nature machine intelligence

REVIEW ARTICLE

https://doi.org/10.1038/s42256-020-00273-z

Deep learning for tomographic image reconstruction

Ge Wang[1]✉, Jong Chul Ye[2] and Bruno De Man[3]

Deep-learning-based tomographic imaging is an important application of artificial intelligence and a new frontier of machine learning. Deep learning has been widely used in computer vision and image analysis, which deal with existing images, improve these images, and produce features from them. Since 2016, deep learning techniques have been actively researched for tomographic imaging, especially in the context of biomedicine, with impressive results and great potential. Tomographic reconstruction produces images of multi-dimensional structures from externally measured 'encoded' data in the form of various tomographic transforms (integrals, harmonics, echoes and so on). In this Review, we provide a general background, highlight representative results with an emphasis on medical imaging, and discuss key issues that need to be addressed in this emerging field. In particular, tomographic imaging is an integral part of modern medicine, and will play a key role in personalized, preventive and precision medicine and make it intelligent, inexpensive and indiscriminate.

When an expecting mother receives an ultrasound exam to have a glimpse of her baby, a real-time dynamic beamforming algorithm is executed. When a patient undergoes a positron emission tomography (PET) scan to evaluate whether a tumour has spread or not, an advanced statistical inference is computed. Computed tomography (CT) scanners visualize internal structures of trauma patients, suitcases or turbine blades, but only after billions of computations are performed. Whether we talk about ultrasound, PET, CT or other tomographic imaging modalities, the measurements produced by the sensors are in a domain that is vastly different from the spatial domain in which the human-interpretable images are formed. The process of computing images from sensor data is what we call tomographic image reconstruction. Reconstruction is the brain of every medical imaging modality.

Background

In CT, PET and single-photon emission computed tomography (SPECT), the sensor data basically reflects line integrals of the image representing an object, and in a simple case the goal of reconstruction is to perform an inverse Radon transform to estimate an image from its line integrals. In magnetic resonance imaging (MRI), the measurements are approximate samples of the Fourier transform of an image, and hence reconstruction aims to estimate the image from its Fourier coefficients, which may be inaccurate, inconsistent and incomplete. In ultrasound imaging, each transducer element records a time signal, where each sample corresponds to a complicated reflection of a transmitted acoustic pulse, and the goal is to reconstruct an image from a large number of such reflections. Finally, in optical imaging the sensor measurements represent complex diffusion and/or interference processes, and the goal is to reconstruct an image of optical properties or light-emitting tracer distributions.

Most image processing tasks use an image as input, and either produce an image that is enhanced and spatially linked to the input image (such as with denoising or semantic segmentation), or produce one or more discrete labels (such as for a feature size measurement or a computer-aided diagnostic report). In the case of image reconstruction, the input is a sensor dataset and the output is an image volume. It is highly non-trivial to compute images from sensor data for the following reasons. Both the sensor datasets and the image volumes contain many millions or sometimes billions of numbers. Moreover, every image voxel value can be defined by thousands of sensor values, and—vice versa—every sensor measurement helps define thousands of image voxel values. Hence, the problem is much more computationally challenging than standard image processing tasks. Moreover, reconstruction may be further complicated by many non-ideal effects, such as patient motion, scattered radiation, non-uniform field gradients, and aberration.

Why does it matter how the reconstruction process is performed? What opportunities are there to further improve the reconstructed image quality? The fundamental answer is that image information is critically important and comes at a cost. The ultimate goal is to perform a diagnostic task with high confidence, while minimizing hardware cost, patient health risk, labour cost, and patient discomfort[1]. This usually boils down to creating images with a certain level of detail or information content, while minimizing scan time, radiation dose, number of sensor elements, and so on. Most often, researchers revert to the easier intermediate criteria of spatial resolution, contrast, noise, image texture, and image artefacts. However, the success of a reconstruction algorithm is ideally assessed using task-based evaluation, such as by cascading reconstruction with image analysis and balancing among all competing factors.

Types of reconstruction. Until recently, most reconstruction approaches could be classified as direct, analytic reconstruction or as iterative reconstruction (Fig. 1). Based on the design of a tomographic imaging device and the knowledge about how it generates sensor data, a forward model can be derived to predict data given an underlying object or subject, either a mathematical model or a numerical (computer) model. Analytic reconstruction is based on the mathematical inverse of the forward model. Examples include

[1]Biomedical Imaging Center, Center for Biotechnology and Interdisciplinary Studies, Department of Biomedical Engineering, School of Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA. [2]Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea. [3]Radiation Imaging Laboratory, Biology and Applied Physics, GE Research, Niskayuna, NY, USA. ✉e-mail: wangg6@rpi.edu

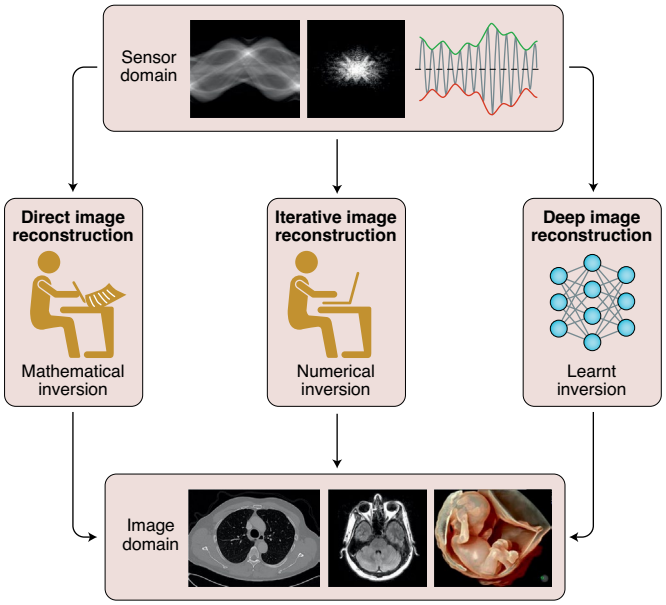

Fig. 1 | Three types of tomographic reconstruction methods. In direct or analytic reconstruction, a mathematical inverse of the forward transform is used to compute images from sensor data. In iterative reconstruction, the current image estimate is iteratively updated such that its forward transform gradually approaches the sensor data. In deep tomographic reconstruction, the inversion does not need to rely on any explicit transform model but is learnt from representative big data. Credit: Molly Freimuth (CT image); Christopher Hardy (MR image); Dawn Fessett (ultrasound image).

filtered back-projection in CT and PET, delay-and-sum beamforming in ultrasound, and inverse Fourier transform in MRI. Iterative reconstruction relies on a numerical forward model combined with a feedback loop. A sensor dataset is calculated based on the forward model and an estimate of the image. The error between this calculation and the actual measurements is back-transformed to the image domain to update the current image estimate. This process is repeated until the error vanishes and the solution converges to an image that well explains the measured data. Iterative reconstruction has been most widely used in PET, SPECT, CT and optical imaging, since the measurements are relatively noisy or a mathematical inverse is unknown or computationally unmanageable. Iterative reconstruction has also been explored for MRI; for example, for reconstruction from data with non-uniform field gradients, where analytical inversion is harder.

Very recently, a third type of reconstruction was conceived: data-driven, deep-learning-based or deep tomographic reconstruction (instead of using shallow networks for reconstruction in early years)[2–9]. This new powerful reconstruction framework naturally evolved from the pervasive shift to validated deep learning techniques in the imaging field at large. Rather than totally relying on an accurate mathematical or physics model, it is empowered by big data with which a deep network can be trained for superior tomographic reconstruction.

Direct reconstruction usually performs poorest in practice, since it starts from an idealized mathematical model. Unfortunately, the real scanner physics deviates from the model in many ways, and the sensor data deviates from what the model would predict. The corresponding errors propagate through the reconstruction process, resulting in image noise, blurring and artefacts. To address this, iterative reconstruction was introduced for various imaging modalities and shined in two ways. First, its reliance on much improved statistical and forward models offers a strong reduction in noise and artefacts[10]. Second, its ability to bring in various types of external prior information greatly enriches the information available to perform a reconstruction and enhances the final image quality. A deep learning framework no longer has to rely on physical insights, although integration of physics within the machine learning pipeline is also a promising area of research. It builds its own model based on lots of training data, and in some cases takes advantage of pre-existing analytic and/or iterative reconstruction algorithms. If the training dataset is too small or not representative, the resulting model would be compromised. However, with larger and more representative datasets, the network model has the potential to exceed the physics-based models and the simplistic image priors traditionally used in iterative reconstruction.

What components of reconstruction can be learnt? It is instructive to look one level deeper and define essential reconstruction elements/aspects that can be learnt or extracted from training data, as opposed to being explicitly defined by the algorithm designer (Fig. 2 and Table 1).

First, there are pre-processing opportunities in the sensor data domain (step 1 in Fig. 2): denoising, deblurring, non-linear corrections and so on. Mathematically, this is written as $\{D_i^{\text{prep}}\} = f_1(\{D_i^{\text{raw}}\})$ where D_i is the sensor data with index i. Direct reconstruction has to rely on separate pre-processing sub-steps to achieve this step. Some iterative reconstruction approaches include a sensor-domain correction step (for example, in-loop beam hardening correction). Similarly, deep reconstruction can learn these sensor-domain processing step and potentially achieve a superior performance through training. Second, there is a geometric relationship from the sensor domain to the image domain (Fig. 2, step 2), which defines how a perturbation in the sensor domain affects the image domain. This transformation of individual sensor data to the image domain is implemented by back-transforming each sensor datum or pre-processed datum (such as after differentiation) to the image domain. As a whole, this second step can be written as $\left\{\left\{I_j^{\text{back}}\right\}_i\right\} = f_2(\{D_i^{\text{prep}}\})$, where I_j is the image with voxel index j: one image for each sensor data index i. For line integrals, the back-transformation of a single non-zero data point is a line in the image space. For Fourier coefficients, the back-transformation of a single non-zero data point is a complex, sinusoidal wave in the image space. Both direct and iterative reconstruction techniques can utilize a known transform to perform forward and backward transforms (for example, projection or back-projection). The explicit use of this transform relationship in deep reconstruction is useful since it facilitates/simplifies the network-based inversion process. We will later refer to some methods that learn this relationship whereas others use an explicit operation for it. Third, the actual inversion step aims to compute an image that explains the data (step 3 in Fig. 2). This is the step where the actual set of equations is solved, either mathematically, numerically or by learning. Mathematically, this step is expressed as $\left\{I_j^{\text{inv}}\right\} = f_3\left(\left\{I_j^{\text{back}}\right\}_i\right)$. For example, this is equivalent to a filtration operation or a weighted summation of all sinusoidal components in direct reconstruction. In iterative reconstruction, this is the iterative update step or the numerical solver. This third aspect of reconstruction is an opportunity for deep reconstruction to shine, since any human-designed optimization approach is by definition sub-optimal, as shown by continuing improvements in this field. Fourth, there is a major benefit in introducing prior information, or information that was not yet contained in the actual sensor data (Fig. 2, step 4). In direct reconstruction, this step could have the rudimentary form of a post-processing filter: $\left\{I_j^{\text{post}}\right\} = f_4\left(\left\{I_j^{\text{inv}}\right\}\right)$. In iterative reconstruction, edge-preserving Markov-random-field regularization is a

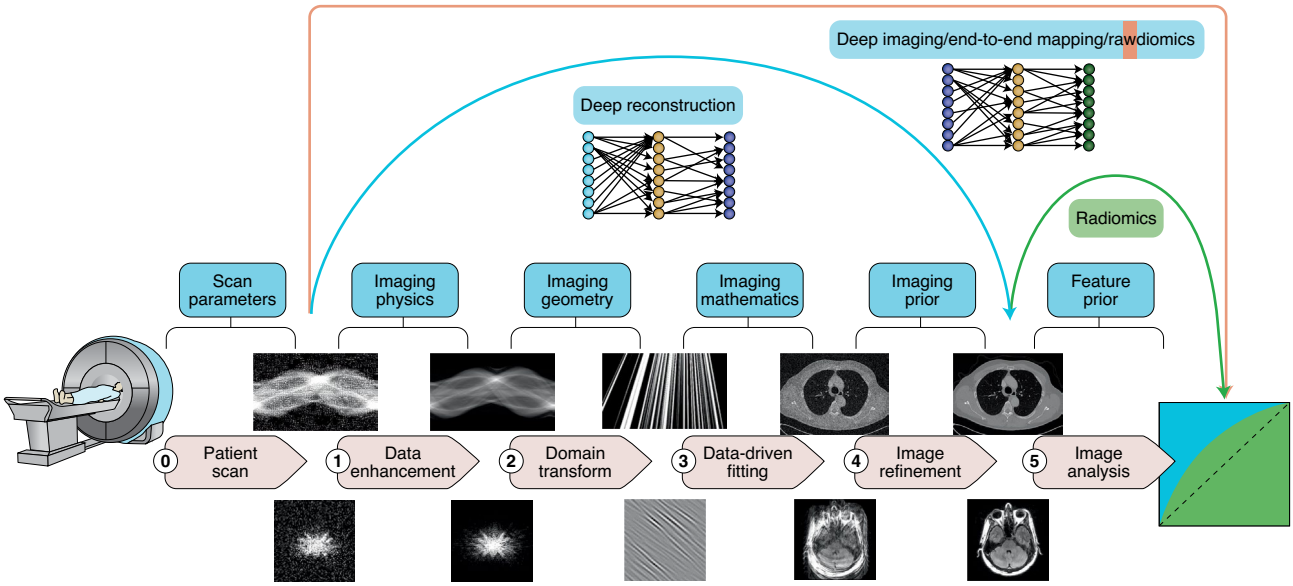

Fig. 2 | The reconstruction process can be broken up into four essential steps, each of which can either be learnt from the training data or be explicitly defined by the algorithm designer. Most deep reconstruction approaches can be classified based on which steps they learn from the data. Multiple steps can be combined together and learnt. Besides the actual image reconstruction, deep learning also offers opportunities to optimize the scan parameters up front and is widely used for analysing images in the back-end. Credit: Molly Freimuth (CT image); Christopher Hardy (MR images).

common types of prior information[11]. Multi-modality approaches such as magnetic resonance (MR)-based partial volume correction in PET are another example[12]. When deep reconstruction learns to match input sensor data to output image data, it does not only learn to solve the data fit problem but also incorporates comprehensive prior information in a data-driven manner. Perhaps the greatest strength of deep reconstruction is that it can learn far more complex prior information than the traditional regularizers used in iterative reconstruction. The ability to ultimately learn both the best possible inverse solver and the most specific prior information makes deep reconstruction very powerful. Deep reconstruction approaches reported in literature may incorporate or skip one or more of the above steps and we will refer to this in the following section and in Table 1. Also note that the four steps described above do not always occur in a sequential order.

Early history. When CT was first developed in the early 1970s, a rudimentary iterative reconstruction was used[13]. A few years later, direct mathematical inverse approaches were proposed for computational reasons[14]. It was not until 2008 that iterative reconstruction was deployed in clinical routine through a series of new product introductions by multiple CT vendors. This resulted in high-quality images with only a fraction of the traditional radiation dose, but it came with increased computational cost and sometimes image texture degradation. The use of a neural network for tomographic reconstruction dates back to around 1990. For example, Floyd et al.[2] applied a neural network for SPECT image reconstruction, and Yan et al.[3] employed a multilayer neural network for MR data truncation artefact removal. These networks were shallow with limited image quality. Modern tomographic networks are deep and are trained by big data on high-performance computers. Their results are impressive, with CT noise reduction exceeding that of iterative reconstruction, preserving good image texture and reducing computational complexity compared to iterative reconstruction[15–17] (Fig. 3, CT).

In MRI, the inverse Fourier transform of the *k*-space measurements has always been the standard approach. Parallel imaging[18–20] was introduced in the late 1990s to overcome the long scan times, associated patient discomfort and motion artefacts, aided by the known sensitivities of multiple radiofrequency receive coils as prior information. This was followed a decade later by compressed sensing (CS)[21], an iterative method that enforces data consistency while regularizing the reconstruction by imposing sparsity or low rank in some transform domain, typically wavelets or total variation. More recently, deep neural networks have been proven to be a powerful method to further reduce MR scan time and motion challenges[22–26] (Fig. 3, MRI).

A similar evolution happened in emission tomography modalities PET and SPECT, where analytic reconstruction techniques such as Fourier re-binning[27] have now been replaced by fully three-dimensional iterative reconstruction techniques (leading to the well-known 'Fully-3D meeting'). In ultrasound imaging, simple delay-and-sum methods have evolved to time-reversal methods[28] and advanced dynamic beamforming methods; iterative methods have not yet become mainstream. In both emission tomography and ultrasound imaging, deep learning is currently an area of intense research and it is most likely that it will soon play a critical role in these imaging modalities too (Fig. 3).

The year 2016 was a special year in the history of deep tomographic reconstruction. At the 2016 Low-Dose X-ray CT Grand Challenge organized by the American Association of Physicists in Medicine (AAPM), Kang et al.[15] won the second place with a deep learning approach, suggesting the potential of deep tomographic reconstruction to surpass mainstream reconstruction algorithms. In parallel, Chen et al.[16] introduced a similar convolutional neural network (CNN) for low-dose CT denoising. In the same year, several deep learning methods were also proposed for MR reconstruction such as variational neural networks[23], multilayer perceptrons[24], unfolded iterative methods[25], and deep learning priors for MRI[26]. The successful demonstration of deep learning for CT and MRI reconstruction, which can outperform the state-of-the-art compressed sensing approaches, have inspired many deep-learning-based image reconstruction activities in other modalities, quickly forming several important technical approaches: image domain learning, unrolled approaches, model-based methods, plug-and-play methods, domain transform approaches, sensor domain learning, GAN based approaches, and so on.

Table 1 | Recent deep learning tomographic reconstruction methods for various biomedical imaging modalities based on their technical contributions

Modality	Architecture	Applications and citations
X-ray CT	Image domain	Low-dose[15–17,48,70,97], sparse view[29,113], limited angle[50], metal artefact[52], dual-energy[143]
	Unrolling	Sparse view[29,58,59,61,62], low-dose[60]
	Sensor domain	Metal artefact[85,86], sinogram analysis[87]
	Model-based/plug-and-play	Low-dose[60,68,70,82,144], sparse view[58,59], interior[83], cone-beam artefact[84]
	Domain transform	Low-dose and sparse view[77–79], interior tomography[83]
	GAN, unsupervised	Low-dose[48,97,110], metal artefact[130]
MRI	Image domain	Accelerated MRI[24,30,49]
	Unrolling	Accelerated MRI[23,25,53–56]
	Sensor domain	Accelerated MRI[88,90,91], artefact removal[89]
	Model-based/plug-and-play	Accelerated MRI[26,63,72], dynamic MRI[65,75]
	Domain transform	Accelerated MRI[76]
	GAN, unsupervised	Accelerated MRI[49,109], contrast synthesis[106,107]
PET/SPECT	Image domain	Low-dose PET[31,32,145,146]
	Model-based/plug-and-play	Low-dose PET[31,67–69]
	Domain transform	Low-dose PET[80,81]
	GAN, unsupervised	Low-dose PET[111], attenuation correction[104,105]
Ultrasound	Image domain	Artefact removal[45], photoacoustic artefact removal[46,47]
	Sensor domain	Radiofrequency interpolation[96]
	Model-based/plug-and-play	Photo-acoustic imaging[64]
	Domain transform	Beamformer[92,94,95]
Optical	Image domain	Super resolution[36–38,103], ghost imaging[42], scattering medium imaging[40,41], mobile phone microscopy[43]
	GAN, unsupervised	Deconvolution microscopy[100,102,103]

State of the art

In this section, we will review representative references in the following complementary technical categories.

Image domain learning. Given the availability of well-established deep learning models from computer vision applications, one of the most straightforward ways of applying deep learning for tomographic reconstruction is to reduce image artefacts as a post-processing step using image domain deep networks (step 4 in Fig. 2).

For example, Jin et al.[29] demonstrated that an encoder–decoder network architecture with a large receptive field like U-Net, which had been mainly used for image segmentation, is also very effective in removing streaking artefacts commonly encountered in sparse-view CT reconstruction. In MRI, some earliest deep learning work[24] was implemented using multilayer perceptrons in the image domain. A magnitude and phase domain encoder–decoder network was proposed to remove aliasing artefacts in accelerated MRI[30]. For PET reconstruction, Gong et al.[31] designed an image domain network to reduce low-dose PET image noise. Qian et al. proposed a deep learning approach for scatter estimation and improved PET image quality[32]. In optical microscopy, artefacts usually appear as image blurring, interference, and so on, which can be easily addressed using image domain deep learning. For single-molecule localization microscopy, deep learning methods[33–36] were implemented using an encoder–decoder architecture to remove image blur from fluorescent probes. Similar image domain networks were extensively used to improve the spatial resolution of microscopy[37,38] or to solve the holographic microscopy problem[39]. Other interesting applications include optical imaging through scattering medium[40,41], ghost imaging[42], mobile phone microscopy[43] and so on.

Ultrasound imaging and photoacoustic imaging[44] also suffer from distinct image artefacts from various sources, such as speckle noise, blur and so on. Accordingly, an ultrasound image denoising method was proposed[45]. Deep learning was also used to mitigate the limited angle and sparse data problems in photoacoustic imaging[46,47].

To further improve the performance of image-domain approaches, researchers have improved network architectures and refined loss functions[48,49]. For example, some of the earliest work for the AAPM low-dose CT Grand Challenge was implemented in the contourlet domain so that the contourlet transform can help identify the noise components, and neural networks can be trained more easily[15]. In the case of limited angle CT, Bubba et al.[50] used the shearlet transform[51] to separate learnable and shearlet-based parts. To deal with severe metal artefacts, which are difficult to remove using a simple image domain network, Zhang et al.[52] uses ensemble learning from multiple types of image domain CNNs. For low-dose CT imaging, Shan et al. proposed a modularized adaptive processing neural network, which brings radiologists-in-the-loop (that is, radiologists help select the number of denoising modules as a hyperparameter in a task-specific fashion) and performs either favourably or comparably relative to commercial iterative reconstruction algorithms[17].

Unrolling approaches. Unfortunately, image-domain learning approaches often suffer from image blurring, especially when the training data is not sufficient. This limitation has led to various other technical contributions, among which unrolling or unfolding methods are one of the most widely investigated approaches. These approaches combine a conventional iterative algorithm to perform the inversion (step 3 in Fig. 2) with a learnt prior (step 4 in Fig. 2).

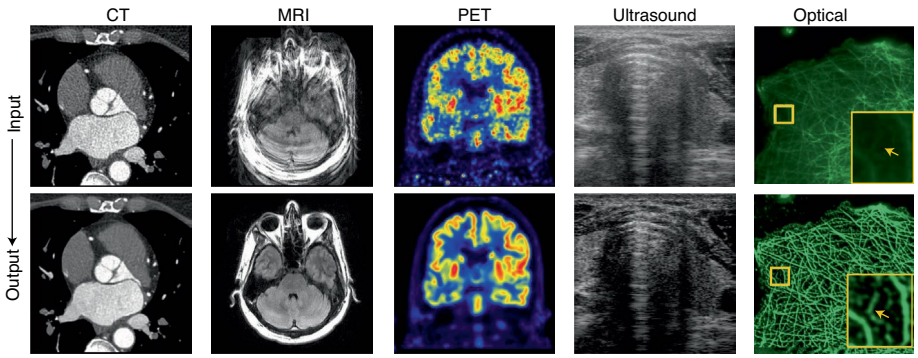

Fig. 3 | Deep learning reconstruction results for various imaging modalities. From left to right: deep reconstruction has shown great promise to suppress CT image noise while preserving image details and texture. Deep learning methods allow the MRI acquisition to be sped up without introducing sampling artefacts. PET fidelity was greatly improved by applying deep learning with MR priors. Deep learning improves image contrast in a low-power ultrasound (US) image. Deep learning recovers spatial resolution and structural details in an ill-posed optical imaging task. Credit: Froedtert & Medical College of Wisconsin (CT images); Christopher Hardy (MR images); Quanzheng Li (PET images).

For example, some of the earliest deep networks for compressed sensing MRI, often referred to as variational networks[23] and ADMM-Net[25], are based on unfolding the steps of an iterative sparse reconstruction, and map each step to a layer/block of a deep neural network. Schlemper et al.[53] used a similar idea for dynamic MRI, Vishnevskiy et al. for 4D flow imaging[54], and Eo et al.[55] proposed the KIKI-Net that alternates between the *k*-space and the image domain. This unrolling strategy was then combined with a recursive neural network for MR reconstruction[56]. A variational network was also successfully used for ultrasound imaging[57].

In CT, Jin et al.[29] adapted a U-Net for sparse view CT as an unrolled iteration for multi-resolution sparse recovery using wavelets. Learned primal-dual[58] and projected gradient[59] were proposed by unfolding the iterative solutions to the sparse recovery problems. Chen et al. unrolled a classic iterative algorithm into a neural network implementing two paths: a classical data-fit update and a prior update with trainable parameters[60]. Furthermore, unrolling techniques were used as an alternative deep learning approach to momentum-type methods for accelerated iterative reconstruction[61,62].

Although the unrolling strategy provides a way to design a new type of deep neural network architectures, it often makes the resulting neural network overly complicated. For example, in ADMM-Net[25], updating Lagrangian parameters is implemented as a number of network layers. Given the success of much simpler and flexible neural network architectures from the computer vision literature, it is not clear whether duplicated layers from unrolling are really or always necessary. A rigorous study of the pros and cons of unrolling is an interesting direction for future research.

Model-based and plug-and-play approaches. Rather than explicitly mapping each iterative step to a layer of an unrolled neural network, model-based and/or plug-and-play approaches incorporate a deep neural network as a prior term in the iterative reconstruction procedure. For example, deep learning MR reconstruction methods by Wang et al.[26] employed a CNN-based prior term for compressed sensing MRI. This idea was extended by Aggarwal et al.[63] as a model-based deep learning framework that systematically uses a pretrained neural network for regularization. A similar idea was employed for photo-acoustic imaging[64]. Traditionally, the model-based methods were widely used for dynamic MRI problems. Recently, Biswas et al.[65] proposed a model-based deep learning prior that incorporates temporal redundancies. Lyu et al. developed an advanced deep neural network incorporating both long short-term memory and generative adversarial network (GAN) for spatial-temporal-based prediction to improve cine cardiac MRI[66]. Model-based reconstruction approaches using deep learning priors were also successful for PET[67,68] and SPECT[69]. Gao et al. developed an anatomy texture prior for high-fidelity CT reconstruction[70].

Another variation is an iterative reconstruction framework with a generalized denoiser, in the class of so-called plug-and-play methods[71]. For example, CNN-based denoisers were used in the plug-and-play approach for MRI reconstruction[72] and optical tomography[73] respectively. Yet another type of related methods is referred to as deep image priors[74]. Rather than using a pretrained model, the neural network architecture itself provides a type of regularization, and the network parameters should be estimated by fitting to the measurement data without separate training. For example, Gong et al.[31] incorporated the deep image prior as a regularization term, estimated during the expectation maximization update for PET. Jin et al.[75] proposed a time-dependent deep image prior to incorporate the temporal redundancy in a dynamic MRI scan.

Although the model-based and plug-and-play approaches have several advantages over the image domain learning by providing better image quality from a smaller set of training data, one of the main limitations of these approaches is that the resulting neural network is not a feed-forward formulation, and computationally expensive. Therefore, their use may be more likely in the applications where high accuracy outweighs reconstruction time.

Domain transform approaches. A number of groups explored directly learning a tomographic mapping from sensor data to an underlying image (steps 2 and 3 in Fig. 2). With the automated transform by manifold approximation (AUTOMAP) technique, Zhu et al.[76] learnt the entire reconstruction process through a manifold encoding–decoding process with multiple fully connected layers. Ye et al.[77] encoded the system geometry in single-view back-projection (step 2 in Fig. 2) while relying on the network to perform the actual tomographic inversion (step 3 in Fig. 2). In the iCT-Net[78], the rotational symmetry of the back-projection was explicitly coded to reduce the dimensionality of the network. Fu et al.[79] developed a hierarchical framework to manage the dimensionality of fully learnt reconstruction and utilized data correlation to reduce the fully connected approach to a sparse-matrix formulation. Following the early successes in CT and MRI, several groups demonstrated the use of deep learning to perform the entire sinogram-to-image PET reconstruction[80,81].

However, a main limitation of the direct mapping from raw data to an image is its critical dependency on big data and expensive computational cost, especially the huge GPU memory requirement.

Care must be taken for the direct mapping to generalize well and avoid instability. Moreover, the dedicated neural network can be limited to only one specific reconstruction geometry, rather than being widely applicable to various scanner architectures and scan protocols.

Sensor domain learning. To mitigate the limitations of the domain transform approaches, some networks embed an analytic transform such as the Radon transform and the Fourier transform as imaging-physics-based knowledge inside the reconstruction network and work in the sensor domain. For example, Wurfl et al. obtained a data-driven ramp filter in end-to-end learning[82]. Han et al. showed that a neural network can be designed in the differentiated back-projection domain so that it can be generalized well for interior tomography[83] and cone-beam artefact removal[84]. For metal artefact reduction, Claus et al.[85] proposed a deep learning method to estimate missing sinogram data, which was further refined with a GAN by Ghani and colleagues[86]. De Man et al.[87] proposed to bypass the reconstruction process entirely and presented a deep learning approach going directly from the CT sinogram domain space to the image feature space. In recent *k*-space deep learning studies[88–90], it was observed that a deep neural network in the *k*-space can estimate the interpolation kernels. Since many MR artefacts come from the *k*-space data inconsistency, *k*-space deep learning is desirable to remove MR artefacts such as the echo planar imaging ghost[89]. The idea of *k*-space learning was recently combined with sparsity and low rank in a model-based approach[91].

In ultrasound imaging, image reconstruction is usually performed using a beamformer, which combines radiofrequency channel data by delay-and-sum. However, delay-and-sum often introduces image artefacts, especially when the number of receive channels, bandwidth and aperture size are not sufficient. To address this, the deep learning approach has been used to directly process the radiofrequency channel data. For example, Luchies and Byram[92] proposed a frequency domain deep learning method to improve the quality of the beam former. Also, a deep neural network was proposed for coherent compound imaging from a small number of illuminating plane waves[93]. For focused B-mode imaging, adaptive and/or compressive deep beam forming methods were developed[94,95]. Additionally, a deep neural network was proposed to interpolate missing radiofrequency-channel data from multi-line acquisitions for accelerated ultrasound scanning[96].

GAN and unsupervised learning approaches. Although most deep learning techniques are supervised methods, which learn the relationship between input data and matching ground-truth or label data, in practice it is often impractical to obtain training data pairs. For example, it is not clinically feasible to acquire both low-dose and high-dose CT scans for the same patient. In microscopy imaging, the ground-truth high-resolution images are usually unknown. In these cases, having matched data pairs is only possible through simulation.

To address this problem, Wolterink et al.[48] proposed a GAN that takes low-dose images and unmatched high-dose images for training. The generator estimates high-dose images from the low-dose image inputs. The discriminator tries to tell whether its inputs are true high-dose images or high-dose images produced by the generator. To prevent plausible artefacts produced by the GAN, Kang et al.[97] adapted a cycle-consistent GAN (cycleGAN)[98]. For microscopy applications, Nguyen et al.[99] developed a conditional GAN (cGAN) to reconstruct video sequences of dynamic live cells as captured with Fourier ptychograpahic microscopy. Lee et al.[100] employed the standard cycleGAN for 3D fluorescent microscopy. In a recent analysis using optimal transport theory[101], it was shown that more robust unsupervised deconvolution is possible using a modified form of the cycleGAN where one of the generators is deterministic, or partially known.[102]. Li et al. addressed the unavailability of missing paired training data by physically realistic modelling of the stimulated emission depletion image formation process for super-resolution[103]. Even when paired data are available, semi-supervised learning with discriminative and mean squared losses is beneficial for training to improve image resolution as demonstrated for deep MR reconstruction[49].

Another important GAN application is image synthesis. For MR-PET applications, the GAN approach was used to generate CT images from MRI images, from which the attenuation background can be obtained[104]. Instead of this indirect approach, a reconstruction network was recently proposed to correct PET images without ever reconstructing the CT images[105]. In MRI, some of contrasts can be missing, such as T_1, T_2, diffusion and so on, and whenever clinically indicated a patient is often called back for one or more scans at additional costs. To address this problem, Dar et al.[106] proposed a cycleGAN to convert MRI images in two contrasts so that a missing contrast image can be generated from the other one. By extending this idea to multiple contrasts, Lee et al.[107] proposed the so-called collaborative GAN (CollaGAN) approach that utilizes acquired multiple contrast images to estimate the corresponding image in a missing contrast. This idea was extended to assess the importance of a specific MR contrast for a given clinical application[107].

Aside from GAN-based approaches, there are other classes of unsupervised learning methods for image reconstruction. Several methods were proposed based on the Noise2Noise idea[108], which can be regarded as self-supervised learning, since images themselves or their noisy versions are used as the reference for network training, as demonstrated in the cases of MRI[109], CT[110] and PET[111] respectively.

Connection to compressed sensing. With the success of deep learning for image reconstruction, a puzzling question from the early development is why deep learning works, and whether the results from deep neural networks are real or cosmetic changes. For the past decade, the most widely used mathematical framework for image reconstruction is 'compressed sensing' or 'sparse representation'[112]. The key idea of compressed sensing is to find an optimal set of sparse basis functions for a given set of measurement data. While deep learning methods are quite different from compressed sensing, the recent studies[113,114] showed that indeed there exists a close relationship between them. Specifically, the classic image reconstruction often requires a top-down design of the basis, after which a sparse subset of the basis functions is selected via computationally expensive optimization. On the other hand, in deep-learning-based image reconstruction, the basis functions can be adaptively generated with the filters learnt from training data. Therefore, rather than resorting to basis engineering to find the optimal basis, deep tomographic reconstruction finds the optimal basis representation in a data-driven way, which makes the algorithmic development conceptually simpler and practically more effective.

Challenges and opportunities

With the advancement of the deep learning field, there are major challenges and opportunities. In the following, we discuss four of them: (1) big data, weakly supervised and unsupervised learning; (2) point of care (POC), hybrid and autonomous imaging; (3) network stability and image quality assessment; and (4) explainable/interpretable and ethical artificial intelligence (AI).

Big data, weakly supervised and unsupervised learning. The availability of big data is currently a general prerequisite for deep tomographic image reconstruction, deep image analysis, and end-to-end mapping from raw data to diagnostic features. The most fundamental challenge with using clinical data for training deep reconstruction networks is that such data and images are limited,

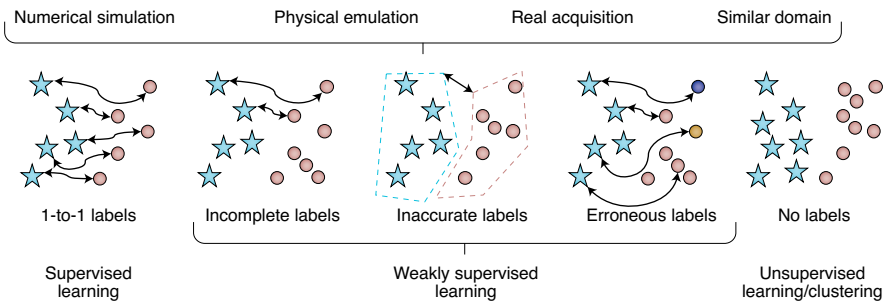

Fig. 4 | Big data acquired from various sources and labelled in different forms. Paired data allow supervised learning. In many scenarios, labels are incomplete, inaccurate and/or erroneous for weakly supervised learning. For unlabelled data, unsupervised learning or clustering must be done and is the most challenging.

with their labels typically unknown, inaccurate and/or insufficient. The ground truth can be approximated by scanning at high dose, over long time, and using sophisticated hardware, but all these are either costly or totally infeasible (say, advanced hardware unavailable or under development). Unlike in the field of computer vision where labelled natural images are readily available, clinical tomographic images and raw data, such as CT sinograms, are generally unavailable.

To address the aforementioned critical need, great efforts have been made to collect and share big data; for example, by the National Institutes of Health (NIH)[115] and Mayo Clinic[116], and on a Radiopaedia list of public datasets[117]. Such datasets may not always be large enough; say only thousands of images in a dataset. On the other hand, proprietary datasets and most specific hospital datasets are not public. Clearly, expansion and integration of such datasets would be beneficial to the field.

Currently, most of the deep learning methods for image reconstruction were developed through supervised learning, but acquiring labels remains a major problem. Given the issues of patient privacy, clinical workload, trade confidentiality and technical limitations, one solution is to simulate training data for deep tomographic imaging[4]. Computational modelling has been constantly improved over the past years. Prior to simulating tomographic data, we need to have realistic patient body models. Image modelling has been studied over decades; for example, virtual patients[118] and XCAT phantoms[119]. Recently, Mueller et al. showed how to synthesize clinical images with deep learning[120]. As far as CT simulation is concerned, there are multiple simulators[121–124]. Monte Carlo simulation tends to have a very high computational complexity, and deep learning can accelerate such simulation since it can be essentially treated as a denoising problem. Although clinical trials remain the gold standard, a Food and Drug Administration team performed a project called Virtual Imaging Clinical Trial for Regulatory Evaluation (VICTRE) to evaluate digital breast tomosynthesis[125]. In VICTRE, the whole imaging chain was simulated, and favourably compared to an existing human trial in a premarket submission for lesion detection. The computational characterization will greatly reduce the costs, and accelerate the translation of AI imaging technologies. Hopefully, many more datasets will be simulated for a majority of imaging tasks, and then the trained networks can be fine-tuned with physical phantoms and real patient scans. Since deep tomographic algorithms are data-driven, and data keep growing, deep imaging algorithms must be constantly evolved in an open environment to absorb new data and distil semantic knowledge from distributed sources. In this context, continuous/incremental learning[126] enables the autonomous refinement of a trained neural network, and federated learning[127,128] learns a shared network model collaboratively while keeping all datasets under each user's control. In this way, users can customize others' codes to fit their specific protocols and data to generate a better network via ensemble learning.

As shown in Fig. 4, clinical datasets are often only weakly labelled and even unlabelled. It is challenging to perform weakly supervised and unsupervised learning. Substantial efforts have been made along this direction; for example, cycleGAN was adapted for low-dose CT denoising[97], self-learning such as Noise2Void for denoising[129], and the disentanglement network was designed for CT metal artefact reduction[130]. The key for such disentanglement is to incorporate various kinds of prior knowledge and enforce consistency among them. However, the barrier is substantial to elevate the dataset-wise correspondence (for example, between metal-free and metal-affected images) to the element-wise correspondence (between a specific metal-affected image and its ground truth), since the former is much weaker than the latter. Many medical imaging problems can be cast as weakly supervised and unsupervised learning tasks in the disentanglement framework. When real-world data is governed by a full array of explanatory factors, many inverse problems are not unique due to lack of information, such as blind image de-blurring. Recently, the ill-posed nature of learning disentangled representations was theoretically and experimentally demonstrated[131]. To enable successful weakly supervised learning, adequate inductive biases or domain-specific priors/constraints are needed (such as a learnt dictionary, low-dimensional manifold, or a deep prior, which can be applied to improve image reconstruction).

POC, hybrid and autonomous imaging. The content in this Review is largely focused on mainstream tomographic scanners in hospitals and clinics. Now, the NIH actively promotes the POC concept so that a patient is better served and could play a greater role in their healthcare. As shown in Fig. 5, POC can be facilitated with mobile and compact devices, wireless connection, and AI techniques for data collection, reconstruction/inversion and analysis. Currently, all major medical imaging modalities already have mobile versions, such as the wearable head PET scanner of West Virginia University, the lightweight MRI unit of Hyperfine, and recently the Evry MRI scanner designed with a 0.5-T magnetic field which can be switched on and off within 15 minutes, as well as the iQ ultrasound probe of the Butterfly Network. Since a POC device is generally a low-cost and compact version of a regular imaging scanner, its performance may be sometimes limited by performance-compromised hardware components. In any case, POC should be empowered by AI to offer best possible imaging services.

Since multi-modality imagers are widely used, including PET-CT, SPECT-CT and PET-MRI, mobile/portable multi-modality imagers are similarly attractive. An example is the concept of simultaneous spiral MR and X-ray (MRX) scans[132]. It is underlined that the aforementioned POC imagers, in individual or hybrid modes,

Fig. 5 | Existing mobile/compact imagers and future possibilities for hybrid imaging. Shown with exemplary robotic C-arm, mobile CT, wearable PET, low-field MRI, and portable ultrasound devices. Credit: Yan Xi (C-arm); Judy Samuelman (CT scanner); Qiyu Peng (wearable PET device); Alex Panther (MRI scanner); Homer Pien (ultrasound unit).

are the precursor of futuristic imaging service. In a good sense, civilization is spanned by decentralization and connectivity. We believe in the trend from 'internet of information' and 'internet of things' to 'internet of service'. We envision a paradigm shift from hospital/clinic-oriented imaging to mobile/integrated services deliverable wherever and whenever needed. Our concept Auto-driving Vehicle-based Affordable Tomography-Analytics Robot (AVATAR)[133] is an example, which could be desirable on disaster scenes, for routine healthcare imaging in rural areas and under-developed countries, as well as during social distancing.

Network stability and image quality assessment. In ref. [134], three kinds of instabilities were reported of deep neural networks for tomographic reconstruction. For some earlier deep networks, weak perturbations could result in strong artefacts, a clear tumour-like change might go undetected, and more data could yield less network performance. This challenge is neither surprising nor worrisome since such critiques will promote further development of this field. In the earliest days of reconstruction, although tomographic reconstruction is theoretically non-unique from a finite number of projections, a regularization technique, such as a limited bandwidth constraint, led to accurately reconstructed cross-sectional images. Then, iterative reconstruction algorithms became popular. While iterative reconstruction could be biased by a strong penalty term, it has been made practical using an optimized imaging protocol. Furthermore, despite the well-recognized status of compressed sensing as a modern iterative reconstruction approach, it still has a chance to produce a false sparse solution; for example, pathological details might be wiped out if total variation is overly minimized. Currently, deep tomographic reconstruction faces a similar situation of not being perfect but holds a great promise to make much progress in the near future. Briefly speaking, the aforementioned instabilities were produced by adversarial attacks, and we have a wide literature base to adapt anti-attacking techniques for robust tomographic reconstruction. Essentially, prevention of adversarial attacks is a common challenge to the entire AI field, and rapid development is under way, with several effective anti-attacking techniques already reported[135]. A more specific key point is that a deep network can be synergized with established analytic and iterative methods including compressed sensing techniques so that the imaging performance will be only improved; for example, deep priors should outperform traditional priors because it is data-driven and domain-specific, and a recently proposed Analytic, Compressive, Iterative Deep (ACID) network produced stable and accurate results eliminating all three kinds of instabilities[136].

Given the very high healthcare relevance, it is critical to address how to evaluate the outcomes of deep tomographic imaging networks. With the complicated nonlinearity of networks and strong dependence on big data, classic image quality criteria such as least square errors and structural similarity metrics are of reduced value. Ultimately, task-based evaluation becomes the most relevant, especially the accuracy, robustness and generalizability in a context of interest. The current practice relies on empirical guidelines for training, validation and testing cycles[137]. Major efforts are needed to understand the learning mechanism and optimize the network architecture and diagnostic performance.

Explainable/interpretable AI and ethical AI. The expandability/interpretability of deep networks is of utmost importance for medical AI in general, and deep tomographic imaging in particular. A recent review on explainable AI[138] covers over 400 papers, including general machine learning techniques and challenges. Complementary to that review, another review[139] specifically targets over 100 papers on the interpretability of deep neural networks including quadratic neural networks. Despite extensive efforts in this area, the current status of research reminds us of the Indian fable 'Blind men and the elephant', in which the elephant is a metaphor of the governing theory of deep learning. Many valid/plausible arguments were made to explain deep networks, which touch different parts of the elephant, but the fundamental secrets of deep learning are still elusive, such as why the non-convex optimization process often yields good results, why the generalizability of deep networks is surprisingly good, and many more. Philosophically speaking, there are two approaches of reasoning: deduction and induction. Accordingly, we have two schools of AI. Deduction is a top-down approach, and corresponds to the rule-based expert systems that used to be popular with the initiative of the development of the fifth-generation computer. On the other hand, induction is bottom-up from data towards information or knowledge. Currently, the mainstream approach of deep learning is inductive or data-driven. However, it is widely recognized that 'strong AI' should be capable of being both data-driven and rule-based. Towards this goal, networks trained with big data could be transformed into knowledge graphs so that the unification of data-driven and rule-based learning can be facilitated. Hopefully, this will eventually lead to semantic tomographic imaging, which means that rich semantic information can be utilized for solving an inverse problem. Once upon a time, people directly worked out answers. Then, applied mathematicians, computer scientists and engineers developed algorithms to compute solutions. Now, data scientists develop algorithms to develop algorithms. Logically, scientists would someday develop algorithms that develop algorithms to develop algorithms, and maybe some of the algorithms could decide what algorithms are needed, how they are useful and when they could fail.

Few people are still unsure that AI-based imaging as a whole workflow from data acquisition via/around intermediate images to

final diagnosis ('rawdiomics', meaning deep mining of critical features from raw data) will be either dominating or indispensable in many applications. Over the past few years, the radiology societies in the United States, China, Europe and other countries all proposed guidelines on the ethical use of AI. A joint statement was published last year[140] to "promote wellbeing, minimize harm, and ensure that the benefits and harms are distributed among the possible stakeholders in a just manner". We advocate that the responsibility is on not only radiologists but also all others of relevance to develop explainable and ethical AI. As mentioned above, the instability issue of deep networks can be effectively addressed[136]. Any balance between sensitivity and specificity in an imaging task should reflect patients' preference. Hence, a humane model will be desirable to enable ethical decisions in the context of ethical conflicts. An interesting study was performed on autonomous vehicle dilemmas for moral decision-making[141]. Similar efforts are needed for AI-based deliberations in imaging tasks.

Conclusion. Looking ahead based on what we have discussed, important deep tomographic imaging developments in the next few years may focus on big data with enlarged size and improved quality for supervised, weakly supervised, self-supervised or un-supervised learning; AI-empowered POC systems for cost-effective, hybrid and autonomous imaging; stabilized networks for accuracy, robustness and regulatory compliance; as well as explainable and ethical AI for optimized imaging results and their clinical acceptance. While there are various hurdles, the main idea is to infuse increasingly more intelligence into machines, and we believe that a synergistic combination of data-driven and rule-based/semantic learning should be the way to go. Towards this holy grail, multidisciplinary investigation and collaboration remains exciting, especially coordinated efforts by academy, industry, hospitals and regulatory agents.

Deep tomographic reconstruction is a rapidly developing field but still in its infancy. For further technical details, the readers are referred to several technical reviews[6–9], a dedicated book[142], and the references cited therein. Given the above overview, insight and outlook, we believe that this field has great utilities and new surprises to offer in the next five to ten years. For example, deep-learning-based reconstruction may enable new types of imaging modalities based on physical phenomena that have not been utilized before. One of the main reasons for the current medical imaging modalities such as CT and MRI being widely used is that their imaging models are linear and allow relatively simple algorithms for image reconstruction. However, there are many exciting nonlinear physical phenomena with great potential of sensitive and specific imaging, which could not be previously explored due to the lack of efficient imaging algorithms. Given the power of deep learning in constructing a direct inverse mapping, deep neural networks may be exactly what we need to capitalize on these nonlinear physical mechanisms and open a new era of imaging research and translation. Finally, tomographic imaging is an integral part of modern medicine, deep-learning-based imaging capabilities, and AI-empowered other clinical and preclinical solutions are complementary and synergistic, and hopefully will be unified to transform the healthcare practice into personalized, preventive and precision medicine that is intelligent, inexpensive and indiscriminate.

Received: 22 April 2020; Accepted: 9 November 2020;
Published online: 10 December 2020

References

1. Barrett, H., Myers, K., Hoeschen, C., Kupinski, M. & Little, M. Task-based measures of image quality and their relation to radiation dose and patient risk. *Phys. Med. Biol.* **60**, R1 (2015).
2. Floyd, C. An artificial neural network for SPECT image reconstruction. *IEEE Trans. Med. Imaging* **10**, 485–487 (1991).
3. Yan, H. & Mao, J. Data truncation artifact reduction in MR imaging using a multilayer neural network. *IEEE Trans. Med. Imaging* **12**, 73–77 (1993).
4. Wang, G. A perspective on deep imaging. *IEEE Access* **4**, 8914–8924 (2016).
5. Wang, G., Ye, J. C., Mueller, K. & Fessler, J. A. Image reconstruction is a new frontier of machine learning. *IEEE Trans. Med. Imaging* **37**, 1289–1296 (2018).
6. McCann, M. T., Jin, K. H. & Unser, M. Convolutional neural networks for inverse problems in imaging: a review. *IEEE Signal Process. Mag.* **34**, 85–95 (2017).
7. Ravishankar, S., Ye, J. C. & Fessler, J. A. Image reconstruction: from sparsity to data-adaptive methods and machine learning. *Proc. IEEE* **108**, 86–109 (2019).
8. Sahiner, B. et al. Deep learning in medical imaging and radiation therapy. *Med. Phys.* **46**, e1–e36 (2019).
9. Gong, K., Berg, E., Cherry, S. R. & Qi, J. Machine learning in PET: from photon detection to quantitative image reconstruction. *Proc. IEEE* **108**, 51–68 (2019).
10. Nuyts, J., De Man, B., Fessler, J. A., Zbijewski, W. & Beekman, F. J. Modelling the physics in the iterative reconstruction for transmission computed tomography. *Phys. Med. Biol.* **58**, R63 (2013).
11. Hebert, T. & Leahy, R. A generalized EM algorithm for 3-D bayesian reconstruction from Poisson data using Gibbs priors. *IEEE Trans. Med. Imagine* **8**, 194–202 (1989).
12. Zaidi, H., Ruest, T., Schoenahl, F. & Montandon, M.-L. Comparative assessment of statistical brain MR image segmentation algorithms and their impact on partial volume correction in PET. *Neuroimage* **32**, 1591–1607 (2006).
13. Hounsfield, G. N. Computerized transverse axial scanning (tomography): part 1. Description of system. *Br. J. Radiol.* **46**, 1016–1022 (1973).
14. Shepp, L. A. & Logan, B. F. The Fourier reconstruction of a head section. *IEEE Trans. Nucl. Sci.* **21**, 21–43 (1974).
15. Kang, E., Min, J. & Ye, J. C. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. *Med. Phys.* **44**, e360–e375 (2017).
16. Chen, H. et al. Low-dose CT via convolutional neural network. *Biomed. Opt. Express* **8**, 679–694 (2017).
17. Shan, H. et al. Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction. *Nat. Mach. Intell.* **1**, 269–276 (2019).
18. Sodickson, D. K. & Manning, W. J. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. *Magn. Reson. Med.* **38**, 591–603 (1997).
19. Pruessmann, K. P., Weiger, M., Scheidegger, M. B. & Boesiger, P. Sense: sensitivity encoding for fast MRI. *Magn. Reson. Med.* **42**, 952–962 (1999).
20. Griswold, M. A. et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn. Reson. Med.* **47**, 1202–1210 (2002).
21. Lustig, M., Donoho, D. & Pauly, J. M. Sparse MRI: the application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.* **58**, 1182–1195 (2007).
22. Brada, R. et al. Towards motion-robust MRI – autonomous motion timing and correction during MR scanning using multi-coil data and a deep-learning neural network. In *Proc. ISMRM Twenty Seventh Annual Meeting* (ISMRM, 2019).
23. Hammernik, K. et al. Learning a variational network for reconstruction of accelerated MRI data. *Magn. Reson. Med.* **79**, 3055–3071 (2018).
24. Kwon, K., Kim, D. & Park, H. A parallel MR imaging method using multilayer perceptron. *Med. Phys.* **44**, 6209–6224 (2017).
25. Sun, J., Li, H. & Xu, Z. Deep ADMM-Net for compressive sensing MRI. In *Advances in Neural Information Processing Systems* 10–18 (NeurIPS, 2016).
26. Wang, S. et al. Accelerating magnetic resonance imaging via deep learning. In *2016 IEEE 13th Int. Symp. Biomedical Imaging* 514–517 (IEEE, 2016).
27. Defrise, M. et al. Exact and approximate rebinning algorithms for 3-D PET data. *IEEE Trans. Med. Imaging* **16**, 145–158 (1997).
28. Fink, M. Time reversal of ultrasonic fields. I. Basic principles. *IEEE Trans. Ultrason. Ferroelec. Freq. Control* **39**, 555–566 (1992).
29. Jin, K. H., McCann, M. T., Froustey, E. & Unser, M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* **26**, 4509–4522 (2017).
30. Lee, D., Yoo, J., Tak, S. & Ye, J. Deep residual learning for accelerated MRI using magnitude and phase networks. *IEEE Trans. Biomed. Eng.* **65**, 1985–1995 (2018).
31. Gong, K., Catana, C., Qi, J. & Li, Q. PET image reconstruction using deep image prior. *IEEE Trans. Med. Imaging* **38**, 1655–1665 (2018).
32. Qian, H., Rui, X. & Ahn, S. Deep learning models for PET scatter estimations. In *2017 IEEE Nuclear Science Symp. Medical Imaging Conf.* 1–5 (IEEE, 2017).
33. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy STORM. *Nat. Methods* **3**, 793–796 (2006).

34. Hess, S. T., Girirajan, T. P. & Mason, M. D. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys. J.* **91**, 4258 (2006).
35. Betzig, E. et al. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006).
36. Nehme, E., Weiss, L. E., Michaeli, T. & Shechtman, Y. DeepSTORM: super-resolution single-molecule microscopy by deep learning. *Optica* **5**, 458–464 (2018).
37. Rivenson, Y. et al. Deep learning microscopy. *Optica* **4**, 1437–1443 (2017).
38. Liu, T. et al. Deep learning-based super-resolution in coherent imaging systems. *Sci. Rep.* **9**, 3926 (2019).
39. Rivenson, Y., Zhang, Y., Günaydın, H., Teng, D. & Ozcan, A. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light Sci. Appl.* **7**, 17141 (2018).
40. Li, S., Deng, M., Lee, J., Sinha, A. & Barbastathis, G. Imaging through glass diffusers using densely connected convolutional networks. *Optica* **5**, 803–813 (2018).
41. Sun, Y., Xia, Z. & Kamilov, U. S. Efficient and accurate inversion of multiple scattering with deep learning. *Opt. Express* **26**, 14678–14688 (2018).
42. Lyu, M. et al. Deep-learning-based ghost imaging. *Sci. Rep.* **7**, 17865 (2017).
43. Rivenson, Y. et al. Deep learning enhanced mobile-phone microscopy. *ACS Photon.* **5**, 2354–2364 (2018).
44. Xu, M. & Wang, L. V. Photoacoustic imaging in biomedicine. *Rev. Sci. Instrum.* **77**, 041101 (2006).
45. Zhou, Z. et al. High spatial–temporal resolution reconstruction of plane-wave ultrasound images with a multichannel multiscale convolutional neural network. *IEEE Trans. Ultrason. Ferroelec. Freq. Control* **65**, 1983–1996 (2018).
46. Antholzer, S., Haltmeier, M. & Schwab, J. Deep learning for photoacoustic tomography from sparse data. *Inverse Prob. Sci. Eng.* **27**, 987–1005 (2019).
47. Davoudi, N., Deán-Ben, X. L. & Razansky, D. Deep learning optoacoustic tomography with sparse data. *Nat. Mach. Intell.* **1**, 453–460 (2019).
48. Wolterink, J. M., Leiner, T., Viergever, M. A. & Išgum, I. Generative adversarial networks for noise reduction in low-dose CT. *IEEE Trans. Med. Imaging* **36**, 2536–2545 (2017).
49. Yang, G. et al. DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. *IEEE Trans. Med. Imaging* **37**, 1310–1321 (2017).
50. Bubba, T. A. et al. Learning the invisible: a hybrid deep learning-shearlet framework for limited angle computed tomography. *Inverse Prob.* **35**, 064002 (2019).
51. Kutyniok, G. & Labate, D. *Shearlets: Multiscale Analysis for Multivariate Data* (Springer, 2012).
52. Zhang, Y. & Yu, H. Convolutional neural network based metal artifact reduction in X-ray computed tomography. *IEEE Trans. Med. Imaging* **37**, 1370–1381 (2018).
53. Schlemper, J., Caballero, J., Hajnal, J. V., Price, A. N. & Rueckert, D. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Trans. Med. Imaging* **37**, 491–503 (2018).
54. Vishnevskiy, V., Walheim, J. & Kozerke, S. Deep variational network for rapid 4D flow MRI reconstruction. *Nat. Mach. Intell.* **2**, 228–235 (2020).
55. Eo, T. et al. KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images. *Magn. Reson. Med.* **80**, 2188–2201 (2018).
56. Qin, C. et al. Convolutional recurrent neural networks for dynamic mr image reconstruction. *IEEE Trans. Med. Imaging* **38**, 280–290 (2018).
57. Vishnevskiy, V., Sanabria, S. J. & Goksel, O. Image reconstruction via variational network for real-time hand-held sound-speed imaging. In *International Workshop on Machine Learning for Medical Image Reconstruction* (eds Knoll, F. et al.) 120–128 (Springer, 2018).
58. Adler, J. & Öktem, O. Learned primal-dual reconstruction. *IEEE Trans. Med. Imaging* **37**, 1322–1332 (2018).
59. Gupta, H., Jin, K. H., Nguyen, H. Q., McCann, M. T. & Unser, M. CNN-based projected gradient descent for consistent CT image reconstruction. *IEEE Trans. Med. Imaging* **37**, 1440–1453 (2018).
60. Chen, H. et al. LEARN: Learned experts' assessment-based reconstruction network for sparse-data CT. *IEEE Trans. Med. Imaging* **37**, 1333–1347 (2018).
61. Cheng, L., Ahn, S., Ross, S. G., Qian, H. & De Man, B. Accelerated iterative image reconstruction using a deep learning based leapfrogging strategy. In *Int. Conf. Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine* 715–720 (Fully 3D, 2017).
62. Chun, I. Y., Huang, Z., Lim, H. & Fessler, J. Momentum-net: fast and convergent iterative neural network for inverse problems. *IEEE Trans. Pattern Anal. Mach. Intell.* https://doi.org/10.1109/TPAMI.2020.3012955 (2020).
63. Aggarwal, H. K., Mani, M. P. & Jacob, M. MoDL: model-based deep learning architecture for inverse problems. *IEEE Trans. Med. Imaging* **38**, 394–405 (2018).
64. Hauptmann, A. et al. Model-based learning for accelerated, limited-view 3-D photoacoustic tomography. *IEEE Trans. Med. Imaging* **37**, 1382–1393 (2018).
65. Biswas, S., Aggarwal, H. K. & Jacob, M. Dynamic MRI using model-based deep learning and SToRM priors: MoDL-SToRM. *Magn. Res. Med.* **82**, 485–494 (2019).
66. Lyu, Q., Shan, H., Xie, Y., Li, D. & Wang, G. Cine cardiac MRI motion artifact reduction using a recurrent neural network. Preprint at https://arxiv.org/abs/2006.12700 (2020).
67. Gong, K. et al. Iterative PET image reconstruction using convolutional neural network representation. *IEEE Trans. Med. Imaging* **38**, 675–685 (2018).
68. Wu, D., Kim, K., El Fakhri, G. & Li, Q. Iterative low-dose CT reconstruction with priors trained by artificial neural network. *IEEE Trans. Med. Imaging* **36**, 2479–2486 (2017).
69. Lim, H., Chun, I. Y., Fessler, J. & Dewaraja, Y. Improved low count quantitative SPECT reconstruction with a trained deep learning based regularizer. *J. Nucl. Med.* **60**, 42 (2019).
70. Gao, Y. et al. A feasibility study of extracting tissue textures from a previous full-dose CT database as prior knowledge for Bayesian reconstruction of current low-dose CT images. *IEEE Trans. Med. Imaging* **38**, 1981–1992 (2019).
71. Buzzard, G. T., Chan, S. H., Sreehari, S. & Bouman, C. A. Plug-and-play unplugged: optimization-free reconstruction using consensus equilibrium. *SIAM J. Imaging Sci.* **11**, 2001–2020 (2018).
72. Liu, J. et al. RARE: Image reconstruction using deep priors learned without ground truth. *IEEE J. Selected Topics Signal Process.* **14**, 1088–1099 (2020).
73. Wu, Z. et al. SIMBA: Scalable inversion in optical tomography using deep denoising priors. *IEEE J. Selected Topics Signal Process.* **14**, 1163–1175 (2020).
74. Ulyanov, D., Vedaldi, A. & Lempitsky, V. Deep image prior. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition* 9446–9454 (IEEE, 2018).
75. Jin, K. H., Gupta, H., Yerly, J., Stuber, M. & Unser, M. Time-dependent deep image prior for dynamic MRI. Preprint at https://arxiv.org/abs/1910.01684 (2019).
76. Zhu, B., Liu, J. Z., Cauley, S. F., Rosen, B. R. & Rosen, M. S. Image reconstruction by domain-transform manifold learning. *Nature* **555**, 487–492 (2018).
77. Ye, D. H., Buzzard, G. T., Ruby, M. & Bouman, C. A. Deep back projection for sparse-view CT reconstruction. In *2018 IEEE Global Conf. Signal and Information Processing* (IEEE, 2018).
78. Li, Y., Li, K., Zhang, C., Montoya, J. & Chen, G.-H. Learning to reconstruct computed tomography images directly from sinogram data under a variety of data acquisition conditions. *IEEE Trans. Med. Imaging* **38**, 2469–2481 (2019).
79. Fu, L. & De Man, B. A hierarchical approach to deep learning and its application to tomographic reconstruction. In *15th Int. Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine* 1107202 (International Society for Optics and Photonics, 2019).
80. Whiteley, W., Luk, W. K. & Gregor, J. DirectPET: full-size neural network PET reconstruction from sinogram data. *J. Med. Imaging* **7**, 032503 (2020).
81. Häggström, I., Schmidtlein, C. R., Campanella, G. & Fuchs, T. J. DeepPET: a deep encoder–decoder network for directly solving the PET image reconstruction inverse problem. *Med. Image Anal.* **54**, 253–262 (2019).
82. Würfl, T. et al. Deep learning computed tomography: learning projection-domain weights from image domain in limited angle problems. *IEEE Trans. Med. Imaging* **37**, 1454–1463 (2018).
83. Han, Y. & Ye, J. C. One network to solve all ROIs: deep learning CT for any ROI using differentiated backprojection. *Med. Phys.* **46**, e855–e872 (2019).
84. Han, Y., Kim, J. & Ye, J. C. Differentiated backprojection domain deep learning for conebeam artifact removal. *IEEE Trans. Med. Imaging* **39**, 3571–3582 (2020).
85. Claus, B. E., Jin, Y., Gjesteby, L. A., Wang, G. & De Man, B. Metal-artifact reduction using deep-learning based sinogram completion: initial results. In *Proc. 14th Int. Meeting Fully Three-Dimensional Image Reconstruction Radiology Nuclear Medicine* 631–634 (Fully 3D, 2017).
86. Ghani, M. U. & Karl, W. C. Fast enhanced CT metal artifact reduction using data domain deep learning. *IEEE Trans. Comput. Imaging* **6**, 181–193 (2019).
87. De Man, Q. et al. A two-dimensional feasibility study of deep learning-based feature detection and characterization directly from CT sinograms. *Med. Phys.* **46**, e790–e800 (2019).
88. Han, Y., Sunwoo, L. & Ye, J. C. k-space deep learning for accelerated MRI. *IEEE Trans. Med. Imaging* **39**, 377–386 (2019).
89. Lee, J., Han, Y., Ryu, J.-K., Park, J.-Y. & Ye, J. C. k-space deep learning for reference-free EPI ghost correction. *Magn. Reson. Med.* **82**, 2299–2313 (2019).
90. Akçakaya, M., Moeller, S., Weingärtner, S. & Uğurbil, K. Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: database-free deep learning for fast imaging. *Magn. Res. Med.* **81**, 439–453 (2019).

91. Pramanik, A., Aggarwal, H. & Jacob, M. Deep generalization of structured low-rank algorithms (Deep-SLR). *IEEE Trans. Med. Imaging* https://doi.org/10.1109/TMI.2020.3014581 (2020).
92. Luchies, A. C. & Byram, B. C. Deep neural networks for ultrasound beamforming. *IEEE Trans. Med. Imaging* **37**, 2010–2021 (2018).
93. Gasse, M. et al. High-quality plane wave compounding using convolutional neural networks. *IEEE Trans. Ultrason. Ferroelec. Freq. Control* **64**, 1637–1639 (2017).
94. Khan, S., Huh, J. & Ye, J. C. Adaptive and compressive beamforming using deep learning for medical ultrasound. *IEEE Trans. Ultrason. Ferroelec. Freq. Control* **67**, 1558–1572 (2020).
95. Luijten, B. et al. Deep learning for fast adaptive beamforming. In *2019 IEEE Int. Conf. Acoustics, Speech and Signal Processing* 1333–1337 (IEEE, 2019).
96. Yoon, Y. H., Khan, S., Huh, J. & Ye, J. C. Efficient B-mode ultrasound image reconstruction from sub-sampled RF data using deep learning. *IEEE Trans. Med. Imaging* **38**, 325–336 (2018).
97. Kang, E., Koo, H. J., Yang, D. H., Seo, J. B. & Ye, J. C. Cycle-consistent adversarial denoising network for multiphase coronary CT angiography. *Med. Phys.* **46**, 550–562 (2019).
98. Zhu, J.-Y., Park, T., Isola, P. & Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proc. IEEE Int. Conf. Computer Vision* 2223–2232 (IEEE, 2017).
99. Nguyen, T., Xue, Y., Li, Y., Tian, L. & Nehmetallah, G. Deep learning approach for Fourier ptychography microscopy. *Opt. Express* **26**, 26470–26484 (2018).
100. Lee, S., Han, S., Salama, P., Dunn, K. W. & Delp, E. J. Three dimensional blind image deconvolution for fluorescence microscopy using generative adversarial networks. In *2019 IEEE 16th Int. Symp. Biomedical Imaging* 538–542 (IEEE, 2019).
101. Villani, C. *Optimal Transport: Old and New* Vol. 338 (Springer, 2008).
102. Lim, S., Lee, S.-E., Chang, S. & Ye, J. C. CycleGAN with a blur kernel for deconvolution microscopy: Optimal transport geometry. *IEEE Trans. Comput. Imaging* **6**, 1127–1138 (2020).
103. Li, M., Shan, H., Pryshchep, S., Lopez, M. M. & Wang, G. Deep adversarial network for super stimulated emission depletion imaging. *J. Nanophoton.* **14**, 016009 (2020).
104. Liu, F., Jang, H., Kijowski, R., Bradshaw, T. & McMillan, A. B. Deep learning MR imaging–based attenuation correction for PET/MR imaging. *Radiology* **286**, 676–684 (2018).
105. Hwang, D. et al. Improving the accuracy of simultaneously reconstructed activity and attenuation maps using deep learning. *J. Nucl. Med.* **59**, 1624–1629 (2018).
106. Dar, S. U. et al. Image synthesis in multi-contrast MRI with conditional generative adversarial networks. *IEEE Trans. Med. Imaging* **38**, 2375–2388 (2019).
107. Lee, D., Moon, W.-J. & Ye, J. C. Assessing the importance of magnetic resonance contrasts using collaborative generative adversarial networks. *Nat. Mach. Intell.* **2**, 34–42 (2020).
108. Lehtinen, J. et al. Noise2Noise: learning image restoration without clean data. In *Int. Conf. Machine Learning* 2965–2974 (PMLR, 2018).
109. Yaman, B. et al. Self-supervised physics-based deep learning MRI reconstruction without fully-sampled data. In *2020 IEEE 17th Int. Symp. Biomedical Imaging* 921–925 (IEEE, 2020).
110. Wu, D., Ren, H. & Li, Q. Self-supervised dynamic CT perfusion image denoising with deep neural networks. *IEEE Trans. Radiation Plasma Med. Sci.* https://doi.org/10.1109/TRPMS.2020.2996566 (2020).
111. Wu, D., Gong, K., Kim, K., Li, X. & Li, Q. Consensus neural network for medical imaging denoising with only noisy training samples. In *Int. Conf. Medical Image Computing and Computer-Assisted Intervention* 741–749 (Springer, 2019).
112. Donoho, D. L. Compressed sensing. *IEEE Trans. Inform. Theory* **52**, 1289–1306 (2006).
113. Ye, J. C., Han, Y. & Cha, E. Deep convolutional framelets: a general deep learning framework for inverse problems. *SIAM J. Imaging Sci.* **11**, 991–1048 (2018).
114. Ye, J. C. & Sung, W. K. Understanding geometry of encoder-decoder CNNs. In *Proc. 36th Int. Conf. Machine Learning* Vol. 97 (eds Chaudhuri, K. & Salakhutdinov, R.) 7064–7073 (PMLR, 2019).
115. *NIH Image Database* (NIH, 2020); https://www.nihlibrary.nih.gov/resources/subject-guides/web-search-thinking-beyond-google/images
116. McCollough, C. H. 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.* **44**, e339–e352 (2017).
117. Gaillard, F. & Murphy, A. Imaging data sets (artificial intelligence). *Radiopaedia* https://radiopaedia.org/articles/imaging-data-sets-artificial-intelligence?lang=us (2020).
118. Xu, X. G., Taranenko, V., Zhang, J. & Shi, C. A boundary-representation method for designing whole-body radiation dosimetry models: pregnant females at the ends of three gestational periods—RPI-P3, -P6 and -P9. *Phys. Med. Biol.* **52**, 7023 (2007).
119. Segars, W. P., Sturgeon, G., Mendonca, S., Grimes, J. & Tsui, B. M. 4D XCAT phantom for multimodality imaging research. *Med. Phys.* **37**, 4902–4915 (2010).
120. Krishna, A. & Mueller, K. Medical CT image generation with style. In *15th Int. Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine* 1107234 (International Society for Optics and Photonics, 2019).
121. De Man, B. et al. CatSim: a new computer assisted tomography simulation environment. In *Medical Imaging 2007: Physics of Medical Imaging* Vol. 6510 65102G (International Society for Optics and Photonics, 2007).
122. Jia, X., Yan, H., Cerviño, L., Folkerts, M. & Jiang, S. B. A GPU tool for efficient, accurate, and realistic simulation of cone beam CT projections. *Med. Phys.* **39**, 7368–7378 (2012).
123. Abadi, E. et al. DukeSim: a realistic, rapid, and scanner-specific simulation framework in computed tomography. *IEEE Trans. Med. Imaging* **38**, 1457–1465 (2018).
124. De Man, B. et al. XCIST (X-ray-based Cancer Imaging Simulation Toolkit): a new open-source project (ITCR grant). In *Institute for Mathematics and its Applications Workshop on Computational Imaging* (IMA, 2019).
125. Sharma, D. et al. *In silico* imaging tools from the VICTRE clinical trial. *Med. Phys.* **46**, 3924–3928 (2019).
126. Parisi, G. I., Kemker, R., Part, J. L., Kanan, C. & Wermter, S. Continual lifelong learning with neural networks: a review. *Neural Netw.* **113**, 54–71 (2019).
127. Konečný, J. et al. Federated learning: Strategies for improving communication efficiency. Preprint at https://arxiv.org/abs/1610.05492 (2016).
128. McMahan, H. B. et al. Communication-efficient learning of deep networks from decentralized data. Preprint at https://arxiv.org/abs/1602.05629 (2016).
129. Krull, A., Buchholz, T.-O. & Jug, F. Noise2Void-learning denoising from single noisy images. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition* 2129–2137 (IEEE, 2019.
130. Liao, H., Lin, W.-A., Zhou, S. K. & Luo, J. ADN: artifact disentanglement network for unsupervised metal artifact reduction. *IEEE Trans. Med. Imaging* **39**, 634–643 (2019).
131. Locatello, F. et al. Challenging common assumptions in the unsupervised learning of disentangled representations. In *Int. Conf. Machine Learning* 4114–4124 (PMLR, 2019).
132. Gjesteby, L., Xi, Y., Kalra, M. K., Yang, Q. & Wang, G. Hybrid imaging system for simultaneous spiral MR and X-ray (MRX) scans. *IEEE Access* **5**, 1050–1061 (2016).
133. Dineley, J. Tackling the silent crisis in cancer care. *Lindau Nobel Laureate Meetings* https://www.lindau-nobel.org/blog-tackling-the-silent-crisis-in-cancer-care-with-innovation/ (2018).
134. Antun, V., Renna, F., Poon, C., Adcock, B. & Hansen, A. C. On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proc. Natl Acad. Sci. USA* https://doi.org/10.1073/pnas.1907377117 (2020).
135. Ozdag, M. Adversarial attacks and defenses against deep neural networks: a survey. *Procedia Comp. Sci.* **140**, 152–161 (2018).
136. Wu, W. et al. Stabilizing deep tomographic reconstruction networks. Preprint at https://arxiv.org/abs/2008.01846 (2020).
137. Park, S. H. & Han, K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. *Radiology* **286**, 800–809 (2018).
138. Arrieta, A. B. et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Inform. Fusion* **58**, 82–115 (2020).
139. Fan, F., Xiong, J. & Wang, G. On interpretability of artificial neural networks. Preprint at https://arxiv.org/abs/2001.02522 (2020).
140. Geis, J. R. et al. Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement. *Radiology* **293**, 436–440 (2019).
141. Awad, E. et al. The Moral Machine experiment. *Nature* **563**, 59–64 (2018).
142. Wang, G., Zhang, Y., Ye, X. & Mou, X. *Machine Learning for Tomographic Imaging* (IOP, 2019).
143. Poirot, M. G. et al. Physics-informed deep learning for dual-energy computed tomography image processing. *Sci. Rep.* **9**, 17709 (2019).
144. Kang, E., Chang, W., Yoo, J. & Ye, J. C. Deep convolutional framelet denosing for low-dose CT via wavelet residual network. *IEEE Trans. Med. Imaging* **37**, 1358–1369 (2018).
145. Ouyang, J., Chen, K. T., Gong, E., Pauly, J., Zaharchuk, G. & Ultra-low-dose, P. E. T. reconstruction using generative adversarial network with feature matching and task-specific perceptual loss. *Med. Phys.* **46**, 3555–3564 (2019).

146. Chen, K. T. et al. Ultra-low-dose 18F-florbetaben amyloid PET imaging using deep learning with multi-contrast MRI inputs. *Radiology* **290**, 649–656 (2019).

Competing interests
The authors declare no competing interests.

Additional information
Correspondence should be addressed to G.W.