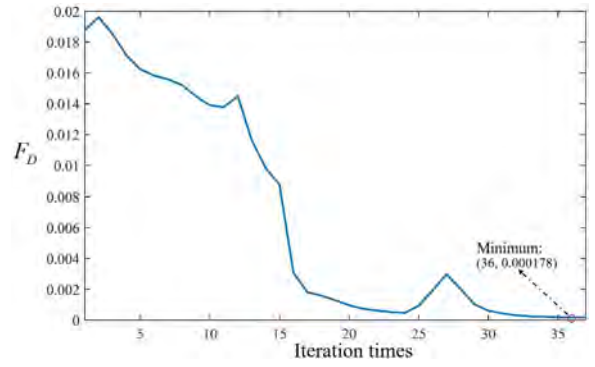

FIGURE 9. The graph of data error $F_D(\bar{\bar{\mathbf{J}}}_{n+1}, \bar{\bar{\xi}}_{n+2})$ in the n-th iteration.
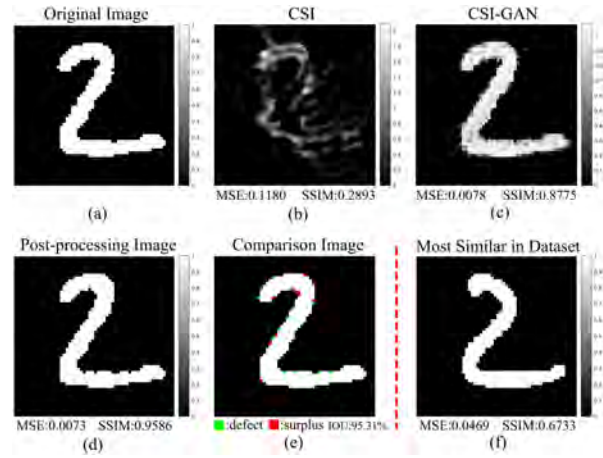

FIGURE 10. The quantitative analysis of the final imaging result for the target scatterer in Fig. 2. (a) The original image of the target scatterer. (b) The imaging result of the CSI method. (c) The imaging result of the proposed CSI-GAN. (d) The post-processing image by binarizing the imaging result of CSI-GAN. (e) The pixel-level comparison between (a) and (d). (f) The most similar image to (a) in the training dataset.

(bottom right corner in Fig. 8) is already very similar to the target image. The graph of data error function $F_D(\bar{\bar{\mathbf{J}}}_{n+1}, \bar{\bar{\xi}}_{n+2})$ in the n-th iteration is shown in Fig. 9. The initial data error at the 0-th iteration in Fig. 9 is the one for the imaging result of the classical CSI method and is proved by experiment that it could not be further improved by the CSI method itself or traps in a local optimum. However, the descending curve of data error in Fig. 9 indicates that the proposed CSI-GAN could haul it out of the quagmire and steadily converge to an ideal physical solution for this inverse scattering problem.

We further conduct a quantitative analysis of the final imaging result for the target scatterer (original image) in Fig. 2 and the results are presented in Fig. 10. The imaging result generated by CSI-GAN (Fig. 10(c)) is very close to the original image with a mean square error (MSE) of 0.0078 and the structural similarity index measure (SSIM) of 0.8775. In order to fully display the recovering capacity of topological details, we plot Fig. 10(d) by binarizing the image in Fig. 10(c) and further improve the SSIM to 0.9586. In Fig. 10(e), the defect and surplus pixels in Fig. 10(d)
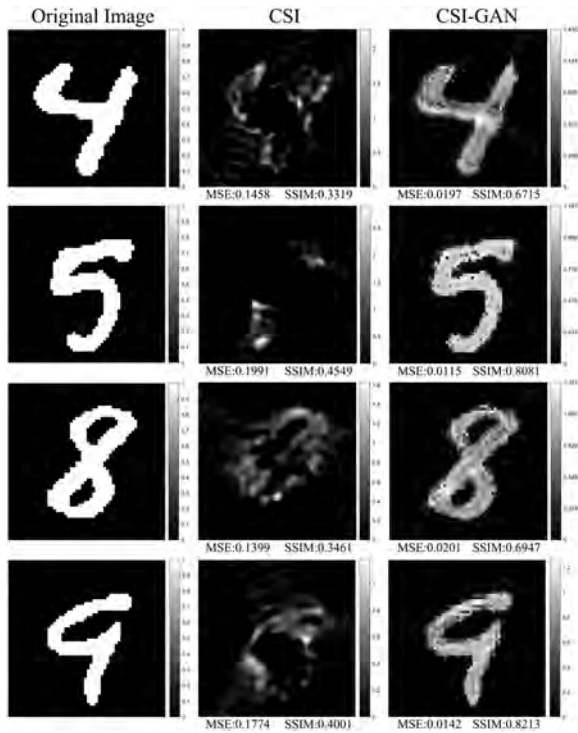
FIGURE 11. The comparison results of the inverse scattering imaging between CSI and the proposed CSI-GAN.

FIGURE 12. The graph of data error F_D and several local-optimum imaging results for the handwritten digit 4 in the iteration process of CSI-GAN.

FIGURE 13. The graph of data error F_D and several maximin imaging results for the handwritten digit 9 in the iteration process of CSI-GAN.

compared with the original image are marked in green and red respectively, showing that 95.31% pixels in Fig. 10(d) are identical to those of the original image, which confirms the powerful precision for the detail-recovery ability of our proposed CSI-GAN. Fig. 10(f) shows the most similar sample (under the criterion of SSIM) with the original image in the training dataset, whose topological posture is distinctly different from the original image. This fact indicates the effectiveness of the physical constraint put on our CSI-GAN, which can guide the output image to converge to the exact image of the target scatterer rather than a similar image existing in the MNIST dataset.

Other testing results are shown in Fig. 11. The imaging results of CSI methods are also given for comparison. As we can see in Fig. 11, the proposed CSI-GAN could generate high-resolution imaging results very close to the ground truth with plenty of semantic features, while the CSI method can only converge to blurry and illegible results.

Next, we will discuss the termination strategy of CSI-GAN iteration process. There are mainly two kinds of cases that may occur in the implementation of CSI-GAN. One is that CSI-GAN continually traps in and jumps out of the local optimum and finally converges to an optimal result, as Fig. 12 shows. Fig. 12 displays the graph of data error F_D and several local-optimum imaging results in the iteration process of CSI-GAN. In this case, the CSI-GAN could be ended when the data error F_D drops to a preset threshold value. Another case is show in Fig. 13 as an example. In this case, the CSI-GAN rapidly converges to an optimal result at the 11-th iteration and
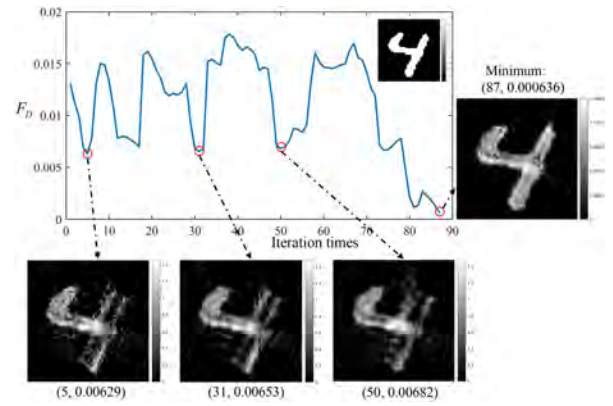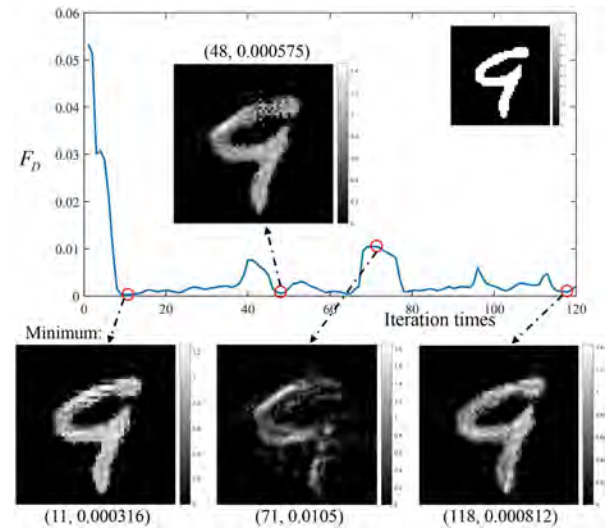
keeps the data error F_D stabilized around the optimal value. The difficulty for the judge of termination for this case is that we could not foresee whether a better result will come out in the subsequent iterations. In other words, we should know whether the CSI-GAN procedure deserves to continue to explore a potential better result when an ideal result is rapidly obtained. Weighing the efficiency and effectiveness together, our final termination strategy is firstly executing the CSI-GAN by 100 iterations. If the optimal data error F_D is lower than a preset threshold, the CSI-GAN is ended and the optimal result is outputted as the final result. Otherwise, if the optimal data error F_D does not meet the requirement, the CSI-GAN will be executed continuously until an acceptable result is obtained. The run time for one iteration of CSI-GAN is about 6s.

V. INFORMATION-METASURFACE-ENABLED INVERSE SCATTERING

It is noticed from the aforementioned discussions that the EMIS requires the time-consuming multi-static multi-view

measurements to achieve the acceptable results. However, for most of practical applications, it becomes a challenging issue for them due to the high-dimensional data or "data crisis". Fortunately, it is well known from the Johnson-Lindenstrauss lemma [70] that the high-dimensional data, if they have a structure representation, can be projected into a low-dimensional feature space with nearly neglectable information loss. That is to say, the essential information of the high-dimensional data can be efficiently retrieved from the remarkably reduced measurements. Given a sample $x \in \mathcal{X}$, where $\mathcal{X}$ is a cloud of Q points in an N-dimensional space, there exists, at least, a projection operator $\mathbf{H} : y = \mathbf{H}(x)$ such that the M-dimensional projection has as small loss of intrinsic information as possible compared to x, where $M<<N$. For illustration, we here restrict ourselves into the case of Born approximation [71]; however, the developed methodologies can be generalized for more general cases in a straightforward manner [63], [72], [73], [74]. Then, the operator $\mathbf{H}$ is linear, and this problem can be represented as $\bar{y} = \overline{\overline{\mathbf{H}}} \cdot \bar{x} + \bar{n}$, where $\bar{y} \in \mathbb{C}^M$ is the M-length measurements, $\bar{x} \in \mathbb{C}^N$ denotes the probed scene, and $\bar{n}$ is the measurement noise. Under the Born approximation, the entry of measurement matrix (H_{ij}) is simply proportional to the fields radiated by the transmitter and receiver antennas at a given point in the scene r_j: $H_{ij} \propto E_i^T(r_j) \cdot E_i^R(r_j)$, where $E_i^T(r_j)$ and $E_i^R(r_j)$ represent the radiation and receiving patterns. Note that a single sensor is assumed for data acquisition, like the setting of computational compressive imager [1]. Each row of the measurement matrix $\overline{\overline{\mathbf{H}}}$ corresponds to a measurement mode, and hence the number of rows equals the number of measurements. Apparently, the entries of $\overline{\overline{\mathbf{H}}}$ is determined by the illumination pattern of information metasurface $E_i^T(r_j)$.

In the terminology of machine learning, the matrix $\overline{\overline{\mathbf{H}}}$ corresponds to the so-called embedding projection operator. We here consider two popular projection operators widely used in the area of machine learning, in particular, the random projection and principal component analysis (PCA) projection. For the random projection, the entries of $\overline{\overline{\mathbf{H}}}$ are drawn from independent random numbers. Note that the random approach imposes no restrictions on the object to be reconstructed. The PCA is a prior task-aware embedding technique, where each row of $\overline{\overline{\mathbf{H}}}$ is trained over many training samples available. Thus, when a set of prior knowledge on the scene under investigation are available, PCA enables the design of efficient measurement modes, reducing remarkably the number of measurements.

Now, we provide a set of results to demonstrate the performance of the above method [71]. To that end, a two-bit information metasurface is considered here, which is composed of 48×48 controllable digital meta-atoms, as shown in Fig. 14(a). More detailed about the metasurface can be found in [71]. In addition, the two-bit information metasurface is trained to get the desired radiation patterns or measurement modes over the MNIST dataset, a widely utilized handwriting dataset in the area of machine learning. Full-wave
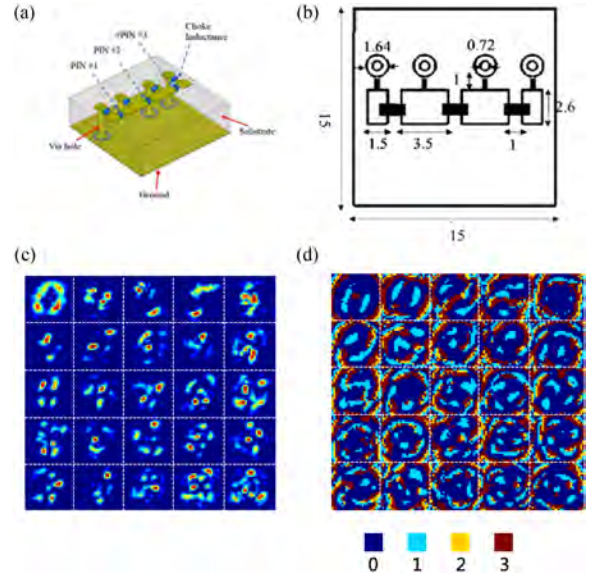
FIGURE 14. Schematic diagrams of information metasurfaces. (a)-(b) Schematic figure and geometrical parameters of the meta-atom for a two-bit information metasurface, respectively. (c)-(d) The first 16 leading PCA radiation patterns and the corresponding coding patterns of information metasurface. .
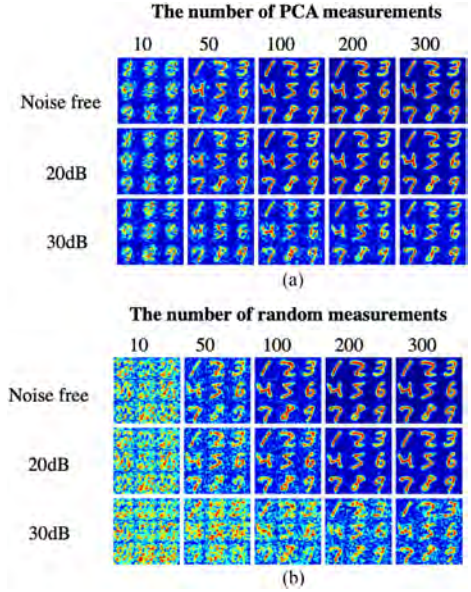
FIGURE 15. The recovered images of nine digit-like objects of the machine-learning imager trained with PCA (a) and random projection (b) with varying numbers of measurements, where different noise levels of 0dB, 20dB and 30dB are considered.

simulation results of the PCA-based radiation patterns are shown in Fig. 14(c), and the corresponding coding patterns of the information metasurface are given in Fig. 14(d). These figures illustrate that the information metasurface can be trained by PCA, implying that the machine-learning-driven information metasurface is able to generate the measurement modes needed by PCA. Fig. 15(a) and (b) illustrate the recovered images of nine digit-like objects of the machine-learning imager trained with the random projection and PCA

with varying numbers of measurements. Here, Fig. 15(a) and (b) present the retrieved images from the PCA and random projection respectively. In each method, the measurement numbers are 10, 50, 100, 200, and 300 from the left to the right. We also evaluate quantitatively the image qualities with respect to the signal-to-noise ratio (SNR). We clearly observe that the quality of image turns better with the increasing measurements, and PCA behave better than the random projection, especially in the low measurement cases.

To summarize, we can see from the above preliminary results that the information metasurface driven by machine learning techniques is able to produce the task-oriented measurement modes, which are very relevant to the probed object, in a very flexible and dynamic manner. In this way, the number of measurements can be remarkably reduced for the high-quality inversion, which is directly responsible for considerably speeding up the data acquisition, and meanwhile maintaining the very low complexity for the digital data processing.

VI. CONCLUSION

In this work, we propose the framework of intelligent EMIS by integrating deep learning techniques for efficient data processing and information metasurface for adaptive data acquisition into the entire EMIS's pipeline, in order to address the fundamental but important challenges arising from the conventional EMIS's strategies. Towards this goal, we firstly propose an unsupervised EMIS framework, i.e., CSI-GAN, which can be trained on the off-the-shelf dataset without paired training samples, leading to the remarkably reduced simulation time during the collection of training dataset. As opposed to the supervised EMIS's solutions relying on a large amount of labeled training data, our CSI-GAN takes the CSI with physical interpretations as the network's backbone, and explores the unsupervised deep learning technique for the update of the contrast images. Moreover, the CSI-GAN is not exclusionary with the supervised deep-learning-based inversion schemes, and the inversion results can be used as the initial solutions for the CSI-GAN to improve the convergence procedure. Of course, compared with those end-to-end inverse scattering deep learning networks trained by paired samples whose imaging results can be directly obtained when given the input electromagnetic data, our CSI-GAN needs several optimization iterations before obtaining the final imaging results, which is inferior to the former in computing overhead and imaging speed. However, our CSI-GAN can be used as a very ideal alternative method to achieve better inverse scattering imaging results when paired training samples are missing or difficult to obtain.

Furthermore, we propose the scheme of adaptive data acquisition with the information metasurface in a cost-efficiency way, remarkably reducing the measurement numbers and thus speeding up the data acquisition but maintaining the reconstruction quality. In the near future, the strategies to improve the performance of the intelligent EMIS could be further explored, for instance, to design more "smart" data acquisition schemes by combing the machine learning method with the information metasurface (reinforcement-learning-driven metasurface) to develop more specialized deep learning networks with more physically meaningful physical properties. We expect that the presented strategy can open a new avenue for the future intelligent EMIS by combining the information measurface and the machine learning techniques.

APPENDIX

A. BACK-PROPAGATION METHOD

Back-propagation (BP) is a simple non-iterative method to solve the EMIS problems and gets the estimate value of the contrast, which is widely used as the initial value for the iterative methods such as CSI. In this section, we provide the concise version of BP written by matrix operations.

The two main steps are as follows:

1) Estimate the value of the normalized contrast current density $\overline{\overline{\mathbf{J}}}$:

$$\overline{\overline{\mathbf{J}}} = \overline{\overline{\mathbf{G}}}_S^H \cdot \overline{\overline{\mathbf{E}}}^{\mathrm{s}} \cdot \overline{\overline{\gamma}} \tag{24}$$

where the superscript H means conjugate transpose, and $\overline{\overline{\gamma}}$ is a complex diagonal matrix with dimension $K \times K$. The value of $\overline{\overline{\gamma}}$ is to minimize the quadratic cost of state equation:

$$\underset{\overline{\overline{\gamma}}}{\operatorname{argmin}} \left\| \overline{\overline{\mathbf{E}}}^{\mathrm{s}} - \overline{\overline{\mathbf{G}}}_S \cdot \left[\overline{\overline{\mathbf{G}}}_S^H \cdot \overline{\overline{\mathbf{E}}}^{\mathrm{s}} \cdot \overline{\overline{\gamma}} \right] \right\|^2 \tag{25}$$

The minimization problem (25) has an analytical solution:

$$\overline{\overline{\gamma}} = \frac{\operatorname{sum}\left[\overline{\overline{\mathbf{E}}}^{\mathrm{s}} \odot \widetilde{\mathbf{E}}^{s*}, 1 \right]}{\operatorname{sum}\left[\widetilde{\mathbf{E}}^{s} \odot \widetilde{\mathbf{E}}^{s*}, 1 \right]} \tag{26}$$

where $\widetilde{\mathbf{E}}^s = \overline{\overline{\mathbf{G}}}_S \cdot \overline{\overline{\mathbf{G}}}_S^H \cdot \overline{\overline{\mathbf{E}}}^{\mathrm{s}}$. The function $\operatorname{sum}[\mathbf{A}, 1]$ means to calculate the summation of each column in matrix $\mathbf{A}$ and form a row vector. The symbol $\odot$ means element-wise multiplication, and superscript "$*$" means element-wise conjugate. The vector division in (26) is element-wise division.

2) Calculate $\overline{\overline{\mathbf{E}}}^{\mathrm{t}}$ using (9) and get the estimation of the contrast $\overline{\chi}$ by means of (8):

$$\overline{\chi} = \frac{\operatorname{sum}[\overline{\overline{\mathbf{J}}} \odot \overline{\overline{\mathbf{E}}}^{\mathrm{t}*}, 2]}{\operatorname{sum}[\overline{\overline{\mathbf{E}}}^{\mathrm{t}} \odot \overline{\overline{\mathbf{E}}}^{\mathrm{t}*}, 2]} \tag{27}$$

where $\operatorname{sum}[\mathbf{A}, 2]$ means to calculate the summation of each row in matrix $\mathbf{A}$ and form a column vector. The vector division in (27) is element-wise division.

B. CONTRAST SOURCE INVERSION METHOD

The classical CSI algorithm [13] plays an important role in our physics-informed unsupervised deep learning frame-work. It is used to extract the physical knowledge and directly integrated in the framework. Here, the classic CSI algorithm written as matrix operations is provided.

We define two error functions named as the data error function $F_D(\bar{\bar{\mathbf{J}}}, \bar{\bar{\xi}})$ and the state error function $F_S(\bar{\bar{\mathbf{J}}})$, which indicate the mismatches in the data equation and the state equation respectively:

$$F_D(\bar{\bar{\mathbf{J}}}, \bar{\bar{\xi}}) = \frac{\left\| \bar{\bar{\xi}} \cdot \bar{\bar{\mathbf{E}}}^{\mathrm{i}} + \bar{\bar{\xi}} \cdot \bar{\bar{\mathbf{G}}}_D \cdot \bar{\bar{\mathbf{J}}} - \bar{\bar{\mathbf{J}}} \right\|^2}{\left\| \bar{\bar{\xi}} \cdot \bar{\bar{\mathbf{E}}}^{\mathrm{i}} \right\|^2} \tag{28}$$

$$F_S(\bar{\bar{\mathbf{J}}}) = \frac{\left\| \bar{\bar{\mathbf{E}}}^{\mathrm{s}} - \bar{\bar{\mathbf{G}}}_S \cdot \bar{\bar{\mathbf{J}}} \right\|^2}{\left\| \bar{\bar{\mathbf{E}}}^{\mathrm{s}} \right\|^2} \tag{29}$$

The CSI method aims to minimize the sum of data error and state error functions by iteratively optimizing the normalized contrast current density matrix $\bar{\bar{\mathbf{J}}}$ and the contrast matrix $\bar{\bar{\xi}}$:

$$\underset{\bar{\bar{\mathbf{J}}}, \bar{\bar{\zeta}}}{\operatorname{argmin}} \, F(\bar{\bar{\mathbf{J}}}, \bar{\bar{\zeta}}) = F_S(\bar{\bar{\mathbf{J}}}) + F_D(\bar{\bar{\mathbf{J}}}, \bar{\bar{\zeta}}) \tag{30}$$

The classic CSI algorithm firstly updates $\bar{\bar{\mathbf{J}}}$ using the Polak-Ribiére conjugate gradient method, and then updates $\bar{\bar{\xi}}$ by orderly calculating (9) and (27). Before giving the details of CSI, we define some auxiliary matrices and values for more succinct expression of CSI formulations:

$$\begin{aligned} \bar{\bar{\mathbf{A}}}_1 &= \bar{\bar{\mathbf{G}}}_S, \bar{\bar{\mathbf{B}}}_1 = \bar{\bar{\mathbf{E}}}^{\mathrm{s}} \\ \bar{\bar{\mathbf{A}}}_2 &= [\bar{\bar{\mathbf{I}}} - \bar{\bar{\xi}} \cdot \bar{\bar{\mathbf{G}}}_D], \bar{\bar{\mathbf{B}}}_2 = \bar{\bar{\xi}} \cdot \bar{\bar{\mathbf{E}}}^{\mathrm{i}} \\ c_1 &= 1 \Big/ \left\| \bar{\bar{\mathbf{E}}}^{\mathrm{s}} \right\|^2, c_2 = 1 \Big/ \left\| \bar{\bar{\xi}} \cdot \bar{\bar{\mathbf{E}}}^{\mathrm{i}} \right\|^2 \end{aligned} \tag{31}$$

where $\bar{\bar{\mathbf{I}}}$ represents the identity matrix. The value of $\bar{\bar{\xi}}$ remains constant in one iterative update of $\bar{\bar{\mathbf{J}}}$, so for the updating process of $\bar{\bar{\mathbf{J}}}$, the objective function $F(\bar{\bar{\mathbf{J}}}, \bar{\bar{\xi}})$ in (30) can be shortened to $F(\bar{\bar{\mathbf{J}}})$:

$$\underset{\bar{\bar{\mathbf{J}}}}{\operatorname{argmin}} \, F(\bar{\bar{\mathbf{J}}}) = c_1 \left\| \bar{\bar{\mathbf{B}}}_1 - \bar{\bar{\mathbf{A}}}_1 \cdot \bar{\bar{\mathbf{J}}} \right\|^2 + c_2 \left\| \bar{\bar{\mathbf{B}}}_2 - \bar{\bar{\mathbf{A}}}_2 \cdot \bar{\bar{\mathbf{J}}} \right\|^2 \tag{32}$$

The main steps of the classical CSI algorithm are as follows.

1) Use the BP method to calculate the initial values of $\bar{\bar{\mathbf{J}}}$ and $\bar{\bar{\xi}}$, represented as $\bar{\bar{\mathbf{J}}}_0$ and $\bar{\bar{\xi}}_0$ respectively.

2) For the n-th iteration ($n \geq 0$), put $\bar{\bar{\mathbf{J}}}_n$ and $\bar{\bar{\xi}}_n$ into (31) to update the auxiliary matrices and values $\bar{\bar{\mathbf{A}}}_2$, $\bar{\bar{\mathbf{B}}}_2$, and c_2. $\bar{\bar{\xi}}_n$ is the diagonalizable matrix of the contrast vector $\bar{\chi}_n$.

3) Calculate the gradient matrix $\bar{\bar{\mathbf{g}}}_n$ of the objective function $F(\bar{\bar{\mathbf{J}}})$ with respect to $\bar{\bar{\mathbf{J}}} = \bar{\bar{\mathbf{J}}}_n$:

$$\bar{\bar{\mathbf{g}}}_n = c_1 \bar{\bar{\mathbf{A}}}_1^H \cdot (\bar{\bar{\mathbf{A}}}_1 \cdot \bar{\bar{\mathbf{J}}}_n - \bar{\bar{\mathbf{B}}}_1) + c_2 \bar{\bar{\mathbf{A}}}_2^H \cdot (\bar{\bar{\mathbf{A}}}_2 \cdot \bar{\bar{\mathbf{J}}}_n - \bar{\bar{\mathbf{B}}}_2) \tag{33}$$

4) Calculate the Polak-Ribière conjugate gradient search direction $\bar{\bar{\mathbf{p}}}_n$:

$$\bar{\bar{\mathbf{p}}}_n = \begin{cases} \bar{\bar{\mathbf{g}}}_n, & n = 0 \\ \bar{\bar{\mathbf{g}}}_n + \bar{\bar{\mathbf{p}}}_{n-1} \cdot \bar{\bar{\beta}}_n, & n \geq 1 \end{cases} \tag{34}$$

where $\bar{\bar{\beta}}_n$ is the diagonalizable matrix of $\bar{\beta}_n$:

$$\bar{\beta}_n = \frac{\mathrm{Re}\left[\mathrm{sum}\left[\bar{\bar{\mathbf{g}}}_n \odot (\bar{\bar{\mathbf{g}}}_n - \bar{\bar{\mathbf{g}}}_{n-1})^*, 1 \right] \right]}{\mathrm{sum}\left[\bar{\bar{\mathbf{g}}}_n \odot \bar{\bar{\mathbf{g}}}_n^*, 1 \right]} \tag{35}$$

In (35), the function Re[∗] means taking the real value part of a complex value vector or matrix, and the vector division is element-wise operator.

5) Do one-dimensional search at the direction $\bar{\bar{\mathbf{p}}}_n$ to minimize the objective function:

$$\underset{\bar{\bar{\alpha}}_n}{\operatorname{argmin}} \, F\left(\bar{\bar{\mathbf{J}}}_{n+1} \right) = F\left(\bar{\bar{\mathbf{J}}}_n + \bar{\bar{\mathbf{p}}}_n \cdot \bar{\bar{\alpha}}_n \right) \tag{36}$$

where $\bar{\bar{\alpha}}_n$ is the diagonalizable matrix of the row vector $\bar{\alpha}_n$, who has analytical solution:

$$\bar{\alpha}_n = -\frac{\mathrm{sum}\left[\bar{\bar{\mathbf{g}}}_n \odot \bar{\bar{\mathbf{p}}}_n^*, 1 \right]}{\left(\begin{array}{l} c_1 \cdot \mathrm{sum}\left[(\bar{\bar{\mathbf{A}}}_1 \cdot \bar{\bar{\mathbf{p}}}_n) \odot (\bar{\bar{\mathbf{A}}}_1 \cdot \bar{\bar{\mathbf{p}}}_n)^*, 1 \right] + \\ c_2 \cdot \mathrm{sum}\left[(\bar{\bar{\mathbf{A}}}_2 \cdot \bar{\bar{\mathbf{p}}}_n) \odot (\bar{\bar{\mathbf{A}}}_2 \cdot \bar{\bar{\mathbf{p}}}_n)^*, 1 \right] \end{array} \right)} \tag{37}$$

The vector division is element-wise operator.

1) Update the normalized contrast current density:

$$\bar{\bar{\mathbf{J}}}_{n+1} = \bar{\bar{\mathbf{J}}}_n + \bar{\bar{\mathbf{p}}}_n \cdot \bar{\bar{\alpha}}_n \tag{38}$$

2) Update the contrast:

$$\bar{\chi}_{n+1} = \frac{\mathrm{sum}[\bar{\bar{\mathbf{J}}}_{n+1} \odot \bar{\bar{\mathbf{E}}}^{\mathrm{t}*}, 2]}{\mathrm{sum}[\bar{\bar{\mathbf{E}}}^{\mathrm{t}} \odot \bar{\bar{\mathbf{E}}}^{\mathrm{t}*}, 2]} \tag{39}$$

3) Execute Steps 2-7 repeatedly, until reaching the maximum number of iterations or getting the value of objective function $F(\bar{\bar{\mathbf{J}}}_n, \bar{\bar{\xi}}_n)$ less than predetermined threshold.

REFERENCES

[1] X. Chen, "Introduction," in *Proc. Comput. Methods Electromagn. Inverse Scattering*, 2017, pp. 1–12.

[2] R. Marklein et al., "Linear and nonlinear inversion algorithms applied in nondestructive evaluation," *Inverse Problems*, vol. 18, no. 6, pp. 1733–1759, Dec. 2002.

[3] R. Zoughi and S. Kharkovsky, "Microwave and millimetre wave sensors for crack detection," *Fatigue Fracture Eng. Mater. Struct.*, vol. 31, no. 8, pp. 695–713, Aug. 2008.

[4] S. Gholizadeh, "A review of non-destructive testing methods of composite materials," in *Proc. 15th Portuguese Conf. Fracture*, 2016, vol. 1, pp. 50–57.

[5] A. Neal, "Ground-penetrating radar and its use in sedimentology: Principles, problems and progress," *Earth-Sci. Rev.*, vol. 66, no. 3-4, pp. 261–330, Aug. 2004.

[6] J. A. Huisman, S. S. Hubbard, J. D. Redman, and A. P. Annan, "Measuring soil water content with ground penetrating radar: A review," *Vadose Zone J.*, vol. 2, no. 4, pp. 476–491, Nov. 2003.

[7] A. Abubakar, T. M. Habashy, V. L. Druskin, L. Knizhnerman, and D. Alumbaugh, "2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements," *Geophysics*, vol. 73, no. 4, pp. F165–F177, Jul./Aug. 2008.

[8] E. J. Bond, X. Li, S. C. Hagness, and B. D. Van Veen, "Microwave imaging via space-time beamforming for early detection of breast cancer," *IEEE Trans. Antennas Propag.*, vol. 51, no. 8, pp. 1690–1705, Aug. 2003.

[9] N. K. Nikolova, "Microwave imaging for breast cancer," *IEEE Microw. Mag.*, vol. 12, no. 7, pp. 78–94, Dec. 2011.

[10] X. Li and S. C. Hagness, "A confocal microwave imaging algorithm for breast cancer detection," *IEEE Microw. Wireless Compon. Lett.*, vol. 11, no. 3, pp. 130–132, Mar. 2001.

[11] X. Zhuge and A. G. Yarovoy, "A sparse aperture MIMO-SAR-based UWB imaging system for concealed weapon detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 509–518, Jan. 2011.

[12] D. M. Sheen, D. L. McMakin, and T. E. Hall, "Three-dimensional millimeter-wave imaging for concealed weapon detection," *IEEE Trans. Microw. Theory Techn.*, vol. 49, no. 9, pp. 1581–1592, Sep. 2001.

[13] P. M. van den Berg and R. E. Kleinman, "A contrast source inversion method," *Inverse Problems*, vol. 13, no. 6, pp. 1607–1620, Dec. 1997.

[14] P. M. van den Berg, A. L. van Broekhoven, and A. Abubakar, "Extended contrast source inversion," *Inverse Problems*, vol. 15, no. 5, pp. 1325–1344, Oct. 1999.

[15] P. M. van den Berg, A. Abubakar, and J. T. Fokkema, "Multiplicative regularization for contrast profile inversion," *Radio Sci.*, vol. 38, no. 2, Apr. 2003.

[16] A. Abubakar, P. M. van den Berg, and J. J. Mallorqui, "Imaging of biomedical data using a multiplicative regularized contrast source inversion method," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 7, pp. 1761–1771, Jul. 2002.

[17] W. C. Chew and Y. M. Wang, "Reconstruction of two-dimensional permittivity distribution using the distorted born iterative method," *IEEE Trans. Med. Imag.*, vol. 9, no. 2, pp. 218–225, Jun. 1990.

[18] L. Li, L. G. Wang, J. Ding, P. K. Liu, M. Y. Xia, and T. J. Cui, "A probabilistic model for the nonlinear electromagnetic inverse scattering: TM case," *IEEE Trans. Antennas Propag.*, vol. 65, no. 11, pp. 5984–5991, Nov. 2017.

[19] P. Rocca, M. Benedetti, M. Donelli, D. Franceschini, and A. Massa, "Evolutionary optimization as applied to inverse scattering problems," *Inverse Problems*, vol. 25, no. 12, Dec. 2009, Art. no. 123003.

[20] P. Rocca, G. Oliveri, and A. Massa, "Differential evolution as applied to electromagnetics," *IEEE Antennas Propag. Mag.*, vol. 53, no. 1, pp. 38–49, Feb. 2011.

[21] M. Pastorino, "Stochastic optimization methods applied to microwave imaging: A review," *IEEE Trans. Antennas Propag.*, vol. 55, no. 3, pp. 538–548, Mar. 2007.

[22] D. Silver et al., "Mastering the game of go without human knowledge," *Nature*, vol. 550, no. 7676, pp. 354–359, Oct. 2017.

[23] G. Hinton et al., "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.

[24] G. Pironkov, S. U. N. Wood, and S. Dupont, "Hybrid-task learning for robust automatic speech recognition," *Comput. Speech Lang.*, vol. 64, Nov. 2020, Art. no. 101103.

[25] A. Graves, A. R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2013, pp. 6645–6649.

[26] K. Lu, "Intelligent recognition system for high precision image significant features in large data background," in *Proc. Int. Conf. Cyber Secur. Intell. Analytics*, 2020, vol. 928, pp. 1056–1062.

[27] T. Tong, X. Mu, L. Zhang, Z. Yi, and P. Hu, "MBVCNN: Joint convolutional neural networks method for image recognition," in *Proc. Mater. Sci., Energy Technol., Power Eng. I*, 2017, vol. 1839, Art. no. 020091.

[28] H. Sun, Q. Zhang, H. Wang, and A. Li, "Research of images recognition method based on RBF neural network," in *Proc. 7th Int. Conf. Syst. Syst. Eng.*, 2012, pp. 24–26.

[29] K. Cho et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, *arXiv:1406.1078v3*.

[30] C. Escolano, M. R. Costa-Jussa, and J. A. R. Fonollosa, "From bilingual to multilingual neural-based machine translation by incremental training," in *Proc. 57th Annu. Meeting Assoc. Comput. Linguistics, Student Res. Workshop*, 2019, pp. 236–242.

[31] S. Kwon, B. H. Go, and J. H. Lee, "A text-based visual context modulation neural model for multimodal machine translation," *Pattern Recognit. Lett.*, vol. 136, pp. 212–218, Aug. 2020.

[32] I. J. Goodfellow et al., "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, vol. 27.

[33] C. Ledig et al., "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. 30th IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 105–114.

[34] Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, "StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 8789–8797.

[35] J. Xiao and X. Bi, "Multi-scale attention generative adversarial networks for video frame interpolation," *IEEE Access*, vol. 8, pp. 94842–94851, 2020.

[36] C. He, Y. Wan, Y. Gu, and F. L. Lewis, "Integral reinforcement learning-based multi-robot minimum time-energy path planning subject to collision avoidance and unknown environmental disturbances," *IEEE Control Syst. Lett.*, vol. 5, no. 3, pp. 983–988, Jul. 2021.

[37] B. Sangiovanni, G. P. Incremona, M. Piastra, and A. Ferrara, "Self-configuring robot path planning with obstacle avoidance via deep reinforcement learning," *IEEE Control Syst. Lett.*, vol. 5, no. 2, pp. 397–402, Apr. 2021.

[38] J. Yoo, D. Jang, H. J. Kim, and K. H. Johansson, "Hybrid reinforcement learning control for a micro quadrotor flight," *IEEE Control Syst. Lett.*, vol. 5, no. 2, pp. 505–510, Apr. 2021.

[39] A. Massa, D. Marcantonio, X. Chen, M. Li, and M. Salucci, "DNNs as applied to electromagnetics, antennas, and propagation—A review," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 11, pp. 2225–2229, Nov. 2019.

[40] 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.

[41] Z. Wei and X. Chen, "Deep-learning schemes for full-wave nonlinear inverse scattering problems," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1849–1860, Apr. 2019.

[42] 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. Interv.*, 2015, vol. 9351, pp. 234–241.

[43] Y. Sun, Z. H. Xia, and U. S. Kamilov, "Efficient and accurate inversion of multiple scattering with deep learning," *Opt. Exp.*, vol. 26, no. 11, pp. 14678–14688, May 2018.

[44] Y. Zhou, Y. Zhong, Z. Wei, T. Yin, and X. Chen, "An improved deep learning scheme for solving 2-D and 3-D inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 69, no. 5, pp. 2853–2863, May 2021.

[45] K. Xu, L. Wu, X. Ye, and X. Chen, "Deep learning-based inversion methods for solving inverse scattering problems with phaseless data," *IEEE Trans. Antennas Propag.*, vol. 68, no. 11, pp. 7457–7470, Nov. 2020.

[46] W. Shao and Y. Du, "Microwave imaging by deep learning network: Feasibility and training method," *IEEE Trans. Antennas Propag.*, vol. 68, no. 7, pp. 5626–5635, Jul. 2020.

[47] Y. Chen, M. Zhong, Z. Guan, and F. Han, "A tailored semiphysics-driven artificial neural network for electromagnetic full-wave inversion," *IEEE Trans. Antennas Propag.*, vol. 70, no. 8, pp. 6207–6217, Aug. 2022.

[48] W. Shao and B. Zhou, "Dielectric breast phantoms by generative adversarial network," *IEEE Trans. Antennas Propag.*, vol. 70, no. 8, pp. 6256–6264, Aug. 2022.

[49] P. Isola, J. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 5967–5976.

[50] Z. Ma, K. Xu, R. Song, C. F. Wang, and X. Chen, "Learning-based fast electromagnetic scattering solver through generative adversarial network," *IEEE Trans. Antennas Propag.*, vol. 69, no. 4, pp. 2194–2208, Apr. 2021.

[51] L. Guo, G. F. Song, and H. S. Wu, "Complex-valued pix2pix-deep neural network for nonlinear electromagnetic inverse scattering," *Electronics*, vol. 10, no. 6, Mar. 2021, Art. no. 752.

[52] R. Song, Y. Huang, X. Ye, K. Xu, C. Li, and X. Chen, "Learning-based inversion method for solving electromagnetic inverse scattering with mixed boundary conditions," *IEEE Trans. Antennas Propag.*, vol. 70, no. 8, pp. 6218–6228, Aug. 2022.

[53] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, vol. 378, pp. 686–707, 2019.

[54] 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.

[55] Z. Wei and X. Chen, "Physics-inspired convolutional neural network for solving full-wave inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 67, no. 9, pp. 6138–6148, Sep. 2019.

[56] R. Guo et al., "Physics embedded deep neural network for solving full-wave inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 70, no. 8, pp. 6148–6159, Aug. 2022.

[57] R. Guo et al., "Physics embedded deep neural network for solving volume integral equation: 2-D case," *IEEE Trans. Antennas Propag.*, vol. 70, no. 8, pp. 6135–6147, Aug. 2022.

[58] Y. Sanghvi, Y. Kalepu, and U. K. Khankhoje, "Embedding deep learning in inverse scattering problems," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 46–56, 2020.

[59] G. Oliveri, M. Salucci, N. Anselmi, and A. Massa, "Compressive sensing as applied to inverse problems for imaging theory, applications, current trends, and open challenges," *IEEE Antennas Propag. Mag.*, vol. 59, no. 5, pp. 34–46, Oct. 2017.

[60] J. Hunt et al., "Metamaterial apertures for computational imaging," *Science*, vol. 339, no. 6117, pp. 310–313, Jan. 2013.

[61] L. Li et al., "Electromagnetic reprogrammable coding-metasurface holograms," *Nature Commun.*, vol. 8, Aug. 2017, Art. no. 197.

[62] H. T. Zhao, Y. Shuang, M. L. Wei, T. J. Cui, P. del Hougne, and L. L. Li, "Metasurface-assisted massive backscatter wireless communication with commodity Wi-Fi signals," *Nature Commun.*, vol. 11, no. 1, Aug. 2020, Art. no. 3926.

[63] L. Li, H. Zhao, C. Liu, L. Li, and T. J. Cui, "Intelligent metasurfaces: Control, communication and computing," *eLight*, vol. 2, no. 1, May 2022, Art. no. 7.

[64] C. Liu, L. Li, and T. J. Cui, "Physics-informed unsupervised deep learning framework for solving full-wave inverse scattering problems," in *Proc. IEEE Conf. Antenna Meas. Appl.*, 2022.

[65] X. Chen, "Reconstructing dielectric scatterers," in *Proc. Comput. Methods Electromagn. Inverse Scattering*, 2017, pp. 123–182.

[66] Y. LeCun, C. Cortes, and C. J. C. Burges, "The MNIST database of handwritten digits," MNIST, 2012. [Online]. Available: http://yann.lecun.com/exdb/mnist/

[67] J. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. 16th IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2242–2251.

[68] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *Proc. 34th Int. Conf. Mach. Learn.*, 2017, vol. 70, pp. 214–223.

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

[70] S. Dasgupta and A. Gupta, "An elementary proof of a theorem of Johnson and Lindenstrauss," *Random Struct. Algorithms*, vol. 22, no. 1, pp. 60–65, Jan. 2003.

[71] L. L. Li et al., "Machine-learning reprogrammable metasurface imager," *Nature Commun.*, vol. 10, Mar. 2019, Art. no. 1082.

[72] L. Li et al., "Intelligent metasurface imager and recognizer," *Light-Sci. Appl.*, vol. 8, Oct. 2019, Art. no. 97.

[73] H. Li et al., "Intelligent electromagnetic sensing with learnable data acquisition and processing," *Patterns*, vol. 1, no. 1, Apr. 2020, Art. no. 100006.

[74] Z. Wang, H. R. Zhang, H. T. Zhao, T. J. Cui, and L. L. Li, "Intelligent electromagnetic metasurface camera: System design and experimental results," *Nanophotonics*, vol. 11, no. 9, pp. 2011–2024, May 2022.

CHE LIU (Member, IEEE) was born in Suzhou, Jiangsu, China, in 1993. He received the B.Eng. degree in information science and technology and the Ph.D. degree from Southeast University, Nanjing, China, in 2015 and 2022, respectively. He is currently a Zhishan Postdoctor with Southeast University.

His research interests include computational electromagnetic, meta-material, and deep learning. He is committed to use artificial intelligence technology solving electromagnetic issues, including ISAR imaging, holographic imaging, inverse scattering imaging, automatic antenna design, and diffraction neural network.

HONGRUI ZHANG was born in Shandong, China, in 1998. He received the B.S. degree from the Beijing Institute of Technology, Beijing, China, in 2020. He is currently working toward the Ph.D. degree with the Department of Electronics, Peking University, Beijing, China. His research interests include intelligent electromagnetic imaging and sensing based on metamaterials.

LIANLIN LI (Senior Member, IEEE) was born in 1980. He received the Ph.D. degree from the Institute of Electronics, Chinese Academy of Sciences, Beijing, China, in 2006. He is currently a hundred talented program Professor with Peking University, Beijing, China. His research interests include super-resolution imaging, microwave imaging, sparse signal processing, and ultrawideband radar systems.

TIE JUN CUI (Fellow, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees from Xidian University, Xi'an, China, in 1987, 1990, and 1993, respectively. In March 1993, he joined the Department of Electromagnetic Engineering, Xidian University, and was promoted to an Associate Professor in November 1993. From 1995 to 1997, he was a Research Fellow with the Institut fur Hochstfrequenztechnik und Elektronik (IHE), University of Karlsruhe, Karlsruhe, Germany. In July 1997, he joined the Center for Computational Electromagnetics, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Champaign, IL, USA, first as a Postdoctoral Research Associate and then as a Research Scientist. In September 2001, he was a Cheung-Kong Professor with the Department of Radio Engineering, Southeast University, Nanjing, China. In January 2018, he became the Chief Professor of Southeast University. He is an Academician of the Chinese Academy of Science. He is the first author of the books *Metamaterials: Theory, Design, and Applications* (Springer, November 2009), *Metamaterials: Beyond Crystals, Noncrystals, and Quasicrystals* (CRC Press, March 2016), and *Information Metamaterials* (Cambridge University Press, 2021). He has authored or coauthored more than 600 peer-reviewed journal articles, which have been cited by more than 51 000 times (H-Factor 114), and licensed more than 150 patents. His research has been selected as one of the most exciting peer-reviewed optics research Optics in 2016 by Optics and Photonics News Magazine, ten Breakthroughs of China Science in 2010, and many Research Highlights in a series of journals. His work has been widely reported by *Nature News, MIT Technology Review, Scientific American, Discover*, and *New Scientists*. He was the recipient of the Research Fellowship from Alexander von Humboldt Foundation, Bonn, Germany, in 1995, Young Scientist Award from the International Union of Radio Science in 1999, Cheung Kong Professor by the Ministry of Education, China, in 2001, and National Science Foundation of China for Distinguished Young Scholars in 2002. He was also the recipient of the Natural Science Award (first-class) from the Ministry of Education, China, in 2011, and National Natural Science Awards of China (second-class, twice) in 2014 and 2018.

Prof. Cui was an Associate Editor for IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and a Guest Editor of *Science China-Information Sciences*, *Science Bulletin*, IEEE JOURNAL ON EMERGING AND SELECTED TOPICS IN CIRCUITS AND SYSTEM, and *Research*. He is the Chief Editor of Metamaterial Short Books in Cambridge University Press, an Editor of *Materials Today Electronics*, an Associate Editor for *Research*, and an Editorial Board Member of *National Science Review*, *eLight*, *PhotoniX*, *Advanced Optical Materials*, *Small Structure*, and *Advanced Photonics Research*. He presented more than 100 Keynote and Plenary Talks in Academic Conferences, Symposiums, or Workshops. From 2019 to 2022, he was ranked in the top 1% for the highly cited articles in the field of Physics by Clarivate Web of Science (Highly Cited Researcher).