KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images

Taejoon Eo[1] | **Yohan Jun[1]** | **Taeseong Kim[1]** | **Jinseong Jang[1]** | **Ho-Joon Lee[2]** |
Dosik Hwang[1]

[1] School of Electrical and Electronic Engineering, Yonsei University, Seoul, Korea

[2] Department of Radiology and Research Institute of Radiological Science, Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea

Correspondence
Dosik Hwang, College of Engineering, Yonsei University, 50, Yonsei-ro, Seodaemun-gu, Seoul, 03722, Korea.
Email: dosik.hwang@yonsei.ac.kr

Purpose: To demonstrate accurate MR image reconstruction from undersampled k-space data using cross-domain convolutional neural networks (CNNs)

Methods: Cross-domain CNNs consist of 3 components: (1) a deep CNN operating on the k-space (KCNN), (2) a deep CNN operating on an image domain (ICNN), and (3) an interleaved data consistency operations. These components are alternately applied, and each CNN is trained to minimize the loss between the reconstructed and corresponding fully sampled k-spaces. The final reconstructed image is obtained by forward-propagating the undersampled k-space data through the entire network.

Results: Performances of K-net (KCNN with inverse Fourier transform), I-net (ICNN with interleaved data consistency), and various combinations of the 2 different networks were tested. The test results indicated that K-net and I-net have different advantages/disadvantages in terms of tissue-structure restoration. Consequently, the combination of K-net and I-net is superior to single-domain CNNs. Three MR data sets, the T_2 fluid-attenuated inversion recovery (T_2 FLAIR) set from the Alzheimer's Disease Neuroimaging Initiative and 2 data sets acquired at our local institute (T_2 FLAIR and T_1 weighted), were used to evaluate the performance of 7 conventional reconstruction algorithms and the proposed cross-domain CNNs, which hereafter is referred to as KIKI-net. KIKI-net outperforms conventional algorithms with mean improvements of 2.29 dB in peak SNR and 0.031 in structure similarity.

Conclusion: KIKI-net exhibits superior performance over state-of-the-art conventional algorithms in terms of restoring tissue structures and removing aliasing artifacts. The results demonstrate that KIKI-net is applicable up to a reduction factor of 3 to 4 based on variable-density Cartesian undersampling.

KEYWORDS
convolutional neural networks, cross-domain deep learning, image reconstruction, k-space completion, MRI acceleration

1 | INTRODUCTION

Magnetic resonance imaging (MRI) is a noninvasive medical imaging technique that provides various contrast mechanisms for visualizing anatomical structures and physiological functions. However, MRI is relatively slow because of its long acquisition time, and therefore is used infrequently for applications that require fast scanning. The long MRI acquisition time is the

result of not being able to simultaneously sample multiple data points; instead, the data should be sequentially sampled in time (i.e., point by point) over the images' Fourier space, referred to as the "k-space." Despite the development of advanced hardware and imaging techniques, such as parallel imaging[1,2] and echo planar imaging,[3] the maximum for magnetic resonance (MR) data collection remains limited. Instead of acquiring the fully sampled MR data in the k-space, the k-space data can be subsampled at a frequency that is lower than the Nyquist rate (i.e., it can be undersampled) to accelerate the acquisition process. However, simple undersampling schemes result in aliasing artifacts in the reconstructed images, and many tissue structures in images are obscured by these artifacts. Therefore, various efforts have focused on developing advanced reconstruction algorithms to improve the image quality of the undersampled MR images. One of the most representative algorithms is compressed sensing (CS), which uses sparsity in specific transform domains.[4-8] As more advanced algorithms, CS algorithms have been combined with parallel imaging[9-11] and low-rank constraint terms.[12,13] More recently, image-adaptive algorithms that enforce sparsity on image patches, such as dictionary-learning algorithms,[14-18] have appeared.

Compressed-sensing MRI (CS-MRI) uses sparse coefficients in global sparsifying transforms, such as wavelets, curvelets, and contourlets transform[4-8]; however, in the l_1 minimization process, the sparse coefficient values tend to be significantly smaller than those of the original coefficient values. This value reduction hides detailed structures and results in blur artifacts in the reconstructed images. Furthermore, high-frequency oscillatory artifacts remain in the reconstructed image when large errors occurring in the transform domain are not properly reduced via minimization.[14] Therefore, CS with global sparsifying transforms is generally limited to a reduction factor of 2.5 to 3 for typical MR images.[14] Dictionary learning MRI (DL-MRI) updates adaptive dictionaries by alternating back and forth between an image domain and the k-space.[14] This image-adaptive sparsifying algorithm can represent better sparsification of images compared with CS-MRI, owing to dictionaries that are learned from the image itself or images from similar classes. Therefore, DL-MRI achieves better results than nonadaptive CS algorithms. However, detailed information of images may disappear if the number of dictionary patches is not sufficiently large or if the reduction factor is too high to generate dictionaries that do not induce blurring.

Meanwhile, several recent studies have demonstrated the applicability of deep-learning techniques to the reconstruction of undersampled MR images[19-21] or CT images.[22] In training, tuples of undersampled images and fully sampled images are fed to convolutional neural networks (CNNs) to learn the relationship between the undersampled images and the corresponding fully sampled images.[19,22] In testing, arbitrary undersampled images obtained with the same protocols are fed to the well-trained CNNs, and the final reconstructed images are obtained as outputs of the CNNs. These studies can be interpreted as attempts to replace 3 main steps of the conventional reconstruction algorithms: (1) selection of image characteristics that are assumed by humans (e.g., sparsity), (2) extraction of features that represent the image characteristics (e.g., wavelet coefficients), and (3) optimization of features (e.g., l_1 minimization). These 3 steps are then replaced with (1) data-driven feature extraction (i.e., deep neural networks such as CNNs) and (2) a unified optimization method (i.e., loss backpropagation).

Existing CNN-based algorithms outperform conventional CS algorithms because of their data-driven feature extraction and high-nonlinearity properties.[19-22] However, the existing algorithms contain 2 major limitations. First, the CNNs used in previous studies are only trained on the image domain (i.e., the CNNs estimate true images from images in which the detailed structures are already distorted or have even disappeared).[19-22] To resolve the problem, we developed a deep CNN that operates on the k-space and can use the maximum possible extent of the k-space itself, which contains the true high-frequency components of the images in its outer area (although some high-frequency components may be missing). Furthermore, an iterative deep-learning approach is introduced. We iteratively (or alternately) applied 2 different CNNs operating on different domains (the k-space and image domain), and data consistency was interleaved among the CNNs. The second limitation of the earlier studies is that the network depth was shallower (3 to 5 layers) than that of most networks used in recent image-restoration studies.[19-22] In some studies, deep CNNs with layer depths greater than 20 afford much more promising results than shallower networks because of their larger receptive fields.[23,24] In the present study, we exploited CNNs with layer depths greater than or equal to 20 and compared their results with those of shallower networks (e.g., 3 layers).

This study proposes a new algorithm that can estimate fully sampled MR data from undersampled MR data using a combination of 4 different CNNs, and is hereafter referred to as the KIKI-net (the network architecture operating on k-space, image, k-space, and image sequentially). We designate this type of network architecture as cross-domain CNNs (CD-CNNs). The proposed KIKI-net can use true sampled points in the k-space to the maximum extent possible, which is highly effective in restoring detailed tissue structures in images as well as in reducing aliasing artifacts.

2 | METHODS

All experiments conducted in the present study were approved by the institutional review board. Written informed consent was obtained from all human subjects.

This section provides the problem formulation, network component details, practical implementation of CD-CNNs, and experiment framework. The network architecture is based on deep CNNs that have been proven to learn pixel-to-pixel regression in the area of computer vision.[25-32] The CD-CNNs consist of 3 components: a deep CNN for k-space completion (KCNN), a deep CNN for image restoration (ICNN), and an interleaved data consistency (IDC).

2.1 | Problem formulation

Let $\mathbf{k} \in \mathbb{C}^{n_{kx} \times n_{ky}}$ denote a 2D complex-valued MR k-space. Our purpose is to reconstruct a fully sampled image $\mathbf{x}$ from the undersampled k-space, $\mathbf{k_u}$, obtained as follows:

$$\mathbf{k_u} = \mathbf{U} \circ \mathbf{k} = \mathbf{U} \circ \mathcal{F}_{2D}(\mathbf{x}) = \mathbf{k_{u,r}} + i\mathbf{k_{u,i}} \tag{1}$$

$$\mathbf{x_u} = \mathcal{F}_{2D}^{-1}(\mathbf{k_u}) = \mathbf{x_{u,r}} + i\mathbf{x_{u,i}} \tag{2}$$

where $\mathbf{k_u} \in \mathbb{C}^{n_{kx} \times n_{ky}}$ denotes the undersampled k-space; $\mathbf{U} \in \mathbb{R}^{n_{kx} \times n_{ky}}$ denotes the binary undersampling mask; $\circ$ denotes element-wise multiplication; $\mathcal{F}_{2D}$ and $\mathcal{F}_{2D}^{-1}$ denote the 2D Fourier transform (FT) and inverse Fourier transform (IFT), respectively; $\mathbf{k_{u,r}} \in \mathbb{R}^{n_{kx} \times n_{ky}}$ and $\mathbf{k_{u,i}} \in \mathbb{R}^{n_{kx} \times n_{ky}}$ denote the real and imaginary channels of $\mathbf{k_u}$, respectively; $\mathbf{x_u}$ denotes the undersampled image; and $\mathbf{x_{u,r}}$ and $\mathbf{x_{u,i}}$ denote the real and imaginary channels of $\mathbf{x_u}$, respectively. To solve the ill-posed problem of reconstructing $\mathbf{x}$ from a small number of samples in the k-space, $\mathbf{k_u}$, we introduce 2 minimization equations: 1 for k-space completion, and the other for image restoration.

$$\mathbf{k}\text{-space completion}: \underset{\widehat{\mathbf{k}}}{\operatorname{argmin}} \|\mathbf{k} - \widehat{\mathbf{k}}\|_2^2$$

$$= \underset{\theta_k}{\operatorname{argmin}} \|\mathbf{k} - H_k(\mathbf{k_u}; \theta_k)\|_2^2 \tag{3}$$

where $\widehat{\mathbf{k}}$ is the estimation of the true k-space, $\mathbf{k}$. In terms of a learning algorithm, $\widehat{\mathbf{k}}$ is estimated by a hypothesis function, H_k, with the input of $\mathbf{k_u}$ and the unknown parameters θ_k. Therefore, the minimization equation changes to find optimal θ_k as shown in the right side of Equation 3.

$$\text{image restoration}: \underset{\widehat{\mathbf{x}}}{\operatorname{argmin}} \|\mathbf{x} - \widehat{\mathbf{x}}\|_2^2 + \lambda \|\mathbf{k_u} - \mathbf{U} \circ \mathcal{F}_{2D}(\widehat{\mathbf{x}})\|_2^2$$

$$= \underset{\theta_x}{\operatorname{argmin}} \|\mathbf{x} - H_x(\mathbf{x_u}; \theta_x)\|_2^2 + \lambda \|\mathbf{k_u} - \mathbf{U} \circ \mathcal{F}_{2D}\left(H_x(\mathbf{x_u}; \theta_x)\right)\|_2^2 \tag{4}$$

where $\widehat{\mathbf{x}}$ is the estimation of the true image, $\mathbf{x}$; and H_x is the hypothesis function to estimate $\widehat{\mathbf{x}}$ with the input of $\mathbf{x_u}$ and the unknown parameters θ_x. The right term of Equation 4 is a regularization term for data consistency, and λ is the regularization parameter. Then, a combination form of Equations 3 and 4, which is the target objective function of this study, is

$$\underset{\theta_k, \theta_x}{\operatorname{argmin}} \left\| \mathbf{x} - H_x\left(\mathcal{F}_{2D}^{-1}\left(H_k(\mathbf{k_u}; \theta_k)\right); \theta_x\right) \right\|_2^2$$

$$+ \lambda \left\| \mathbf{k_u} - \mathbf{U} \circ \mathcal{F}_{2D}\left(H_x\left(\mathcal{F}_{2D}^{-1}\left(H_k(\mathbf{k_u}; \theta_k)\right); \theta_x\right)\right) \right\|_2^2 \tag{5}$$

We have determined that simultaneously obtaining θ_k and θ_x by solving Equation 5 is highly constrained in terms of network design due to high computational complexities, overfitting problems, and memory shortages. Therefore, we introduced an iterative deep-learning approach that obtains θ_k and θ_x by alternately solving Equations 3 and 4 until the final loss is saturated. The detailed process of our approach is presented in the following sections.

2.2 | Deep CNN for k-space completion

To solve the minimization equation in Equation 3, we introduced a deep CNN as the hypothesis function H_k, which completes unacquired points in the k-space using acquired points in the k-space. This CNN is denoted as KCNN. Equation 3 can then be rewritten as

$$\underset{\theta_k}{\operatorname{argmin}} \|\mathbf{k} - H_k(\mathbf{k_{in}}; \theta_k)\|_2^2 = \underset{\theta_{KCNN}}{\operatorname{argmin}} \|\mathbf{k} - H_{KCNN}(\mathbf{k_{in}}; \theta_{KCNN})\|_2^2 \tag{6}$$

where H_{KCNN} is the CNN-based hypothesis function; θ_{KCNN} represents the parameters of the CNN; $\mathbf{k_{in}}$ is the input k-space, which is $\mathbf{k_u}$ for the first KCNN. The network architecture of KCNN is presented in Figure 1A, and the forward-pass equations of KCNN, namely $H_{KCNN}(\mathbf{k_{in}}; \theta_{KCNN})$, are as follows.

The KCNN component consists of 3 network components: a feature-extraction net, an inference net, and a reconstruction net. In the feature-extraction net, features of the undersampled k-space data are extracted by a pair of convolution and activation layers. Feature maps are independently extracted from the real and imaginary k-spaces and are then concatenated. The forward-pass equation of the feature extraction net is

$$\mathbf{F_r} = \sigma(\mathbf{W_{F,r}} * \mathbf{k_{in,r}} + \mathbf{b_{F,r}}) \tag{7}$$

$$\mathbf{F_i} = \sigma(\mathbf{W_{F,i}} * \mathbf{k_{in,i}} + \mathbf{b_{F,i}}) \tag{8}$$

$$\mathbf{F} = [\mathbf{F_r}, \mathbf{F_i}] \tag{9}$$

where $\mathbf{k_{in,r}}$ and $\mathbf{k_{in,i}}$ denote the real and imaginary channels of $\mathbf{k_{in}}$ and $\mathbf{W_{F,r}} \in \mathbb{R}^{w \times h \times 1 \times c}$, and $\mathbf{W_{F,i}} \in \mathbb{R}^{w \times h \times 1 \times c}$ denote the weight matrices of the 2 corresponding convolution layers of the feature extraction net. For the weight matrices, 1, w, h, and c respectively denote the number of channels and the width, height, and number of weight matrices. Values of $\mathbf{b_{F,r}} \in \mathbb{R}^c$ and $\mathbf{b_{F,i}} \in \mathbb{R}^c$ denote the bias matrices; $\mathbf{F_r} \in \mathbb{R}^{n_{kx} \times n_{ky} \times c}$ and $\mathbf{F_i} \in \mathbb{R}^{n_{kx} \times n_{ky} \times c}$ denote the feature maps extracted from the real and imaginary channels of the k-space, respectively; and σ denotes activation function. The 2 feature maps, $\mathbf{F_r}$ and $\mathbf{F_i}$, are concatenated along the channel

A Architecture of KCNN

B Architecture of ICNN
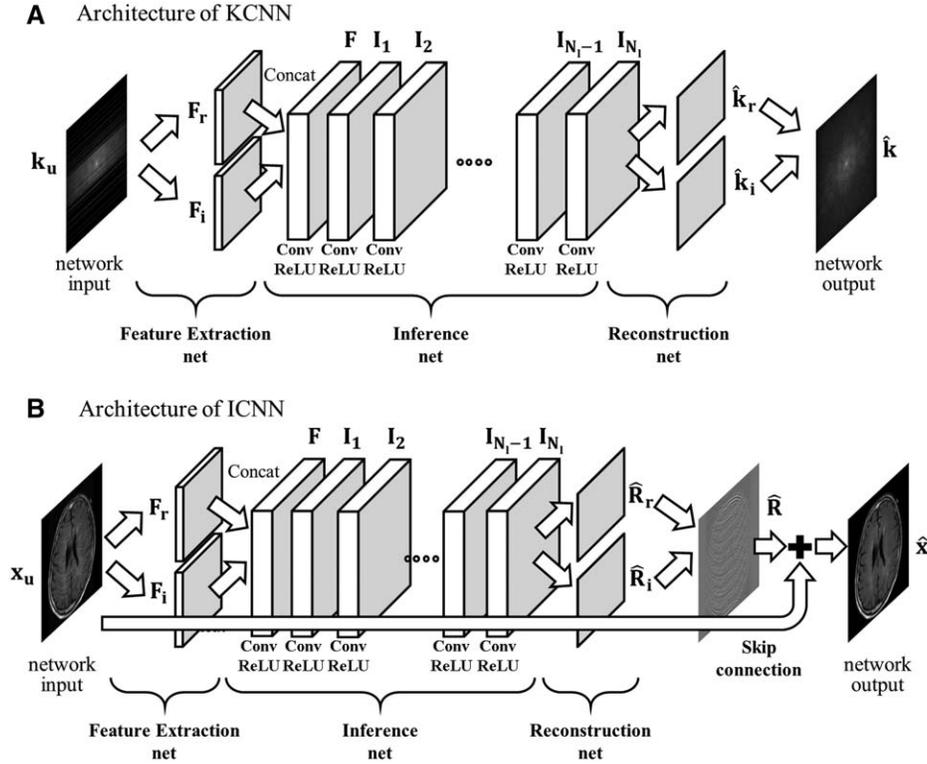
FIGURE 1 Network architecture of the deep convolutional neural network (CNN) for k-space completion (KCNN) (A) and the deep CNN for image restoration (ICNN) (B)

axis to form $\mathbf{F} \in \mathbb{R}^{n_{kx} \times n_{ky} \times (2c)}$ such that they are fed into the next net.

The inference net infers and fills the empty points of the feature maps. These are gradually filled while the feature maps pass through multiple convolution and activation layers. The forward-pass equation of the inference net is

$$\mathbf{I}_1 = \sigma(\mathbf{W}_{\mathbf{I}_1} * \mathbf{F} + \mathbf{b}_{\mathbf{I}_1}) \tag{10}$$

$$\mathbf{I}_n = \sigma(\mathbf{W}_{\mathbf{I}_n} * \mathbf{I}_{n-1} + \mathbf{b}_{\mathbf{I}_n}) \tag{11}$$

where $n = 2, \ldots, N_l - 1$, N_l; and N_l is the depth of layers. Values of $\mathbf{W}_{\mathbf{I}_1} \in \mathbb{R}^{w \times h \times (2c) \times c}$ and $\mathbf{W}_{\mathbf{I}_n} \in \mathbb{R}^{w \times h \times c \times c}$ denote the first and nth convolution matrices of the inference net, respectively; and $\mathbf{b}_{\mathbf{I}_1} \in \mathbb{R}^c$ and $\mathbf{b}_{\mathbf{I}_n} \in \mathbb{R}^c$ are the first and the nth bias values of the inference net, respectively. The final, fully filled feature maps are obtained through N_l different convolution and activation layers.

The reconstruction net receives the fully filled feature maps, $\mathbf{I}_{N_l} \in \mathbb{R}^{n_{kx} \times n_{ky} \times c}$, as input and forms the final network output with respect to the completed k-space. The following are the forward-pass equations for the reconstruction net:

$$\widehat{\mathbf{k}_r} = \mathbf{W}_{\mathbf{R},\mathbf{r}} * \mathbf{I}_{N_l} + \mathbf{b}_{\mathbf{R},\mathbf{r}} \tag{12}$$

$$\widehat{\mathbf{k}_i} = \mathbf{W}_{\mathbf{R},\mathbf{i}} * \mathbf{I}_{N_l} + \mathbf{b}_{\mathbf{R},\mathbf{i}} \tag{13}$$

$$H_{\text{KCNN}}(\mathbf{k}_{\mathbf{in}}; \boldsymbol{\theta}_{\text{KCNN}}) = \widehat{\mathbf{k}}_{\text{KCNN}} = \widehat{\mathbf{k}_r} + i\widehat{\mathbf{k}_i} \tag{14}$$

where $\boldsymbol{\theta}_{\text{KCNN}} = \{(\mathbf{W}_{\mathbf{F},\mathbf{r}}, \mathbf{b}_{\mathbf{F},\mathbf{r}}), (\mathbf{W}_{\mathbf{F},\mathbf{i}}, \mathbf{b}_{\mathbf{F},\mathbf{i}}), (\mathbf{W}_{\mathbf{I}_1}, \mathbf{b}_{\mathbf{I}_1}), \ldots,$

$(\mathbf{W}_{\mathbf{I}_N}, \mathbf{b}_{\mathbf{I}_N}), (\mathbf{W}_{\mathbf{R},\mathbf{r}}, \mathbf{b}_{\mathbf{R},\mathbf{r}}), (\mathbf{W}_{\mathbf{R},\mathbf{i}}, \mathbf{b}_{\mathbf{R},\mathbf{i}})\}$. The completed real and imaginary k-spaces, namely $\widehat{\mathbf{k}_r}$ and $\widehat{\mathbf{k}_i}$, are reconstructed by a 1×1 convolution ($\mathbf{W}_{\mathbf{R},\mathbf{r}}$, $\mathbf{b}_{\mathbf{R},\mathbf{r}}$, $\mathbf{W}_{\mathbf{R},\mathbf{i}}$, and $\mathbf{b}_{\mathbf{R},\mathbf{i}}$) layer from the output of the inference net, $\mathbf{I}_{\mathbf{N}}$. The output of KCNN, $H_{\text{KCNN}}(\mathbf{k}_{\mathbf{in}}; \boldsymbol{\theta}_{\text{KCNN}}) = \widehat{\mathbf{k}}_{\text{KCNN}}$, is obtained by combining the 2 k-space channels.

2.3 | Deep CNN for image restoration

Restoration of degraded structures and removal of remaining artifacts in the KCNN-reconstructed image involve constructing another deep CNN operating on the image domain. The ICNN is adapted from conventional deep CNNs that were developed for image restoration, including super-resolution[23,24,33] and compression-artifact removal.[28,34] The role of ICNN is solving the left term of Equation 4, which can be rewritten as

$$\underset{\boldsymbol{\theta}_x}{\arg\min} \|\mathbf{x} - H_x(\mathbf{x}_{\mathbf{in}}; \boldsymbol{\theta}_x)\|_2^2 = \underset{\boldsymbol{\theta}_{\text{ICNN}}}{\arg\min} \|\mathbf{x} - H_{\text{ICNN}}(\mathbf{x}_{\mathbf{in}}; \boldsymbol{\theta}_{\text{ICNN}})\|_2^2$$

$$\tag{15}$$

where $\mathbf{x}_{\mathbf{in}}$ is the input image to be restored; $\boldsymbol{\theta}_{\text{ICNN}}$ is the ICNN parameters; and H_{ICNN} is the CNN-based hypothesis function. The ICNN network architecture, which corresponds to H_{ICNN} of Equation 15, is presented in Figure 1B. Particularly, the ICNN consists of network components similar to those in KCNN (feature extraction, inference, and

reconstruction nets). Because MR images are complex-valued, we divided the input image, $\mathbf{x_{in}}$, into real and imaginary channels ($\mathbf{x_{in,r}}$ and $\mathbf{x_{in,i}}$, respectively) as in the KCNN. In the ICNN, a skip-connection layer is added to the reconstruction net.[23,24] The forward-pass equations for H_{ICNN} are as follows:

$$\mathbf{F_r} = \sigma(\mathbf{W_{F,r}} * \mathbf{x_{in,r}} + \mathbf{b_{F,r}}) \tag{16}$$

$$\mathbf{F_i} = \sigma(\mathbf{W_{F,i}} * \mathbf{x_{in,i}} + \mathbf{b_{F,i}}) \tag{17}$$

$$\mathbf{F} = [\mathbf{F_r}, \mathbf{F_i}] \tag{18}$$

$$\mathbf{I_1} = \sigma(\mathbf{W_{I_1}} * \mathbf{F} + \mathbf{b_{I_1}}) \tag{19}$$

$$\mathbf{I_n} = \sigma(\mathbf{W_{I_n}} * \mathbf{I_{n-1}} + \mathbf{b_{I_n}}) \tag{20}$$

$$\widehat{\mathbf{R_r}} = \mathbf{W_{R,r}} * \mathbf{I_{N_l}} + \mathbf{b_{R,r}} \tag{21}$$

$$\widehat{\mathbf{R_i}} = \mathbf{W_{R,i}} * \mathbf{I_{N_l}} + \mathbf{b_{R,i}} \tag{22}$$

$$\widehat{\mathbf{R}} = \widehat{\mathbf{R_r}} + i\widehat{\mathbf{R_i}} \tag{23}$$

$$H_{ICNN}(\mathbf{x_{in}}; \boldsymbol{\theta_{ICNN}}) = \widehat{\mathbf{x}}_{ICNN} = \mathbf{x_{in}} + \widehat{\mathbf{R}} \tag{24}$$

where $n = 2, \ldots, N_l - 1$, N_l denotes the order of the layers; $\widehat{\mathbf{R}}$ denotes the reconstructed residual image; $\widehat{\mathbf{x}}$ denotes the final output image; and $\boldsymbol{\theta_{ICNN}}$ of Equation 24 denotes the CNN parameters $\{(\mathbf{W_F}, \mathbf{b_F}), (\mathbf{W_{I_1}}, \mathbf{b_{I_1}}), \ldots, (\mathbf{W_{I_{N_l}}}, \mathbf{b_{I_{N_l}}}), (\mathbf{W_R}, \mathbf{b_R})\}$. The other variables represent the same network elements as those in the KCNN. The inference net infers the feature maps in which detailed features are restored, and artifacts are reduced by applying N_l convolution/activation layers to the extracted feature maps, $\mathbf{F} \in \mathbb{R}^{n_x \times n_y \times c}$. The reconstruction net predicts the residual image, $\mathbf{R} = \mathbf{x} - \mathbf{x_{in}}$, through a single convolution layer and forms the final reconstructed image, $\widehat{\mathbf{x}}_{ICNN}$, as in Equation 24. The ICNN uses the skip connection only to learn the sparse residual image, which results in faster and more effective training.[23,24]

2.4 | Interleaved data consistency

The originally sampled k-space data can change while passing through ICNN, because ICNN optimizes only the left term of Equation 4. To ensure data consistency, which is the right term of Equation 4, we have to optimize Equation 4 with fixed H_{ICNN} and $\boldsymbol{\theta_{ICNN}}$, which are obtained using ICNN. Then, the closed-form solution of Equation 4 is[14]

$$\widehat{\mathbf{k_D}}(k_x, k_y) = \begin{cases} \dfrac{\widehat{\mathbf{k}}_{ICNN}(k_x, k_y) + \lambda \mathbf{k_u}(k_x, k_y)}{1 + \lambda} & \text{if } U(k_x, k_y) = 1 \\ \widehat{\mathbf{k}}_{ICNN}(k_x, k_y) & \text{if } U(k_x, k_y) = 0 \end{cases} \tag{25}$$

where $\widehat{\mathbf{k}}_{ICNN} = \mathcal{F}_{2D}(\widehat{\mathbf{x}}_{ICNN})$, which is the FT of the ICNN output; $\widehat{\mathbf{k_D}}$ denotes the reconstructed k-space with data consistency; and k_x and k_y denote k-space indices. The final output of IDC is $\widehat{\mathbf{x_D}} = \mathcal{F}_{2D}^{-1}(\widehat{\mathbf{k_D}})$.
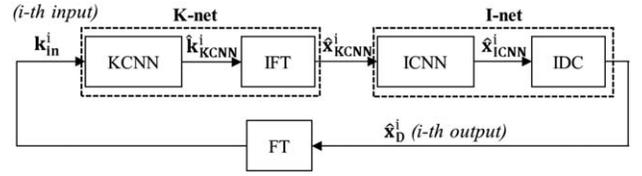
FIGURE 2 Block diagram for data flow and intermediate operations of cross-domain CNNs (CD-CNNs)

2.5 | Cross-domain CNNs

The 3 different network components defined previously (namely, KCNN and ICNN with IDC), which solve Equations 3 and 4, are iteratively applied to solve Equation 5. We named this iterative deep-learning approach as CD-CNN. The data flow and intermediate operations of the one block of CD-CNNs at the ith iteration are illustrated in Figure 2. As depicted in Figure 2, the one block of CD-CNNs consists of 2 parts: K-net (the combination of KCNN and IFT) and I-net (the combination of ICNN and IDC). The input k-space data at ith iteration, $\mathbf{k_{in}^i}$, passes through a network operating on the k-space (i.e., KCNN). The value of $\mathbf{k_{in}^1}$, the first input data, is $\mathbf{k_u}$, which is the undersampled k-space data; and $\mathbf{k_{in}^i}$ with i larger than 2 is the output at the previous (i.e., $i-1$) iteration. The output of KCNN, $\widehat{\mathbf{k}}_{KCNN}^i$, is inverse Fourier transformed to obtain $\widehat{\mathbf{x}}_{KCNN}^i$. We designate this KCNN+FT operation, which gives the first output image, as K-net. Then, $\widehat{\mathbf{x}}_{KCNN}^i$ is fed to the next network, ICNN, which yields $\widehat{\mathbf{x}}_{ICNN}^i$. At the final step of the iteration, $\widehat{\mathbf{x}}_{ICNN}^i$ is fed to IDC to yield $\widehat{\mathbf{x}}_D^i$, which is the ith iteration output of the CD-CNNs. The ICNN+IDC operation, which gives the second output image, is designated as I-net. The one-block CD-CNN procedure is iterated until the loss between the fully sampled image $\mathbf{x}$ and the final output image $\widehat{\mathbf{x}}_D^i$ is saturated.

Training was performed with an incremental manner rather than an end-to-end manner. The KIKI-net consists of CNNs of which layer depth is more than 100 and the interleaved operations among CNNs, FT, and IDC. Therefore, training the KIKI-net with an end-to-end manner is likely to involve problems of nonlocal minimum or overfitting and memory shortages because of the large number of the parameters to be learned (more than 3.5 million in the networks we used). To separately train each CNN while not involving operations of FT and IDC in training, only one last network (i.e., KCNN or ICNN) was trained while previously trained networks were fixed. For example, when training the one block of KIKI-net (i.e., KI-net), KCNN was trained first and the ICNN was trained after the KCNN training. More specifically, to train the first KCNN, all undersampled k-space data and their corresponding fully sampled k-space data were fed to KCNN as inputs and outputs. To train the next ICNN, IFTs of these KCNN outputs, which are the K-net outputs, and

their corresponding fully sampled images were fed to ICNN as inputs and outputs, respectively. The next block receives the outputs of the previous block (i.e., KI-net) as inputs, and was trained in the same manner as the previous block.

More details for training, including loss function and optimizer, as well as network specifications, including network depths, filter sizes, activation function for CNNs, parameters of stochastic gradient descent optimizer, deep-learning libraries, and training/testing times, are provided in the supporting information.

2.6 | Experimental framework

Three different MR data sets were used: T_2 fluid-attenuated inversion recovery (T_2-FLAIR) brain real-valued data set provided by the Alzheimer's Disease Neuroimaging Initiative (ADNI) [35] and 2 complex-valued data sets, T_2-FLAIR and T_1-weighted data set, which were acquired at our local institute. Details of data acquisition, including scanner information, sequence parameters, and the number of slices used for training/testing, are provided in the supporting information.

Undersampled k-space data were retrospectively obtained by subsampling the fully sampled k-space data. Before undersampling, all MR images were normalized to a maximum magnitude of 1. A Cartesian random undersampling scheme in a phase-encoding (i.e., anterior–posterior) direction was used for undersampled k-space data-set generation. Reduction factors were 2, 3, and 4. The binary undersampling masks are presented in Supporting Information Figure S1. Fully sampled images were used as label data during the training.

The proposed KIKI-net's reconstruction performance was compared with the following 7 conventional algorithms: baseline zero-filling, CS-MRI,[4] DL-MRI,[14] block-matching and 3D filtering (BM3D) MRI,[36] a CNN-based algorithm by Wang et al[19] (denoted as Wang's algorithm), PANO (patch-based nonlocal operator),[18] and FDLCP (fast dictionary learning method on classified patches).[16] The detailed parameters of the conventional algorithms are provided in the supporting information.

The reconstructed images were evaluated using 2 numerical metrics: peak SNR (PSNR) and structure similarity (SSIM).[37] Particularly, PSNR was calculated as the ratio in decibels (dB) of the peak intensity value of the reference image to the RMS error between the reconstructed and reference images. The SSIM, an image quality metric, was used to evaluate structure similarity and detailed features in the 2 images. The SSIM is known to be better correlated with the perception of the human visual system than PSNR.[37] The patch size used to calculate SSIM was 11.
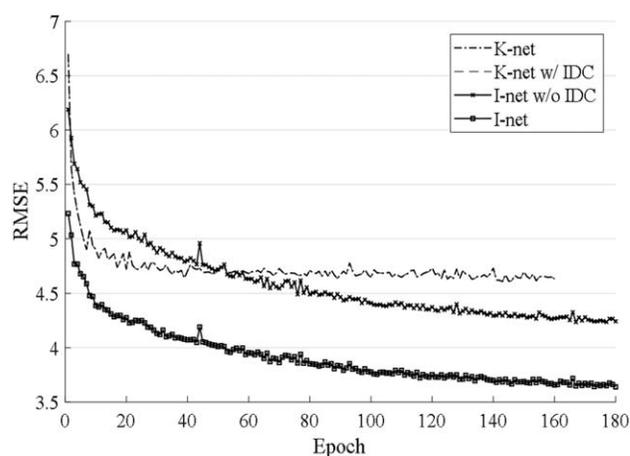
FIGURE 3 Root mean square error (RMSE) versus of epochs for K-net without interleaved data consistency (IDC), K-net with IDC, I-net without IDC, and I-net during training

3 | RESULTS

3.1 | K-net versus I-net

To evaluate the efficacy of each network component, we compared the results from K-net and I-net. Figure 3 depicts RMS error versus the number of epochs for 4 different networks with the same network capacity (i.e., K-net, K-net with IDC, I-net without IDC, and I-net). During training, both K-net and I-net stably converged before 180 epochs in terms of RMS error. The RMS error values of K-net and K-net with IDC were exactly the same, indicating that K-net performs its own data consistency. In comparison with single networks, I-net showed a performance superior to K-net. Moreover, the combination of IDC into ICNN resulted in greater improvements. However, Figure 4 shows that I-net was not always superior to K-net in all image areas. Figure 4 depicts the true fully sampled image (A), zero-filling image (B), and the images reconstructed with K-net (C), K-net with IDC (D), I-net without IDC (E), and I-net at $R = 4$ (F). Figure 4G-L and Figure 4M-R are the magnified images of solid and dotted boxes, respectively, found in Figure 4A-F. As shown in Figure 4C,D, K-net and K-net with IDC result in the same performance. In Figure 4I,J, K-nets successfully removed the oscillatory artifacts owing to undersampling, as indicated in Figure 4H. They also faintly restored the real structures in Figure 4G as depicted by the dotted circles in Figure 4I,J. Contrarily, the I-nets failed to remove the artifacts, and furthermore sharpened them as depicted by dotted circles in Figure 4K,L. Moreover, the recovered shapes from K-net depicted by dotted circles in Figure 4I,J are not observed in Figure 4K,L. In contrast, another result from I-net, the dotted circles in Figure 4Q,R indicate that I-nets are capable of restoring structures that could not be restored by K-net as shown in Figure 4O,P. In addition, I-nets were superior to K-nets in restoring detailed structures, because
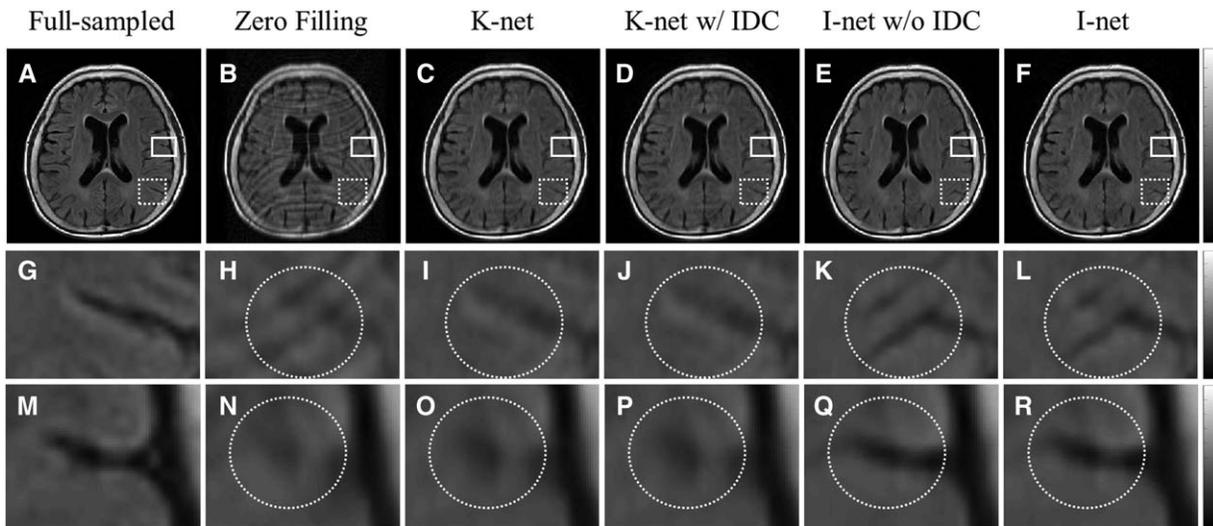
FIGURE 4 Reconstruction results of K-net and I-net at $R = 4$ undersampling: true fully sampled image (A); zero-filling image (B); and images reconstructed with K-net (C), K-net with IDC (D), I-net without IDC (E), and I-net (F). G-L, M-R, Magnified images of solid boxes and dotted boxes in (A) to (F), respectively

of their structure-sharpening characteristics. These results indicate that K-net and I-net have different advantages/disadvantages in terms of tissue-structure restoration.

3.2 | Single-domain CNNs versus CD-CNNs

To evaluate the efficacy of CD-CNNs, which are a combination of K-net and I-net, the CD-CNN's performance was compared with that of iterative CNNs operating on only a single domain. This kind of CNN was designated as single-domain CNNs (SD-CNNs). We compared the results of 2 SD-CNNs (IIII-net and KKKK-net) and 2 CD-CNNs (IKIK-net and KIKI-net) under the same conditions of the number of CNN iterations and network capacity. Figure 5 depicts the fully sampled image (A), the magnified fully sampled image (B), and the magnified images reconstructed with zero-filling

(C), IIII-net (D), KKKK-net (E), IKIK-net (F), and KIKI-net (G) at $R = 4$ for the boxed region of interest in Figure 5A. For the first example images, the CD-CNN images (Figure 5F1,G1) depict better reconstructions than the SD-CNN images (Figure 5D1,E1) in terms of restoring detailed tissues as shown in the ellipsoids in Figure 5D1-G1. IKIK-net and KIKI-net showed similar performances. In the case of the second example images, the IIII-net image depicts that high-frequency aliasing artifacts in Figure 5C2 are accentuated as a realistic structure, as shown in the circle in Figure 5D2. In contrast, the 2 CD-CNNs well removed the aliasing artifacts while not shaping the artifacts that look like realistic structures, as shown in Figure 5F2,G2.

The quantitative evaluations of the different I-net and K-net combinations (IIII-net, KKKK-net, IKIK-net, and KIKI-net) are listed in Table 1, which depicts the average PSNR/
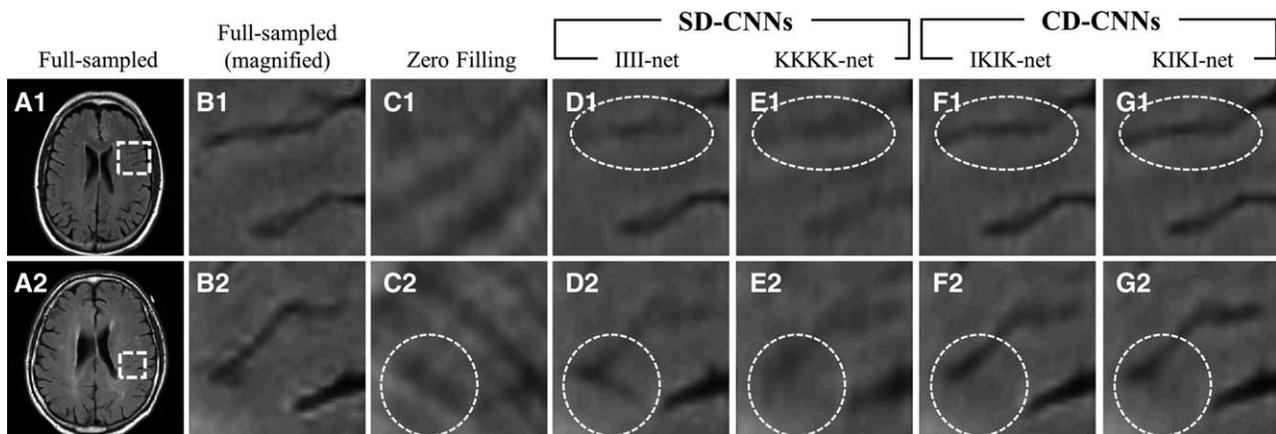
FIGURE 5 Reconstruction results from single-domain CNNs (SD-CNNs) and CD-CNNs at $R = 4$ undersampling: fully sampled image (A); magnified image of (A) of boxed region of interest (B); zero-filling image (C); and the image reconstructed with IIII-net (D), KKKK-net (E), IKIK-net (F), and KIKI-net (G)

TABLE 1 Average peak SNR/structure similarities of different combinations of networks on T_2-FLAIR_ADNI data sets undersampled at $R = 3$

Order of nets	First net PSNR/SSIM	Second net PSNR/SSIM	Third net PSNR/SSIM	Fourth net PSNR/SSIM
I-, I-, I-, I-	37.10/0.9716	38.12/0.9766	38.41/0.9782	38.43/0.9791
K-, K-, K-, K-	34.19/0.9636	34.84/0.9688	35.46/0.9688	35.46/0.9688
I-, K-, I-, K-	37.10/0.9716	38.98/0.9771	39.59/0.9809	39.58/0.9810
K-, I-, K-, I-	34.19/0.9636	38.64/0.9768	40.30/0.9821	**40.35/0.9833**

Note: PSNR, peak SNR; SSIM, structure similarity.
The highest PSNR and SSIM values are bold faced.

SSIMs of test images reconstructed with the 4 CNN combinations at $R = 3$. For the first network, I-net had a better performance than K-net. For the second network, the 2 CD-CNNs (KI-net and IK-net) exhibited a better performance than II-net and KK-net. For the third and fourth networks, the CD-CNNs consistently showed better performance than the other combinations that consisted of SD-CNNs (i.e., III-net, KKK-net, IIII-net, and KKKK-net). Therefore, CD-CNNs were more effective than SD-CNNs. The KIKI-net resulted in a slightly better performance than IKIK-net; however, the differences were not significant. The KIKI-net showed minor improvements compared with KI-net (1.71 dB in PSNR, 0.0065 in SSIM). Although the results were not presented in this study, no significant improvement was observed in the iteration after KIKI-net (e.g., KIKIK, KIKIKI). Therefore, we fixed the iteration of CD-CNNs as 2 (KI-net is 1 iteration of CD-CNNs).

3.3 | Comparison with conventional algorithms

Qualitative and quantitative comparisons are provided for the conventional algorithms and the proposed KIKI-net. Figure 6 depicts the resultant images from the ADNI data set (T_2-FLAIR_ADNI). In this figure, the true and reconstructed images (A1) with zero-filling (B1), CS-MRI (C1), DL-MRI (D1), BM3D-MRI (E1), Wang's algorithm (F1), PANO (G1), FDLCP (H1), and KIKI-net (I1) are shown. Figure 6A2-I2 and Figure 6A3-I3 are magnified images of the solid and dotted-boxed regions of interest in Figure 6A1-I1, respectively. The undersampling factor was 4. In the CS-MRI images (Figure 6C1), oscillatory undersampling artifacts noted in the zero-filling image in (Figure 6B1) still remained because of high acceleration, whereas the artifacts were sufficiently removed at $R = 2, 3$, which are depicted in Supporting Information Figure S6. Moreover, as depicted by the magnified image in Figure 6C2,C3, the algorithm failed to restore details observed in the fully sampled images. The DL-MRI, BM3D-MRI, and Wang's algorithm successfully reduced the undersampling artifacts despite the high reduction factor; however, some artifacts

are still observed in Figure 6D1-F1, and they failed to restore the details depicted in Figure 6D2-F2 and Figure 6D3-F3. Two state-of-the-art algorithms, PANO (Figure 6G1) and FDLCP (Figure 6H1), removed most of the aliasing artifacts, and were outstanding in restoring details. However, some details were not fully recovered as shown in the ellipsoids in Figure 6G2,H2,G3,H3. In contrast, KIKI-net successfully restored those details as shown in the ellipsoids in Figure 6I2,I3.

Figures 7 and 8 depict the true and reconstructed images (A) with zero-filling (B), CS-MRI (C), DL-MRI (D), Wang's algorithm (E), PANO (F), FDLCP (G), and KIKI-net (H) for the 2 sets of complex-valued k-space data acquired at our local institute (T_2-FLAIR_ours for Figure 7 and T_1-weighted_ours for Figure 8). Figure 7I-P shows magnified images of the boxed regions of interest in Figure 7A-H, and Figure 7Q-W shows the error maps of Figure 7I-P. The undersampling factors were 3 and 4 for T_2-FLAIR_ours and T_1-weighted_ours, respectively. Because the BM3D-MRI algorithm does not support complex-valued data,[21] the results for this algorithm are not included in these figures. As depicted in these Figures, KIKI-net outperforms the other algorithms for the complex-valued data sets, which shows similar aspects to the results in Figure 6. In Figure 7P, KIKI-net shows better performance than the other algorithms (Figure 7J-O) in improving conspicuity of thin vessels and boundaries between gray and white matter. Figure 8P shows how KIKI-net can restore details that were severely blurred and distorted in the highly undersampled image (Figure 8J) and not be fully restored in the images reconstructed with other algorithms (Figure 8K-O). However, even though KIKI-net best restored the details as close to the fully sampled image as possible compared with the others, some of detailed structures such as vessels were still blurred and distorted in the shape, as shown in the center of Figure 8P, because of the high undersampling factor, which appears to be a limit of our algorithm at $R \geq 4$ in Cartesian undersampling for the complex-valued data. In the case of the other reduction factors, results for the 3 data sets are provided in Supporting Information Figures S6, S7, and S8. Reconstructed phase images from different reconstruction methods for the 2
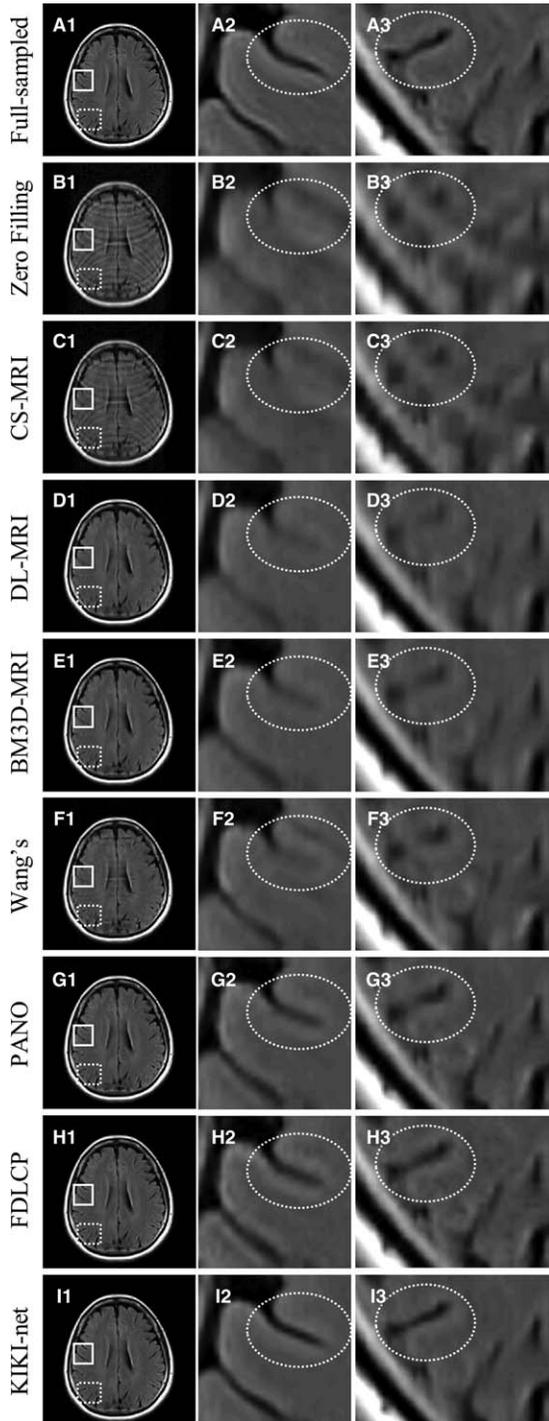
FIGURE 6 Reconstruction results from conventional algorithms and KIKI-net at $R = 4$ undersampling for the T_2 fluid-attenuated inversion recovery Alzheimer's Disease Neuroimaging Initiative data set (T_2-FLAIR_-ADNI): fully sampled image (A1); zero-filling image (B1); image reconstructed with compressed-sensing MRI (CS-MRI) (C1); image reconstructed with dictionary learning (DL) MRI (D1); image reconstructed with block-matching and 3D filtering (BM3D) MRI (E1); image reconstructed with Wang's algorithm (F1); image reconstructed with PANO (patch-based nonlocal operator) (G1); image reconstructed with FDLCP (fast dictionary learning method on classified patches) (H1); and image reconstructed with our proposed algorithm, KIKI-net (I1). A2-I2, A3-I3, Magnified images for the solid boxes and dotted boxes in (A1) to (I1), respectively

complex-valued data sets are also provided in Supporting Information Figures S9 and S10.

The quantitative evaluations are summarized in Table 2 along with the average PSNR/SSIM values at $R = 2$, 3, and 4 on the 3 data sets (T_2-FLAIR_ADNI, T_2-FLAIR_-ours, and T_1-weighted_ours). Among the 7 conventional algorithms, PANO and FDLCP exhibited high PSNRs and SSIMs for the 3 data sets. The PANO algorithm exhibited the highest PSNR for T_2-FLAIR_ADNI and T_1_weighted_ours data sets at $R = 2$, and the highest SSIM for the T_1_weighted_ours data set at $R = 2$, which are slightly better results than KIKI-net (0.34 dB in PSNR and 0.0045 in SSIM). Except for these 2 cases, our proposed algorithm, KIKI-net, exhibited the highest values with mean improvements of 2.29 dB in PSNR and 0.031 in SSIM compared with the highest values of the other algorithms.

4 | DISCUSSION

The current study presents a novel CD-CNN, referred to as KIKI-net for reconstructing undersampled MR images. The KIKI-net consists of the following 3 components: KCNN for k-space completion, ICNN for removing artifacts and restoring image details, and IDC for regularizing and activating network learning. Additionally, to iteratively perform each optimization, each network is alternately applied and independently trained. Experimental results with various reduction factors and 3 different data sets demonstrated that KIKI-net outperforms conventional reconstruction algorithms including CS-MRI, DL-MRI, BM3D-MRI, a CNN-based algorithm that operates only on the image domain, PANO, and FDLCP. KIKI-net is outstanding in terms of restoring detailed tissue structures that would disappear in other algorithms as well as in simultaneously removing aliasing artifacts.

To effectively solve the 2 minimization equations for k-space completion and image restoration (Equations 3 and 4), K-net (KCNN with IFT) and I-net (ICNN with IDC) were alternately applied such that each term was iteratively minimized. The number of layers linearly increases in CD-CNNs with the number CD-CNN iterations. For example, the KIKI-net used in this study, which combines 4 25-layer CNNs, corresponds to a CNN that consists of 100 layers (25 layers $\times$ 4); however, if this extremely deep KIKI-net is trained in an end-to-end manner, there are several factors that degrade the training performance, such as the difficulty of hyperparameter tuning,[38-40] the vanishing gradient problem[41,42] due to a larger number of parameters, and graphic processing unit memory shortage. Furthermore, FT and IFT must be included in the networks, which increases the computational complexity of the network and worsen these
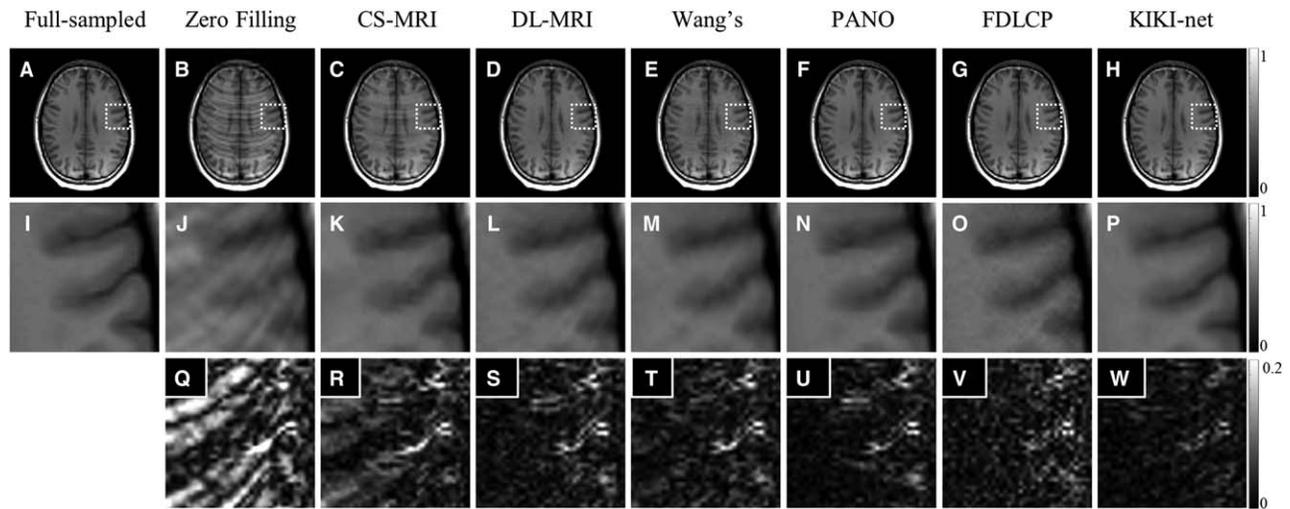
FIGURE 7 Reconstruction results from conventional algorithms and KIKI-net at $R = 3$ undersampling for the T_1-weighted_ours data set: fully sampled image (A); zero-filling image (B); image reconstructed with CS-MRI (C); image reconstructed with DL-MRI (D); image reconstructed with Wang's algorithm (E); image reconstructed with PANO (F); image reconstructed with FDLCP (G); and image reconstructed with our proposed algorithm, KIKI-net (H). I-P, Magnified images for the dotted boxes in (A) to (H), respectively. Q-W, Error maps of (I) to (P)

problems. To overcome these issues, we trained each 25-layer network independently. The FT used in IDC was not performed during training but was used only to convert network outputs to the next network inputs.

We found that K-net and I-net are different in their tissue-structure restoring roles. K-net is effective in restoring structures that have disappeared because of undersampling and removing high-frequency oscillatory artifacts. However, K-net shows low performance in improving the clarity of detailed tissues. I-net is effective in improving the structure sharpness and clarity. However, it can make the artifacts sharp; and to more realistic structures, it mistakes the oscillatory artifacts on images as real structures.

When the 2 networks are combined, the disadvantages of the individual networks are effectively complemented. The deeper network depth of the proposed (25-layer $\times$ 4) CD-CNNs than that of existing CNN-based reconstruction algorithms was also an important factor that improves the reconstruction performance. In our experiments on network depths and the number of filters of each layer, performance continued to improve until the network capacity reached the current CD-CNN capacity (Supporting Information Figures S4 and S5).

Although KIKI-net performed well in the 3 data sets presented in this study, there was a large performance drop in results for data sets at our local institute compared with
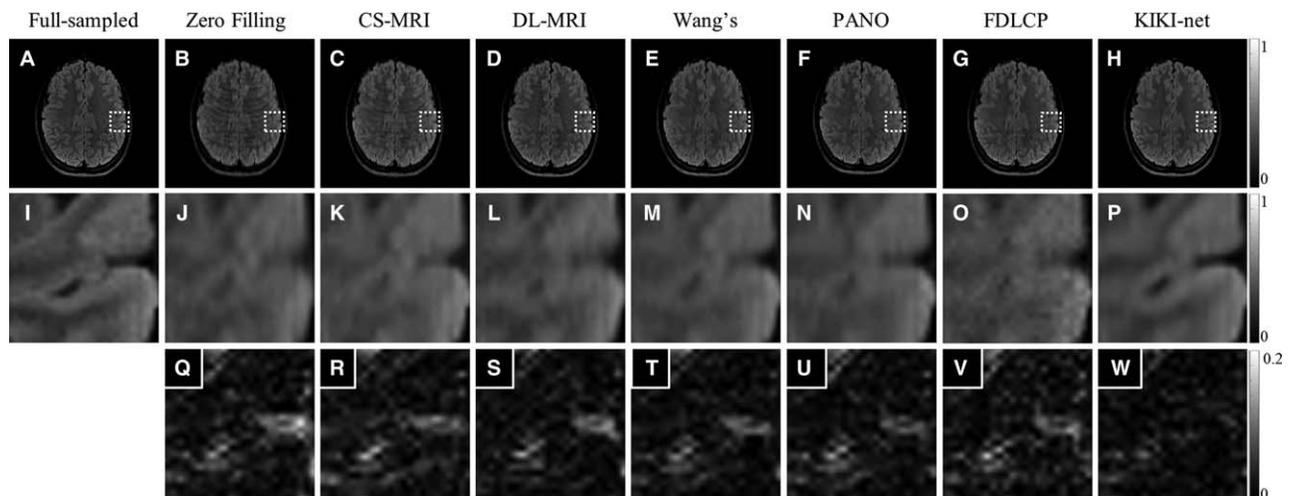
FIGURE 8 Reconstruction results from conventional algorithms and KIKI-net at $R = 4$ undersampling for the T_2-FLAIR_ours data set: fully sampled image (A); zero-filling image (B); image reconstructed with CS-MRI (C); image reconstructed with DL-MRI (D); image reconstructed with Wang's algorithm (E); image reconstructed with PANO (F); image reconstructed with FDLCP (G); and image reconstructed with our proposed algorithm, KIKI-net (H). I-P, Magnified images for the dotted boxes in (A) to (H), respectively. Q-W, Error maps of (I) to (P)

TABLE 2 Quantitative results of conventional algorithms and KIKI-net; average peak SNR/structural similarities for $R = 2$, 3, and 4 on the 3 data sets

		Data set		
		T₂-FLAIR_ADNI	T₁-weighted_ours	T₂-FLAIR_ours
R	Algorithm	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
2	Zero filling	30.49/0.9359	34.08/0.9531	27.50/0.8823
	CS-MRI	36.06/0.9806	41.68/0.9798	30.83/0.9177
	DL-MRI	36.11/0.9814	39.82/0.9776	29.98/0.9029
	BM3D-MRI	41.65/0.9814	—	—
	Wang's	41.20/0.9804	41.49/0.9785	30.39/0.9148
	PANO	**45.83**/0.9856	**43.09/0.9836**	32.99/0.9218
	FDLCP	44.37/0.9845	41.71/0.9752	32.03/0.9166
	KIKI-net	45.49/**0.9901**	43.07 /0.9830	**35.32/0.9655**
3	Zero filling	28.98/0.8913	29.58/0.9223	26.39/0.8566
	CS-MRI	31.36/0.9514	34.17/0.9598	27.98/0.8836
	DL-MRI	33.03/0.9550	36.53/0.9708	28.81/0.8878
	BM3D-MRI	37.11/0.9677	—	—
	Wang's	36.52/0.9623	38.44/0.9684	28.85/0.8865
	PANO	38.17/0.9724	38.54/0.9695	30.59/0.9020
	FDLCP	38.37/0.9744	38.25/0.9612	30.64/0.9068
	KIKI-net	**40.35/0.9833**	**39.88/0.9794**	**33.29/0.9479**
4	Zero filling	26.74/0.8572	24.43/0.8749	23.62/0.8191
	CS-MRI	27.54/0.9269	27.91/0.9165	25.01/0.8477
	DL-MRI	32.06/0.9433	34.99/0.9593	25.62/0.8487
	BM3D-MRI	34.62/0.9515	—	—
	Wang's	33.84/0.9419	35.82/0.9577	24.84/0.8436
	PANO	34.94/0.9554	36.04/0.9603	27.16/0.8803
	FDLCP	35.28/0.9601	36.02/0.9586	26.78/0.8720
	KIKI-net	**36.50/0.9669**	**37.08/0.9720**	**31.60/0.9315**

Note: PSNR, peak SNR; SSIM, structural similarity.
The highest PSNR and SSIM values are bold faced.

the results for the ADNI data set. The similar performance drops were also observed for the conventional algorithms. The performance drops appeared to be caused by relatively high noise levels of the data sets at our local institute. The CNNs could not learn to predict the true noiseless images, because the label data for training were also noisy. In our additional experiments on different noise levels for the ADNI data set (Supporting Information Figures S2 and S3), it was observed that as the noise level increased, the output images became more blurred in order to reduce the errors caused by random noise, not undersampling. Consequently, the blurring resulted in the performance drop (lower PSNR).

Future experiments are required to extend KIKI-net applicability. The present study focused on variable-density Cartesian trajectory MR acquisition; however, a similar concept of completing k-space with K-net and I-net can also be applied with appropriate modifications of KIKI-net with respect to non-Cartesian acquisition, including radial-trajectory and spiral-trajectory acquisition. Moreover, to achieve higher acceleration, KIKI-net can be combined with parallel imaging by an appropriate modification of KIKI-net, such that it can fully use multicoil data acquired with undersampling.

ORCID

Jinseong Jang http://orcid.org/0000-0002-0042-9304
Ho-Joon Lee http://orcid.org/0000-0003-0831-6184

REFERENCES

[1] Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: sensitivity encoding for fast MRI. *Magn Reson Med.* 1999;42:952-962.

[2] Griswold MA, Jakob PM, Heidemann RM, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med.* 2002;47:1202-1210.

[3] Stehling MK, Turner R, Mansfield P. Echo-planar imaging: magnetic resonance imaging in a fraction of a second. *Science.* 1991;254:43-50.

[4] Lustig M, Donoho D, Pauly JM. Sparse MRI: the application of compressed sensing for rapid MR imaging. *Magn Reson Med.* 2007;58:1182-1195.

[5] Haldar JP, Hernando D, Liang Z-P. Compressed-sensing MRI with random encoding. *IEEE Trans Med Imaging.* 2011;30:893-903.

[6] Qu X, Cao X, Guo D, Hu C, Chen Z. Combined sparsifying transforms for compressed sensing MRI. *Electron Lett.* 2010;46:121-123.

[7] Qu X, Zhang W, Guo D, Cai C, Cai S, Chen Z. Iterative thresholding compressed sensing MRI based on contourlet transform. *Inverse Probl Sci Eng.* 2010;18:737-758.

[8] Qu X, Guo D, Chen Z, Cai C. Compressed sensing MRI based on nonsubsampled contourlet transform. In Proceedings of IEEE International Symposium on IT in Medicine and Education, Beijing, China, 2008. pp. 693-696.

[9] Liang D, Liu B, Wang J, Ying L. Accelerating SENSE using compressed sensing. *Magn Reson Med.* 2009;62:1574-1584.

[10] Murphy M, Alley M, Demmel J, Keutzer K, Vasanawala S, Lustig M. Fast l1-SPIRiT compressed sensing parallel imaging MRI: scalable parallel implementation and clinically feasible runtime. *IEEE Trans Med Imaging.* 2012;31:1250-1262.

[11] Otazo R, Kim D, Axel L, Sodickson DK. Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI. *Magn Reson Med.* 2010;64:767-776.

[12] Majumdar A. Improving synthesis and analysis prior blind compressed sensing with low-rank constraints for dynamic MRI reconstruction. *Magn Reson Imaging.* 2015;33:174-179.

[13] Otazo R, Candès E, Sodickson DK. Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. *Magn Reson Med.* 2015;73:1125-1136.

[14] Ravishankar S, Bresler Y. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans Med Imaging.* 2011;30:1028-1041.

[15] Ravishankar S, Bresler Y. Efficient blind compressed sensing using sparsifying transforms with convergence guarantees and application to magnetic resonance imaging. *SIAM J Imaging Sci.* 2015;8:2519-2557.

[16] Zhan Z, Cai J-F, Guo D, Liu Y, Chen Z, Qu X. Fast multiclass dictionaries learning with geometrical directions in MRI reconstruction. *IEEE Trans Biomed Eng.* 2016;63:1850-1861.

[17] Akçakaya M, Basha TA, Goddu B, t al. Low-dimensional-structure self-learning and thresholding: regularization beyond compressed sensing for MRI Reconstruction. *Magn Reson Med.* 2011;66:756-767.

[18] Qu X, Hou Y, Lam F, Guo D, Zhong J, Chen Z. Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator. *Med Image Anal.* 2014;18:843-856.

[19] Wang S, Su Z, Ying L, et al. Accelerating magnetic resonance imaging via deep learning. In Proceedings of the IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, 2016. pp. 514-517.

[20] Lee D, Yoo J, Ye JC. Deep residual learning for compressed sensing MRI. In Proceedings of IEEE 14th International Symposium on Biomedical Imaging (ISBI), Melbourne, Australia, 2017. pp. 15-18.

[21] Yang Y, Sun J, Li H, Xu Z. ADMM-Net: a deep learning approach for compressive sensing MRI. arXiv preprint, 2017. arXiv:1705.06869.

[22] Jin KH, McCann MT, Froustey E, Unser M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans Image Process.* 2017;26:4509-4522.

[23] Kim J, Kwon Lee J, Mu Lee K. Accurate image super-resolution using very deep convolutional networks. In Proceedings the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, 2016. pp. 1646-1654.

[24] Kim J, Kwon Lee J, Mu Lee K. Deeply-recursive convolutional network for image super-resolution. In Proceedings the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, 2016. pp. 1637-1645.

[25] Xie J, Xu L, Chen E. Image denoising and inpainting with deep neural networks. In Proceedings of Neural Information and Processing Systems, Lake Tahoe, NV, 2012. pp. 341-349.

[26] Jain V, Seung S. Natural image denoising with convolutional networks. In Proceedings of Neural Information and Processing Systems, Vancouver, Canada, 2009. pp. 769-776.

[27] Xu L, Ren JS, Liu C, Jia J. Deep convolutional neural network for image deconvolution. In Proceedings of Neural Information and Processing Systems, Montréal, Canada, 2014. pp. 1790-1798.

[28] Dong C, Deng Y, Change Loy C, Tang X. Compression artifacts reduction by a deep convolutional network. In Proceedings of IEEE International Conference of Computer Vision (ICCV), Chile, 2015. pp. 576-584.

[29] Iizuka S, Simo–Serra E, Ishikawa H. Let there be color! Joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification. *ACM Trans Graph.* 2016;35:110.

[30] Wang L, Huang Z, Gong Y, Pan C. Ensemble based deep networks for image super-resolution. *Pattern Recogn.* 2017;68:191-198.

[31] Park H, Lee KM. Look wider to match image patches with convolutional neural networks. *IEEE Signal Process Lett.* 2017;24:1788-1792.

[32] Lyu G, Yin H, Yu X, Luo S. A local characteristic image restoration based on convolutional neural network. *IEICE Trans Inf Syst.* 2016;99:2190-2193.

[33] Dong C, Loy CC, He K, Tang X. Image super-resolution using deep convolutional networks. *IEEE Trans Pattern Anal Mach Intell.* 2016;38:295-307.

[34] Anagha R, Kavya B, Namratha M, Singasani C, Hamsa J. A machine learning approach for removal of JPEG compression artifacts: a survey. *Int J Comput Appl.* 2016;138:24–28.

[35] Jack CR, Bernstein MA, Fox NC, et al. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J Magn Reson Imaging.* 2008;27:685-691.

[36] Eksioglu EM. Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3D–MRI. *J Math Imaging Vision.* 2016;56:430-440.

[37] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process.* 2004;13:600-612.

[38] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015; 521:436-444.

[39] Sadeghi Z, Testolin A. Learning representation hierarchies by sharing visual features: a computational investigation of Persian character recognition with unsupervised deep learning. *Cogn Process.* 2017;18:273-284.

[40] Drokin I. Hybrid pre training algorithm of Deep Neural Networks. *ITM Web of Conferences.* 2016;6:02007.

[41] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans Neural Netw.* 1994;5:157-166.

[42] Hochreiter S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *Int J Uncertain Fuzz Knowledge-Based Syst.* 1998;6:107-116.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the supporting information tab for this article.

FIGURE S1 Undersampling masks used in the present study. Masks at $R = 2$ (A), $R = 3$ (B), and $R = 4$ (C) for 256×256 images are shown

FIGURE S2 Reconstruction results of KIKI-net with varying noise conditions of training/testing data. From left to right, fully sampled images, noisy fully sampled (label) images, noisy zero-filling (input) images at $R = 3$, and the output images of KIKI-net are shown. From top to bottom: sigma values of Gaussian noise are 0, 0.01, and 0.04

FIGURE S3 Mean peak SNR (PSNR) versus sigma values of white Gaussian noise for 3 different images. The PSNR of noisy zero-filling (input) images at $R = 3$ (blue Xs), noisy fully sampled images (red stars), and the output images of KIKI-net (yellow circles) are shown

FIGURE S4 Reconstruction results of KIKI-net versus the network depths: fully sampled image (A), zero-filling image (B), image reconstructed with KIKI-net of 20-layer (5 layers × 4) (C), 40-layer (10 layers × 4) (D), 60-layer (15 layers × 4) (E), 80-layer (20 layers × 4) (F), and 100-layer (25 layers × 4) (G). The PSNRs are represented at the bottom right of each figure. H-N, Magnified images of boxed regions of interest in (A) to (F)

FIGURE S5 Reconstruction results of KIKI-net versus the number of channels for each filter: fully sampled image (A), zero-filling image (B), image reconstructed with KIKI-net of 8 filters (C), 16 filters (D), 32 filters (E), and 64 filters (F). The PSNRs are presented at the bottom right of each figure. G-L, Magnified images of boxed regions of interest in (A) to (F)

FIGURE S6 Reconstruction results from conventional algorithms and KIKI-net at $R = 2$ undersampling (A1-Z1) and $R = 3$ undersampling (A2-Z2) for the T_2-FLAIR_-ADNI data set: fully sampled image (A); zero-filling image (B); image reconstructed with CS-MRI (C); image reconstructed with DL-MRI (D); image reconstructed with BM3D-MRI (E); image reconstructed with Wang's algorithm (F); image reconstructed with PANO (G); image reconstructed with FDLCP (H); and image reconstructed with our proposed algorithm, KIKI-net (I). J-R, Magnified images for the dotted boxes in (A) to (I), respectively. S-Z, Error maps of (J) to (R)

FIGURE S7 Reconstruction results from conventional algorithms and KIKI-net at $R = 2$ undersampling (A1-W1) and $R = 4$ undersampling (A2-W2) for the T_1-weighted_ours data set: fully sampled image (A); zero-filling images (B); image reconstructed with CS-MRI (C); image reconstructed with DL-MRI (D); image reconstructed with Wang's algorithm (E); image reconstructed with PANO (F); image reconstructed with FDLCP (G); and image reconstructed with our proposed algorithm, KIKI-net (H). I-P, Magnified images for the dotted boxes in (A) to (H), respectively. Q-W, Error maps of (I) to (P)

FIGURE S8 Reconstruction results from conventional algorithms and KIKI-net at $R = 2$ undersampling (A1-W1) and $R = 3$ undersampling (A2-W2) for the T_2-FLAIR_ours data set: fully sampled image (A); zero-filling image (B); image reconstructed with CS-MRI (C); image reconstructed with DL-MRI (D); image reconstructed with Wang's algorithm (E); image reconstructed with PANO (F); image reconstructed with FDLCP (G); and image reconstructed with our proposed algorithm, KIKI-net (H). I-P, Magnified images for the dotted boxes in (A) to (H), respectively. Q-W, Error maps of (I) to (P)

FIGURE S9 Reconstructed phase images from conventional algorithms and KIKI-net at $R = 2$ undersampling (A1-H1), $R = 3$ undersampling (A2-H2), and $R = 4$ undersampling (A3-H3) for the T_1-weighted_ours data set: fully sampled image (A); zero-filling image (B); image reconstructed with CS-MRI (C); image reconstructed with DL-MRI (D); image reconstructed with Wang's algorithm (E); image reconstructed with PANO (F); image reconstructed with FDLCP (G); and image reconstructed with our proposed algorithm, KIKI-net (H). The RMS error values are presented in the right bottom of the corresponding images

FIGURE S10 Reconstructed phase images from conventional algorithms and KIKI-net at $R = 2$ undersampling (A1-H1), $R = 3$ undersampling (A2-H2), and $R = 4$ undersampling (A3-H3) for the T_2-FLAIR_ours data set: fully sampled image (A); zero-filling image (B); image reconstructed with CS-MRI (C); image reconstructed with DL-MRI (D); image reconstructed with Wang's algorithm (E); image reconstructed with PANO (F); image reconstructed with FDLCP (G); and image reconstructed with our proposed algorithm, KIKI-net (H). The RMS error values are presented in the right bottom of the corresponding images