

# k-Space Deep Learning for Accelerated MRI

Yoseo Han, Leonard Sunwoo<sup>✉</sup>, and Jong Chul Ye<sup>✉</sup>, Senior Member, IEEE

**Abstract**—The annihilating filter-based low-rank Hankel matrix approach (ALOHA) is one of the state-of-the-art compressed sensing approaches that directly interpolates the missing  $k$ -space data using low-rank Hankel matrix completion. The success of ALOHA is due to the concise signal representation in the  $k$ -space domain, thanks to the duality between structured low-rankness in the  $k$ -space domain and the image domain sparsity. Inspired by the recent mathematical discovery that links convolutional neural networks to Hankel matrix decomposition using data-driven framelet basis, here we propose a fully data-driven deep learning algorithm for  $k$ -space interpolation. Our network can be also easily applied to non-Cartesian  $k$ -space trajectories by simply adding an additional regridding layer. Extensive numerical experiments show that the proposed deep learning method consistently outperforms the existing image-domain deep learning approaches.

**Index Terms**—Compressed sensing MRI, deep learning, Hankel structured low-rank completion, convolution framelets.

## I. INTRODUCTION

RECENTLY, inspired by the tremendous success of deep learning [1]–[3], many researchers have investigated deep learning approaches for MR reconstruction problems and successfully demonstrated significant performance gain [4]–[10].

In particular, Wang *et al.* [4] used the deep learning reconstruction either as an initialization or a regularization term. Deep network architecture using unfolded iterative compressed sensing (CS) algorithm was also proposed to learn a set of regularizers and associated filters [5]. These works were followed by novel extension using deep residual learning [6], domain adaptation [7], data consistency layers [9], etc. An extreme form of the neural network called Automated Transform by

Manuscript received May 16, 2019; accepted June 29, 2019. Date of publication July 5, 2019; date of current version January 31, 2020. This work was supported by the Korea Science and Engineering Foundation under Grant NRF2016R1A2B3008104. (*Corresponding author:* Jong Chul Ye.)

Y. Han is with the Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea (e-mail: hanyoseob@kaist.ac.kr).

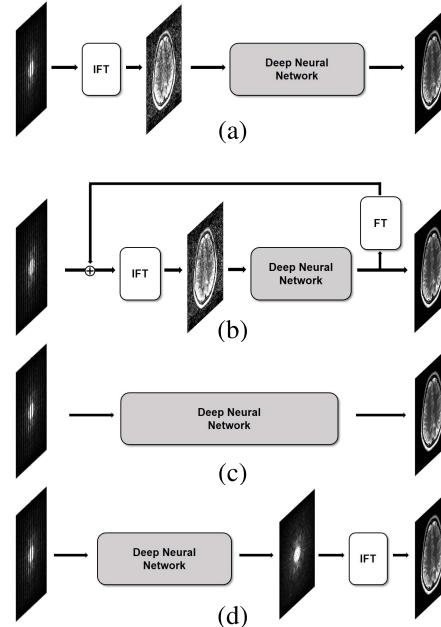
L. Sunwoo is with the Department of Radiology, College of Medicine, Seoul National University, Seoul 110744, South Korea, and also with Seoul National University Bundang Hospital, Seongnam 13620, South Korea.

J. C. Ye is with the Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea, and also with the Department of Mathematical Sciences, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea (e-mail: jong.ye@kaist.ac.kr).

This article has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the author.

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2019.2927101



**Fig. 1.** Deep learning frameworks for accelerated MRI: (a) image domain learning [6], [7], (b) cascaded network [4], [5], [9], (c) AUTOMAP [10], and (d) the proposed  $k$ -space learning. IFT: Inverse Fourier transform.

Manifold Approximation (AUTOMAP) [10] even attempts to estimate the Fourier transform itself using fully connected layers. All these pioneering works have consistently demonstrated superior reconstruction performances over the compressed sensing approaches [11]–[16] at significantly lower run-time computational complexity.

Although the end-to-end recovery approach like AUTOMAP [10] may directly recover the image without ever interpolating the missing  $k$ -space samples (see Fig. 1(c)), it works only for the sufficiently small size images due to its huge memory requirement for fully connected layers. Accordingly, most of the popular deep learning MR reconstruction algorithms are either in the form of image domain post-processing as shown in Fig. 1(a) [6], [7], or iterative updates between the  $k$ -space and the image domain using a cascaded network as shown in Fig. 1(b) [4], [5], [9], [17].

One of the main purposes of this paper is to reveal that the aforementioned approaches are not all the available options for MR deep learning, but there exists another effective deep learning approach. In fact, as illustrated in Fig. 1(d), the proposed deep learning approach directly interpolates the missing  $k$ -space data so that accurate reconstruction can be obtained by simply taking the Fourier transform of the interpolated  $k$ -space data. In contrast to AUTOMAP [10], our

network is implemented in the form of convolutional neural network (CNN) without requiring fully connected layer, so the GPU memory requirement for the proposed  $k$ -space deep learning is minimal.

In fact, our  $k$ -space deep learning scheme is inspired by the success of structured low-rank Hankel matrix approaches [13]–[16], [18], exploiting concise signal representation in the  $k$ -space domain thanks to the duality between the structured low-rankness in the  $k$ -space and the image domain sparsity. In addition, the recent theory of deep convolutional framelets [19] showed that an encoder-decoder network can be regarded as a signal representation that emerges from Hankel matrix decomposition. Therefore, by synergistically combining these findings, we propose a novel  $k$ -space deep learning algorithms that perform better and generalize well. We further show that our deep learning approach for  $k$ -space interpolation can handle general  $k$ -space sampling patterns beyond the Cartesian trajectory, such as radial, spiral, etc. Moreover, our theory and empirical results also show that multi-channel calibration-free  $k$ -space interpolation can be easily realized using the proposed framework by simply stacking multi-coil  $k$ -space data along the channel direction as an input to feed in the neural network.

We are aware of recent  $k$ -space neural network approaches called the scan-specific robust artificial neural networks for  $k$ -space interpolation (RAKI) [20]. Unlike the proposed method, RAKI considers scan-specific learning without training data, so the neural network weights needs to be recalculated for each  $k$ -space input data. On the other hand, the proposed method exploits the generalization capability of the neural network. More specifically, even with the same trained weights, the trained neural network can generate diverse signal representation depending on the input thanks to the combinatorial nature of the ReLU nonlinearity. This makes the neural network expressive and generalized well to the unseen test data. The theoretical origin of the expressiveness will be also discussed in this paper. Our source code can be found at <https://github.com/hanyoseob/k-space-deep-learning>.

After the original version of this work was available on Arxiv, there appear several deep learning algorithms exploiting  $k$ -space learning [21], [22]. These works are based on hybrid formulation (see Fig. 1(b)) and utilize the deep neural network as a regularization term for  $k$ -space denoising. On the other hand, the our method exploits the nature of deep neural network as a generalizable and expressive  $k$ -space representation to directly interpolate the missing  $k$ -space data.

## II. MATHEMATICAL PRELIMINARIES

### A. Notations

In this paper, matrices are denoted by bold upper case letters, i.e.  $\mathbf{A}, \mathbf{B}$ , whereas the vectors are represented by bold lower cases letters, i.e.  $\mathbf{x}, \mathbf{y}$ . In addition,  $[\mathbf{A}]_{ij}$  refers to the  $(i, j)$ -th element of the matrix  $\mathbf{A}$ , and  $x[i]$  denotes the  $i$ -th element of the vector  $\mathbf{x}$ . The notation  $\bar{\mathbf{v}} \in \mathbb{R}^d$  for a vector  $\mathbf{v} \in \mathbb{R}^d$  denotes its flipped version, i.e. the indices of  $\mathbf{v}$  are time-reversed such that  $\bar{\mathbf{v}}[n] = \mathbf{v}[-n]$ . The  $N \times N$  identity matrix is denoted as  $\mathbf{I}_N$ , while  $\mathbf{1}_N$  is an  $N$ -dimensional vector with 1's. The superscript  $T$  and  $\top$  for a matrix or vector denote

the transpose and Hermitian transpose, respectively.  $\mathbb{R}$  and  $\mathbb{C}$  denote the real and imaginary fields, respectively.  $\mathbb{R}_+$  refers to the nonnegative real numbers.

### B. Forward Model for Accelerated MRI

The spatial Fourier transform of an arbitrary smooth function  $x : \mathbb{R}^2 \rightarrow \mathbb{R}$  is defined by

$$\hat{x}(\mathbf{k}) = \mathcal{F}[x](\mathbf{k}) := \int_{\mathbb{R}^d} e^{-i\mathbf{k} \cdot \mathbf{r}} x(\mathbf{r}) d\mathbf{r},$$

with spatial frequency  $\mathbf{k} \in \mathbb{R}^2$  and  $i = \sqrt{-1}$ . Let  $\{\mathbf{k}_n\}_{n=1}^N$ , for some integer  $N \in \mathbb{N}$ , be a collection of finite number of sampling points of the  $k$ -space confirming to the Nyquist sampling rate. Accordingly, the discretized  $k$ -space data  $\hat{\mathbf{x}} \in \mathbb{C}^N$  is introduced by

$$\hat{\mathbf{x}} = [\hat{x}[0] \quad \cdots \quad \hat{x}[N-1]], \quad \text{where } \hat{x}[i] = \hat{x}(\mathbf{k}_i). \quad (1)$$

For a given under-sampling pattern  $\Lambda$  for accelerated MR acquisition, let the downsampling operator  $\mathcal{P}_\Lambda : \mathbb{C}^N \rightarrow \mathbb{C}^N$  be defined as

$$[\mathcal{P}_\Lambda[\hat{\mathbf{x}}]]_i = \begin{cases} \hat{x}[i], & i \in \Lambda \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

Then, the under-sampled  $k$ -space data is given by

$$\hat{\mathbf{y}} := \mathcal{P}_\Lambda[\hat{\mathbf{x}}] \quad (3)$$

### C. Low-Rank Hankel Matrix Approaches

From the undersampled data in (3), CS-MRI [11], [12] attempts to find the feasible solution that has minimum non-zero support in some sparsifying transform domain. This can be achieved by finding a function  $z : \mathbb{R}^2 \rightarrow \mathbb{R}$  such that

$$\begin{aligned} & \min_z \|\mathcal{T}z\|_1 \\ & \text{subject to } \mathcal{P}_\Lambda[\hat{\mathbf{x}}] = \mathcal{P}_\Lambda[\hat{\mathbf{z}}] \end{aligned} \quad (4)$$

where  $\mathcal{T}$  denotes the image domain sparsifying transform and

$$\hat{\mathbf{z}} = [\hat{z}(\mathbf{k}_0) \quad \cdots \quad \hat{z}(\mathbf{k}_{N-1})]^T. \quad (5)$$

This optimization problem usually requires iterative update between the  $k$ -space and the image domain after the discretization of  $z(\mathbf{r})$  [11], [12].

On the other hand, in recent structured low-rank matrix completion algorithms [13]–[16], [18], the compressed sensing problems was solved either by imposing the low-rank structured matrix penalty [14], [16] or by converting it a direct  $k$ -space interpolation problem using low-rank structure matrix completion [13], [15]. More specifically, let  $\mathbb{H}_d(\hat{\mathbf{x}})$  denote a Hankel matrix constructed from the  $k$ -space measurement  $\hat{\mathbf{x}}$  in (1), where  $d$  denotes the matrix pencil size (for more details on the construction of Hankel matrices and their relation to the convolution, see Section I in Supplementary Material). Then, if the underlying signal  $x(\mathbf{r})$  in the image domain is sparse and described as the signal with the finite rate of innovations (FRI) with rate  $s$  [23], the associated Hankel matrix  $\mathbb{H}_d(\hat{\mathbf{x}})$  with  $d > s$  is low-ranked [15], [16], [18]. Therefore, if some of  $k$ -space data are missing, we can construct an appropriate weighted Hankel matrix with missing elements such that the

missing elements are recovered using low-rank Hankel matrix completion approaches [24]:

$$(P) \quad \min_{\widehat{\mathbf{z}} \in \mathbb{C}^N} \text{RANK } \mathbb{H}_d(\widehat{\mathbf{z}}) \quad (6)$$

subject to  $\mathcal{P}_\Lambda[\widehat{\mathbf{x}}] = \mathcal{P}_\Lambda[\widehat{\mathbf{z}}]$ .

While the low-rank Hankel matrix completion problem (*P*) can be solved in various ways, one of the main technical huddles is its relatively large computational complexity for matrix factorization and large memory requirement for storing Hankel matrix. Although several new approaches have been proposed to solve these problems [25], the following section shows that a deep learning approach is a novel and efficient way to solve this problem.

### III. MAIN CONTRIBUTION

#### A. ALOHA as a Signal Representation

Consider the following image regression problem under the low-rank Hankel matrix constraint:

$$\min_{\widehat{\mathbf{z}} \in \mathbb{C}^N} \|x - \mathcal{F}^{-1}[\widehat{\mathbf{z}}]\|^2 \quad (7)$$

subject to  $\text{RANK } \mathbb{H}_d(\widehat{\mathbf{z}}) = s, \quad \mathcal{P}_\Lambda[\widehat{\mathbf{x}}] = \mathcal{P}_\Lambda[\widehat{\mathbf{z}}], \quad (8)$

where  $s$  denotes the estimated rank. In the above formulation, the cost in (7) is defined in the image domain to minimize the errors in the image domain, whereas the low-rank Hankel matrix constraint in (8) is imposed in the  $k$ -space.

Now, we convert the complex-valued constraint in (8) to a real-valued constraint. The procedure is as follows. First, the operator  $\Re : \mathbb{C}^N \rightarrow \mathbb{R}^{N \times 2}$  is defined as

$$\Re[\widehat{\mathbf{z}}] := [\text{Re}(\widehat{\mathbf{z}}) \quad \text{Im}(\widehat{\mathbf{z}})], \quad \forall \widehat{\mathbf{z}} \in \mathbb{C}^N \quad (9)$$

where  $\text{Re}(\cdot)$  and  $\text{Im}(\cdot)$  denote the real and imaginary part of the argument. Similarly, we define its inverse operator  $\Re^{-1} : \mathbb{R}^{N \times 2} \rightarrow \mathbb{C}^N$  as

$$\Re^{-1}[\widehat{\mathbf{Z}}] := \widehat{\mathbf{z}}_1 + i\widehat{\mathbf{z}}_2, \quad \forall \widehat{\mathbf{Z}} := [\mathbf{z}_1 \ \mathbf{z}_2] \in \mathbb{R}^{N \times 2} \quad (10)$$

Then, as shown in Section II in Supplementary Material, we can approximately convert (8) to an optimization problem with real-valued constraint:

$$(P_A) \quad \min_{\widehat{\mathbf{z}} \in \mathbb{C}^N} \|x - \mathcal{F}^{-1}[\widehat{\mathbf{z}}]\|^2$$

subject to  $\text{RANK } \mathbb{H}_{d|2}(\Re[\widehat{\mathbf{z}}]) = Q \leq 2s,$   
 $\mathcal{P}_\Lambda[\widehat{\mathbf{x}}] = \mathcal{P}_\Lambda[\widehat{\mathbf{z}}]. \quad (11)$

In the recent theory of deep convolutional framelets [19], this low-rank constraint optimization problem was addressed using learning-based signal representation. More specifically, for any  $\widehat{\mathbf{z}} \in \mathbb{C}^N$ , let the Hankel structured matrix  $\mathbb{H}_{d|2}(\Re[\widehat{\mathbf{z}}])$  have the singular value decomposition  $\mathbf{U}\Sigma\mathbf{V}^\top$ , where  $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_Q] \in \mathbb{R}^{N \times Q}$  and  $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_Q] \in \mathbb{R}^{2d \times Q}$  denote the left and the right singular vector bases matrices, respectively;  $\Sigma = (\sigma_{ij}) \in \mathbb{R}^{Q \times Q}$  is the diagonal matrix with singular values. Now, consider matrix pair  $\Psi, \tilde{\Psi} \in \mathbb{R}^{2d \times Q}$

$$\Psi := \begin{pmatrix} \psi_1^1 & \cdots & \psi_Q^1 \\ \psi_1^2 & \cdots & \psi_Q^2 \end{pmatrix} \quad \text{and} \quad \tilde{\Psi} := \begin{pmatrix} \tilde{\psi}_1^1 & \cdots & \tilde{\psi}_Q^1 \\ \tilde{\psi}_1^2 & \cdots & \tilde{\psi}_Q^2 \end{pmatrix} \quad (12)$$

that satisfy the low-rank projection constraint:

$$\Psi \tilde{\Psi}^\top = \mathbf{P}_R(\mathbf{V}), \quad (13)$$

where  $\mathbf{P}_R(\mathbf{V})$  denotes the projection matrix to the range space of  $\mathbf{V}$ . We further introduce the generalized pooling and unpooling matrices  $\Phi, \tilde{\Phi} \in \mathbb{R}^{N \times M}$  [19] that satisfies the condition

$$\tilde{\Phi} \Phi^\top = \mathbf{I}_N, \quad (14)$$

Using Eqs. (13) and (14), we can obtain the following matrix equality:

$$\mathbb{H}_{d|2}(\Re[\widehat{\mathbf{z}}]) = \tilde{\Phi} \Phi^\top \mathbb{H}_{d|2}(\Re[\widehat{\mathbf{z}}]) \Psi \tilde{\Psi}^\top = \tilde{\Phi} \mathbf{C} \tilde{\Psi}^\top, \quad (15)$$

where

$$\mathbf{C} := \Phi^\top \mathbb{H}_{d|2}(\Re[\widehat{\mathbf{z}}]) \Psi \in \mathbb{R}^{N \times Q} \quad (16)$$

By taking the generalized inverse of Hankel matrix, (15) can be converted to the framelet basis representation [19]. Moreover, one of the most important observations in [19] is that the resulting framelet basis representation can be equivalently represented by single layer encoder-decoder convolution architecture:

$$\Re[\widehat{\mathbf{z}}] = (\tilde{\Phi} \mathbf{C}) \circledast g(\tilde{\Psi}), \quad \text{where } \mathbf{C} = \Phi^\top (\Re[\widehat{\mathbf{z}}] \circledast h(\Psi)) \quad (17)$$

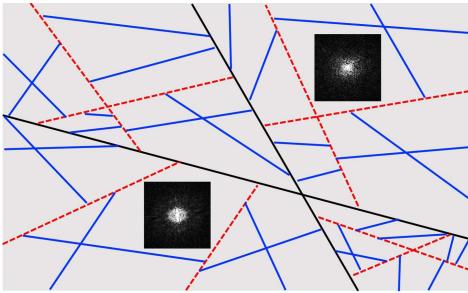
and  $\circledast$  denotes the multi-channel input multi-channel output convolution. The second and the first part of (17) correspond to the encoder and decoder layers with the corresponding convolution filters  $h(\Psi) \in \mathbb{R}^{2d \times Q}$  and  $g(\tilde{\Psi}) \in \mathbb{R}^{dQ \times 2}$ :

$$h(\Psi) := \begin{pmatrix} \overline{\psi}_1^1 & \cdots & \overline{\psi}_Q^1 \\ \overline{\psi}_1^2 & \cdots & \overline{\psi}_Q^2 \end{pmatrix}, \quad g(\tilde{\Psi}) := \begin{pmatrix} \tilde{\psi}_1^1 & \tilde{\psi}_1^2 \\ \vdots & \vdots \\ \tilde{\psi}_Q^1 & \tilde{\psi}_Q^2 \end{pmatrix},$$

which are obtained by reordering the matrices  $\Psi$  and  $\tilde{\Psi}$  in (12). Specifically,  $\overline{\psi}_i^1 \in \mathbb{R}^d$  (resp.  $\overline{\psi}_i^2 \in \mathbb{R}^d$ ) denotes the  $d$ -tap encoder convolutional filter applied to the real (resp. imaginary) component of the  $k$ -space data to generate the  $i$ -th channel output. In addition,  $g(\tilde{\Psi})$  is a reordered version of  $\tilde{\Psi}$  so that  $\tilde{\psi}_i^1 \in \mathbb{R}^d$  (resp.  $\tilde{\psi}_i^2 \in \mathbb{R}^d$ ) corresponds to the  $d$ -tap decoder convolutional filter to generate the real (resp. imaginary) component of the  $k$ -space data by convolving with the  $i$ -th channel input. We can further use recursive application of encoder-decoder representation for the resulting framelet coefficients  $\mathbf{C}$  in (17). In Corollary 4 of our companion paper [26], we showed that the recursive application of the encoder-decoder operations across the layers increases the net length of the convolutional filters.

Since (17) is a general form of the signals that are associated with a Hankel structured matrix, we are interested in using it to estimate bases for  $k$ -space interpolation. Specifically, we consider a complex-valued signal space  $\mathcal{H}$  determined by the filters  $\Psi$  and  $\tilde{\Psi}$ :

$$\mathcal{H}(\Psi, \tilde{\Psi}) = \left\{ \mathbf{z} \in \mathbb{C}^N \mid \Re[\mathbf{z}] = \Phi^\top (\mathbf{C} \circledast g(\tilde{\Psi})), \right. \\ \left. \mathbf{C} = (\tilde{\Phi} \Re[\mathbf{z}] \circledast h(\Psi)) \right\}. \quad (18)$$



**Fig. 2.** An example of input space partitioning for the case of two-channel three-layer ReLU neural network. Depending on input  $k$ -space data, the corresponding partition and its associated linear representation are selected.

Then, the ALOHA formulation  $P_A$  can be equivalently represented by

$$(P'_A) \quad \min_{\hat{\mathbf{z}} \in \mathcal{H}(\Psi, \tilde{\Psi})} \min_{\Psi, \tilde{\Psi}} \|x - \mathcal{F}^{-1}[\hat{\mathbf{z}}]\|^2 \\ \text{subject to } \mathcal{P}_\Lambda[\hat{\mathbf{x}}] = \mathcal{P}_\Lambda[\hat{\mathbf{z}}],$$

In other words, ALOHA is to find the optimal filter  $\Psi, \tilde{\Psi}$  and the associated  $k$ -space signal  $\hat{\mathbf{z}} \in \mathcal{H}(\Psi, \tilde{\Psi})$  that satisfies the data consistency conditions. In contrast to the standard CS approaches in which signal presentation in the image domain is separately applied from the data-consistency constraint in the  $k$ -space, the success of ALOHA over CS can be contributed to more efficient signal representation in the  $k$ -space domain that simultaneously take care of the data consistency in the  $k$ -space domain.

### B. Generalization and Depth

To allow training for neural networks, the problem formulation in  $(P'_A)$  should be decoupled into two steps: the learning phase to estimate  $\Psi, \tilde{\Psi}$  from the training data, and the inference phase to estimate the interpolate signal  $\hat{\mathbf{z}}$  for the given filter set  $\Psi, \tilde{\Psi}$ . Although we have revealed the relations between ALOHA and encoder-decoder architecture, the derivation is for specific input signal and it is not clear how the relations would translate when training is performed over multiple training data set, and the trained network can be generalized to the unseen test data. Given that the sparsity prior in dictionary learning enables the selection of appropriate basis functions from the dictionary for each given input, one may conjecture that there should be similar mechanisms in deep neural networks that enable adaptation to the specific input signals.

In Section IV of Supplementary Material, we show that the ReLU nonlinearities indeed plays a critical role in the adaptation and generalization. In fact, in our companion paper [26], we have shown that ReLU offers combinatorial convolution frame basis selection depending on each input image. More specifically, thanks to ReLU, a trained filter set produce large number of partitions in the input space as shown in Fig. 2, in which each region shares the same linear signal representation. Therefore, depending on each  $k$ -space input data, a particular region and its associated linear representation are selected. Moreover, we show that the number of input space partition and the associated linear

representation increases exponentially with the depth, channel and the skipped connection. By synergistically exploiting the efficient signal representation in the  $k$ -space domain, this enormous expressivity from the same filter sets can make the  $k$ -space deep neural network more powerful than the conventional image domain learning.

For the more details on the theoretical aspect of deep neural networks, see Section IV of Supplementary Material or our companion paper [26].

### C. Extension to Parallel Imaging

In [15], we formally showed that when  $\{\hat{\mathbf{x}}_i\}_{i=1}^P$  denote the  $k$ -space measurements from  $P$  receiver coils, the following extended Hankel structured matrix is low-ranked:

$$\mathbb{H}_{d|P}(\hat{\mathbf{X}}) = [\mathbb{H}_d(\hat{\mathbf{x}}_1) \quad \cdots \quad \mathbb{H}_d(\hat{\mathbf{x}}_P)] \quad (19)$$

where

$$\hat{\mathbf{X}} = [\hat{\mathbf{x}}_1 \quad \cdots \quad \hat{\mathbf{x}}_P] \in \mathbb{C}^{N \times P}.$$

Thus, similar to the single channel cases, the date-driven decomposition of the extended Hankel matrix in (19) can be represented by stacking the each  $k$ -space data along the channel direction and applies the deep neural network for the given multi-channel data. Therefore, except the number of input and output channels, the network structure for parallel imaging data is identical to the single channel  $k$ -space learning.

### D. Sparsification

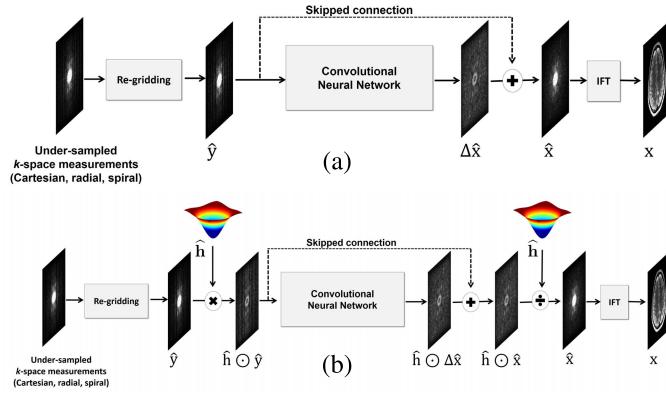
To further improve the performance of the structured matrix completion approach, in [18], we showed that even if the image  $x(\mathbf{r})$  may not be sparse, it can be often converted to a sparse signal.

For example, the outmost skipped connection for the residual learning is another way to make the signal sparse. Note that fully sampled  $k$ -space data  $\hat{\mathbf{x}}$  can be represented by

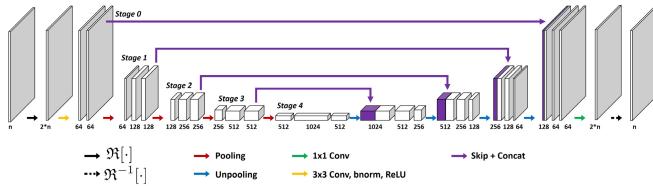
$$\hat{\mathbf{x}} = \hat{\mathbf{y}} + \Delta\hat{\mathbf{x}},$$

where  $\hat{\mathbf{y}}$  is the undersampled  $k$ -space measurement in (3), and  $\Delta\hat{\mathbf{x}}$  is the residual part of  $k$ -space data that should be estimated. In practice, some of the low-frequency part of  $k$ -space data including the DC component are acquired in the undersampled measurement so that the image component from the residual  $k$ -space data  $\Delta\hat{\mathbf{x}}$  are mostly high frequency signals, which are sparse. Therefore,  $\Delta\hat{\mathbf{x}}$  has low-rank Hankel matrix structure, which can be effectively processed using the deep neural network. This can be easily implemented using a skipped connection before the deep neural network as shown in Fig. 3(a). However, the skipped connection also works beyond the sparsification. In our companion paper [26] (which is also repeated in Section IV of Supplementary Material), we showed that the skipped connection at the inner layers makes the frame basis more expressive. Therefore, we conjecture that the skipped connections play dual roles in our  $k$ -space learning.

Second, we can convert a signal to an innovation signal using a shift-invariant transform represented by the whitening filter  $h$  such that the resulting innovation signal  $z = h * x$  becomes an FRI signal [23]. For example, many MR images



**Fig. 3.** Overall reconstruction flows of the proposed method with (a) skipped connection, and (b) skipped connection and weighting layer. IFT denotes the inverse Fourier transform. For the parallel imaging, the input and output are multi-coil *k*-space data, after stacking them along the channel direction.



**Fig. 4.** A network backbone of the proposed method. The input and output are complex-valued.

can be sparsified using finite difference or wavelet transform [15]. This implies that the Hankel matrix from the weighted *k*-space data,  $\hat{z}(\mathbf{k}) = \hat{h}(\mathbf{k})\hat{x}(\mathbf{k})$  are low-ranked, where the weight  $\hat{h}(\mathbf{k})$  is determined from the finite difference or Haar wavelet transform [15], [18]. Thus, the deep neural network is applied to the weighted *k*-space data to estimate the missing spectral data  $\hat{h}(\mathbf{x})\hat{x}(\mathbf{k})$ , after which the original *k*-space data is obtained by dividing with the same weight, i.e.  $\hat{x}(\mathbf{k}) = \hat{z}(\mathbf{k})/\hat{h}(\mathbf{k})$ . This can be easily implemented using a weighting and unweighting layer as shown in Fig. 3(b).

In this paper, we consider these two strategies to investigate which strategy is better for different sampling trajectories.

### E. Extension to General Sampling Patterns

Since the Hankel matrix formulation in ALOHA is based on the Cartesian coordinate, we add extra regridding layers to handle the non-Cartesian sampling trajectories. Specifically, for radial and spiral trajectories, the non-uniform fast Fourier transform (NUFFT) was used to perform the regridding to the Cartesian coordinate. For Cartesian sampling trajectory, the regridding layer using NUFFT is not necessary, and we instead perform a zero-filling in the unacquired *k*-space regions as an initialization step.

## IV. IMPLEMENTATION

### A. Network Backbone

The network backbone follows the U-Net [3] which consists of convolution, batch normalization, rectified linear unit (ReLU), and contracting path connection with concatenation as shown in Fig. 4. Here, the input and output are the complex-valued *k*-space data, while  $\Re[\cdot]$  and  $\Re^{-1}[\cdot]$

denote the operators in (9) and (10), respectively, that convert complex valued input to two-channel real value signals and vice versa. For parallel imaging, multi-coil *k*-space data are given as input and output after stacking them along channel direction. Specially, in our parallel imaging experiments, we use eight coils *k*-space data.

The yellow arrow is the basic operator that consists of  $3 \times 3$  convolutions followed by a rectified linear unit (ReLU) and batch normalization. The same operation exists between the separate blocks at every stage, but the yellow arrows are omitted for visibility. A red and blue arrows are  $2 \times 2$  average pooling and average unpooling operators, respectively, located between the stages. A violet arrow is the skip and concatenation operator. A green arrow is the simple  $1 \times 1$  convolution operator generating interpolated *k*-space data from multichannel data.

### B. Network Training

We use the  $l_2$  loss in the image domain for training. For this, the Fourier transform operator is placed as the last layer to convert the interpolated *k*-space data to the complex-valued image domain so that the loss values are calculated for the reconstructed image. Stochastic gradient descent (SGD) optimizer was used to train the network. For the IFT layer, the adjoint operation from SGD is also Fourier transform. The size of mini batch is 4, and the number of epochs in single and multi coil networks is 1000 and 500, respectively. The initial learning rate is  $10^{-5}$ , which gradually dropped to  $10^{-6}$  until 300-th epochs. The regularization parameter was  $\lambda = 10^{-4}$ .

The labels for the network were the images generated from direct Fourier inversion from fully sampled *k*-space data. The input data for the network was the regridded down-sampled *k*-space data from Cartesian, radial, and spiral trajectories. The details of the downsampling procedure will be discussed later. For each trajectory, we train the network separately.

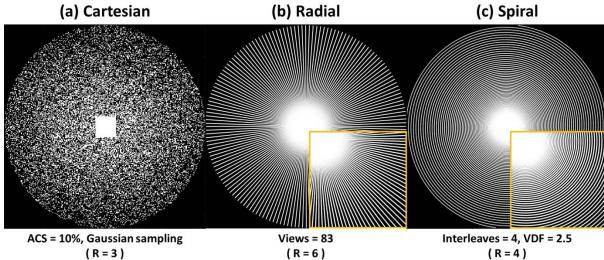
The proposed network was implemented using MatConvNet toolbox in MATLAB R2015a environment [27]. Processing units used in this research are Intel Core i7-7700 central processing unit and GTX 1080-Ti graphics processing unit. Training time lasted about 5 days.

## V. MATERIAL AND METHODS

### A. Data Acquisition

The evaluations were performed on single coil and multi coils *k*-space data for various *k*-space trajectories such as Cartesian, radial, and spiral cases.

For the Cartesian trajectory, knee *k*-space dataset (<http://mridata.org/>) were used. The raw data were acquired from 3D fast-spin-echo (FSE) sequence with proton density weighting included fat saturation comparison by a 3.0T whole body MR system (Discovery MR 750, DV22.0, GE Healthcare, Milwaukee, WI, USA). The repetition time (TR) and echo time (TE) were 1550 ms and 25 ms, respectively. There were 320 slices in total, and the thickness of each slice was 0.6 mm. The field of view (FOV) defined  $160 \times 128 \text{ mm}^2$  and the size of acquisition matrix is  $320 \times 256$ . The voxel size was 0.5 mm. The number of coils is 8. Eight coils *k*-space data was used for



**Fig. 5.** Various under-sampling patterns: (a) Cartesian undersampling at  $R = 3$ , (b) radial undersampling at  $R = 6$ , and (c) spiral undersampling at  $R = 4$ . Magnified views are provided for radial and spiral trajectories.

multi-coil  $k$ -space deep learning. In addition, to evaluate the performance of the algorithm for the single coil experiment, coil compression (<http://mrsrl.stanford.edu/tao/software.html>) was applied to obtain a single coil  $k$ -space data. For the Cartesian trajectory as shown in Fig. 5(a), the input  $k$ -space was downsampled to a Gaussian pattern using x4 acceleration factor in addition to the 10% auto-calibration signal (ACS) line. Therefore, the net acceleration factor is about 3 ( $R = 3$ ). Among the 20 cases of knee data, 18 cases were used for training, 1 case for validation, and the other for test.

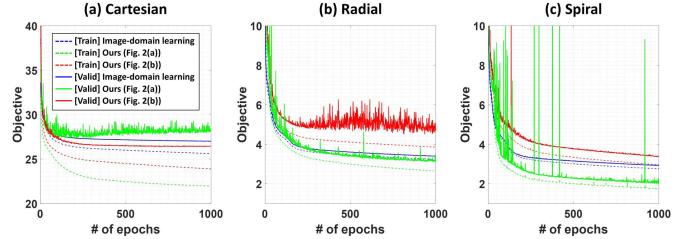
For radial and spiral sampling patterns, a synthesized  $k$ -space data from Human Connectome Project (HCP) MR dataset (<https://db.humanconnectome.org>) were used. Specifically, the multi-coil radial and spiral  $k$ -space data are generated using MRI simulator (<http://bigwww.epfl.ch/algorithms/mri-reconstruction/>). The T2 weighted brain images contained within the HCP were acquired Siemens 3T MR system using a 3D spin-echo sequence. The TR and TE were 3200 ms and 565 ms, respectively. The number of coils was 8, but the final reconstruction was obtained as the absolute of the sum. The FOV was  $224 \times 224$  mm<sup>2</sup>, and the size of acquisition matrix was  $320 \times 320$ . The voxel size was 0.7 mm. The total of 199 subject datasets was used in this paper. Among the 199 subject, 180 were used for network training, 10 subject for validation, and the other subject for test. Fig. 5(b) shows the down-sampled  $k$ -space radial sampling patterns. The down-sampled radial  $k$ -space consists of only 83 spokes, which corresponds to  $R = 6$  acceleration factor compared to the 503 spokes for the fully sampled data that were used as the ground-truth. On the other hand, Fig. 5(c) shows the down-sampled spiral sampling pattern, composed of 4 interleaves that corresponds to  $R = 4$  acceleration compared to the full spiral trajectory with 16 interleaves. The spiral  $k$ -space trajectory was obtained with a variable density factor (VDF) of 2.5.

### B. Performance Evaluation

For quantitative evaluation, the normalized mean square error (NMSE) value was used, which is defined as

$$NMSE = \frac{\sum_{i=1}^M \sum_{j=1}^N [x^*(i, j) - x(i, j)]^2}{\sum_{i=1}^M \sum_{j=1}^N [x^*(i, j)]^2}, \quad (20)$$

where  $x$  and  $x^*$  denote the reconstructed images and ground truth, respectively.  $M$  and  $N$  are the number of pixel for row



**Fig. 6.** Objective functions of a single coil for (a) Cartesian, (b) radial, and (c) spiral trajectories. Dashed and solid lines indicate an objective function of the train and validation phase, respectively. [Train]: training phase, [Valid]: validation phase.

and column. We also use the peak signal to noise ratio (PSNR), which is defined by

$$PSNR = 20 \cdot \log_{10} \left( \frac{NM \|x^*\|_\infty}{\|x - x^*\|_2} \right). \quad (21)$$

We also use the structural similarity (SSIM) index [28], defined as

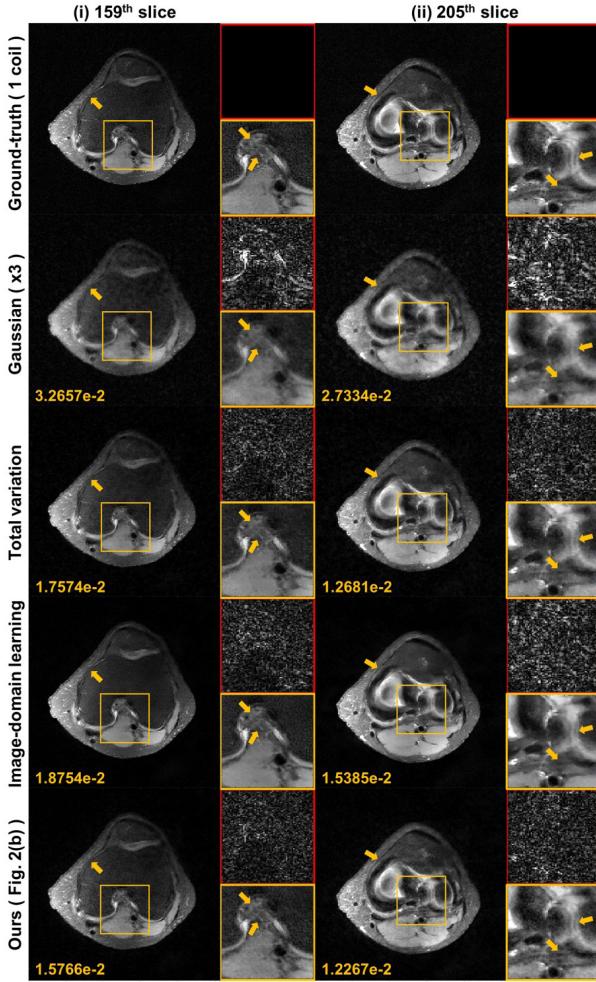
$$SSIM = \frac{(2\mu_x\mu_{x^*} + c_1)(2\sigma_{xx^*} + c_2)}{(\mu_x^2 + \mu_{x^*}^2 + c_1)(\sigma_x^2 + \sigma_{x^*}^2 + c_2)}, \quad (22)$$

where  $\mu_x$  is a average of  $x$ ,  $\sigma_x^2$  is a variance of  $x$  and  $\sigma_{xx^*}$  is a covariance of  $x$  and  $x^*$ . There are two variables to stabilize the division such as  $c_1 = (k_1 L)^2$  and  $c_2 = (k_2 L)^2$ .  $L$  is a dynamic range of the pixel intensities.  $k_1$  and  $k_2$  are constants by default  $k_1 = 0.01$  and  $k_2 = 0.03$ .

For extensive comparative study, we also compared with the following algorithms: total variation (TV) penalized CS, ALOHA [15], and four types of CNN models including the variational model [5], a cascade model [9], the cross-domain model called KIKI network [29], and an image-domain model [7]. In particular, [7] is a representative example of Fig. 1(a). Specifically, the image domain residual learning using the standard U-Net backbone in Fig. 4 was used. Unlike the proposed network, the input and output are an artifact corrupted image and artifact-only image, respectively [8]. In addition, the variational model [5] and the cascade model [9] represent Fig. 1(b). The cross-domain model is formed by linking the  $k$ -space model in Fig. 1(d) and the image-domain model in Fig. 1(a). Unfortunately, Fig. 1(c) does not scale well due to the enormous memory requirement, so cannot be used in the comparative study. For fair comparison, the cascade [9] and cross-domain [29] networks were modified for parallel imaging. All the neural networks were trained using exactly the same data set. For ALOHA [15] and the proposed method in Fig. 3(b), the total variation based  $k$ -space weighting was used.

## VI. RESULTS

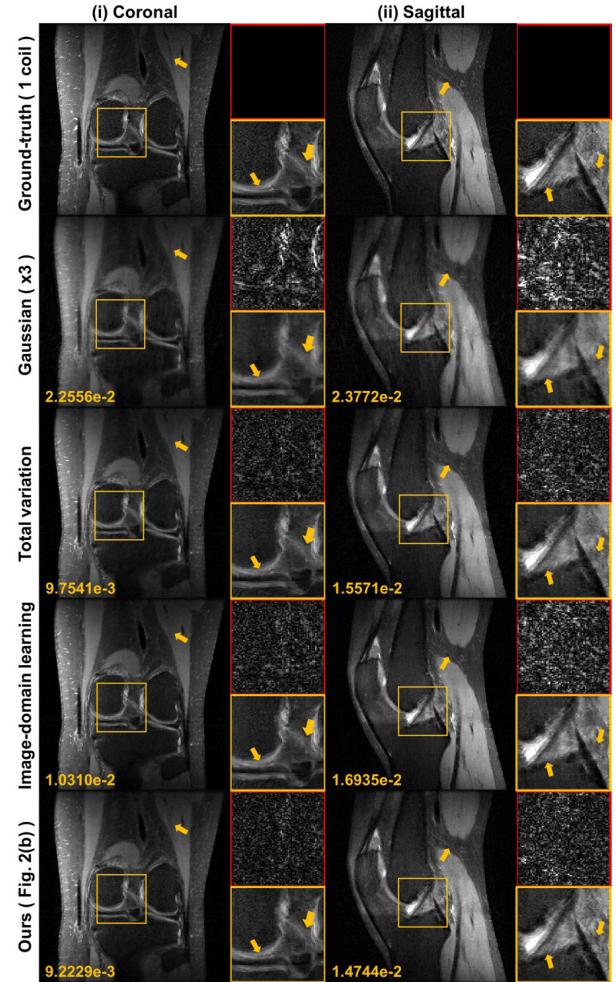
To evaluate the performance of sparsifications in the single coil, the proposed method was trained using the sparsifications as shown in as Figs. 3(a)(b). Fig. 6 shows the objective functions along the trajectories such as (a) Cartesian, (b) radial, and (c) spiral trajectory. In the Cartesian trajectory, the proposed network in Fig. 3(b) produces the lowest curve in the validation phase (see red curves in Fig. 6(a)). The proposed network in Fig. 3(a) shows the best convergence during the training,



**Fig. 7.** Reconstruction results from Cartesian trajectory at  $R = 3$  in single coil. The difference images were amplified five times. Yellow and red boxes illustrate the enlarged and difference views, respectively. The number written to the images is the NMSE value.

but the generalization at the test phase was not good (see green curves in Fig. 6(a)). In the non-Cartesian trajectories, the best convergence appears using the proposed network with only skipped connection in Fig. 3(a) (see green curves in Fig. 6(b)(c)). Based on these convergence behavior and generalization, the proposed network was trained with different sparsification schemes. The network in Fig. 3(b) was trained for the Cartesian trajectory and the network in Fig. 3(a) was used for the non-Cartesian trajectories such as radial and spiral trajectories.

Fig. 7 shows the results of single coil reconstruction from Cartesian trajectory using the architecture with skipped connection and weighting layer as shown in Fig. 3(a). While all the algorithms provide good reconstruction results, the proposed method most accurately recovered high frequency edges and textures as shown in the enlarged images and difference images of Fig. 7. Fig. 8 shows the reformed images along the (i) coronal and (ii) sagittal directions. Again, the reformatted coronal and sagittal images by the proposed method preserved the most detailed structures of underlying images without any artifact along the slice direction. The quantitative comparison in Table I in terms of average PSNR, NMSE,



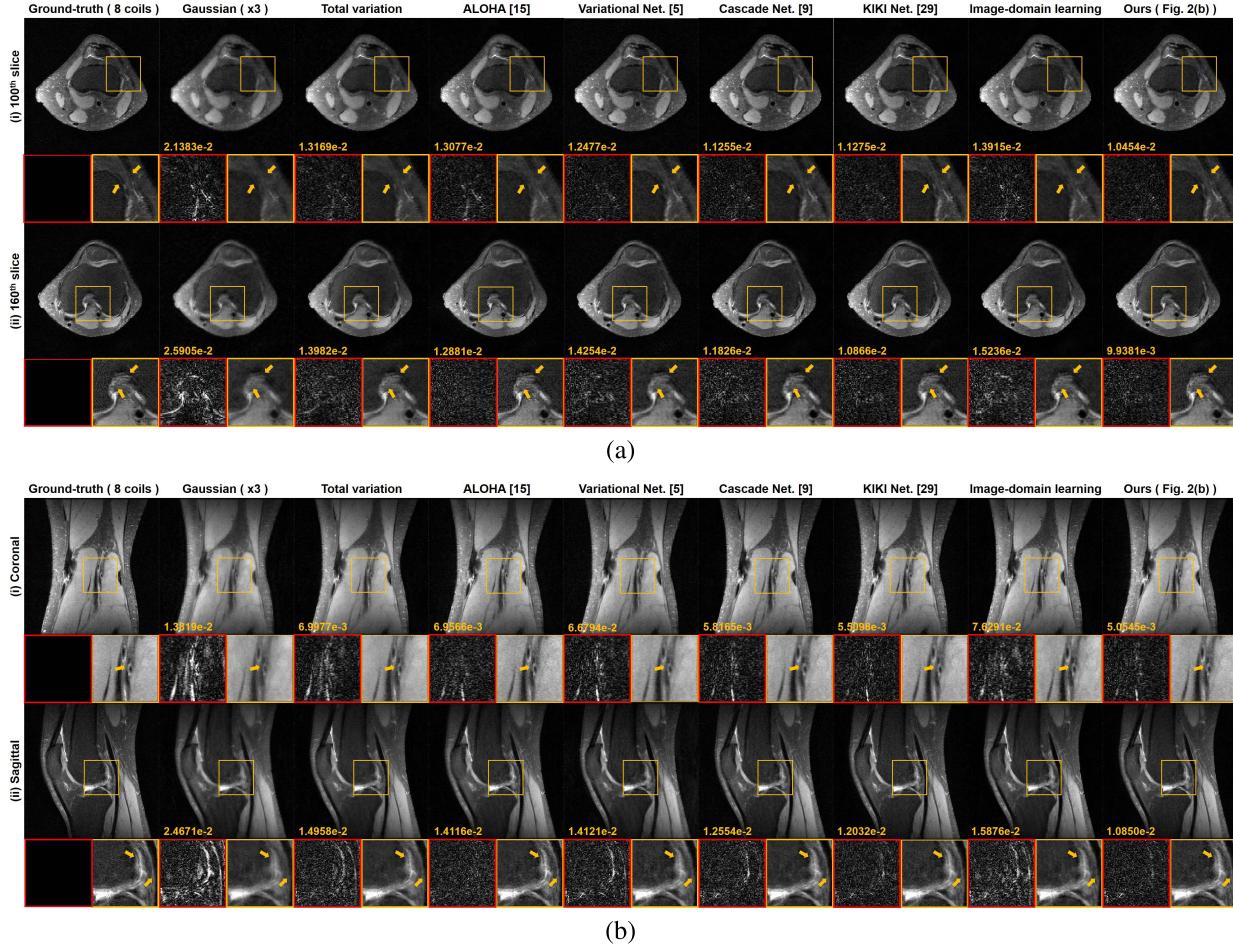
**Fig. 8.** (i) Coronal and (ii) sagittal reformed reconstruction results from Cartesian trajectory at  $R = 3$  in single coil. The results from first to last rows indicate ground-truth, downsampled, total variation, image-domain learning and the proposed method. Yellow and red boxes illustrate the enlarged and difference views, respectively. The difference images were amplified five times. The number written to the images is the NMSE value.

TABLE I  
QUANTITATIVE COMPARISON FROM CARTESIAN TRAJECTORY AT  $R = 3$  IN SINGLE COIL

Metric	Input ( 1 coil, $\times 3$ )	Total variation	Image-domain learning	Ours ( Fig. 3(b) )
PSNR [dB]	33.9058	35.7112	35.5419	<b>35.9586</b>
NMSE ( $\times 10^{-2}$ )	2.7779	1.9067	1.9567	<b>1.7921</b>
SSIM	0.7338	0.7548	0.7447	<b>0.7664</b>
Times (sec/slice)	-	0.1344	0.0272	0.0377

and SSIM values also confirmed that the proposed  $k$ -space interpolation method produced the best quantitative values in all area. The computation time of the proposed method is slightly slower than the image-domain learning because of the weighting and Fourier transform operations, but it is still about 3.5 times faster than the total variation penalized compressed sensing (CS) approach.

Fig. 9 shows the parallel imaging results from eight coil measurement. Because the ALOHA [15] and the proposed method directly interpolate missing  $k$ -space, these methods clearly preserve textures detail structures as shown in Fig. 9. However, ALOHA [15] is more than 100 times slower than



**Fig. 9.** Reconstruction results from Cartesian trajectory at  $R = 3$  in multi coils: (a) axial, (b-i) coronal and (b-ii) sagittal reconstruction results. Yellow and red boxes illustrate the enlarged and difference views, respectively. The difference images were amplified five times. The number written to the images is the NMSE value.

TABLE II  
QUANTITATIVE COMPARISON FROM CARTESIAN TRAJECTORY AT  $R = 3$  IN 8 COILS PARALLEL IMAGING

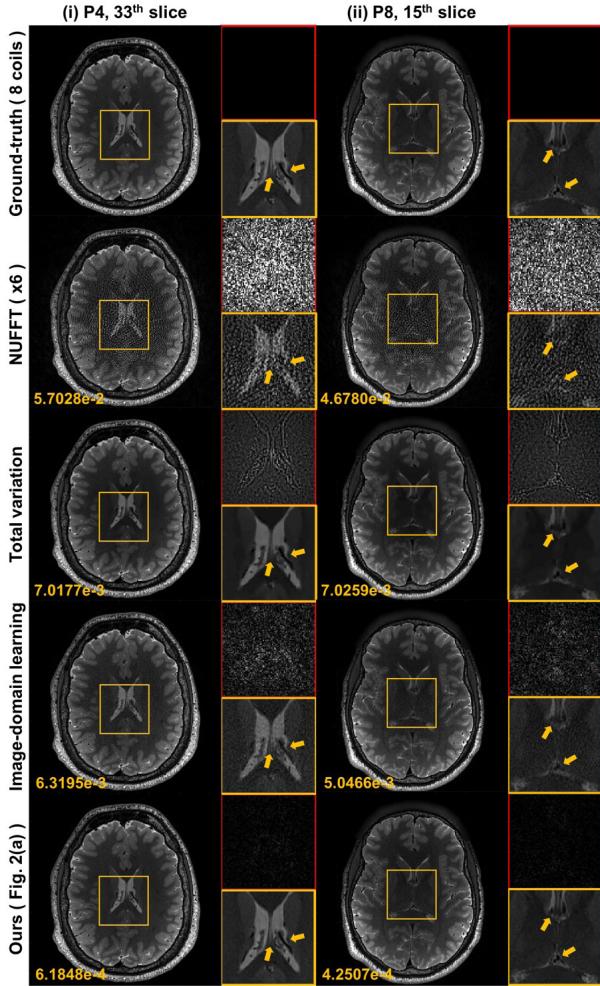
Metric	Input ( 8 coils, $\times 3$ )	Total variation	ALOHA [15]	Variational Net. [5]	Cascade Net. [9]	KIKI Net. [29]	Image-domain learning	Ours ( Fig. 3(b) )
PSNR [dB]	34.2021	36.0689	36.1013	36.1428	36.6055	36.6847	35.8497	<b>36.9931</b>
NMSE ( $\times 10^{-2}$ )	2.3384	1.5826	1.5812	1.5379	1.4196	1.4086	1.6497	<b>1.3154</b>
SSIM	0.7710	0.7878	0.7914	0.7966	0.7982	0.8018	0.7776	<b>0.8087</b>
Times (sec/slice)	-	1.0938	16.5554	0.1688	0.2219	0.5438	0.1188	0.1438

the  $k$ -space deep learning as shown in Table II. All CNN methods except the imaging-domain model outperform than CS methods. Although the cascade model [9] was performed with data consistency step, the method did not completely overcome the limitations of image-domain learning. In the cross-domain learning [29], they proposed to train  $k$ -space and image-domain sequentially. The cross-domain network consists of deeper layers (100 layers; 25 layers  $\times$  4 individual models) than the proposed method, but the performance was worse than our method. As shown Table II, the proposed method shows best performance in terms of average PSNR, NMSE, and SSIM values.

Fig. 10 shows the reconstruction images from x6 accelerated radial sampling patterns using the architecture in Fig. 3(a) for 8 coils parallel imaging. The results for single coil are shown in Fig. S1 in Supplementary Material. The proposed  $k$ -space deep learning provided realistic image quality and

preserves the detailed structures as well as the textures, but the image domain network failed to remove the noise signals and the total variation method did not preserve the realistic textures and sophisticated structures. Our method also provides much smaller NMSE values, as shown at the bottom of each Fig. 10 and Fig. S1 in Supplementary Material. Average PSNR, NMSE and SSIM values are shown in Table III and Table S1 in Supplementary Material for multi coils and single coil cases, respectively. The average values were calculated across all slices and 9 subjects. The proposed  $k$ -space deep learning provided the best quantitative values.

Fig. 11 shows the reconstruction images from x4 accelerated spiral trajectory for 8 coils parallel imaging and Fig. S2 in Supplementary Material illustrate the single coil results, respectively. Similar to the radial sampling patterns, the proposed method provides significantly improved image reconstruction results, and the average PSNR, NMSE and SSIM values



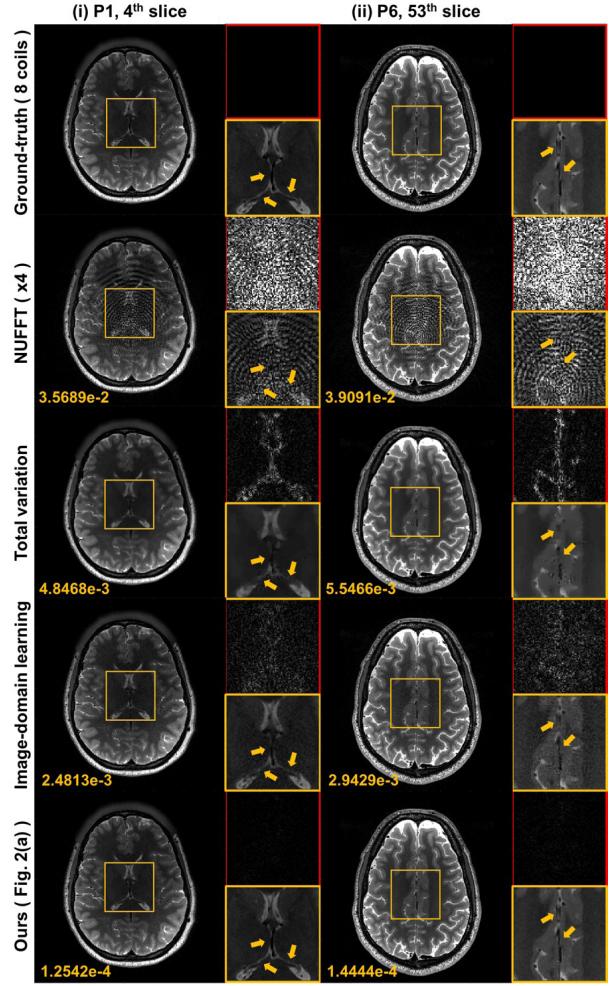
**Fig. 10.** Reconstruction results from radial trajectory at  $R = 6$  in 8 coils parallel imaging. The difference images were amplified five times. Yellow and red boxes illustrate the enlarged and difference views, respectively. The number written to the images is the NMSE value.

**TABLE III**  
QUANTITATIVE COMPARISON FROM RADIAL UNDERSAMPLING  
AT  $R = 6$  IN 8 COILS PARALLEL IMAGING

Metric	Input ( 8 coils, $\times 6$ )	Total variation	Image-domain learning	Ours ( Fig. 3(a) )
PSNR [dB]	30.6396	38.2357	39.9140	<b>50.8136</b>
NMSE ( $\times 10^{-2}$ )	3.6137	0.5998	0.4224	<b>0.0353</b>
SSIM	0.5927	0.9528	0.9020	<b>0.9908</b>

in Table IV and Table S2 in Supplementary Material also confirm that the proposed method consistently outperform other method for all patients.

To evaluate the improvements of the proposed method, a radiologist (with 9 years of experience) thoroughly reviewed the reconstructed images. For radial trajectory images, the pyramidal tract (arrows in Fig. S3(i) of the Supplementary Material), a bundle of nerve fibers conducting motor signal from the motor cortex to the brainstem or to the spinal cord, was evaluated. It was noted that high signal intensity of the pyramidal tract was exaggerated on the TV method, which could be misdiagnosed as an abnormal finding. In addition, the ability to discriminate a pair of internal cerebral veins were



**Fig. 11.** Reconstruction results from spiral trajectory at  $R = 4$  in 8 coils parallel imaging. The difference images were amplified five times. Yellow and red boxes illustrate the enlarged and difference views, respectively. The number written to the images is the NMSE value.

**TABLE IV**  
QUANTITATIVE COMPARISON FROM SPIRAL UNDERSAMPLING  
AT  $R = 4$  IN 8 COILS PARALLEL IMAGING

Metric	Input ( 8 coils, $\times 4$ )	Total variation	Image-domain learning	Ours ( Fig. 3(a) )
PSNR [dB]	30.3733	38.7646	41.0581	<b>53.5643</b>
NMSE ( $\times 10^{-2}$ )	4.0373	0.6130	0.3404	<b>0.0201</b>
SSIM	0.6507	0.9487	0.8881	<b>0.9940</b>
Times (sec/slice)	-	7.8818	0.1375	0.1431

evaluated. While the TV and image-domain learning methods can not differentiate the two veins, the proposed method is able to show the two internal cerebral veins separately (arrows on Fig. S3(ii) in Supplementary Material). With regard to spiral trajectory images (Fig. S4 in the Supplementary Material), the TV method shows bright dot-like artifacts along several slices (Fig. S4(i)). When the small T2 hyperintensity lesions in the left frontal lobe were evaluated, the TV method fails to demonstrate the lesions. The image domain learning preserves the lesions, but the margin of the lesions is blurry in the noisy background. The small lesions are clearly depicted on the proposed method (arrows in Fig. S4(ii)). Overall, the quality of image reconstruction of the proposed method was superior to that of other methods.

## VII. DISCUSSION

In order to improve the performance of ALOHA, the matrix pencil size should be significantly large, which is not possible in standard ALOHA formulation due to the large memory requirement and extensive computational burden. We believe that one of the important reasons that the proposed k-space deep learning provides better performance than ALOHA is that the cascaded convolution results in much longer filter length. Nevertheless, with the same set of trained filters, our network can be adapted to different input images due to the combinatorial nature of ReLU nonlinearity. We believe that this also contributes to the benefits of k-space learning over ALOHA. Moreover, efficient signal representation in  $k$ -space domain, which is the key idea of ALOHA, can synergistically work with the expressivity of the neural network to enable better performance than image domain learning.

## VIII. CONCLUSION

Inspired by a link between the ALOHA and deep learning, this paper showed that fully data-driven  $k$ -space interpolation is feasible by using  $k$ -space deep learning and the image domain loss function. The proposed  $k$ -space interpolation network outperformed the existing image domain deep learning for various sampling trajectory. As the proposed  $k$ -space interpolation framework is quite effective and also supported by novel theory, so we believe that this opens a new area of researches for many Fourier imaging problems.

## REFERENCES

- [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
- [3] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, Nov. 2015, pp. 234–241.
- [4] S. Wang *et al.*, "Accelerating magnetic resonance imaging via deep learning," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2016, pp. 514–517.
- [5] K. Hammernik *et al.*, "Learning a variational network for reconstruction of accelerated MRI data," *Magn. Reson. Med.*, vol. 79, no. 6, pp. 3055–3071, Jun. 2018.
- [6] D. Lee, J. Yoo, S. Tak, and J. Ye, "Deep residual learning for accelerated MRI using magnitude and phase networks," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 9, pp. 1985–1995, Sep. 2018.
- [7] Y. Han, J. Yoo, H. H. Kim, H. J. Shin, K. Sung, and J. C. Ye, "Deep learning with domain adaptation for accelerated projection-reconstruction MR," *Magn. Reson. Med.*, vol. 80, no. 3, pp. 1189–1205, Sep. 2018. doi: [10.1002/mrm.27106](https://doi.org/10.1002/mrm.27106).
- [8] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.
- [9] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for dynamic MR image reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 2, pp. 491–503, Feb. 2018.
- [10] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, pp. 487–492, Mar. 2018.
- [11] M. Lustig, D. Donoho, and J. Pauly, "Sparse MRI: The application of compressed sensing for rapid mr imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, Dec. 2007.
- [12] H. Jung, K. Sung, K. Nayak, E. Kim, and J. Ye, "K-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI," *Magn. Reson. Med.*, vol. 61, no. 1, pp. 103–116, Jan. 2009.
- [13] P. J. Shin *et al.*, "Calibrationless parallel imaging reconstruction based on structured low-rank matrix completion," *Magn. Reson. Med.*, vol. 72, no. 4, pp. 959–970, Oct. 2014.
- [14] J. P. Haldar, "Low-rank modeling of local  $k$ -space neighborhoods (LORAKS) for constrained MRI," *IEEE Trans. Med. Imag.*, vol. 33, no. 3, pp. 668–681, Mar. 2014.
- [15] K. H. Jin, D. Lee, and J. C. Ye, "A general framework for compressed sensing and parallel MRI using annihilating filter based low-rank Hankel matrix," *IEEE Trans. Comput. Imaging*, vol. 2, no. 4, pp. 480–495, Dec. 2016.
- [16] G. Ongie and M. Jacob, "Off-the-grid recovery of piecewise constant images from few Fourier samples," *SIAM J. Imag. Sci.*, vol. 9, no. 3, pp. 1004–1041, 2016.
- [17] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL: Model-based deep learning architecture for inverse problems," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 394–405, Feb. 2019.
- [18] J. C. Ye, J. M. Kim, K. H. Jin, and K. Lee, "Compressive sampling using annihilating filter-based low-rank interpolation," *IEEE Trans. Inf. Theory*, vol. 63, no. 2, pp. 777–801, Feb. 2017.
- [19] J. C. Ye, Y. Han, and E. Cha, "Deep convolutional framelets: A general deep learning framework for inverse problems," *SIAM J. Imag. Sci.*, vol. 11, no. 2, pp. 991–1048, 2018.
- [20] M. Akçakaya, S. Moeller, S. Weingärtner, and K. Uğurbil, "Scan-specific robust artificial-neural-networks for  $k$ -space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging," *Magn. Reson. Med.*, vol. 81, no. 1, pp. 439–453, Jun. 2019.
- [21] A. Pramanik, H. K. Aggarwal, and M. Jacob, "Off-the-grid model based deep learning (O-MODL)," 2018, [arXiv:1812.10747](https://arxiv.org/abs/1812.10747). [Online]. Available: <https://arxiv.org/abs/1812.10747>
- [22] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL-MUSSELS: Model-based deep learning for multi-shot sensitivity encoded diffusion MRI," 2018, [arXiv:1812.08115](https://arxiv.org/abs/1812.08115). [Online]. Available: <https://arxiv.org/abs/1812.08115>
- [23] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.
- [24] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, p. 717, Dec. 2009.
- [25] G. Ongie and M. Jacob, "A fast algorithm for convolutional structured low-rank matrix recovery," *IEEE Trans. Comput. Imag.*, vol. 3, no. 4, pp. 535–550, Dec. 2017.
- [26] J. C. Ye and W. K. Sung, "Understanding geometry of encoder-decoder CNNs," 2019, [arXiv:1901.07647](https://arxiv.org/abs/1901.07647). [Online]. Available: <https://arxiv.org/abs/1901.07647>
- [27] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for Matlab," in *Proc. 23rd ACM Int. Conf. Multimedia*, Oct. 2015, pp. 689–692.
- [28] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [29] T. Eo, Y. Jun, T. Kim, J. Jang, H.-J. Lee, and D. Hwang, "KIKI-net: Cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images," *Magn. Reson. Med.*, vol. 80, no. 5, pp. 2188–2201, Nov. 2018.