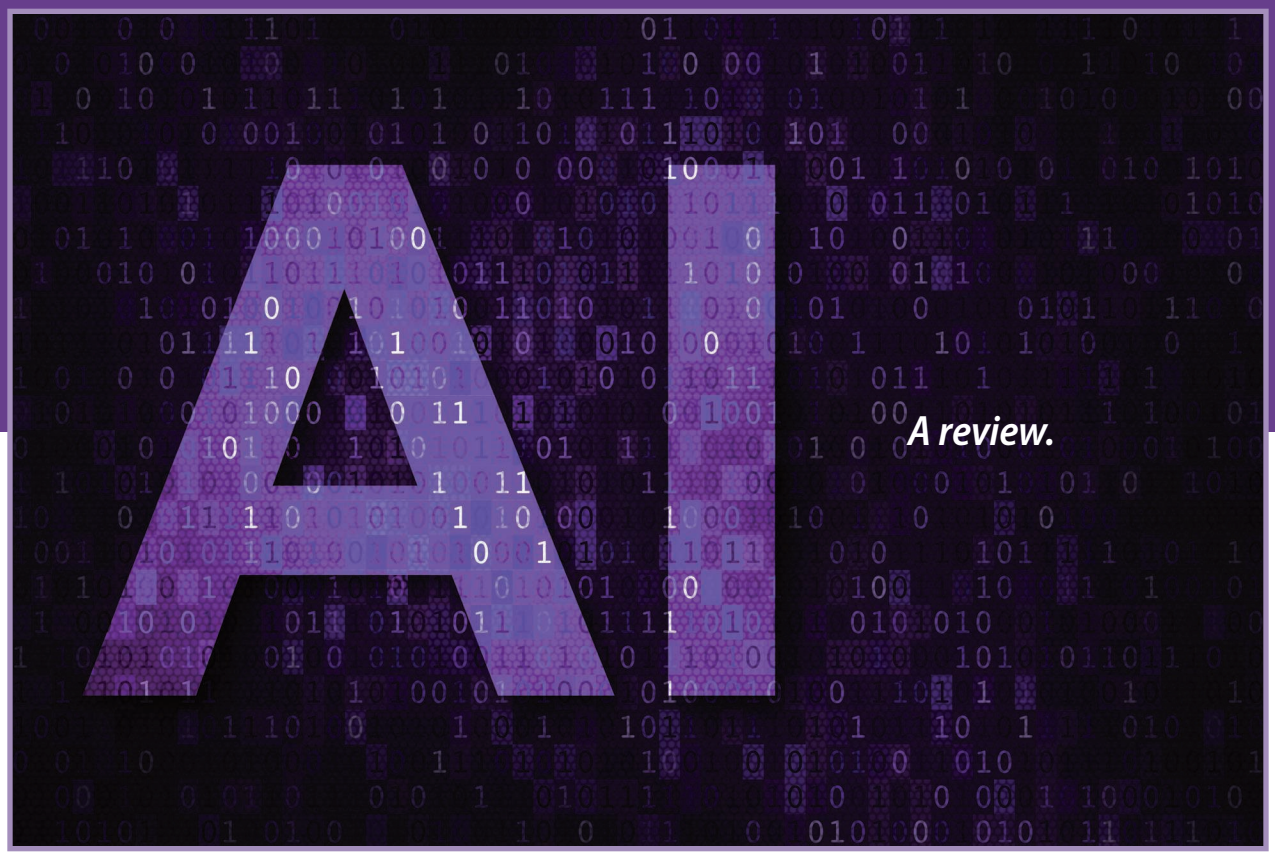

Machine Learning in Electromagnetics With Applications to Biomedical Imaging

Maokun Li, Rui Guo, Ke Zhang, Zhichao Lin, Fan Yang, Shenheng Xu, Xudong Chen, Andrea Massa, and Aria Abubakar

Biomedical imaging is a relevant noninvasive technique aimed at generating an image of the biological structure under analysis. The arising visual representation of the characteristics of the object is affected by both the measurement process and reconstruction algorithm. This procedure can be considered as a hybridization of data information, measurement physics, and prior information.

Recently, machine learning techniques, especially deep learning, have been widely used in image and signal processing [13]. With the help of big data, massive parallelization, and computational optimization, the learning

Digital Object Identifier 10.1109/MAP.2020.3043469
Date of current version: 14 January 2021

1045-9243/21©2021IEEE

and generalization abilities of machine learning techniques have increased significantly.

This can also be fruitfully exploited in reconstruction algorithms in biomedical imaging. With prior information derived from training data and embedded in the neural network, both reconstruction speed and accuracy can be further improved. Deep learning-based biomedical imaging is a rapidly developing field of research. In this article, we provide a review of the state-of-the-art progress in this framework, and we tentatively suggest that the hybridization of physical laws and machine learning may be a good research direction to further improve the efficiency and accuracy in biomedical imaging algorithms.

With the help of big data, massive parallelization, and computational optimization, the learning and generalization abilities of machine learning techniques have increased significantly.

INTRODUCTION

Biomedical imaging is an essential tool for medical diagnosis and treatment. Typical imaging modalities include medical X-ray imaging, magnetic resonance imaging (MRI), and ultrasonography [1]. Many emerging medical imaging modalities are also developed, such as positron emission tomography, microwave imaging, electrical impedance tomography (EIT), and photo/thermal-acoustic imaging [2]–[4]. In many biomedical imaging technologies, electromagnetic (EM) fields and waves play an important role, serving directly as an interrogating/sensing tool.

In these modalities, the imaging process can be generally formulated as an inverse problem. It can be solved using linear or nonlinear inversion algorithms that convert the measured data into images pictorially representing the distribution of physical parameters. The physics of measurement, i.e., the forward modeling that emulates the measurement process, plays an important role.

Inverse problems are usually ill posed because the measured data cannot fully determine the physical parameters, and the reconstructed models are also prone to noise in the data and errors in the measurement [5]. As a remedy, regularization methods are necessary to stabilize the inversion by bringing additional information to guide the inversion to a reasonable solution [6]–[8]. From a general viewpoint, the image reconstruction is a hybridization of measurement physics, measured data, and prior information. These three types of information are very different; therefore, designing a robust, accurate, and efficient inversion algorithm for biomedical imaging is not straightforward.

Furthermore, many biomedical EM applications require real-time imaging to monitor the dynamics of human physical conditions. This poses another challenge for biomedical imaging using EM waves. Within the limits of computational resources, the design of real-time imaging algorithms requires a good balance between image quality and computational complexity.

Traditional imaging algorithms formulate the underlying inversion problem as the minimization between measured and simulated data [9]. Iterative or noniterative methods can be used to search the optimal model parameters such that the simulated data matches the measured values [10], [11]. Such a process can be computationally intensive, especially when the forward modeling has a large computational complexity.

To meet the need of real-time imaging, linearized algorithms, such as the Born approximation, are often used [2], [12]. These algorithms avoid the online computation of the forward model, but their accuracy may deteriorate for complex or high-contrast targets where multiple scattering cannot be ignored.

In modern technology, artificial neural networks (ANNs) with millions of parameters can be trained to provide the deep neural network (DNN) with good learning and generalization capabilities.

Machine learning methods have also been applied to biomedical imaging. They can help to incorporate prior information into the inversion process through the training procedure. Prior information "learned" from data can be more versatile than that provided using traditional methods, such as regularization. For example, information about human body structures can be learned from training data to improve the quality of reconstructed models. Therefore, it can help to improve the quality of the reconstruction.

Moreover, after the training process, the online computation for DNNs can be very efficient since their architecture is suitable for massive parallelization on high-performance computing platforms like GPUs. Such an advantage can also help real-time imaging. Accordingly, the application of machine learning techniques to biomedical imaging has recently become an important research topic.

Deep learning techniques can be applied to the imaging process in several stages. They can be directly applied to preprocess measured data [14]–[16]. For instance, filters for noise removal or signal classification can be constructed with a trained network based on previous data. In the imaging process, end-to-end DNNs can be trained to transform the measured data directly to images [17], [18]. In these approaches, the physics of measurement is implicitly embedded in the network, and

the computation can be accelerated with the help of massive parallelization.

Deep learning techniques can also be exploited for constructing better regularization methods to improve the image quality [19]–[21]. Furthermore, physical simulations can be hybridized together with the neural network [22]–[25]. With the contribution of both physics and prior information learned from data, these methods have good learning and generalization capabilities while also improving the inversion accuracy. The integration of physics and machine learning techniques reveals a possible approach to connecting physics and human experience to solve imaging problems in the real world.

THE THEORY OF MACHINE LEARNING

Machine learning is a scientific discipline to design and develop algorithms that allow computers to evolve behaviors based on empirical data. More specifically, the algorithms build a mathematical model based on training data to make predictions or decisions without being explicitly programmed to perform the task [26], [27].

Machine learning methods can be mainly divided into three kinds:

- *Supervised learning*: The machine is given a training data set that contains input–output pairs to learn to produce the output when an input is given. This is an object-oriented technique that learns the model from training examples through constrained optimization.
- *Unsupervised learning*: No desired output is provided in the training set. The goal of the machine is to build optimal representations or find implicit patterns of the input data according to a human's criteria. Two classic examples are clustering and dimensionality reduction.
- *Reinforcement learning*: The machine interacts with its environments by producing actions that will, in turn, give rewards or punishments back to the machine. The goal of reinforcement learning is to maximize the future rewards it receives during the lifetime of the task.

Most recent developments of biomedical applications that use machine learning belong to the supervised or unsupervised learning categories. Here, we do not aim to highlight all learning techniques one by one. Instead, we briefly discuss several representative machine learning methods that we may come across in the "Applications in Biomedical Signal Processing" and "Applications in Image Reconstruction" section. More fundamental knowledge as well as technical details about machine learning methods can be found in [27]–[29], which are very good tutorials for machine learning beginners.

SUPPORT VECTOR MACHINE

A support vector machine (SVM) is a classic supervised learning model that constructs a hyperplane in a high-dimensional space to discriminate two types of data or estimate the trends of data. For both classification and regression problems, the formula of the hyperplane can be written as

$$f(\boldsymbol{x}) = \boldsymbol{w}^T \boldsymbol{\varphi}(\boldsymbol{x}) + b, \tag{1}$$

where $\boldsymbol{w}$ and b are the unknown normal vector and bias coefficient, respectively, and $\boldsymbol{\varphi}(\cdot)$ denotes the nonlinear function (known as the *kernel*) that maps the input data $\boldsymbol{x}$ to a high-dimensional space, where the transformed data can be distinguished or estimated by a linear hyperplane (see Figure 1).

The training phase of an SVM can be formulated as an optimization process:

$$(\boldsymbol{w}, b) = \arg\left\{\min_{\boldsymbol{w},b}\left(\frac{1}{2}\|\boldsymbol{w}\|^2 + \lambda \sum_{i=1}^{m} \ell(f(\boldsymbol{x}_i), y_i)\right)\right\}, \tag{2}$$

where λ is the regularization coefficient, and $\ell(\cdot)$ denotes the loss function. In the classification problem, the value of $\ell(\cdot)$ is proportional to the number of error samples, while, in the regression problem, its value is proportional to the distance between each point and the hyperplane. The minimization problem in (2) can be efficiently solved in its dual form [30].

DICTIONARY LEARNING

Dictionary learning refers to the strategy that learns a set of sparse bases through a training stage [31]. The learned bases are

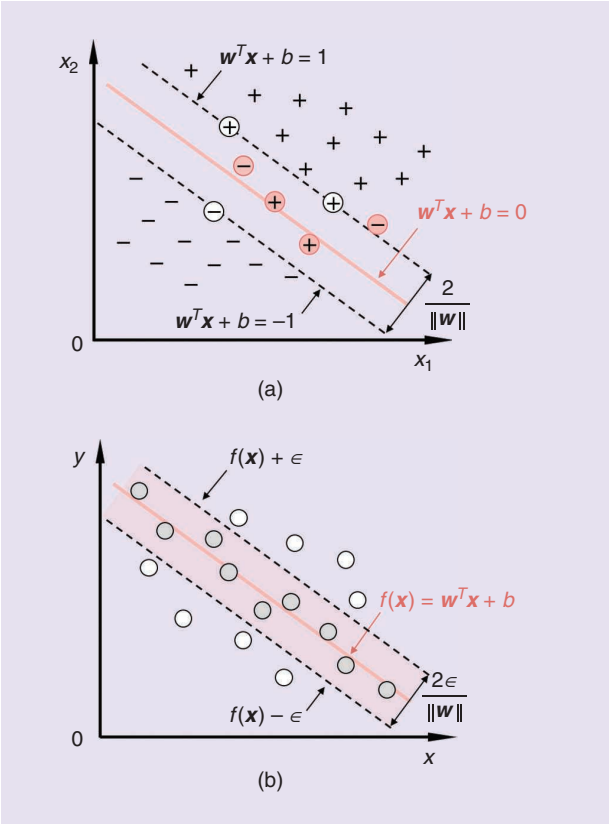

FIGURE 1. A schematic of SVM hyperplanes for (a) classification and (b) regression problems. In (a), the black circles on the demarcation plane are the so-called support vectors, and the red circles are some tolerable error samples. In (b), ϵ denotes the tolerable deviation between the transformed data and the hyperplane.

called the *dictionary*, which is often overcomplete. Compared with the conventional sparse-coding strategy, which uses a prespecified basis, the learning approach can encode more prior knowledge embedded in the training data.

Given a set of training signals $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_N] \in \mathbb{R}^{M \times N}$, we assume that there exists a dictionary $\mathbf{D} \in \mathbb{R}^{M \times K}$ that can sparsely encode all signals of $\mathbf{S}$. The training process can be mathematically expressed as

$$(\mathbf{D}, \mathbf{\Gamma}) = \arg\left\{ \min_{\mathbf{D}, \mathbf{\Gamma}} \| \mathbf{S} - \mathbf{D}\mathbf{\Gamma} \|_F^2 \,\middle|\, \forall i, \| \boldsymbol{\gamma}_i \|_0 \le K_0 \right\}, \tag{3}$$

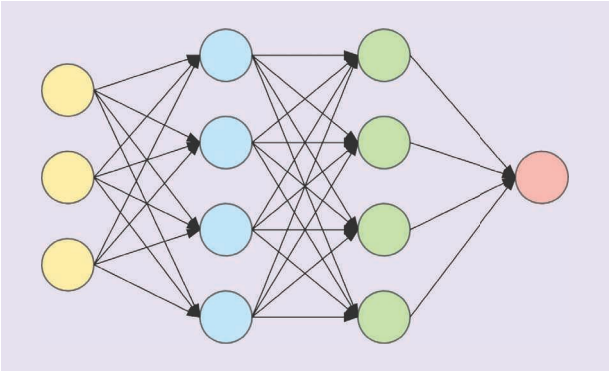

FIGURE 2. The structure of an FCN. The first layer is the input layer, the last layer is the output layer, and the layers between them are all hidden layers. The number of units in the output layer is not necessarily equal to one.

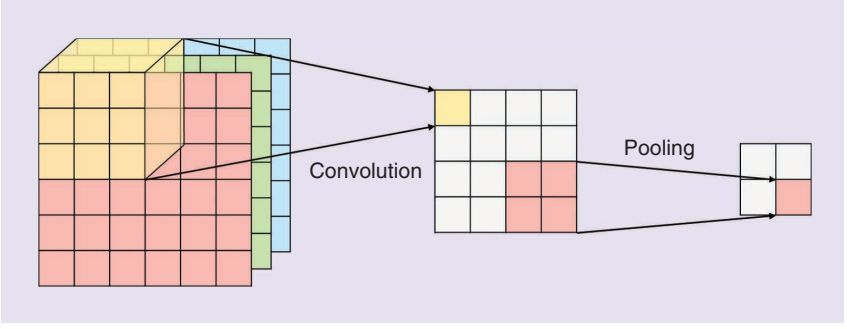

FIGURE 3. A schematic of one convolution layer and one pooling layer. The kernels of the convolution layer are trainable, while the pooling layer has no training parameter.

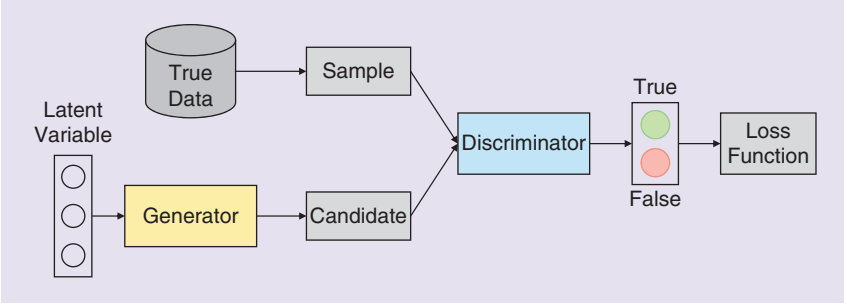

FIGURE 4. The structure of a GAN, which consists of a generator and discriminator. Generally, the generator is a deconvolutional neural network, and the discriminator is a CNN.

where the meaning of | is "be subject to"; $\|\cdot\|_F$ and $\|\cdot\|_0$ are the matrix Frobenius and L_0 norms, respectively; K_0 is the sparsity; and $\mathbf{\Gamma} \in \mathbb{R}^{K \times N}$ is the sparse representation, with each column $\boldsymbol{\gamma}_i$ being the coding coefficients corresponding to each $\mathbf{s}_i$. A common way to solve (3) is to iteratively alternate the optimization with regard to $\mathbf{D}$ and $\mathbf{\Gamma}$ until convergence [32], [33].

DNNs

DNNs are considered as a subset of ANNs. Some well-known and well-developed network structures include the fully connected neural network (FCN) [34], convolutional neural network (CNN) [35], and recurrent neural network [36].

The universal approximation theorem theoretically verified the ability of an FCN to approximate arbitrary continuous functions [34]. An FCN has multiple hidden layers with nonlinear modules embodied in each (see Figure 2). For each node, $y_i = f_i(\Sigma_{i,j} W_{i,j} x_j + b_i) = f_i(W_i^T x + b_i)$, where W is the weight matrix, b is the bias vector, and $f_i(\cdot)$ is the nonlinear activation function that ensures the nonlinear fitting ability of the network. However, when the dimensions of the input or output vector increase, the number of unknowns to train in the FCN increases very quickly and makes it nearly impossible to compute due to the limitations of the hardware.

A CNN greatly reduces the number of training parameters using convolutional layers. The two main advantages of convolutional layers over fully connected layers are parameter sharing and sparsity of connections. In addition, a CNN often uses pooling layers to reduce the size of their representation to speed up computation as well as make some of the features it detects more robust. As shown in Figure 3, convolving n_k kernels of size $f_h \times f_w \times n_c$ with a $n_h \times n_w \times n_c$ input layer will result in a feature map of size $(n_h - f_h + 1) \times (n_w - f_w + 1) \times n_k$, and the pooling layer can halve the input size. In the image processing field, various CNN architectures have been developed, such as VGG [37], residual neural network [38], and U-Net [39].

GENERATIVE ADVERSARIAL NETWORKS

Generative adversarial networks (GANs) [40] show great power to approximate real images and are capable of creating realistic-looking visual representations. The main framework of GANs is illustrated in Figure 4, which includes two parts: a generator and discriminator. The generator generates "fake" candidates, while the discriminator distinguishes them from the true data distribution. Specifically, the generator transforms the latent variables sampled from the prior distribution into a synthesized image, while the discriminator receives both the

true and synthesized images, and it discriminates which one is fake by itself.

In such a situation, the training objective of the generator is to increase the error rate of the discriminator, which is trained to recognize the synthesized samples as much as possible. Theoretically, the generator and discriminator will eventually reach a balance where the error rate of the discriminator is equal to 0.5. One promising application of GANs in the field of image reconstruction is to train the generator as the network for inversion and the discriminator to evaluate the difference between the reconstructed image and real data.

Machine learning is a scientific discipline to design and develop algorithms that allow computers to evolve behaviors based on empirical data.

APPLICATIONS IN BIOMEDICAL SIGNAL PROCESSING

Biomedical signals usually reflect the physiological activities of certain tissues or organs in the human body, and they are widely used or studied in clinical practice for purposes such as diagnosis and monitoring. Some biomedical signals are directly measured on the human body without an external excitation, such as electrocardiogram (ECG), electroencephalogram (EEG), and electromyography signals. In contrast, other biomedical signals are the response of a part of the human body in the presence of external excitations, such as microwave and electrical currents. Examples are the scattered fields of a woman's breast when excited by microwaves and the induced voltages around a human thorax when excited by currents.

In recent decades, machine learning has been extensively explored for biomedical signal processing problems, such as signal classification, signal extraction, and noise reduction. What follows is a brief summary of published works.

Biomedical signal classification is useful in a number of applications, such as the detection of certain diseases [41]–[48], recognition of limb or body movement [49]–[52], mental tasks classification [53], and fall motion detection [14]. In the application of machine learning to biomedical signal processing, a large proportion of published works are related to the classification of biomedical signals.

Traditionally, the classification task is divided into two substeps: 1) the processing of raw signal, including preprocessing, feature extraction/selection, and so on, and 2) the classification of the features obtained in step one using a classifier. In the first step, preprocessing usually includes noise reduction, normalization, and so on. Feature extraction/selection seeks the essential features that best represent the signals. The quality of features usually determines the performance of classifiers to a large extent, and irrelevant features may degrade the overall performance of classifiers. Various features have been used by researchers, such as those based on discrete Fourier transform [14], [51], discrete wavelet transform [42], [44], [45], principal component analysis [46], [47], entropy [41], and other methods [43], [50], [53].

For the second step, a variety of machine learning-based classifiers have been developed, such as those based on decision trees [45], SVM [44], [46], [47], [49], neural networks [14], [41], [42], [51], [53], and other methods [50]. Recently, these feature extraction/selection processes have done automatically by neural networks. This enables the training of an end-to-end network for biomedical signal classification [48], [52], [54].

Biomedical signal extraction is the process of attaining the desired signals from measured composite signals. Recently, machine learning-based methods have been demonstrated to be effective for some extraction problems. For example, in fetal ECG extraction, a number of methods have been proposed based on the principle of the adaptive noise canceler (ANC) with the integration of machine learning models, such as adaptive neuron–fuzzy inference systems [55] and neural networks [56]–[58]. Another example is the extraction of EEG signals from the original EEG signals, which are interfered with by ocular, muscular, and/or cardiac signals. For this problem, the ANC approach was proven effective, and many works were based on the combination of ANC and machine learning models [59]–[61].

Biomedical signal noise reduction aims at reducing the noise or artifacts in the signal, such as the white noise, baseline wander, and movement artifacts. Recently, methods based on neural networks, such as ANNs [62], [63] and denoising autoencoders [64], [65], have shown promising results on various biomedical signals and for a wide variety of noise sources.

In summary, biomedical signals are usually corrupted by noise and other interferences [66], and their interpretation traditionally calls for well-trained human experts. The application of machine learning methods to biomedical signal processing may help remove interferences in the signals on one hand and, on the other hand, enable automatic interpretation of the measured signal.

APPLICATIONS IN IMAGE RECONSTRUCTION

Image reconstruction is the process of converting measurement (data) to images (models). We tentatively classify image reconstruction techniques into direct, iterative, and machine learning approaches.

Direct imaging (e.g., back projection and the inverse Fourier transform [12]) uses approximate analytical expressions to complete the conversion. Such methods are based on the Born approximation of the EM scattering equations to linearize the imaging problem. The reconstruction is efficient, but some image artifacts generally arise due to the undersampled data or neglect of multiple scattering effects.

The iterative approach seeks the model that minimizes data mismatch within a constrained model space using the gradient

descent method [67]. More specifically, the data mismatch measures the discrepancy between simulated and measured data, while the model space is usually constrained by regularization set according to image priors. Some well-known regularizers includes smoothest constraint [68], total variation [6], wavelet transform [7], dictionary learning [8], and so on. It is worth pointing out that the iterative approach usually generates more accurate images than the direct approach, but the process is time consuming. Furthermore, the regularization represented by rigorous mathematical forms is sometimes not sufficient to reconstruct an accurate image.

The machine learning method mines the implicit features embedded in the training data set by using the learned parameters to guide image reconstruction. Since the prior knowledge of the groundtruth is encoded in the learned parameters, the accuracy and efficiency of reconstruction can be enhanced. Machine learning methods can be categorized into two main categories. The first is a direct data-driven approach, and the second method learns prior knowledge in conventional inversion algorithms.

DIRECT DATA-DRIVEN APPROACHES

With the learning approach, it is straightforward to build a direct map connecting the data and model spaces based on a model set and the corresponding simulated model response. In 2002, a radial basis-function neural network (RBFNN) was trained to solve microwave medical imaging problems [69]. Due to the limitation of computing power at that time, the output of the RBFNN contains only three parameters, namely, the position and size of proliferated marrow inside the bone associated with a limp. RBFNN was also applied to EIT in [70] and [71] by setting the input as boundary voltage measurements and the output as conductivity in discrete triangle pixels.

In 2006, an SVM-based approach was used for microwave breast cancer detection [72], where the SVM input is the electric field intensity measured at receiver locations, while the output is the probability maps of tumor presence. The SVM was also applied to the classification between hemorrhagic and ischemic brain strokes as well as their localization [73]. In recent years, researchers began to train end-to-end DNNs for the EIT problem [17], [74], [75]. We borrow a figure from [17] to illustrate this type of work, shown in Figure 5. Although the network structures in the three studies are different, they all showed promising results with a properly designed training data set.

Similar approaches have been applied to other medical imaging modalities, such as MRI and MR–electrical properties tomography (MR-EPT) [18], [76]. In [76], the authors presented a mathematical description of the manifold learning process. In [18], a conditional GAN was used for data-to-image reconstruction, and the arising method turned out to be robust to the noise and outperform conventional MR-EPT techniques. Another work for MR-EPT can be found in [77], where a 3D patch-based CNN was trained to predict conductivity maps from B1 phase data.

Despite the good results in using neural networks to approximate the imaging process that maps data to a model, it is still unclear whether the neural network can fully learn the underlying physics of the measurement and behave the same as or outperform traditional imaging methods. This will affect the learning and generalization abilities; i.e., we do not know the boundary of the prediction and when it will fail for a new data set. A poorly designed network may easily fail for a new data set. Therefore, such methods are applicable only when we are confident that the training and testing models share the same space.

One way to make trained networks more adaptive is to hybridize the learning techniques with traditional imaging methods that are based on physical modeling. For example, we can preprocess the network input by some inversion operators [78], such as back projection or even nonlinear inversion. After that, the task for the network can be focused

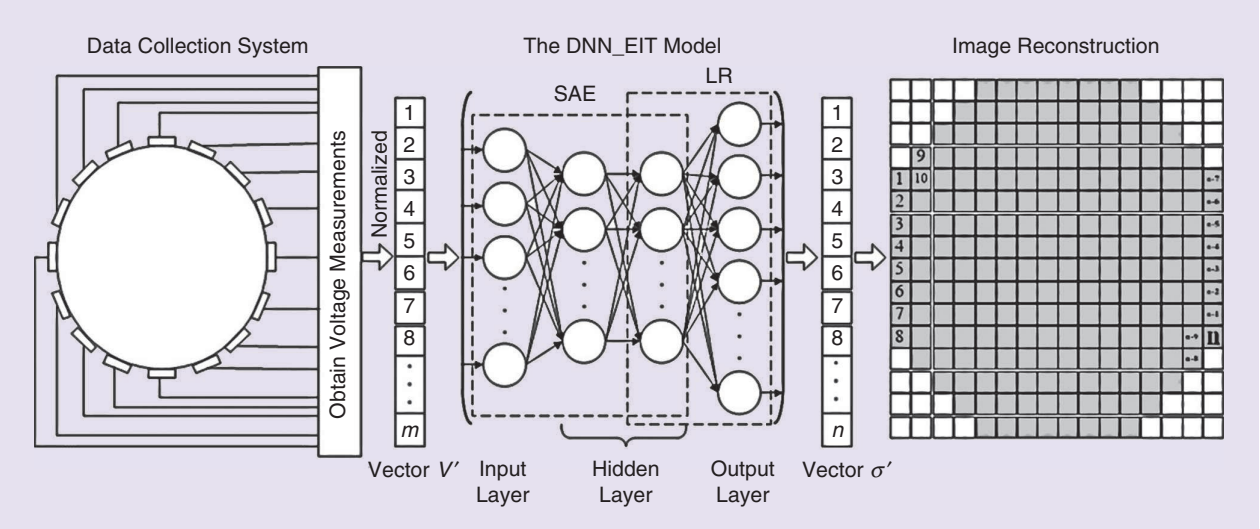

FIGURE 5. The DNN for EIT in [17]. The inputs are measured EIT voltages and the outputs are conductivity pixels. SAE: stacked autoencoder; LR: logistic regression. (© 2019 SAGE Publications; used with permission.)

on building an image-to-image map to improve the quality of reconstruction rather than doing all of the task of physics-based inversion.

Following this idea, some approaches similar to image enhancement/denoising have been studied. For example, in EIT, the authors of [79] took the conductivity model from one-step Gauss–Newton inversion as the input of a neural network to output the conductivity model with a much higher resolution. The network here behaves as a postprocessor to alleviate artifacts resulting from imperfect inversion.

In [80], Hamilton and Hauptmann used U-Net for real-time EIT by setting the U-Net input to the recovered approximate conductivity from the D-bar approach [81], the output being the high-quality image of conductivity distribution. In [82]–[84], CNNs were used for the superresolution of microwave imaging, with the inputs of the CNNs being inverted models from conventional algorithms, while the outputs are the groundtruth.

In computerized tomography (CT), related works reconstruct high-quality images from low-resolution images [10], [85]–[88]. An example from [10] is shown in Figure 6. In [88], the authors tried transfer learning from a 2D trained network to 3D CNN for low-dose CT by using the Wasserstein GAN to improve the performance of the network.

In [89], residual CNNs were trained to learn the aliasing artifacts brought by down-sampled MR brain data. In [90], a GAN with pixelwise, frequency-domain, and perceptual loss was proposed for high-resolution MRI reconstruction. In [91], a cycle GAN was proposed for MRI reconstruction, the cyclic loss being a combination of undersampled frequency loss and the fully reconstructed image loss.

Also for MRI, Han et al. [92] provided a new perspective that a pretrained network using a large number of CT data sets can be fine-tuned with only a few radial MR data sets. This is potentially a good approach when training with a small data set since a large, labeled data set is not always available due to privacy, legal, and business-related concerns [93].

The data-driven strategies described have been successfully tested to remove image artifacts, such as edge blurring and intensity biases. However, the robustness of the trained neural network is still a concern.

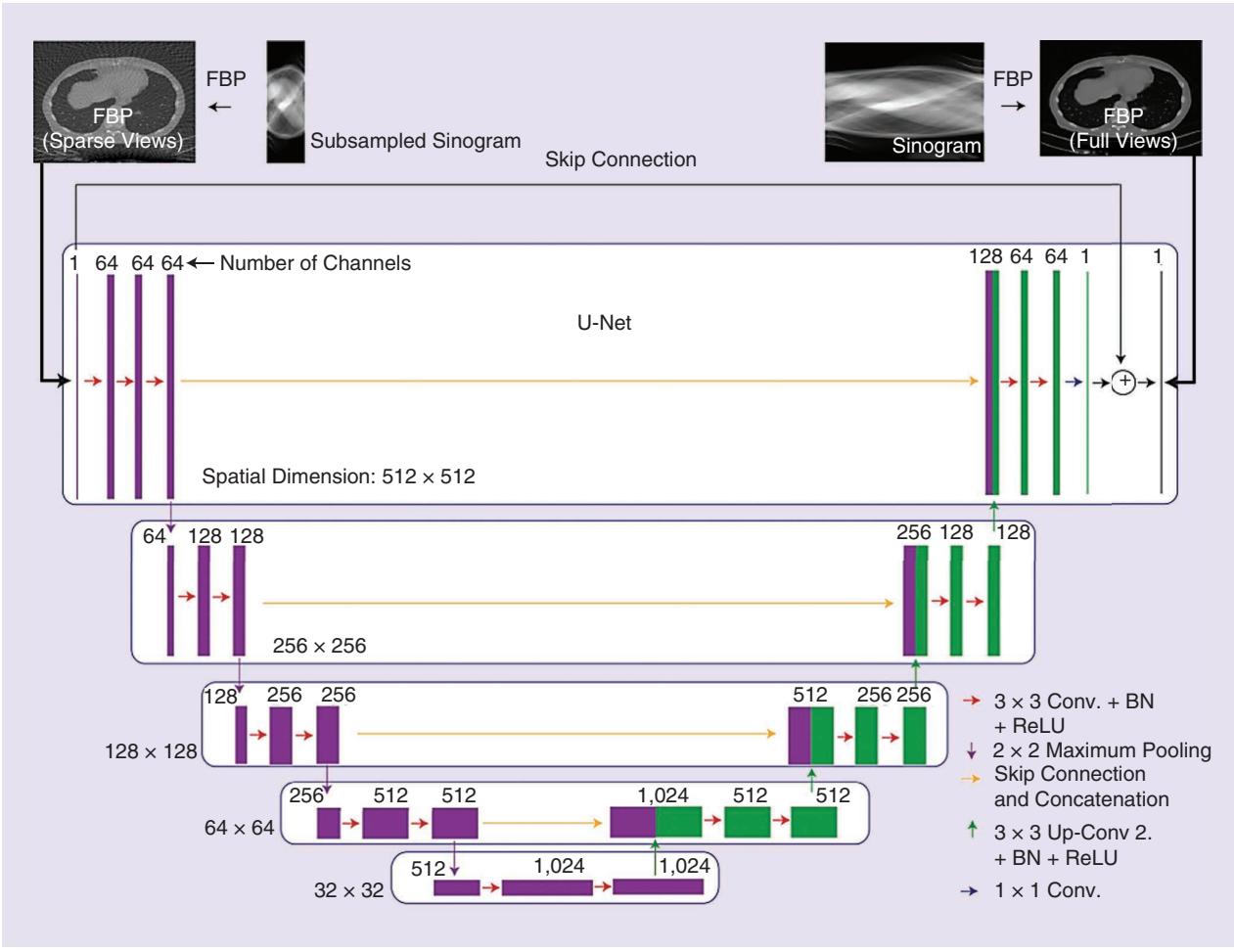

FIGURE 6. The architecture of the deep convolutional network in [10]. The input is a sparse-view CT image, and the output is a full-view CT image. FBP: filtered back projection; BN: batch normalization; Conv.:convolution; Up-Conv.: up-convolution. (Source: [10].)

In [94], Huang et al. investigated whether some noise or perturbations will mislead a neural network to failure in detecting an existing lesion, showing that the trained neural network, especially U-Net, is susceptible to Poisson noise. Hence, the estimation of the uncertainty (error) of the output of neural networks is crucial for reliable clinical applications.

In [95], Bayesian deep learning was applied for MRI reconstruction by also generating the corresponding uncertainty map. The experiments showed that Bayesian learning could alleviate network overfitting and perform competitively when the data are further away from the training distribution.

A reconsideration of data consistence after postprocessing may alleviate the susceptibility of the learning-based algorithm. Here, we would like to introduce two works of Wei et al. on EIT; see [96] and [97]. The task of the neural networks is to recover the complete equivalent current from the degraded equivalent current that is the major part of the complete current. After that, the current corrupted with noise can be accurately recovered. The current domain can be either "dominant current" or "induced current."

An important part of the method is to preprocess and postprocess the input and output currents with Green's function (i.e., mapping data to current or mapping current to permittivity). Since Green's function contains the physics of EM wave propagation, the inversion algorithm is robust to both noise and boundary inaccuracies. Similar works in microwave imaging can also be found in [98] and [99].

LEARNING PRIOR KNOWLEDGE IN CONVENTIONAL INVERSION ALGORITHMS

When our prior knowledge has large uncertainties, and we are not sure whether testing scenarios are involved in the training set, we may need physics to guide learning-based inversion. One way to achieve this goal is to incorporate the learned prior knowledge into conventional inversion algorithms, where the feedback from data consistence helps to guide the reconstruction.

The most common way of incorporating prior information into inversion is the regularization method. Extensive studies on dictionary learning have been done. The optimal sparse-signal transform obtained in the training process can be used as a regularizer for image reconstruction. Here, we list only a few reports from the literature due to space limits.

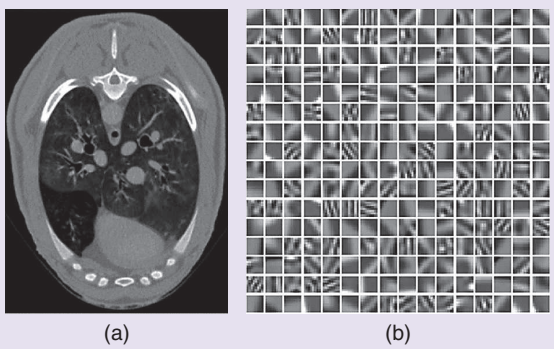

FIGURE 7. The construction of a global dictionary in [100]. (a) A sheep lung image is used to extract the training patches. (b) The learned dictionary consists of 256 atoms. (Source: [100].)

For instance, a two-step alternative updating algorithm was proposed for dictionary learning-based MRI in [8]. In the first step, the reconstructed image is fixed, and the algorithm learns the sparse representation from a set of image patches. In the second step, the learned dictionary is fixed as the regularization in the cost function, where an image update can be achieved by solving a least-squares problem.

In [100], the learned dictionary (see Figure 7) is incorporated into a statistical iterative reconstruction framework as regularization. The authors showed two ways that the dictionary can be either predetermined before an image reconstruction task or adaptively defined during the reconstruction. In [101], the constraints of dictionary learning and time sparsity are imposed for dynamic MR data reconstruction. For more details about dictionary learning as applied in biomedicine, please refer to [11] and the references therein.

The sparse-coding step in dictionary learning is NP-hard [33]. Zheng et al. applied squares-transform learning for low-dose CT reconstruction [102], [103]. The proposed iterative algorithm can run faster than a synthesis dictionary learning approach. Moreover, neural networks are also used to learn the regularization in the conventional framework of CT [19], [104]. In [19], nonlinear sparse representations were learned from normal-dose CT images with an autoencoder architecture. In [104], a CNN provides a reference model for iterative deterministic optimization so that the reconstructed image satisfies both data consistence and prior knowledge.

In addition, the quality of data can be improved by machine learning when the data to be inverted are roughly estimated or undersampled. In [105], Shah et al. used a CNN to recover the total electric field from the Born approximated field for 2D breast microwave imaging. In [106], the input of U-Net was the folded CT image recovered from undersampled data, while the output was a high-resolution CT image. The authors then used the output image to generate new data, and they supplemented the new data to the original undersampled data for the final reconstruction.

Furthermore, machine learning techniques can also be applied to learn the descent direction in iterative inversion algorithms to accelerate their convergence [22]. One example is the supervised descent method (SDM) [107]. In [23], [24], and [108], SDM was applied to both microwave imaging and EIT. In the training process, SDM learns the descent directions pointing from an initial model to the true models from the training data set. In the prediction, the learned descent directions and data residual between simulated and measured data work together to compute the model update. Since physical modeling is involved in online prediction, this method has shown a good generalization ability. In Figure 8, we present the absolute reconstruction [109] with SDM inversion using experimental human thoracic data [24].

Within the framework of conventional learning, Shen et al. applied deep reinforcement learning to automatically tuning regularization parameters in iterative CT reconstruction [21]. After

training, the parameter-tuning policy network can automatically compute the regularization factors. With the rapid development in reinforcement learning, we may benefit from it by discovering more interesting and powerful algorithms for solving imaging and inverse problems.

Generally speaking, it is worthwhile to notice that the latest publications show a trend of integrating physical operators into neural networks to improve the stability of data inversion since we know that physical laws can provide effective knowledge to reconstruct a model from measurements. In [110], Sun et al. developed the deep alternating direction method of multipliers (ADMM) network for MRI reconstruction, which is based on the ADMM algorithm. The iterative ADMM procedure is mapped on a data flow graph, where the Fourier transform and forward operator are predefined, but the filters and hyperparameters in ADMM are trainable.

In [25], [111], and [112], the authors developed iterative neural networks for CT or MRI reconstruction. The model update comes from two parts: one part is from gradient descent on the least square of data fitness, while the other part is from neural networks that would otherwise be represented by empirical regularization. An illustration of the architecture is shown in Figure 9.

In [113]–[115], the authors mapped an analytic reconstruction pipeline in CT to neural networks by specifying layers that implement the forward modeling and filtering operators. The network can automatically learn strategies such as compensation weights and apodization windows.

In [116]–[118], analytical data consistence layers are added intermediately in cascade neural networks for CT or MRI reconstruction. It is also promising in [116] that CNNs can learn spatiotemporal correlations when reconstructing the frames of sequences.

The work of Adler and Öktem [22], [119] showed that the steepest-descent and primal–dual methods for optimization can be unfolded to iterative neural networks that play the role of learning updating operators, with the (possibly nonlinear) forward operator and regularization as well as their derivatives explicitly written in the architecture. The architecture of a primal–dual network is shown in Figure 10, where the forward operator and its adjoint are denoted by $\mathcal{T}$ and $\mathcal{T}*$, respectively.

In [120], self-supervised learning was used in the physics-based neural network for MRI reconstruction, where the loss function is defined in the data space rather than the image space. This scheme avoids the necessity of using high-resolution images as training labels and, thus, may be a good way for training when lacking sufficient training data.

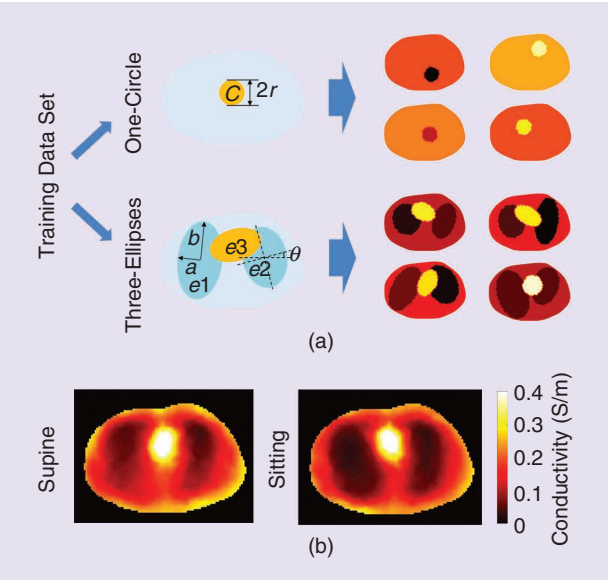

FIGURE 8. EIT with the SDM. (a) The design of the training data set. (b) A human thorax reconstruction with SDM inversion. C: circle; e: ellipse.

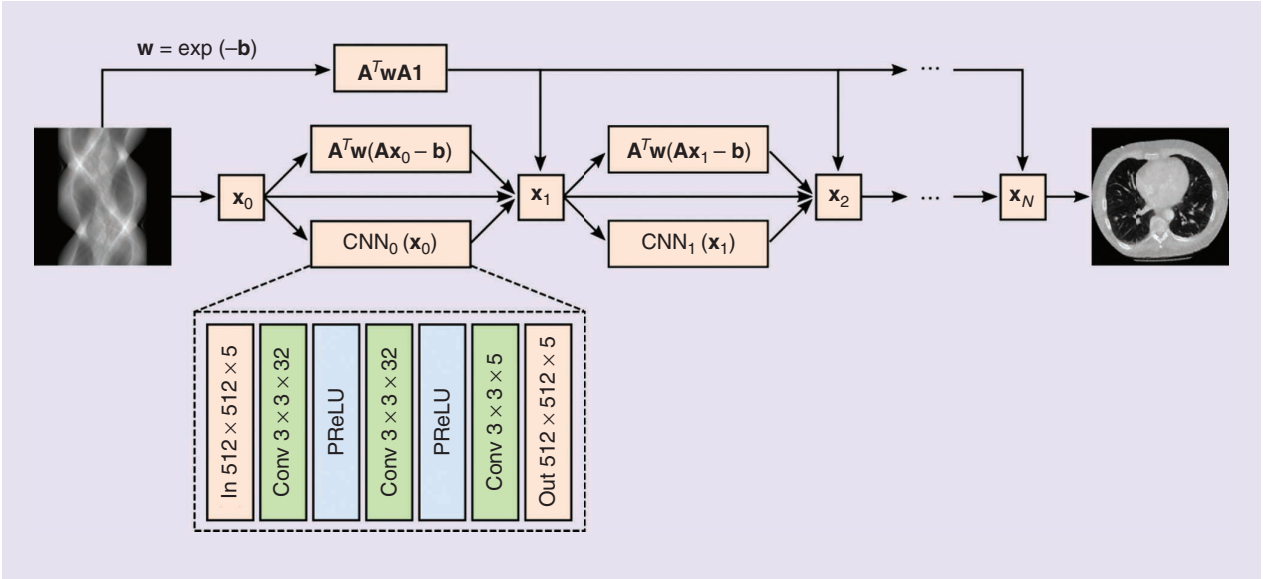

FIGURE 9. The architecture of the DNN in [25] where the data residual is incorporated. The input is reconstructed from the projection data by FBP. PRelu: parametric rectified linear unit.

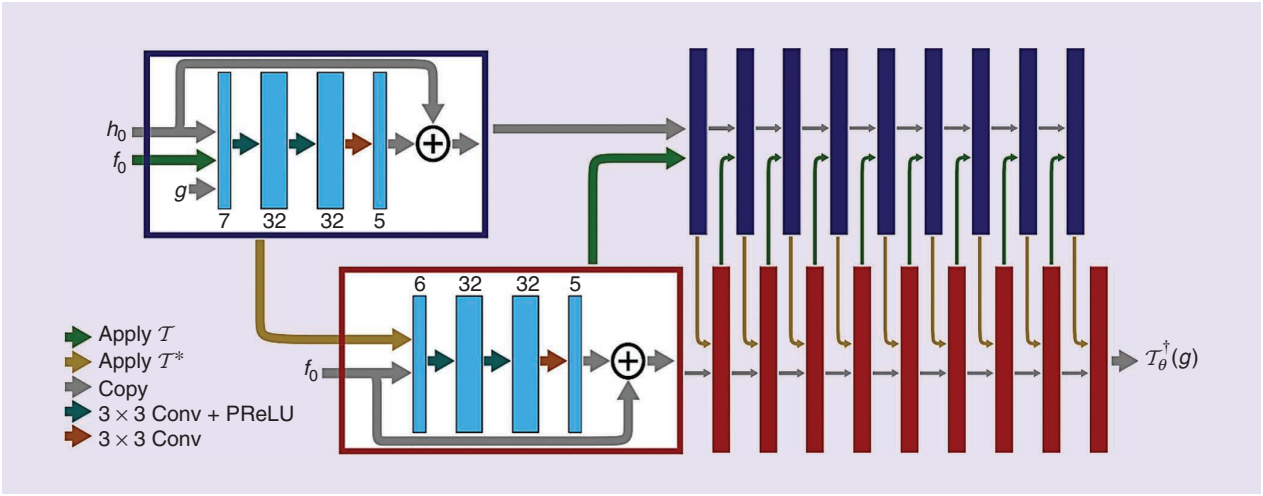

FIGURE 10. The architecture of the primal-dual network in [119] where the forward operator and its adjoint are incorporated. The forward operator and its adjoint are denoted by $\mathcal{T}$ and $\mathcal{T}^*$, respectively. (Source: [119].)

DISCUSSION AND CONCLUSIONS

The application of machine learning techniques to EM imaging for biomedical applications is a rapidly developing research area. New algorithms are developed on a daily basis. We are aware that some recent and new developments in this field of research are not listed in this review article. However, we are confident that this is a key and vast framework and that many more open problems are to be investigated.

Machine learning methods, especially deep learning, can extract useful information from large data sets and keep this information in the neural network or other types of representations. In many applications, the neural network can be used as a "black box." We specify the input and output parameters. The network parameters can be optimized during the training process. This workflow can also help with the design of imaging algorithms, in which the input represents the measured data, and the output is the reconstructed image. Since deep learning has an improved learning capability and generalization ability, it has been proven effective in some applications.

Inversion algorithms convert the measured data to images that represent the distribution of physical properties. This process can be considered as a combination of information from measurement, physical laws, and prior information. In this process, physical laws represent the relationship between physical parameters and measured data. This provides important information to constrain the reconstructed images. Compared with the information learned directly from data, physical laws are more abstract and have a better generalization ability than machine learning methods. Therefore, incorporating physics-based algorithms in the machine learning process can improve the accuracy of reconstructed images. Physics can also improve the learning capability and generalization ability of machine learning-based inversion algorithms.

Like other imaging algorithms, biomedical imaging algorithms extend human sensory functions. Images help doctors better understand a patient's situation. Recently, deep learning techniques have also been developed to analyze medical images for diagnosis purposes. In this workflow, the images are first constructed and then used as input for deep learning algorithms. Usually, the measured data contains richer information than the reconstructed images. Therefore, it would be better if diagnosis could be computed directly based on measured data. This would be difficult for humans because the data are abstract and complex, but it could be feasible for computers. Combining machine learning-based imaging and diagnosis may open some interesting problems and bring innovative medical solutions.

As sensor performance improves and storage costs decrease, the acquired data volume will keep expanding. Machine learning algorithms provide us with a feasible solution to use information from data. They can also help biomedical imaging, in which traditional algorithms are usually based on the physics of measurement. A proper integration of these two methods in the inversion algorithms will lead to more accurate images and more efficient computation.

ACKNOWLEDGMENTS

This work was supported in part by the National Key R&D Program of China (2018YFC0603604); National Science Foundation of China (61571264 and 61971263); Guangzhou Science and Technology Plan (201804010266); Beijing Innovation Center for Future Chip; and Research Institute of Tsinghua, Pearl River Delta.

AUTHOR INFORMATION

Maokun Li (maokunli@tsinghua.edu.cn) is an associate professor at the Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing, 100086, China. His research interests include fast algorithms in computational electromagnetics and their applications in antenna modeling, inverse problems, and so on. He is a Senior Member of IEEE.

Rui Guo (guor17@mails.tsinghua.edu.cn) is a Ph.D. student at the Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing, 100086, China. His research interests include the theories and applications of machine learning in inverse scattering problems. He is a Graduate Student Member of IEEE.

Ke Zhang (kzhang320@mail.tsinghua.edu.cn) is a postdoctoral researcher at the Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing, 100086, China. His research interests include electrical impedance tomography, biomedical imaging, and noninvasive physiological measurement.

Zhichao Lin (lzc19@mails.tsinghua.edu.cn) is a Ph.D. student at the Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing, 100086, China. His research interests include the theory and applications of deep learning in electrical impedance tomography.

Fan Yang (fan_yang@tsinghua.edu.cn) is a professor at the Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing, 100086, China. His research interests include antennas, surface electromagnetics, computational electromagnetics, and applied electromagnetic systems. He is a Fellow of IEEE.

Shenheng Xu (shxu@tsinghua.edu.cn) is an associate professor at the Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing, 100086, China. His research interests include novel designs of high-gain antennas, artificial electromagnetic structures, and electromagnetic and antenna theories. He is a Member of IEEE.

Xudong Chen (elechenx@nus.edu.sg) is a professor in the Department of Electrical and Computer Engineering, National University of Singapore, 117583, Singapore. His research interests focus on electromagnetic inverse problems and computational imaging. He is a Fellow of IEEE and fellow of the Electromagnetics Academy.

Andrea Massa (andrea.massa@unitn.it) is a professor at the University of Trento, Trento, 38123, Italy; holder of a Chang-Jiang Chair Professorship at University of Electronic Science and Technology of China, Chengdu, 610097, China, and visiting professor at Tsinghua University, Beijing, 100086, China. His research interests include the theory and applications of electromagnetics, inverse scattering, antenna arrays, and optimization techniques. He is a Fellow of IEEE.

Aria Abubakar (aabubakar@slb.com) is the head of data science for Digital Subsurface Solutions, Schlumberger, Houston, Texas, 77056, USA. He has been working in numerical modeling and optimization for geoscience and biomedical applications as well as multiphysics data integration. He is a Senior Member of IEEE.

REFERENCES

[1] R. Acharya, R. Wasserman, J. Stevens, and C. Hinojosa, "Biomedical imaging modalities: A tutorial," *Comput. Med. Imag. Graph.*, vol. 19, no. 1, pp. 3–25, 1995. doi: 10.1016/0895-6111(94)00043-3.

[2] H. Ammari, *An Introduction to Mathematics of Emerging Biomedical Imaging*, vol. 62. Hoboken, NJ: Springer-Verlag, 2008.

[3] Z. Q. Zhang and Q. H. Liu, "Three-dimensional nonlinear image reconstruction for microwave biomedical imaging," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 3, pp. 544–548, 2004. doi: 10.1109/TBME.2003.821052.

[4] P. Beard, "Biomedical photoacoustic imaging," *Interface Focus*, vol. 1, no. 4, pp. 602–631, 2011. doi: 10.1098/rsfs.2011.0028.

[5] M. Bertero and M. Piana, "Inverse problems in biomedical imaging: Modeling and methods of solution," in *Complex Systems in Biomedicine*. Hoboken, NJ: Springer-Verlag, 2006, pp. 1–33.

[6] A. Abubaker and P. M. Van Den Berg, "Total variation as a multiplicative constraint for solving inverse problems," *IEEE Trans. Image Process.*, vol. 10, no. 9, pp. 1384–1392, 2001. doi: 10.1109/83.941862.

[7] R. Scapaticci, P. Kosmas, and L. Crocco, "Wavelet-based regularization for robust microwave imaging in medical applications," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, pp. 1195–1202, 2014. doi: 10.1109/TBME.2014.2381270.

[8] S. Ravishankar and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1028–1041, 2010. doi: 10.1109/TMI.2010.2090538.

[9] T. M. Habashy and A. Abubakar, "A general framework for constraint minimization for the inversion of electromagnetic measurements," *Progr. Electromagn. Res.*, vol. 46, pp. 265–312, 2004. doi: 10.2528/PIER03100702.

[10] 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, 2017. doi: 10.1109/TIP.2017.2713099.

[11] T. Tong, J. Caballero, K. Bhatia, and D. Rueckert, "Dictionary learning for medical image denoising, reconstruction, and segmentation," in *Machine Learning and Medical Imaging*. Amsterdam, The Netherlands: Elsevier, 2016, pp. 153–181.

[12] W. C. Chew, *Waves and Fields in Inhomogeneous Media*. New York: Wiley, 1995.

[13] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. doi: 10.1038/nature14539.

[14] B. Jokanovic, M. Amin, and F. Ahmad, "Radar fall motion detection using deep learning," in *Proc. 2016 IEEE Radar Conf. (RadarConf)*, pp. 1–6. doi: 10.1109/RADAR.2016.7485147.

[15] Y. Gandole, "Noise reduction of biomedical signal using artificial neural network model," *Int. J. Eng. Technol.*, vol. 2, no. 1, p. 8, 2012.

[16] K. Antczak, "Deep recurrent neural networks for ECG signal denoising," 2018, arXiv:1807.11551.

[17] X. Li et al., "A novel deep neural network method for electrical impedance tomography," *Trans. Inst. Meas. Control.*, vol. 41, no. 14, pp. 4035–4049, 2019. doi: 10.1177/0142331219845037.

[18] S. Mandija, E. F. Meliadò, N. R. Huttinga, P. R. Luijten, and C. A. van den Berg, "Opening a new window on MR-based Electrical Properties Tomography with deep learning," *Sci. Rep.*, vol. 9, no. 1, p. 8895, 2019. doi: 10.1038/s41598-019-45382-x.

[19] D. Wu, K. Kim, G. El Fakhri, and Q. Li, "Iterative low-dose CT reconstruction with priors trained by artificial neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2479–2486, 2017. doi: 10.1109/TMI.2017.2753138.

[20] H. M. Yao, E. Wei, and L. Jiang, "Two-step enhanced deep learning approach for electromagnetic inverse scattering problems," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 11, pp. 2254–2258, 2019. doi: 10.1109/LAWP.2019.2925578.

[21] C. Shen, Y. Gonzalez, L. Chen, S. B. Jiang, and X. Jia, "Intelligent parameter tuning in optimization-based iterative CT reconstruction via deep reinforcement learning," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1430–1439, 2018. doi: 10.1109/TMI.2018.2823679.

[22] J. Adler and O. Öktem, "Solving ill-posed inverse problems using iterative deep neural networks," *Inverse Probl.*, vol. 33, no. 12, p. 124,007, 2017. doi: 10.1088/1361-6420/aa9581.

[23] R. Guo, X. Song, M. Li, F. Yang, S. Xu, and A. Abubakar, "Supervised descent learning technique for 2-D microwave imaging," *IEEE Trans. Antennas Propag.*, vol. 67, no. 5, pp. 3550–3554, 2019. doi: 10.1109/TAP.2019.2902667.

[24] K. Zhang, R. Guo, M. Li, F. Yang, S. Xu, and A. Abubakar, "Supervised descent learning for thoracic electrical impedance tomography," *IEEE Trans. Biomed. Eng.*, early access. doi: 10.1109/TBME.2020.3027827.

[25] D. Wu, K. Kim, B. Dong, and Q. Li, "End-to-end abnormality detection in medical imaging," 2017, arXiv:1711.02074.

[26] J. R. Koza, F. H. Bennett, D. Andre, and M. A. Keane, "Automated design of both the topology and sizing of analog electrical circuits using genetic programming," in *Artificial Intelligence in Design'96*. Hoboken, NJ: Springer-Verlag, 1996, pp. 151–170.

[27] C. M. Bishop, *Pattern Recognition and Machine Learning*. Hoboken, NJ: Springer-Verlag, 2006.

[28] Z. Ghahramani, "Unsupervised learning," in *Summer School on Machine Learning*. Hoboken, NJ: Springer-Verlag, 2003, pp. 72–112.
[29] M. Längkvist, L. Karlsson, and A. Loutfi, "A review of unsupervised feature learning and deep learning for time-series modeling," *Pattern Recognit. Lett.*, vol. 42, pp. 11–24, June 2014. doi: 10.1016/j.patrec.2014.01.008.
[30] C.-J. Hsieh, K.-W. Chang, C.-J. Lin, S. S. Keerthi, and S. Sundararajan, "A dual coordinate descent method for large-scale linear SVM," in *Proc. 25th Int. Conf. Machine Learn.*, 2008, pp. 408–415. doi: 10.1145/1390156.1390208.
[31] K. Kreutz-Delgado, J. F. Murray, B. D. Rao, K. Engan, T.-W. Lee, and T. J. Sejnowski, "Dictionary learning algorithms for sparse representation," *Neural Comput.*, vol. 15, no. 2, pp. 349–396, 2003. doi: 10.1162/089976603762552951.
[32] K. Engan, S. O. Aase, and J. H. Husoy, "Frame based signal compression using method of optimal directions (MOD)," in *Proc. 1999 IEEE Int. Symp. Circuits Syst.*, 1999, vol. 4, pp. 1–4. doi: 10.1109/ISCAS.1999.779928.
[33] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, 2006. doi: 10.1109/TSP.2006.881199.
[34] K. Hornik, M. Stinchcombe, and H. White, "Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks," *Neural Netw.*, vol. 3, no. 5, pp. 551–560, 1990. doi: 10.1016/0893-6080(90)90005-6.
[35] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," in *The Handbook Brain Theory Neural Networks*. Cambridge, MA: MIT Press, 1995, pp. 255–258.
[36] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," 2015, arXiv:1506.00019.
[37] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, arXiv:1409.1556.
[38] 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. doi: 10.1109/CVPR.2016.90.
[39] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.
[40] I. Goodfellow et al., "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
[41] S. P. Kumar, N. Sriraam, P. Benakop, and B. Jinaga, "Entropies based detection of epileptic seizures with artificial neural network classifiers," *Expert Syst. Appl.*, vol. 37, no. 4, pp. 3284–3291, 2010. doi: 10.1016/j.eswa.2009.09.051.
[42] E. D. Übeyli, "Combined neural network model employing wavelet coefficients for EEG signals classification," *Digit. Signal Process.*, vol. 19, no. 2, pp. 297–308, 2009. doi: 10.1016/j.dsp.2008.07.004.
[43] P. Fergus, D. Hignett, A. Hussain, D. Al-Jumeily, and K. Abdel-Aziz, "Automatic epileptic seizure detection using scalp EEG and advanced artificial intelligence techniques," *BioMed Res. Int.*, vol. 2015, p. 986,736, 2015. doi: 10.1155/2015/986736.
[44] C. Venkatesan, P. Karthigaikumar, A. Paul, S. Satheeskumaran, and R. Kumar, "ECG signal preprocessing and SVM classifier-based abnormality detection in remote healthcare applications," *IEEE Access*, vol. 6, pp. 9767–9773, Jan. 2018. doi: 10.1109/ACCESS.2018.2794346.
[45] E. Gokgoz and A. Subasi, "Comparison of decision tree algorithms for EMG signal classification using DWT," *Biomed. Signal Process. Control*, vol. 18, pp. 138–144, Apr. 2015. doi: 10.1016/j.bspc.2014.12.005.
[46] D. Byrne, M. O'Halloran, E. Jones, and M. Glavin, "Support vector machine-based ultrawideband breast cancer detection system," *J. Electromagn. Waves Appl.*, vol. 25, no. 13, pp. 1807–1816, 2011. doi: 10.1163/156939311797454015.
[47] J. Sacristán, B. L. Oliveira, and S. Pistorius, "Classification of electromagnetic signals obtained from microwave scattering over healthy and tumorous breast models," in *Proc. 2016 IEEE Can. Conf. Electr. Comput. Eng. (CCECE)*, pp. 1–5. doi: 10.1109/CCECE.2016.7726761.
[48] U. R. Acharya, H. Fujita, O. S. Lih, Y. Hagiwara, J. H. Tan, and M. Adam, "Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network," *Inf. Sci.*, vol. 405, pp. 81–90, Sept. 2017. doi: 10.1016/j.ins.2017.04.012.
[49] A. Alkan and M. Günay, "Identification of EMG signals using discriminant analysis and SVM classifier," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 44–47, 2012. doi: 10.1016/j.eswa.2011.06.043.
[50] S.-H. Park and S.-P. Lee, "EMG pattern recognition based on artificial intelligence techniques," *IEEE Trans. Rehabil. Eng.*, vol. 6, no. 4, pp. 400–405, 1998. doi: 10.1109/86.736154.
[51] Y. Matsumura, Y. Mitsukura, M. Fukumi, N. Akamatsu, Y. Yamamoto, and K. Nakaura, "Recognition of EMG signal patterns by neural networks," in *Proc. 9th Int. Conf. Neural Inf. Process. (ICONIP'02)*, 2002, vol. 2, pp. 750–754. doi: 10.1109/ICONIP.2002.1198158.
[52] M. Atzori, M. Cognolato, and H. Muller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," *Front. Neurorobot.*, vol. 10, p. 9, Sept. 2016. doi: 10.3389/fnbot.2016.00009.
[53] N.-Y. Liang, P. Saratchandran, G.-B. Huang, and N. Sundararajan, "Classification of mental tasks from EEG signals using extreme learning machine," *Int. J. Neural Syst.*, vol. 16, no. 1, pp. 29–38, 2006. doi: 10.1142/S0129065706000482.
[54] A. Supratak, H. Dong, C. Wu, and Y. Guo, "Deepsleepnet: A model for automatic sleep stage scoring based on raw single-channel EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 11, pp. 1998–2008, 2017. doi: 10.1109/TNSRE.2017.2721116.
[55] K. Assaleh, "Extraction of fetal electrocardiogram using adaptive neuro-fuzzy inference systems," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 1, pp. 59–68, 2006. doi: 10.1109/TBME.2006.883728.
[56] M. A. Hasan, M. I. Ibrahimy, and M. B. I. Reaz, "Fetal ECG extraction from maternal abdominal ECG using neural network," *J. Softw. Eng. Appl.*, vol. 2, no. 5, pp. 330–334, 2009. doi: 10.4236/jsea.2009.25043.
[57] A. M. Kaleem and R. D. Kokate, "An efficient adaptive filter for fetal ECG extraction using neural network," *J. Intell. Syst.*, vol. 28, no. 4, pp. 589–600, 2019. doi: 10.1515/jisys-2017-0031.
[58] Y. Ma, Y. Xiao, G. Wei, and J. Sun, "Fetal ECG extraction using adaptive functional link artificial neural network," in *Proc. 2014 Asia Pacific Signal Inf. Process. Assoc. Annu. Summit Conf.*, pp. 1–4. doi: 10.1109/APSIPA.2014.7041680.
[59] A. Erfanian and B. Mahmoudi, "Real-time ocular artifact suppression using recurrent neural network for electro-encephalogram based brain-computer interface," *Med. Biol. Eng. Comput.*, vol. 43, no. 2, pp. 296–305, 2005. doi: 10.1007/BF02345969.
[60] A. Jafarifarmand and M. A. Badamchizadeh, "Artifacts removal in EEG signal using a new neural network enhanced adaptive filter," *Neurocomputing*, vol. 103, pp. 222–231, Mar. 2013. doi: 10.1016/j.neucom.2012.09.024.
[61] J. Hu, C. sheng Wang, M. Wu, Y. xiao Du, Y. He, and J. She, "Removal of EOG and EMG artifacts from EEG using combination of functional link neural network and adaptive neural fuzzy inference system," *Neurocomputing*, vol. 151, no. 151, pp. 278–287, 2015. doi: 10.1016/j.neucom.2014.09.040.
[62] J. Mateo, A. M. Torres, M. A. García, and J. L. Santos, "Noise removal in electroencephalogram signals using an artificial neural network based on the simultaneous perturbation method," *Neural Comput. Appl.*, vol. 27, no. 7, pp. 1941–1957, 2016. doi: 10.1007/s00521-015-1988-7.
[63] S. Poungponsri and X. Yu, "An adaptive filtering approach for electrocardiogram (ECG) signal noise reduction using neural networks," *Neurocomputing*, vol. 117, pp. 206–213, Oct. 2013. doi: 10.1016/j.neucom.2013.02.010.
[64] H.-T. Chiang, Y.-Y. Hsieh, S.-W. Fu, K.-H. Hung, Y. Tsao, and S.-Y. Chien, "Noise reduction in ECG signals using fully convolutional denoising autoencoders," *IEEE Access*, vol. 7, pp. 60,806–60,813, Apr. 2019. doi: 10.1109/ACCESS.2019.2912036.
[65] P. Xiong, H. Wang, M. Liu, F. Lin, Z. Hou, and X. Liu, "A stacked contractive denoising auto-encoder for ECG signal denoising," *Physiol. Meas.*, vol. 37, no. 12, pp. 2214–2230, 2016. doi: 10.1088/0967-3334/37/12/2214.
[66] E. N. Bruce, *Biomedical Signal Processing and Signal Modeling*. New York: Wiley, 2001.
[67] X. Chen, *Computational Methods for Electromagnetic Inverse Scattering*. Hoboken, NJ: Wiley, 2018.
[68] S. C. Constable, R. L. Parker, and C. G. Constable, "Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data," *Geophysics*, vol. 52, no. 3, pp. 289–300, 1987. doi: 10.1190/1.1442303.
[69] I. T. Rekanos, "Neural-network-based inverse-scattering technique for online microwave medical imaging," *IEEE Trans. Magn.*, vol. 38, no. 2, pp. 1061–1064, 2002. doi: 10.1109/20.996272.
[70] P. Wang, H. l. Li, L.-l. Xie, and Y.-c. Sun, "The implementation of FEM and RBF neural network in EIT," in *Proc. 2009 2nd Int. Conf. Intell. Netw. Intell. Syst.*, pp. 66–69. doi: 10.1109/ICINIS.2009.26.
[71] M. Michalikova, R. Abed, M. Prauzek, and J. Koziorek, "Image reconstruction in electrical impedance tomography using neural network," in *Proc. 2014 Cairo Int. Biomed. Eng. Conf. (CIBEC)*, pp. 39–42. doi: 10.1109/CIBEC.2014.7020959.
[72] A. Kerhet, M. Raffetto, A. Boni, and A. Massa, "A SVM-based approach to microwave breast cancer detection," *Eng. Appl. Artific. Intell.*, vol. 19, no. 7, pp. 807–818, 2006. doi: 10.1016/j.engappai.2006.05.010.
[73] M. Salucci, A. Gelmini, J. Vrba, I. Merunka, G. Oliveri, and P. Rocca, "Instantaneous brain stroke classification and localization from real scattering data," *Microw. Opt. Technol. Lett.*, vol. 61, no. 3, pp. 805–808, 2019. doi: 10.1002/mop.31639.
[74] C. Tan, S. Lv, F. Dong, and M. Takei, "Image reconstruction based on convolutional neural network for electrical resistance tomography," *IEEE Sensors J.*, vol. 19, no. 1, pp. 196–204, 2018. doi: 10.1109/JSEN.2018.2876411.

[75] Y. Gao et al., "EIT-CDAE: A 2-D electrical impedance tomography image reconstruction method based on auto encoder technique," in *Proc. 2019 IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, pp. 1–4. doi: 10.1109/BIOCAS.2019.8918979.
[76] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, p. 487, 2018. doi: 10.1038/nature25988.
[77] N. Hampe, U. Katscher, C. A. van den Berg, K. K. Tha, and S. Mandija, "Deep learning brain conductivity mapping using a patch-based 3D U-net," 2019, arXiv:1908.04118.
[78] M. T. McCann, K. H. Jin, and M. Unser, "Convolutional neural networks for inverse problems in imaging: A review," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 85–95, 2017. doi: 10.1109/MSP.2017.2739299.
[79] S. Martin and C. T. Choi, "A post-processing method for three-dimensional electrical impedance tomography," *Sci. Rep.*, vol. 7, no. 1, p. 7212, 2017. doi: 10.1038/s41598-017-07727-2.
[80] S. J. Hamilton and A. Hauptmann, "Deep D-Bar: Real-time electrical impedance tomography imaging with deep neural networks," *IEEE Trans. Med. Imag.*, vol. 37, no. 10, pp. 2367–2377, 2018. doi: 10.1109/TMI.2018.2828303.
[81] A. I. Nachman, "Global uniqueness for a two-dimensional inverse boundary value problem," *Ann. Math.*, vol. 143, no. 1, pp. 71–96, 1996. doi: 10.2307/2118653.
[82] P. Shah and M. Moghaddam, "Super resolution for microwave imaging: A deep learning approach," in *Proc. 2017 IEEE Int. Symp. Antennas Propag. USNC/URSI Nat. Radio Sci. Meeting*, pp. 849–850. doi: 10.1109/APUSNCURSINRSM.2017.8072467.
[83] L. Wang and J. Xu, "Deep convolutional neural networks for breast image analysis on holographic microwave imaging," in *Proc. ASME Int. Mech. Eng. Congr. Exposition*, 2018. doi: 10.1115/IMECE2018-86765.
[84] V. Khoshdel, A. Ashraf, and J. LoVetri, "Enhancement of multimodal microwave-ultrasound breast imaging using a deep-learning technique," *Sensors*, vol. 19, no. 18, p. 4050, 2019. doi: 10.3390/s19184050.
[85] H. Chen et al., "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524–2535, 2017. doi: 10.1109/TMI.2017.2715284.
[86] H. Chen et al., "Low-dose CT via convolutional neural network," *Biomed. Opt. Express*, vol. 8, no. 2, pp. 679–694, 2017. doi: 10.1364/BOE.8.000679.
[87] Z. Zhang, X. Liang, X. Dong, Y. Xie, and G. Cao, "A sparse-view CT reconstruction method based on combination of DenseNet and deconvolution," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1407–1417, 2018. doi: 10.1109/TMI.2018.2823338.
[88] H. Shan et al., "3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1522–1534, 2018. doi: 10.1109/TMI.2018.2832217.
[89] D. Lee, J. Yoo, and J. C. Ye, "Deep residual learning for compressed sensing MRI," in *Proc. 2017 IEEE 14th Int. Symp. Biomed. Imag. (ISBI 2017)*, pp. 15–18. doi: 10.1109/ISBI.2017.7950457.
[90] G. Yang et al., "DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1310–1321, 2017. doi: 10.1109/TMI.2017.2785879.
[91] T. M. Quan, T. Nguyen-Duc, and W.-K. Jeong, "Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1488–1497, 2018. doi: 10.1109/TMI.2018.2820120.
[92] Y. Han, J. Yoo, H. H. Kim, H. J. Shin, K. Sung, and J. C. Ye, "Deep learning with domain adaptation for accelerated projection-reconstruction MR," *Magn. Reson. Med.*, vol. 80, no. 3, pp. 1189–1205, 2018. doi: 10.1002/mrm.27106.
[93] G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler, "Image reconstruction is a new frontier of machine learning," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1289–1296, 2018. doi: 10.1109/TMI.2018.2833635.
[94] Y. Huang, T. Würfl, K. Breininger, L. Liu, G. Lauritsch, and A. Maier, "Some investigations on robustness of deep learning in limited angle tomography," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Hoboken, NJ: Springer-Verlag, 2018, pp. 145–153.
[95] J. Schlemper et al., "Bayesian deep learning for accelerated MR image reconstruction," in *International Workshop on Machine Learning for Medical Image Reconstruction*, pp. 64–71. Hoboken, NJ: Springer-Verlag, 2018.
[96] Z. Wei and X. Chen, "Induced-current learning method for nonlinear reconstructions in electrical impedance tomography," *IEEE Trans. Med. Imag.*, vol. 39, no. 5, pp. 1326–1334, 2019. doi: 10.1109/TMI.2019.2948909.
[97] Z. Wei, D. Liu, and X. Chen, "Dominant-current deep learning scheme for electrical impedance tomography," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 9, pp. 2546–2555, 2019. doi: 10.1109/TBME.2019.2891676.
[98] 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, 2019. doi: 10.1109/TAP.2019.2922779.
[99] Y. Sanghvi, Y. Kalepu, and U. K. Khankhoje, "Embedding deep learning in inverse scattering problems," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 46–56, May 2019. doi: 10.1109/TCI.2019.2915580.
[100] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang, "Low-dose X-ray CT reconstruction via dictionary learning," *IEEE Trans. Med. Imag.*, vol. 31, no. 9, pp. 1682–1697, 2012.
[101] J. Caballero, A. N. Price, D. Rueckert, and J. V. Hajnal, "Dictionary learning and time sparsity for dynamic MR data reconstruction," *IEEE Trans. Med. Imag.*, vol. 33, no. 4, pp. 979–994, 2014. doi: 10.1109/TMI.2014.2301271.
[102] X. Zheng, Z. Lu, S. Ravishankar, Y. Long, and J. A. Fessler, "Low dose CT image reconstruction with learned sparsifying transform," in *Proc. 2016 IEEE 12th Image, Video, Multidimensional Signal Process. Workshop (IVMSP)*, pp. 1–5. doi: 10.1109/IVMSPW.2016.7528219.
[103] X. Zheng, S. Ravishankar, Y. Long, and J. A. Fessler, "PWLS-ULTRA: An efficient clustering and learning-based approach for low-dose 3D CT image reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1498–1510, 2018. doi: 10.1109/TMI.2018.2832007.
[104] B. Chen, K. Xiang, Z. Gong, J. Wang, and S. Tan, "Statistical iterative CBCT reconstruction based on neural network," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1511–1521, 2018. doi: 10.1109/TMI.2018.2829896.
[105] P. Shah, G. Chen, and M. Moghaddam, "Learning nonlinearity of microwave imaging through deep learning," in *Proc. 2018 IEEE Int. Symp. Antennas Propag. USNC/URSI Nat. Radio Sci. Meeting*, pp. 699–700. doi: 10.1109/APUSNCURSINRSM.2018.8609005.
[106] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, "Deep learning for undersampled MRI reconstruction," *Phys. Med. Biol.*, vol. 63, no. 13, p. 135,007, 2018. doi: 10.1088/1361-6560/aac71a.
[107] X. Xiong and F. De la Torre, "Supervised descent method and its applications to face alignment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 532–539. doi: 10.1109/CVPR.2013.75.
[108] M. Li, K. Zhang, R. Guo, F. Yang, S. Xu, and A. Abubakar, "Supervised descent method for electrical impedance tomography," in *Proc. 2019 Photon. Electromagn. Res. Symp.—Fall (PIERS-Fall)*, pp. 2342–2348. doi: 10.1109/PIERS-Fall48861.2019.9021506.
[109] G. Hahn et al., "Imaging pathologic pulmonary air and fluid accumulation by functional and absolute EIT," *Physiol. Meas.*, vol. 27, no. 5, p. S187, 2006. doi: 10.1088/0967-3334/27/5/S16.
[110] Y. Yang, J. Sun, H. Li, and Z. Xu, "Deep ADMM-Net for compressive sensing MRI," in *Proc. 30th Int. Conf. Neural Inf. Process. Syst.*, 2016, pp. 10–18. doi: 10.5555/3157096.3157098.
[111] K. Hammernik et al., "Learning a variational network for reconstruction of accelerated MRI data," *Magn. Reson. Med.*, vol. 79, no. 6, pp. 3055–3071, 2018. doi: 10.1002/mrm.26977.
[112] H. Chen et al., "LEARN: Learned experts' assessment-based reconstruction network for sparse-data CT," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1333–1347, 2018. doi: 10.1109/TMI.2018.2805692.
[113] T. Würfl, F. C. Ghesu, V. Christlein, and A. Maier, "Deep learning computed tomography," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2016, pp. 432–440. doi: 10.1007/978-3-319-46726-9_50.
[114] C. Syben, B. Stimpel, J. Lommen, T. Würfl, A. Dörfler, and A. Maier, "Deriving neural network architectures using precision learning: Parallel-to-fan beam conversion," in *Proc. German Conf. Pattern Recognit.*, 2018, pp. 503–517. doi: 10.1007/978-3-030-12939-2_35.
[115] T. Würfl et al., "Deep learning computed tomography: Learning projection-domain weights from image domain in limited angle problems," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1454–1463, 2018. doi: 10.1109/TMI.2018.2833499.
[116] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for dynamic MR image reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 2, pp. 491–503, 2017. doi: 10.1109/TMI.2017.2760978.
[117] A. Kofler, M. Haltmeier, C. Kolbitsch, M. Kachelrieß, and M. Dewey, "A U-nets cascade for sparse view computed tomography," in *Proc. Int. Workshop Machine Learn. Med. Image Reconstruct.*, 2018, pp. 91–99. doi: 10.1007/978-3-030-00129-2_11.
[118] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL: Model-based deep learning architecture for inverse problems," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 394–405, 2018. doi: 10.1109/TMI.2018.2865356.
[119] J. Adler and O. Öktem, "Learned primal-dual reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1322–1332, 2018. doi: 10.1109/TMI.2018.2799231.
[120] B. Yaman, S. A. H. Hosseini, S. Moeller, J. Ellermann, K. Uğurbil, and M. Akçakaya, "Self-supervised physics-based deep learning MRI reconstruction without fully-sampled data," in *Proc. 2020 IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, pp. 921–925. doi: 10.1109/ISBI45749.2020.9098514.