Check for updates

Deep learning for biomedical image reconstruction: a survey

Hanene Ben Yedder[1] · Ben Cardoen[1] · Ghassan Hamarneh[1]

Abstract

Medical imaging is an invaluable resource in medicine as it enables to peer inside the human body and provides scientists and physicians with a wealth of information indispensable for understanding, modelling, diagnosis, and treatment of diseases. Reconstruction algorithms entail transforming signals collected by acquisition hardware into interpretable images. Reconstruction is a challenging task given the ill-posedness of the problem and the absence of exact analytic inverse transforms in practical cases. While the last decades witnessed impressive advancements in terms of new modalities, improved temporal and spatial resolution, reduced cost, and wider applicability, several improvements can still be envisioned such as reducing acquisition and reconstruction time to reduce patient's exposure to radiation and discomfort while increasing clinics throughput and reconstruction accuracy. Furthermore, the deployment of biomedical imaging in handheld devices with small power requires a fine balance between accuracy and latency. The design of fast, robust, and accurate reconstruction algorithms is a desirable, yet challenging, research goal. While the classical image reconstruction algorithms approximate the inverse function relying on expert-tuned parameters to ensure reconstruction performance, deep learning (DL) allows automatic feature extraction and real-time inference. Hence, DL presents a promising approach to image reconstruction with artifact reduction and reconstruction speed-up reported in recent works as part of a rapidly growing field. We review state-of-the-art image reconstruction algorithms with a focus on DL-based methods. First, we examine common reconstruction algorithm designs, applied metrics, and datasets used in the literature. Then, key challenges are discussed as potentially promising strategic directions for future research.

Keywords Image reconstruction · Modality · Deep learning · Inverse problem · Analytical approach · Iterative approach · Limited data representation

✉ Hanene Ben Yedder
  hbenyedd@sfu.ca

  Ben Cardoen
  bcardoen@sfu.ca

  Ghassan Hamarneh
  hamarneh@sfu.ca

[1] School of Computing Science, Simon Fraser University, Burnaby, Canada

1 Introduction

Biomedical image reconstruction translates signals acquired by a wide range of sensors into images that can be used for diagnosis and discovery of biological processes in cell and organ tissue. Each biomedical imaging modality leverages signals in different bands of the electromagnetic spectrum, e.g. from gamma rays (Positron emission tomography PET/SPECT)), X-rays (computed tomography (CT)), visible light (microscopy, endoscopy), infrared (thermal images), and radio-frequency (Nuclear magnetic resonance imaging (MRI)), as well as pressure sound waves (in ultrasound (US) imaging) (Webb and Kagadis 2003). Reconstruction algorithms transform the collected signals into a 2, 3, or 4-dimensional image.

The accuracy of each reconstruction is critical for discovery and diagnosis. Robustness to noise and generalization cross modality specifications' (e.g., sampling pattern, rate, etc.) and imaging devices parameters' allow a reconstruction algorithm to be used in wider applications. The time required for each reconstruction determines the number of subjects that can be diagnosed as well as the suitability of the technique in operating theatres and emergency situations. The number of measurements needed for a high quality reconstruction impacts the exposure a patient or sample will have to endure. Finally, the hardware requirements define whether a reconstruction algorithm can be used only in a dedicated facility or in portable devices thus dictating the flexibility of deployment.

The study of image reconstruction is an active area of research in modern applied mathematics, engineering and computer science. It forms one of the most active interdisciplinary fields of science (Fessler 2017) given that improvement in the quality of reconstructed images offers scientists and clinicians an unprecedented insight into the biological processes underlying disease. Figure 1 provides an illustration of the reconstruction problem and shows a typical data flow in a medical imaging system.

Over the past few years, researchers have begun to apply machine learning techniques to biomedical reconstruction to enable real-time inference and improved image quality in a clinical setting. Here, we first provide an overview of the image reconstruction problem and outline its characteristics and challenges (Sect. 1.1) and then outline the purpose, scope, and the layout of this review (Sect. 1.2).
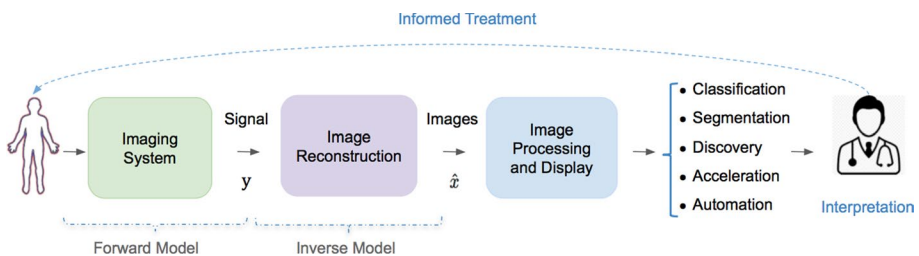

Fig. 1 Data flow in a medical imaging and image interpretation system. Forward model encodes the physics of the imaging system. The inverse model transforms the collected signals by the acquisition hardware into a meaningful image. The success of diagnosis, evaluation, and treatment rely on accurate reconstruction, image visualization and processing algorithms

1.1 Inverse problem and challenges

1.1.1 From output to input

Image reconstruction is the process of forming an interpretable image from the raw data (signals) collected by the imaging device. It is known as an inverse problem where given a set of measurements, the goal is to determine the original structure influencing the signal collected by a receiver given some signal transmission medium properties (Fig. 2). Let $y \in \mathbb{C}^M$ represent a set of raw acquired sensor measurements subject to some unknown noise (or perturbation) vector $\mathcal{N} \in \mathbb{C}^M$ intrinsic to the acquisition process. The objective is to recover the spatial-domain (or spatio-temporal) unknown image $x \in \mathbb{C}^N$ such that:

$$y = \mathcal{A}(\mathcal{F}(x), \mathcal{N}) \tag{1}$$

where $\mathcal{F}(\cdot)$ is the forward model operator that models the physics of image-formation, which can include signal propagation, attenuation, scattering, reflection and other transforms, e.g. Radon or Fourier transform. $\mathcal{F}(\cdot)$ can be a linear or a nonlinear operator depending on the imaging modality. $\mathcal{A}$ is an aggregation operation representing the interaction between noise and signal, in the assumption of additive noise $\mathcal{A} = +$.

While imaging systems are usually well approximated using mathematical models that define the forward model, an exact analytic inverse transform $\mathcal{A}^{-1}(\cdot)$ is not always possible.
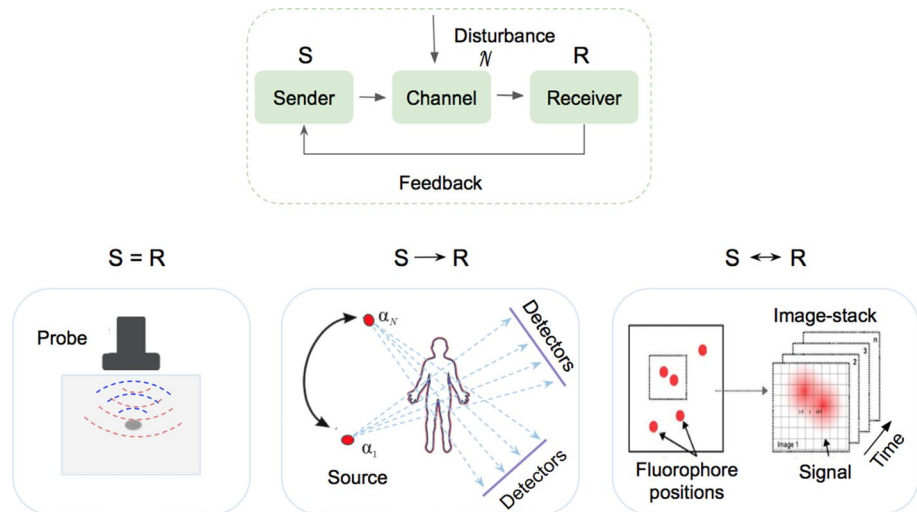

Fig. 2 Propagation of signals from sender to receiver. While passing through a transmission channel, signals s pick up noise (assuming additive) along the way, until the acquired sensor data $y = s + \mathcal{N}$ reaches the receiver. The properties of the received signal may, via a feedback loop, affect properties of future signal transmission. Sender and receiver modeling differ within modalities. For example the lower figure illustrates in the left ultrasound probe used to send and collect signals (S = R); in the middle: X-ray signal propagates through the subject toward the detector (S → R); Single molecule localization microscopy is an instance where S and R are components of a feedback loop (S ↔ R). The signal, optical measurements of nano-meter precise fluorescent emissions, is reconstructed in a discrete (temporal) X/Y/Z domain. An example of a feedback loop in SMLM is the auto-tuning of the laser power in response to an increase in density (R) that can compromise the spatio-temporal sparsity required for accurate reconstruction (Cardoen et al. 2019)

Reconstruction approaches usually resort to iteratively approximate the inverse function and often involve expert-tuned parameters and prior domain knowledge considerations to optimize reconstruction performance.

1.1.2 An ill-posed problem

A basic linear and finite-dimensional model of the sampling process leads us to study a discrete linear system of the imaging problem of the form:

$$y = Fx + \mathcal{N}. \tag{2}$$

For full sampling, in MRI for instance, we have M = N and the matrix $F \in \mathbb{C}^{M \times N}$ is a well-conditioned discrete Fourier transform matrix. However, when $(M < N)$, there will be many images x that map to the same measurements y, making the inverse an ill-posed problem (Fig. 3). Mathematically, the problem is typically under-determined as there would be fewer equations to describe the model than unknowns. Thus, one challenge for the reconstruction algorithm is to select the best solution among a set of potential solutions (McCann and Unser 2019). One way to reduce the solution space is to leverage domain specific knowledge by encoding priors, i.e. regularization.

Sub-sampling in MRI or sparse-view/limited-angle in CT are examples of how reducing data representation $(M < N)$ can accelerate acquisition or reduce radiation exposure. Additional gains can be found in lowered power, cost, time, and device complexity, albeit at the cost of increasing the degree of ill-posedness and complexity of the reconstruction problem. This brings up the need for sophisticated reconstruction algorithms with high feature extraction power to make the best use of the collected signal, capture modality-specific imaging features, and leverage prior knowledge. Furthermore, developing high-quality
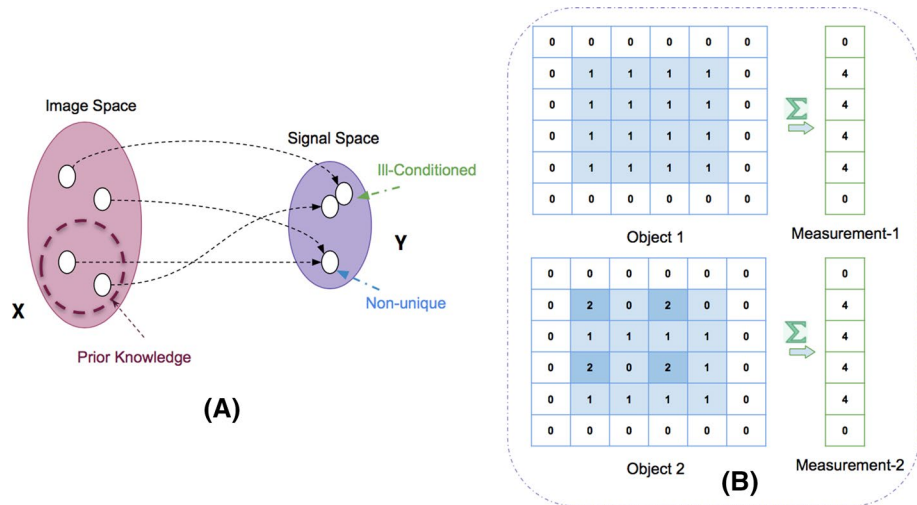

Fig. 3 (**A**) A problem is ill-conditioned when two different objects produce very close observed signals. When the observed signals are identical and hence identical reconstructed images, the inverse solution is non-unique. Prior knowledge can be leveraged to rule out certain solutions that conflict with the additional knowledge about the object beyond the measurement vectors. (**B**) A use case toy example of two objects with the same acquired signals. Prior knowledge about homogeneity of object rules out the second object

reconstruction algorithms requires not only a deep understanding of both the physics of the imaging systems and the biomedical structures but also specially designed algorithms that can account for the statistical properties of the measurements and tolerate errors in the measured data.

1.2 Scope of this survey

The field of biomedical image reconstruction has undergone significant advances over the past few decades and can be broadly classified into two categories: conventional methods (analytical and optimization based methods) and data-driven or learning-based methods. Conventional methods (discussed in Sect. 2) are the most dominant and have been extensively studied over the last few decades with a focus on how to improve their results (Vandenberghe et al. 2001; Jhamb et al. 2015; Assili 2018) and reduce their computational cost (Despres and Jia 2017).

Researchers have recently investigated deep learning (DL) approaches for various biomedical image reconstruction problems (discussed in Sect. 3) inspired by the success of deep learning in computer vision problems and medical image analysis (Goceri and Goceri 2017). This topic is relatively new and has gained a lot of interest over the last few years, as shown in Fig. 4a and listed in Table 1, and forms a very active field of research with numerous special journal issues (Wang 2016; Wang et al. 2018; Ravishankar et al. 2019). MRI and CT received the most attention among studied modalities, as illustrated in Fig. 4b, given their widespread clinical usage, the availability of analytic inverse transform and the availability of public (real) datasets. Several special issues journals devoted to MRI reconstruction have been recently published (Knoll et al. 2020; Liang et al. 2020; Sandino et al. 2020).

Fessler (2017) wrote a brief chronological overview on image reconstruction methods highlighting an evolution from basic analytical to data-driven models. McCann et al. (2017)) summarized works using CNNs for inverse problems in imaging. Later, Lucas et al. (2018) provided an overview of the ways deep neural networks may be employed to solve inverse problems in imaging (e.g., denoising, superresolution, inpainting). As illustrated in Fig. 4 since their publications a great deal of work has been done warranting a review. Recently, McCann and Unser (2019) wrote a survey on image reconstruction approaches
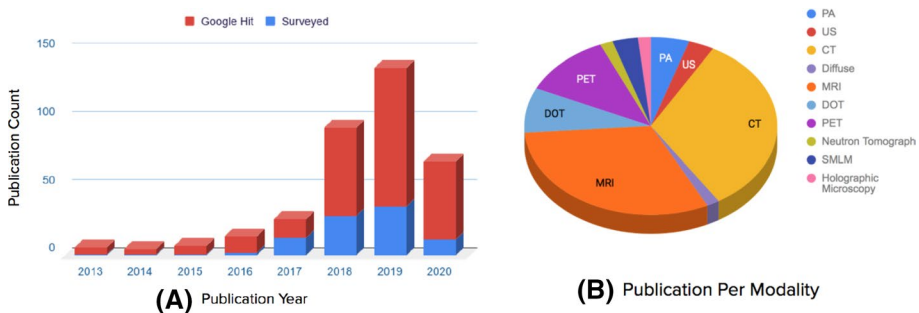

Fig. 4 (**A**) The marked increase in publications on biomedical image reconstruction and deep learning in the past 10 years. In red: results obtained by PubMed query that can be found at: https://bit.ly/recon_hit. Search query identified 310 contributions which were filtered according to quality, venue, viability and idea novelty, resulting in a total of 95 contributions as representatives covered in this survey (blue). (**B**) The pie chart represents the frequency of studies per modality covered in this survey

Table 1 Comparison between different papers that have applied machine learning to the biomedical image reconstruction problem. Mod.: modality; Samp.: signal sampling (F: full sampling, SS: sub-sampling, LA: limited angle data, SV: sparse view, LD: low dose data); TA: consideration of temporal aspect; D: dimension (PC: point clouds); Site: type of tissue (soft tissue like brain and breast, mix: different part of the body, Abdom: abdominal area, Bio Samp: biological samples, Sim: simulated data); Arch: architecture model (we denote by CNN(s) models that use handcrafted convolutional architectures); E2E: whether the approach is end-to-end or combined (Pos: DL as post-processing, Pre: DL as pre-processing, R2T: Raw-to-task); Loss: used loss function (Comb.loss: combined loss, Perc: perceptual loss, Adver: adversarial loss, CE: cross entropy, Jacc: Jaccard loss, MMD: maximum mean discrepancy); Reg: regularization; Inp: input data (M.Spec Meas: multi-spectral measurements, RF: radio frequencies measurements, PA Meas: photo-acoustic measurements); Out: output data (Chrom Maps: chromophore concentration map, Ref index: refractive index, Img: intensity image, Seg: segmentation mask); Metrics: evaluation metrics (M: MSE or RMSE, P: PSNR, S: SSIM, D: lesion detection metric like Dice and Jaccard, C: contrast to noise ratio and IT: reconstruction inference speed) Aug: used augmentation technique (A: affine, E: elastic, R: rotation transformations); Data: used training data (Sim: simulated data, Clini: clinical data, Phan: phantom data) Pub. Data: public data used

References	Mod.	Year	Samp.	MS	TA	D	Site	Arch.	E2E	Loss	Reg.	Inp.	Out.	Aug.	Metrics M|P|S|D|C|IT	Data	Size	Pub. Data
Cai et al.	PA	2018	F	✓		2D	Mix	ResU-Net	✓	MSE		M.Spec Meas	Chrom Maps		✓|✓|-|-|-|✓	Sim	~2000	LowDoseCT
Yoon et al.	US	2019	SS			3D	Abdom	Decoder	✓	MSE	L2	RF	Img		-|✓|-|-|-|✓	Clini	500	
Wu et al.	US	2018	F		✓	2D	Mix	CNN	✓	MAE		RF	Elast dist		✓|✓|-|-|✓|-	Sim Phan Clini	~100	
Gupta et al.	CT	2018	SV			2D	Mix	ResU-Net	Pos	Comb. loss		FBP	Img		-|✓|✓|-|-|-|-	Clini	~1000	LowDoseCT
Sun et al.	Mult. Scat	2018	F			2D	Bio Samp.	ResU-Net	Pos	MSE		FBP	Ref index		-|✓|-|-|-|-	Sim Phan	1550	Fresnel
Zhu et al.	MRI	2018	SS			2D	Brain	Decoder	✓	MSE	L2	k-space	Img	A	✓|✓|-|-|-|-	Imag-Net Clini	~10K	MGH-USCHCP
Chen et al.	CT	2017	LD			2D	Mix	Residual AE	Pos	MSE		FBP	Img	A+R	✓|✓|✓|✓|-|-	Sim Clini	~10K	LowDoseCT
Würfl et al.	CT	2018	LA			3D	Mix	Unrolling Iterative	✓	MSE		Sinogram	Img		-|✓|✓|-|-|-	Clini	~10K	LowDoseCT
Feng et al.	DOT	2019	F			2D	Soft	MLP	✓	MSE		Scatter-ing	Chrom Maps		✓|✓|✓|-|-|-	Sim	~20K	
Yoo et al.	DOT	2019	F			3D	Mix	Decoder	✓	MSE		Scatter-ing	Chrom Maps		-|-|-|-|-|-	Sim Clini	~1500	
Cui et al.	PET	2019	F			3D	Mix	3DU-Net	Pos	MSE		Sinogram	Img		-|-|-|✓|-|-	Sim Clini	30	

Table 1 (continued)

References	Mod.	Year	Samp.	MS	TA	D	Site	Arch.	E2E	Loss	Reg.	Inp.	Out.	Aug.	Metrics MIPISIDICIIS	Data	Size	Pub. Data
Hyun et al.	MRI	2018	SS			2D	Brain	U-Net	Pre	MSE		k-space	Full k-space		✓|-|-|✓ -|-|-	Clini	1400	
Waibel et al.	PA	2018	LV			2D	Vessels	U-Net	✓	MAE		PA Meas	Maps	R	✓|-|-|-|- -|-	Sim	3600	
Cheng et al.	MRI	2018	SS			2D	Soft	CNN	Pre	MSE		k-space	Img		✓|✓|✓|✓|-| -|✓	Clini	65K	
Liang et al.	CT	2018	SV			2D	Mix	ResNet	Pre	MSE		Sparse Sinogram	Sinog.		✓|-|-|-|-|-	Clini	17K	giveascan
Jin et al.	CT	2017	SV			2D	Mix	U-Net	Pos	MSE		FBP	Img	A	-|✓|-|-|-|-	Sim Clini	1000	
Wolterink et al.	CT	2017	LD			2D	Thorax cardiac	GAN	Pos	Adver	L2	FBP	Img		✓|-|-|-| -|✓	Sim Clini	500	
Yang et al.	MRI	2017	SS			2D	Brain	Condi GAN.	Pos	Adver Perc		k-space	Img	A+E	✓|✓|-|-|- -|✓	Clini	21K	MICCAI2013
Qin et al.	MRI	2019	SS	✓		3D+t	Cardiac	RCNN	✓	MSE		k-space	Img	A+E +R	✓|✓|✓|-|- -|✓	Sim Clini	NA	
Sun et al.	MRI	2016	SS			2D	Brain Chest	Unrolling ADMM	✓	NMSE		k-space	Img		✓|✓|-|-|- -|✓	Clini	150	MICCAI2013
Schlemper et al.	MRI	2018	SS		✓	2D	Cardiac	CNNs	✓	MSE		k-space	Img	A+E	✓|✓|-|-|- -|✓	Sim	300	
Nehme et al.	SMLM	2018				2D	Sim	Encoder Decoder	✓	MSE	L1	Fluo-phore emission	SR Img		✓|-|-|-|-|✓	Sim	100K	SMLM
Schlemper et al.	MRI	2017	SS			2D	Cardiac	CNNs	✓	MSE		k-space	Img	A	✓|-|-|-|-|✓	Sim	300	
Boublil et al.	CT	2015	LD			2D	Mix	MLP	Pos	MSE		FBP	img		✓|✓|✓|✓| -|-	Clini	100K	
Thaler et al.	CT	2018	SV			2D	Mix	wGAN	✓	Adver		LD Sinogram	Img	R+A	✓|-|-|✓ -|-|-	Clini	NA	

Table 1 (continued)

References	Mod.	Year	Samp.	MS	TA	D	Site	Arch.	E2E	Loss	Reg.	Inp.	Out.	Aug.	Metrics MIP|SI|DI|CI|IS	Data	Size	Pub. Data
Zhou et al.	CT	2019	LA			3D	Mix	GAN	✓	MSE Adver.	Cor coef	LD Sinogram FBP	Img		✓|-|-|✓|-|-✓	Clini	5436	LowDoseCT
Yedder et al.	DOT	2018	LA			2D	Breast	Decoder	✓	MSE	L2	Scatter-ing	Chrom Maps		✓|✓|✓|✓|-|-✓	Phan	4000	
Wang et al.	MRI	2016	SS			2D	Brain	Decoder	Pos	MSE		IFT	Img		-|-|-|-|-|-	Clini	500	
Sun et al.	MRI	2019	SS			2D	Brain	Enc-Dec	R2T	MSE CE		IFT	Img seg	A+E	✓|✓|-|✓|-|-	Clini	192	MRBrainS
Yedder et al.	DOT	2019	LA			2D	Breast	Decoder	✓	MSE Jacc.	L2	Scatter-ing	Chrom Maps		✓|✓|✓|✓|-|-✓	Phan	20K	
Boyd et al.	SMLM	2018	SS			PC	NA	CNN	✓	MMD		Fluo-phore emission	3D Localiz	A	✓|-|-|✓|-|✓|-	Sim	100K	SMLM
Häggström et al.	PET	2019	F			2D	Sim	Encoder Decoder	✓	MSE		Sinogram	Img	A+R	✓|✓|✓|-|-|-✓	Sim	291K	
Oksuz et al.	MRI	2019	SS		✓	2D+t	Cardiac	3D CNN RCNN	✓	MSE CE		k-space	Img		✓|✓|✓|✓|-|-|-	Sim Clini	300	mridata.org
Rajagopal et al.	CT	2019	SS			3D	Phan	Unrolling Iterative	✓	MSE		Sinogram	Img		✓|-|-|-|-|-|-	Phan	50K	
Huang et al.	CT	2019	LA			2D	abdomen	U-Net Iterative	✓	MSE		FBP	Artifact Img		✓|-|-|-|-|-|-	Clini	500	
Wang et al.	DOT	2019	LA			2D	Phan	Stacked AE	✓	MSE		Scatter-ing	Chrom Maps		-|-|-|✓|-|-	Sim	NA	

where they presented a toolbox of operators that can be used to build imaging systems models and showed how a forward model and sparsity-based regularization can be used to solve reconstruction problems. While their review is more focused on the mathematical foundations of conventional methods, they briefly discussed data-driven approaches, their theoretical underpinning, and performance. Similarly, Arridge et al. (2019) gave a comprehensive survey of algorithms that combine model and data-driven approaches for solving inverse problems with focus on deep neural based techniques and pave the way towards providing a solid mathematical theory. Zhang and Dong (2019) provided a conceptual review of some recent DL-based methods for CT with a focus on methods inspired by iterative optimization approaches and their theoretical underpinning from the perspective of representation learning and differential equations.

This survey provides an overview of biomedical image reconstruction methods with a focus on DL-based strategies, discusses their different paradigms (e.g., image domain, sensor domain (raw data) or end to end learning, architecture, loss, etc.) and how such methods can help overcome the weaknesses of conventional non-learning methods. To summarize the research done to date, we provide a comparative, structured summary of works representative of the variety of paradigms on this topic, tabulated according to different criteria, discusses the pros and cons of each paradigm as well as common evaluation metrics and training dataset challenges. The theoretical foundation was not emphasized in this work as it was comprehensively covered in the aforementioned surveys. A summary of the current state of the art and outline of what we believe are strategic directions for future research are discussed.

The remainder of this paper is organized as follows: in Sect. 2 we give an overview of conventional methods discussing their advantages and limitations. We then introduce the key machine learning paradigms and how they are being adapted in this field complementing and improving on conventional methods. A review of available data-sets and performance metrics is detailed in Sect. 3. Finally, we conclude by summarizing the current state-of-the-art and outlining strategic future research directions (Sect. 6).

2 Conventional image reconstruction approaches

A wide variety of reconstruction algorithms have been proposed during the past few decades, having evolved from analytical methods to iterative or optimization-based methods that account for the statistical properties of the measurements and noise as well as the hardware of the imaging system (Fessler 2017). While these methods have resulted in significant improvements in reconstruction accuracy and artifact reduction, are in routine clinical use currently, they still present some weaknesses. A brief overview of these methods' principles is presented in this section outlining their weaknesses.

2.1 Analytical methods

Analytical methods are based on a continuous representation of the reconstruction problem and use simple mathematical models for the imaging system. Classical examples are the inverse of the Radon transform such as filtered back-projection (FBP) for CT and the inverse Fourier transform (IFT) for MRI. These methods are usually computationally inexpensive (in the order of ms) and can generate good image quality in the absence of noise and under the assumption of full sampling/all angles projection (McCann and Unser 2019).

They typically consider only the geometry and sampling properties of the imaging system while ignoring the details of the system physics and measurement noise (Fessler 2017).

When dealing with noisy or incomplete measured data e.g., reducing the measurement sampling rate, analytical methods results deteriorate quickly as the signal becomes weaker. Thus, the quality of the produced image is compromised. Thaler et al. (2018) provided examples of a CT image reconstruction using the FBP method for different limited projection angles and showed that analytical methods are unable to recover the loss in the signal (Fig. 5), resulting in compromised diagnostic performance.

2.2 Iterative methods

Iterative reconstruction methods, based on more sophisticated models for the imaging system's physics, sensors and noise statistics, have attracted a lot of attention over the past few decades. They combine the statistical properties of data in the sensor domain (raw measurement data), prior information in the image domain, and sometimes parameters of the imaging system into their objective function (McCann and Unser 2019). Compared to analytical methods iterative reconstruction algorithms offer a more flexible reconstruction framework and better robustness to noise and incomplete data representation problems at the cost of increased computation (Ravishankar et al. 2019).

Iterative reconstruction methods involve minimizing an objective function that usually consists of a data term and a regularization terms imposing some prior:
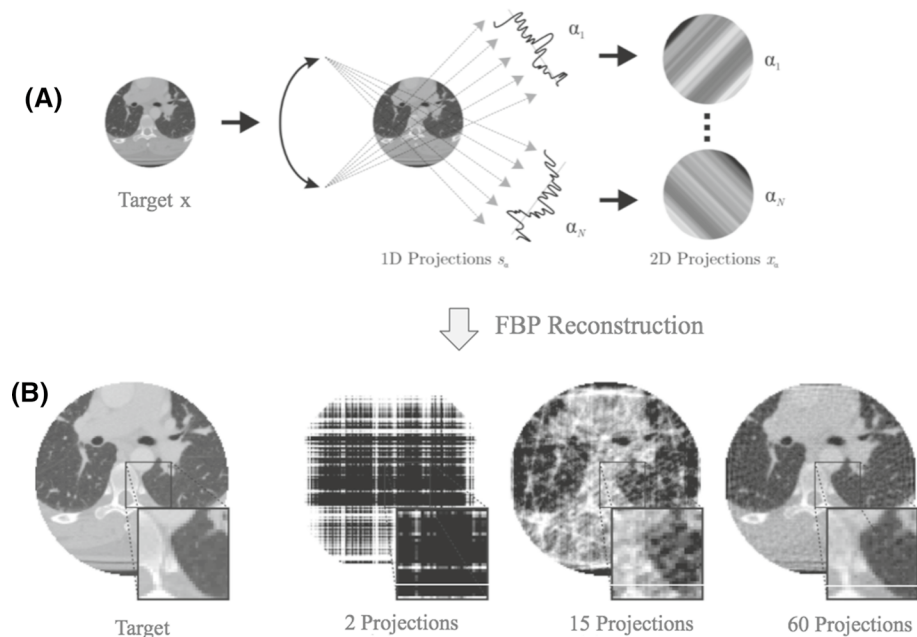

Fig. 5 CT Image reconstruction from sparse view measurements. (**A**) Generation of 2D projections from a target 2D CT slice image × for a number of N fixed angles α_i. (**B**) Image reconstructed using conventional filtered back projection method for different number of a projection angles N. Modified figures from (Thaler et al. 2018) Reprinted by permission from Springer Nature (Thaler et al. 2018)

$$\hat{x}^* = \underset{\hat{x}}{\operatorname{argmin}} \, ||\mathcal{F}(\hat{x}) - y|| + \lambda \mathcal{R}(\hat{x}) \tag{3}$$

where $||\mathcal{F}(\hat{x}) - y||$ is a data fidelity term that measures the consistency of the approximate solution $\hat{x}$, in the space of acceptable images e.g 2D, 3D and representing the physical quantity of interest, to the measured signal y, which depends on the imaging operator and could include images, Fourier samples, line integrals, etc. $\mathcal{R}(\cdot)$ is a regularization term encoding the prior information about the data, and λ is a hyper-parameter that controls the contribution of the regularization term. The reconstruction error is minimized iteratively until convergence. We note that the solution space after convergence does not need to be singular, for example in methods where a distribution of solutions is sampled (Bayesian neural networks (Yang et al. 2018; Luo et al. 2019)). In such cases, the "$\hat{x}^* \in$" notation replaces the current "$\hat{x}^* =$". However, for consistency with recent art we use the simplified notation assuming a singular solution space at convergence.

The regularization term is often the most important part of the modeling and what researchers have mostly focused on in the literature as it vastly reduces the solution space by accounting for assumptions based on the underlying data (e.g., smoothness, sparsity, spatio-temporal redundancy). The interested reader can refer to (Dong and Shen 2015; McCann and Unser 2019) for more details on regularization modeling. Figure 6 shows an example of an iterative approach workflow for diffuse optical tomography (DOT) imaging.

Several image priors were formulated as sparse transforms to deal with incomplete data issues. The sparsity idea, representing a high dimensional image x by only a small number of nonzero coefficients, is one dominant paradigm that has been shown to drastically improve the reconstruction quality especially when the number of measurements M or theirs signal to noise ratio (SNR) is low. Given the assumption that an image can be represented with a few nonzero coefficients (instead of its number of pixels), it is possible to recover it from a much smaller number of measurements. A popular choice for a sparsifying transform is total variation (TV) that is widely studied in academic literature. The interested reader is referred to Rodríguez (2013) for TV based algorithms modeling details. While TV imposes a strong assumption on the gradient sparsity via the non-smooth
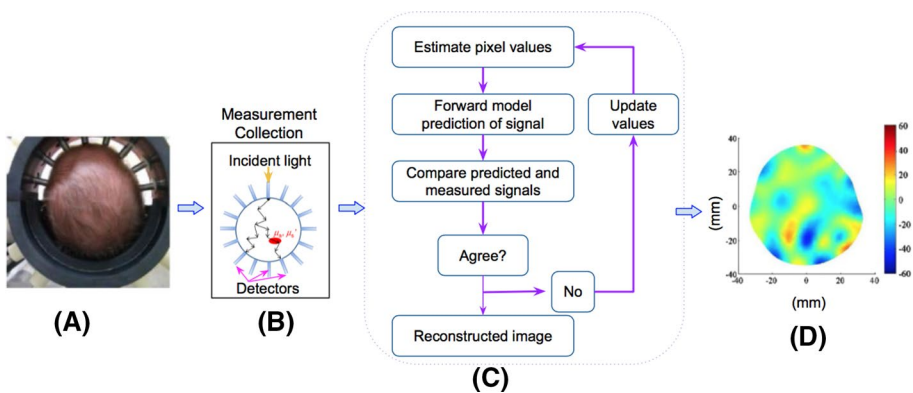

Fig. 6 Iterative image reconstruction workflow example. (**A**) Diffuse optical tomography (DOT) fibers brain probe consisting of a set of fibers for illumination and outer bundle fibers for detection (Hoshi and Yamada 2016). (**B**) Probe scheme and light propagation modelling in the head, used by the forward model. (**C**) Iterative approach pipeline. (**D**) DOT reconstructed image shows the total Hemoglobin (Hb) concentrations in the brain. (Figure licensed under CC-BY 4.0 Creative Commons license)

absolute value that is more suited to piece-wise constant images, TV tends to cause artifacts such as blurred details and undesirable patchy texture in the final reconstructions (Fig. 7 illustrates an example of artifacts present in TV-based reconstruction). Recent work aimed at exploiting richer feature knowledge to overcome TV's weaknesses, for example TV-type variants (Zhang et al. 2016a), non-local means (NLM) (Zhang et al. 2016b), wavelet approaches (Gao et al. 2011), and dictionary learning (Xu et al. 2012). Non-local means filtering methods, widely used for CT (Zhang et al. 2017), are operational in the image domain and allow the estimation of the noise component based on multiple patches extracted at different locations in the image (Sun et al. 2016).

While sparsity measures the first order sparsity, i.e the sparsity of vectors (the number of non-zero elements), some models exploited alternative properties such as the low rank of the data, especially when processing dynamic or time-series data, e.g. dynamic and functional MRI (Liang 2007; Singh et al. 2015) where data can have a correlation over time. Low rank can be interpreted as the measure of the second order (i.e., matrix) sparsity (Lin 2016). Low rank is of particular interest to compression given its requirement to a full utilization of the spatial and temporal correlation in images or videos leading to a better exploitation of available data. Recent work combined low rank and sparsity properties for improved reconstruction results (Zhao et al. 2012). Guo et al. (2014); Otazo et al. (2015) decomposed the dynamic image sequence into the sum of a low-rank component, that can capture the background or slowly changing parts of the dynamic object, and sparse component that can capture the dynamics in the foreground such as local motion or contrast changes. The interested reader is referred to (Lin 2016; Ravishankar et al. 2019) for low rank based algorithms modeling survey.

The difficulties of solving the image reconstruction problem motivated the design of highly efficient algorithms for large scale, nonsmooth and nonconvex optimization problems such as the alternating direction method of multipliers (ADMM) (Boyd et al. 2011; Gabay and Mercier 1976), primal-dual algorithm (Chambolle and Pock 2011b), iterative shrinkage-thresholding algorithm (ISTA) (Daubechies et al. 2004), to name just a few. For instance, using the augmented Lagrangian function (Sun et al. 2016) of Eq. 3, the ADMM algorithm solves the reconstruction problem by breaking it into smaller subproblems that are easier to solve. Despite the problem decomposition, a large number of iterations is still
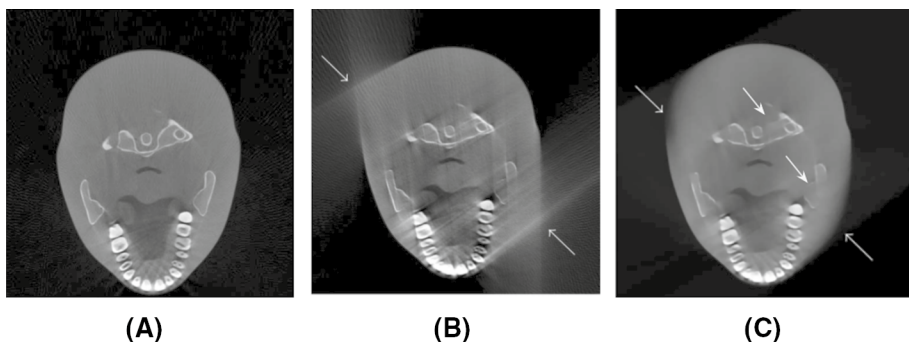

(A)   **(B)**   **(C)**

Fig. 7 Figure Reconstruction results of a limited-angle CT real head phantom (120 projections) using different conventional reconstruction methods. (**A**) Reference image reconstructed from 360 projections, (**B**) Image reconstruction using analytical FBP, (**C**) Images reconstructed by iterative POCS-TV with initial zero image (Guo et al. 2016). Arrows point out to artifacts present in the reconstructed images. (Figure licensed under CC-BY 4.0 Creative Commons license)

required for convergence to a satisfactory solution. Furthermore, performance is defined to a large extent by the choice of transformation matrix and the shrinkage function which remain challenging to choose (McCann and Unser 2019).

ISTA is based on a simpler gradient-based algorithm where each iteration leverages hardware accelerated matrix-vector multiplications representing the physical system modeling the image acquisition process and its transpose followed by a shrinkage/soft-threshold step (Daubechies et al. 2004). While the main advantage of ISTA relies on its simplicity, it has also been known to converge quite slowly (Zhang and Dong 2019). Convergence was improved upon by a fast iterative soft-thresholding algorithm (FISTA)(Beck and Teboulle 2009) based on the idea of Nesterov (1983).

Overall, although iterative reconstruction methods showed substantial accuracy improvements and artifact reductions over the analytical ones, they still face three major weaknesses: First, iterative reconstruction techniques tend to be vendor-specific since the details of the scanner geometry and correction steps are not always available to users and other vendors. Second, there are substantial computational overhead costs associated with popular iterative reconstruction techniques due to the load of the projection and back-projection operations required at each iteration. The computational cost of these methods is often several orders of magnitude higher than analytical methods, even when using highly-optimized implementations. A trade-off between real-time performance and quality is made in favor of quality in iterative reconstruction methods due to their non-linear complexity of quality in function of the processing time. Finally, the reconstruction quality is highly dependent on the regularization function form and the related hyper-parameters settings as they are problem-specific and require non-trivial manual adjustments. Over-imposing sparsity ($\mathcal{L}\infty$ penalties) for instance can lead to cartoon-like artifacts (McCann and Unser 2019). Proposing a robust iterative algorithm is still an active research area (Sun et al. 2019c; Moslemi et al. 2020).

3 Deep learning based image reconstruction

To further advance biomedical image reconstruction, a more recent trend is to exploit deep learning techniques for solving the inverse problem to improve resolution accuracy and speed up reconstruction results. As a deep neural network represents a complex mapping, it can detect and leverage features in the input space and build increasingly abstract representations useful for the end-goal. Therefore, it can better exploit the measured signals by extracting more contextual information for reconstruction purposes. In this section, we summarize works using DL for inverse problems in imaging.

Learning-based image reconstruction is driven by data where a training dataset is used to tune a parametric reconstruction algorithm that approximates the solution to the inverse problem with a significant one-time, offline training cost that is offset by a fast inference time. There is a variety of these algorithms, with some being closely related to conventional methods and others not. While some methods considered machine learning as a reconstruction step by combining a traditional model with deep modeling to recover the missing details in the input signal or enhance the resulting image (Sect. 3.1), some others considered a more elegant solution to reconstruct an image from its equivalent initial measurements directly by learning all the parameters of a deep neural network, in an end-to-end fashion, and therefore approximating the underlying physics of the inverse problem (Sect. 3.2), or even going further and solving for the target task directly (Sect. 3.3). Figure 8

shows a generic example of the workflow of these approaches. Table 1 surveys various papers based on these different paradigms and provides a comparison in terms of used data (Table 1-Column "Mod.","Samp.","D"), architecture (Table 1-Column "TA","Arch."), loss and regularization (Table 1-Column "Loss","Reg."), augmentation (Table 1-Column "Aug."), etc .

3.1 Deep learning as processing step: two step image reconstruction models

Complementing a conventional image reconstruction approach with a DL-model enables improved accuracy while reducing the computational cost. The problem can be addressed either in the sensor domain (pre-processing) or the image domain (post-processing) (Fig. 8a, Table 1-Column "E2E").

3.1.1 A pre-processing step (sensor domain)

The problem is formulated as a regression in the sensor domain from incomplete data representation (e.g., sub-sampling, limited view, low dose) to complete data (full dose or view) using DL methods and has led to enhanced results (Hyun et al. 2018; Liang et al. 2018; Cheng et al. 2018). The main goal is to estimate, using a DL model, missing parts of the signal that have not been acquired during the acquisition phase in order to input a better signal to an analytical approach for further reconstruction.
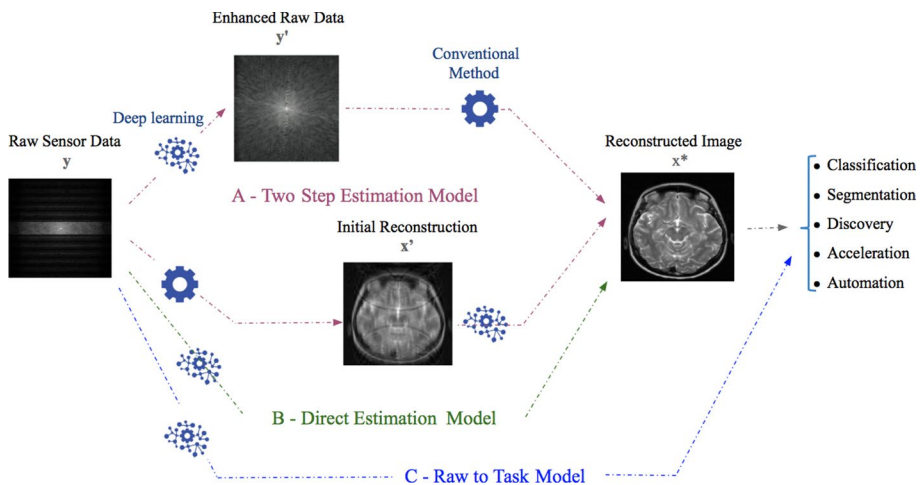

Fig. 8 Visualization of common deep learning-based reconstruction paradigms from raw sensor data. (**A**) A two-step processing model is shown where deep learning complements conventional methods (Sect. 3.1). A typical example would be to pre-process the raw sensor data using a conventional approach f_1, enhance the resulting image with a deep learning model f_{θ_2} then perform task processing using f_{θ_3}. The overall function will thus be $f_{\theta_3} \circ f_{\theta_2} \circ f_1$. Vice versa, one can pre-process the raw sensor data using a DL-based approach f_{θ_2} to enhance the collected raw data and recover missing details then apply conventional method f_1 to reconstruct image and finally perform task processing using f_{θ_3}. The overall function will be $f_{\theta_3} \circ f_1 \circ f_{\theta_2}$ in this case. (**B**) An end-to-end model is shown: The image is directly estimated from the raw sensor data with a deep learning model 3.2 followed by downstream processing tasks. The overall function will be $f_{\theta_3} \circ f_{\theta_2}$. (**C**) Task results can be inferred with or without explicit image reconstruction (Sect. 3.3) using f_{θ_3} function

Hyun et al. (2018) proposed a k-space correction based U-Net to recover the unsampled data followed by an IFT to obtain the final output image. They demonstrated artifact reduction and morphological information preservation when only 30% of the k-space data is acquired. Similarly, Liang et al. (2018) proposed a CT angular resolution recovery based on deep residual convolutional neural networks (CNNs) for accurate full view estimation from unmeasured views while reconstructing images using filtered back projection. Reconstruction demonstrated speed-up with fewer streaking artifacts along with the retrieval of additional important image details. Unfortunately, since noise is not only present in the incomplete data acquisition case, but also in the full data as well, minimizing the error between the reference and the predicted values can cause the model to learn to predict the mean of these values. As a result, the reconstructed images can suffer from lack of texture detail.

Huang et al. (2019b) argue that DL-based methods can fail to generalize to new test instances given the limited training dataset and DL's vulnerability to small perturbations especially in noisy and incomplete data case. By constraining the reconstructed images to be consistent with the measured projection data, while the unmeasured information is complemented by learning based methods, reconstruction quality can be improved. DL predicted images are used as prior images to provide information in missing angular ranges first followed by a conventional reconstruction algorithm to integrate the prior information in the missing angular ranges and constrain the reconstruction images to be consistent to the measured data in the acquired angular range.

Signal regression in the sensor domain reduces signal loss enabling improved downstream results from the coupled analytic method. However, the features extracted by DL methods are limited to the sensor domain only while analytical methods' weaknesses are still present in afterword processing.

3.1.2 A post-processing step (image domain)

The regression task is to learn the mapping between the low-quality reconstructed image and its high-quality counterpart. Although existing iterative reconstruction methods improved the reconstructed image quality, they remain computationally expensive and may still result in reconstruction artifacts in the presence of noise or incomplete information, e.g. sparse sampling of data (Cui et al. 2019; Singh et al. 2020). The main difficulty arises from the non-stationary nature of the noise and serious streaking artifacts due to information loss (Chen et al. 2012; Al-Shakhrah and Al-Obaidi 2003). Noise and artifacts are challenging to isolate as they may have strong magnitudes and do not obey specific model distributions in the image domain (Wang et al. 2016). The automatic learning and detection of complex patterns offered by deep neural networks can outperform handcrafted filters in the absence of complete information (Yang et al. 2017a; Singh et al. 2020).

Given an initial reconstruction estimate from a direct inverse operator e.g., FBP (Sun et al. 2018; Gupta et al. 2018; Chen et al. 2017a), IFT (Wang et al. 2016), or few iterative approach steps (Jin et al. 2017; Cui et al. 2019; Xu et al. 2017), deep learning is used to refine the initialized reconstruction and produce the final reconstructed image. For example, Chen et al. (2017b) used an autoencoder to improve FBP results on a limited angle CT projection. Similarly, Jin et al. (2017) enhanced FBP results on a sparse-view CT via subsequent filtering by a U-Net to reduce artifacts. U-Net and other hour-glass shaped architectures rely on a bottleneck layer to encode a low-dimensional representation of the target image.

Generative adversarial networks (GAN) (Goodfellow et al. 2014) were leveraged to improve the quality of reconstructed images. Wolterink et al. (2017) proposed to train an adversarial network to estimate full-dose CT images from low-dose CT ones and showed empirically that an adversarial network improves the model's ability to generate images with reduced aliasing artifacts. Interestingly, they showed that combining squared error loss with adversarial loss can lead to a noise distribution similar to that in the reference full-dose image even when no spatially aligned full-dose and low dose scans are available.

Yang et al. (2017a) proposed a deep de-aliasing GAN (DAGAN) for compressed sensing MRI reconstruction that resulted in reduced aliasing artifacts while preserving texture and edges in the reconstruction. Remarkably, a combined loss function based on content loss (consisting of a pixel-wise image domain loss, a frequency domain loss and a perceptual VGG loss) and adversarial loss were used. While frequency domain information was incorporated to enforce similarity in both the spatial (image) and the frequency domains, a perceptual VGG coupled to a pixel-wise loss helped preserve texture and edges in the image domain.

Combining DL and conventional methods reduce the computational cost but has its own downsides. For instance, the features extracted by DL methods are highly impacted by the results of the conventional methods, especially in case of limited measurements and the presence of noise where the initially reconstructed image may contain significant and complex artifacts that may be difficult to remove even by DL models. In addition, the information missing from the initial reconstruction is challenging to be reliably recovered by post-processing like many inverse problems in the computer vision literature such as image inpainting. Furthermore, the number of iterations required to obtain a reasonable initial image estimate using an iterative method can be hard to define and generally requires a long run-time (in the order of several min) to be considered for real-time scanning. Therefore, the post-processing approach may be more suitable to handle initial reconstructions that are of relatively good quality and not drastically different from the high-quality one.

3.2 End-to-end image reconstruction: direct estimation

An end-to-end solution leverages the image reconstruction task directly from sensor-domain data using a deep neural network by mapping sensor measurements to image domain while approximating the underlying physics of the inverse problem (Fig. 8b). This direct estimation model may represent a better alternative as it benefits from the multiple levels of abstraction and the automatic feature extraction capabilities of deep learning models.

Given pairs of measurement vectors $y \in \mathbb{C}^M$ and their corresponding ground truth images $x \in \mathbb{C}^N$ (that produce y), the goal is to optimize the parameters $\theta \in \mathbb{R}^d$ of a neural network in an end-to-end manner to learn the mapping between the measurement vector y and its reconstructed tomographic image x, which recovers the parameters of underlying imaged tissue. Therefore, we seek the inverse function $\mathcal{A}^{-1}(\cdot)$ that solves:

$$\theta^* = \underset{\theta}{\arg\min} \, \mathcal{L}\big(\mathcal{A}^{-1}(y,\theta),x\big) + \lambda\mathcal{R}(\mathcal{A}^{-1}(y,\theta)) \tag{4}$$

where $\mathcal{L}$ is the loss function of the network that, broadly, penalizes the dissimilarity between the estimated reconstruction and the ground truth. The regularization term $\mathcal{R}$, often introduced to prevent over-fitting, can apply penalties on layer parameters (weights) or layer activity (output) during optimization. $\mathcal{L}1/\mathcal{L}2$ norms are the most common choices.

λ is a hyper-parameter that controls the contribution of the regularization term. Best network parameters (θ^*) depend on hyper-parameters, initialization and a network architecture choice.

Recently, several paradigms have emerged for end-to-end DL-based reconstruction the most common of which are generic DL models and DL models that unroll an iterative optimization.

3.2.1 Generic models

Although some proposed models rely on multilayer perceptron (MLP) feed-forward artificial neural network (Pelt and Batenburg 2013; Boublil et al. 2015; Feng et al. 2018; Wang et al. 2019), CNNs remain the most popular generic reconstruction models mainly due to their data filtering and features extraction capabilities. Specifically, encoder-decoder (Nehme et al. 2018; Häggström et al. 2019), U-Net (Waibel et al. 2018), ResNet (Cai et al. 2018) and decoder like architecture (Yoon et al. 2018; Wu et al. 2018; Zhu et al. 2018) are the most dominant architectures as they rely on a large number of stacked layers to enrich the level of features extraction. A set of skip connections enables the later layers to reconstruct the feature maps with both the local details and the global texture and facilitates stable training when the network is very deep. Figure 9 illustrates some of the architectures that are widely adopted in medical image reconstruction.

The common building blocks of neural network architectures are convolutional layers, batch normalization layers, and rectified linear units (ReLU). ReLU is usually used to enforce information non-negativity properties, given that the resulting pixels values represent tissue properties e.g., chromophores concentration maps (Yoo et al. 2020), refractive index (Sun et al. 2016), and more examples in Table 1. Batch normalization is used to reduce the internal covariate shift and accelerates convergence. The resulting methods can produce relatively good reconstructions in a short time and can be adapted to other modalities but require a large training dataset and good initialization parameters. Table 1-Column "E2E" (check-marked) summarizes papers by architecture, loss, and regularization for 2D, 3D, 4D, and different modalities.

Zhu et al. (2018) proposed a manifold learning framework based decoder neural network to emulate the fast-Fourier transform (FFT) and learn an end-to-end mapping between k-space data and image domains where they showed artifact reduction and reconstruction speed up. However, when trained with an $\mathcal{L}1$ or $\mathcal{L}2$ loss only, a DL-based reconstructed image still exhibits blurring and information loss, especially when used with incomplete data. Similarly, Yedder et al. (2018) proposed a decoder like model for DOT image reconstruction. While increased reconstruction speed and lesion localization accuracy are shown, some artifacts are still present in the reconstructed image when training with $\mathcal{L}2$ loss only. This motivated an improved loss function in their follow-up work (Yedder et al. 2019) where they suggested combining $\mathcal{L}2$ with a Jaccard loss component to reduce reconstructing false-positive pixels.

Thaler et al. (2018) proposed a Wasserstein GAN (WGAN) based architecture for improving the image quality for 2D CT image slice reconstruction from a limited number of projection images using a combination of $\mathcal{L}1$ and adversarial losses. Similarly, Ouyang et al. (2019) used a GAN based architecture with a task-specific perceptual and $\mathcal{L}1$ losses to synthesize PET images of high quality and accurate pathological features. Some attempts were made to reconstruct images observed by the human visual system directly from brain responses using GAN (Shen et al. 2019; St-Yves and Naselaris 2018). Shen
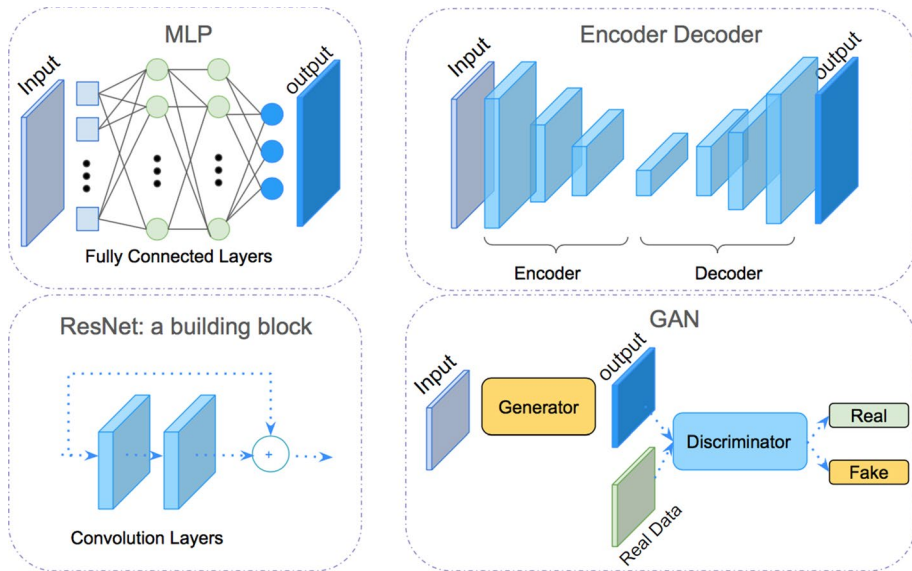
Fig. 9 Common network architectures used for image reconstruction. From left to right: a multilayer perceptron network based on fully connected layers; an encoder-decoder architecture based convolutional layers; a residual network, e.g. ResNet (He et al. 2016), utilizing skip connections (skip connection is crucial in facilitating stable training when the network is very deep); and a generative adversarial network (GAN). A decoder like architecture includes only the decoder part of the encoder-decoder architecture and may be preceded by a fully connected layer to map measurements to the image space depending on the input data size. A U-Net (Ronneberger et al. 2015) resembles the encoder-decoder architecture while it uses skip connections between symmetric depths. A GAN (Goodfellow et al. 2014) includes a generator and descriminor that contest with each other in a min-max game. The generator learns to create more realistic data by incorporating feedback from the discriminator. We refer the interested reader to (Alom et al. 2019; Khan et al. 2020) for a an in depth survey of various types of deep neural network architectures

et al. (2019) trained a GAN for functional magnetic resonance imaging (fMRI) reconstruction to directly map between brain activity and perception (the perceived stimulus). An adversarial training strategy which consists of three networks: a generator, a comparator, and a discriminator was adopted. The comparator is a pre-trained neural network on natural images used to calculate a perceptual loss term. The loss function is based on $\mathcal{L}2$ distance in image space, perceptual loss (similarity in feature space) and adversarial loss.

To further enhance results and reduce artifacts due to motion and corruption of k-space signal, Oksuz et al. (2019) proposed a recurrent convolutional neural network (RCNN) to reconstruct high quality dynamic cardiac MR images while automatically detecting and correcting motion-related artifacts and exploiting the temporal dependencies within the sequences. Proposed architecture included two sub-networks trained jointly: an artifact detection network that identifies potentially corrupted k-space lines and an RCNN for reconstruction.

To relax the requirement of a large number of training samples, a challenging requirement in a medical setting, simulating data was proposed as an alternative source of training data. However, creating a realistic synthetic dataset is a challenging task in itself as it requires careful modeling of the complex interplay of factors influencing real-world acquisition environment. To bridge the gap between the real and in silico worlds, transfer

learning provides a potential remedy as it helps transfer the measurements from the simulation domain to the real domain by keeping the general attenuation profile while accounting for real-world factors such as scattering, etc.

Yedder et al. (2019) proposed a supervised sensor data distribution adaptation based MLP to take advantage of cross-domain learning and reported accuracy enhancement in detecting tissue abnormalities. Zhou et al. (2019) proposed unsupervised CT sinograms adaptation, based on CycleGAN and content consistent regularization, to further alleviate the need for real measurement-reconstruction pairs. Interestingly, the proposed method integrated the measurement adaptation network and the reconstruction network in an end-to-end network to jointly optimize the whole network.

Waibel et al. (2018) investigated two different variants of DL-based methods for photoacoustique image reconstruction from limited view data. The first method is a post-processing DL method while the second one is an end-to-end reconstruction model. Interestingly, they showed empirically that an end-to-end model achieved qualitative and quantitative improvements compared to reconstruction with a post-processing DL method. Zhang and Liang (2020) studied the importance of fully connected layers, commonly used in end-to end model (Zhu et al. 2018; Yedder et al. 2019), to realize back-projection data from the sensor domain to the image domain and showed that while a back-projection can be learned through neural networks it is constrained by the memory requirements induced by the non-linear number of weights in the fully connected layers. Although several generic DL architectures and loss functions have been explored to further enhance reconstruction results in different ways (resolution, lesion localization, artifact reduction, etc.), a DL-based method inherently remains a black-box that can be hard to interpret. Interpretability is key not only for trust and accountability in a medical setting but also to correct and improve the DL model.

3.2.2 Unrolling iterative methods

Unrolling conventional optimization algorithms into a DL model has been suggested by several works (Qin et al. 2018; Schlemper et al. 2017; Würfl et al. 2018; Sun et al. 2016; Adler and Öktem 2018; Hosseini et al. 2019) in order to combine the benefits of traditional iterative methods and the expressive power of deep models (Table 1-Column "E2E"). Rajagopal et al. (2019) proposed a theoretical framework demonstrating how to leverage iterative methods to bootstrap network performance while preserving network convergence and interpretability featured by the conventional approaches.

Deep ADMM-Net (Sun et al. 2016) was the first proposed model reformulating the iterative reconstruction ADMM algorithm into a deep network for accelerating MRI reconstruction, where each stage of the architecture corresponds to an iteration in the ADMM algorithm. In its iterative scheme, the ADMM algorithm requires tuning of a set of parameters that are difficult to determine adaptively for a given data set. By unrolling the ADMM algorithm into a deep model, the tuned parameters are now all learnable from the training data. The ADMM-Net was later further improved to Generic-ADMM-Net (Yang et al. 2017b) where a different variable splitting strategy was adopted in the derivation of the ADMM algorithm and demonstrated state-of-the-art results with a significant margin over the BM3D-based algorithm (Dabov et al. 2007). Similarly, the PD-Net (Adler and Öktem 2018) adopted neural networks to approximate the proximal operator by unrolling the primal-dual hybrid gradient algorithm (Chambolle and Pock 2011a) and demonstrated performance boost compared with conventional and two step image reconstruction models.

Figure 10 shows results of their proposed method in a simplified setting of a two-dimensional CT image reconstruction problem.

In like manner, Schlemper et al. (2017) proposed a cascade convolutional network that embeds the structure of the dictionary learning-based method while allowing end-to-end parameter learning. The proposed model enforces data consistency in the sensor and image domain simultaneously, reducing the aliasing artifacts due to sub-sampling. An extension for dynamic MR reconstructions (Schlemper et al. 2018) exploits the inherent redundancy of MR data.

While, the majority of the aforementioned methods used shared parameters over iterations only, Qin et al. (2018) proposed to propagate learnt representations across both iteration and time. Bidirectional recurrent connections over optimization iterations are used to share and propagate learned representations across all stages of the reconstruction process and learn the spatio–temporal dependencies. The proposed deep learning based iterative algorithm can benefit from information extracted at all processing stages and mine the temporal information of the sequences to improve reconstruction accuracy. The advantages of leveraging temporal information was also demonstrated in single molecule localization microscopy (Cardoen et al. 2019). An LSTM was able to learn an unbiased emission density prediction in a highly variable frame sequence of spatio-temporally separated fluorescence emissions. In other words, joint learning over the temporal domain of each sequence and across iterations leads to improved de-aliasing.

3.2.3 Semi-supervised and unsupervised learning methods

The majority of deep learning methods for image reconstruction are based on supervised learning where the mapping between signal and reconstructed image is learned. However, the performance of supervised learning is, to a large degree, determined by the size of the available training data which is constrained in a medical setting. In an effort to work around this requirement, Gong et al. (2018) applied the deep image prior method (Ulyanov et al. 2018) to PET image reconstruction. The network does not need prior training matching pairs, as it substitutes the target image with a neural network representation of the target domain but rather requires the patient's own prior information. The maximum-likelihood estimation of the target image is formulated as a constrained optimization problem and is solved using the ADMM algorithm. Gong et al. (2019) have extended the approach to the
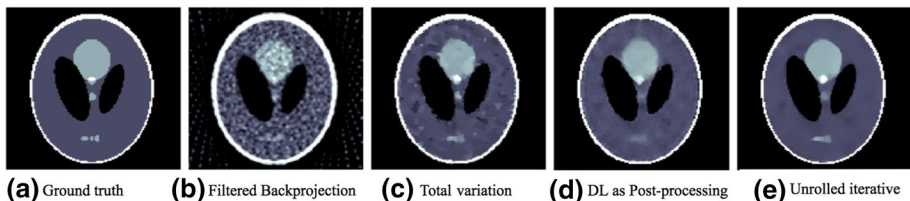

(a) Ground truth **(b)** Filtered Backprojection **(c)** Total variation **(d)** DL as Post-processing **(e)** Unrolled iterative

Fig. 10 Reconstructions of the Shepp–Logan phantom in a two-dimensional CT image reconstruction problem using different reconstruction methods. (**a**) Reference image.(**b**)–(**c**) The standard FBP and TV-regularized reconstruction. (**d**) FBP reconstruction followed by a U-Net architecture. (**e**) The learned primal-dual scheme (Adler and Öktem 2018). One can observe the artifacts in the conventional methods removed by the DL-based approaches. Overall, the unrolled-iterative method gave the best results and outperformed even the DL as a post-processing step method. The interested reader is referred to (Arridge et al. 2019) for more details about the training setting. (Reproduced with permission of Cambridge University Press)

direct parametric PET reconstruction where acquiring high-quality training data is more difficult. However, to obtain high-quality reconstructed images without prior training datasets, registered MR or CT images are necessary. While rigid registration is sufficient for brain regions, accurate registration is more difficult for other regions.

Instead of using matched high-dose and a low-dose CT, Kang et al. (2019) proposed to employ the GAN loss to match the probability distribution where a cycleGAN architecture based cyclic loss and identity loss for multiphase cardiac CT problems is used. Similarly, Li et al. (2020b) used an unsupervised data fidelity enhancement network that uses an unsupervised network to fine-tune a supervised network to different CT scanning protocol properties. Meng et al. (2020) proposed to use only a few supervised sinogram-image pairs to capture critical features (i.e., noise distribution, tissue characteristics) and transfer these features to larger unlabeled low-dose sinograms. A weighted least-squares loss model based TV regularization term and a KL divergence constraint between the residual noise of supervised and unsupervised learning is used while a FBP is employed to reconstruct CT images from the obtained sinograms.

3.3 Raw-to-task methods: task-specific learning

Typical data flow in biomedical imaging involves solving the image reconstruction task followed by preforming an image processing task (e.g., segmentation, classification) (Fig 11a). Although each of these two tasks (reconstruction and image processing) is solved separately, the useful clinical information extracted for diagnosis by the second task is highly dependent on the reconstruction results. In the raw-to-task paradigm, task results are directly inferred from raw data, where image reconstruction and processing are lumped together and reconstructed image may not be necessarily outputted (Fig. 8c, Fig. 11b)

Jointly solving for different tasks using a unified model is frequently considered in the computer vision field, especially for image restoration (Sbalzarini 2016), and has lead to improved results than solving tasks sequentially (Paul et al. 2013). The advantages explain the recent attention this approach received in biomedical image reconstruction (Sun et al. 2019a; Huang et al. 2019a). For instance, a unified framework allows joint parameters tuning for both tasks and feature sharing where the two problems regularize each other when considered jointly. In addition, when mapping is performed directly from the sensor domain, the joint task can even leverage sensor domain features for further results enhancing while it can be regarded as a task-based image quality metric that is learned from the
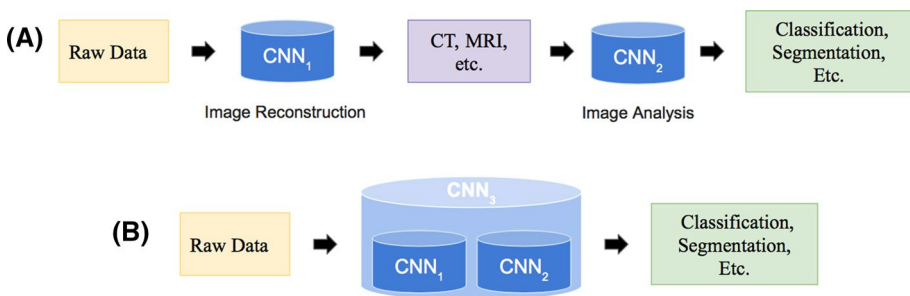

Fig. 11 (**A**) Biomedical image processing workflow usually involves two steps optimized independently (reconstruction and image analysis) for diagnosis purposes. (**B**) Jointly solving these tasks using a unified model allows joint parameters tuning and feature sharing

data. Furthermore, Sbalzarini (2016) argues that solving ill-posed problems in sequence can be more efficient than in isolation. Since the output space of a method solving an inverse problem is constrained by forming the input space of the next method, the overall solution space contracts. Computational resources can then be focused on this more limited space improving runtime, solution quality or both.

Image reconstruction is not always necessary for optimal performance. For example, in single-pixel imaging recent work (Latorre-Carmona et al. 2019) illustrates that an online algorithm can ensure on-par performance in image classification without requiring the full image reconstruction. A typical pipeline in processing compressive sensing requires image reconstruction before classical computer vision (CV) algorithms process the images in question. Recent work (Braun et al. 2019) has shown that CV algorithms can be embedded into the compressive sensing domain directly, avoiding the reconstruction step altogether. In discrete tomography the full range of all possible projections is not always feasible or available due to time, resource, or physical constraints. A hybrid approach in discrete tomography (Pap et al. 2019) finds a minimal number of projections required to obtain a close approximation of the final reconstruction image. While an image is still reconstructed after the optimal number of projections is found, determining this optimal set required no reconstructions, a significant advantage compared to an approach where either an iterative or combinatorial selection of reconstructions to determine the optimal set of projections.

4 Medical training datasets

The performance of learning-based methods is dictated to a large extent by the size and diversity of the training dataset. In a biomedical setting, the need for large, diverse, and generic datasets is non-trivial to satisfy given constraints such as patient privacy, access to acquisition equipment and the problem of divesting medical practitioners to annotate accurately the existing data. In this section, we will discuss how researchers address the trade-offs in this dilemma and survey the various publicly available dataset type used in biomedical image reconstruction literature.

Table 1-Columns "Data", "Site" and "Size" summarise details about dataset used by different surveyed papers, which are broadly classified into clinical (real patient), physical phantoms, and simulated data. The sources of used datasets have been marked in the last column of Table 1-Columns "Pub". Data" in case of their public availability to other researchers. Since phantom data are not commonly made publicly available, the focus was mainly given to real and simulated data whether they are publicly available or as part of challenges. Used augmentation techniques have been mentioned in Table 1-Columns "Aug". Remarkably, augmentation is not always possible in image reconstruction task especially in sensor domain given the non-symmetries of measurements in some case, the nonlinear relationship between raw and image data, and the presence of other phenomena(e.g., scattering). We herein survey the most common source of data and discuss their pros and cons.

4.1 Real-world datasets

Some online platforms (e.g., Lung Cancer Alliance , Mridata , MGH-USC HCP (2016), and Biobank) made the initiative to share datasets between researchers for image reconstruction task. Mridata, for example, is an open platform for sharing clinical MRI raw

k-space datasets. The dataset is sourced from acquisitions of different manufacturers, allowing researchers to test the sensitivity of their methods to overfitting on a single machine's output while may require the application of transfer-learning techniques to handle different distributions. As of writing, only a subset of organs for well known modalities e.g., MRI and CT are included (Table 1-Columns "Site"). Representing the best reconstruction images acquired for a specific modality, the pairs of signal-image form a gold standard for reconstruction algorithms. Releasing such data, while extremely valuable for researchers, is a non-trivial endeavour where legal and privacy concerns have to be taken into account by, for example, de-anonymization of the data to make sure no single patient or ethnographically distinct subset of patients can ever be identified in the dataset. Source of real-word used datasets on surveyed papers has been marked in Table 1-Columns "Pub. Data" where the sizes remain relatively limited to allow a good generalization of DL-based methods.

4.2 Physics-based simulation

Physics-based simulation (Schweiger and Arridge 2014; Harrison 2010; Häggström et al. 2016) provides an alternative source of training data that allows generating a large and diversified dataset. The accuracy of a physical simulation with respect to real-world acquisitions increases at the cost of an often super-linear increase in computational resources. In addition, creating realistic synthetic datasets is a nontrivial task as it requires careful modeling of the complex interplay of factors influencing real-world acquisition environment. With a complete model of the acquisition far beyond computational resources, a practitioner needs to determine how close to reality the simulation needs to be in order to allow the method under development to work effectively. Transfer learning provides a potential remedy to bridge the gap between real and in silico worlds and alleviates the need for a large clinical dataset (Zhu et al. 2018; Yedder et al. 2019). In contrast, the approach of not aiming for complete realism but rather using the simulation as a tool to sharpen the research question can be appropriate. Simulation is a designed rather than a learned model. For both overfitting to available data is undesirable. The assumptions underlying the design of the simulation are more easily verified or shown not to hold if the simulation is not fit to the data, but represents a contrasting view. For example, simulation allows the recreation and isolation of edge cases where a current approach is performing sub-par. As such simulation is a key tool for hypothesis testing and validation of methods during development. For DL-based methods the key advantage simulation offers is the almost unlimited volume of data that can be generated to augment limited real-world data in training. With the size of datasets as one of the keys determining factors for DL-based methods leveraging simulation is essential. Surveyed papers that used simulated data as a training or augmentation data have been marked in Table 1-Columns "Pub".

4.3 Challenge datasets

There are only a few challenge (competition) datasets for image reconstruction task e.g., LowDoseCT (2014) FastMRI (Zbontar et al. 2018), Fresnel (Geffrin et al. 2005) and SMLM challenge (Holden and Sage 2016) that includes raw measurements. Low-quality signals can be simulated by artificially subsampling the full-dose acquired raw signal while keeping their corresponding high resolution images pair (Gong et al. 2018; Xu et al. 2017). While this method offers an alternative source of training data, the downsampling

is only one specific artificial approximation to the actual acquisition of low-dose imaging and may not reflect the real physical condition. Hence, performance can be compromised by not accounting for the discrepancy between the artificial training data and real data. Practitioners can leverage techniques such as transfer learning to tackle the discrepancy. Alternatively, researchers collect high-quality images from other medical imaging challenges, e.g., segmentation (MRBrainS challenge (Mendrik et al. 2015), Bennett and Simon (2013)), and use simulation, using a well known forward model, to generate full and/or incomplete sensor domain pairs. Here again, only a subset of body scans and diseases for well-studied modalities are publicly available as highlighted in Table 1-Columns "Site" and "Pub. Data".

5 Reconstruction evaluation metrics

5.1 Quality

Measuring the performance of the reconstruction approaches is usually performed using three metrics, commonly applied in computer vision, in order to access the quality of the reconstructed images. These metrics are the root mean squared error (RMSE) or normalized mean squared error, structural similarity index (SSIM) or its extended version multiscale structural similarity (Wang et al. 2003), and peak signal to noise ratio (PSNR).

While RMSE measures the pixel-wise intensity difference between ground truth and reconstructed images by evaluating pixels independently, ignoring the overall image structure, SSIM, a perceptual metric, quantifies visually perceived image quality and captures structural distortion. SSIM combines luminance, contrast, and image structure measurements and ranges between [0,1] where the higher SSIM value the better and SSIM = 1 means that the two compared images are identical.

PSNR (Eq. 5) is a well-known metric for image quality assessment which provides similar information as the RMSE but in units of dB while measuring the degree to which image information rises above background noise. Higher values of PSNR indicate a better reconstruction.

$$PSNR = 20.\log_{10}\left(\frac{X_{max}}{\sqrt{MSE}}\right) \qquad (5)$$

where X_{max} is the maximum pixel value of the ground truth image.

Illustrating modality specific reconstruction quality is done by less frequently used metrics such as contrast to noise ratio (CNR) (Wu et al. 2018) for US. Furthermore, modalities such as US, SMLM or some confocal fluorescence microscopy can produce a raw image where the intensity distribution is exponential or long-tailed. Storing the raw image in a fixed-precision format would lead to unacceptable loss of information due to uneven sampling of the represented values. Instead, storing and more importantly processing the image in logarithmic or dB-scale preserves the encoded acquisition better. The root mean squared logarithmic error (RMSLE) is then a logical extension to use as an error metric (Beattie 2018).

Normalized mutual information (NMI) is a metric used to determine the mutual information shared between two variables, in this context image ground truth and reconstruction (Wu et al. 2018; Zhou et al. 2019). When there is no shared information NMI is 0,

whereas if both are identical a score of 1 is obtained. To illustrate NMI's value, consider two images X, Y with random values. When generated from two different random sources, X and Y are independent, yet RMSE (X,Y) can be quite small. When the loss function minimizes RMSE, such cases can induce stalled convergence at a suboptimal solution due to a constant gradient. NMI, on the other hand, would return zero (or a very small value), as expected. RMSE minimizes average error which may make it less suitable for detailed distribution matching tasks in medical imaging such as image registration where NMI makes for a more effective optimization target (Wells III et al. 1996).

The intersection over union, or Jaccard index, is leveraged to ensure detailed accurate reconstruction (Yedder et al. 2018; Sun et al. 2019a). In cases where the object of interest is of variable size and small with respect to the background, an RMSE score is biased by matching the background rather than the target of interest. In a medical context, it is often small deviations (e.g. tumors, lesions, deformations) that are critical to diagnosis. Thus, unlike computer vision problems where little texture changes might not alter the overall observer's satisfaction, in medical reconstruction, further care should be taken to ensure that practitioners are not misled by a plausible but not necessarily correct reconstruction. Care should be taken to always adjust metrics with respect to their expected values under an appropriate random model (Gates and Ahn 2017). The understanding of how a metric responds to its input should be a guideline to its use. As one example, the normalization method in NMI has as of writing no less than 6 (Gates and Ahn 2017) alternatives with varying effect on the metric. Table 1-Columns "Metrics" surveys the most frequently used metrics on surveyed papers.

5.2 Inference speed

With reconstruction algorithms constituting a key component in time-critical diagnosis or intervention settings, the time complexity is an important metric in selecting methods. Two performance criteria are important in the context of time: throughput measures how many problem instances can be solved over a time period, and latency measures the time needed to process a single problem instance. In a non-urgent medical setting, a diagnosing facility will value throughput more than latency. In an emergency setting where even small delays can be lethal, latency is critical above all. For example, if a reconstruction algorithm is deployed on a single device it is not unexpected for there to be waiting times for processing. As a result latency, if the waiting time is included, will be high and variable, while throughput is constant. In an emergency setting there are limits as to how many devices can be deployed, computing results on scale in a private cloud on the other hand can have high throughput, but higher latency as there will be a need to transfer data offsite for processing. In this regard it is critical for latency sensitive applications to allow deployment on mobile (low-power) devices. To minimize latency (including wait-time), in addition to parallel deployment, the reconstruction algorithm should have a predictable and constant inference time, which is not necessarily true for iterative approaches.

Unfortunately, while some papers reported their training and inference times, (Table 1-Columns "Metrics-IS") it is not obvious to compare their time complexity given the variability in datasets, sampling patterns, hardware, and DL frameworks. A dedicated study, out of the scope of this survey, needs to be conducted for a fair comparison. Overall, the offline training of DL methods bypasses the laborious online optimization procedure of conventional methods, and has the advantage of lower inference time over all but the simplest analytical. In addition, when reported, inference times are usually in the order

of milliseconds per image making them real time capable. The unrolled iterative network, used mainly in the MR image reconstruction (Sun et al. 2016; Yang et al. 2017b) as the forward and backward projections can be easily performed via FFT, might be computationally expensive for other modality where the forward and backward projections can not be computed easily. Therefore, the total training time can be much longer. In addition when data is acquired and reconstructed in 3D, the GPU memory becomes a limiting factor especially when multiple unrolling modules are used or the network architecture is deep.

6 Conclusion, discussion and future direction

Literature shows that DL-based image reconstruction methods have gained popularity over the last few years and demonstrated image quality improvements when compared to conventional image reconstruction techniques especially in the presence of noisy and limited data representation. DL-based methods address the noise sensitivity and incompleteness of analytical methods and the computational inefficiency of iterative methods. In this section we will illustrate the main trends in DL-based medical image reconstruction: a move towards task-specific learning where image reconstruction is no longer required for the end task, and the focus on a number of strategies to overcome the inherent sparsity of manually annotated training data.

6.1 Discussion

Learning: Unlike conventional approaches that work on a single image in isolation and require prior knowledge, DL-based reconstruction methods leverage the availability of large-scale datasets and the expressive power of DL to impose implicit constraints on the reconstruction problem. While DL-based approaches do not require prior knowledge, their performance can improve with it. By not being dependent on prior knowledge, DL-based methods are more decoupled from a specific imaging (sub)modality and thus can be more generalizable. The ability to integrate information from multiple sources without any preprocessing is another advantage of deep neural networks. Several studies have exploited GANs for cross-modal image generation (Li et al. 2020a; Kearney et al. 2020) as well as to integrate prior information (Lan et al. 2019; Bhadra et al. 2020). Real-time reconstruction is offered by DL-based methods by performing the optimization or learning stage offline, unlike conventional algorithms that require iterative optimization for each new image. The diagnostician can thus shorten diagnosis time increasing the throughput in patients treated. In operating theatres and emergency settings this advantage can be life saving.

Interpretability: While the theoretical understanding and the interpretability of conventional reconstruction methods are well established and strong (e.g., one can prove a method's optimality for a certain problem class), it is weak for the DL-based methods (due to the black-box nature of DL) despite the effort in explaining the operation of DL-based methods on many imaging processing tasks. However, one may accept the possibility that interpretabilty is secondary to performance as fully understanding DL-based approaches may never become practical.

Complexity: On the one hand, conventional methods can be straightforward to implement, albeit not necessarily to design. On the other, they are often dependent on parameters requiring manual intervention for optimal results. DL-based approaches can be challenging to train with a large if not intractable hyper-parameter space (e.g., learning rate,

initialization, network design). In both cases, the hyper-parameters are critical to results and require a large time investment from the developer and the practitioner. In conclusion, there is a clear need for robust self-tuning algorithms, for both DL-based and conventional methods.

Robustness: Conventional methods can provide good reconstruction quality when the measured signal is complete and the noise level is low, their results are consistent across datasets and degrade as the data representation and/or the signal to noise ratio is reduced by showing noise or artifacts (e.g., streaks, aliasing). However, a slight change in the imaging parameters (e.g., noise level, body part, signal distribution, adversarial examples, and noise) can severely reduce the DL-based approaches' performances and might lead to the appearance of structures that are not supported by the measurements (Antun et al. 2019; Gottschling et al. 2020). DL based approaches still leave many unsolved technical challenges regarding their convergence and stability that in turn raise questions about their reliability as a clinical tool. A careful fusion between DL-based and conventional approaches can help mitigate these issues and achieve the performance and robustness required for biomedical imaging.

Speed: DL-based methods have the advantage in processing time over all but the most simple analytical methods at inference time. As a result, latency will be low for DL-based methods. However, one must be careful in this analysis. DL-based methods achieve fast inference by training for a long duration, up to weeks, during development. If any changes to the method are needed and retraining is required, even partial, a significant downtime can ensue. Typical DL-based methods are not designed to be adjusted at inference time. Furthermore, when a practitioner discovers that, at diagnosis time, the end result is subpar, an iterative method can be tuned by changing its hyper-parameters. For a DL-based approach, this is non-trivial if not infeasible.

A final if not less important distinction is adaptive convergence. At deployment, a DL-based method has a fixed architecture and weights with a deterministic output. Iterative methods can be run iterations until acceptable performance is achieved. This is a double-edged sword as convergence is not always guaranteed and the practitioner might not know exactly how many more iterations are needed.

Training Dataset: Finally, the lack of large scale biomedical image datasets for training due to privacy, legal, and intellectual property related concerns, limits the application of DL-based methods on health care. Training DL-based models often requires scalable high performance hardware provided by cloud based offerings. However, deploying on cloud computing and transmitting the training data risks the security, authenticity, and privacy of that data. Training on encrypted data offers a way to ensure privacy during training (Gilad-Bachrach et al. 2016). More formally a homomorphic encryption algorithm (Rivest et al. 1978) can ensure evaluation (reconstruction) on the encrypted data results are identical after decryption to reconstruction on the non-encrypted data. In practice, this results in an increase in dataset size as compression becomes less effective, a performance penalty is induced by the encryption and decryption routines, and interpretability and debugging the learning algorithm becomes more complex since it operates on human unreadable data.

The concept of federated learning, where improvements of a model (weights) are shared between distributed clients without having to share datasets, has seen initial success in ensuring privacy while enabling improvements in quality (Geyer et al. 2017). However a recent work by Zhu et al. (2019) has shown that if an attacker has access to the network architecture and the shared weights, the training data can be reconstructed with high fidelity from the gradients alone. Data sharing security in a federated setting still presents a concern that requires further investigations.

Simulating a suitable training set also remains a challenge that requires careful tuning and more realistic physical models to improve DL-based algorithm generalization.

Performance: Reconstruction is sensitive to missing raw measurements (false negative, low recall) and erroneous signal (false positive, low precision). Sensitivity in reconstruction is critical to ensure all present signals are reconstructed but high recall is not sufficient as in this setting a reconstructed image full of artifacts and noise can still have high recall. Specificity may be high in reconstructed images, where little to no artifacts or noise is present, but at the cost of omitting information. In the following we discuss how the indicators of both are affected by training regimes and dataset size.

Deep learning reconstruction quality varies when using different loss metrics. For instance, when a basic $\mathcal{L}1$ or $\mathcal{L}2$ loss in image space is used solely, reconstruction results tend to be worse quality because it generates an average of all possible reconstructions having the same distance in image space, hence the reconstructed images might still present artifacts (Shen et al. 2019; Yedder et al. 2018). Adding feature loss in high dimensional space like perceptual loss and frequency domain loss helps to better constrain the reconstruction to be perceptually more similar to the original image (Yang et al. 2017a; Shen et al. 2019; Huang et al. 2019b). Furthermore, reconstruction quality varies with dataset size. To analyze the effect of training-dataset size on reconstruction quality, Shen et al. (2019) trained their reconstruction model with a variable number of training samples while gradually increasing the dataset size from 120 to 6,000 samples. Interestingly, they reported that while the reconstruction quality improved with the number of training samples, the quality increase trend did not converge which suggests that better reconstruction quality might be achieved if larger datasets are available. Similarly, Sun and Kamilov (2018) studied the effect of decreasing the number of measurements and varying the noise level on multiple scattering based deep learning reconstruction methods and concluded that better feature extraction from higher dimensional space are required for very low sampling cases. In addition, data is usually imbalanced, with the unhealthy (including anomalies) samples being outnumbered by the healthy ones resulting in reduced performance and compromised reconstruction of tumor related structures (Wen et al. 2019).

While deep networks can take the whole image and produce outputs in a single forward pass, input size can sometimes be a limiting factor as excessive memory cost becomes a challenge for current GPU architectures especially for three-dimensional (3D) images with relatively high resolution. Patch-based training divides the image into possibly overlapping patches and produces outputs for each sub-image in independent forward passes while requiring a reduced amount of memory and time on GPU. For patch-based training, each iteration can take a set of patches (minibatch) for a more stable training where a proper choice of batch size is important for a network convergence and generalization (Goceri and Gooya 2018). However, patch-based learning is mostly relevant when training in the same domain, which is the case of deep learning as a pre/post-processing step (Cheng et al. 2018; Wu et al. 2019; Xu et al. 2017), given the difficulty to map part of the raw signal to a corresponding image patch in end-to-end learning.

Task-specific Learning: Prediction tasks, such as classification or prediction of treatment directly from measurements, that is, omitting the image reconstruction task, is a nontrivial approach that, under the right conditions, can be very worthwhile. Since an explicit encoding of the image is no longer needed, the architecture can be reduced in complexity or redesigned to focus specifically on the task. In addition, focusing on task-specific performance ensures all computational resources are committed to task-specific training rather than expending focus to perhaps unnecessary image reconstruction as an intermediate step. Indeed, as we move towards artificial intelligence based task the field is still undecided if

the reconstructed image is required at all for state-of-the-art performance in classification, diagnosis, or treatment prediction as it becomes less important to have a spatio-temporal image representation (i.e. the domain that humans are familiar with). However, one cannot deny the human reliance on visual evidence. Removing the image reconstruction task then can lead to lower interpretability, and perhaps makes it harder for clinicians to trust the outcome. Trust can be regained by increased performance and stability. We observe that from an optimization point of view solving a sequence of two ill-posed problems (reconstruction and task prediction) can be easier compared to solving the entire composite problem at once. When the output of the first ill-posed problem is the input of the second, the search space can be dramatically reduced (Sbalzarini 2016).

Finally, a stumbling block towards adaptation of task-specific training is the lack of standardized datasets. Learning the translation from measurement domain to task domain leads to a field of widely diverse combinations of measurement versus tasks domains, making it less likely to find a standardized challenge dataset offering at least a subset of combinations. The motivation behind task-specific learning can be as simple as a lack of image training data, forcing the practitioner to remove the image reconstruction task altogether by necessity. An contrasting approach to task-specific learning and image-reconstruction aided learning are multi-output or synthesizing networks. For example, consider MRI where the acquisition device can generate T1 or T2 images, or a combination of both, by changing acquisition parameters. A classification network that expects both as input is then limited in application unless an intermediate network synthesizes the 'missing' images. In other words, rather than omitting the reconstruction, synthesizing networks reconstruct a set of images from a single input (Sharma and Hamarneh 2019). One can then generalize this concept to asking what reconstruction space is ideal for the end task, rather than limiting the method to human interpretable reconstruction (images).

6.2 Future directions

The future of the field will be to produce higher quality images given the most limited resource budget such as radiation dose, scanning time, and scanner complexity as well as from online real patient inputs using algorithms with the fewest hand-tuned parameters and lowest power consumption.

There remain a vast range of challenges and opportunities in the field. So far, most approaches focus on CT and MR image reconstruction while only a handful of approaches exist for the reconstruction of the remaining modalities. Hence, the applicability of deep learning models to these problems is yet to be fully explored. In addition, proposed deep learning architectures are often generic and are not fully optimized for specific applications. For instance, how to optimally exploit spatio-temporal redundancy, or how to exploit multi-spectral data. By addressing these core questions and designing network layers to efficiently learn such data representation, the network architecture can gain a boost in performance and reliability.

Joint multi-modal image reconstruction, such as DOT/CT (Baikejiang et al. 2017) and PET/MRI (Wang et al. 2015), has been proposed on several iterative approaches to take advantage of both imaging modalities and led to improving the overall imaging performance, especially avoiding the spatio-temporal artifacts due to the scanning with different devices at different times and positions. The idea relies on leveraging the information provided by feature similarity between multiple modalities. While this direction has great interest, it has only just begun to receive consideration (Cui et al. 2019) and

remains a direction for further exploration. Collecting suitably calibrated and registered data on hybrid multi-modality imaging systems remains a key challenge as well.

Attention driven sampling received increasing interest recently especially in a limited data representation context. While adapting the sampling to the reconstruction algorithm showed improved image quality compared to conventional sampling strategies (Jin et al. 2019), it could be computationally more expensive with unknown convergence behavior. The development of efficient sampling learning algorithms would be a promising research direction. In optical coherence tomography the application of wavelet based compressive sensing has been shown capable to reconstruct with as little as 20% of the samples (Lebed et al. 2013). The potential demonstrated through handcrafted sampling strategies indicates that DL-based methods could also exploit this opportunity.

Deep learning models require computationally powerful hardware (GPU) to provide online reconstruction and achieve state of the art performance. Network pruning and sparsifying, recently proposed for computer vision tasks (Alford et al. 2018; Aghasi et al. 2017; Huang and Wang 2018), presents a promising direction yet to be explored on image reconstruction tasks in order to allow DL-based model processing on CPU and mobile devices. This will be of great interest to emergent mobile scanners (e.g., DOT (Shokoufi and Golnaraghi 2016), US (Georgeson and Safai 2017), CT (Rykkje et al. 2019)).

Reducing the encoding size in bits of network weights without altering the quality of the prediction has been demonstrated on several references deep learning image processing networks Sun et al. (2019b). A reported decrease in training time of 30-60% is only one benefit. By reducing network weights from 64 or 32 bits to 8 or smaller (i.e., weights quantization) the network requires a smaller memory footprint. As a result, networks that are at the time of writing too large to fit on a single GPU can be reduced to fit on a single GPU. Conversely, networks too large to deploy on edge devices (handheld scanner, mobile phones, etc.) can become easily deployable in the field without changing their architecture. Finally and not least important, the reduction in training time results in a significant reduction in the environmental impact of the training procedure.

As performance evaluation metrics are more computer vision task-oriented, they tend to be insufficient in the biomedical imaging context as they do not provide real diagnostic accuracy significance. We advocate a shift toward task-oriented evaluations given that reconstructed images are usually used for a specific diagnosis or treatment purpose. While such a measure may be expensive, especially if they require human expert feedback, they will be critical in creating algorithms that can advance the biomedical imaging practice.

The application of novel DL approaches to medical image reconstruction is spanning an increasingly large number of modalities. A clear trend is present in recent art moving towards unsupervised approaches, inspired by a constraint in realistic or real world training data. An extension of this principle is illustrated by task-specific training where the reconstruction task is omitted altogether in favor of a singular focus on the actual end task such as classification or diagnosis. Related to the sparsity of manually annotated training data is generalizability with works focusing on sacrificing challenge dataset performance in favor of generalizability. Hybrid learning, the concept of reconstructing an image using a secondary modality as prior shows promise in this regard as well. With interpretability in DL still a hot topic we observe the beginning of a trend that moves away from interpretability by omitting any intermediate results as showcased in task-specific learning. The potential for reduced trust by the end user is offset by 'trust by results': if a diagnosis is more accurate without a reconstructed image, is the image needed at all? Novel concepts

such as federated learning and homomorphic encryption to protect distributed clinical data are quickly finding adoption in the field.

References

Adler J, Öktem O (2018) Learned primal-dual reconstruction. IEEE Trans Med Imaging 37(6):1322–1332

Aghasi A, Abdi A, Nguyen N, Romberg J (2017) Net-trim: convex pruning of deep neural networks with performance guarantee. In: Advances in neural information processing systems, pp 3177–3186

Alford S, Robinett R, Milechin L, Kepner J (2018) Pruned and structurally sparse neural networks. arXiv preprint arXiv:1810.00299

Alom MZ, Taha TM, Yakopcic C, Westberg S, Sidike P, Nasrin MS, Hasan M, Van Essen BC, Awwal AA, Asari VK (2019) A state-of-the-art survey on deep learning theory and architectures. Electronics 8(3):292

Al-Shakhrah I, Al-Obaidi T (2003) Common artifacts in computerized tomography: a review. Appl Radiol 32(8):25–32

Antun V, Renna F, Poon C, Adcock B, Hansen AC (2019) On instabilities of deep learning in image reconstruction-does AI come at a cost? arXiv preprint arXiv:1902.05300

Arridge S, Maass P, Öktem O, Schönlieb C-B (2019) Solving inverse problems using data-driven models. Acta Numerica 28:1–174

Assili S (2018) A review of tomographic rconstruction techniques for computed tomography. arXiv preprint arXiv:1808.09172

Baikejiang R, Zhang W, Li C (2017) Diffuse optical tomography for breast cancer imaging guided by computed tomography: a feasibility study. J X-ray Sci Technol 25(3):341–355

Beattie B (2018) Improvements in the robustness and accuracy of bioluminescence tomographic reconstructions of distributed sources within small animals. PhD thesis, Columbia University

Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imaging Sci 2(1):183–202

Bennett L, Simon W (2013) Cardiac atlas project standard challenge-MICCAI 2013 grand challenge, 2013. http://masiweb.vuse.vanderbilt.edu/workshop2013/index.php/

Bhadra S, Zhou W, Anastasio MA (2020) Medical image reconstruction with image-adaptive priors learned by use of generative adversarial networks. In: Medical imaging 2020: physics of medical imaging, vol 11312, p 113120V. International Society for Optics and Photonics

Biobank U (2014) About uk biobank. https://www.ukbiobank.ac.uk/about-biobank-uk. Accessed 17 June 2020.

Boublil D, Elad M, Shtok J, Zibulevsky M (2015) Spatially-adaptive reconstruction in computed tomography using neural networks. IEEE Trans Med Imaging 34(7):1474–1485

Boyd S, Parikh N, Chu E, Peleato B, Eckstein J et al (2011) Distributed optimization and statistical learning via the alternating direction method of multipliers. Found Trends® Mach Learn 3(1):1–122

Boyd N, Jonas E, Babcock HP, Recht B (2018) Deeploco: fast 3D localization microscopy using neural networks. BioRxiv 267096

Braun H, Turaga P, Spanias A, Katoch S, Jayasuriya S, Tepedelenlioglu C (2019) Reconstruction-free compressive vision for surveillance applications. Synth Lect Signal Process 14(1):1–100

Cai C, Deng K, Ma C, Luo J (2018) End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging. Opt Lett 43(12):2752–2755

Cardoen B, Yedder HB, Sharma A, Chou KC, Nabi IR, Hamarneh G (2019) ERGO: efficient recurrent graph optimized emitter density estimation in single molecule localization microscopy. IEEE Trans Med Imaging, 39(6):1942–1956

Chambolle A, Pock T (2011a) A first-order primal-dual algorithm for convex problems with applications to imaging. J Math Imaging Vis 40(1):120–145

Chambolle A, Pock T (2011b) A first-order primal-dual algorithm for convex problems with applications to imaging. J Math Imaging Vis 40(1):120–145

Chen Y, Yang Z, Hu Y, Yang G, Zhu Y, Li Y, Chen W, Toumoulin C et al (2012) Thoracic low-dose CT image processing using an artifact suppressed large-scale nonlocal means. Phys Med Biol 57(9):2667

Chen H, Zhang Y, Kalra MK, Lin F, Chen Y, Liao P, Zhou J, Wang G (2017a) Low-dose CT with a residual encoder-decoder convolutional neural network. IEEE Trans Med Imaging 36(12):2524–2535

Chen H, Zhang Y, Zhang W, Liao P, Li K, Zhou J, Wang G (2017b) Low-dose CT via convolutional neural network. Biomed Opt Express 8(2):679–694

Cheng JY, Chen F, Alley MT, Pauly JM, Vasanawala SS (2018) Highly scalable image reconstruction using deep neural networks with bandpass filtering. arXiv preprint arXiv:1805.03300

Cui J, Gong K, Guo N, Kim K, Liu H, Li Q (2019) CT-guided PET parametric image reconstruction using deep neural network without prior training data. In: Medical imaging 2019: physics of medical imaging, vol 10948, p 109480Z. International Society for Optics and Photonics

Dabov K, Foi A, Katkovnik V, Egiazarian K (2007) Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans Image Process 16(8):2080–2095

Daubechies I, Defrise M, De Mol C (2004) An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Commun Pure Appl Math J Issued Courant Inst Math Sci 57(11):1413–1457

Despres P, Jia X (2017) A review of GPU-based medical image reconstruction. Physica Medica 42:76–92

Dong B, Shen Z (2015) Image restoration: a data-driven perspective. In: Proceedings of the international congress of industrial and applied mathematics (ICIAM), pp 65–108. Citeseer

Fan Q, Witzel T, Nummenmaa A, Van Dijk KR, Van Horn JD, Drews MK, Somerville LH, Sheridan MA, Santillana RM, Snyder J et al (2016) MGH-USC human connectome project datasets with ultra-high b-value diffusion MRI. Neuroimage 124:1108–1114

Feng J, Sun Q, Li Z, Sun Z, Jia K (2018) Back-propagation neural network-based reconstruction algorithm for diffuse optical tomography. J Biomed Opt 24(5):051407

Fessler JA (2017) Medical image reconstruction: a brief overview of past milestones and future directions. arXiv preprint arXiv:1707.05927

Gabay D, Mercier B (1976) A dual algorithm for the solution of nonlinear variational problems via finite element approximation. Comput Math Appl 2(1):17–40

Gao H, Yu H, Osher S, Wang G (2011) Multi-energy CT based on a prior rank, intensity and sparsity model (prism). Inverse Probl 27(11):115012

Gates AJ, Ahn Y-Y (2017) The impact of random models on clustering similarity. J Mach Learn Res 18(1):3049–3076

Geffrin J-M, Sabouroux P, Eyraud C (2005) Free space experimental scattering database continuation: experimental set-up and measurement precision. Inverse Probl 21(6):S117

Georgeson G, Safai M (2017) Portable X-ray backscattering imaging system including a radioactive source, May 23. US Patent 9,658,173

Geyer RC, Klein T, Nabi M (2017) Differentially private federated learning: a client level perspective. arXiv preprint arXiv:1712.07557

Gilad-Bachrach R, Dowlin N, Laine K, Lauter K, Naehrig M, Wernsing J (2016) Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy. In: International conference on machine learning, pp 201–210

Goceri E, Goceri N (2017) Deep learning in medical image analysis: recent advances and future trends. International Conferences Computer Graphics, Visualization, Computer Vision and Image Processing (CGVCVIP 2017), Portugal.

Goceri E, Gooya A (2018) On the importance of batch size for deep learning. In: International conference on mathematics. Istanbul, Turkey

Gong K, Catana C, Qi J, Li Q (2018) Pet image reconstruction using deep image prior. IEEE Trans Med imaging 38(7):1655–1665

Gong K, Catana C, Qi J, Li Q (2019) Direct patlak reconstruction from dynamic pet using unsupervised deep learning. In: 15th International meeting on fully three-dimensional image reconstruction in radiology and nuclear medicine, vol 11072, p 110720R. International Society for Optics and Photonics

Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. In: Advances in neural information processing systems, pp 2672–2680

Gottschling NM, Antun V, Adcock B, Hansen AC (2020) The troublesome kernel: why deep learning for inverse problems is typically unstable. arXiv preprint arXiv:2001.01258

Guo H, Qiu C, Vaswani N (2014) An online algorithm for separating sparse and low-dimensional signal sequences from their sum. IEEE Trans Signal Process 62(16):4284–4297

Guo J, Qi H, Xu Y, Chen Z, Li S, Zhou L (2016) Iterative image reconstruction for limited-angle CT using optimized initial image. Comput Math Methods Med 2016 2016:1–9

Gupta H, Jin KH, Nguyen HQ, McCann MT, Unser M (2018) CNN-based projected gradient descent for consistent CT image reconstruction. IEEE Trans Med Imaging 37(6):1440–1453

Häggström I, Beattie BJ, Schmidtlein CR (2016) Dynamic PET simulator via tomographic emission projection for kinetic modeling and parametric image studies. Med Phys 43(6Part1):3104–3116

Häggström I, Schmidtlein CR, Campanella G, Fuchs TJ (2019) Deeppet: a deep encoder-decoder network for directly solving the PET image reconstruction inverse problem. Med Image Anal 54:253–262

Harrison RL (2010) Monte carlo simulation of emission tomography and other medical imaging techniques. In: AIP conference proceedings, vol 1204, pp 126–132. AIP

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778

Holden S, Sage D (2016) Imaging: super-resolution fight club. Nat Photonics 10(3):152

Hoshi Y, Yamada Y (2016) Overview of diffuse optical tomography and its clinical applications. J Biomed Opt 21(9):091312

Hosseini SAH, Yaman B, Moeller S, Hong M, Akçakaya M (2019) Dense recurrent neural networks for inverse problems: History-cognizant unrolling of optimization algorithms. arXiv preprint arXiv :1912.07197

Huang Y, Preuhs A, Lauritsch G, Manhart M, Huang X, Maier A (2019b) Data consistent artifact reduction for limited angle tomography with deep learning prior. In: International workshop on machine learning for medical image reconstruction, pp 101–112. Springer

Huang Z, Wang N (2018) Data-driven sparse structure selection for deep neural networks. In: Proceedings of the European conference on computer vision (ECCV), pp 304–320

Huang Q, Yang D, Yi J, Axel L, Metaxas D (2019a) FR-Net: joint reconstruction and segmentation in compressed sensing cardiac MRI. In: International conference on functional imaging and modeling of the heart, pp 352–360. Springer

Hyun CM, Kim HP, Lee SM, Lee S, Seo JK (2018) Deep learning for undersampled MRI reconstruction. Phys Med Biol 63(13):135007

Jhamb TK, Rejathalal V, Govindan V (2015) A review on image reconstruction through MRI k-space data. Int J Image Graph Signal Process 7(7):42

Jin KH, Unser M, Yi KM (2019) Self-supervised deep active accelerated MRI. arXiv preprint arXiv :1901.04547

Jin KH, McCann MT, Froustey E, Unser M (2017) Deep convolutional neural network for inverse problems in imaging. IEEE Trans Image Process 26(9):4509–4522

Kang E, Koo HJ, Yang DH, Seo JB, Ye JC (2019) Cycle-consistent adversarial denoising network for multiphase coronary CT angiography. Med Phys 46(2):550–562

Kearney V, Ziemer BP, Perry A, Wang T, Chan JW, Ma L, Morin O, Yom SS, Solberg TD (2020) Attention-aware discrimination for MR-to-CT image translation using cycle-consistent generative adversarial networks. Radiol Artif Intell 2(2):e190027

Khan A, Sohail A, Zahoora U, Qureshi AS (2020) A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev. https://doi.org/10.1007/s10462-020-09825-6

Knoll F, Hammernik K, Zhang C, Moeller S, Pock T, Sodickson DK, Akcakaya M (2020) Deep-learning methods for parallel magnetic resonance imaging reconstruction: a survey of the current approaches, trends, and issues. IEEE Signal Process Mag 37(1):128–140

Lan H, Zhou K, Yang C, Cheng J, Liu J, Gao S, Gao F (2019) Ki-gan: knowledge infusion generative adversarial network for photoacoustic image reconstruction in vivo. In: International conference on medical image computing and computer-assisted intervention, pp 273–281. Springer

Latorre-Carmona P, Traver VJ, Sánchez JS, Tajahuerce E (2019) Online reconstruction-free single-pixel image classification. Image Vis Comput 86:28–37

Lebed E, Lee S, Sarunic MV, Beg MF (2013) Rapid radial optical coherence tomography image acquisition. J Biomed Opt 18(3):036004

Li D, Du C, He H (2020a) Semi-supervised cross-modal image generation with generative adversarial networks. Pattern Recogn 100:107085

Liang Z-P (2007) Spatiotemporal imaging with partially separable functions. In: 2007 4th IEEE international symposium on biomedical imaging: from nano to macro, pp 988–991. IEEE

Liang D, Cheng J, Ke Z, Ying L (2020) Deep magnetic resonance image reconstruction: inverse problems meet neural networks. IEEE Signal Process Mag 37(1):141–151

Liang K, Yang H, Kang K, Xing Y (2018) Improve angular resolution for sparse-view CT with residual convolutional neural network. In: Medical imaging 2018: physics of medical imaging, vol 10573, p 105731K. International Society for Optics and Photonics

Li D, Li S, Zhu M, Gao Q, Bian Z, Huang H, Zhang S, Huang J, Zeng D, Ma J (2020b) Unsupervised data fidelity enhancement network for spectral ct reconstruction. In: Medical imaging 2020: physics of medical imaging, vol 11312, p 113124D. International Society for Optics and Photonics

Lin Z (2016) A review on low-rank models in data analysis. Big Data Inf Anal 1(2&3):139–161

LowDoseCT (2014) Low dose CT grand challenge. https://www.aapm.org/GrandChallenge/LowDoseC/

Lucas A, Iliadis M, Molina R, Katsaggelos AK (2018) Using deep neural networks for inverse problems in imaging: beyond analytical methods. IEEE Signal Process Mag 35(1):20–36

Lung Cancer Alliance. Give a scan, Fact Sheet No. 282, 2017. http://www.giveascan.org. Accessed 20 Feb 2020

Luo G, Zhao N, Jiang W, Cao P (2019) MRI reconstruction using deep bayesian inference. arXiv preprint arXiv:1909.01127

McCann MT Unser M (2019) Algorithms for biomedical image reconstruction. arXiv preprint arXiv:1901.03565

McCann MT, Jin KH, Unser M (2017) Convolutional neural networks for inverse problems in imaging: a review. IEEE Signal Process Mag 34(6):85–95

Mendrik AM, Vincken KL, Kuijf HJ, Breeuwer M, Bouvy WH, De Bresser J, Alansary A, De Bruijne M, Carass A, El-Baz A et al (2015) MRbrains challenge: online evaluation framework for brain image segmentation in 3T MRI scans. Comput Intell Neurosci 2015:1

Meng M, Li S, Yao L, Li D, Zhu M, Gao Q, Xie Q, Zhao Q, Bian Z, Huang J et al (2020) Semi-supervised learned sinogram restoration network for low-dose ct image reconstruction. In: Medical imaging 2020: physics of medical imaging, vol 11312, p 113120B. International Society for Optics and Photonics

Moslemi V, Erfanian V, Ashoor M (2020) Estimation of optimized timely system matrix with improved image quality in iterative reconstruction algorithm: a simulation study. Heliyon 6(1):e03279

Mridata. mridata.org. http://mridata.org. Accessed 30 Nov 2019

Nehme E, Weiss LE, Michaeli T, Shechtman Y (2018) Deep-storm: super-resolution single-molecule microscopy by deep learning. Optica 5(4):458–464

Nesterov YE (1983) A method for solving the convex programming problem with convergence rate o $(1/k^2)$. Dokl Akad Nauk SSSR 269:543–547

Oksuz I, Clough J, Ruijsink B, Puyol-Antón E, Bustin A, Cruz G, Prieto C, Rueckert D, King AP, Schnabel JA (2019) Detection and correction of cardiac MRI motion artefacts during reconstruction from k-space. In: International conference on medical image computing and computer-assisted intervention, pp 695–703. Springer

Otazo R, Candès E, Sodickson DK (2015) Low-rank plus sparse matrix decomposition for accelerated dynamic mri with separation of background and dynamic components. Magn Reson Med 73(3):1125–1136

Ouyang J, Chen KT, Gong E, Pauly J, Zaharchuk G (2019) Ultra-low-dose PET reconstruction using generative adversarial network with feature matching and task-specific perceptual loss. Med Phys 46(8):3555–3564

Pap G, Lékó G, Grósz T (2019) A reconstruction-free projection selection procedure for binary tomography using convolutional neural networks. In: International conference on image analysis and recognition, pp 228–236. Springer

Paul G, Cardinale J, Sbalzarini IF (2013) Coupling image restoration and segmentation: a generalized linear model/bregman perspective. Int J Comput Vis 104(1):69–93

Pelt DM, Batenburg KJ (2013) Fast tomographic reconstruction from limited data using artificial neural networks. IEEE Trans Image Process 22(12):5238–5251

Qin C, Schlemper J, Caballero J, Price AN, Hajnal JV, Rueckert D (2018) Convolutional recurrent neural networks for dynamic MR image reconstruction. IEEE Trans Med Imaging 38(1):280–290

Rajagopal A, Stier N, Dey J, King MA, Chandrasekaran S (2019) Towards deep iterative-reconstruction algorithms for computed tomography (CT) applications. In: Medical imaging 2019: physics of medical imaging, vol 10948, p 1094856. International Society for Optics and Photonics

Ravishankar S, Ye JC, Fessler JA (2019) Image reconstruction: From sparsity to data-adaptive methods and machine learning. arXiv preprint arXiv:1904.02816

Rivest RL, Adleman L, Dertouzos ML et al (1978) On data banks and privacy homomorphisms. Found Secure Comput 4(11):169–180

Rodríguez P (2013) Total variation regularization algorithms for images corrupted with different noise models: a review. J Electr Comput Eng 2013:10

Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention, pp 234–241. Springer

Rykkje A, Carlsen JF, Nielsen MB (2019) Hand-held ultrasound devices compared with high-end ultrasound systems: a systematic review. Diagnostics 9(2):61

Sandino CM, Cheng JY, Chen F, Mardani M, Pauly JM, Vasanawala SS (2020) Compressed sensing: From research to clinical practice with deep neural networks: shortening scan times for magnetic resonance imaging. IEEE Signal Process Mag 37(1):117–127

Sbalzarini IF (2016) Seeing is believing: quantifying is convincing: computational image analysis in biology. In: Focus on bio-image informatics, pp 1–39. Springer

Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D (2018) A deep cascade of convolutional neural networks for dynamic MR image reconstruction. IEEE Trans Med Imaging 37(2):491–503

Schlemper J, Caballero J, Hajnal JV, Price A, Rueckert D (2017) A deep cascade of convolutional neural networks for MR image reconstruction. In: International conference on information processing in medical imaging, pp 647–658. Springer

Schweiger M, Arridge SR (2014) The toast++ software suite for forward and inverse modeling in optical tomography. J Biomed Opt 19(4):040801

Sharma A, Hamarneh G (2019) Missing MRI pulse sequence synthesis using multi-modal generative adversarial network. IEEE Trans Med Imaging. 39(4):1170–1183

Shen G, Dwivedi K, Majima K, Horikawa T, Kamitani Y (2019) End-to-end deep image reconstruction from human brain activity. Front Comput Neurosci 13:21

Shokoufi M, Golnaraghi F (2016) Development of a handheld diffuse optical breast cancer assessment probe. J Innov Opt Health Sci 9(02):1650007

Singh R, Digumarthy SR, Muse VV, Kambadakone AR, Blake MA, Tabari A, Hoi Y, Akino N, Angel E, Madan R et al (2020) Image quality and lesion detection on deep learning reconstruction and iterative reconstruction of submillisievert chest and abdominal CT. Am J Roentgenol. 214(3):566–573.

Singh V, Tewfik AH, Ress DB (2015) Under-sampled functional MRI using low-rank plus sparse matrix decomposition. In: 2015 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp 897–901. IEEE

St-Yves, G, Naselaris T (2018) Generative adversarial networks conditioned on brain activity reconstruct seen images. In: 2018 IEEE international conference on systems, man, and cybernetics (SMC), pp 1054–1061. IEEE

Sun Y, Xia Z, Kamilov US (2018) Efficient and accurate inversion of multiple scattering with deep learning. Opt Express 26(11):14678–14688

Sun Y, Wohlberg B, Kamilov US (2019c) An online plug-and-play algorithm for regularized image reconstruction. IEEE Trans Comput Imaging 5(3):395–408

Sun X, Choi J, Chen C-Y, Wang N, Venkataramani S, Srinivasan VV, Cui X, Zhang W, Gopalakrishnan K (2019b) Hybrid 8-bit floating point (HFP8) training and inference for deep neural networks. In: Advances in neural information processing systems, pp 4901–4910

Sun L, Fan Z, Ding X, Huang Y, Paisley J (2019a) Joint CS-MRI reconstruction and segmentation with a unified deep network. In: International conference on information processing in medical imaging, pp 492–504. Springer

Sun Y, Kamilov US (2018) Stability of scattering decoder for nonlinear diffractive imaging. arXiv preprint arXiv:1806.08015

Sun J, Li H, Xu Z et al (2016) Deep ADMM-Net for compressive sensing MRI. In: Advances in neural information processing systems, pp 10–18

Thaler F, Hammernik K, Payer C, Urschler M, Štern D (2018) Sparse-view CT reconstruction using wasserstein GANs. In: International workshop on machine learning for medical image reconstruction, pp 75–82. Springer

Ulyanov D, Vedaldi A, Lempitsky V(2018) Deep image prior. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 9446–9454

Vandenberghe S, D'Asseler Y, Van de Walle R, Kauppinen T, Koole M, Bouwens L, Van Laere K, Lemahieu I, Dierckx R (2001) Iterative reconstruction algorithms in nuclear medicine. Comput Med Imaging Graph 25(2):105–111

Waibel D, Gröhl J, Isensee F, Kirchner T, Maier-Hein K, Maier-Hein L (2018) Reconstruction of initial pressure from limited view photoacoustic images using deep learning. In: Photons plus ultrasound: imaging and sensing 2018, vol 10494, p 104942S. International Society for Optics and Photonics

Wang G (2016) A perspective on deep imaging. IEEE Access 4:8914–8924

Wang G, Kalra M, Murugan V, Xi Y, Gjesteby L, Getzin M, Yang Q, Cong W, Vannier M (2015) Vision 20/20: simultaneous CT-MRI—next chapter of multimodality imaging. Med Phys 42(10):5879–5889

Wang G, Ye JC, Mueller K, Fessler JA (2018) Image reconstruction is a new frontier of machine learning. IEEE Trans Med Imaging 37(6):1289–1296

Wang H, Wu N, Cai Y, Ren L, Zhao Z, Han G, Wang J (2019) Optimization of reconstruction accuracy of anomaly position based on stacked auto-encoder neural networks. IEEE Access 7:116578–116584

Wang Z, Simoncelli EP, Bovik AC (2003) Multiscale structural similarity for image quality assessment. In: The thrity-seventh asilomar conference on signals, systems & computers, vol 2, pp 1398–1402. IEEE

Wang S, Su Z, Ying L, Peng X, Zhu S, Liang F, Feng D, Liang D (2016) Accelerating magnetic resonance imaging via deep learning. In: 2016 IEEE 13th international symposium on biomedical imaging (ISBI), pp 514–517. IEEE

Webb A, Kagadis GC (2003) Introduction to biomedical imaging. Med Phys 30(8):2267–2267

Wells WM III, Viola P, Atsumi H, Nakajima S, Kikinis R (1996) Multi-modal volume registration by maximization of mutual information. Med Image Anal 1(1):35–51

Wen B, Ravishankar S, Pfister L, Bresler Y (2019) Transform learning for magnetic resonance image reconstruction: from model-based learning to building neural networks. arXiv preprint arXiv:1903.11431

Wolterink JM, Leiner T, Viergever MA, Išgum I (2017) Generative adversarial networks for noise reduction in low-dose CT. IEEE Trans Med Imaging 36(12):2536–2545

Wu D, Kim K, Li Q (2019) Computationally efficient deep neural network for computed tomography image reconstruction. Med Phys 46(11):4763–4776

Wu S, Gao Z, Liu Z, Luo J, Zhang H, Li S (2018) Direct reconstruction of ultrasound elastography using an end-to-end deep neural network. In: International conference on medical image computing and computer-assisted intervention, pp 374–382. Springer

Würfl T, Hoffmann M, Christlein V, Breininger K, Huang Y, Unberath M, Maier AK (2018) Deep learning computed tomography: learning projection-domain weights from image domain in limited angle problems. IEEE Trans Med Imaging 37(6):1454–1463

Xu J, Gong E, Pauly J, Zaharchuk G (2017) 200x low-dose PET reconstruction using deep learning. arXiv preprint arXiv:1712.04119

Xu Q, Yu H, Mou X, Zhang L, Hsieh J, Wang G (2012) Low-dose X-ray CT reconstruction via dictionary learning. IEEE Trans Med Imaging 31(9):1682–1697

Yang G, Yu S, Dong H, Slabaugh G, Dragotti PL, Ye X, Liu F, Arridge S, Keegan J, Guo Y et al (2017a) DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. IEEE Transactions on Medical Imaging 37(6):1310–1321

Yang B, Ying L, Tang J (2018) Artificial neural network enhanced bayesian pet image reconstruction. IEEE Trans Med imaging 37(6):1297–1309

Yang Y, Sun J, Li H, Xu Z (2017b) ADMM-Net: a deep learning approach for compressive sensing MRI. CoRR, arXiv:abs/1705.06869

Yedder HB, BenTaieb A, Shokoufi M, Zahiremami A, Golnaraghi F, Hamarneh G (2018) Deep learning based image reconstruction for diffuse optical tomography. In: International workshop on machine learning for medical image reconstruction, pp 112–119. Springer

Yedder HB, Shokoufi M, Cardoen B, Golnaraghi F, Hamarneh G (2019) Limited-angle diffuse optical tomography image reconstruction using deep learning. In: International conference on medical image computing and computer-assisted intervention, pp 66–74. Springer

Yoon YH, Khan S, Huh J, Ye JC (2018) Efficient b-mode ultrasound image reconstruction from sub-sampled RF data using deep learning. IEEE Trans Med Imaging 38(2):325–336

Yoo J, Sabir S, Heo D, Kim KH, Wahab A, Choi Y, Lee S-I, Chae EY, Kim HH, Bae YM et al. (2020) Deep learning diffuse optical tomography. IEEE Trans Med Imaging. 39(4):877–887

Zbontar J, Knoll F, Sriram A, Muckley MJ, Bruno M, Defazio A, Parente M, Geras KJ, Katsnelson J, Chandarana H et al (2018) FastMRI: an open dataset and benchmarks for accelerated MRI. arXiv preprint arXiv:1811.08839

Zhang Y, Wang Y, Zhang W, Lin F, Pu Y, Zhou J (2016a) Statistical iterative reconstruction using adaptive fractional order regularization. Biomed Opt Express 7(3):1015–1029

Zhang Y, Xi Y, Yang Q, Cong W, Zhou J, Wang G (2016b) Spectral CT reconstruction with image sparsity and spectral mean. IEEE Trans Comput Imaging 2(4):510–523

Zhang H, Zeng D, Zhang H, Wang J, Liang Z, Ma J (2017) Applications of nonlocal means algorithm in low-dose X-ray CT image processing and reconstruction: a review. Med Phys 44(3):1168–1185

Zhang H, Dong B (2019) A review on deep learning in medical image reconstruction. arXiv preprint arXiv:1906.10643

Zhang Q, Liang D (2020) Visualization of fully connected layer weights in deep learning ct reconstruction. arXiv preprint arXiv:2002.06788

Zhao B, Haldar JP, Christodoulou AG, Liang Z-P (2012) Image reconstruction from highly undersampled (k, t)-space data with joint partial separability and sparsity constraints. IEEE Trans Med Imaging 31(9):1809–1820

Zhou B, Lin X, Eck B (2019) Limited angle tomography reconstruction: synthetic reconstruction via unsupervised sinogram adaptation. In: International conference on information processing in medical imaging, pp 141–152. Springer

Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS (2018) Image reconstruction by domain-transform manifold learning. Nature 555(7697):487

Zhu L, Liu Z, Han S (2019) Deep leakage from gradients. In: Advances in neural information processing systems, pp 14747–14756