

# A review and experimental evaluation of deep learning methods for MRI reconstruction

Arghya Pal

Department of Psychiatry and Radiology, Harvard Medical School, Boston, MA, USA

apal5@bwh.harvard.edu

Yogesh Rathi

Department of Psychiatry and Radiology, Harvard Medical School, Boston, MA, USA

yogesh@bwh.harvard.edu

## Abstract

Following the success of deep learning in a wide range of applications, neural network-based machine-learning techniques have received significant interest for accelerating magnetic resonance imaging (MRI) acquisition and reconstruction strategies. A number of ideas inspired by deep learning techniques for computer vision and image processing have been successfully applied to nonlinear image reconstruction in the spirit of compressed sensing for accelerated MRI. Given the rapidly growing nature of the field, it is imperative to consolidate and summarize the large number of deep learning methods that have been reported in the literature, to obtain a better understanding of the field in general. This article provides an overview of the recent developments in neural-network based approaches that have been proposed specifically for improving parallel imaging. A general background and introduction to parallel MRI is also given from a classical view of k-space based reconstruction methods. Image domain based techniques that introduce improved regularizers are covered along with k-space based methods which focus on better interpolation strategies using neural networks. While the field is rapidly evolving with plenty of papers published each year, in this review, we attempt to cover broad categories of methods that have shown good performance on publicly available data sets. Limitations and open problems are also discussed and recent efforts for producing open data sets and benchmarks for the community are examined.

**Keywords:** MRI Reconstruction, Deep Learning, machine learning, k-space reconstruction, parallel MRI

## 1. Introduction

Magnetic Resonance Imaging (MRI) is an indispensable clinical and research tool used to diagnose and study several diseases of the human body. It has become a *sine qua non* in various fields of radiology, medicine, and psychiatry. Unlike computed tomography (CT), it can provide detailed images of the soft tissue and does not require any radiation, thus making it less risky to the subjects. MRI scanners sample a patient's anatomy in the frequency domain that we will call "k-space". The number of rows/columns that are acquired in k-space is proportional to the quality (and spatial resolution) of the reconstructed MR image. To get higher spatial resolution, longer scan time is required due to the increased number of k-space points that need to be sampled (Fessler, 2010). Hence, the subject has to stay still in the MRI scanner for the duration of the scan to avoid signal drops and motion artifacts. Many researchers have been trying to reduce the number of k-space lines

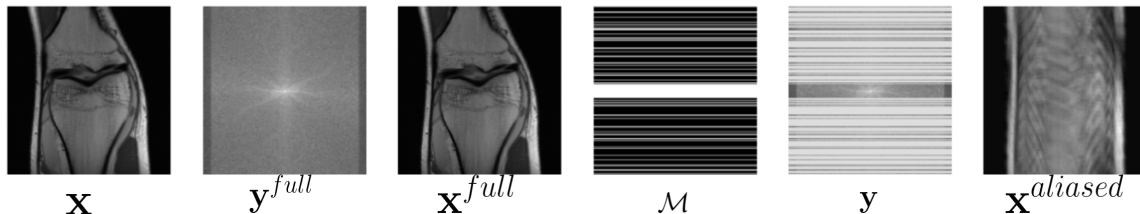


Figure 1: **(Left to Right)** An example of a fastMRI (Zbontar et al., 2018) knee image  $\mathbf{x}$ , fully sampled k-space  $\mathbf{y}^{full}$ , corresponding reconstructed image  $\mathbf{x}^{full}$  from  $\mathbf{y}^{full}$ , the sampling mask  $\mathcal{M}$  of fastMRI that we apply on fully sampled k-space  $\mathbf{y}^{full}$ , the sampled k-space  $\mathbf{y}$ , and the corresponding aliased reconstructed image  $\mathbf{x}^{aliased}$ .

to save scanning time, which leads to the well-known problem of “aliasing” as a result of the violation of the Nyquist sampling criteria (Nyquist, 1928). Reconstructing high-resolution MR images from the undersampled or corrupted measurements was a primary focus of various sparsity promoting methods, wavelet-based methods, edge-preserving methods, and low-rank based methods. This paper reviews the literature on solving the inverse problem of MR image reconstruction from noisy measurements using Deep Learning (DL) methods, while providing a brief introduction to classical optimization based methods. We shall discuss more about this in Sec. 1.1.

A DL method learns a non-linear function  $f : \mathcal{Y} \rightarrow \mathcal{X}$  from a set of all possible mapping functions  $\mathcal{F}$ . The accuracy of the mapping function can be measured using some notion of a loss function  $l : \mathcal{Y} \times \mathcal{X} \rightarrow [0, \infty)$ . The empirical risk (Vapnik, 1991),  $\hat{L}(f)$ , can be estimated as  $\hat{L}(f) = \frac{1}{2} \sum_{i=1}^m l(f(\mathbf{y}_i), \mathbf{x}_i)$  and the generalization error of a mapping function  $f(\cdot)$  can be measured using some notion of accuracy measurement. MR image reconstruction using deep learning, in its simplest form, amounts to learning a map  $f$  from the undersampled k-space measurement  $\mathcal{Y} \in \mathbb{C}^{N_1 \times N_2}$ , or  $\mathcal{Y} \in \mathbb{R}^{N_1 \times N_2 \times 2}$  to an unaliased MR image  $\mathcal{X} \in \mathbb{C}^{N_1 \times N_2}$ , or  $\mathcal{X} \in \mathbb{R}^{N_1 \times N_2 \times 2}$ , where  $N_1$ ,  $N_2$  are the height and width of the complex valued image. In several real-world cases, higher dimensions such as time, volume, etc., are obtained and accordingly the superscripts of  $\mathcal{Y}$  and  $\mathcal{X}$  change to  $\mathbb{C}^{N_1 \times N_2 \times N_3 \times N_4 \times \dots}$ . For the sake of simplicity, we will use assume  $\mathcal{Y} \in \mathbb{C}^{N_1 \times N_2}$  and  $\mathcal{X} \in \mathbb{C}^{N_1 \times N_2}$ .

In this survey, we focus on two broad aspects of DL methods, i.e. (i) generative models, which are data generation processes capturing the underlying density of data distribution; and (ii) non-generative models, that learn complex feature representations of images intending to learn the inverse mapping from k-space measurements to MR images. Given the availability and relatively broad access to open-source platforms like Github, PyTorch (Paszke et al., 2019), and TensorFlow (Abadi et al., 2015), as well as large curated datasets and high-performance GPUs, deep learning methods are actively being pursued for solving the MR image reconstruction problem with reduced number of samples while avoiding artifacts and boosting the signal-to-noise ratio (SNR).

In Sec. 1.1, we briefly discuss the mathematical formulation that utilizes k-space measurements from multiple receiver coils to reconstruct an MR image. Furthermore, we discuss some challenges of the current reconstruction pipeline and discuss the DL methods (in Sec.

1.2) that have been introduced to address these limitations. We finally discuss the open questions and challenges to deep learning methods for MR reconstruction in sections 2.1, 2.2, and 3.

### 1.1 Mathematical Formulation for Image Reconstruction in Multi-coil MRI

Before discussing undersampling and the associated aliasing problem, let us first discuss the simple case of reconstructing an MR image,  $\mathbf{x} \in \mathbb{C}^{N_1 \times N_2}$ , from a fully sampled k-space measurement,  $\mathbf{y}^{full} \in \mathbb{C}^{N_1 \times N_2}$ , using the Fourier transform  $\mathcal{F}(\cdot)$ :

$$\mathbf{y}^{full} = \mathcal{F}\mathbf{x} + \eta, \quad (1)$$

where  $\eta \sim \mathcal{N}(0, \Sigma)$  is the associated measurement noise typically assumed to have a Gaussian distribution (Virtue and Lustig, 2017) when the k-space measurement is obtained from a single receiver coil.

Modern MR scanners support parallel acquisition using an array of overlapping receiver coils  $n$  modulated by their sensitivities  $S_i$ . So Eqn. 1 changes to:  $\mathbf{y}_i^{full} = \mathcal{F}S_i\mathbf{x} + \eta$ , where  $i = \{1, 2, \dots, n\}$  is the number of receiver coils. We use  $\mathbf{y}_i$  for the undersampled k-space measurement coming from the  $i^{th}$  receiver coil. To speed up the data acquisition process, multiple lines of k-space data (for cartesian sampling) are skipped using a binary sampling mask  $\mathcal{M} \in \mathbb{C}^{N_1 \times N_2}$  that selects a subset of k-space lines from  $\mathbf{y}^{full}$  in the phase encoding direction:

$$\mathbf{y}_i = \mathcal{M} \odot \mathcal{F}S_i\mathbf{x} + \eta. \quad (2)$$

An example of  $\mathbf{y}^{full}$ ,  $\mathbf{y}$ ,  $\mathcal{M}$  is shown in Fig 1.

To estimate the MR image  $\mathbf{x}$  from the measurement, a data fidelity loss function is typically used to ensure that the estimated data is as close to the measurement as possible. A typical loss function is the squared loss, which is minimized to estimate  $\mathbf{x}$ :

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \sum_i \|\mathbf{y}_i - \mathcal{M} \odot \mathcal{F}S_i\mathbf{x}\|_2^2 = \arg \min_{\mathbf{x}} \|\mathbf{y} - A\mathbf{x}\|_2^2. \quad (3)$$

We borrow this particular formulation from (Sriram et al., 2020a; Zheng et al., 2019). This squared loss function is quite convenient if we wish to compute the error gradient during optimization.

However, the formulation in Eqn. 3 is under-determined if data is undersampled and does not have a unique solution. Consequently, a regularizer  $\mathcal{R}(\mathbf{x})$  is typically added to solve such an ill-conditioned cost function:<sup>1</sup>

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \sum_i \lambda_i \mathcal{R}_i(\mathbf{x}). \quad (4)$$

Please note that each  $\mathcal{R}_i(\mathbf{x})$  is a separate regularizer, while the  $\lambda_i$ s are hyperparameters that control the properties of the reconstructed image  $\hat{\mathbf{x}}$  while avoiding over-fitting. Eqn.

1. The regularization term,  $\mathcal{R}(\mathbf{x})$  is related to the prior,  $p(\mathbf{x})$ , of a maximum *a priori* (MAP) estimation of  $\mathbf{x}$ , i.e.  $\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} (-\log p(\mathbf{y}|\mathbf{x}) - \log p(\mathbf{x}))$ . In fact, in Ravishankar et al. (2019) the authors loosely define  $\beta\mathcal{R}(\mathbf{x}) = -\log p(\mathbf{x})$ , which promotes some desirable image properties such as spatial smoothness, sparsity in image space, edge preservation, etc. with a view to get a unique solution.

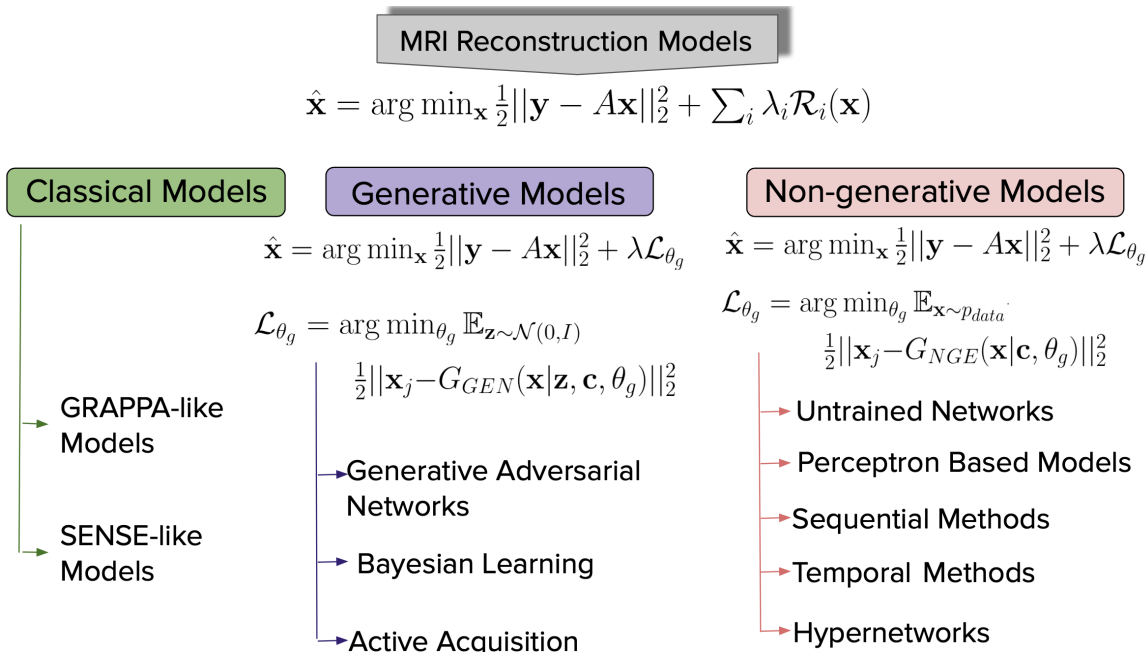


Figure 2: **MRI Reconstruction Methods:** We shall discuss various MR reconstruction methods with a main focus on Deep Learning (DL) based methods. Depending on the optimization function, a DL method can be classified into a Generative (discussed in Sec. 4) or a Non-Generative model (discussed in Sec. 5). However, for the sake of completeness, we shall also discuss classical k-space (GRAPPA-like) and image space based (SENSE-like) methods.

3 along with the regularization term can be optimized using various methods, such as (i) the Morozov formulation,  $\hat{\mathbf{x}} = \min\{\mathcal{R}(\mathbf{x}); \text{such that, } \|A\mathbf{x} - \mathbf{y}\| \leq \delta\}$ ; (ii) the Ivanov formulation, i.e.  $\hat{\mathbf{x}} = \min\{\|A\mathbf{x} - \mathbf{y}\|; \text{such that, } \mathcal{R}(\mathbf{x}) \leq \epsilon\}$ ; or (iii) the Tikonov formulation,  $\hat{\mathbf{x}} = \min\{\|A\mathbf{x} - \mathbf{y}\| + \lambda \mathcal{R}(\mathbf{x})\}$ , discussed in (Oneto et al., 2016).

In general, the Tikonov formulation can be designed using a physics based, sparsity promoting, dictionary learning, or a deep learning based model. But there are several factors that can cause loss in data quality (especially small anatomical details) such as inaccurate modeling of the system noise, complexity, generalizability etc. To overcome these limitations, it is essential to develop inverse mapping methods that not only provide good data fidelity but also generalize well to unseen and unexpected data. In the next section, we shall describe how DL methods can be used as priors or regularizers for MR reconstruction.

## 1.2 Deep Learning Priors for MR Reconstruction

We begin our discussion by considering DL methods with learn-able parameters  $\theta$ . The learn-able parameters  $\theta$  can be trained using some notion of a learning rule that we shall

discuss in Sec. 3. A DL training process helps us to find a function  $G_{DL}(\cdot)$  that acts as a regularizer to Eqn. 4 with an overarching concept of an inverse mapping, i.e; (please note that, we shall follow (Zheng et al., 2019) to develop the optimization formulation)

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda \mathcal{L}_\theta; \text{ where, } \mathcal{L}_\theta = \arg \min_{\theta} \sum_j \|\mathbf{x}_j - G_{DL}(\mathbf{x}|\mathbf{z}, \mathbf{c}, \theta)\|_2^2 \quad (5)$$

and  $\mathbf{z}$  is a latent variable capturing the statistical regularity of the data samples, while  $\mathbf{c}$  is a conditional random variable that depends on a number of factors such as: undersampling of the k-space (Shaul et al., 2020; Oksuz et al., 2019a; Shitrit and Raviv, 2017), the resolution of the image (Yang et al., 2017; Yuan et al., 2020), or the type of DL network used (Lee et al., 2019). Based on the nature of the learning, there are two types of DL methods known as *generative models*, and *non-generative models*. We shall start with a basic understanding of DL methods to a more in-depth study of different architectures in Secs. 4 and 5.

In generative modeling, the random variable  $\mathbf{z}$  is typically sampled from a noisy Gaussian distribution,  $\mathbf{z} \sim \mathcal{N}(0, I)$  with or without the presence of the conditional random variable, i.e.;

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda \mathcal{L}_{\theta_g}; \mathcal{L}_{\theta_g} = \arg \min_{\theta_g} \mathbb{E}_{\mathbf{z} \sim \mathcal{N}(0, I)} \frac{1}{2} \|\mathbf{x}_j - G_{GEN}(\mathbf{x}|\mathbf{z}, \mathbf{c}, \theta_g)\|_2^2 \quad (6)$$

There are various ways to learn the parameters of Eqn. 6. For instance, the Generative Adversarial Network (GAN) (Goodfellow et al., 2014) allows us to learn the generator function  $G_{GEN}(\cdot)$  using an interplay between two modules, while the Variational Autoencoders (VAEs) (Kingma and Welling, 2013) learns  $G_{GEN}(\cdot)$  by optimizing the evidence lower bound (ELBO), or by incorporating a prior in a Bayesian Learning setup as described in Section 4.2. It is shown in the literature that a generative model can efficiently de-alias an MR image that had undergone a  $4\times$  or  $8\times$  undersampling in k-space (Zbontar et al., 2018).

The non-generative models on the other hand do not learn the underlying latent representation, but instead learn a mapping from the measurement space to the image space. Hence, the random variable  $\mathbf{z}$  is not required. The cost function for a non-generative model is given by:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda \mathcal{L}_{\theta_g}; \mathcal{L}_{\theta_g} = \arg \min_{\theta_g} \mathbb{E}_{\mathbf{x} \sim p_{data}} \frac{1}{2} \|\mathbf{x}_j - G_{NGE}(\mathbf{x}|\mathbf{c}, \theta_g)\|_2^2. \quad (7)$$

The function  $G_{NGE}(\cdot)$  is a non-generative mapping function that could be a Convolutional Neural Network (CNN) (Zheng et al., 2019; Akçakaya et al., 2019; Sriram et al., 2020b), a Long Short Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997), or any other similar deep learning model. The non-generative models show a significant improvement in image reconstruction quality compared to classical methods. We shall describe the generative and the non-generative modeling based approaches in detail in Secs. 4 and 5 respectively. Below, we give a brief overview of the classical or the non-DL based approaches for MR image reconstruction.

### 1.3 Classical Methods for MR Reconstruction

In the literature, several approaches can be found that perform an inverse mapping to reconstruct the MR image from k-space data. Starting from analytic methods (Fessler and

Sutton, 2003; Laurette et al., 1996), there are several works that provide ways to do MR reconstruction, such as the physics based image reconstruction methods (Roeloffs et al., 2016; Tran-Gia et al., 2016; Maier et al., 2019; Tran-Gia et al., 2013; Hilbert et al., 2018; Sumpf et al., 2011; Ben-Eliezer et al., 2016; Zimmermann et al., 2017; Schneider et al., 2020), the sparsity promoting compressed sensing methods (Feng and Bresler, 1996; Bresler and Feng, 1996; Candès et al., 2006), and low-rank based approaches (Haldar, 2013). All these methods can be roughly categorized into two categories, i.e. (i) GRAPPA-like methods: where prior assumptions are imposed on the k-space; and (ii) SENSE-like methods: where an image is reconstructed from the k-space while jointly unaliasing (or dealiasing) the image using sparsity promoting and/or edge preserving image regularization terms.

A few k-space methods estimate the missing measurement lines by learning kernels from an already measured set of k-space lines from the center of the k-space (i.e., the auto-calibration or ACS lines). These k-space based methods include methods such as SMASH (Sodickson, 2000), VDAUTOSMASH (Heidemann et al., 2000), GRAPPA and its variations (Bouman and Sauer, 1993; Park et al., 2005; Seiberlich et al., 2008). The k-t GRAPPA (Huang et al., 2005) takes advantage of the correlations in the k-t space and interpolates the missing data. On the other hand, sparsity promoting low rank based methods are based on the assumption that, when the image reconstruction follows a set of constraints (such as sparsity, smoothness, parallel imaging, etc.), the resultant k-space should follow a structure that has low rank. The low rank assumption has been shown to be quite successful in dynamic MRI (Liang, 2007), functional MRI (Singh et al., 2015), and diffusion MRI (Hu et al., 2019). In this paper we give an overview of the low-rank matrix approaches (Haldar, 2013; Jin et al., 2016; Lee et al., 2016; Ongie and Jacob, 2016; Haldar and Zhuo, 2016; Haldar and Kim, 2017) in Sec. 2.1. While in k-t SLR (Lingala et al., 2011), a spatio-temporal total variation norm is used to recover the dynamic signal matrix.

The image space based reconstruction methods, such as, the model based image reconstruction algorithms, incorporate the underlying physics of the imaging system and leverage image priors such as neighborhood information (e.g. total-variation based sparsity, or edge preserving assumptions) during image reconstruction. Another class of works investigated the use of compressed sensing (CS) in MR reconstruction after its huge success in signal processing (Feng and Bresler, 1996; Bresler and Feng, 1996; Candès et al., 2006). Compressed sensing requires incoherent sampling and sparsity in the transform domain (Fourier, Wavelet, Ridgelet or any other basis) for nonlinear image reconstruction. We also describe dictionary learning based approaches that are a special case of compressed sensing using an overcomplete dictionary. The methods described in (Gleichman and Eldar, 2011; Ravishankar and Bresler, 2016; Lingala and Jacob, 2013; Rathi et al., 2011; Michailovich et al., 2011) show various ways to estimate the image and the dictionary from limited measurements.

#### 1.4 Main Highlights of This Literature Survey

The main contributions of this paper are:

- We give a holistic overview of MR reconstruction methods, that includes a family of classical k-space based image reconstruction methods as well as the latest developments using deep learning methods.

- We provide a discussion of the basic DL tools such as activation functions, loss functions, and network architecture, and provide a systematic insight into generative modeling and non-generative modeling based MR image reconstruction methods and discuss the advantages and limitations of each method.
- We compare eleven methods that includes classical, non-DL and DL methods on fastMRI dataset and provide qualitative and quantitative results in Sec. 7.
- We conclude the paper with a discussion on the open issues for the adoption of deep learning methods for MR reconstruction and the potential directions for improving the current state-of-the-art methods.

## 2. Classical Methods for Parallel Imaging

This section reviews some of the classical k-space based MR image reconstruction methods and the classical image space based MR image reconstruction methods.

### 2.1 Inverse Mapping using k-space Interpolation

Classical k-space based reconstruction methods are largely based on the premise that the missing k-space lines can be interpolated (or extrapolated) based on a weighted combination of all acquired k-space measurement lines. For example, in the SMASH (Sodickson, 2000) method, the missing k-space lines are estimated using the spatial harmonics of order  $m$ . The k-space signal can then be written as:

$$\mathbf{y}(k_1, k_2) = \sum_{i=1}^n w_i^m \mathbf{y}_i(k_1, k_2) = \int \int dn_1 dn_2 \sum_{i=1}^n w_i^m S_i \mathbf{x} e^{-jk_1 n_1 - j(k_2 + m \Delta k_2) n_2} \quad (8)$$

where,  $w_i^m$  are the spatial harmonics of order  $m$ ,  $\Delta k_2 = \frac{2\pi}{FOV}$  is the minimum k-space interval ( $FOV$  stands for field-of-view), and  $\mathbf{y}(k_1, k_2)$  is the k-space measurement of an image  $\mathbf{x}$ ,  $(k_1, k_2)$  are the co-ordinates in k-space along the phase encoding (PE) and frequency encoding (FE) directions, and  $j$  represents the imaginary number. From Eqn. 8, one can note that the  $m^{th}$ -line of k-space can be generated using  $m$ -number of spatial harmonics and hence we can estimate convolution kernels to approximate the missing k-space lines from the acquired k-space lines. So, in SMASH a k-space measurement  $\mathbf{x}$ , also known as a composite signal in common parlance, is basically a linear combination of  $n$  number of component signals (k-space measurements) coming from  $n$  receiver coils modulated by their sensitivities  $S_i$ , i.e.

$$\mathbf{y} = \sum_{i=1}^n w_i \mathcal{M} \odot \mathcal{F} S_i \mathbf{x} + \eta. \quad (9)$$

We borrow the mathematical notation from (Sodickson, 2000) and represent the composite signal in Eqn. 9 as follows:

$$\mathbf{y}(k_1, k_2) = \sum_{i=1}^n w_i \mathbf{y}_i(k_1, k_2) = \int \int dn_1 dn_2 \sum_{i=1}^n w_i \mathcal{M} S_i \mathbf{x} e^{-jk_1 n_1 - jk_2 n_2}. \quad (10)$$

However, SMASH requires the exact estimation of the sensitivity of the receiver coils to accurately solve the reconstruction problem.

To address this limitation, AUTO-SMASH (Jakob et al., 1998) assumed the existence of a fully sampled block of k-space lines called *autocalibration lines* (ACS) at the center of the k-space (the low frequency region) and relaxed the requirement of the exact estimation of receiver coil sensitivities. The AUTO-SMASH formulation can be written as:

$$\mathbf{y}(k_1, k_2 + m\Delta k_2) = \sum_{i=1}^n w_i^0 \mathbf{y}_i^{ACS}(k_1, k_2 + m\Delta k_2) = \int \int dn_1 dn_2 \sum_{i=1}^n w_i^0 S_i \mathbf{x} e^{-jk_1 n_1 - j(k_2 + m\Delta k_2) n_2} \quad (11)$$

We note that the AUTO-SMASH paper theoretically proved that it can learn a kernel that is,  $\sum_{i=1}^n w_i^m \mathbf{y}_i(k_1, k_2) = \sum_{i=1}^n w_i^m \mathbf{y}_i^{ACS}(k_1, k_2 + m\Delta k_2)$ , that is, it can learn a linear shift invariant convolutional kernel to interpolate missing k-space lines from the knowledge of the fully sampled k-space lines of ACS region. The variable density AUTO-SMASH (VD-AUTO-SMASH) (Heidemann et al., 2000) further improved the reconstruction process by acquiring multiple ACS lines in the center of the k-space. The composite signal  $\mathbf{x}$  is estimated by adding each individual component  $\mathbf{y}_i$  by deriving linear weights  $w_i^m$  and thereby estimating the missing k-space lines. The more popular, generalized auto calibrating partially parallel acquisitions (GRAPPA) (Bouman and Sauer, 1993) method uses this flavour of VD-AUTO-SMASH, i.e. the shift-invariant linear interpolation relationships in k-space, to learn the coefficients of a convolutional kernel from the ACS lines. The missing k-space are estimated as a linear combination of observed k-space points coming from all receiver coils. The weights of the convolution kernel are estimated as follows: a portion of the k-space lines in the ACS region are artificially masked to get a simulated set of acquired k-space points  $\mathbf{y}^{ACS_1}$  and missing k-space points  $\mathbf{y}^{ACS_2}$ . Using the acquired k-space lines  $\mathbf{y}^{ACS_1}$ , we can estimate the weights of the GRAPPA convolution kernel  $K$  by minimizing the following cost function:

$$\hat{K} = \arg \min_K \|\mathbf{y}^{ACS_2} - K \circledast \mathbf{y}^{ACS_1}\|_2^2 \quad (12)$$

where  $\circledast$  represents the convolution operation. The GRAPPA method has shown very good results for uniform undersampling, and is the method used in product sequences by Siemens and GE scanners. There are also recent methods (Xu et al., 2018; Chang et al., 2012) that show ways to learn non-linear weights of a GRAPPA kernel.

The GRAPPA method regresses k-space lines from a learned kernel without assuming any specific image reconstruction constraint such as sparsity, limited support, or smooth phase as discussed in (Kim et al., 2019). On the other hand, low-rank based methods assume an association between the reconstructed image and the k-space structure, thus implying that the convolution-structured Hankel or Toeplitz matrices leveraged from the k-space measurements must show a distinct null-space vector association with the kernel. As a result, any low-rank recovery algorithm can be used for image reconstruction. The simultaneous autocalibrating and k-space estimation (SAKE) (Shin et al., 2014) algorithm used the block Hankel form of the local neighborhood in k-space across all coils for image reconstruction. Instead of using correlations across multiple coils, the low-rank matrix modeling of local k-space neighborhoods (LORAKS) (Haldar, 2013) utilized the image phase



constraint and finite image support (in image space) to produce very good image reconstruction quality. The LORAKS method does not require any explicit calibration of k-space samples and can work well even if some of the constraints such as sparsity, limited support, and smooth phase are not strictly satisfied. The AC-LORAKS (Haldar, 2015) improved the performance of LORAKS by assuming access to the ACS measurements, i.e.:

$$\mathbf{y} = \arg \min_{\mathbf{y}} \|\mathbf{y} - \mathcal{P}(\mathcal{M}\mathbf{A}\mathbf{x})N\|_2^2 \quad (13)$$

where  $\mathcal{P}(\cdot)$  is a mapping function that transforms the k-space measurement to a structured low-rank matrix, and the matrix  $N$  is the null space matrix. The mapping  $\mathcal{P}(\cdot)$  basically takes care of the constraints such as sparsity, limited support, and smooth phase. In the PRUNO (Zhang et al., 2011) method, the mapping  $\mathcal{P}(\cdot)$  only imposes limited support and parallel imaging constraints. On the other hand, the number of nullspace vectors in  $N$  is set to 1 in the SPIRiT method (Lustig and Pauly, 2010). The ALOHA method (Lee et al., 2016) uses the weighted k-space along with transform-domain sparsity of the image. Different from them, the method of (Otazo et al., 2015) uses a spatio-temporal regularization.

## 2.2 Image Space Rectification based Methods

These methods directly estimate the image from k-space by imposing prior knowledge about the properties of the image (e.g., spatial smoothness). Leveraging image prior through linear interpolation works well in practice but largely suffers from sub-optimal solutions and as a result the practical cone beam algorithm (Laurette et al., 1996) was introduced that improves image quality in such a scenario. The sensitivity encoding (SENSE) method (Pruessmann et al., 1999) is an image unfolding method that unfolds the periodic repetitions from the knowledge of the coil. In SENSE, the signal in a pixel location  $(i, j)$  is a weighted sum of coil sensitivities, i.e.;

$$I_k(i, j) = \sum_{k=1}^{N_c} \sum_{j=1}^{N_2} S_{kj} \mathbf{x}(i, j), \quad (14)$$

where  $N_2$  is the height of image  $\mathbf{x} \in R^{N_1 \times N_2}$ ,  $N_c$  is the number of coils, and  $S_k$  is the coil sensitivity of the  $k^{th}$  coil. The  $I_k$  is the  $k^{th}$  coil image that has aliased pixels at a certain position, and  $i$  is a particular row and  $j = \{1, \dots, N_2\}$  is the column index counting from the top of the image to the bottom. The  $S$  is the sensitivity matrix that assembles the corresponding sensitivity values of the coils at the locations of the involved pixels in the full FOV image  $\mathbf{x}$ . The coil images  $I_k$ , the sensitivity matrix  $S$ , and the image  $\mathbf{x}$  in Eqn. 14 can be re-written as;

$$I = S\mathbf{x}. \quad (15)$$

By knowing the complex sensitivities at the corresponding positions, we can compute the generalized inverse of the sensitivity matrix:

$$\mathbf{x} = (\hat{S}^H \hat{S})^{-1} \hat{S}^H \cdot I. \quad (16)$$

Please note that,  $I$  represents the complex coil image values at the chosen pixel and has length  $N_c$ . In k-t SENSE and k-t BLAST (Tsao et al., 2003) the information about the spatio-temporal support is obtained from the training dataset that helps to reduce aliasing.

The physics based methods allow statistical modeling instead of simple geometric modeling present in classical methods and reconstruct the MR images using the underlying physics of the imaging system (Roeloffs et al., 2016; Tran-Gia et al., 2016; Maier et al., 2019; Tran-Gia et al., 2013; Hilbert et al., 2018; Sumpf et al., 2011; Ben-Eliezer et al., 2016; Zimmermann et al., 2017; Schneider et al., 2020). These types of methods sometimes use very simplistic anatomical knowledge based priors (Chen et al., 1991; Gindi et al., 1993; Cao and Levin, 1997) or “pixel neighborhood” (Szeliski, 2010) information via a Markov Random Field based regularization (Sacco, 1990; Besag, 1986).

A potential function based regularization takes the form  $\mathcal{R}(\mathbf{x}) = \sum_{i=2}^n \psi(\mathbf{x}_i - \mathbf{x}_{i-1})$ , where the function,  $\psi(\cdot)$ , could be a hyperbolic, Gaussian (Bouman and Sauer, 1993) or any edge-preserving function (Thibault et al., 2007). The Total Variation (TV) could also be thought of as one such potential function. The (Rasch et al., 2018) shows a variational approach for the reconstruction of subsampled dynamic MR data, which combines smooth, temporal regularization with spatial total variation regularization.

Different from Total Variation (TV) approaches, (Bostan et al., 2012) proposed a stochastic modeling approach that is based on the solution of a stochastic differential equation (SDE) driven by non-Gaussian noise. Such stochastic modeling approaches promote the use of nonquadratic regularization functionals by tying them to some generative, continuous-domain signal model.

The Compressed Sensing (CS) based methods impose sparsity in the image domain by modifying Eqn. 2 to the following:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\Gamma\mathbf{x}\|_1, \quad (17)$$

where  $\Gamma$  is an operator that makes  $\mathbf{x}$  sparse. The  $l_1$  norm is used to promote sparsity in the transform or image domain. The  $l_1$  norm minimization can be pursued using a basis pursuit or greedy algorithm (Boyd et al., 2004). However, use of non-convex *quasi*-norms (Chartrand, 2007; Zhao and Hu, 2008; Chartrand and Staneva, 2008; Saab et al., 2008) show an increase in robustness to noise and image non-sparsity. The structured sparsity theory (Boyer et al., 2019) shows that only  $\mathcal{O}(M + M \log N)$  measurements are sufficient to reconstruct MR images when  $M$ -sparse data with size  $N$  are given. The kt-SPARSE approach of (Lustig et al., 2006) uses a spatio-temporal regularization for high SNR reconstruction.

Iterative sparsity based methods (Ravishankar and Bresler, 2012; Liu et al., 2015, 2016) assume that the image can be expressed as a linear combination of the columns (atoms) from a dictionary  $\Gamma$  such that  $\mathbf{x} = \Gamma^T \mathbf{h}$  and  $\mathbf{h}$  is the coefficient vector. Hence Eqn. 4 becomes:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \mathcal{R}(\mathbf{x}); \quad \mathcal{R}(\mathbf{x}) = \min_{\mathbf{h}} \frac{1}{2} \|\mathbf{x} - \Gamma^T \mathbf{h}\|_2^2 + \alpha \|\mathbf{h}\|_1. \quad (18)$$

The SOUP-DIL method (Bruckstein et al., 2009) uses an exact block coordinate descent scheme for optimization while a few methods (Chen et al., 2008; Lauzier et al., 2012; Chen et al., 2008) assume to have a prior image  $\mathbf{x}_0$ , i.e.  $\mathcal{R}(\mathbf{x} - \mathbf{x}_0)$ , to optimize Eqn. 4. The method in (Caballero et al., 2014) optimally sparsifies the spatio-temporal data by training an overcomplete basis of atoms.

The method in (Hongyi Gu, 2021) shows a DL based approach to leverage wavelets for reconstruction.

The transform based methods are a generalization of the CS approach that assumes a sparse approximation of the image along with a regularization of the transform itself, i.e.,  $\Gamma\mathbf{x} = \mathbf{h} + \epsilon$ , where  $\mathbf{h}$  is the sparse representation of  $\mathbf{x}$  and  $\epsilon$  is the modeling error. The method in (Ravishankar and Bresler, 2015) proposed a regularization as follows:

$$\mathcal{R}(x) = \min_{\Gamma, \mathbf{h}} \frac{1}{2} \|\mathbf{x} - \Gamma^T \mathbf{h}\|_2^2 + \alpha Q(\Gamma), \quad (19)$$

where  $Q(\Gamma) = -\log |\det \Gamma| + 0.5 \|\Gamma\|^2$  is the transform regularizer. In this context, the STROLLR method (Wen et al., 2018) used a global and a local regularizer.

In general, Eqn. 5 is a non-convex function and cannot be optimized directly with gradient descent update rules. The unrolled optimization algorithm procedure decouples the data consistency term and the regularization term by leveraging variable splitting in Eqn 4 as follows:

$$\min_{x, \mathbf{z}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \mu \|\mathbf{x} - \mathbf{h}\|_2^2 + \mathcal{R}(\mathbf{h}), \quad (20)$$

where the regularization is decoupled using a quadratic penalty on  $\mathbf{x}$  and an auxiliary random variable  $\mathbf{z}$ . Eqn 20 is optimized via alternate minimization of

$$\mathbf{h}_i = \min_h \lambda \|\mathbf{x}_{i-1} - \mathbf{h}\|_2^2 + \mathcal{R}(\mathbf{h}) \quad (21)$$

and the data consistency term:

$$\mathbf{x}_i = \min_x \|\mathbf{A}\mathbf{x}_{i-1} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{h}_i - \mathbf{x}\|_2^2 \quad (22)$$

where the  $\mathbf{h}_i, \mathbf{h}_{i-1}$  are the intermediate variables at iteration  $i$ . The alternating direction method of multiplier networks (ADMM net) introduce a set of intermediate variables,  $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n$ , and eventually we have a set of dictionaries,  $\Gamma_1, \Gamma_2, \dots, \Gamma_n$ , such that,  $\mathbf{h}_i = \Gamma_i \mathbf{x}$ , collectively promote sparsity. The basic ADMM net update (Yang et al., 2018) is as follows:

$$\begin{aligned} \arg \min_x \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \sum_i \|\Gamma_i \mathbf{x} + \beta_i - \mathbf{z}_i\|_2^2; \arg \min_{\mathbf{z}_i} \sum_i [\lambda_i g(\mathbf{h}_i) \|\Gamma_i \mathbf{x} + \beta_i - \mathbf{h}_i\|_2^2] \\ \beta_i \leftarrow \beta_{i-1} + \alpha_i (\Gamma_i \mathbf{x} - \mathbf{h}_i) \end{aligned} \quad (23)$$

where  $g(\cdot)$  can be any sparsity promoting operator and  $\beta$  is called a multiplier. The iterative shrinkage thresholding algorithm (ISTA) solves this CS optimization problem as follows:

$$\mathbf{h}_{i+1} = \mathbf{x}_i - \Phi^H (\Phi \mathbf{x}_i - \mathbf{y}); \quad \mathbf{x}_{i+1} = \arg \min_x \frac{1}{2} \|\mathbf{x} - \mathbf{h}_{i+1}\|_2^2 + \lambda \|\Gamma \mathbf{x}\|_1. \quad (24)$$

Later in this paper, we shall show how ISTA and ADMM can be organically used within the modern DL techniques in Sec. 5.3.

### 3. Review of Deep Learning Building Blocks

In this section, we will describe basic building blocks that are individually or collectively used to develop complex DL methods that work in practice. Any DL method, by design, has three major components: the network structure, the training process, and the dataset on which the DL method is trained and tested. We shall discuss each one of them below in detail.

#### 3.1 Various Deep Learning Frameworks

**Perceptron:** The journey of DL started in the year 1943 when Pitts and McCulloch (McCulloch and Pitts, 1943) gave the mathematical model of a biological neuron. This mathematical model is based on the “all or none” behavioral dogma of a biological neuron. Soon after, Rosenblatt provided the perceptron learning algorithm (Rosenblatt, 1957) which is a mathematical model based on the behaviour of a neuron. The perceptron resembles the structure of a neuron with dendrites, axons and a cell body. The basic perceptron is a binary classification algorithm of the following form:

$$f(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{w} \cdot \mathbf{x} + b > 0, \\ 0 & \text{otherwise} \end{cases} \quad (25)$$

where  $\mathbf{x}_i$ 's are the components of an image vector  $\mathbf{x}$ ,  $w_i$ 's are the corresponding weights that determine the slope of the classification line, and  $b$  is the bias term. This setup collectively resembles the “all or none” working principle of a neuron. However, in the famous book of Minsky and Papert (Minsky and Papert, 2017) called “Perpetron” it was shown that the perceptron can't classify non-separable points such as an exclusive-OR (XOR) function.

**Multilayer Perceptron (MLP):** It was understood that the non-separability problem of perceptron can be overcome by a multilayer perceptron (Minsky and Papert, 2017) but the research stalled due to the unavailability of a proper training rule. In the year 1986, Rumelhart *et al.* (Rumelhart et al., 1986) proposed the famous “backpropagation algorithm” that provided fresh air to the study of neural network. A Multilayer Perceptron (MLP) uses several layers of multiple perceptrons to perform nonlinear classification. A MLP is comprised of an input layer, an output layer, and several densely connected in-between layers called *hidden layers*:

$$\begin{aligned} \mathbf{h}^1 &= \psi^1\left(\sum W^1 \mathbf{x} + b^1\right); \quad \mathbf{h}^i = \psi^i\left(\sum W^i \mathbf{h}^{i-1} + b^i\right), \text{ where, } i \in \{2, \dots, n-1\} \\ \mathbf{y} &= \psi^n\left(\sum W^n \mathbf{h}^{n-1} + b^{n-1}\right). \end{aligned} \quad (26)$$

Along with the hidden layers and the input-output layer, the MLP learns features of the input dataset and uses them to perform classification. The dense connection among hidden layers, input and output layers often creates a major computational bottleneck when the input dimension is very high.

**Neocognitron or the Convolutional Neural Network (CNN):** The dense connection (also known as global connection) of a MLP was too flexible a model and prone to overfitting and sometimes had large computational overhead. To cope with this situation, a local sliding window based network with shared weights was proposed in early 1980s called neocognitron

network (Fukushima and Miyake, 1982) and later popularized as Convolutional Neural Network (CNN) in the year 1998 (LeCun et al., 1989). Similar to the formulation of (Liang et al., 2020a), we write the feedforward process of a CNN as follows:

$$\begin{aligned} C_0 &= \mathbf{x} \\ C_i &= \psi_{i-1}(K_{i-1} \star C_{i-1}); \quad i \in \{1, \dots, n-1\}, \quad C_n = \psi_{n-1}(K_{n-1} \star C_{n-1}), \end{aligned} \quad (27)$$

where  $C_i \in R^{h \times w \times d}$  is the  $i^{\text{th}}$  hidden layer comprised of  $d$ -number of feature maps each of size  $h \times w$ ,  $K_i$  is the  $i^{\text{th}}$  kernel that performs a convolution operation on  $C_i$ , and  $\psi_i(\cdot)$  are activation functions to promote non-linearity. We show a vanilla kernel operation in Eqn. 27. Please note that, the layers of the CNN can either be a fully connected dense layer, a max-pooling layer that downsizes the input, or a dropout layer to perform regularization that is not shown in Eqn. 27.

**Recurrent Neural Networks (RNNs):** A CNN can learn hidden features of a dataset using its inherent deep structure and local connectivities through a convolution kernel. But they are not capable of learning the time dependence in signals. The recurrent neural network (RNN) (Rumelhart et al., 1986) at its basic form is a time series neural network with the following form:

$$\begin{aligned} \mathbf{h}_0 &= \mathbf{x} \\ \mathbf{h}_t &= \psi_t(W_{\mathbf{hh}}\mathbf{h}_{t-1} + W_{\mathbf{xh}}\mathbf{x}_t), \quad t \in \{1, \dots, n-1\}; \quad \mathbf{h}_n = \psi_n(W_{\mathbf{hy}}\mathbf{h}_{n-1}), \end{aligned} \quad (28)$$

where  $t$  is the time and the RNN takes the input  $\mathbf{x}$  in a sequential manner. However, the RNN suffers from the problem of “vanishing gradient”. The vanishing gradient is observed when gradients from output layer of a RNN trained with gradient based optimization method changes parameter values by a very small amount, thus effecting no change in parameter learning. The Long Short Term Memory (LSTM) network (Hochreiter and Schmidhuber, 1997) uses memory gates, and sigmoid and/or tanh activation function and later ReLU activation function (see Sec. 3.2 for activation functions) to control the gradient signals and overcome the vanishing gradient problem.

**Transformer Networks:** Although the LSTM has seen tremendous success in DL and MR reconstruction, there are a few problems associated with the LSTM model (Vaswani et al., 2017) such as: (i) the LSTM networks perform a sequential processing of input; and (ii) the short attention span of the hidden states that may not learn a good contextual representation of input. Such shortcomings are largely mitigated using a recent advancement called the Transformer Network (Vaswani et al., 2017). A transformer network has a self-attention mechanism<sup>2</sup>, a positional embedding and a non-sequential input processing setup and empirically this configuration outperforms the LSTM networks by a large margin (Vaswani et al., 2017).

---

2. **Self-attention:** The attention mechanism provides a way to know which part of the input is to be given more focus. The self-attention, on the other hand, measures the contextual relation of an input by allowing it to interact with itself. Let’s assume that we are at  $\mathbf{h}_i^{\text{th}} \in R^{c \times N}$  layer that has  $C$  number of channels and  $N$  number of locations on a feature map. We get two feature vectors,  $f(\mathbf{h}_i) = W_f \mathbf{h}_i$  and  $g(\mathbf{h}_i) = W_g \mathbf{h}_i$ , by transforming the  $\mathbf{h}_i^{\text{th}}$  layer to a vector (typically done with a  $1 \times 1$  convolution). The contextual similarity of these two vectors  $f(\mathbf{h}_i)$  and  $g(\mathbf{h}_i)$  is measured by;

$$\beta_{k,l} = \frac{\exp(\mathbf{s}_{kl})}{\sum_{l=1}^N \exp(\mathbf{s}_{kl})}, \text{ where } \mathbf{s}_{kl} = f(\mathbf{h}_i)^T g(\mathbf{h}_i)$$

### 3.2 Activation Functions

The activation function,  $\psi(\cdot)$ , operates on a node or a layer of a neural network and provides a boolean output, probabilistic output or output within a range. The step activation function was proposed by McCulloch & Pitts in (McCulloch and Pitts, 1943) with the following form:  $\psi(\mathbf{x}) = 1$ , if  $\mathbf{x} \geq 0.5$  and 0 otherwise. Several initial works also used the hyperbolic tangent function,  $\psi(\mathbf{x}) = \tanh(\mathbf{x})$ , as an activation function that provides value within the range  $[-1, +1]$ . The sigmoid activation function,  $\psi(\mathbf{x}) = \frac{1}{1+e^{-x}}$ , is a very common choice and provides only positive values within the range  $[0, 1]$ . However, one major disadvantage of the sigmoid activation function is that its derivative,  $\psi'(\mathbf{x}) = \psi(\mathbf{x})(1 - \psi(\mathbf{x}))$ , quickly saturates to zero which leads to the *vanishing gradient* problem. This problem was addressed by adding a Rectified Linear Unit (ReLU) to the network (Brownlee, 2019),  $\psi(\mathbf{x}) = \max(0, \mathbf{x})$  with the derivative  $\psi'(\mathbf{x}) = 1$ , if  $\mathbf{x} > 0$  or 0 elsewhere.

### 3.3 Network Structures

**The VGG Network:** In late 2015, Zisserman *et al.* published their seminal paper with the title “very deep convolutional networks for large-scale image recognition” (VGG) (Simonyan and Zisserman, 2014) that presents a 16-layered network called VGG network. Each layer of the VGG network has an increasing number of channels. The network was shown to achieve state-of-the-art level performance in Computer Vision tasks such as classification, recognition, etc.

**The ResNet Model:** A recently developed model called Residual Networks or ResNet (He et al., 2016), modifies the layer interaction shown in Eqn. 27 to the following form,  $C_i = \psi_{i-1}(K_{i-1} \star C_{i-1}) + C_{i-2}$  where,  $i \in \{2, \dots, n-1\}$ , and provides a “shortcut connection” to the hidden layers. The identity mapping using the shortcut connections has a large positive impact on the performance and stability of the networks.

**UNet:** A UNet architecture (Ronneberger et al., 2015) was proposed to perform image segmentation task in biomedical images. The total end-to-end architecture pictorially resembles the english letter “U” and has a encoder module and a decoder module. Each encoder layer is comprised of unpadded convolution, a rectified linear unit (ReLU in Sec. 3.2), and a pooling layer, which collectively downsample the image to some latent space. The decoder has the same number of layers as the encoder. Each decoder layer upsamples the data from its previous layer until the input dimension is reached. This architecture has been shown to provide good quantitative results on several datasets.

**Autoencoders:** The autoencoders (AE) are a type of machine learning models that capture the patterns or regularities of input data samples in an unsupervised fashion by mapping the target values to be equal to the input values (i.e. identity mapping). For example, given a data point  $\mathbf{x}$  randomly sampled from the training data distribution  $p_{data}(\mathbf{x})$ , a standard AE learns a low-dimensional representation  $\mathbf{z}$  using an encoder network,  $\mathbf{z} \sim D(\mathbf{z}|\mathbf{x}, \theta_d)$ , that is parameterized by  $\theta_d$ . The low-dimensional representation  $\mathbf{z}$ , also called the latent representation, is subsequently projected back to the input dimension using a decoder network,  $\tilde{\mathbf{x}} \sim G_{GEN}(\mathbf{x}|\mathbf{z}, \theta_g)$ , that is parameterized by  $\theta_g$ . The model parameters,

---

The output is  $\mathbf{o} = (\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_N)$ , where  $\mathbf{o}_k = v(\sum_{i=1}^N \beta_{k,i} m(\mathbf{h}_{il}))$ , and  $m(h_{il}) = W_m h_i$ ,  $v(h_{il}) = W_v \mathbf{h}_{il}$ . Here, the  $W_v, W_m, W_f, W_g$  are learnable matrices that collectively provide the self-attention vector  $\mathbf{o}\mathbf{o} = (\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_N)$  for a given layer  $\mathbf{h}_i$ .

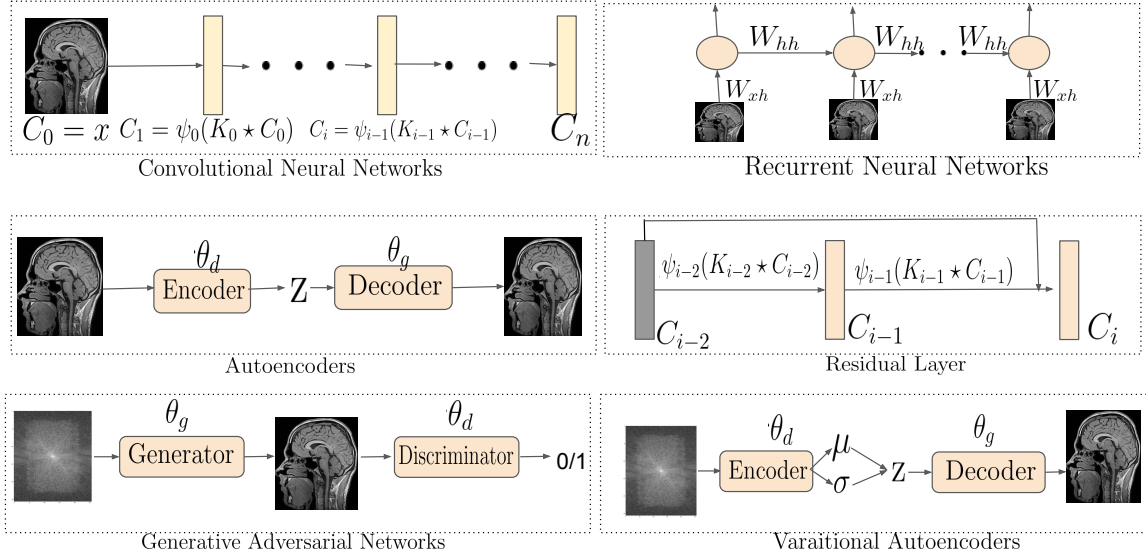


Figure 3: **Deep Learning Models:** We visually demonstrate various DL models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Autoencoders, Residual Layer, Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) discussed in Sec. 3. The **Convolutional Neural Network** is comprised with many layers  $\{C_0, C_1, \dots, C_n\}$  and the network gets input for layer  $C_{i+1}$  from its previous layer  $C_i$  after imposing non-linearity, i.e.  $C_{i+1} = \psi_i(K_i \star C_i)$ , using a non-linear function  $\psi_i$ . The **Recurrent Neural Network** is a time varying memory network. On the other hand, the **Autoencoders** perform non-linear dimensionality reduction by projecting the input  $\mathbf{x}$  to a lower dimensional variable  $\mathbf{z}$  using an encoder network and projects back  $\mathbf{z}$  to the image space  $\mathbf{x}$  using a decoder network. The ResNet uses the **Residual Layer** to solve problems like vanishing gradient, slow convergence, etc. The **Generative Adversarial Network** and the **Variational Autoencoder** are generative models that are discussed in Sec. 4.

i.e.  $(\theta_d, \theta_g)$ , are trained using the standard back propagation algorithm with the following optimization objective:

$$\mathcal{L}_{\theta_d, \theta_g} = \arg \min_{\theta_g, \theta_d} \mathbb{E}_{\mathbf{z} \sim D(\mathbf{x}, \theta_d)} \frac{1}{2} \|\mathbf{x} - G_{GEN}(\mathbf{x}|\mathbf{z}, \theta_g)\|_2^2. \quad (29)$$

From a Bayesian learning perspective, an AE learns a posterior density,  $p(\mathbf{z}|\mathbf{x})$ , using the encoder network,  $G_{GEN}(\mathbf{z}|\mathbf{x}, \theta_g)$ , and a decoder network,  $D(\mathbf{x}|\mathbf{z}, \theta_d)$ . The vanilla autoencoder, in a way, can be thought of as a non-linear principal component analysis (PCA) (Hinton and Salakhutdinov, 2006) that progressively reduces the input dimension using the encoder network and finds regularities or patterns in the data samples.

**Variational Autoencoder Networks:** Variational Autoencoder (VAE) is basically an autoencoder network that is comprised of an encoder network  $G_{GEN}(\mathbf{z}|\mathbf{x}, \theta_g)$  that estimates

the posterior distribution  $p(\mathbf{z}|\mathbf{x})$  and the inference  $p(\mathbf{x}|\mathbf{z})$  with a decoder network  $D(z, \theta_d)$ . However, the posterior  $p(\mathbf{z}|\mathbf{x})$  is intractable<sup>3</sup> and several methods have been proposed to approximate the inference using techniques such as the Metropolis-Hasting (Metropolis et al., 1953) and variational inference (VI) algorithms (Kingma and Welling, 2013). The main essence of VI algorithm is to estimate an intractable probability density, i.e.  $p(\mathbf{z}|\mathbf{x})$  in our case, from a class of tractable probability densities,  $\mathcal{Z}$ , with an objective to finding a density,  $q(\mathbf{z})$ , almost similar to  $p(\mathbf{z}|\mathbf{x})$ . We then sample from the tractable density  $q(\mathbf{z})$  instead of  $p(\mathbf{z}|\mathbf{x})$  to get an approximate estimation, i.e.

$$q^*(z) = \arg \min_{q(\mathbf{z}) \in \mathcal{Q}} D_{KL}(q(\mathbf{z})||p(\mathbf{z}|\mathbf{x})). \quad (30)$$

Here,  $D_{KL}$  is the KL divergence (Kullback and Leibler, 1951). The VI algorithm, however, typically never converges to the global optimal solution but provides a very fast approximation. The VAE consists of three components: (i) the encoder network  $E_{\theta_e}(\mathbf{z}|\mathbf{x})$  that observes and encodes a data point  $\mathbf{x}$  from the training dataset  $\mathcal{D}(\mathbf{x})$  and provides the mean and variance of the approximate posterior, i.e.  $\mathcal{N}(\mu_{\theta}(\mathbf{x}_i), \sigma_{\theta}(\mathbf{x}_i))$  from a batch of  $n$  data points  $\mathbf{x}_i|_{i=1}^n$ ; (ii) a prior distribution  $G(z)$ , typically an isotropic Gaussian  $\mathcal{N}(0, I)$  from which  $\mathbf{z}$  is sampled, and (iii) a generator network  $G_{\theta_g}(\mathbf{x}|\mathbf{z})$  that generates data points given a sample from the latent space  $\mathbf{z}$ . The VAE, however, cannot directly optimize the VI, i.e.  $q^*(z) = \arg \min_{q(\mathbf{z}) \in \mathcal{Q}} D_{KL}(q(\mathbf{z})||p(\mathbf{z}|\mathbf{x}))$ , as the KL divergence requires an estimate of the intractable density  $p(\mathbf{x})$ , i.e.  $D_{KL}(q(\mathbf{z})||p(\mathbf{z}|\mathbf{x})) = \mathbb{E}[\log q(\mathbf{z})] - \mathbb{E}[\log p(\mathbf{z}, \mathbf{x})] + \mathbb{E}[\log p(\mathbf{x})]$ . As a result, VAE estimates the evidence lower bound (ELBO) that is similar to KL divergence, i.e.:

$$\begin{aligned} ELBO(q) &= \mathbb{E}[\log p(\mathbf{z}, \mathbf{x})] - \mathbb{E}[\log q(\mathbf{z})] = \mathbb{E}[\log p(\mathbf{z})] + \mathbb{E}[\log p(\mathbf{x}|\mathbf{z})] - \mathbb{E}[\log q(\mathbf{z})] \\ &= \mathbb{E}[\log p(\mathbf{x}|\mathbf{z})] - D_{KL}(p(\mathbf{z}|\mathbf{x})||q(\mathbf{z})) = \mathbb{E}_{\mathbf{z} \sim p(\mathbf{z}|\mathbf{x})}[\log p(\mathbf{x}|\mathbf{z})] - D_{KL}(p(\mathbf{z}|\mathbf{x})||q(\mathbf{z})) \end{aligned} \quad (31)$$

Since  $ELBO(q) \leq \log p(\mathbf{x})$ , optimizing Eqn. (31) provides a good approximation of the marginal density  $p(\mathbf{x})$ . Please note that, Eqn. 31 is similar to Eqn. 4 where the first term in Eqn. 31 is the data consistency term and the KL divergence term acts as a regularizer.

**Generative Adversarial Networks:** A vanilla Generative Adversarial Networks (GAN) setup, by design, is an interplay between two neural networks called the *generator*, that is  $G_{GEN}(\mathbf{x}|\mathbf{z}, \theta_g)$ , and the *discriminator*  $D_{\theta_d}(\cdot)$  parameterized by  $\theta_g$  and  $\theta_d$  respectively. The  $G_{GEN}(\mathbf{x}|\mathbf{z}, \theta_g)$  samples the latent vector  $z \in \mathbb{R}^{n \times 1 \times 1}$  and generates  $\mathbf{x}^{gen}$ . While the discriminator, on the other hand, takes  $\mathbf{x}$  (or  $\mathbf{x}^{gen}$ ) as input and provides a  $\{real, fake\}$  decision on  $\mathbf{x}$  being sampled from a real data distribution or from  $G_{GEN}(\mathbf{x}|\mathbf{z}, \theta_g)$ . The parameters are trained using a game-theoretic adversarial objective, i.e.:

$$\mathcal{L}_{\theta_d, \theta_g} = -\mathbb{E}_{\mathbf{x} \sim p_{data}}[\log D(\mathbf{x}|\theta_d)] - \mathbb{E}_{\mathbf{z} \sim \mathcal{N}(0, I)}[\log(1 - D(G_{GEN}(\mathbf{x}|\mathbf{z}, \theta_g), \theta_d))]. \quad (32)$$

---

3. **Intractability:** We can learn the density of the latent representation from the data points themselves, i.e.  $p(\mathbf{z}|\mathbf{x}) = p(\mathbf{z})p(\mathbf{x}|\mathbf{z})/p(\mathbf{x})$ , by expanding the Bayes theorem of conditional distribution. In this equation, the numerator is computed for a single realization of data points. However, the denominator is the marginal distribution of data points,  $p(\mathbf{x})$ , which are complex and hard to estimate; thus, leading to intractability of estimating the posterior.



As the training progresses, the generator  $G_{GEN}(\mathbf{z}|\theta_g)$  progressively learns a strategy to generate realistic looking images, while the discriminator  $D(\mathbf{x}, \theta_d)$  learns to discriminate the generated and real samples.

### 3.4 Loss Functions

In Sec. 1.1, we mentioned the loss function,  $L$ , that estimates the empirical loss and the generalization error. Loss functions that are typically used in MR image reconstruction using DL methods are the Mean Squared Error (MSE), Peak Signal to Noise Ratio (PSNR) and the Structural Similarity Loss function (SSIM), or an  $l_1$  loss to optimize Eqns. 3, 4, 5. The MSE loss between an image  $\mathbf{x}$  and its noisy approximation  $\hat{\mathbf{x}}$  is defined as,  $MSE(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{n} \sum_{i=1}^n (\hat{\mathbf{x}}_i - \mathbf{x}_i)^2$ , where  $n$  is the number of samples. The root MSE (RMSE) is essentially the squared root of the MSE,  $\sqrt{MSE(\mathbf{x}, \hat{\mathbf{x}})}$ . The  $l_1$  loss,  $l_1(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{n} \sum_{i=1}^n |\hat{\mathbf{x}}_i - \mathbf{x}_i|$ , provides the absolute difference and is typically used as a regularization term to promote sparsity. The PSNR is defined using the MSE,  $PSNR(\mathbf{x}, \hat{\mathbf{x}}) = 20 \log_{10}(MAX_{\mathbf{x}} - 10 \log_{10}(MSE(\mathbf{x}, \hat{\mathbf{x}}))$ , where  $MAX_{\mathbf{x}}$  is the highest pixel value attained by the image  $\mathbf{x}$ . The PSNR metric captures how strongly the noise in the data affects the fidelity of the approximation with respect to the maximum possible strength of the signal (hence the name peak signal to noise ratio). The main concern with  $MSE(\mathbf{x}, \hat{\mathbf{x}})$  and  $PSNR(\mathbf{x}, \hat{\mathbf{x}})$  is that they penalize large deviations much more than smaller ones (Zhao et al., 2016) (e.g., outliers are penalized more than smaller anatomical details). The SSIM loss for a pixel  $i$  in the image  $\mathbf{x}$  and the approximation  $\hat{\mathbf{x}}$  captures the perceptual similarity of two images:  $SSIM(\mathbf{x}_i, \hat{\mathbf{x}}_i) = \frac{2\mu_{\mathbf{x}}\mu_{\hat{\mathbf{x}}} + c_1}{\mu_{\mathbf{x}}^2 + \mu_{\hat{\mathbf{x}}}^2 + c_1} \frac{2\sigma_{\mathbf{x}\hat{\mathbf{x}}} + c_2}{\sigma_{\mathbf{x}}^2 + \sigma_{\hat{\mathbf{x}}}^2 + c_2}$ , here  $c_1, c_2$  are two constants, and  $\mu$  is the mean and  $\sigma$  is the standard deviation.

**The VGG loss:** It is shown in (Johnson et al., 2016) that the deeper layer feature maps (feature maps are discussed in Eqn. 27) of a VGG-16 network, i.e. a VGG network that has 16 layers, can be used to compare perceptual similarity of images. Let us assume that, the  $L^{th}$  layer of a VGG network has distinct  $N_L$  feature maps each of size  $M_L \times M_L$ . The matrix  $F^L \in \mathbb{R}^{N_L \times M_L}$ , stores the activations  $F_{i,j}^L$  of the  $i^{th}$  filter at position  $j$  of layer  $L$ . Then, the method computes feature correlation using:  $C_{i,j}^L = \sum_k F_{i,k}^L F_{j,k}^L$ , where any  $F_{n,o}^m$  conveys the activation of the  $n^{th}$  filter at position  $o$  in layer  $m$ . The correlation  $C_{i,j}^L$  is considered as a VGG-loss function.

## 4. Inverse Mapping using Deep Generative Models

Based on how the generator network,  $G_{GEN}(\mathbf{x}|\mathbf{z}, \theta_g)$ , is optimized in Eqn. 6, we get different manifestations of deep generative networks such as Generative Adversarial Networks, Bayesian Networks, etc. In this section, we shall discuss specifics on how these networks are used in MR reconstruction.

### 4.1 Generative Adversarial Networks (GANs)

We gave a brief introduction to GAN in Sec. 3.3 and in this section we shall take a closer look on the different GAN methods used to learn the inverse mapping from k-space measurements to the MR image space.

**Inverse Mapping from k-space:** The current GAN based k-space methods can be broadly classified into two categories: (i) methods that directly operate on the k-space

$\mathbf{y}$  and reconstruct the image  $\mathbf{x}$  by learning a non-linear mapping and (ii) methods that impute the missing k-space lines  $\mathbf{y}^{missing}$  in the undersampled k-space measurements. In the following paragraphs, we first discuss the recent GAN based direct k-space to MR image generation methods followed by the undersampled k-space to full k-space generation methods.

The direct k-space to image space reconstruction methods, as shown in Fig. 3 (a), are based on the premise that the missing k-space lines can be estimated from the acquired k-space lines provided we have a good non-linear interpolation function, i.e.;

$$\mathcal{L}_{\theta_d, \theta_g} = -\mathbb{E}_{\mathbf{x} \sim p_{data}}[\log D(\mathbf{x}|\theta_d)] - \mathbb{E}_{\mathbf{z} \sim \mathcal{N}(0, I)}[\log(1 - D(G_{GEN}(\mathbf{x}|y, \theta_g), \theta_d))]. \quad (33)$$

This GAN framework was used in (Oksuz et al., 2018) for correcting motion artifacts in cardiac imaging using a generic AUTOMAP network<sup>4</sup>. Such AUTOMAP-like generator architectures not only improve the reconstruction quality but help in other downstream tasks such as MR image segmentation (Oksuz et al., 2019a, 2020, 2019b). However, while the ‘‘AUTOMAP as generator’’ based methods solve the broader problem of motion artifacts, but they largely fail to solve the banding artifacts along the phase encoding direction. To address this problem, a method called MRI Banding Removal via Adversarial Training (Defazio et al., 2020) leverages a perceptual loss along with the discriminator loss in Eqn. 32. The perceptual loss ensures data consistency, while, the discriminator loss checks whether: (i) the generated image has a horizontal (0) or a vertical (1) banding; and (ii) the generated image resembles the real image or not. With a 4x acceleration, a 12-layered UNet generator and a ResNet discriminator, the methodology has shown remarkable improvements (Defazio et al., 2020) on fastMRI dataset.

Instead of leveraging the k-space regularization within the parameter space of a GAN (Oksuz et al., 2018, 2019a), the k-space data imputation using GAN directly operates on the k-space measurements to regularize Eqn. 32. To elaborate, these type of methods estimate the missing k-space lines by learning a non-linear interpolation function (similar to GRAPPA) within an adversarial learning framework, i.e.

$$\mathcal{L}_{\theta_d, \theta_g} = -\mathbb{E}_{\mathbf{x} \sim p_{data}}[\log D(\mathbf{x}|\theta_d)] - \mathbb{E}_{\mathbf{z} \sim \mathcal{N}(0, I)}[\log(1 - D(\mathcal{F}G_{GEN}(\mathbf{y}^{full}|\mathbf{y}, \theta_g), \theta_d))]. \quad (34)$$

The accelerated magnetic resonance imaging (AMRI) by adversarial neural network method (Shitrit and Raviv, 2017) aims to generate the missing k-space lines,  $\mathbf{y}^{missing}$  from  $y$  using a conditional GAN,  $\mathbf{y}^{missing} \sim G(\mathbf{y}^{missing}|\mathbf{z}, c = y)$ . The combined  $y, \mathbf{y}^{missing}$  is Fourier transformed and passed to the discriminator. The AMRI method showed improved PSNR value with good reconstruction quality and no significant artifacts as shown in Fig. 4. Later, in subsampled brain MRI reconstruction by generative adversarial neural networks method (SUBGAN) (Shaul et al., 2020), the authors discussed the importance of temporal context and how that mitigates the noise associated with the target’s movement. The UNet-based generator in SUBGAN takes three adjacent subsampled k-space slices  $\mathbf{y}_{i-1}, \mathbf{y}_i, \mathbf{y}_{i+1}$  taken

4. **AUTOMAP** (Zhu et al., 2018): This is a two stage network resembling the unrolled optimization like methods (Schlemper et al., 2017). The first sub network ensures the data consistency, while, the other sub network helps in refinement of the image. The flexibility of AUTOMAP enables it to learn the k-space to image space mapping from alternate domains instead of strictly from a paired k-space to MR image training dataset.

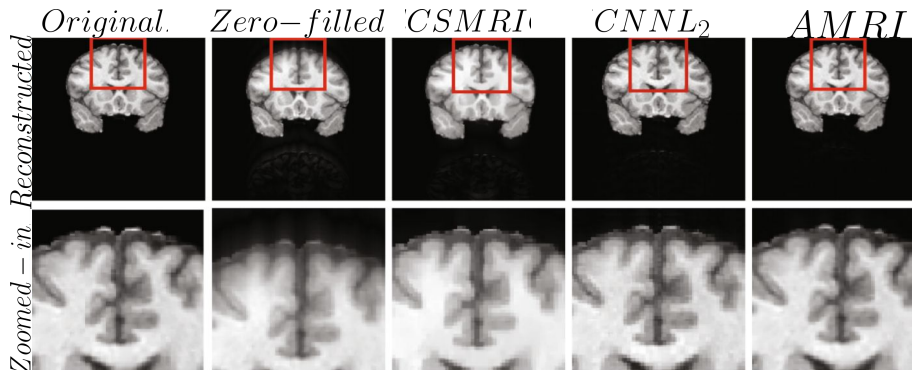


Figure 4: **Reconstruction using k-space Imputation:** We show qualitative comparison of a few methods with the AMRI method (Shitrit and Raviv, 2017). The first row are the reconstructed images and the second row images are zoomed-in version of the red boxes shown in first row. The Zero-filled, compressed sensing MRI (CSMRI), CNN network trained with  $l_2$  loss (CNN-L2) have lesser detail as compared to the AMRI (proposed) method. The images are borrowed from AMRI (Shitrit and Raviv, 2017) paper.

at timestamps  $t_{i-1}, t_i, t_{i+1}$  and provides the reconstructed image. The method achieved a performance boost of  $\sim 2.5$  in PSNR with respect to the other state-of-the-art GAN methods while considering 20% of the original k-space samples on IXI dataset (Rowland et al., 2004). We also show reconstruction quality of SUBGAN on fastMRI dataset in Fig. 5. Another method called multi-channel GAN (Zhang et al., 2018b) advocates the use raw of k-space measurements from all coils and has shown good k-space reconstruction and lower background noise compared to classical parallel imaging methods like GRAPPA and SPIRiT. However, we note that this method achieved  $\sim 2.8$  dB lower PSNR than the GRAPPA and SPIRiT methods.

Despite their success in MR reconstruction from feasible sampling patterns of k-space, the previous models we have discussed so far have the following limitations: (i) they need unaliased images for training, (ii) they need paired k-space and image space data, or (iii) the need fully sampled k-space data. In contrast, we note a recent work called unsupervised MR reconstruction with GANs (Cole et al., 2020) that only requires the undersampled k-space data coming from the receiver coils and optimizes a network for image reconstruction. Different from AutomapGAN (Oksuz et al., 2019a), in this setup the generator provides the undersampled k-space (instead of the MR image as in case of AutomapGAN) after applying Fourier transform, sensitivity encoding and a random-sampling mask  $\mathcal{M}_1$  on the generated image, i.e.  $\mathbf{y}^{gen} = \mathcal{M}_1 \mathcal{F} \mathcal{S}(G(\mathbf{x}|y, \theta_g))$ . The discriminator takes the k-space measurements instead of an MR image and provides thee learned signal to the generator.

**Image space Rectification Methods:** Image space rectification methods operate on the image space and learn to reduce noise and/or aliasing artifacts by updating Eqn. 4 to the

following form:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{x} - G_{GEN}(\mathbf{x}|\mathbf{x}^{low}, \theta_g)\|_2^2 + \lambda \|\mathbf{y}^t - \mathcal{F}G_{GEN}(\mathbf{x}|\mathbf{x}^{low}, \theta_g)\|_2^2 + \zeta \mathcal{L}_{\theta_d, \theta_g}; \quad (35)$$

$$\mathcal{L}_{\theta_d, \theta_g} = -\mathbb{E}_{\mathbf{x} \sim p_{data}} [\log D(\mathbf{x}|\theta_d)] - \mathbb{E}_{z \sim \mathcal{N}(0, I)} [\log(1 - D(G_{GEN}(\mathbf{x}|\mathbf{x}^{low}, \theta_g), \theta_d))].$$

The GAN framework for deep de-aliasing (Yang et al., 2017) regularizes the reconstruction by adopting several image priors such as: (i) image content information like object boundary, shape, and orientation, along with using a perceptual loss function:  $\frac{1}{2} \|\mathbf{x} - \mathbf{x}^{low}\|_2^2$ , (ii) data consistency is ensured using a frequency domain loss  $\frac{1}{2} \|\mathbf{y}^t - \mathbf{y}^u\|_2^2$ , and (iii) a VGG loss (see Sec. 3.4) which enforces semantic similarity between the reconstructed and the ground truth images. The method demonstrated a 2dB improvement in PSNR score on IXI dataset with 30% undersampling. However, it was observed that the finer details were lost during the process of de-aliasing with a CNN based GAN network. In the paper called self-attention and relative average discriminator based GAN (SARAGAN) (Yuan et al., 2020), the authors show that fine details tend to fade away due to smaller size of the convolution kernels, leading to poor performance. Consequently, the SARAGAN method adopts a relativistic discriminator (Jolicoeur-Martineau, 2018) along with a self-attention network (see Sec. 3.1 for self-attention) to optimize the following equation that is different from Eqn. 35:

$$\mathcal{L}_{\theta_d, \theta_g} = -\mathbb{E}_{\mathbf{x} \sim p_{data}} [\text{sigmoid}(D(\mathbf{x}) - D(G(\mathbf{x}^{low})))] - \mathbb{E}_{z \sim \mathcal{N}(0, I)} [\text{sigmoid}(D(G(\mathbf{x}^{low})) - D(\mathbf{x}))] \quad (36)$$

where,  $\text{sigmoid}(\cdot)$  is the sigmoid activation function discussed in Sec. 3.2. This method showed excellent performance in the MICCAI 2013 grand challenge brain MRI reconstruction dataset and got an SSIM of 0.9951 and PSNR of  $45.7536 \pm 4.99$  with 30% sampling rate. Among other methods, sparsity based constraints are imposed as a regularizer to Eqn. 35 in compressed sensing GAN (GANCS) (Mardani et al., 2018a), RefineGAN (Quan et al., 2018), and the structure preserving GAN (Deora et al., 2020; Lee et al., 2018) methods. Some qualitative results using the RefineGAN method are shown in Fig. 5. On the other hand, methods like PIC-GAN (Lv et al., 2021) and (MGAN) (Zhang et al., 2018a) use a SENSE-like reconstruction strategy that combines MR images reconstructed from parallel receiver coils using a GAN framework. Such methods have also shown good performance with low normalized mean squared error in the knee dataset.

**Combined k-space and image space methods:** Thus far, we have discussed k-space (GRAPPA-like GAN methods) and image space (SENSE-like GAN methods) MR reconstruction methods that work in isolation. However, both these strategies can be combined together to leverage the advantages of both methods. Recently, a method called sampling augmented neural network with incoherent structure for MR image reconstruction (SANTIS) (Liu et al., 2019) was proposed that leverages a cycle consistency loss,  $\mathcal{L}_{cyc}$  in addition to a GAN loss,  $\mathcal{L}_{\theta_d, \theta_g}$  i.e.,

$$\mathcal{L}_{full} = \mathcal{L}_{\theta_d, \theta_g} + \mathcal{L}_{cyc} = \lambda_1 \mathbb{E}[\|\mathbf{x} - G(\mathbf{x}^{low}, \theta_g)\|_2^2] + \lambda_{GAN} \left( \mathbb{E}[\log D(\mathbf{x}, \theta_d)] \right. \\ \left. + \mathbb{E}[\log(1 - D(G(\mathbf{x}^{low}, \theta_g), \theta_d))] \right) + \lambda_2 \mathbb{E}[\|y - F(G(\mathbf{x}^{low}, \theta_g), \theta_f)\|_2^2], \quad (37)$$

where the function  $F(\cdot)$  is another generator network that projects back the MR image to k-space. The method achieved an SSIM value of 91.96 on the 4x undersampled knee

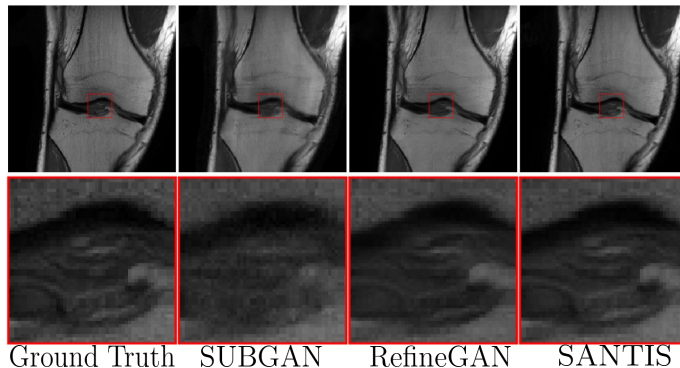


Figure 5: **Comparison of GAN Methods:** Qualitative comparison of (i) k-space interpolation based method, i.e. SUBGAN; (ii) Image space rectification method, i.e. RefineGAN; and (iii) the combined k-space and image space rectification method, i.e. SANTIS.

FastMRI dataset (see Fig. 5 and Table 1). In the collaborative GAN method (CollaGAN) (Lee et al., 2019), instead of cycle consistency between k-space and the image domain from a single image, they consider a collection of domains such as T1-weighted and T2-weighted data and try to reconstruct the MR images with cycle consistency in all domains. The InverseGAN (Narnhofer et al., 2019) method performs cycle consistency using a single network that learns both the forward and inverse mapping from and to k-space.

## 4.2 Bayesian Learning

Bayes’s theorem expresses the posterior  $p(\mathbf{x}|y)$  as a function of the k-space data likelihood,  $p(y|\mathbf{x})$  and the prior  $p(\mathbf{x})$  with the form  $p(\mathbf{x}|y) \propto p(y|\mathbf{x})p(\mathbf{x})$  also known as the “product-of-the-experts” in DL literature. In (Tezcan et al., 2018), the prior is estimated with a Monte Carlo sampling technique which is computationally intensive. To overcome the computational cost, several authors have proposed to learn a non-linear mapping from undersampled k-space to image space using VAEs. In these VAE based methods (Tezcan et al., 2019; Gaillochet et al., 2020; Van Essen et al., 2012), the networks are trained on image patches  $m_j$  obtained from k-space measurement  $\mathbf{y}_i$  and the VAE network is optimized using these patches with the following cost function:  $\sum_{j=1}^N ELBO(m_j)$ ,

$$\arg \min_m \left[ \|\mathbf{E}\mathbf{x} - y\|_2^2 - \sum_{\mathbf{x} \in \Omega(m)} ELBO(\mathbf{x}) \right]; ELBO(\mathbf{x}) = \mathbb{E}_{D_{\theta_d}(z|\mathbf{x})} \left[ \log G_{\theta_g}(\mathbf{x}|\mathbf{z}) + \log \frac{p(z)}{G_{\theta_g}(\mathbf{x}|\mathbf{z})} \right]. \quad (38)$$

These methods have mainly been evaluated on the Human Connectome Project (HCP) (Van Essen et al., 2012) dataset and have shown good performance on 4x undersampled images (see Fig 6).

Different from them, PixelCNN+ considers each pixel as random variable and estimates the joint distribution of pixels over an image  $\mathbf{x}$  as the product of conditional distribution, i.e.

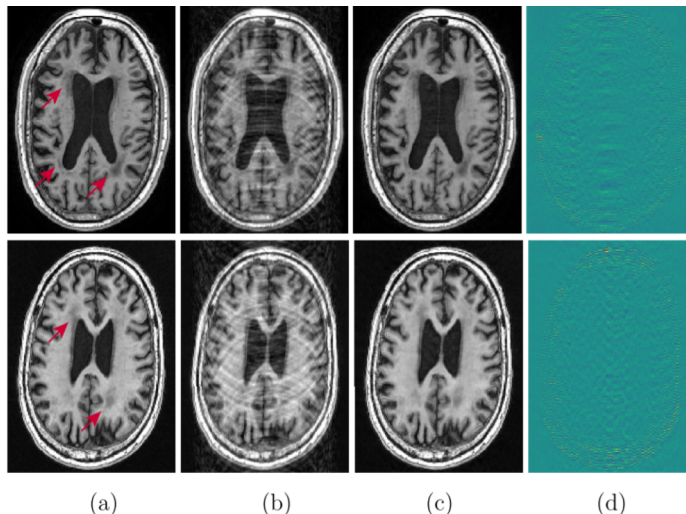


Figure 6: **MR Reconstruction using VAE Based Methods:** Qualitative comparison of (a) original; (b) zero-filled; and (c) VAE based DDP method (Tezcan et al., 2019). Also shown are the difference maps in (d) between DDP and the original images. These results are borrowed from the DDP (Tezcan et al., 2019) paper.

$p(\mathbf{x}) = \prod_{i=1}^{n^2} p(\mathbf{x}_i | \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{i-1})$ . The method proposed in (Luo et al., 2020) considers a generative regression model called PixelCNN+ (Oord et al., 2016) to estimate the prior  $p(\mathbf{x})$ . This method demonstrated very good performance, i.e. they achieved more than 3 dB PSNR improvement than the current state-of-the-art methods like GRAPPA, variational networks (see Sec. 5.3) and SPIRIT algorithms. The Recurrent Inference Machines (RIM) for accelerated MRI reconstruction (Lonning et al., 2018) is an general inverse problem solver that performs a step-wise reassessments of the maximum a posteriori and infers the inverse transform of a forward model. Despite showing good results, the overall computational cost and running time is very high compared to GAN or VAE based methods.

### 4.3 Active Acquisition Methods

**Combined k-space and image methods:** All of the above methods consider a fixed k-space sampling that is predetermined by the user. This sampling process is isolated from the reconstruction pipeline. Recent works have investigated if the sampling process itself can be included as a part of the reconstruction optimization framework. A basic overview of these works can be described as follows:

- The algorithm has access to the fully sampled training MR images  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$
- The encoder,  $G_{\theta_g}(\cdot)$ , learns the sampling pattern by optimizing parameter  $\theta_g$ .
- The decoder,  $D_{\theta_d}(\cdot)$ , is the reconstruction algorithm that is parameterized by  $\theta_d$

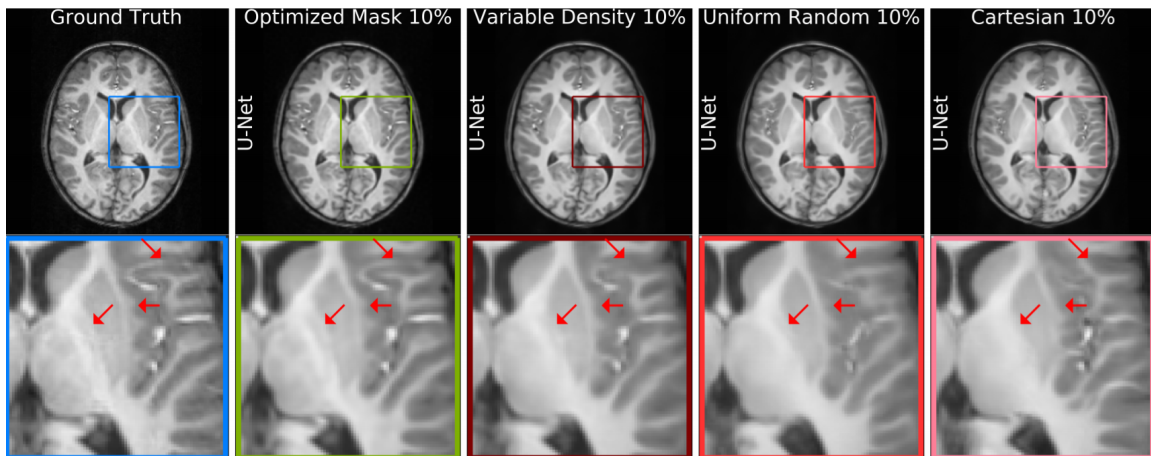


Figure 7: **MR Reconstruction using Active Acquisition Method:** Variable mask learned by the active acquisition method LOUPE (Bahadir et al., 2019) can improve the image quality of a UNet based MR image reconstruction task. The leftmost image is the ground truth, the second image shows the reconstruction using an optimized mask, the next image is the variable mask, that is followed by the uniformly random and Cartesian mask. The images are borrowed from the LOUPE (Bahadir et al., 2019) paper.

- The encoder  $G_{\theta_g}(\cdot)$  is optimized by minimizing the empirical risk on the training MR images,  $\frac{1}{N} \sum_{i=1}^N L_q(\mathbf{x}_i, D_{\theta_d}(G_{\theta_g}(\mathcal{F}(\mathbf{x}))))$ , where  $L_q$  is some arbitrary loss of the decoder.

This strategy was used in LOUPE (Bahadir et al., 2019) where a network was learnt to optimize the under-sampling pattern such that  $G_{\theta_g}(\cdot)$  provided a probabilistic sampling mask  $\mathcal{M}(\cdot)$  assuming each line in k-space as an independent Bernoulli random variable by optimizing:

$$\arg \min_{\theta_g, \theta_d} \mathbb{E}_{\mathcal{M} \sim G_{\theta_g}} [ \|D_{\theta_d}(\mathcal{M}\mathcal{F}\mathbf{x}) - \mathbf{x}\|_1 + \lambda\mathcal{M} ], \quad (39)$$

where  $D_{\theta_d}$  is an anti-aliasing deep neural network. Experiments on  $T_1$ -weighted structural brain MRI scans show that the LOUPE method improves PSNR by  $\sim 5\%$  with respect to the state-of-the-art methods that is shown in Fig. 7, second column. A follow-up work to LOUPE (Bahadir et al., 2020) imposed a hard sparsity constraint on  $G_{\theta_g}(\cdot)$  to ensure robustness to noise. In the deep active acquisition method (Zhang et al., 2019b),  $G_{\theta_g}(\cdot)$  is termed the evaluator and  $D_{\theta_d}(\cdot)$  is the reconstruction network. Given a zero-filled MR image,  $D_{\theta_d}(\mathbf{x}^{ZF})$  provides the reconstructed image and the uncertainty map. The evaluator  $G_{\theta_g}(\cdot)$  decomposes the reconstructed image and ground truth image into spectral maps and provides a score to each k-space line of the reconstructed image. Based on the scores, the methodology decides to acquire the appropriate k-space locations from the MR scanner. The Deep Probabilistic Subsampling (DPS) method in (Huijben et al., 2019) develops a task-adaptive probabilistic undersampling scheme using a softmax based approach followed by MR reconstruction. On the other hand, the work on joint model

based deep learning (J-MoDL) (Aggarwal and Jacob, 2020) optimized both sampling and reconstruction using Eqns 21 and 22 to jointly optimize a data consistency network and a regularization network. The data consistency network is a residual network that acts as a denoiser, while the regularization network decides the sampling scheme. The PILOT (Weiss et al., 2021) method also jointly optimizes the k-space sampling and the reconstruction. The network has a sub-sampling layer to decide the importance of a k-space line, while the regridding and the task layer jointly reconstruct the image. The optimal k-space lines are chosen either using the greedy traveling salesman problem or imposing acquisition machine constraints. Joint optimization of k-space sampling and reconstruction also appeared in recent methods such as (Heng Yu, 2021; Guanxiang Luo, 2021).

## 5. Inverse Mapping using Non-generative Models

In this section we discuss non-generative models that use the following optimization framework:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \mathcal{L}_{\theta_g}; \quad \mathcal{L}_{\theta_g} = \arg \min_{\theta_g} \mathbb{E}_{\mathbf{x} \sim p_{data}} \frac{1}{2} \|\mathbf{x}_j - G_{NGE}(\mathbf{x}|c, \theta_g)\|_2^2. \quad (40)$$

The non-generative models also have a data consistency term and a regularization term similar to Eqn. 4. As discussed earlier in section 1.2, however, the non-generative models do not assume any underlying distribution of the data and learn the inverse mapping by parameter optimization using Eqn. 7. Below, we discuss the different types of non-generative models.

### 5.1 Perceptron Based Models

The work in (Kwon et al., 2017; Cohen et al., 2018) developed a multi-level perceptron (MLP) based learning technique that learns a nonlinear relationship between the k-space measurements, the aliased images, and the desired unaliased images. The input to an MLP is the real and imaginary part of an aliased image and the k-space measurement, and the output is the corresponding unaliased image. We show a visual comparison of this method (Kwon et al., 2017) with the SPIRiT and GRAPPA methods in Fig. 8. This method showed better performance with lower RMSE at different undersampling factors.

### 5.2 Untrained Networks

So far, we have talked about various deep learning architectures and their training strategies using a given training dataset. The most exciting question one can ask “is it always necessary to train a DL network to obtain the best result at test time?”, or “can we solve the inverse problem using DL similar to classical methods that do not necessarily require a training phase to learn the parameter priors”? We note several state-of-the-art methods that use ACS lines or other k-space lines of the k-space measurement  $y$  to train a DL network instead of an MR image as a ground truth. The robust artificial neural network for k-space interpolation (RAKI) (Akçakaya et al., 2019) trains a CNN by using the ACS lines. The RAKI methodology shares some commonality with GRAPPA. However, the main distinction is the linear estimation of the convolution kernel in GRAPPA which is replaced with a non-linear kernel in CNN. The CNN kernels are optimized using the following



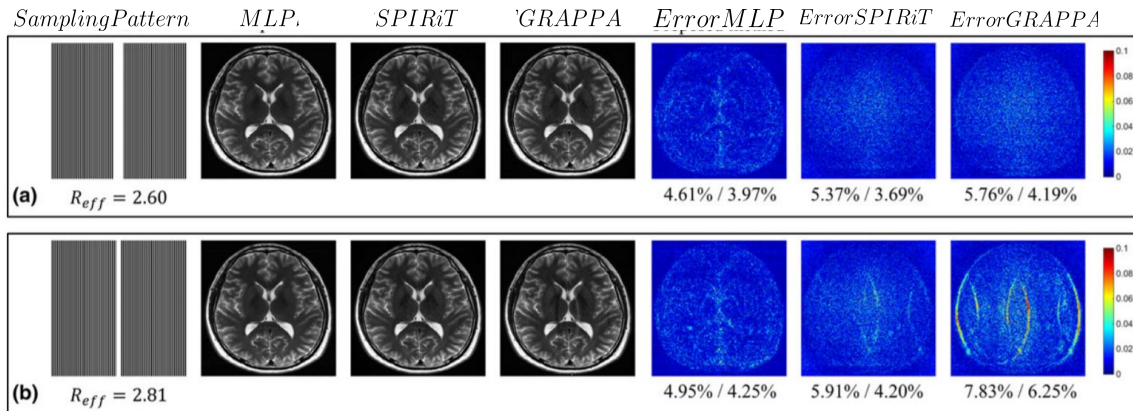


Figure 8: **Visual Comparison with Perceptron, SPIRiT and GRAPPA methods.** We note the dealiasing capability of the perceptron based method at different sampling rates, i.e. row (a) and row (b), for a  $T_2$ -weighted image are visibly better than the classical methods. The figure is borrowed from (Kwon et al., 2017) paper.

objective function:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \mathcal{L}_{\theta_g}; \quad \mathcal{L}_{\theta_g} = \|\mathbf{y}^{ACS} - G_{NGE}(\tilde{\mathbf{y}}^{ACS}; \theta_g)\|_F^2 \quad (41)$$

where,  $y^{ACS}$  are the acquired ACS lines, and  $G_{NGE}(\cdot)$  is the CNN network that performs MRI reconstruction. The RAKI method has shown 0%, 0%, 11%, 28%, and 41% improvement in RMSE score with respect to GRAPPA on phantom images at  $\{2x, 3x, 4x, 5x, 6x\}$  acceleration factors respectively. A followup work called residual RAKI (rRAKI) (Zhang et al., 2019a) improves the RMSE score with the help of a residual network structure. The LORAKI (Kim et al., 2019) method is based on the low rank assumption of LORAKS (Haldar, 2013). It uses a recurrent CNN network to combine the auto-calibrated LORAKS (Haldar, 2013) and the RAKI (Akçakaya et al., 2019) methods. On five different slices of a  $T_2$ -weighted dataset, the LORAKI method has shown good improvement in SSIM scores compared to GRAPPA, RAKI, AC-LORAKS among others. Later, the sRAKI-RNN (Hosseini et al., 2019b) method proposed a unified framework that performs regularization through calibration and data consistency using a more simplified RNN network than LORAKI.

Deep Image Prior (DIP) and its variants (Ulyanov et al., 2018; Cheng et al., 2019; Gandelsman et al., 2019) have shown outstanding results on computer vision tasks such as denoising, in-painting, super resolution, domain translation etc. A vanilla DIP network uses a randomly weighted autoencoder,  $D_{\theta_d}(G_{\theta_g}(\mathbf{z}))$ , that reconstructs a clean image  $x \in \mathbb{R}^{W \times H \times 3}$  given a fixed noise vector  $\mathbf{z} \in \mathbb{R}^{W \times H \times D}$ . The network is optimized using the “ground truth” *noisy* image  $\hat{\mathbf{x}}$ . A manual or user chosen “early stopping” of the optimization is required as optimization until convergence overfits to noise in the image. A recent work called Deep Decoder (Heckel and Hand, 2018) shows that an under-parameterized decoder

network,  $D_{\theta_d}(\cdot)$ , is not expressive enough to learn the high frequency components such as noise and can nicely approximate the denoised version of the image. The Deep Decoder uses pixel-wise linear combinations of channels and shared weights in spatial dimensions that collectively help it to learn relationships and characteristics of nearby pixels. It has been recently understood that such advancements can directly be applied to MR image reconstruction (Mohammad Zalbagi Darestani, 2021). Given a set of k-space measurements  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$  from receiver coils, an un-trained network  $G_{\theta}(\mathbf{z})$  uses an iterative first order method to estimate parameters  $\hat{\theta}$  by optimizing;

$$\min \mathcal{L}_{\theta} = \frac{1}{2} \|\mathbf{y}_i - \mathcal{M}\mathcal{F}G(\mathbf{z}; \theta)\|_2^2. \quad (42)$$

The network is initialized with random weight  $\theta_0$  and then optimized using Eqn. 42 to obtain  $\hat{\theta}$ . The work in (Dave Van Veen, 2021) introduces a feature map regularization:  $\frac{1}{2} \sum_{j=1}^L \|D_j \mathbf{y}_i - \mathcal{M}\mathcal{F}G_{j,i}(\mathbf{z}; \theta)\|_2^2$ , in Eqn. 42 where  $D_j$  matches the features of  $j^{th}$  layer. This term encourages fidelity between the network’s intermediate representations and the acquired k-space measurements. The works in (Heckel, 2019; Heckel and Soltanolkotabi, 2020) provide theoretical guarantees on recovery of image  $x$  from the k-space measurements. Recently proposed method termed “Scan-Specific Artifact Reduction in k-space” or SPARK (Arefeen et al., 2021) trains a CNN to estimate and correct k-space errors made by an input reconstruction technique. The results of this method are also quite impressive given that only ACS lines are used for training the CNN. Along similar lines, the authors in (Yoo et al., 2021) used the Deep Image Prior setup for dynamic MRI reconstruction.

In the self supervised approach, a subset of the undersampled k-space lines are typically used to validate the DL network in addition to the acquired undersampled k-space lines that are used to optimize the network. Work in this direction divides the total k-space lines into two portions: (i) k-space lines for data consistency  $\mathbf{y}^{dc}$  and (ii) k-space lines  $\mathbf{y}^{loss}$  for regularization. In (Yaman et al., 2020), the authors use a multi-fold validation set of k-space data to optimize the DL network. Other methods such as SRAKI (Hosseini et al., 2020, 2019a) use the self-supervision to reconstruct the images. A deep Reinforcement learning based approach is studied in (Jin et al., 2019) that deploys a reconstruction network and an active acquisition network. Another method such as (Yaman et al., 2021b) provides an unrolled optimization algorithm to estimate the missing k-space lines. Other methods that fall under this umbrella include the transformer based method (Korkmaz et al., 2021) and a scan specific optimization method (Yaman et al., 2021a; Tamir et al., 2019).

### 5.3 Convolutional Neural Networks

Spatial models are mostly dominated by the various flavours of CNNs such as complex-valued CNN (Wang et al., 2020b; Cole et al., 2019), unrolled optimization using CNN (Schlemper et al., 2017), variational networks (Hammernik et al., 2018), etc. Depending on how the MR images are reconstructed, we divide all CNN based spatial methods into the following categories.

**Inverse mapping from k-space:** The AUTOMAP (Zhu et al., 2018) learns a reconstruction mapping using a network having three fully connected layers (3 FCs) and two convolutional layers (2 Convs) with an input dimension of  $128 \times 128$ . Any image of size more than  $128 \times 128$  is cropped and subsampled to  $128 \times 128$ . The final model yielded

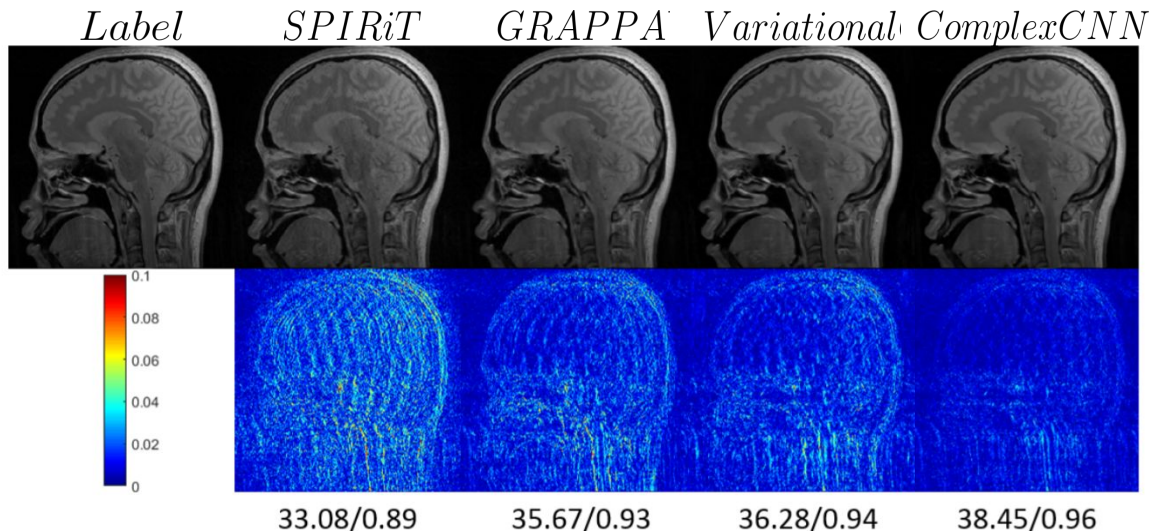


Figure 9: **Comparison of a Complex-valued CNN with other state-of-the-art methods:** We show visual comparison of SPIRiT (Lustig and Pauly, 2010), GRAPPA (Bouman and Sauer, 1993), VariationalNet, and the ComplexCNN. We note from the difference map that the ComplexCNN performed well with respect to the other state-of-the-art methods. The PSNR and SSIM values are given at the bottom for each method.

a PSNR of 28.2 on FastMRI knee dataset outperforming the previous validation baseline of 25.9 on the same dataset. Different from these methods, there are a few works (Wang et al., 2020b; Cole et al., 2019) that have used CNN networks with complex valued kernels to reconstruct MR images from complex valued k-space measurements. The method in (Wang et al., 2020a) uses a complex valued ResNet (that is a type of CNN) network and is shown to obtain good results on 12-channel fully sampled k-space dataset (see Fig. 9 for a visual comparison with other methods). Another method uses a Laplacian pyramid-based complex neural network (Liang et al., 2020b) for MR image reconstruction.

**Inverse Mapping for Image Rectification:** In CNN based sequential spatial models such as DeepADMM net models (Sun et al., 2016; Schlemper et al., 2017) and Deep Cascade CNN (DCCNN) (Schlemper et al., 2017), the regularization is done in image space using the following set of equations:

$$\arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \|\mathbf{x} + \beta - \mathbf{z}\|_2^2; \quad \arg \min_{\mathbf{z}_i} \sum_i [\lambda_i g(\Gamma_i \mathbf{z}_i) \|\mathbf{x} + \beta - \mathbf{z}\|_2^2] \quad (43)$$

$$\beta \leftarrow \beta + \alpha(\mathbf{x} - \mathbf{z}).$$

Here,  $\alpha$  and  $\beta$  are Lagrange multipliers. The ISTA net (Zhang and Ghanem, 2018) modifies the above image update rule as follows,  $\mathbf{x}_i = \arg \min_{\mathbf{x}} \frac{1}{2} \|C(\mathbf{x}) - C(\mathbf{z}_i)\|_2^2 + \lambda \|C(\mathbf{x})\|_1$ , using a CNN network  $C(\cdot)$ . Note that the DeepADMM network demonstrated good performance when the network was trained on brain data but tested on chest data. Later, MODL

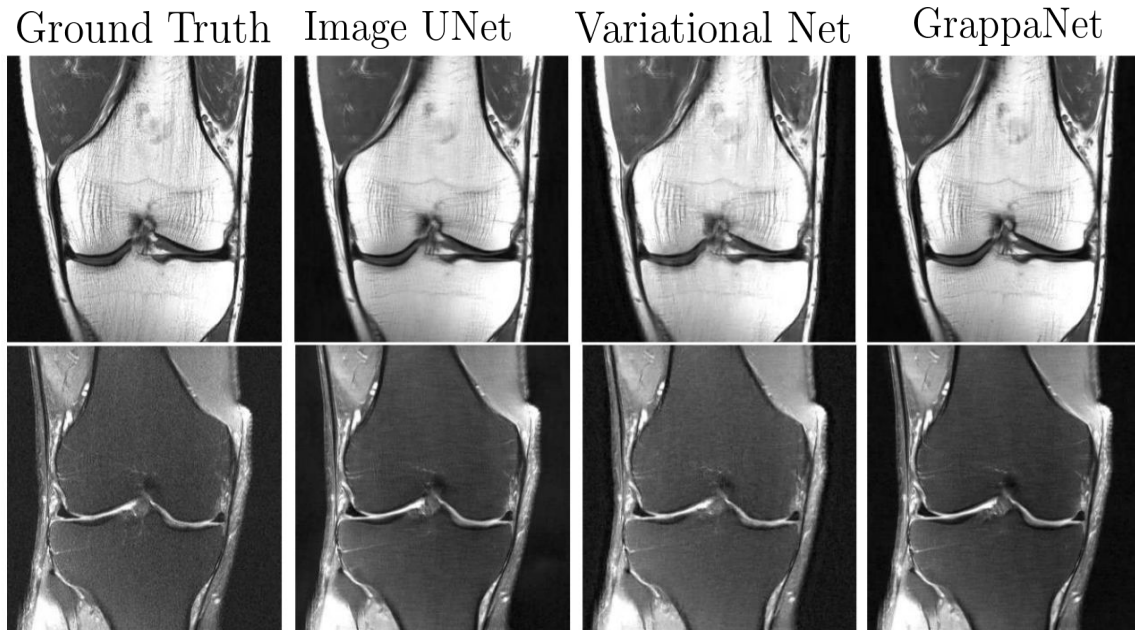


Figure 10: **GrappaNet Comparison with VariationalNET, UNet methods**

(Aggarwal et al., 2019) proposed a model based MRI reconstruction where they used a convolution neural network (CNN) based regularization prior. Later, a dynamic MRI using MODL based deep learning was proposed by (Biswas et al., 2019). The optimization, i.e.  $\arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|C(\mathbf{x})\|_2^2$ , denoises the alias artifacts and noise using a CNN network  $C(\cdot)$  as a regularization prior, and  $\lambda$  is a trainable parameter. To address this concern, a full end-to-end CNN model called GrappaNet (Sriram et al., 2020b) was developed, which is a nonlinear version of GRAPPA set within a CNN network. The CNN network has two sub networks; the first sub network,  $f_1(\mathbf{y})$ , fills the missing k-space lines using a non-linear CNN based interpolation function similar to GRAPPA. Subsequently, a second network,  $f_2$ , maps the filled k-space measurement to the image space. The GrappaNet model has shown excellent performance (40.74 PSNR, 0.957 SSIM) on the FastMRI dataset and is one of the best performing methods. A qualitative comparison is shown in Fig. 10. Along similar lines, a deep variational network (Hammernik et al., 2018) is used to MRI reconstruction. Other works, such as (Wang et al., 2016; Cheng et al., 2018; Aggarwal et al., 2018) train the parameters of a deep network by minimizing the reconstruction error between the image from zero-filled k-space and the image from fully sampled k-space. The cascaded CNN network learns spatio-temporal correlations efficiently by combining convolution and data sharing approaches in (Schlemper et al., 2017). The (Seegoolam et al., 2019) method proposed to use a CNN network to estimate motions from undersampled MRI sequences that is used to fuse data along the entire temporal axis.

## 5.4 Recurrent Neural Networks

**Inverse mapping from k-space:** We note that a majority of the iterative temporal networks, a.k.a the recurrent neural network models, are k-space to image space reconstruction methods and typically follow the optimization described in Section 5.3. The temporal methods, by design, are classified into two categories, namely (i) regularization methods, and (ii) variable splitting methods.

Several state-of-the-art methods have considered temporal methods as a way to regularize using the iterative hard threshold (IHT) method from (Blumensath and Davies, 2009) that approximates the  $l_0$  norm. Mathematically, the IHT update rule is as follows:

$$\mathbf{x}_{t+1} = H_k[\mathbf{x}_t - \alpha\Phi^T(\Phi\mathbf{x}_t - \mathbf{y})], \quad (44)$$

where  $\alpha$  is the step-size parameter,  $H_k[\cdot]$  is the operator that sets all but  $k$ -largest values to zero (proxy for  $l_0$  operation), and the dictionary  $\Phi$  satisfies the restricted isometry property (RIP)<sup>5</sup>. The work in (Xin et al., 2016) shows that this hard threshold operator resembles the memory state of the LSTM network. Similar to the clustering based sparsity pattern of IHT, the gates of LSTM inherently promotes sparsity. Along similar lines, the Neural Proximal Gradient Descent work (Mardani et al., 2018b) envisioned a one-to-one correspondence between the proximal gradient descent operation and the update of a RNN network. Mathematically, an iteration of a proximal operator  $P_f$  given by:  $\mathbf{x}_{t+1} = P_f(\mathbf{x}_t + \alpha\Phi^H(\mathbf{y} - \Phi\mathbf{x}))$ , resembles the LSTM update rule:

$$\mathbf{s}_{t+1} = g(\mathbf{x}_t; \mathbf{y}); \quad \mathbf{x}_{t+1} = P_f(\mathbf{s}_{t+1}), \quad (45)$$

where  $g(\mathbf{x}_t; \mathbf{y}) = \mathbf{x}_t + \alpha\Phi^H(\mathbf{y} - \Phi\mathbf{x})$  is the update step,  $\mathbf{s}_{t+1}$  is the hidden state and  $\Phi$  is a dictionary. Different from these, a local-global recurrent neural network is proposed in (Guo et al., 2021) that uses two recurrent networks, one network to capture high frequency components, and another network to capture the low frequency components. The method in (Oh et al., 2021) uses a bidirectional RNN and replaces the dense network structure of (Zhu et al., 2018) while removing aliasing artifacts in the reconstructed image.

The Convolutional Recurrent Neural Networks or CRNN (Qin et al., 2018) method proposed a variable splitting and alternate minimisation method using a RNN based model. Recovering finer details was the main challenge of the PyramidRNN (Wang et al., 2019) that proposed to reconstruct images at multiple scales. Three CRNNs are deployed to reconstruct images at different scales, i.e.  $\mathbf{x}^0 = CRNN(\mathbf{x}^{zero}, \mathbf{y}, \theta_1)$ ,  $\mathbf{x}^1 = CRNN(\mathbf{x}^0, \mathbf{y}, \theta_2)$ ,  $\mathbf{x}^2 = CRNN(\mathbf{x}^1, \mathbf{y}, \theta_3)$ , and the final data consistency is performed after  $\mathbf{x}^0, \mathbf{x}^1, \mathbf{x}^2$  are combined using another CNN. The CRNN is used as a recurrent neural network in the variational approach of VariationNET (Sriram et al., 2020a), i.e.

$$\mathbf{x}_{t+1} = \mathbf{x}_t - \alpha\Phi^H(\mathbf{y} - \Phi\mathbf{x}) + D(\mathbf{x}_t), \quad (46)$$

where  $D(\cdot)$  is a CRNN network that provides MR reconstruction. In this unrolled optimization method, the CRNN is used as a proximal operator to reconstruct the MR image. The

---

5. **Restricted Isometry Property (RIP):** The projection from the measurement matrix  $E$  in Eqn. 17 should preserve the distance between two MR images  $\mathbf{x}_1$  and  $\mathbf{x}_2$  bounded by factors of  $1 - \delta$  and  $1 + \delta$ , i.e.  $(1 - \delta)\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2 \leq \|E(\mathbf{x}_1 - \mathbf{x}_2)\|_2^2 \leq (1 + \delta)\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$ , where  $\delta$  is a small constant.

Acceleration	Model	NMSE	PSNR	SSIM	Acceleration	Model	NMSE	PSNR	SSIM
4-fold	Zero Filled	0.0198	32.51	0.811	8-fold	Zero Filled	0.0352	29.60	0.642
	SENSE (Pruessmann et al., 1999)	0.0154	32.79	0.816		SENSE (Pruessmann et al., 1999)	0.0261	31.65	0.762
	GRAPPA (Griswold et al., 2002)	0.0104	27.79	0.816		GRAPPA(Griswold et al., 2002)	0.0202	25.31	0.782
	RfineGAN(Quan et al., 2018)	0.0138	34.00	0.901		RefineGAN(Quan et al., 2018)	0.0221	32.09	0.792
	DeepADMM (Sun et al., 2016)	0.0055	34.52	0.895		DeepADMM(Sun et al., 2016)	0.0201	36.37	0.810
	LORAKI (Kim et al., 2019)	0.0091	35.41	0.871		LORAKI (Kim et al., 2019)	0.0181	36.45	0.882
	VariationNET (Sriram et al., 2020a)	0.0049	38.82	0.919		VariationNET (Sriram et al., 2020a)	0.0211	36.63	0.788
	GrappaNet (Sriram et al., 2020b)	0.0026	40.74	0.957		GrappaNet (Sriram et al., 2020b)	0.0071	36.76	0.922
	J-MoDL (Aggarwal and Jacob, 2020)	0.0021	41.53	0.961		J-MoDL (Aggarwal and Jacob, 2020)	0.0065	35.08	0.928
	Deep Decoder (Heckel and Hand, 2018)	0.0132	31.67	0.938		Deep Decoder (Heckel and Hand, 2018)	0.0079	29.654	0.929
DIP (Ulyanov et al., 2018)	0.0113	30.46	0.923	DIP (Ulyanov et al., 2018)	0.0076	29.18	0.912		

Table 1: Normalized mean squared error (NMSE), PSNR and SSIM over the test data with 1000 samples using eight different methods at 4-fold and 8-fold acceleration factors.

VariationNET is a followup work of Deep Variational Network of (Hammernik et al., 2018) that we discussed in Sec. 5.3. The VariationalNET unrolls an iterative algorithm involving a CRNN based recurrent neural network based regularizer, while the Deep Variational Network of (Hammernik et al., 2018) unrolls an iterative algorithm involving a receptive field based convolutional regularizer.

## 5.5 Hypernetwork Models

Hypernetworks are meta-networks that regress optimal weights of a task network (often called as data network (Pal and Balasubramanian, 2019) or main network (Ha et al., 2016)). The data network  $G_{NGE}(\mathbf{x}^{low}, \theta_g)$  performs the mapping from aliased or the low resolution images  $\mathbf{x}^{low}$  to the high resolution MR images  $\mathbf{x}^{gen}$ . The hypernetwork  $H(\alpha, \theta_{hyp})$  estimates weights  $\theta_g$  of the network  $G_{NGE}(\theta_g)$  given the random variable  $\alpha$  sampled from a prior distribution  $\alpha \sim p(\alpha)$ . The end-to-end network is trained by optimizing:

$$\arg \min_{\psi} \mathbb{E}_{\theta_g \sim H(\alpha, \theta_{hyp})} [ \|G_{NGE}(\mathbf{x}^{low}, \theta_g) - \mathbf{x}\|_2^2 + \mathcal{R}(G_{NGE}(\mathbf{x}^{low}, \theta_g)) ]. \quad (47)$$

In (Wang et al., 2021), the prior distribution  $p(\alpha)$  is a uniform distribution  $\mathcal{U}[-1, +1]$  (and the process is called Uniform Hyperparameter Sampling) or the sampling can be based on the data density (called data-driven hyperparameter sampling). Along similar lines, the work in (Ramanarayanan et al., 2020) trained a dynamic weight predictor (DWP) network that provides layer wise weights to the data network. The DWP generates the layer wise weights given the context vector  $\gamma$  that comprises of three factors such as the anatomy under study, undersampling mask pattern and the acceleration factor.

## 6. Comparison of state-of-the-art methods

Given the large number of DL methods being proposed, it is imperative to compare these methods on a standard publicly available dataset. Many of these methods have shown their effectiveness on various real world datasets using different quantitative metrics such as SSIM, PSNR, RMSE, etc. There is, however, a scarcity of qualitative and quantitative comparison of these methods on a single dataset. While the fastMRI challenge allowed comparison of several methods, yet, several recent methods from the categories discussed above were not part of the challenge. Consequently, we compare a few representative MR

reconstruction methods both qualitatively and quantitatively on the fastMRI knee dataset (Zbontar et al., 2018). We note that, doing a comprehensive comparison of all the methods mentioned in this review is not feasible due to non-availability of the code as well as the sheer magnitude of the number of methods (running into hundreds). We compared the following representative models:

- Zero filled image reconstruction method
- Classical image space based SENSE method (Pruessmann et al., 1999)
- Classical k-space based GRAPPA method (Griswold et al., 2002)
- Unrolled optimization based method called DeepADMM (Sun et al., 2016)
- Low rank based LORAKI (Kim et al., 2019)
- Generative adversarial network based RefineGAN (Quan et al., 2018) network
- Variational network called VariationNET (Sriram et al., 2020a)
- The deep k-space method GrappaNet (Sriram et al., 2020b)
- Active acquisition based method J-MoDL (Aggarwal and Jacob, 2020)
- Untrained network model Deep Decoder (Heckel and Hand, 2018) and
- Deep Image Prior DIP (Ulyanov et al., 2018) method.

The fastMRI knee dataset consists of raw k-space data from 1594 scans acquired on four different MRI machines. We used the official training, validation and test data split in our experiments. We did not use images with a width greater than 372 and we note that such data is only 7% of the training data split. Both the 4x and 8x acceleration factors were evaluated.

We used the original implementation<sup>6</sup> of GrappaNet, VariationaNET, SENSE, GRAPPA, DeepADMM, Deep Decoder, and DIP method. Similar to GrappaNET, we always use the central 30 k-space lines to compute the training target. Treating the real and imaginary as two distinct channels, we dealt with the complex valued input, i.e. we have 30 channel input k-space measurements for the 15-coil complex-valued k-space. Where applicable, the models were trained with a linear combination of  $L_1$  and SSIM loss, i.e.

$$J(\hat{\mathbf{x}}, \mathbf{x}) = -SSIM(G(\mathbf{y}), \mathbf{x}) + \lambda \|G(\mathbf{y}) - \mathbf{x}\|_1 \quad (48)$$

where  $\lambda$  is a hyperparameter,  $G(\mathbf{y})$  is the model prediction, and  $\mathbf{x}$  is the ground truth.

---

6. Below are the official implementations of various methods we discussed:

**VariationaNET:** <https://github.com/VLOGroup/mri-variationalnetwork/>

**GrappaNET:** <https://github.com/facebookresearch/fastMRI.git>

**RefineGAN:** <https://github.com/tmquan/RefineGAN.git>

**DeepADMM:** <https://github.com/yangyan92/Deep-ADMM-Net.git>

**SENSE, GRAPPA:** <https://mrirecon.github.io/bart/>

**Deep Decoder:** <https://github.com/MLI-lab/ConvDecoder.git>

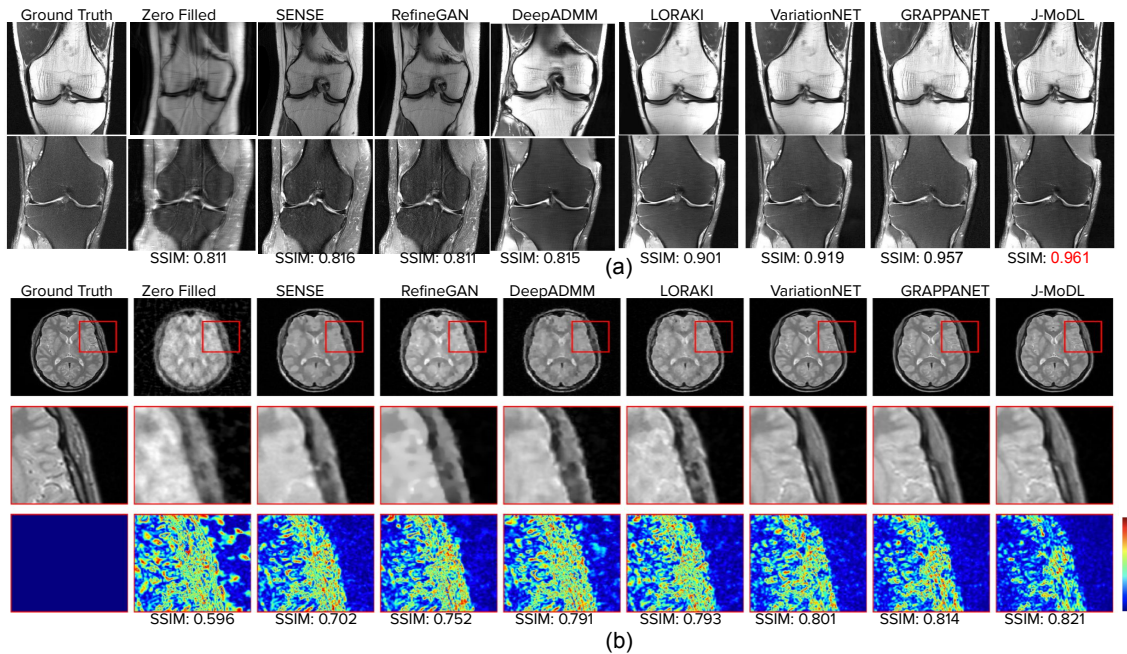


Figure 11: **Qualitative Comparison:** Comparison for 8x acceleration using “Zero Filled”, SENSE, DeepADMM (Sun et al., 2016), LORAKI (Kim et al., 2019), RefineGAN (Quan et al., 2018), VariationNET (Sriram et al., 2020a), GrappaNet (Sriram et al., 2020b), J-MoDL (Aggarwal and Jacob, 2020) results on fastMRI dataset. We show qualitative results for (a) knee and (b) brain datasets and also report the corresponding SSIM scores.

Quantitative results are shown in Table 1 for several metrics such as NMSE, PSNR, and SSIM scores. We observe that GrappaNET, J-MoDL and VariationNET outperformed the baseline methods by a large margin. We note that the zero-filled and SENSE reconstructions in Fig 11 (a), (b) show a large amount of over-smoothing. The reconstruction of SENSE and zero-filled model also lack a majority of the high frequency detail that is clinically relevant, but fine details are visible in case of GrappaNET, VariationNET, J-MoDL, and RefineGAN methods. The comparable performance of Deep Decoder and DIP advocates the importance of letting untrained neural network figure out how to perform k-space to MR image reconstruction. The J-MoDL method makes heavy use of training data and the joint optimization of k-space lines and the reconstruction of MR images to get good results both for  $4\times$  and  $8\times$  as shown in Table 1. On the other hand, the Deep Decoder and DIP methods achieve good performance using untrained networks as discussed in Sec. 5.2, which is advantageous as it generalizes to any MR reconstruction scenario.

## 7. Discussion

In this paper, we discussed and reviewed several classical reconstruction methods, as well as deep generative and non-generative methods to learn the inverse mapping from k-space



to image space. Naturally, one might ask the following questions given the review of several papers above: “are DL methods free from errors?”, “do they always generalize well?”, and “are they robust?”. To better understand the above mentioned rhetorical questions, we need to discuss several aspects of the performance of these methods such as (i) correct reconstruction of minute details of pathology and anatomical structures; (ii) risk quantification; (iii) robustness; (iv) running time complexity; and (v) generalization.

Due to the blackbox-like nature of DL methods, the reliability and risk quantification associated with them are often questioned. In a recent paper on “risk quantification in Deep MRI reconstruction” (Edupuganti et al., 2020), the authors strongly suggest for quantifying the risk and reliability of DL methods and note that it is very important for accurate patient diagnoses and real world deployment. The paper also shows how Stein’s Unbiased Risk Estimator (SURE) (Metzler et al., 2018) can be used as a way to assess uncertainty of the DL model:

$$SURE = -n\sigma^2 + \|\hat{\mathbf{x}} - \mathbf{x}\|^2 + \sigma^2 \text{trace}\left(\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}}\right), \quad (49)$$

where the second term represents the end-to-end network sensitivity to small input perturbations. This formulation works even when there is no access to the ground truth data  $\mathbf{x}$ . In this way, we can successfully measure the risk associated with a DL model. In addition to the SURE based method, there are a few other ways to quantify the risk and reliability associated with a DL model. The work “On instabilities of deep learning in image reconstruction” (Antun et al., 2019) uses a set of pretrained models such as AUTOMAP (Zhu et al., 2018), DAGAN (Yang et al., 2017), or Variational Network (Sriram et al., 2020a) with noisy measurements to quantify their stability. This paper as well as several others (Narnhofer et al., 2021; Antun et al., 2019) have discussed how the stability of the reconstruction process is related to the network architecture, training set and also the subsampling pattern.

Whether a DL method can capture high frequency components or not is also another area of active research in MR reconstruction. The robustness of a DL based MR reconstruction method is also studied in various papers such as (Raj et al., 2020; Cheng et al., 2020a; Calivá et al., 2020; Zhang et al., 2021). For example, the (Cheng et al., 2020a; Calivá et al., 2020; Cheng et al., 2020b) works perceived adversarial attack as a way to capture minute details during sMR reconstruction thus showing a significant improvement in robustness compared to other methods. The work proposed to train a deep learning network with a loss that pays special attention to small anatomical details. The methodology progressively adds minuscule perturbation  $\delta$  to the input  $x$  not perceivable to human eye but may shift the decision of a DL system. The method in (Raj et al., 2020) uses a generative adversarial framework that entails a perturbation generator network  $G(\cdot)$  to add minuscule distortion on  $k$ -space measurement  $y$ . The work in (Zhang et al., 2021) proposed to incorporate fast gradient based attack on a zero filled input image and train the deep learning network not to deviate much under such attack. The FINE (Zhang et al., 2020) methodology, on the other hand, has used pre-trained recon network that was fine-tuned using data consistency to reduce generalization error inside unseen pathologies. Please refer to the paper (Darestani et al., 2021) which provides a summary of robustness of different approaches for image reconstruction.

Optimizing a DL network is also an open area of active research. The GAN networks suffer from a lack of proper optimization of the structure of the network (Goodfellow et al.,

2014). On the other hand, the VAE and Bayesian methods suffer from the large training time complexities. We see several active research groups and papers (Salimans et al., 2016; Bond-Taylor et al., 2021) in Computer Vision and Machine Learning areas pondering upon these questions and coming up with good solutions. Also, the work in (Hammernik et al., 2017) has shown the effectiveness of capturing perceptual similarity using the  $l_2$  and SSIM loss to include local structure in the reconstruction process. Recently, work done in (Maarten Terpstra, 2021) shows that the standard  $l_2$  loss prefers or biases the reconstruction with lower image magnitude and hence they propose a new loss function between ground truth  $\mathbf{x}^{gt}$  and the reconstructed complex-valued images  $\mathbf{x}$ , i.e.  $L_{perp} = P(\mathbf{x}, \mathbf{x}^{gt}) + l_1(\mathbf{x}, \mathbf{x}^{gt})$  where  $P(\mathbf{x}, \mathbf{x}^{gt}) = |\text{Real}(\mathbf{x})\text{Imaginary}(\mathbf{x}^{gt}) - \text{Real}(\mathbf{x}^{gt})\text{Imaginary}(\mathbf{x})|/|\mathbf{x}^{gt}|$  which favours the finer details during the reconstruction process. The proposed loss function achieves better performance and faster convergence on complex image reconstruction tasks.

Regarding generalization, we note that some DL based models have shown remarkable generalization capabilities, for example: the AUTOMAP work was trained on natural images but generalized well for MR reconstruction. However, current hardware (memory) limitations preclude from using this method for high resolution MR reconstruction. On the other hand, some of the latest algorithms that show exceptional performance on the knee dataset have not been extensively tested on other low SNR data. In particular, these methods also need to be tested on quantitative MR modalities to better assess their performance.

Another bottleneck for using these DL methods is the large amount of training data required. While GAN and Bayesian networks produce accurate reconstruction of minute details of anatomical structures if sufficient data are available at the time of training, it is not clear as to how much training data is required and whether the networks can adapt quickly to change in resolution and field-of-view. Further, these works have not been tested in scenarios where changes in MR acquisition parameters such as relaxation time (TR), echo time (TE), spatial resolution, number of channels, and undersampling pattern are made at test time and are different from the training dataset.

Most importantly, MRI used for diagnostic purposes should be robust and accurate in reconstructing images of pathology (major and minor). While training based methods have demonstrated their ability to reconstruct normal looking images, extensive validation of these methods on pathological datasets is needed for adoption in clinical settings. To this end, the MR community needs to collaborate and collect normative and pathological datasets for testing. We specifically note that, the range of pathology can vary dramatically in the anatomy being imaged (e.g., the size, shape, location and type of tumor). Thus, extensive training and unavailability of pathological images may present significant challenges to methods that are data hungry. In contrast, untrained networks may provide MR reconstruction capabilities that are significantly better than the current state-of-the-art but do not perform as well as highly trained networks (but generalize well to unknown scenarios).

Finally, given the exponential rate at which new DL methods are being proposed, several standardized datasets with different degrees of complexity, noise level (for low SNR modalities) and ground truth availability are required to perform a fair comparison between methods. Additionally, fully-sampled raw data (with different sampling schemes) needs to be made available to compare for different undersampling factors. Care must be taken not to obtain data that have already been “accelerated” with the use of standard GRAPPA-like methods, which might bias the results (Efrat Shimron, 2021).

Nevertheless, recent developments using new DL methods point to the great strides that have been made in terms of data reconstruction quality, risk quantification, generalizability and reduction of running time complexity. We hope that this review of DL methods for MR image reconstruction will give researchers a unique viewpoint and summarize in succinct terms the current state-of-the-art methods. We however humbly note that, given the large number of methods presented in the literature, it is impossible to cite and categorize each one of them. As such, in this review, we collected and described broad categories of methods based on the type of methodology used for MR reconstruction.

## Acknowledgments

This work was supported by NIH grant: R01MH116173 (PIs: Setsompop, Rathi).

## Ethical Standards

This work used data from human subjects that is openly available (fastMRI) and was acquired following all applicable regulations as required by the local IRB.

## Conflicts of Interest

None.

## References

- Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL <https://www.tensorflow.org/>. Software available from tensorflow.org.
- Hemant K Aggarwal, Merry P Mani, and Mathews Jacob. Modl: Model-based deep learning architecture for inverse problems. *IEEE transactions on medical imaging*, 38(2):394–405, 2018.
- Hemant K. Aggarwal, Merry P. Mani, and Mathews Jacob. Modl: Model-based deep learning architecture for inverse problems. *IEEE Transactions on Medical Imaging*, 38(2):394–405, 2019. doi: 10.1109/TMI.2018.2865356.
- Hemant Kumar Aggarwal and Mathews Jacob. J-modl: Joint model-based deep learning for optimized sampling and reconstruction. *IEEE Journal of Selected Topics in Signal Processing*, 14(6): 1151–1162, 2020.

- Mehmet Akçakaya, Steen Moeller, Sebastian Weingärtner, and Kâmil Uğurbil. Scan-specific robust artificial-neural-networks for k-space interpolation (raki) reconstruction: Database-free deep learning for fast imaging. *Magnetic resonance in medicine*, 81(1):439–453, 2019.
- Vegard Antun, Francesco Renna, Clarice Poon, Ben Adcock, and Anders C Hansen. On instabilities of deep learning in image reconstruction-does ai come at a cost? *arXiv preprint arXiv:1902.05300*, 2019.
- Yamin Arefeen, Onur Beker, Heng Yu, Elfar Adalsteinsson, and Berkin Bilgic. Scan specific artifact reduction in k-space (spark) neural networks synergize with physics-based reconstruction to accelerate mri. *arXiv preprint arXiv:2104.01188*, 2021.
- Cagla D Bahadir, Alan Q Wang, Adrian V Dalca, and Mert R Sabuncu. Deep-learning-based optimization of the under-sampling pattern in mri. *IEEE Transactions on Computational Imaging*, 6:1139–1152, 2020.
- Cagla Deniz Bahadir, Adrian V Dalca, and Mert R Sabuncu. Learning-based optimization of the under-sampling pattern in mri. In *International Conference on Information Processing in Medical Imaging*, pages 780–792. Springer, 2019.
- Noam Ben-Eliezer, Daniel K Sodickson, Timothy Shepherd, Graham C Wiggins, and Kai Tobias Block. Accelerated and motion-robust in vivo t<sub>2</sub> mapping from radially undersampled data using bloch-simulation-based iterative reconstruction. *Magnetic resonance in medicine*, 75(3):1346–1354, 2016.
- Julian Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society: Series B (Methodological)*, 48(3):259–279, 1986.
- Sampurna Biswas, Hemant K Aggarwal, and Mathews Jacob. Dynamic mri using model-based deep learning and storm priors: Modl-storm. *Magnetic resonance in medicine*, 82(1):485–494, 2019.
- Thomas Blumensath and Mike E Davies. Iterative hard thresholding for compressed sensing. *Applied and computational harmonic analysis*, 27(3):265–274, 2009.
- Sam Bond-Taylor, Adam Leach, Yang Long, and Chris G Willcocks. Deep generative modelling: A comparative review of vaes, gans, normalizing flows, energy-based and autoregressive models. *arXiv preprint arXiv:2103.04922*, 2021.
- Emrah Bostan, Ulugbek Kamilov, and Michael Unser. Reconstruction of biomedical images and sparse stochastic modeling. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 880–883. Ieee, 2012.
- Charles Bouman and Ken Sauer. A generalized gaussian image model for edge-preserving map estimation. *IEEE Transactions on image processing*, 2(3):296–310, 1993.
- Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.
- Claire Boyer, Jérémie Bigot, and Pierre Weiss. Compressed sensing with structured sparsity and structured acquisition. *Applied and Computational Harmonic Analysis*, 46(2):312–350, 2019.
- Yoram Bresler and Ping Feng. Spectrum-blind minimum-rate sampling and reconstruction of 2-d multiband signals. In *Proceedings of 3rd IEEE International Conference on Image Processing*, volume 1, pages 701–704. IEEE, 1996.

- Jason Brownlee. A gentle introduction to the rectified linear unit (relu). *Machine learning mastery*, 6, 2019.
- Alfred M Bruckstein, David L Donoho, and Michael Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM review*, 51(1):34–81, 2009.
- Jose Caballero, Anthony N Price, Daniel Rueckert, and Joseph V Hajnal. Dictionary learning and time sparsity for dynamic mr data reconstruction. *IEEE transactions on medical imaging*, 33(4):979–994, 2014.
- Francesco Calivá, Kaiyang Cheng, Rutwik Shah, and Valentina Pedoia. Adversarial robust training in mri reconstruction. *arXiv preprint arXiv:2011.00070*, 2020.
- Emmanuel J Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory*, 52(2):489–509, 2006.
- Yue Cao and David N Levin. Using prior knowledge of human anatomy to constrain mr image acquisition and reconstruction: half k-space and full k-space techniques. *Magnetic resonance imaging*, 15(6):669–677, 1997.
- Yuchou Chang, Dong Liang, and Leslie Ying. Nonlinear grappa: A kernel approach to parallel mri reconstruction. *Magnetic Resonance in Medicine*, 68(3):730–740, 2012.
- Rick Chartrand. Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signal Processing Letters*, 14(10):707–710, 2007.
- Rick Chartrand and Valentina Staneva. Restricted isometry properties and nonconvex compressive sensing. *Inverse Problems*, 24(3):035020, 2008.
- C-T Chen, XIAOLONG Ouyang, Win H Wong, Xiaoping Hu, VE Johnson, C Ordonez, and CE Metz. Sensor fusion in image reconstruction. *IEEE Transactions on Nuclear Science*, 38(2):687–692, 1991.
- Guang-Hong Chen, Jie Tang, and Shuai Leng. Prior image constrained compressed sensing (piccs). In *Photons Plus Ultrasound: Imaging and Sensing 2008: The Ninth Conference on Biomedical Thermoacoustics, Optoacoustics, and Acousto-optics*, volume 6856, page 685618. International Society for Optics and Photonics, 2008.
- Joseph Y Cheng, Feiyu Chen, Marcus T Alley, John M Pauly, and Shreyas S Vasanawala. Highly scalable image reconstruction using deep neural networks with bandpass filtering. *arXiv preprint arXiv:1805.03300*, 2018.
- Kaiyang Cheng, Francesco Calivá, Rutwik Shah, Misung Han, Sharmila Majumdar, and Valentina Pedoia. Addressing the false negative problem of deep learning mri reconstruction models by adversarial attacks and robust training. In *Medical Imaging with Deep Learning*, pages 121–135. PMLR, 2020a.
- Kaiyang Cheng, Francesco Calivá, Rutwik Shah, Misung Han, Sharmila Majumdar, and Valentina Pedoia. Addressing the false negative problem of deep learning mri reconstruction models by adversarial attacks and robust training. In Tal Arbel, Ismail Ben Ayed, Marleen de Bruijne, Maxime Descoteaux, Herve Lombaert, and Christopher Pal, editors, *Proceedings of the Third Conference on Medical Imaging with Deep Learning*, volume 121 of *Proceedings of Machine Learning Research*, pages 121–135. PMLR, 06–08 Jul 2020b. URL <https://proceedings.mlr.press/v121/cheng20a.html>.

- Zezhou Cheng, Matheus Gadelha, Subhransu Maji, and Daniel Sheldon. A bayesian perspective on the deep image prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5443–5451, 2019.
- Ouri Cohen, Bo Zhu, and Matthew S Rosen. Mr fingerprinting deep reconstruction network (drone). *Magnetic resonance in medicine*, 80(3):885–894, 2018.
- Elizabeth K Cole, John Pauly, and J Cheng. Complex-valued convolutional neural networks for mri reconstruction. In *In proceedings of the 27th Annual Meeting of ISMRM, Montreal, Canada*, page 4714, 2019.
- Elizabeth K Cole, John M Pauly, Shreyas S Vasanaawala, and Frank Ong. Unsupervised mri reconstruction with generative adversarial networks. *arXiv e-prints*, pages arXiv–2008, 2020.
- Mohammad Zalbagi Darestani, Akshay Chaudhari, and Reinhard Heckel. Measuring robustness in deep learning based compressive sensing. *arXiv preprint arXiv:2102.06103*, 2021.
- et al Dave Van Veen. Using untrained convolutional neural networks to accelerate mri in 2d and 3d. In *Proceedings of the 29th Annual Meeting of ISMRM*, 2021.
- Aaron Defazio, Tullie Murrell, and Michael Recht. Mri banding removal via adversarial training. *Advances in Neural Information Processing Systems*, 33, 2020.
- Puneesh Deora, Bhavya Vasudeva, Saumik Bhattacharya, and Pyari Mohan Pradhan. Structure preserving compressive sensing mri reconstruction using generative adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 522–523, 2020.
- Vineet Edupuganti, Morteza Mardani, Shreyas Vasanaawala, and John M Pauly. Risk quantification in deep mri reconstruction. In *NeurIPS 2020 Workshop on Deep Learning and Inverse Problems*, 2020.
- et al. Efrat Shimron. Subtle inverse crimes: Naively using publicly available images could make reconstruction results seem misleadingly better! In *Proceedings of the 29th Annual Meeting of ISMRM*, 2021.
- Ping Feng and Yoram Bresler. Spectrum-blind minimum-rate sampling and reconstruction of multi-band signals. In *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*, volume 3, pages 1688–1691. IEEE, 1996.
- Jeffrey A Fessler. Model-based image reconstruction for mri. *IEEE signal processing magazine*, 27(4):81–89, 2010.
- Jeffrey A Fessler and Bradley P Sutton. Nonuniform fast fourier transforms using min-max interpolation. *IEEE transactions on signal processing*, 51(2):560–574, 2003.
- Kunihiko Fukushima and Sei Miyake. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In *Competition and cooperation in neural nets*, pages 267–285. Springer, 1982.
- Mélanie Gaillochet, Kerem Can Tezcan, and Ender Konukoglu. Joint reconstruction and bias field correction for undersampled mr imaging. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 44–52. Springer, 2020.

- Yosef Gandelsman, Assaf Shocher, and Michal Irani. "double-dip": Unsupervised image decomposition via coupled deep-image-priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11026–11035, 2019.
- Gene Gindi, Mindy Lee, Anand Rangarajan, and I George Zubal. Bayesian reconstruction of functional images using anatomical information as priors. *IEEE Transactions on Medical Imaging*, 12(4):670–680, 1993.
- Sivan Gleichman and Yonina C Eldar. Blind compressed sensing. *IEEE Transactions on Information Theory*, 57(10):6958–6975, 2011.
- Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014.
- Mark A Griswold, Peter M Jakob, Robin M Heidemann, Mathias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase. Generalized autocalibrating partially parallel acquisitions (grappa). *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 47(6):1202–1210, 2002.
- et al. Guanxiong Luo. Joint estimation of coil sensitivities and image content using a deep image prior. In *Proceedings of the 29th Annual Meeting of ISMRM*, 2021.
- Pengfei Guo, Jeya Maria Jose Valanarasu, Puyang Wang, Jinyuan Zhou, Shanshan Jiang, and Vishal M Patel. Over-and-under complete convolutional rnn for mri reconstruction. *arXiv preprint arXiv:2106.08886*, 2021.
- David Ha, Andrew M Dai, and Quoc V Le. Hypernetworks. *ICLR*, 2016.
- Justin P Haldar. Low-rank modeling of local  $k$ -space neighborhoods (loraks) for constrained mri. *IEEE transactions on medical imaging*, 33(3):668–681, 2013.
- Justin P Haldar. Autocalibrated loraks for fast constrained mri reconstruction. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pages 910–913. IEEE, 2015.
- Justin P Haldar and Tae Hyung Kim. Computational imaging with loraks: Reconstructing linearly predictable signals using low-rank matrix regularization. In *2017 51st Asilomar Conference on Signals, Systems, and Computers*, pages 1870–1874. IEEE, 2017.
- Justin P Haldar and Jingwei Zhuo. P-loraks: low-rank modeling of local  $k$ -space neighborhoods with parallel imaging data. *Magnetic resonance in medicine*, 75(4):1499–1514, 2016.
- Kerstin Hammernik, Florian Knoll, Daniel K Sodickson, and Thomas Pock. L2 or not l2: impact of loss function design for deep learning mri reconstruction. In *Proceedings of the 25th Annual Meeting of ISMRM, Honolulu, HI*, 2017.
- Kerstin Hammernik, Teresa Klatzer, Erich Kobler, Michael P Recht, Daniel K Sodickson, Thomas Pock, and Florian Knoll. Learning a variational network for reconstruction of accelerated mri data. *Magnetic resonance in medicine*, 79(6):3055–3071, 2018.
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

- Reinhard Heckel. Regularizing linear inverse problems with convolutional neural networks. *arXiv preprint arXiv:1907.03100*, 2019.
- Reinhard Heckel and Paul Hand. Deep decoder: Concise image representations from untrained non-convolutional networks. In *International Conference on Learning Representations*, 2018.
- Reinhard Heckel and Mahdi Soltanolkotabi. Compressive sensing with un-trained neural networks: Gradient descent finds a smooth approximation. In *International Conference on Machine Learning*, pages 4149–4158. PMLR, 2020.
- R Heidemann, Mark A Griswold, Axel Haase, and Peter M Jakob. Variable density auto-smash imaging. In *Proc of 8th Scientific Meeting ISMRM, Denver*, page 274, 2000.
- et al Heng Yu. eraki: Fast robust artificial neural networks for k-space interpolation (raki) with coil combination and joint reconstruction. In *Proceedings of the 29th Annual Meeting of ISMRM*, 2021.
- Tom Hilbert, Tilman J Sumpf, Elisabeth Weiland, Jens Frahm, Jean-Philippe Thiran, Reto Meuli, Tobias Kober, and Gunnar Krueger. Accelerated t2 mapping combining parallel mri and model-based reconstruction: Grappatini. *Journal of Magnetic Resonance Imaging*, 48(2):359–368, 2018.
- Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006.
- Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.
- et al Hongyi Gu. Compressed sensing mri revisited: Optimizing l1-wavelet reconstruction with modern data science tools. In *Proceedings of the 29th Annual Meeting of ISMRM*, 2021.
- Seyed Amir Hossein Hosseini, Steen Moeller, Sebastian Weingärtner, Kâmil Uğurbil, and Mehmet Akçakaya. Accelerated coronary mri using 3d spirit-raki with sparsity regularization. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 1692–1695. IEEE, 2019a.
- Seyed Amir Hossein Hosseini, Chi Zhang, Kâmil Uğurbil, Steen Moeller, and Mehmet Akçakaya. sraki-rnn: accelerated mri with scan-specific recurrent neural networks using densely connected blocks. In *Wavelets and Sparsity XVIII*, volume 11138, page 111381B. International Society for Optics and Photonics, 2019b.
- Seyed Amir Hossein Hosseini, Chi Zhang, Sebastian Weingärtner, Steen Moeller, Matthias Stuber, Kamil Ugurbil, and Mehmet Akçakaya. Accelerated coronary mri with sraki: A database-free self-consistent neural network k-space reconstruction for arbitrary undersampling. *Plos one*, 15(2):e0229418, 2020.
- Yuxin Hu, Evan G Levine, Qiyuan Tian, Catherine J Moran, Xiaole Wang, Valentina Taviani, Shreyas S Vasanawala, Jennifer A McNab, Bruce A Daniel, and Brian L Hargreaves. Motion-robust reconstruction of multishot diffusion-weighted images without phase estimation through locally low-rank regularization. *Magnetic resonance in medicine*, 81(2):1181–1190, 2019.
- Feng Huang, James Akao, Sathya Vijayakumar, George R Duensing, and Mark Limkeman. k-t grappa: A k-space implementation for dynamic mri with high reduction factor. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 54(5):1172–1184, 2005.



- Iris AM Huijben, Bastiaan S Veeling, and Ruud JG van Sloun. Deep probabilistic subsampling for task-adaptive compressed sensing. In *International Conference on Learning Representations*, 2019.
- Peter M Jakob, Mark A Grisowld, Robert R Edelman, and Daniel K Sodickson. Auto-smash: a self-calibrating technique for smash imaging. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 7(1):42–54, 1998.
- Kyong Hwan Jin, Dongwook Lee, and Jong Chul Ye. A general framework for compressed sensing and parallel mri using annihilating filter based low-rank hankel matrix. *IEEE Transactions on Computational Imaging*, 2(4):480–495, 2016.
- Kyong Hwan Jin, Michael Unser, and Kwang Moo Yi. Self-supervised deep active accelerated mri. *arXiv preprint arXiv:1901.04547*, 2019.
- Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision*, pages 694–711. Springer, 2016.
- Alexia Jolicoeur-Martineau. The relativistic discriminator: a key element missing from standard gan. In *International Conference on Learning Representations*, 2018.
- Tae Hyung Kim, Pratyush Garg, and Justin P Haldar. Loraki: Autocalibrated recurrent neural networks for autoregressive mri reconstruction in k-space. *arXiv preprint arXiv:1904.09390*, 2019.
- DP Kingma and M Welling. Auto-encoding variational bayes. iclr 2014 2014. *arXiv preprint arXiv:1312.6114*, 2013.
- Yilmaz Korkmaz, Salman UH Dar, Mahmut Yurt, Muzaffer Özbey, and Tolga Çukur. Un-supervised mri reconstruction via zero-shot learned adversarial transformers. *arXiv preprint arXiv:2105.08059*, 2021.
- Solomon Kullback and Richard A Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.
- Kinam Kwon, Dongchan Kim, and HyunWook Park. A parallel mr imaging method using multilayer perceptron. *Medical physics*, 44(12):6209–6224, 2017.
- I Laurette, PM Koulibaly, L Blanc-Feraud, P Charbonnier, JC Nosmas, M Barlaud, and J Darcourt. Cone-beam algebraic reconstruction using edge-preserving regularization. In *Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, pages 59–73. Springer, 1996.
- Pascal Theriault Lauzier, Jie Tang, and Guang-Hong Chen. Prior image constrained compressed sensing: Implementation and performance evaluation. *Medical physics*, 39(1):66–80, 2012.
- Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551, 1989.
- Dongwook Lee, Kyong Hwan Jin, Eung Yeop Kim, Sung-Hong Park, and Jong Chul Ye. Acceleration of mr parameter mapping using annihilating filter-based low rank hankel matrix (aloha). *Magnetic resonance in medicine*, 76(6):1848–1864, 2016.
- Dongwook Lee, Jaejun Yoo, Sungho Tak, and Jong Chul Ye. Deep residual learning for accelerated mri using magnitude and phase networks. *IEEE Transactions on Biomedical Engineering*, 65(9):1985–1995, 2018.

- Dongwook Lee, Junyoung Kim, Won-Jin Moon, and Jong Chul Ye. Collagan: Collaborative gan for missing image data imputation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2487–2496, 2019.
- Dong Liang, Jing Cheng, Ziwen Ke, and Leslie Ying. Deep magnetic resonance image reconstruction: Inverse problems meet neural networks. *IEEE Signal Processing Magazine*, 37(1):141–151, 2020a.
- Haoyun Liang, Yu Gong, Hoel Kervadec, Cheng Li, Jing Yuan, Xin Liu, Hairong Zheng, and Shanshan Wang. Laplacian pyramid-based complex neural network learning for fast mr imaging. In *Medical Imaging with Deep Learning*, pages 454–464. PMLR, 2020b.
- Zhi-Pei Liang. Spatiotemporal imaging with partially separable functions. In *2007 4th IEEE international symposium on biomedical imaging: from nano to macro*, pages 988–991. IEEE, 2007.
- Sajan Goud Lingala and Mathews Jacob. Blind compressive sensing dynamic mri. *IEEE transactions on medical imaging*, 32(6):1132–1145, 2013.
- Sajan Goud Lingala, Yue Hu, Edward DiBella, and Mathews Jacob. Accelerated dynamic mri exploiting sparsity and low-rank structure: kt slr. *IEEE transactions on medical imaging*, 30(5):1042–1054, 2011.
- Fang Liu, Alexey Samsonov, Lihua Chen, Richard Kijowski, and Li Feng. Santis: sampling-augmented neural network with incoherent structure for mr image reconstruction. *Magnetic resonance in medicine*, 82(5):1890–1904, 2019.
- Yunsong Liu, Jian-Feng Cai, Zhifang Zhan, Di Guo, Jing Ye, Zhong Chen, and Xiaobo Qu. Balanced sparse model for tight frames in compressed sensing magnetic resonance imaging. *PloS one*, 10(4):e0119584, 2015.
- Yunsong Liu, Zhifang Zhan, Jian-Feng Cai, Di Guo, Zhong Chen, and Xiaobo Qu. Projected iterative soft-thresholding algorithm for tight frames in compressed sensing magnetic resonance imaging. *IEEE transactions on medical imaging*, 35(9):2130–2140, 2016.
- K Lonning, P Putzky, M Caan, and M Welling. Recurrent inference machines for accelerated mri reconstruction. 2018.
- Guanxiong Luo, Na Zhao, Wenhao Jiang, Edward S Hui, and Peng Cao. Mri reconstruction using deep bayesian estimation. *Magnetic resonance in medicine*, 84(4):2246–2261, 2020.
- Michael Lustig and John M Pauly. Spirit: iterative self-consistent parallel imaging reconstruction from arbitrary k-space. *Magnetic resonance in medicine*, 64(2):457–471, 2010.
- Michael Lustig, Juan M Santos, David L Donoho, and John M Pauly. kt sparse: High frame rate dynamic mri exploiting spatio-temporal sparsity. In *Proceedings of the 13th annual meeting of ISMRM, Seattle*, volume 2420, 2006.
- Jun Lv, Chengyan Wang, and Guang Yang. Pic-gan: A parallel imaging coupled generative adversarial network for accelerated multi-channel mri reconstruction. *Diagnostics*, 11(1):61, 2021.
- et al. Maarten Terpstra. Rethinking complex image reconstruction: perpendicular loss for improved complex image reconstruction with deep learning. In *Proceedings of the 29th Annual Meeting of ISMRM*, 2021.

- Oliver Maier, Jasper Schoormans, Matthias Schloegl, Gustav J Strijkers, Andreas Lesch, Thomas Benkert, Tobias Block, Bram F Coolen, Kristian Bredies, and Rudolf Stollberger. Rapid t1 quantification from high resolution 3d data with model-based reconstruction. *Magnetic resonance in medicine*, 81(3):2072–2089, 2019.
- Morteza Mardani, Enhao Gong, Joseph Y Cheng, Shreyas S Vasanaawala, Greg Zaharchuk, Lei Xing, and John M Pauly. Deep generative adversarial neural networks for compressive sensing mri. *IEEE transactions on medical imaging*, 38(1):167–179, 2018a.
- Morteza Mardani, Qingyun Sun, David Donoho, Vardan Papyan, Hatf Monajemi, Shreyas Vasanaawala, and John Pauly. Neural proximal gradient descent for compressive imaging. *Advances in Neural Information Processing Systems*, 31:9573–9583, 2018b.
- Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.
- Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.
- Christopher A Metzler, Ali Mousavi, Reinhard Heckel, and Richard G Baraniuk. Unsupervised learning with stein’s unbiased risk estimator. *arXiv preprint arXiv:1805.10531*, 2018.
- Oleg Michailovich, Yogesh Rathi, and Sudipto Dolui. Spatially regularized compressed sensing for high angular resolution diffusion imaging. *IEEE transactions on medical imaging*, 30(5):1100–1115, 2011.
- Marvin Minsky and Seymour A Papert. *Perceptrons: An introduction to computational geometry*. MIT press, 2017.
- Reinhard Heckel Mohammad Zalbagi Darestani. Can un-trained networks compete with trained ones for accelerated mri? In *Proceedings of the 29th Annual Meeting of ISMRM*, 2021.
- Dominik Narnhofer, Kerstin Hammernik, Florian Knoll, and Thomas Pock. Inverse GANs for accelerated MRI reconstruction. In Dimitri Van De Ville, Manos Papadakis, and Yue M. Lu, editors, *Wavelets and Sparsity XVIII*, volume 11138, pages 381 – 392. International Society for Optics and Photonics, SPIE, 2019. doi: 10.1117/12.2527753. URL <https://doi.org/10.1117/12.2527753>.
- Dominik Narnhofer, Alexander Effland, Erich Kobler, Kerstin Hammernik, Florian Knoll, and Thomas Pock. Bayesian uncertainty estimation of learned variational mri reconstruction. *arXiv preprint arXiv:2102.06665*, 2021.
- Harry Nyquist. Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, 47(2):617–644, 1928.
- Changheun Oh, Dongchan Kim, Jun-Young Chung, Yeji Han, and HyunWook Park. A k-space-to-image reconstruction network for mri using recurrent neural network. *Medical Physics*, 48(1):193–203, 2021.
- Ilkay Oksuz, James Clough, Aurelien Bustin, Gastao Cruz, Claudia Prieto, Rene Botnar, Daniel Rueckert, Julia A Schnabel, and Andrew P King. Cardiac mr motion artefact correction from k-space using deep learning-based reconstruction. In *International Workshop on Machine Learning for Medical Image Reconstruction*, pages 21–29. Springer, 2018.

- Ilkay Oksuz, James Clough, Wenjia Bai, Bram Ruijsink, Esther Puyol-Antón, Gastao Cruz, Claudia Prieto, Andrew P King, and Julia A Schnabel. High-quality segmentation of low quality cardiac mr images using k-space artefact correction. In *International Conference on Medical Imaging with Deep Learning*, pages 380–389. PMLR, 2019a.
- Ilkay Oksuz, James Clough, Bram Ruijsink, Esther Puyol-Antón, Aurelien Bustin, Gastao Cruz, Claudia Prieto, Daniel Rueckert, Andrew P King, and Julia A Schnabel. Detection and correction of cardiac mri motion artefacts during reconstruction from k-space. In *International conference on medical image computing and computer-assisted intervention*, pages 695–703. Springer, 2019b.
- Ilkay Oksuz, James R Clough, Bram Ruijsink, Esther Puyol Anton, Aurelien Bustin, Gastao Cruz, Claudia Prieto, Andrew P King, and Julia A Schnabel. Deep learning-based detection and correction of cardiac mr motion artefacts during reconstruction for high-quality segmentation. *IEEE Transactions on Medical Imaging*, 39(12):4001–4010, 2020.
- Luca Oneto, Sandro Ridella, and Davide Anguita. Tikhonov, ivanov and morozov regularization for support vector machine learning. *Machine Learning*, 103(1):103–136, 2016.
- Greg Ongie and Mathews Jacob. Off-the-grid recovery of piecewise constant images from few fourier samples. *SIAM Journal on Imaging Sciences*, 9(3):1004–1041, 2016.
- Aaron van den Oord, Nal Kalchbrenner, Oriol Vinyals, Lasse Espeholt, Alex Graves, and Koray Kavukcuoglu. Conditional image generation with pixcnn decoders. *arXiv preprint arXiv:1606.05328*, 2016.
- Ricardo Otazo, Emmanuel Candes, and Daniel K Sodickson. Low-rank plus sparse matrix decomposition for accelerated dynamic mri with separation of background and dynamic components. *Magnetic resonance in medicine*, 73(3):1125–1136, 2015.
- Arghya Pal and Vineeth N Balasubramanian. Zero-shot task transfer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2189–2198, 2019.
- Jaeseok Park, Qiang Zhang, Vladimir Jellus, Orlando Simonetti, and Debiao Li. Artifact and noise suppression in grappa imaging using improved k-space coil calibration and variable density sampling. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 53(1):186–193, 2005.
- Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019. URL <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.
- Klaas P Pruessmann, Markus Weiger, Markus B Scheidegger, and Peter Boesiger. Sense: sensitivity encoding for fast mri. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 42(5):952–962, 1999.
- Chen Qin, Jo Schlemper, Jose Caballero, Anthony N Price, Joseph V Hajnal, and Daniel Rueckert. Convolutional recurrent neural networks for dynamic mr image reconstruction. *IEEE transactions on medical imaging*, 38(1):280–290, 2018.

- Tran Minh Quan, Thanh Nguyen-Duc, and Won-Ki Jeong. Compressed sensing mri reconstruction using a generative adversarial network with a cyclic loss. *IEEE transactions on medical imaging*, 37(6):1488–1497, 2018.
- Ankit Raj, Yoram Bresler, and Bo Li. Improving robustness of deep-learning-based image reconstruction. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 7932–7942. PMLR, 13–18 Jul 2020. URL <https://proceedings.mlr.press/v119/raj20a.html>.
- Sriprabha Ramanarayanan, Balamurali Murugesan, Keerthi Ram, and Mohanasankar Sivaprakasam. Mac-reconnet: A multiple acquisition context based convolutional neural network for mr image reconstruction using dynamic weight prediction. In *Medical Imaging with Deep Learning*, pages 696–708. PMLR, 2020.
- Julian Rasch, Ville Kolehmainen, Riikka Nivajärvi, Mikko Kettunen, Olli Gröhn, Martin Burger, and Eva-Maria Brinkmann. Dynamic mri reconstruction from undersampled data with an anatomical prescan. *Inverse problems*, 34(7):074001, 2018.
- Yogesh Rathi, O Michailovich, Kawin Setsompop, Sylvain Bouix, Martha Elizabeth Shenton, and C-F Westin. Sparse multi-shell diffusion imaging. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 58–65. Springer, 2011.
- Saiprasad Ravishankar and Yoram Bresler. Learning sparsifying transforms. *IEEE Transactions on Signal Processing*, 61(5):1072–1086, 2012.
- Saiprasad Ravishankar and Yoram Bresler. Efficient blind compressed sensing using sparsifying transforms with convergence guarantees and application to magnetic resonance imaging. *SIAM Journal on Imaging Sciences*, 8(4):2519–2557, 2015.
- Saiprasad Ravishankar and Yoram Bresler. Data-driven learning of a union of sparsifying transforms model for blind compressed sensing. *IEEE Transactions on Computational Imaging*, 2(3):294–309, 2016.
- Saiprasad Ravishankar, Jong Chul Ye, and Jeffrey A Fessler. Image reconstruction: From sparsity to data-adaptive methods and machine learning. *Proceedings of the IEEE*, 108(1):86–109, 2019.
- Volkert Roeloffs, Xiaoqing Wang, Tilman J Sumpf, Markus Untenberger, Dirk Voit, and Jens Frahm. Model-based reconstruction for t1 mapping using single-shot inversion-recovery radial flash. *International Journal of Imaging Systems and Technology*, 26(4):254–263, 2016.
- Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- Frank Rosenblatt. *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory, 1957.
- A Rowland, M Burns, T Hartkens, J Hajnal, D Rueckert, and Derek LG Hill. Information extraction from images (ixi): Image processing workflows using a grid enabled image database. *Proceedings of DiDaMIC*, 4:55–64, 2004.
- David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.

- Rayan Saab, Rick Chartrand, and Ozgur Yilmaz. Stable sparse approximations via nonconvex optimization. In *2008 IEEE international conference on acoustics, speech and signal processing*, pages 3885–3888. IEEE, 2008.
- Maddalena Sacco. Stochastic relaxation, gibbs distributions and bayesian restoration of images. *Seconda Universit*, 1990.
- Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. *Advances in neural information processing systems*, 29: 2234–2242, 2016.
- Jo Schlemper, Jose Caballero, Joseph V Hajnal, Anthony N Price, and Daniel Rueckert. A deep cascade of convolutional neural networks for dynamic mr image reconstruction. *IEEE transactions on Medical Imaging*, 37(2):491–503, 2017.
- Manuel Schneider, Thomas Benkert, Eddy Solomon, Dominik Nickel, Matthias Fenchel, Berthold Kiefer, Andreas Maier, Hersh Chandarana, and Kai Tobias Block. Free-breathing fat and r2\* quantification in the liver using a stack-of-stars multi-echo acquisition with respiratory-resolved model-based reconstruction. *Magnetic resonance in medicine*, 84(5):2592–2605, 2020.
- Gavin Seegoolam, Jo Schlemper, Chen Qin, Anthony Price, Jo Hajnal, and Daniel Rueckert. Exploiting motion for deep learning reconstruction of extremely-undersampled dynamic mri. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 704–712. Springer, 2019.
- Nicole Seiberlich, Felix Breuer, Martin Blaimer, Peter Jakob, and Mark Griswold. Self-calibrating grappa operator gridding for radial and spiral trajectories. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 59(4):930–935, 2008.
- Roy Shaul, Itamar David, Ohad Shitrit, and Tammy Riklin Raviv. Subsampled brain mri reconstruction by generative adversarial neural networks. *Medical Image Analysis*, 65:101747, 2020.
- Peter J Shin, Peder EZ Larson, Michael A Ohliger, Michael Elad, John M Pauly, Daniel B Vigneron, and Michael Lustig. Calibrationless parallel imaging reconstruction based on structured low-rank matrix completion. *Magnetic resonance in medicine*, 72(4):959–970, 2014.
- Ohad Shitrit and Tammy Riklin Raviv. Accelerated magnetic resonance imaging by adversarial neural network. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 30–38. Springer, 2017.
- Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- Vimal Singh, Ahmed H Tewfik, and David B Ress. Under-sampled functional mri using low-rank plus sparse matrix decomposition. In *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 897–901. IEEE, 2015.
- Daniel K Sodickson. Spatial encoding using multiple rf coils: Smash imaging and parallel mri. *Methods in biomedical magnetic resonance imaging and spectroscopy*, pages 239–250, 2000.
- Anuroop Sriram, Jure Zbontar, Tullie Murrell, Aaron Defazio, C Lawrence Zitnick, Nafissa Yakubova, Florian Knoll, and Patricia Johnson. End-to-end variational networks for accelerated mri reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 64–73. Springer, 2020a.

- Anuroop Sriram, Jure Zbontar, Tullie Murrell, C Lawrence Zitnick, Aaron Defazio, and Daniel K Sodickson. Grappanet: Combining parallel imaging with deep learning for multi-coil mri reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14315–14322, 2020b.
- Tilman J Sumpf, Martin Uecker, Susann Boretius, and Jens Frahm. Model-based nonlinear inverse reconstruction for t2 mapping using highly undersampled spin-echo mri. *Journal of Magnetic Resonance Imaging*, 34(2):420–428, 2011.
- Jian Sun, Huibin Li, Zongben Xu, et al. Deep admn-net for compressive sensing mri. *Advances in neural information processing systems*, 29, 2016.
- Richard Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.
- Jonathan I Tamir, Stella X Yu, and Michael Lustig. Unsupervised deep basis pursuit: Learning inverse problems without ground-truth data. *arXiv preprint arXiv:1910.13110*, 2019.
- Kerem C Tezcan, Christian F Baumgartner, Roger Luechinger, Klaas P Pruessmann, and Ender Konukoglu. Mr image reconstruction using deep density priors. *IEEE transactions on medical imaging*, 38(7):1633–1642, 2018.
- Kerem C. Tezcan, Christian F. Baumgartner, Roger Luechinger, Klaas P. Pruessmann, and Ender Konukoglu. {MR} image reconstruction using deep density priors. In *International Conference on Medical Imaging with Deep Learning – Extended Abstract Track*, London, United Kingdom, 08–10 Jul 2019. URL <https://openreview.net/forum?id=ryxKXECaK4>.
- Jean-Baptiste Thibault, Ken D Sauer, Charles A Bouman, and Jiang Hsieh. A three-dimensional statistical approach to improved image quality for multislice helical ct. *Medical physics*, 34(11):4526–4544, 2007.
- Johannes Tran-Gia, Daniel Stäb, Tobias Wech, Dietbert Hahn, and Herbert Köstler. Model-based acceleration of parameter mapping (map) for saturation prepared radially acquired data. *Magnetic resonance in medicine*, 70(6):1524–1534, 2013.
- Johannes Tran-Gia, Sotirios Bisdas, Herbert Köstler, and Uwe Klose. A model-based reconstruction technique for fast dynamic t1 mapping. *Magnetic resonance imaging*, 34(3):298–307, 2016.
- Jeffrey Tsao, Peter Boesiger, and Klaas P Pruessmann. k-t blast and k-t sense: dynamic mri with high frame rate exploiting spatiotemporal correlations. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 50(5):1031–1042, 2003.
- Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 9446–9454, 2018.
- David C Van Essen, Kamil Ugurbil, Edward Auerbach, Deanna Barch, Timothy EJ Behrens, Richard Bucholz, Acer Chang, Liyong Chen, Maurizio Corbetta, Sandra W Curtiss, et al. The human connectome project: a data acquisition perspective. *Neuroimage*, 62(4):2222–2231, 2012.
- V Vapnik. Principles of risk minimization for learning theory. *Advances in Neural Information Processing Systems*, 4, 1991.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008, 2017.

- Patrick Virtue and Michael Lustig. The empirical effect of gaussian noise in undersampled mri reconstruction. *Tomography*, 3(4):211–221, 2017.
- Alan Q Wang, Adrian V Dalca, and Mert R Sabuncu. Regularization-agnostic compressed sensing mri reconstruction with hypernetworks. *arXiv preprint arXiv:2101.02194*, 2021.
- Puyang Wang, Eric Z Chen, Terrence Chen, Vishal M Patel, and Shanhui Sun. Pyramid convolutional rnn for mri reconstruction. *arXiv preprint arXiv:1912.00543*, 2019.
- Shanshan Wang, Zhenghang Su, Leslie Ying, Xi Peng, Shun Zhu, Feng Liang, Dagan Feng, and Dong Liang. Accelerating magnetic resonance imaging via deep learning. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 514–517. IEEE, 2016.
- Shanshan Wang, Huitao Cheng, Leslie Ying, Taohui Xiao, Ziwen Ke, Hairong Zheng, and Dong Liang. Deepcomplexmri: Exploiting deep residual network for fast parallel mr imaging with complex convolution. *Magnetic Resonance Imaging*, 68:136–147, 2020a.
- Shanshan Wang, Huitao Cheng, Leslie Ying, Taohui Xiao, Ziwen Ke, Hairong Zheng, and Dong Liang. Deepcomplexmri: Exploiting deep residual network for fast parallel mr imaging with complex convolution. *Magnetic Resonance Imaging*, 68:136–147, 2020b.
- Tomer Weiss, Ortal Senouf, Sanketh Vedula, Oleg Michailovich, Michael Zibulevsky, and Alex Bronstein. Pilot: Physics-informed learned optimized trajectories for accelerated mri. *MELBA*, pages 1–23, 2021.
- Bihan Wen, Yanjun Li, and Yoram Bresler. The power of complementary regularizers: Image recovery via transform learning and low-rank modeling. *arXiv preprint arXiv:1808.01316*, 2018.
- Bo Xin, Yizhou Wang, Wen Gao, David Wipf, and Baoyuan Wang. Maximal sparsity with deep networks? *Advances in Neural Information Processing Systems*, 29:4340–4348, 2016.
- Lin Xu, Qian Zheng, and Tao Jiang. Improved parallel magnertic resonance imaging reconstruction with complex proximal support vector regression. *Scientific reports*, 8(1):1–9, 2018.
- Burhaneddin Yaman, Seyed Amir Hossein Hosseini, S. Moeller, J. Ellermann, K. Uğurbil, and M. Akçakaya. Multi-mask self-supervised learning for physics-guided neural networks in highly accelerated mri. *ArXiv*, abs/2008.06029, 2020.
- Burhaneddin Yaman, Seyed Amir Hossein Hosseini, and Mehmet Akçakaya. Scan-specific mri reconstruction using zero-shot physics-guided deep learning. *arXiv e-prints*, pages arXiv–2102, 2021a.
- Burhaneddin Yaman, Seyed Amir Hossein Hosseini, and Mehmet Akçakaya. Zero-shot self-supervised learning for mri reconstruction. *arXiv preprint arXiv:2102.07737*, 2021b.
- Guang Yang, Simiao Yu, Hao Dong, Greg Slabaugh, Pier Luigi Dragotti, Xujiang Ye, Fangde Liu, Simon Arridge, Jennifer Keegan, Yike Guo, et al. Dagan: Deep de-aliasing generative adversarial networks for fast compressed sensing mri reconstruction. *IEEE transactions on medical imaging*, 37(6):1310–1321, 2017.
- Yan Yang, Jian Sun, Huibin Li, and Zongben Xu. Admm-csnet: A deep learning approach for image compressive sensing. *IEEE transactions on pattern analysis and machine intelligence*, 42(3):521–538, 2018.
- Jaeyun Yoo, Kyong Hwan Jin, Harshit Gupta, Jerome Yerly, Matthias Stuber, and Michael Unser. Time-dependent deep image prior for dynamic mri. *IEEE Transactions on Medical Imaging*, 2021.



- Zhenmou Yuan, Mingfeng Jiang, Yaming Wang, Bo Wei, Yongming Li, Pin Wang, Wade Menpes-Smith, Zhangming Niu, and Guang Yang. Sara-gan: Self-attention and relative average discriminator based generative adversarial networks for fast compressed sensing mri reconstruction. *Frontiers in Neuroinformatics*, 14:58, 2020. ISSN 1662-5196. doi: 10.3389/fninf.2020.611666. URL <https://www.frontiersin.org/article/10.3389/fninf.2020.611666>.
- Jure Zbontar, Florian Knoll, Anuroop Sriram, Tullie Murrell, Zhengnan Huang, Matthew J Muckley, Aaron Defazio, Ruben Stern, Patricia Johnson, Mary Bruno, et al. fastmri: An open dataset and benchmarks for accelerated mri. *arXiv preprint arXiv:1811.08839*, 2018.
- Chi Zhang, Seyed Amir Hossein Hosseini, Steen Moeller, Sebastian Weingärtner, Kamil Ugurbil, and Mehmet Akcakaya. Scan-specific residual convolutional neural networks for fast mri using residual raki. In *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, pages 1476–1480. IEEE, 2019a.
- Chi Zhang, Jinghan Jia, Burhaneddin Yaman, Steen Moeller, Sijia Liu, Mingyi Hong, and Mehmet Akcakaya. On instabilities of conventional multi-coil mri reconstruction to small adversarial perturbations. *arXiv preprint arXiv:2102.13066*, 2021.
- Jian Zhang and Bernard Ghanem. Ista-net: Interpretable optimization-inspired deep network for image compressive sensing. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1828–1837, 2018.
- Jian Zhang, Chunlei Liu, and Michael E Moseley. Parallel reconstruction using null operations. *Magnetic resonance in medicine*, 66(5):1241–1253, 2011.
- Jinwei Zhang, Zhe Liu, Shun Zhang, Hang Zhang, Pascal Spincemaille, Thanh D Nguyen, Mert R Sabuncu, and Yi Wang. Fidelity imposed network edit (fine) for solving ill-posed image reconstruction. *Neuroimage*, 211:116579, 2020.
- Pengyue Zhang, Fusheng Wang, Wei Xu, and Yu Li. Multi-channel generative adversarial network for parallel magnetic resonance image reconstruction in k-space. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 180–188. Springer, 2018a.
- Pengyue Zhang, Fusheng Wang, Wei Xu, and Yu Li. Multi-channel generative adversarial network for parallel magnetic resonance image reconstruction in k-space. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 180–188. Springer, 2018b.
- Zizhao Zhang, Adriana Romero, Matthew J Muckley, Pascal Vincent, Lin Yang, and Michal Drozdal. Reducing uncertainty in undersampled mri reconstruction with active acquisition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2049–2058, 2019b.
- Hang Zhao, Orazio Gallo, Iuri Frosio, and Jan Kautz. Loss functions for image restoration with neural networks. *IEEE Transactions on computational imaging*, 3(1):47–57, 2016.
- Tiejun Zhao and Xiaoping Hu. Iterative grappa (igrappa) for improved parallel imaging reconstruction. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 59(4):903–907, 2008.
- Hao Zheng, Faming Fang, and Guixu Zhang. Cascaded dilated dense network with two-step data consistency for mri reconstruction. *Advances in Neural Information Processing Systems*, 32:1744–1754, 2019.

Bo Zhu, Jeremiah Z Liu, Stephen F Cauley, Bruce R Rosen, and Matthew S Rosen. Image reconstruction by domain-transform manifold learning. *Nature*, 555(7697):487–492, 2018.

Markus Zimmermann, Zaheer Abbas, Krzysztof Dzieciol, and N Jon Shah. Accelerated parameter mapping of multiple-echo gradient-echo data using model-based iterative reconstruction. *IEEE transactions on medical imaging*, 37(2):626–637, 2017.