Deep Tomographic Image Reconstruction: Yesterday, Today, and Tomorrow—Editorial for the 2nd Special Issue "Machine Learning for Image Reconstruction"

***Abstract*— As a follow-up to the first IEEE TRANSACTIONS ON MEDICAL IMAGING (TMI) special issue on the theme of deep tomographic reconstruction, the second special issue is assembled to reflect the status and momentum of this rapidly emerging field. In this editorial, we provide a brief background illustrating the motivation for the development of network-based, data-driven, and learning-oriented reconstruction methods, summarize the included papers, and report our verification of the shared deep learning codes. Finally, we discuss several important research topics to facilitate further investigation and collaboration.**

***Index Terms*— Tomography, image reconstruction, machine learning, artificial intelligence, deep learning, deep reconstruction.**

I. INTRODUCTION

TOMOGRAPHIC image reconstruction is an important aspect of medical imaging and plays a key role in modern medicine. Classic image reconstruction algorithms assume idealized imaging models and datasets. This allows the use of Fourier inversion and other analytic solutions. However, when clinical needs and technical factors are taken into consideration, both imaging scanners and raw data are far from being ideal and often become difficult to model precisely. For example, CT sinograms are noisy when using low-dose scans, incomplete due to a limited scanning range or presence of metal implants, and inconsistent due to patient motion, including motion of the heart. In these cases, iterative solutions are helpful because they can be designed to handle imperfect datasets and improve image quality. Such iterative methods utilize prior knowledge of the imaging model and the image content. For example, prior knowledge about the physics of data acquisition (e.g., the Poisson nature of X-ray, gamma-ray, and optical photon emission, and Rician distribution of MRI image noise) can be captured by a maximum likelihood formulation. Prior knowledge about the image content may be accounted for by sparsity, low-rank, and dimension-reduction priors. These priors can be combined, and form the core of iterative image reconstruction algorithms.

Despite remarkable successes of iterative algorithms, they suffer from several limitations in practice. For example, compressed sensing methods [1] allow sparse view reconstructions when certain conditions are satisfied, which are often associated with the number of projections on an order of hundreds. However, these sparsity-oriented algorithms produce strong artifacts when the number of projections is below that threshold. As another example, the amount of k-space data must be made rather small for fast MRI scans, and in this case, even if a contemporary compressed sensing method is used, image artifacts cannot be fully suppressed. Furthermore, applications of these methods necessitate tuning of hyper-parameters and are computationally demanding.

The main reasons for degradation in the image quality of the above iterative algorithms include (a) the limited image redundancies that classical image priors (e.g., sparsity or low-rank) can represent, and (b) inaccuracies in the prior knowledge about the imaging physics. Indeed, extensive prior knowledge may be available, but not easy to be extracted and utilized for tomographic reconstruction. For example, the forward model of an imaging system is typically approximate. Tomographic imaging models are treated as linear systems for tractable iterative algorithms. Nonlinearities and uncertain factors may be too complicated to be modeled in a closed form. In addition, the formulated prior knowledge on the image content is usually over-simplistic, including non-negativity, sparsity, low-rank, and even low-dimensional manifolds. Importantly, existing medical images themselves collectively represent high-dimensional complicated distributions, and cannot be analytically expressed to aid tomographic reconstruction. In other words, the gap between the hidden knowledge and the tomographic need is large, demanding a paradigm shift in image reconstruction methods.

Deep learning, as a mainstream of artificial intelligence (AI) and machine learning, enables a paradigm shift for medical imaging in general, and tomographic reconstruction in particular. This data-driven approach relies on learning of the system model mismatches and subtleties as well as the image manifold shapes and relationships from big data using a properly designed network-based reconstruction scheme. This data-driven approach for image reconstruction leads to a new category of image reconstruction methods, being significantly different from analytic and iterative algorithms, and yet such a deep reconstruction network is trained using an iterative algorithm, and, after the training process, used in a feed-forward fashion like an analytic algorithm.

Since 2016, deep tomographic reconstruction has rapidly become the main trend of medical image reconstruction research and development [2]. In June 2018, the first special

Digital Object Identifier 10.1109/TMI.2021.3115547

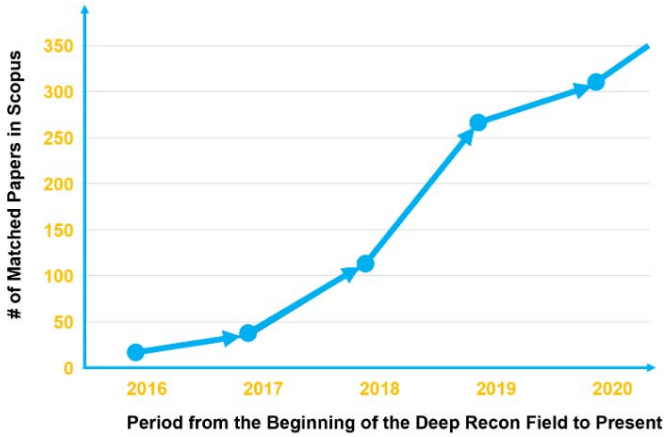

Fig. 1. Field of deep tomographic reconstruction growing rapidly as measured by the number of publications per year.

issue was published on image reconstruction using machine learning especially deep learning methods [3]. That first special issue had a major impact on the field, as partially evidenced by its bibliometric measures, including a total number of citations of 2,601 and a mean number of citations of 130 by Scopus, which is an expertly curated bibliometric database and the largest one of its kind, covering more than 5,000 publishers. We further surveyed the deep tomographic reconstruction literature using Scopus. With the dedicated search rule ("deep learning" AND "medical" AND "image" AND "reconstruct*") within the union of titles, abstracts, and keywords, we had 998 hits. It can be clearly seen that the deep reconstruction field has enjoyed an exponential growth since 2016 with over 300 papers published in the last year alone, as shown in Figure 1.

This remarkable progress in deep tomographic reconstruction can be attributed to several factors [4]. These include big datasets, fast computing platforms, advanced deep learning tools, as well as accumulated reconstruction expertise on imaging modalities. In this context, impressive or even breakthrough imaging results are constantly produced across a wide spectrum of imaging scanners, scanning modes, and clinical tasks, some of which are now published in this special issue [5].

II. Special Issue Summary

This special issue consists of 24 papers. The first 11 papers are focused on the recovery of CT images using deep learning. The remaining 13 papers address issues in a variety of other modalities as well as reconstruction robustness, also known as the hallucination problem.

A. Deep CT Reconstruction

In [A1], Bai *et al.* develop a deep interactive denoiser (DID) by introducing a lightweight test-time optimization process that can run on top of any existing DL-based denoiser to generate images with different noise-resolution tradeoffs. This allows clinicians to interact with the denoisers to efficiently review image candidates with various tradeoffs between resolution and noise, and quickly pick the desired one. In [A2], He *et al.* propose a down-sampled imaging geometric modeling approach for the data acquisition process and incorporate it into a hierarchical neural network, which simultaneously combines geometric modeling knowledge of the CT imaging system and prior knowledge gained from a data-driven training process for accurate CT image reconstruction. In [A3], Ye *et al.* propose a unified supervised-unsupervised (SUPER) learning framework for X-ray CT image reconstruction. The learning formulation combines unsupervised learning-based prior together with (supervised) deep network-based prior based on a fixed point iteration analysis. The proposed training algorithm is an approximate scheme for bilevel supervised optimization, with a network-based regularizer in the lower level and a reconstruction loss in the upper level. In [A4], Wu *et al.* consider the problem of reconstructing high-quality CT images from very few projections. To this end, they introduce a Dual-domain Residual-based Optimization NEtwork (DRONE) that consists of embedding, refinement, and awareness modules. A sparse sinogram is extended, and then the sparse-view artifacts are suppressed by the embedding module; image details are recovered in the residual data and image domains by the refinement module; and finally, the results are regularized by the awareness module. In [A5], Yang *et al.* propose an unsupervised continuous kernel conversion method using a cycle-consistent generative adversarial network (cycleGAN) with adaptive instance normalization (AdaIN) for CT reconstruction using different kernels. By interpolating the AdaIN code, the authors demonstrate that the network can convert images along the interpolation path between two CT images using different filter kernels to highlight different structures. In [A6], Tao *et al.* develop a deep neural network to reconstruct a CT image from view-by-view backprojections (VVBP-Tensor), where the network inputs slices of the VVBP-Tensor as feature maps and outputs the reconstructed image. Numerous experiments reveal that the proposed VVBP-Tensor domain learning framework obtains significant improvement over the image, projection, and hybrid projection-image domain learning schemes.

The next set of papers pertain to CT recovery in the presence of various artifacts and advanced scanning methods. In [A7], Huang *et al.* propose a plug-and-play (PnP) method for truncation correction in CT. Their approach uses a deep learning method to reconstruct a prior image for the extrapolation of missing data first, and then a conventional reconstruction is applied to obtain the final image. With such a PnP method, the region outside the field-of-view (FOV) is restored, which highly relies on the learned image, while the region inside the FOV is reconstructed with full fidelity. In [A8], Zhi *et al.* propose two CNN based models to recover 4-D cone-beam CT images. The challenge is that 4-D CBCT images are degraded by severe streaking artifacts and noise because the phase-resolved image is from an extremely sparse-view CT procedure wherein a few under-sampled projections are used for the reconstruction of each phase. By incorporating the prior image reconstructed from the entire projection data and the correlation among the phase-resolved images, they are able to suppress streaking artifacts and noise while restoring the required features. In [A9], Xia *et al.* develop a approach to

process data obtained under different scanning protocols in CT. Their approach allows training only once, by training a reconstruction network with data originating from multiple alternative geometries and dose levels simultaneously. The geometry and dose levels are parameterized and fed into two multilayer perceptrons (MLPs). The outputs of the MLPs are used to modulate the feature maps of the CT reconstruction network, which condition the network outputs on different geometries and dose levels. In [A10], Hayes *et al.* consider the single source multi-row detector helical CT reconstruction problem. A longstanding challenge is the reconstruction of images from projection data acquired with a helical pitch greater than 1.5. The authors demonstrate that a synergistic use of advanced techniques in conventional helical filtered backprojection, compressed sensing, and more recent deep learning methods can be made to enable accurate reconstruction up to pitches of four without significant artifacts. Finally, in [A11], Zhang *et al.* propose the Comprehensive Learning Enabled Adversarial Reconstruction (CLEAR) method using a progressive improvement strategy for subtle structure enhanced low-dose CT imaging. A generator is established on the comprehensive domain which can extract more features than the one built on degraded CT images and directly map raw projections to high-quality CT images. Then, a multi-level loss is assigned to the generator to push the network components to be updated towards high-quality reconstruction, preserving consistency between generated images and gold-standard counterparts.

The second half of this special issue is focused on MRI, PET, X-ray, ultrasound, microscopy, and reconstruction robustness.

B. Deep MRI

Moving on to deep learning for MRI, in [A12], Zou *et al.* propose a patient-specific deep generative model for the reconstruction of dynamic cardiac MRI. Unlike traditional generative models that require extensive training data, the generator parameters and the time-dependent latent variables that drive the generator are learned from the under-sampled data of the specific subject. The latent variables capture the important dynamic components in the data, including cardiac and respiratory phases. Once learned, the representation can be used to generate data in the desired cardiac/respiratory phases. In [A13], Lahiri *et al.* propose the Blind Primed Supervised Learning (BLIPS) approach for MR image reconstruction by integrating blind and supervised learning for MR image reconstruction. Image features learned by deep supervised learning-based reconstruction algorithms from paired training data are shown to be different from those learned using shallower blind learning-based reconstruction methods such as dictionary learning. Adopting BLIPS techniques reduces the demands on training data for learning-based reconstruction, and improves performance.

In [A14], Oh *et al.* propose an unpaired deep learning method for MR motion artifact removal that does not require matched motion-free and motion artifact images. In this method, outliers in k-space are first rejected by random sampling. Then, the neural network reconstructs fully sampled MR images. Finally, by aggregating multiple reconstructed images from different downsampled images, motion-corrected images are acquired. In [A15], Cheng *et al.* propose Learned DC for fast MR imaging that implicitly learns data consistency with deep networks, corresponding to the actual probability distribution of system noise. The data consistency term and the prior knowledge are both embedded in the weights of the networks, which provides an implicit learning reconstruction model.

C. Deep PET

In [A16], Zhou *et al.* consider low-dose gated PET where gating is utilized to reduce respiratory motion blurring and facilitate motion correction methods. Reducing injection doses causes increased noise levels in gated images that could corrupt motion estimation and subsequent correction, leading to inferior image quality. The authors propose MDPET, a unified motion correction and denoising adversarial network for motion-compensated low-noise imaging from low-dose gated PET data. The denoising network is unified with the motion estimation network to simultaneously correct the motion and predict a motion-compensated denoised PET reconstruction.

D. Deep X-Ray Imaging

In [A17], Fotouhi *et al.* combine information from 2-D X-ray acquisitions in parallax-free domains and enable the stitching of X-ray images with no constraints on the motion of the C-arm. They leverage the Fourier slice theorem to aggregate information from multiple transmission images. The details of the stitched image are subsequently restored using a deep learning strategy that exploits similarity measures designed around frequency as well as spatial image content.

E. Deep Ultrasound Imaging

The two papers in this category focus on ultrasound imaging. In [A18], Huang *et al.* work on an array transducer for tumor tracking in image-guided radiotherapy, which reduces user dependence and anatomical changes. Due to its flexible geometry, conventional delay-and-sum (DAS) beamforming produces B-mode images with considerable defocusing and distortion. To address this problem, they propose an end-to-end deep learning approach that modifies the conventional DAS beamformer when the transducer geometry is unknown. The proposed approach reduces the distortion and improves the lateral resolution and contrast of the reconstructed B-mode images. In [A19], Chen *et al.* introduce ApodNet, which is a deep learning approach for high frame rate synthetic transmit aperture (STA) ultrasound imaging. It contains an encoder to train a set of optimized binary weights as the apodization of high-frame-rate plane wave transmissions, and a decoder to recover the complete dataset from the channel data and achieve two-way dynamic focusing. ApodNet improves image quality when compared with compressed-sensing (CS) based STA at the same frame rate.

F. Deep Microscopic Reconstruction

The first two papers in this part consider deep microscopic imaging. In [A20], Chen *et al.* exploit morphological priors from neurons for training a deep neural network to extract neuron signals from optical microscopy images. They integrate a deep segmentation network in a generative adversarial network (GAN), expecting the segmentation process to be weakly supervised by pseudo-labels at the pixel level while utilizing the supervision of previously reconstructed neurons at the morphology level. In their reconstruction GAN, the segmentation network extracts neuron signals from raw images, and the discriminator network encourages the extracted neurons to follow the morphology distribution of reconstructed neurons. In [A21], Zhang *et al.* produce images with fluorescence molecular tomography (FMT), which is a promising and high sensitivity modality that can reconstruct the 3-D distribution of interior fluorescent sources. They present a 3-D fusion dual-sampling convolutional neural network (UHR-DeepFMT) to achieve ultra-high spatial resolution reconstruction. Experiments demonstrate that the proposed UHR-DeepFMT method achieves an ultra-high spatial resolution, higher than 0.5mm. Finally, in [A22], Li *et al.* propose a fast quantitative differential phase-contrast imaging method with a half-circle pupil. Typically, to generate an isotropic phase, many images are required. Here, a deep learning method is applied to alleviate the slow data acquisition process and generate sufficiently accurate phases from a small number of measurements. The model is tested on seven different types of biological cells.

G. Deep Image Reconstruction

In [A23], Lee and Jeong propose an image denoising method, Interdependent Self-Cooperative Learning (ISCL), that leverages unpaired learning by combining cyclic adversarial learning with self-supervised residual learning. Unlike the existing unpaired image denoising methods relying on matching data distributions in different domains, the two architectures in ISCL, designed for different tasks, complement each other and boost the learning process. The paper includes various biomedical image denoising experiments. Finally, in [A24], Bhadra *et al.* address a critical issue of tomographic image reconstruction, that is, the robustness of the imaging performance with respect to perturbation. They introduce the concept of a hallucination map in order to understand the effect of the prior in regularized reconstruction methods. Clearly, this is a major direction for further research.

III. Testing of the Methods

Eighteen of 24 accepted papers publicly released the associated source codes for reproducible research, among which eight for CT, four for MRI, two for ultrasound imaging, two for optical microscopy, one for PET, and one for electron microscopy. Also, 11 of these 18 groups used PyTorch, five of them used TensorFlow, one of them used Matlab, and one only provided the trained network without any runnable code. In the cases of trained networks or the trainable code available, we tested them systematically.

A. Guideline for Verification

In our evaluation, we made the best effort to verify the shared codes objectively and quantitatively. Specifically, we took the following steps.

First, we checked the consistency of the shared network. The work can be divided into two parts: (1) In the studies where the training code, testing code, trained network, and testing dataset were made available, all the items were tested. This means that the training process of the network was evaluated, and the trained network was tested on the testing dataset; and (2) In the studies where training and testing codes are available, we learned a trained network based on the same dataset used by the authors (referred to as the reproduced network here), and evaluated whether the reproduced network performed as well as the authors' trained network reported in their paper.

Second, to assess the network generalizability, we can use some commonly used new test items, such as a dataset different from the dataset mentioned in the associated paper; a dataset acquired/simulated at a different dose-level; a different imaging geometry (e.g., a different source to detector distance for CT imaging); a different region of interest (ROI) for interior tomography; and a different amount of data (e.g., a different number of views or a different number of k-space samples). Given the time limitation, we only tested the generalizability of the deep CT networks at different dose levels to have a feeling regarding the generalizability of deep CT reconstruction and as a precursor for further efforts along this direction. When we report on our experimental results, the papers are referenced as items No. 1–18.

B. Shared Codes and Code Sharing

In our evaluation, it was found that there were four problems with the 18 available codes: (1) four source codes cannot work correctly due to lack of codes to run the project, such as the code for loading data, and the code to call the trained network; (2) one code failed to calculate the quantitative metrics due to the lack of reference images; (3) one code did not run successfully because the dependency information of the software platform was not precisely specified; and (4) one code cannot run because the authors did not provide the trained network, and the dataset is private. As a result, seven of the aforementioned codes cannot work because of missing information.

In addition to verifying the reproducibility and generalizability of the deep networks, another purpose of this evaluation is to expose problems in code sharing. Based on our experience, we would like to suggest the following points for code sharing: (1) training code; (2) testing code; (3) training dataset (if possible); (4) testing dataset (provide at least the test data used in the paper); (5) the trained network; (6) the code for loading the original data (both for training and testing); (7) images/data for quantitative evaluation; (8) the code for calculating the quantitative metrics if there exist various versions that would bring out different evaluation results (e.g., Structural Similarity (SSIM)); (9) the dependency of the code (anything that the code needs to work correctly); and (10) the

TABLE I
QUANTITATIVE RESULTS OF CONSISTENCE VERIFICATION OF THE SELECTED DEEP RECONSTRUCTION NETWORKS

No.	Metrics	Visual Inspection	Results from the Paper	Our Test/Difference		Remarks
				RN	AN	
1	RMSE	Y	26.7	NA	24.5/8.24%	A,C
2	PSNR	Y	33.5	NA	31.3/6.57%	A,C
3	PSNR	Y	36.75	NA	40.28/9.61%	A,C
4	SSIM	Y	0.8491	0.8022/5.52%	0.8008/5.69%	A,B
5	PSNR	Y	26.63	NA	25.35/4.81%	A,D
6	RMSE	Y	42.1	45.6/8.31%	NA	B
7	CNR	Y	7.44	NA	5.41/27.28%	A,D
8	SSIM	Y	0.76	0.70/7.89%	0.73/3.95%	A,B
9	PSNR	Y	28.34	NA	28.50/0.56%	A,E
10	PSNR	Y	27.0	27.1/0.37%	NA	B
11	SSIM	Y	0.86	NA	0.79/8.14%	A,C
12	Codes incomplete to run the project.					
13	Codes incomplete to run the project.					
14	Codes incomplete to run the project.					
15	No reference to calculate the quantitative metrics.					
16	Dependency information of the software platform is insufficient.					
17	No code other than the trained network.					
18	Codes incomplete, dataset protected.					

Note: NA: Not available;
RN: Using reproduced network;
AN: Using author-provided network;
Y: Visual inspection consistent with the results in the paper;
A: Author-provided network;
B: Network trained by us (reproduced network);
C: Network cannot be trained due to no code for loading data;
D: Network cannot be trained due to data privacy;
E: Network cannot be trained due to lack of the training code.

README file that contains sufficient information for the readers to know about the project, code, and instructions.

C. Results Using Trained/Reproduced Networks

We compared the results reported in the papers with what we obtained using either the trained networks provided by the authors and/or our reproduced networks. In the case that there was no trained network provided by the authors or the author-provided network failed to work correctly, we just produced the results using our reproduced networks. The comparative results are presented in the following table using the quantitative indexes consistent with what were used in the papers, including PSNR (Peak to Noise Ratio), CNR (Contrast to Noise Ratio), RMSE (Root Mean Square Error), and visually based remarks.

When evaluating the codes, we selected one of the representative results from the paper. It can be seen in Table I that the test results are comparable in terms of visual inspection and generally consistent quantitatively with what was reported in the paper. Nevertheless, we still found certain numerical fluctuations, and possibly some biases, for the following reasons: 1) the exact ROI position was not always given in the paper; 2) the authors just gave the average metrics of a set of test data; 3) the test samples were not always specified; and 4) the training dataset was not always the same as that used in the paper.

D. Generalizability Testing

To test generalizability, we only tested the generalizability of six trained deep CT reconstruction networks with different noise levels. L1–L5 represent the noise background from low to high. It is worth noting that the authors usually used test data at different dose levels in various papers. Hence, we estimated the number of photons emitted from the X-ray source at the dose level involved in the test data. Then, we add Poisson noise by assuming that the noise level is the same, or 1.25, 1.5, 1.75, and 2 times of the reference level (corresponding to L1–L5), respectively. The PNSR results are shown in Table II. For comparison, we used the same metric to evaluate the generalizability. The mean and SD (Standard Deviation) values are listed in the rightmost column. Specifically, the SD of PSNR for the six shared networks varies from 0.07 to 1.75. Among these codes, code No. 4 enjoys the smallest SD of 0.07, while code No. 5 is subject to the largest SD of 1.75.

Some comments on these two extremes are in order. For code No. 4, other compromising factors dominate the image quality, making the impact of noise on images relatively small. In the case of code No. 5, the network was trained on a normal-dose dataset for CT reconstruction, and consequently, the network has a poor anti-noise ability. All of the other four codes were designed for low-dose CT reconstruction. The SD of PSNR for these four codes is less than 1.0, which to some extent demonstrates the robustness of the codes against varying

TABLE II
QUANTITATIVE RESULTS OF CONSISTENCY VERIFICATION OF THE SELECTED DEEP RECONSTRUCTION NETWORKS

No.	L1	L2	L3	L4	L5	Mean±SD
1	29.3	29.02	28.65	27.8	26.8	28.31±0.91
2	31.6	31.76	30.99	30.14	29.81	30.86±0.77
3	33.61	33.27	32.48	32.27	31.07	32.54±0.88
4	33.64	33.59	33.57	33.51	33.45	33.55±0.07
5	25.02	24.29	23.26	21.87	20.14	22.92±1.75
6	31.3	31.09	30.87	30.61	30.22	30.82±0.38

noise levels. To various degrees, the generalizability of these networks is not perfect and could be further improved. Such phenomena are not surprising, given the inherent randomness in the training strategy and the well-known instability of deep networks (see the last section of this editorial on future directions for deep reconstruction research).

IV. CHALLENGES AND PERSPECTIVES

A. Generative Models in Relation to Compressive Sensing

The theoretical guarantees and efficient algorithms in compressive sensing have generated significant excitement in several biomedical tomographic imaging areas including MRI [6]. While deep learning methods often offer improved performance over CS methods, they are less theoretically understood. Recently, several authors have shown that algorithms that iteratively project intermediate tomographic solutions to the range of a generative model learned from big data enjoy robustness guarantees, in addition to the benefits achievable using CS algorithms [7], [8]. Learning projection-based deep-learning algorithms as well as optimizing sampling patterns for the best performance are important further directions.

The theoretical robustness guarantees of generative models discussed above have prompted researchers to use pre-learned generative models for image recovery, which is similar in concept to CS algorithms. Current schemes often alternate between enforcement of data consistency and projection to the range of the generator; the projection step replaces the proximal operator used in CS algorithms to enforce sparsity. While such generative models may be ideally suited to model the physics (e.g., learning the spectra [9] or diffusion signal [10] in MRI), a challenge with directly using generative models to represent images is the limited generalizability [11]. Specifically, the restriction of the images to the range space of the generator could make it challenging to recover arbitrary images, which are not necessarily always in the range [8]. More investigation is needed to reconcile the tradeoff between accuracy, robustness, and generalizability.

Another trend is to use untrained generative models for image recovery. Deep image priors [12] and their extensions have been used in a variety of applications with initial success. These schemes use the inherent bias of convolutional networks to natural image content, rather than noise, to recover features from partial data. While the performance of the early methods was not competitive compared to the ones trained with extensive data, recent advances are promising. Similar methods are now emerging for the recovery of dynamic data, where the same generator is shared across time-frames [13], [14]. These approaches enable the recovery of images in a low-dimensional manifold, offering improved performance compared to low-rank methods for dynamic imaging. More efforts are needed to understand these models and enhance their performance.

Finally, there has been growing interest in unrolling or unfolding methods, which are based on unrolling iterations of a given iterative method while learning parameters related to the model and physics of the problem [15]–[18].

B. Robustness of Deep Learning Methods

It is well known that deep learning classification algorithms are sensitive to adversarial perturbations. Recently, several researchers have reported that deep image reconstruction algorithms are also vulnerable to perturbations to the measurement data [19]. It is argued that deep-learning algorithms possess a fundamental tradeoff between increased performance and compromised robustness. However, these studies do not consider the restrictive settings considered in the context of generative models, where robustness is guaranteed [7], [8]. More work is needed to integrate these complementary algorithmic ingredients; for example, adversarial training strategies have also been suggested to improve the robustness of deep learning algorithms [20].

A special challenge in the context of image reconstruction is the sensitivity to model-mismatch [21]. In particular, the network may be trained assuming a specific forward model (e.g., Fourier sampling in MRI) but when the acquisition scheme differs from the assumed model (e.g., scanner non-idealities, differences between scanners, change in sampling patterns in MRI), the reconstruction performance has been reported to degrade more significantly than traditional approaches such as compressed sensing [21]. A possibility is to train the networks with different expected acquisition schemes (e.g., different sampling patterns as in [22]) to minimize the vulnerability to model mismatches. However, this may come at the expense of performance, or it may be challenging to foresee the different acquisition schemes during training. Model adaptation [21] and domain adaptation [23] seem promising directions to mitigate the above challenges.

C. UnSupervised/Semi-Supervised Training

Large fully sampled and high-quality datasets are becoming available to train deep networks [24], [25] in certain settings, which are great research resources. However, such datasets are not always widely available or are impossible to collect in challenging applications such as ultra-high-resolution imaging or imaging of moving objects due to dose, time, or cost restrictions. Similarly, in image deblurring problems, one may not have access to blur-free images. One possibility in this context is to use transfer learning to adapt learned models from another domain or a simulated dataset. Another interesting opportunity is to develop unsupervised or weakly supervised learning methods. Blind-spot-based methods such

as Noise2Void [26] and statistical approaches such as Stein's unbiased risk estimate (SURE) [27] have been proposed to learn deep denoisers without noise-free data. Similar methods are now emerging to solve inverse problems, which is more challenging due to the quality and incompleteness of data coupled with noise. Extensions of blind-spot methods (e.g., SSDU [28]) and statistical approaches (e.g., ENSURE [29], [30]) are encouraging along this direction.

D. End-to-End Mapping

Early deep learning methods relied on training a generic model (e.g., UNET) to recover images from tomographic projections. Motivated by LISTA [15] and related works [22], [31], many of the better-performing algorithms in the fast-MRI challenge [25] are model-based techniques. These methods rely on the unrolling of CS-like iterative algorithms, which alternate between enforcement of data consistency and projection/denoising with a learned model [16]. Such model-based techniques have also been applied successfully to ultrasound imaging [18], [32], [33]. Since the forward model and its adjoint that capture imaging physics are used within the data-consistency blocks, smaller learnable models are often sufficient to offer good performance. This translates to reducing the training data burden. A challenge with these schemes in some settings can be the large memory demand to hold the unrolling steps in large-scale tasks. However, recent exciting results show that such algorithms can be implemented and trained using the concept of implicit layers [34], [35]. While these approaches are inspiring, the convergence issues and tradeoffs between computation and performance need to be further addressed.

The auto-differentiation toolboxes within modern deep learning software have recently been capitalized to optimize the acquisition schemes in different imaging areas [36]–[39]. Many of these methods use differentiable forward models as a layer in the deep network, followed by joint optimization of the parameters. The joint optimization is observed to improve image quality. Similarly, while the main focus of this special issue is on medical image reconstruction, the data preprocessing task is often the first step in the reconstruction pipeline and can be performed via deep learning such as photon-counting data calibration. The reconstructed images are often post-processed (e.g., segmentation, registration, and classification) to extract information, which is then used for diagnosis. The quality of these images can impact the downstream tasks. Similar to image reconstruction, deep learning methods have been making rapid advances in the above steps as well.

There are unprecedented opportunities to combine the different tasks into a unified framework. Early works in this regard show the benefit of synergizing related tasks together including image reconstruction [40], [41]. Surprisingly, the addition of the tasks can also lead to improvement in image quality [40]. The coordination of these tasks may optimize the design of a system that is the most sensitive to the goal at hand (e.g., can improve the early detection of a disease).

E. Explainability

The use of a deep learning model to achieve the final objective (e.g., diagnosis) from noisy measurements has shown to be successful in many cases. However, a challenge with this approach is the general lack of explainability. In particular, it is often difficult to understand why or how the deep network is providing specific results. This is not desirable in a medical imaging setting, especially when the model fails. In addition, large models may require large training datasets and may suffer from a lack of generalizability. A potential alternative may be to combine well-understood modules with task-specific objectives (e.g., reconstruction, followed by segmentation, and finally diagnosis) with associated penalties for each of the objectives. For more details on network interpretation methods, please read a recent review article [42]. Methods such as unrolling also attempt to address this challenge.

F. Reproducibility

The rapid advancement of the deep reconstruction field also faces challenges in reproducibility. It is encouraging to see that more and more groups are making their codes and data publicly and freely available, which is really helpful in improving reproducibility. Similarly, open challenges (e.g., [24], [25]) ensure consistent testing and benchmarking of algorithms. Image processing toolkits, such as Project-MONAI, are instrumental in image post-processing. Similarly, an open-source computational imaging toolkit, where there is commitment from researchers across medical imaging modalities, could be a powerful platform to bring together researchers in this field. For instance, the availability of consistent forward models and image display routines could be helpful in translating the findings from one modality to another and eventually to clinical practice.

V. Conclusion

Recent years have shown that the interest in deep learning is steadily growing. We thus expect further improvements of deep imaging techniques and extensions in biomedical applications. Translation of these methods still demands major efforts to address the challenges discussed above. Given the synergy among and commitments from major stakeholders in academia, industry, healthcare, regulation administration, and patients, we are optimistic that the field will continue growing and move in the directions required for implementation.

Ge Wang, *Guest Editor*
Department of Biomedical Engineering
Rensselaer Polytechnic Institute
Troy, NY 12180 USA
e-mail: wangg6@rpi.edu

Mathews Jacob, *Guest Editor*
Department of Electrical and
Computer Engineering
University of Iowa
Iowa City, IA 52242 USA
e-mail: mathews-jacob@uiowa.edu

XUANQIN MOU, *Guest Editor*
Institute of Image Processing
and Pattern Recognition
Xi'an Jiaotong University
Xi'an, Shaanxi 710054, China
e-mail: xqmou@mail.xjtu.edu.cn

YONGYI SHI
Institute of Image Processing
and Pattern Recognition
Xi'an Jiaotong University
Xi'an, Shaanxi 710054, China
e-mail: xjshiyongyi@163.com

YONINA C. ELDAR, *Guest Editor*
Faculty of Mathematics
and Computer Science
Weizmann Institute of Science
Rehovot 7610001, Israel
e-mail: yonina.eldar@weizmann.ac.il

APPENDIX
RELATED ARTICLES

[A1] T. Bai *et al.*, "Deep interactive denoiser (DID) for X-ray computed tomography," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 2965–2975, Nov. 2021.

[A2] J. He *et al.*, "Downsampled imaging geometric modeling for accurate CT reconstruction via deep learning," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 2976–2985, Nov. 2021.

[A3] S. Ye, Z. Li, M. T. McCann, Y. Long, and S. Ravishankar, "Unified supervised-unsupervised (SUPER) learning for X-ray CT image reconstruction," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 2986–3001, Nov. 2021.

[A4] W. Wu *et al.*, "DRONE: Dual-domain residual-based optimization NEtwork for sparse-view CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3002–3014, Nov. 2021.

[A5] S. Yang, E. Y. Kim, and J. C. Ye, "Continuous conversion of CT kernel using switchable CycleGAN with AdaIN," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3015–3029, Nov. 2021.

[A6] X. Tao, Y. Wang, L. Lin, Z. Hong, and J. Ma, "Learning to reconstruct CT images from the VVBP-tensor," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3030–3041, Nov. 2021.

[A7] Y. Huang, A. Preuhs, M. Manhart, G. Lauritsch, and A. Maier, "Data extrapolation from learned prior images for truncation correction in computed tomography," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3042–3053, Nov. 2021.

[A8] S. Zhi, M. Kachelriess, F. Pan, and X. Mou, "CycN-Net: A convolutional neural network specialized for 4D CBCT images refinement," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3054–3064, Nov. 2021.

[A9] W. Xia *et al.*, "CT reconstruction with PDF: Parameter-dependent framework for data from multiple geometries and dose levels," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3065–3076, Nov. 2021.

[A10] J. W. Hayes *et al.*, "High pitch helical CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3077–3088, Nov. 2021.

[A11] Y. Zhang *et al.*, "CLEAR: Comprehensive learning enabled adversarial reconstruction for subtle structure enhanced low-dose CT imaging," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3089–3101, Nov. 2021.

[A12] Q. Zou, A. H. Ahmed, P. Nagpal, S. Kruger, and M. Jacob, "Dynamic imaging using a deep generative SToRM (Gen-SToRM) model," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3102–3112, Nov. 2021.

[A13] A. Lahiri, G. Wang, S. Ravishankar, and J. A. Fessler, "Blind primed supervised (BLIPS) learning for MR image reconstruction," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3113–3124, Nov. 2021.

[A14] G. Oh, J. E. Lee, and J. C. Ye, "Unsupervised MR motion artifact deep learning using outlier-rejecting bootstrap aggregation," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3125–3139, Nov. 2021.

[A15] J. Cheng *et al.*, "Learning data consistency and its application to dynamic MR imaging," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3140–3153, Nov. 2021.

[A16] B. Zhou, Y.-J. Tsai, X. Chen, J. S. Duncan, and C. Liu, "MDPET: A unified motion correction and denoising adversarial network for low-dose gated PET," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3154–3164, Nov. 2021.

[A17] J. Fotouhi, X. Liu, M. Armand, N. Navab, and M. Unberath, "Reconstruction of orthographic mosaics from perspective X-ray images," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3165–3177, Nov. 2021.

[A18] X. Huang, M. A. L. Bell, and K. Ding, "Deep learning for ultrasound beamforming in flexible array transducer," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3178–3189, Nov. 2021.

[A19] Y. Chen, J. Liu, X. Luo, and J. Luo, "ApodNet: Learning for high frame rate synthetic transmit aperture ultrasound imaging," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3190–3204, Nov. 2021.

[A20] X. Chen, C. Zhang, J. Zhao, Z. Xiong, Z.-J. Zha, and F. Wu, "Weakly supervised neuron reconstruction from optical microscopy images with morphological priors," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3205–3216, Nov. 2021.

[A21] P. Zhang, G. Fan, T. Xing, F. Song, and G. Zhang, "UHR-DeepFMT: ultra-high spatial resolution reconstruction of fluorescence molecular tomography based on 3D fusion dual-sampling deep neural network," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3217–3228, Nov. 2021.

[A22] A.-C. Li, S. Vyas, Y.-H. Lin, Y.-Y. Huang, H.-M. Huanga, and Y. Luo, "Patch-based U-Net model for isotropic quantitative differential phase contrast imaging," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3229–3237, Nov. 2021.

[A23] K. Lee and W.-K. Jeong, "ISCL: Interdependent self-cooperative learning for unpaired image denoising," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3238–3248, Nov. 2021.

[A24] S. Bhadra, V. A. Kelkar, F. J. Brooks, and M. A. Anastasio, "On hallucinations in tomographic image reconstruction," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3249–3260, Nov. 2021.

REFERENCES

[1] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. New York, NY, USA: Cambridge Univ. Press, 2012.

[2] G. Wang, "A perspective on deep imaging," *IEEE Access*, vol. 4, pp. 8914–8924, 2016.

[3] 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, Jun. 2018.

[4] G. Wang, J. C. Ye, and B. De Man, "Deep learning for tomographic image reconstruction," *Nature Mach. Intell.*, vol. 2, no. 12, pp. 737–748, Dec. 2020.

[5] G. Wang, Y. Zhang, X. Ye, and X. Mou, *Machine Learning for Tomographic Imaging*. Bristol, U.K.,: IOP Publishing, 2019.

[6] M. Jacob, J. C. Ye, L. Ying, and M. Doneva, "Computational MRI: Compressive sensing and beyond," *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 21–23, Jan. 2020.

[7] A. Bora, A. Jalal, E. Price, and A. G. Dimakis, "Compressed sensing using generative models," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 537–546.

[8] A. Raj, Y. Li, and Y. Bresler, "GAN-based projector for faster recovery with convergence guarantees in linear inverse problems," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 5602–5611.

[9] F. Lam, Y. Li, and X. Peng, "Constrained magnetic resonance spectroscopic imaging by learning nonlinear low-dimensional models," *IEEE Trans. Med. Imag.*, vol. 39, no. 3, pp. 545–555, Mar. 2020.

[10] M. Mani, V. A. Magnotta, and M. Jacob, "QModeL: A plug-and-play model-based reconstruction for highly accelerated multi-shot diffusion MRI using learned priors," *Magn. Reson. Med.*, vol. 86, no. 2, pp. 835–851, Aug. 2021.

[11] G. Daras, J. Dean, A. Jalal, and A. G. Dimakis, "Intermediate layer optimization for inverse problems using deep generative models," 2021, *arXiv:2102.07364*. [Online]. Available: http://arxiv.org/abs/2102.07364

[12] V. Lempitsky, A. Vedaldi, and D. Ulyanov, "Deep image prior," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 9446–9454.

[13] Q. Zou, A. H. Ahmed, P. Nagpal, S. Priya, R. Schulte, and M. Jacob, "Generative SToRM: A novel approach for joint alignment and recovery of multi-slice dynamic MRI," 2021, *arXiv:2101.08196*. [Online]. Available: http://arxiv.org/abs/2101.08196

[14] J. Yoo, K. Hwan Jin, H. Gupta, J. Yerly, M. Stuber, and M. Unser, "Time-dependent deep image prior for dynamic MRI," 2019, *arXiv:1910.01684*. [Online]. Available: http://arxiv.org/abs/1910.01684

[15] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. 27th Int. Conf. Int. Conf. Mach. Learn.*, Jun. 2010, pp. 399–406.
[16] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Process. Mag.*, vol. 38, no. 2, pp. 18–44, Mar. 2021.
[17] Y. Li, M. Tofighi, J. Geng, V. Monga, and Y. C. Eldar, "Efficient and interpretable deep blind image deblurring via algorithm unrolling," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 666–681, Jan. 2020.
[18] R. J. van Sloun, R. Cohen, and Y. C. Eldar, "Deep learning in ultrasound imaging," *Proc. IEEE*, vol. 108, no. 1, pp. 11–29, Jan. 2020.
[19] V. Antun, F. Renna, C. Poon, B. Adcock, and A. Hansen, "On instabilities of deep learning in image reconstruction and the potential costs of AI," *Proc. Nat. Acad. Sci. USA*, vol. 117, pp. 30088–30095, Dec. 2020.
[20] F. Caliva, K. Cheng, R. Shah, and V. Pedoia, "Adversarial robust training of deep learning MRI reconstruction models," in *Proc. MELBA*, vol. 1, 2021, pp. 1–32.
[21] D. Gilton, G. Ongie, and R. Willett, "Model adaptation for inverse problems in imaging," 2020, *arXiv:2012.00139*. [Online]. Available: http://arxiv.org/abs/2012.00139
[22] 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, Feb. 2019.
[23] 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, Sep. 2018.
[24] C. H. McCollough *et al.*, "Low-dose CT for the detection and classification of metastatic liver lesions: Results of the 2016 low dose CT grand challenge," *Med. Phys.*, vol. 44, no. 10, pp. e339–e352, Oct. 2017.
[25] F. Knoll *et al.*, "Advancing machine learning for MR image reconstruction with an open competition: Overview of the 2019 fastMRI challenge," *Magn. Reson. Med.*, vol. 84, no. 6, pp. 3054–3070, Dec. 2020.
[26] A. Krull, T.-O. Buchholz, and F. Jug, "Noise2Void—Learning denoising from single noisy images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 2129–2137.
[27] S. Soltanayev and S. Y. Chun, "Training deep learning based denoisers without ground truth data," 2018, *arXiv:1803.01314*. [Online]. Available: http://arxiv.org/abs/1803.01314
[28] B. Yaman, S. A. H. Hosseini, S. Moeller, J. Ellermann, K. Uğurbil, and M. Akçakaya, "Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data," *Magn. Reson. Med.*, vol. 84, no. 6, pp. 3172–3191, 2020.
[29] Y. C. Eldar, "Generalized SURE for exponential families: Applications to regularization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 471–481, Feb. 2009.
[30] H. K. Aggarwal, A. Pramanik, and M. Jacob, "Ensure: Ensemble Stein's unbiased risk estimator for unsupervised learning," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 1160–1164.
[31] K. Hammernik *et al.*, "Learning a variational network for reconstruction of accelerated MRI data," *Magn. Reson. Med.*, vol. 79, no. 6, pp. 3055–3071, 2017.
[32] O. Solomon *et al.*, "Deep unfolded robust PCA with application to clutter suppression in ultrasound," *IEEE Trans. Med. Imag.*, vol. 39, no. 4, pp. 1051–1063, Apr. 2020.
[33] B. Luijten *et al.*, "Adaptive ultrasound beamforming using deep learning," 2019, *arXiv:1909.10342*. [Online]. Available: http://arxiv.org/abs/1909.10342
[34] M. Kellman *et al.*, "Memory-efficient learning for large-scale computational imaging," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 1403–1414, Sep. 2020.
[35] D. Gilton, G. Ongie, and R. Willett, "Deep equilibrium architectures for inverse problems in imaging," 2021, *arXiv:2102.07944*. [Online]. Available: http://arxiv.org/abs/2102.07944
[36] V. Sitzmann *et al.*, "End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging," *ACM Trans. Graph.*, vol. 37, no. 4, pp. 1–13, Aug. 2018.
[37] C. A. Metzler, H. Ikoma, Y. Peng, and G. Wetzstein, "Deep optics for single-shot high-dynamic-range imaging," 2019, *arXiv:1908.00620*. [Online]. Available: http://arxiv.org/abs/1908.00620
[38] C. D. Bahadir, A. V. Dalca, and M. R. Sabuncu, "Learning-based optimization of the under-sampling pattern in MRI," in *Proc. 26th Int. Conf. Inf. Process. Med. Imag.*, Hong Kong, Jun. 2019, pp. 780–792.
[39] H. K. Aggarwal and M. Jacob, "J-MoDL: Joint model-based deep learning for optimized sampling and reconstruction," *IEEE J. Sel. Topics Signal Process.*, vol. 14, no. 6, pp. 1151–1162, Oct. 2020.
[40] A. Pramanik, X. Wu, and M. Jacob, "Joint calibrationless reconstruction and segmentation of parallel MRI," 2021, *arXiv:2105.09220*. [Online]. Available: http://arxiv.org/abs/2105.09220
[41] L. Sun, Z. Fan, X. Ding, Y. Huang, and J. Paisley, "Joint CS-MRI reconstruction and segmentation with a unified deep network," in *Proc. 26th Int. Conf. Inf. Process. Med. Imag.* Hong Kong: Springer, Jun. 2019, pp. 492–504.
[42] F.-L. Fan, J. Xiong, M. Li, and G. Wang, "On interpretability of artificial neural networks: A survey," *IEEE Trans. Radiat. Plasma Med. Sci.*, doi: 10.1109/TRPMS.2021.3066428.