
MAJOR PAPER

Reconstruction of Compressed-sensing MR Imaging Using Deep Residual Learning in the Image Domain

Shohei Ouchi* and Satoshi Ito

Purpose: A deep residual learning convolutional neural network (DRL-CNN) was applied to improve image quality and speed up the reconstruction of compressed sensing magnetic resonance imaging. The reconstruction performances of the proposed method was compared with iterative reconstruction methods.

Methods: The proposed method adopted a DRL-CNN to learn the residual component between the input and output images (i.e., aliasing artifacts) for image reconstruction. The CNN-based reconstruction was compared with iterative reconstruction methods. To clarify the reconstruction performance of the proposed method, reconstruction experiments using 1D-, 2D-random under-sampling and sampling patterns that mix random and non-random under-sampling were executed. The peak-signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) were examined for various numbers of training images, sampling rates, and numbers of training epochs.

Results: The experimental results demonstrated that reconstruction time is drastically reduced to 0.022 s per image compared with that for conventional iterative reconstruction. The PSNR and SSIM were improved as the coherence of the sampling pattern increases. These results indicate that a deep CNN can learn coherent artifacts and is effective especially for cases where the randomness of k -space sampling is rather low. Simulation studies showed that variable density non-random under-sampling was a promising sampling pattern in 1D-random under-sampling of 2D image acquisition.

Conclusion: A DRL-CNN can recognize and predict aliasing artifacts with low incoherence. It was demonstrated that reconstruction time is significantly reduced and the improvement in the PSNR and SSIM is higher in 1D-random under-sampling than in 2D. The requirement of incoherence for aliasing artifacts is different from that for iterative reconstruction.

Keywords: *compressed sensing, reconstruction, deep learning*

Introduction

MRI and X-ray computed tomography are commonly used medical diagnostic tools. MRI scanners produce high-quality images without the damaging ionizing radiation of

X-rays. However, the scan time of MRI is much longer than that of X-ray computed tomography. Numerous techniques have thus been proposed to speed up MRI scan time. The application of compressed sensing (CS)^{1,2} to MRI image acquisition (CS-MRI)³ has attracted much interest in recent years. CS is a signal recovery theory, proposed by Donoho et al.,¹ that allows image reconstruction from fewer sampled signals than those required by the Nyquist–Shannon sampling theorem. With the application of CS, scan time can be shortened by changing the imaging sequence and algorithm for image reconstruction. CS greatly differs from conventional rapid imaging methods such as parallel imaging^{4,5} in that it does not require additional hardware for MRI (parallel imaging requires multiple coils and a signal receiving system). CS-MRI has the problems of image quality degradation due to the assumption of sparsity in the image space, insufficient randomness of measurement matrix and noises

Department of Innovation Systems Engineering, Graduate School of Engineering, Utsunomiya University, Tochigi, Japan

*Corresponding author: Department of Innovation Systems Engineering, Doctoral Degree Program, Graduate School of Engineering, Utsunomiya University, 7-1-2, Yoto, Utsunomiya, Tochigi 321-8585, Japan. Phone/Fax: +81-28-689-6276, E-mail: ultuchiboy@gmail.com

*Related work was presented and one-page abstract was published in Japanese at The 46th Annual Meeting of the Japanese Society for Magnetic Resonance in Medicine held on September 7th - 9th, 2018 in Ishikawa, Japan.

©2020 Japanese Society for Magnetic Resonance in Medicine
This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives International License.

Received: April 23, 2019 | Accepted: April 25, 2020

superimposed on the images. An iterative reconstruction method is used to solve the minimization problem. The reconstruction time required per image is relatively long compared with that for standard Fourier-transform-based methods. There is a great demand for fast reconstruction, because up to several hundred images are taken in diagnostic imaging.

Deep learning has received much attention because of its excellent performance in speech and image recognition.^{6,7} Deep learning has been applied to medical image processing, including anatomic classification,⁸ super-resolution,⁹ and MRI image reconstruction.^{10–14} It takes a lot of time to train a convolutional neural network (CNN), because deep learning involves the computation of a function that maps some inputs to their corresponding outputs using a huge number of training images. However, once learning has been completed, rapid reconstruction with no iterative processing can be achieved.

The first application of CNN to MRI was the de-aliasing of alias-superimposed images obtained in parallel imaging, which is classified as image domain learning.¹¹ They applied a multi-layer perceptron to reduce aliasing artifacts generated by under-sampling in k -space. Wang et al.¹⁰ used an image domain CNN for reconstruction from a randomly under-sampled signal; the images obtained from the CNN are used as a constrained reconstruction model in conventional CS iterative reconstruction. Lee et al.¹² proposed a deep learning network for the reconstruction of MRI images in which a multi-scale network structure called U-Net is utilized to cope with globally distributed artifact patterns and phase image reconstruction. Hyun et al.¹³ applied a CNN for image reconstruction using a uniformly subsampled signal and showed that high-quality images without remarkable artifacts can be obtained. As an alternative to image domain learning, k -space learning has been proposed, with methods such as automated transform by manifold approximation (AUTOMAP), in which the transformation from the source signal (k -space signal) to the target image domain can be achieved via data-driven supervised learning.¹⁴ Images can be reconstructed directly from an under-sampled k -space signal with AUTOMAP. However, the required number of parameters scales quadratically with input size, limiting practical applications. Therefore, training in the image domain is practical in the sense that fewer parameters make it easier to train the network and the network is less prone to overfitting.¹⁴

We applied deep residual learning (DRL) in image domain learning using a CNN to the image reconstruction problem using an under-sampled MRI signal. Zhang et al.¹⁵ proposed a CNN in which DRL is adopted for removing noise from noisy images. DRL is integrated to speed up the training process and enhance denoising. Inspired by Zhang et al.'s network, we propose a CNN reconstruction method that uses DRL with a modified version of Zhang et al.'s network. Our method predicts aliased image components caused by under-sampling using patch-based learning and then obtains an alias-free image by subtracting the predicted alias components from the input alias-superimposed image.

A CNN has been previously applied to learning artifacts, with a large receptive field used to cover all artifacts.¹² Patch-based learning can detect local features in an image; Zhang et al.'s network shows great image denoising performance.

Since reconstruction process of DRL-CNN is different with conventional iterative reconstruction methods; i.e., sparsified transform functions are not used in DRL-CNN or non-iterative process, therefore, reconstruction performance may differ in both methods. In this study, the performance comparison and evaluation of CS image reconstruction between DRL-CNN and iterative reconstruction methods were investigated using several under-sampling patterns. To clarify the reconstruction performance of DRL-CNN, reconstruction experiments using 1D-, 2D-random under-sampling and sampling patterns that mix random and non-random under-sampling were executed using DRL-CNN and iterative methods, such as the iterative soft thresholding algorithm (ISTA),¹⁶ the Split Bregman method,¹⁷ and the alternating direction method of multipliers (ADMM).¹⁸ To the best of our knowledge, reconstruction experiments using a mixture of random and non-random under-sampling pattern have not been executed. These examinations reveal the characteristics of the DRL-CNN and suggest the appropriate under-sampling patterns.

Materials and Methods

As mentioned in before, DRL-CNN was adopted in proposed method. First, the characteristics of the DRL-CNN and the details of the DRL-CNN reconstruction is presented in *Network structure of DRL-CNN* section. Next, the mathematical explanation of image reconstruction using the output of CNN is described in *Image reconstruction in DRL-CNN* section. Finally, iterative reconstruction methods compared with proposed method is explained in *Comparison with iterative reconstruction methods* section.

Network structure of DRL-CNN

The proposed DRL-CNN was designed to predict the residual image, i.e., the difference between the alias-superimposed image and the ground truth image, which is different from a typical CNN that predict the de-aliased image directly. According to He et al.,¹⁹ residual mapping is much easier for a CNN to learn than unreferenced mapping. Zhang et al. adopted residual learning together with batch normalization in a deep CNN and applied it to the image denoising problem. They quantitatively and qualitatively showed that this method has good image denoising performance. Our proposed CNN was a modified version of Zhang et al.'s denoising CNN. We used a leaky rectified linear unit (ReLU)²⁰ instead of a ReLU. A leaky ReLU provides the slope of the negative part of the function as an argument, which makes it possible to perform the back-propagation of a CNN and is expected to improve learning.

Figure 1 shows the network used in this study. The input layer has 64 filters of size $3 \times 3 \times 1$ as a convolutional (Conv) layer and a leaky ReLU. The middle layers have a Conv layer

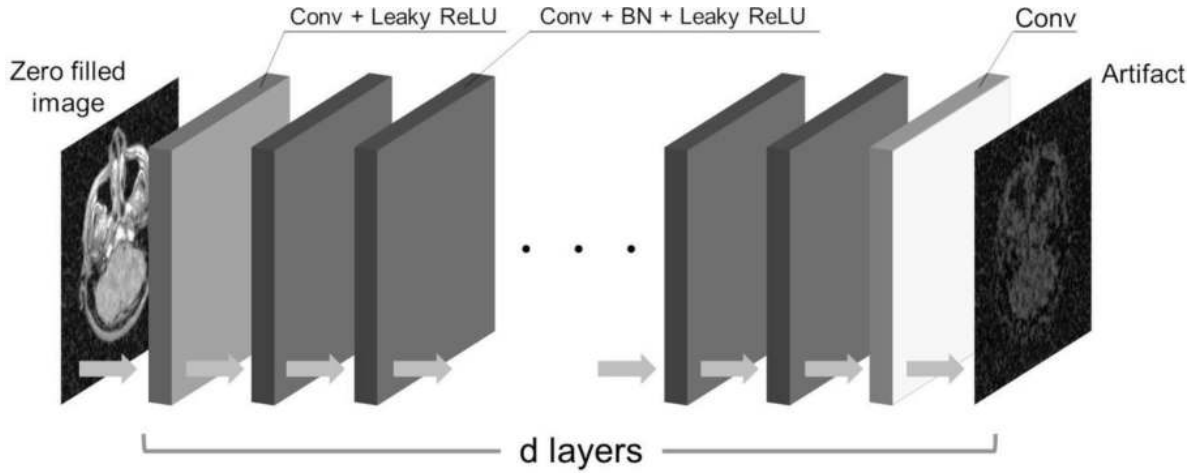


Fig. 1 The network structure of DRL-CNN. The input layer has 64 filters of size $3 \times 3 \times 1$ by Conv and the Leaky ReLU. The intermediate layers have Conv with 64 filters of size $3 \times 3 \times 64$, batch normalization, Leaky ReLU. The output layer has Conv with a filter of size $3 \times 3 \times 64$. Proposed method predicts aliased image components caused by under-sampling using patch-based learning and then obtains an aliased-free image by subtracting the predicted alias components from the input alias super-imposed image. Conv, convolution; DRL-CNN, deep residual learning convolutional neural network; Leaky ReLU, leaky rectified linear units.

with 64 filters of size $3 \times 3 \times 64$, batch normalization, and a leaky ReLU. The output layer has a Conv layer with a filter of size $3 \times 3 \times 64$. The kernel size of the Conv layer was set to 3×3 . All these parameters were determined according to the visual geometry group (VGG) network by Simonyan and Zisserman.²¹ In the de-aliasing of MRI images, the size of the output image should be the same as that of the input aliased image. Therefore, simple zero data padding at the boundary before the Conv operation was carried out to make the feature map of the middle layers have the same size as that of the input image. Zhang et al. showed that this simple zero padding strategy has good denoising performance without boundary artifacts. With these layer configurations, the receptive field of the network, which is defined as the region in the input space that a particular CNN is looking at, was given by $(2d + 1) \times (2d + 1)$, where d is the total number of layers, for a kernel size of 3×3 .

Image reconstruction in DRL-CNN

The image reconstruction procedure used in this study can be divided into two steps: reconstruction step by DRL-CNN and data consistency step, as shown in Fig. 2.

(a) Reconstruction by DRL-CNN: Let the under-sampled signal and original image (ground truth) vector be $y \in \mathbb{C}^N$, $x \in \mathbb{R}^N$, respectively and the full Fourier encoding matrix be $F \in \mathbb{C}^{N \times N}$, then Fx represents the full k -space data. Consider a diagonal matrix representing the under-sampling mask as $U \in \mathbb{R}^{N \times N}$, then the under-sampled k -space signal can be expressed as UFx and zero-filled image x_z is obtained by Eq. (1),

$$x_z = F^H UFx, \quad (1)$$

where H represents the Hermitian transpose operation. Aliasing artifacts v caused by under-sampling will be super-imposed on the zero-filled image x_z . In this time, x_z can be expressed as Eq. (2):

$$x_z = x + v. \quad (2)$$

When the original mapping is more like an identity mapping, the residual mapping will be much easier to be optimized.¹⁹ Zero-filled image x_z is much more like the fully scanned image x and close to identity mapping. Therefore, residual learning formulation is more suitable for image de-aliasing in DRL-CNN. Let $R(x_z)$ be the residual mapping to predict v ($R(x_z) \approx v$), reconstructed image x' is obtained by Eq. (3):

$$x' = x_z - R(x_z). \quad (3)$$

(b) Error reduction of the k -space signal: It is reasonable to keep the measured signal data y in non-zero-filled part, and replace the zero-padded part with predicted k -space data Fx' . This data consistency step will reduce the mean squared error in k -space and the quality of resultant image is expected to be improved.¹³ Hereinafter, this technique is called the data consistency step in k -space, and the DRL-CNN resulting from performing this process is called the DRL-CNN-K.

$$y' = F^H \{y + (I - U)Fx'\}. \quad (4)$$

The loss function used to update the network parameter is the mean squared error $l(\Theta)$ between the true and estimated

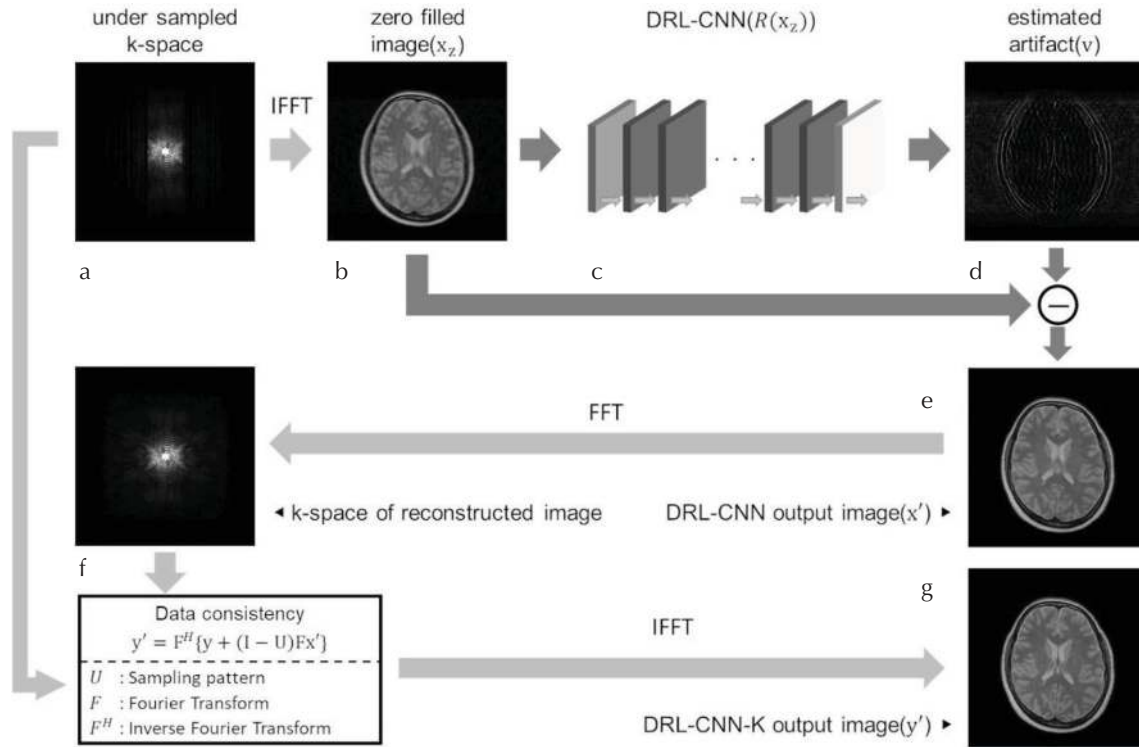


Fig. 2 The image reconstruction procedure is divided into two steps: reconstruction step using DRL-CNN and error reduction step in the k -space domain. k -Space signal is under-sampled according to the sampling pattern (a). The under-sampled signal is replaced with zero data on points where the signal is not acquired. At this point, aliasing artifacts appear on the zero-filled image (b). DRL-CNN is used to estimate the aliasing artifacts (c and d). Reconstructed image is obtained by subtracting the estimated artifact from the zero-filled image (e). Inverse Fourier-transform is applied to obtained image and then the calculated k -space signals on the sampling points are replaced with the acquired true signal (a) to improve the data consistency (f). The updated signal is applied Fourier-transform to obtain updated image (g). DRL-CNN, deep residual learning convolutional neural network; DRL-CNN-K, deep residual learning convolutional neural network in k -space.

artifacts. The loss function for the i -th image in the total number of N training images is defined as Eq. (5):

$$l(\Theta) = \frac{1}{2N} \sum_{i=1}^N \|R(x_{zi}; \Theta) - (x_{zi} - x_i)\|_2^2, \quad (5)$$

where Θ is the network parameter.

Parallel imaging which install multiple receiver coils is widely used in practical use of MR scan, therefore, CS and parallel imaging will be used together in practical applications of CNN. This paper focuses solely on under-sampling signal supposing the use of single receiver coil MRI to evaluate the difference in obtained image quality between DRL-CNN and conventional iterative reconstruction methods.

MR images used in this studies are proton-density-weighted (PDW), T_1 -weighted (T_1W), and T_2 -weighted (T_2W) images of the head included in the information eXtraction from image (IXI) dataset (<http://brain-development.org/ixi-dataset/>).²² All these are absolute value images. The IXI dataset openly provides a large number of healthy images. For this study, we used 2500 images (100 images per person), that are sampled from various sequences captured using a 1.5T MRI system (Intera, Philips Healthcare, Best, The Netherlands) at Guy's Hospital.

We split the 2500 MR images into five data set including 25 subjects each to apply fivefold cross-validation. Four of five data set were used for training and one for testing. Average of five times evaluation was used for peak-signal-to-noise ratio (PSNR) and structural similarity index (SSIM)²³ value of each examination.

The imaging conditions were as follows: for T_1W images, TR = 9.813 ms, TE = 4.603 ms, number of phase encoding steps was 192, and flip angle was 8°; for T_2W images, TR = 8178.34 ms, TE = 100 ms, number of phase encoding steps was 187, echo train length = 16, and flip angle was 90°; and for PDW images, TR = 8178.34 ms, TE = 8 ms, number of phase encoding steps was 187, echo train length = 16, and flip angle was 90°. If the size of images is smaller than 256×256 pixels, then and zero data were filled in the both end of the image to be 256×256 pixels image.

The fully sampled k -space data were obtained by taking a Fourier transform of the model images. We consider 2D-random under-sampling, where k -space signals are randomly under-sampled in both the phase- and frequency-encoding direction, as well as 1D-random under-sampling, where only the phase-encoding direction is randomly under-sampled. Since the signal energy is concentrated at the center of the k -space, signal under-sampling was not applied at

those regions. Signal inside a circular region with a radius of 14 pixels for 2D-random under-sampling, and inside 50 lines for 1D-random under-sampling at the center of the k -space were acquired without under-sampling. For both under-sampling methods, sampling rates of 30%, 40%, and 50% were examined. Variable density pseudorandom sampling masks whose sampling density varies in proportion to the Gaussian distribution were created in each sampling rates. In general, 2D-random sampling is not used for 2D image acquisition, because signal acquisition in the frequency-encoding direction is rapid. However, it is worth considering how the learning efficiency of aliasing artifacts by CNN varies depending on the number of under-sampling dimensions. So, we examined 1D- and 2D-random under-sampling in 2D image acquisition.

Comparison with iterative reconstruction methods

Generally, CS-MRI is formulated as a penalized inverse problem including the L1–L2 norm, and the solution for the object function shown as Eq. (6) is taken as the reconstructed MR image.

$$\operatorname{argmin}_x \|UFx - y\|_2^2 + \lambda \|\Psi x\|_1, \quad (6)$$

According to the CS theory, signal sparsity is an important prior to remove the aliasing artifacts due to under-sampling in k -space, therefore, sparsity is introduced by applying sparsifying transform function Ψ to the image. Since L1-minimization problem is non-smooth and non-differentiable function, many algorithms have been proposed so far. There are three types of algorithms including gradient-based algorithm,³ variable splitting algorithm^{16–18} and operator splitting algorithm.²⁴ ISTA¹⁶ is considered as one of the operator splitting algorithm and well known as reconstruction algorithm of CS. ADMM¹⁸ and Split Bregman algorithm¹⁷ are utilized as variable splitting algorithms that came to be used to improve the quality of reconstructed images in the development period of CS-MRI. ADMM considers the augmented Lagrangian function of a penalized inverse problem, and splits variables into subgroups, which can be alternatively optimized by solving a few simple sub problems. The Split Bregman algorithm is a concept from functional analysis for finding extrema of convex functional. The splitting technique is used to decouple the L1- and L2-components in the object function to minimize. Obtained image quality differs depending on the reconstruction algorithm even the object function to be optimized is the same. In general, ADMM provides superior results than previously reported gradient-based method and ISTA.¹⁶ So, we evaluated ISTA which is widely used and known as CS reconstruction algorithm, and two ADMM method, i.e., Split Bregman method (hereinafter referred to as Split Bregman) and constrained split augmented Lagrangian shrinkage algorithm for balanced model (C-SALSA-B)¹⁸ to compare the results with proposed method. Algorithms 1 and 2 show the reconstruction steps for

Algorithm 1

C-SALSA-B

- 1: $n = 1$
 - 2: Repeat
 - 3: $\alpha_{n+1} = \frac{\mu}{\mu + \rho} \Psi F^* U^* (y + h_n) + \gamma_c (z_n + d_n)$
 $+ \Psi F^* \left[(1 - \gamma_c) I - \frac{\mu}{\mu + \rho} U^* U \right] F \Psi^* (z_n + d_n)$
 - 4: $z_{n+1} = \operatorname{Soft}_{\tau_s} (\alpha_{n+1} - d_n)$
 - 5: $h_{n+1} = h_n - \delta_h (UF \Psi^* \alpha_{n+1} - y)$
 - 6: $d_{n+1} = d_n - \delta_d (\alpha_{n+1} - z_{n+1})$
 - 7: $n = n + 1$
 - 8: Until converge
- Output: $x = \Psi^* \alpha_n$

Algorithm 2

Split Bregman

- 1: $n = 1$
 - 2: Repeat
 - 3: $x_{n+1} = \frac{1}{2\mu + 1} F^* y + \left(I - \frac{1}{2\mu + 1} F^* U^* UF \right) \Psi^* (\alpha_n - b_n)$
 - 4: $\alpha_{n+1} = \operatorname{Soft}_{\tau_b} (\Psi x_{n+1} + b_n)$
 - 5: $b_{n+1} = b_n + \Psi x_{n+1} - \alpha_{n+1}$
 - 6: $n = n + 1$
 - 7: Until converge
- Output: $x = \Psi^* \alpha_n$

C-SALSA-B and Split Bregman respectively, where, $\operatorname{Soft}_{\tau}()$ means soft thresholding using threshold value τ . Wavelet transform with Daubechies (number of coefficients 6) was used as a sparsifying transform function in Split Bregman and C-SALSA-B. Obtained images were compared in terms of the PSNR and SSIM.

Images in the dataset were divided into smaller patches for training the DRL-CNN. Zhang et al.¹⁵ set the patch size to 40×40 when the noise level is known and to 50×50 when the noise level is unknown. The present study aims to remove aliasing artifacts due to signal under-sampling. The appearance of artifacts depends on the under-sampling pattern and may not resemble the noise in the image. The relationship between patch size and DRL-CNN performance was examined by varying the patch size from 31×31 to 71×71 .

Datasets were created for each sampling pattern, sampling rate, and number of images used for training. Adam, an adaptive learning rate optimization algorithm designed for training deep neural networks,²⁵ was used as an optimizer of DRL-CNN training. The batch size of the input dataset was 128 and the learning rate was gradually decreased from 1×10^{-3} to 1×10^{-5} . Data augmentation was never applied in the training because we focused on the relationship between the number of training images and DRL-CNN performance.

Training and testing were performed using the MatConvNet package²⁶ in MATLAB R2017b (MathWorks, Inc., Natick, MA, USA) on a computer equipped with an Intel Core i7-7700 (3.60 GHz) CPU and an NVIDIA GeForce GTX 1080 Ti GPU. The training required 0.4 h for 80 images, 1.2 h for 240 images, and 8 h for 2000 images.

Results

Figure 3 shows the PSNR and SSIM values of the reconstructed images for various patch sizes for 1D-random sampling. The dataset used in this examination contained 300 PDW images and a single sampling pattern was used for 40% under-sampling. The stride for patching was standardized to about one-third of the patch size. As shown in Fig. 3, good results were obtained when the patch size was 61×61 . Therefore, the patch size was set to 61×61 and the number of network layers was set to 30 in subsequent examinations. The relationship between PSNR, SSIM, and the number of training images is summarized in Table 1. The number of training images was increased from 80 to 2000 for the PDW, T_1W , and T_2W image models. Table 1 shows the fivefold cross-validation results for datasets with 100 to 2500 images. Here, the sampling rate was set to 40% using the same under-sampling mask. Figure 4 shows the PSNR and SSIM values for various numbers of training PDW images for 1D- and 2D-random under-sampling. Figure 5 shows the PSNR and SSIM values for various numbers of training epochs for

1D- and 2D-random under-sampling using 2000 PDW training images. Figure 6 shows the reconstructed images for the DRL-CNN and the DRL-CNN-K for 1D-random under-sampling using the trained network for 80, 240, and 2000 images. Figures 6a, 6b, and 6q respectively show the fully scanned image, zero-filled image, and error image of Fig. 6b. Figures 6r–6w show the error images corresponding to the DRL-CNN images in Figs. 6c, 6e, and 6g and the DRL-CNN-K images in Figs. 6d, 6f, and 6h, respectively. For both 1D- and 2D-random under-sampling, the PSNR and SSIM values for the DRL-CNN increase with the number of training images. As shown in Fig. 6, aliasing artifacts tend to decrease with increasing number of training images. Furthermore, structure preservation is improved with the use of the data consistency step in the DRL-CNN-K. The PSNR and SSIM values increase with increasing training epoch number, as shown in Fig. 5.

A comparison of reconstructed images obtained with the DRL-CNN, the DRL-CNN-K, and conventional iterative methods, namely ISTA, C-SALSA-B, and Split Bregman are shown in Figs. 7 and 8 for 1D- and 2D-random under-sampling, respectively. The reconstruction parameter for ISTA was soft threshold level = $2\sigma_n$, those for C-SALSA-B are $\gamma = 0.5$, $\mu = 1$, $\rho = 1$, $\delta_h = \delta_d = 1$, and $\tau_s = 2\sigma_n$, and those for Split Bregman are $\mu = 1$ and $\tau_b = 2\sigma_n$, where σ_n is the estimated standard deviation of noise in the image. The aliasing artifacts in the images are shown in Figs. 7e-1 to 7j-1 and Figs. 8e-1 to 8j-1 for a sampling rate of 30%. As shown in Fig. 8, aliasing artifacts are much fewer for the CNN-based reconstruction method.

Each image includes an enlarged view of the area indicated by the red box. Figures 7 and 8 show that the reconstructed image becomes increasingly similar to the fully scanned image with increasing sampling rate for both sampling patterns. The results of the quantitative evaluation of the sampling rate using 2000 PDW images are summarized in Fig. 9. The PSNR and SSIM values for the iterative

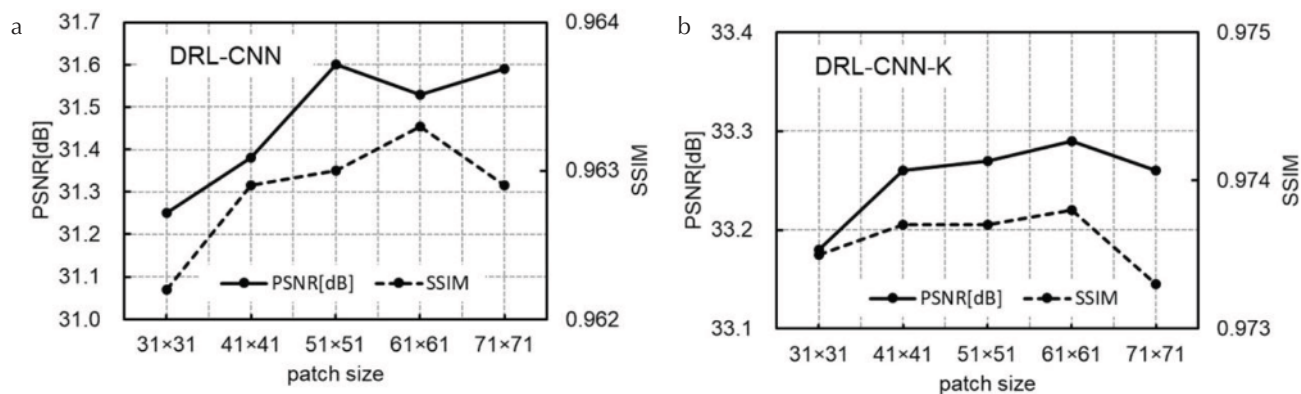


Fig. 3 The relationship between the patch size and PSNR (and SSIM) of reconstructed images for 1D-random under-sampling. The dataset used in this examination contained 300 PDW images and a single sampling pattern was used for 40% under-sampling. (a) DRL-CNN, (b) DRL-CNN-K. Best results were obtained when the patch size is 61×61 . DRL-CNN, deep residual learning convolutional neural network; DRL-CNN-K, deep residual learning convolutional neural network in k -space; PDW, proton-density-weighted; PSNR, peak-signal-to-noise ratio; SSIM, structural similarity index.

Table 1 The PSNR and SSIM results of PDW, T_1W and T_2W images for the different sampling patterns, sampling rate and the number of training images

Pattern	Sequence	Number of training images	DRL-CNN		DRL-CNN-K	
			PSNR [dB]	SSIM	PSNR [dB]	SSIM
2D-random	PDW	80	31.09	0.958	37.80	0.989
		240	31.69	0.958	38.69	0.990
		2000	32.44	0.964	39.61	0.992
	T_1W	80	31.43	0.927	37.02	0.974
		240	31.73	0.931	37.54	0.976
		2000	32.51	0.938	38.40	0.980
	T_2W	80	31.75	0.960	38.29	0.990
		240	32.09	0.963	38.96	0.991
		2000	32.99	0.967	39.85	0.992
1D-random	PDW	80	30.36	0.949	31.97	0.963
		240	31.33	0.955	32.99	0.968
		2000	32.61	0.964	34.26	0.974
	T_1W	80	31.52	0.927	33.00	0.945
		240	32.11	0.935	33.57	0.951
		2000	33.21	0.945	34.69	0.959
	T_2W	80	30.44	0.953	32.08	0.966
		240	31.19	0.958	32.82	0.970
		2000	32.50	0.965	34.14	0.976

CNN, convolutional neural network; DRL, deep residual learning; DRL-CNN-K, deep residual learning convolutional neural network in k -space; PDW, proton-density-weighted; PSNR, peak-signal-to-noise ratio; SSIM, structural similarity index.

reconstruction methods are the averages of 100 images randomly picked up from 2500 images.

For 2D-random sampling, the reconstructed image obtained with ISTA is over-smoothed (i.e., fine details are missing). C-SALSA-B and Split Bregman produce images with higher quality. The DRL-CNN-K produces images that are comparable to those for C-SALSA-B and Split Bregman for each sampling rate. Similar results were obtained in the quantitative evaluation shown in Fig. 9a, where the DRL-CNN-K has PSNR values comparable to those for C-SALSA-B and Split Bregman.

For 1D-random sampling, Fig. 9c shows that the DRL-CNN has higher PSNR values than those for Split Bregman and C-SALSA-B for sampling rates lower than 40%. The PSNR is further improved for DRL-CNN-K, which adds a data consistency step. The comparison of reconstructed images in Fig. 8 shows that many more aliasing artifacts remain for the iterative reconstruction methods when the sampling rate is 30% or 40%. Figure 8 also shows that details are lost and obvious artifacts are visible for the iterative reconstruction methods. In contrast, the structure and small contrast are well preserved with the DRL-CNN. The preservation of the brain structure is further improved with the DRL-CNN-K.

To investigate the relationship between the obtained image quality and the incoherence of signal under-sampling,

a new sampling scheme in which non-random and random under-sampling were mixed in k -space was applied in reconstruction experiments using the DRL-CNN-K. As shown in Fig. 10, regular under-sampling was applied within $\pm L$ lines from the center of k -space except for the central 50 lines, where standard non-skipping sampling was executed. Random under-sampling was applied to the rest of k -space. The sampling density was kept constant in all cases except for the central 50 lines. The obtained average PSNR values for 100 test images for various values of parameter L are shown in Fig. 10b, with the sampling rate set to 30%, 40%, and 50%. The PSNR values increase with increasing parameter L ; i.e., the coherence of signal under-sampling increases even though it decreases for the conventional iterative reconstruction methods.

Reconstruction experiments using variable density non-random under-sampling patterns were performed to further improve the obtained image quality in DRL-CNN. We used a Gaussian distribution for the sampling density of non-random under-sampling. Figures 11a and 11g show the sampling patterns of variable density under-sampling for the reduction factor of 30% and 40%, respectively used in our experiments. In signal under-sampling, the k -space was divided into small segments and uniform signal under-sampling were executed in each segment at a given sampling density. Sampling density and width of each segmented are indicated in

Figs. 11a and 11g. Reconstructed images using the same MR image model as Fig. 8 are shown in Fig. 11. Comparison of PSNR with uniform under-sampling are shown in Fig. 12. As seen in Fig. 11, reconstruction errors of uniform under-sampling Figs. 11f and 11i are larger than that of variable density non-random under-sampling Figs. 11c and 11i and the details of the object are much more preserved in the images obtained with variable density random under-sampling.

The reconstruction time for the DRL-CNN was 1.78 s using the CPU only and 0.022 s using the CPU and GPU. The reconstruction times were 4.03, 13.67, and 13.95 s for ISTA, C-SALSA-B, and Split Bregman, respectively. Because C-SALSA-B and Split Bregman have more steps in the reconstruction procedure than does ISTA, their reconstruction times are longer than that of ISTA. The reconstruction time for the DRL-CNN-K is almost the same as that for the

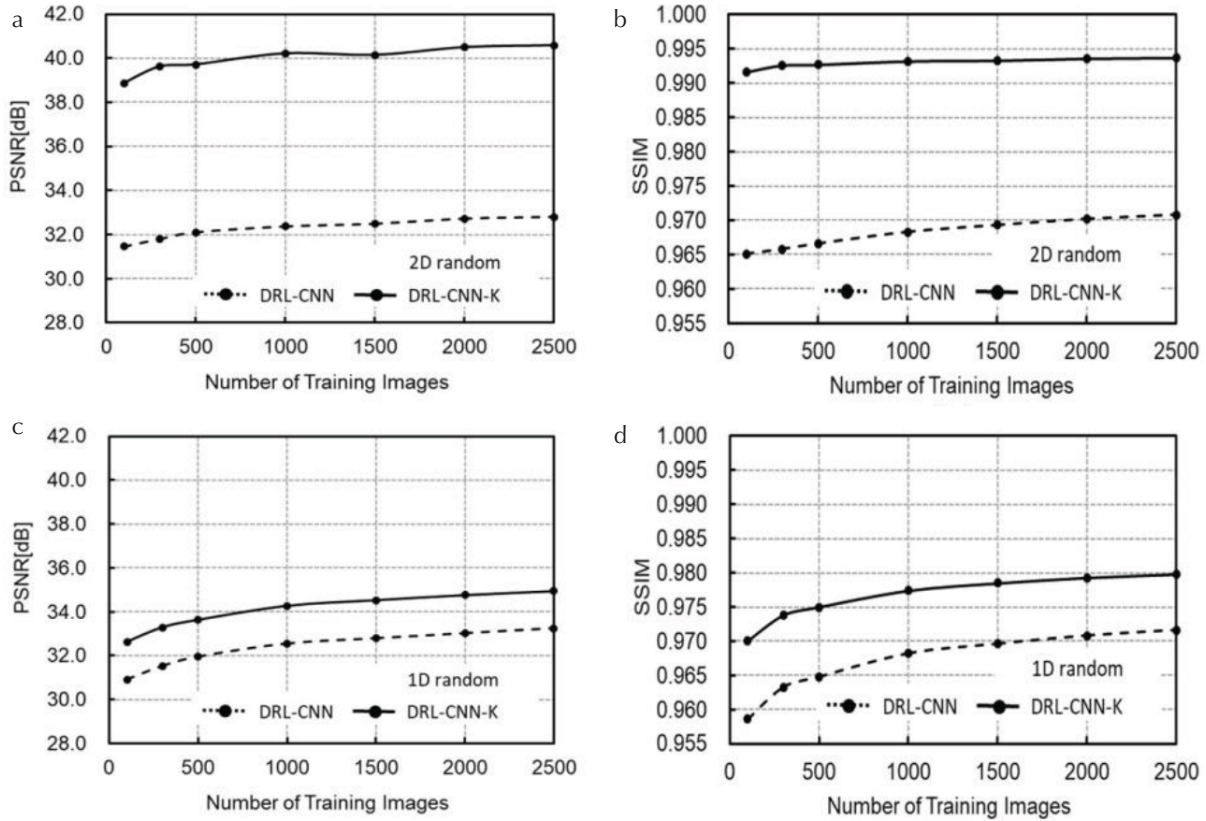


Fig. 4 The relationship between the number of training images and PSNR (and SSIM). The number of PDW images used in the dataset was increased from 100 to 2500. (a–d) show PSNR and SSIM results for 2D- and 1D-random under-sampling using 40% signal, respectively. For both 1D- and 2D-random under-sampling, PSNR and SSIM improve with the increase of training images. The improvement of PSNR with reference to the number of training images appears to be greater for the 1D-random under-sampling. DRL-CNN, deep residual learning convolutional neural network; DRL-CNN-K, deep residual learning convolutional neural network in *k*-space; PDW, proton-density-weighted; PSNR, peak-signal-to-noise ratio; SSIM, structural similarity index.

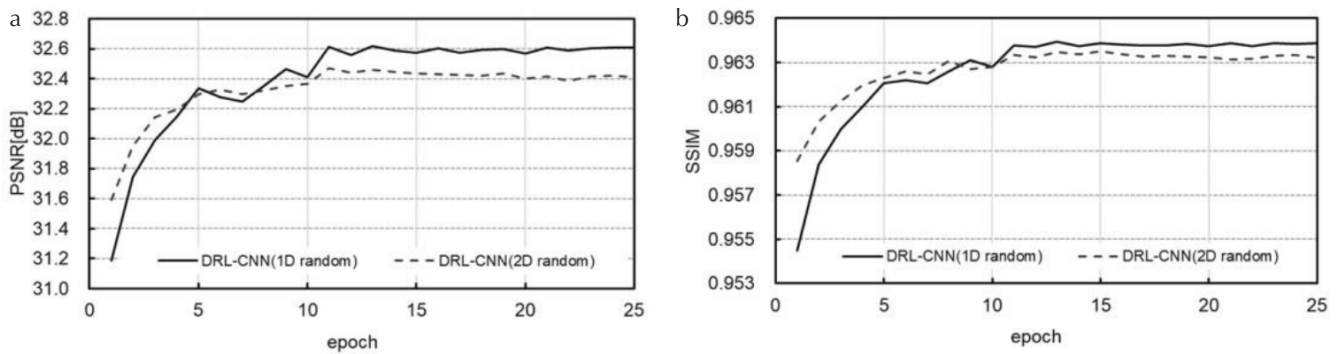


Fig. 5 The relationship between the increase of epoch number and PSNR (and SSIM) (a) and (b) shows PSNR and SSIM results, respectively. The sampling rate was 40% and 2000 images are used for training. PSNR and SSIM improve with the increase of epoch number. DRL-CNN, deep residual learning convolutional neural network; PSNR, peak-signal-to-noise ratio; SSIM, structural similarity index.

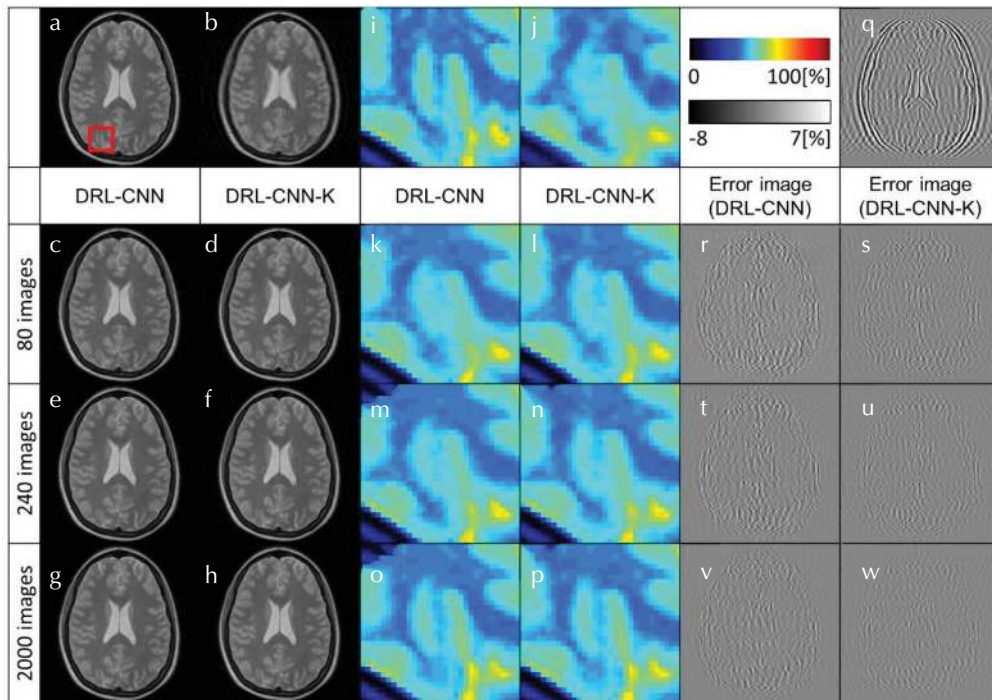


Fig. 6 Improvement of image quality as the increase of the number of training images in 1D-random under-sampling (sampling rate 40%). The number of PDW training images is increased from 100 to 2500 with fivefold cross-validation. (a) Fully sampled image, (b) zero-filled image. (c), (e), and (g) are images with DRL-CNN for the number of training image 80, 240 and 2000, respectively. (d), (f), and (h) are images with DRL-CNN-K and (i-p) are enlarged images corresponding to (a-h), respectively. (q-w) are error images corresponding to (b-h). It is shown that as the number of training images increases, the quality of reconstructed images improves. DRL-CNN, deep residual learning convolutional neural network; DRL-CNN-K, deep residual learning convolutional neural network in *k*-space; PDW, proton-density-weighted.

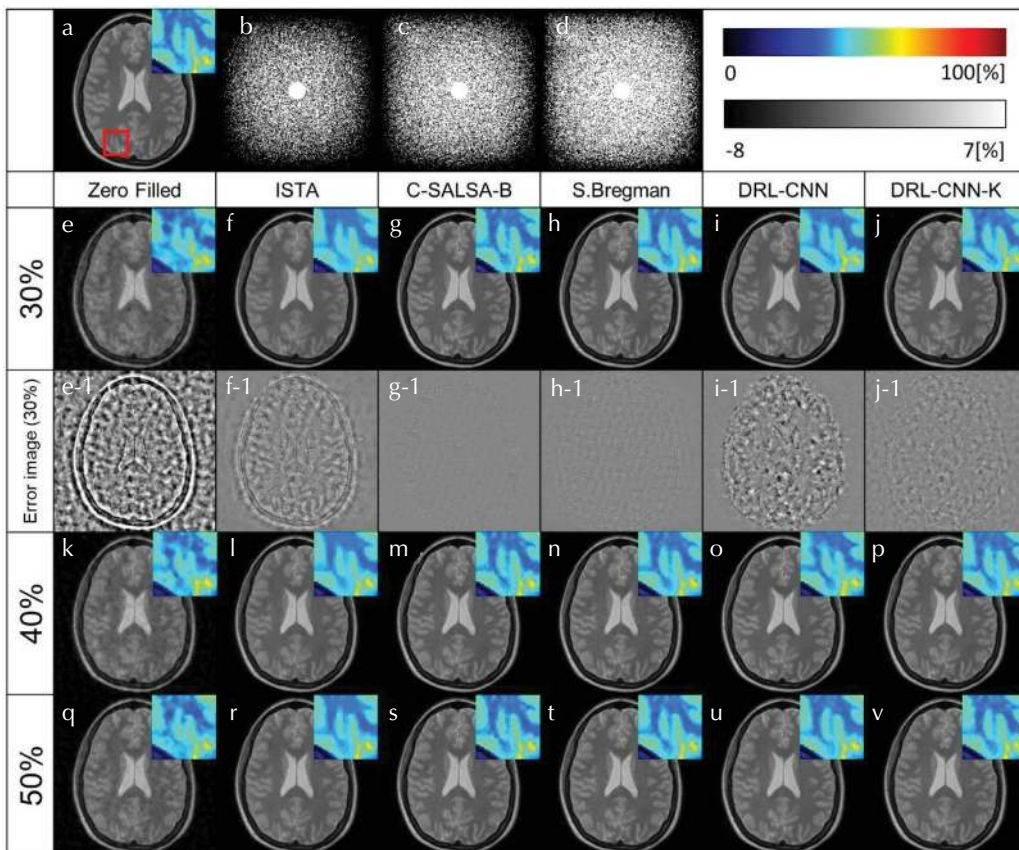


Fig. 7 Comparison of reconstructed images for 2D-random under-sampling. Sampling rate is set as 30%, 40%, and 50% and the dataset used in this examination contains 2500 PDW images. (a) Fully sampled image, (b-d) are sampling patterns for 30%, 40%, and 50% respectively. (e-j) are zero-filled, ISTA, C-SALSA-B, Split Bregman, DRL-CNN, DRL-CNN-K using 30% signal. (k-p) and (q-v) corresponding images to (e-j) for 40% and 50% signal respectively. (e-1-j-1) are error images of (e-j), respectively. C-SALSA-B, constrained split augmented Lagrangian shrinkage algorithm for balanced model; DRL-CNN, deep residual learning convolutional neural network; DRL-CNN-K, deep residual learning convolutional neural network in *k*-space; ISTA, iterative soft thresholding algorithm; PDW, proton-density-weighted.

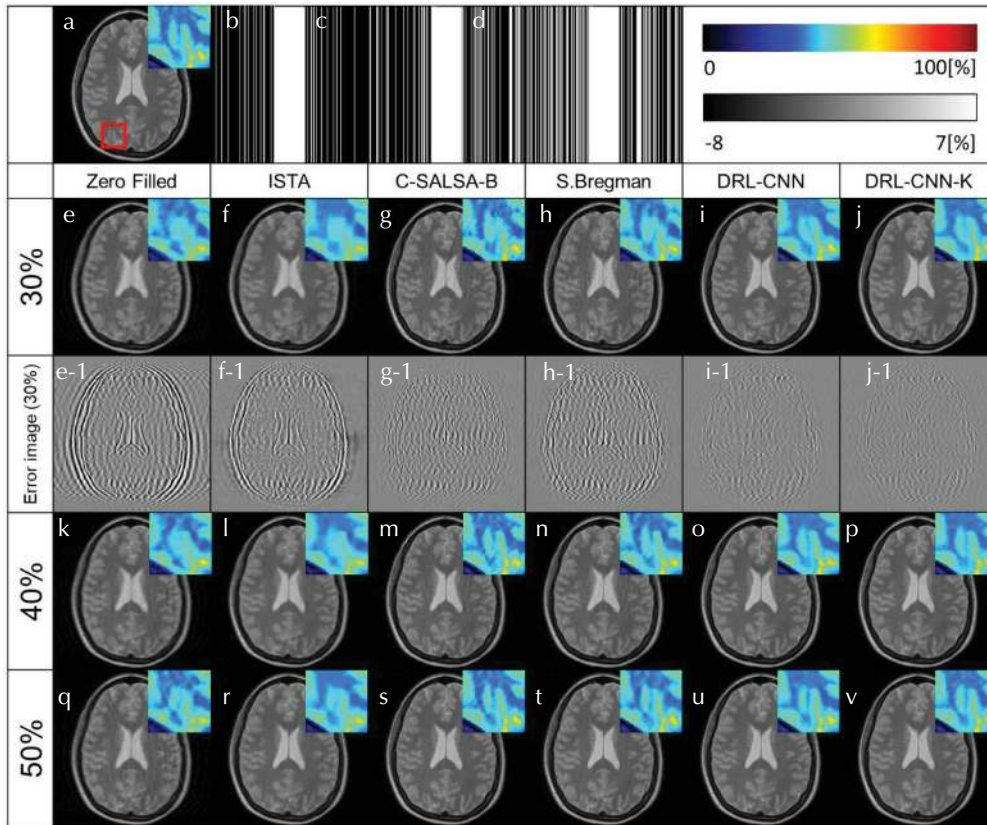


Fig. 8 Comparison of reconstructed images for 1D-random under-sampling. Sampling rate is set as 30%, 40%, and 50% and the dataset used in this examination contains 2500 PDW images. (a) Fully sampled image, (b–d) are sampling patterns for 30%, 40%, and 50% respectively. (e–j) are zero-filled, ISTA, C-SALSA-B, Split Bregman, DRL-CNN, DRL-CNN-K using 30% signal. (k–p) and (q–v) corresponding images to (e–j) for 40% and 50% signal respectively. (e-1–j-1) are error images of (e–j), respectively. C-SALSA-B, constrained split augmented Lagrangian shrinkage algorithm for balanced model; DRL-CNN, deep residual learning convolutional neural network; DRL-CNN-K, deep residual learning convolutional neural network in k -space; ISTA, iterative soft thresholding algorithm; PDW, proton-density-weighted.

DRL-CNN because the data consistency step takes a very short time.

Discussion

We will first focus on the patch size of the training images for the DRL-CNN. In Fig. 3, the PSNR and SSIM values tend to increase with increasing patch size for 1D-random under-sampling. Good results were obtained using a patch size of 61×61 . In this study, the target of prediction by the DRL-CNN is aliasing artifacts due to under-sampling, which have properties different from those of Gaussian noise.¹⁵ Thus, increasing the patch size makes it easier for the CNN to capture the occurrence pattern of aliasing artifacts, leading to improved reconstruction performance. However, as the patch size increases, the CNN layers become deeper, and the number of patches that can be cut out may decrease. The optimal patch size for image reconstruction using the DRL-CNN was found to be 61×61 , which is larger than that used for noise processing. The optimal patch size depends on the environment because aliasing artifact appearance depends on the sampling rate and under-sampling method. The feature map of the DRL-CNN has the same size as that of the input image. Lee et al.¹² reported that a multi-scale network such as U-Net is more robust than a single-resolution network because of its large receptive field for capturing the aliasing

artifact pattern. A comparison of single- and multi-scale DRL-CNNs will be conducted in future studies.

Figure 4 and Table 1 show that the PSNR tends to improve with increasing number of training images, regardless of the number of random sampling dimensions and the weight of the MRI parameter (proton density, T_1 , or T_2). The improvement in PSNR with increasing number of training images is larger for 1D-random sampling. As described later, this can be explained by the incoherence of artifacts being lower for 1D-random sampling than for 2D-random sampling. It is difficult for the CNN to learn artifact appearance rules in 2D under-sampling where aliasing artifacts have a random-noise-like distribution. Therefore, learning effectiveness reaches its peak early. It is common for the PSNR and SSIM values to improve with increasing number of training images for a deep CNN; however, it was found that the PSNR improvement was different between 1D and 2D under-sampling in k -space.

Next, we consider the relationship between the sampling rate and the quality of reconstructed images. Looking at 2D-random sampling first, the PSNR and SSIM values for the DRL-CNN are smaller than those for ISTA, as shown in Figs. 7 and 9; however, those metric can be improved by adopting DRL-CNN-K. These results indicate that superior performance in terms of CS reconstruction can be obtained by simply adding a data consistency step in signal space.

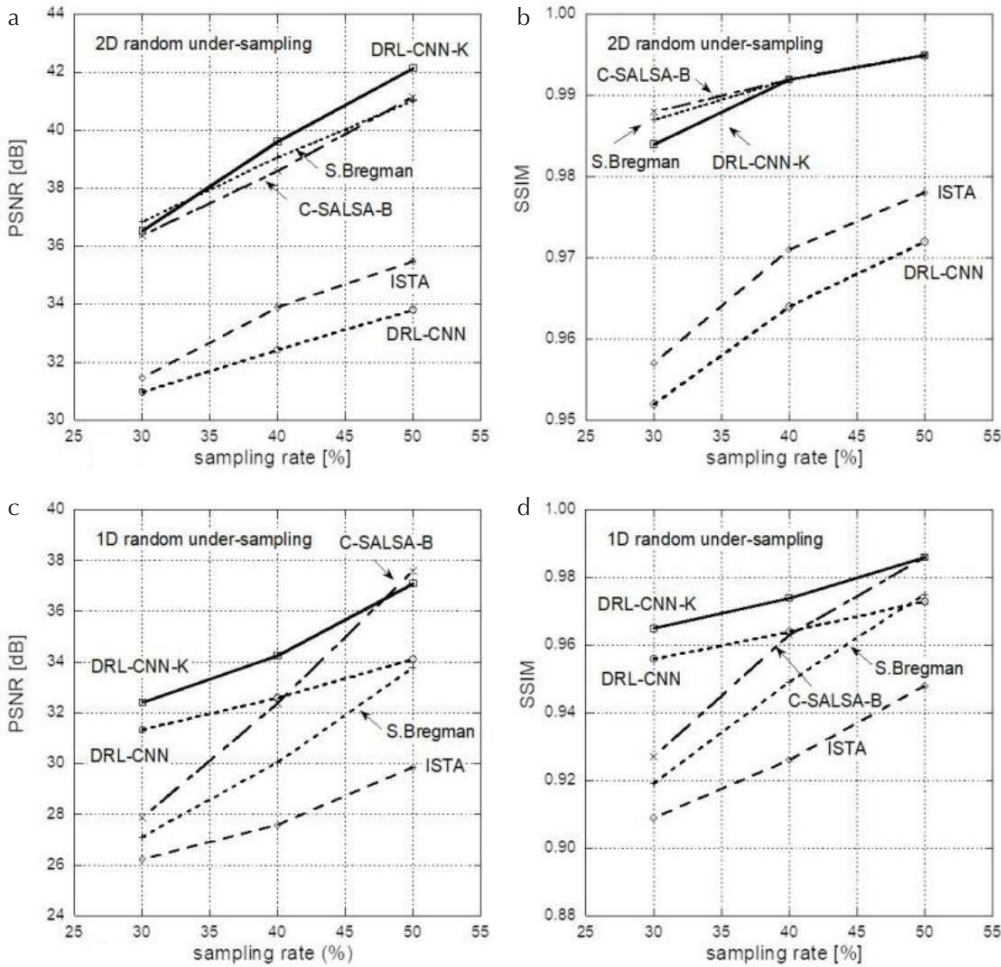


Fig. 9 Comparison of PSNR and SSIM with iterative reconstruction methods, i.e., ISTA, C-SALSA-B and Split Bregman for 30–50% sampling rate. (a) and (b) PSNR and SSIM vs. sampling rate for 2D-random under-sampling, (c) and (d) PSNR and SSIM vs. sampling rate for 1D-random under-sampling. Comparative PSNRs with C-SALSA-B and split Bregman are obtained in DRL-CNN-K in 2D-random under-sampling. Focusing on 1D-random under-sampling, DRL-CNN shows higher PSNR than split Bregman and C-SALSA-B when sampling rate is smaller than 40%. PSNR is further improved in DRL-CNN-K by adding data consistency step to DRL-CNN. C-SALSA-B, constrained split augmented Lagrangian shrinkage algorithm for balanced model; DRL-CNN, deep residual learning convolutional neural network; DRL-CNN-K, deep residual learning convolutional neural network in k -space; ISTA, iterative soft thresholding algorithm; PSNR, peak-signal-to-noise ratio; SSIM, structural similarity index.

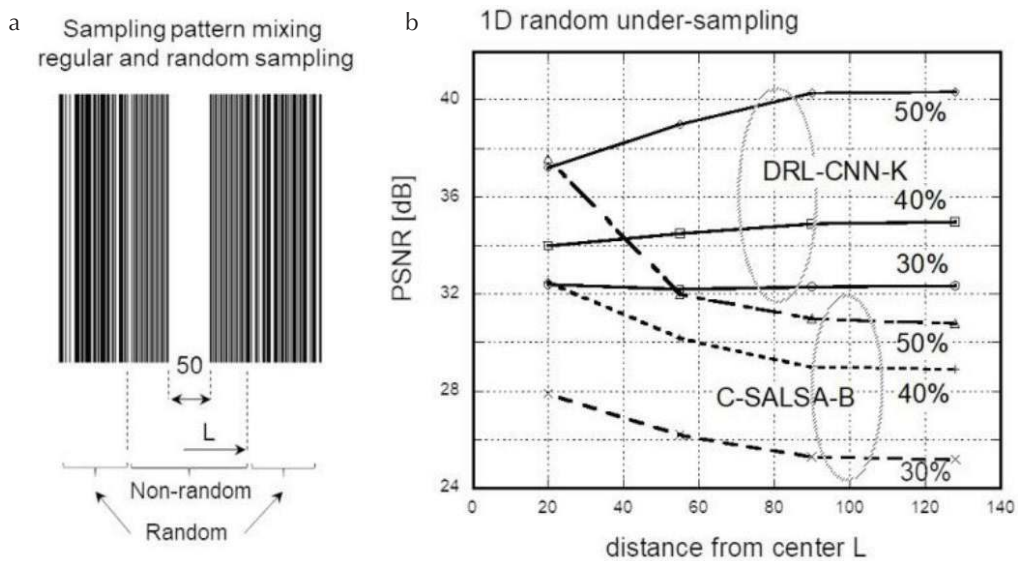


Fig. 10 The relationship between image PSNR and incoherency of sampling pattern. A new sampling scheme in which non-random and random under-sampling were mixed in a k -space and they were tested for reconstruction experiments using DRL-CNN-K and C-SALSA-B. As shown in (a), regular under-sampling is applied within $\pm L$ lines from to the center of k -space except the central 50 lines. Random under-sampling is applied to the rest of k -space. (b) shows the PSNR increase with increasing parameter L ; i.e., the coherence of signal under-sampling increases even though it decreases for the conventional iterative reconstruction methods. C-SALSA-B, constrained split augmented Lagrangian shrinkage algorithm for balanced model; DRL-CNN-K, deep residual learning convolutional neural network in k -space; PSNR, peak-signal-to-noise ratio.

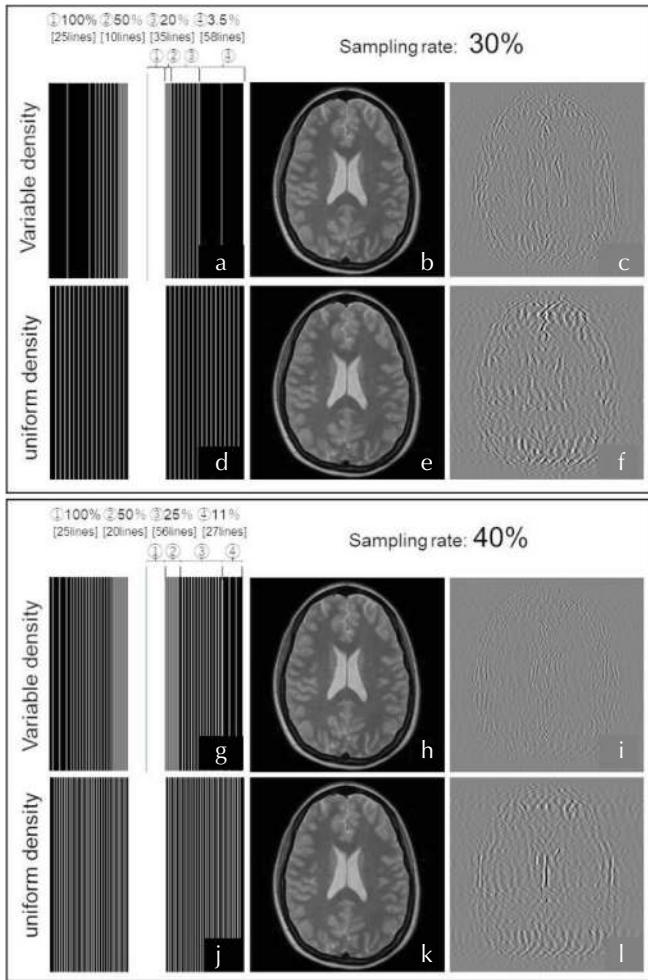


Fig. 11 Comparison of reconstructed images between variable density and uniform density in non-random under-sampling. Sampling rate is set as 30% and 40%. (a) and (g) are variable density non-random under-sampling patterns and (d) and (j) are uniform density non-random under-sampling patterns, respectively. (b), (e), (h), and (k) are reconstructed images using DRL-CNN-K. (c), (f), (i), and (l) are error images corresponding to (b), (e), (h), and (k). It is shown from error images that the reconstruction error of uniform under-sampling is larger than that of variable density non-random under-sampling and the details of the object are much more preserved in the images obtained with variable density random under-sampling. DRL-CNN-K, deep residual learning convolutional neural network in k -space.

The reconstructed image for the DRL-CNN had some errors that depended on the sampling rate, and its Fourier transform also had errors in k -space. The k -space signal estimated by the DRL-CNN is replaced by the acquired true signal at the sampled points, which reduces the errors distributed in k -space and thus improves the PSNR of the image. The image quality improvement obtained from data consistency is proportional to the total number of sampling points replaced by the true signal. Because we assume a real-valued function for the spin density distribution, the k -space signal has Hermitian symmetry; i.e., $y(k_x, k_y) = y(-k_x, -k_y)^*$,

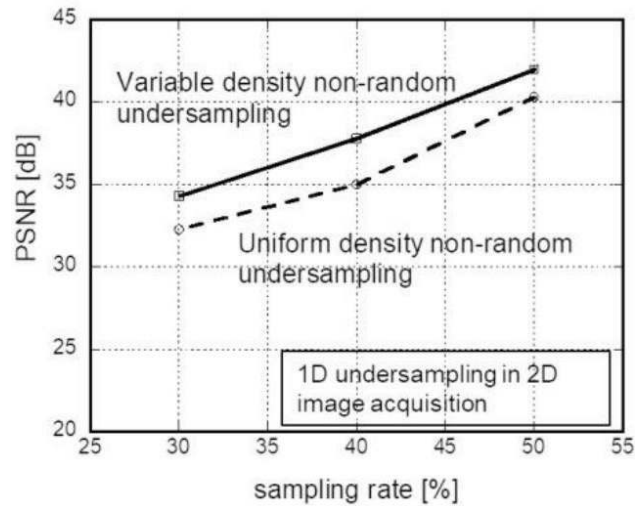


Fig. 12 Comparison of obtained image PSNR between variable density and uniform density in non-random under-sampling. The PSNR of variable density non-random under-sampling is higher than that of uniform density non-random under-sampling. PSNR, peak-signal-to-noise ratio.

where $*$ denotes the complex conjugate of a complex function. Therefore, the replacement of the two sampling points $y(k_x, k_y)$ and $y(-k_x, -k_y)^*$ corresponds to the replacement of the same sampling point signal, and therefore, the replacement of the signal at $y(-k_x, -k_y)^*$ does not contribute to image quality improvement. The probability of acquiring both symmetrical points (k_x, k_y) and $(-k_x, -k_y)$ in signal sampling at a given sampling rate is much smaller in 2D-random under-sampling than in 1D-random under-sampling because the randomness of sampling points is higher in the former sampling. Therefore, the improvement in PSNR and SSIM is greater in 2D-random under-sampling, as shown in Figs. 9a and 9b. The data consistency step is independent of CNN learning, so there is no need to retrain the CNN when this step is added. Moreover, the computation cost of this step is very small.

The obtained image quality and PSNR in 1D-random sampling are better than those in 2D-random sampling for the DRL-CNN, as shown in Figs. 4, 6, 7, and 9, which is inconsistent with CS theory established under the randomness of the measurement matrix. Results similar to those in Fig. 9 were obtained using different sampling patterns but the same sampling rates. In general, 2D-random under-sampling should provide better quality images compared with 1D-random under-sampling from the standard CS viewpoint because the randomness of 2D-random under-sampling is higher than that of 1D-random under-sampling. However, CS with random under-sampling has some limitations in preserving structure details in an image, especially for a low sampling rate, because noise-like artifacts are superimposed onto the reconstructed image. Several studies have utilized regular under-sampling in the application of CNN to CS image reconstruction.²⁷ As a simple example, consider a point image in the image domain.

Sample its Fourier-transformed signal with equal space skipping. The fold-over point images are placed at the position defined by the position of the original point image and skipping distance. For this simple example, it will be easier for the CNN to learn the rule of the fold-over effect in regular under-sampling than in random under-sampling because the aliasing is very simple and easy to recognize. For regular under-sampling, the blurring of the reconstructed image is smaller compared with that obtained in random under-sampling. The results shown in Figs. 9a and 9c are consistent with these insights; the PSNR in 1D-random under-sampling is higher than that in 2D-random under-sampling and performance is excellent especially for a low sampling rate, for which noticeable aliasing artifacts tend to appear. The relationship between the obtained image quality and the incoherence of signal under-sampling in Fig. 10 indicates that the PSNR and SSIM are improved as the coherence of the sampling pattern increases. We can conclude from these results that artifacts with low incoherence are easier to be learned by the CNN, which contributes to the higher image quality for the DRL-CNN.

This is a major difference between CS iterative reconstruction and image domain learning using a CNN. These results indicate that a deep CNN can learn coherent artifacts and is effective especially for cases where the randomness of k -space sampling is rather low, such as in 1D-random under-sampling in 2D image acquisition. Based on the results shown in Fig. 10, we conducted reconstruction experiments using variable density non-random under-sampling to improve the image quality in CS-MRI. Variable density random under-sampling has been widely used in CS-MRI in which dense sampling is executed in the central region of k -space, since most of MR signal energy concentrated on that space. Figure 12 shows that obtained image PSNRs are further improved by adopting variable density sampling in non-random signal under-sampling. Since the signal distribution of k -space depends on the distribution of the imaging subject, the optimal function that define the sampling density in k -space may vary depending on the position of cross-section or imaging conditions. Careful consideration should be paid for the choice of sampling density function.

For the application of the proposed method to clinical examinations, we must consider other practical issues such as echo signal decay by T_2 relaxation in fast spin echo sequences or inhomogeneous image noise in Sensitivity Encoding (SENSE) imaging. The CNN can learn signal decay or noise in addition to artifact appearance. This will be considered in future studies.

The sparsity of images is assumed in CS theory and iterative reconstruction. When objects to be imaged have low sparsity in the image space, sparsity can be introduced by applying a sparsifying transform function or total variation (TV)²⁸ to the image. The preservation of fine structure and the degree of artifacts in the reconstructed images generally depend on the sparsifying transform function, which must thus be carefully chosen. In contrast, there is no need for a

sparsifying transform function for the DRL-CNN, which simplifies the reconstruction process and avoids the over-smoothing and unnatural appearance introduced by a sparsifying transform function or TV. These are major advantages of the DRL-CNN.

The reconstruction time was about 14 s for C-SALSA-B and Split Bregman in our study. This was shortened for the DRL-CNN to 1.78 s using only the CPU and 0.022 s using the CPU and GPU, irrespective of the number of training images. Iterative reconstruction solves L1–L2 minimization problem with some constraints, therefore, it requires a lot of time to obtain images. The reconstruction time of CNN-based method varies depending on the depth of network, the number of filter kernel and filter size and so on. Since it does not have iterative process in reconstruction process, the reconstruction time becomes much shorter compared with iterative reconstruction and has the potential to overcome the issue of CS long reconstruction time.

One issue with deep learning reconstruction is that training takes a long time. Different artifacts and different sampling patterns in k -space affect the performance of the CNN, because the network is implemented in the image domain. One way to address artifact pattern differences is to incorporate images with many different sampling patterns into the training dataset. Even though this will increase training time, it will improve the robustness of the network.

Conclusion

This paper presented a DRL-CNN approach for image reconstruction using an under-sampled MRI signal in which aliasing artifacts due to under-sampling are learned by a patch-based CNN. The relationship between the obtained image quality and the number of under-sampling dimensions in 2D image acquisition was investigated. A comparison of the proposed DRL-CNN with conventional iterative reconstruction methods (ISTA, ADMM, and Split Bregman iteration) shows that the PSNR and SSIM values for the DRL-CNN are smaller than those for iterative reconstruction in 2D-random under-sampling and comparable in 1D-random under-sampling. The DRL-CNN has a drastically reduced reconstruction time. The improvement in PSNR and SSIM obtained by adding a data consistency step in the DRL-CNN is greater in 2D-random under-sampling. With this step, the overall performance is comparable to that for iterative reconstruction in 2D-random under-sampling and superior in 1D-random under-sampling. The highly incoherent aliasing artifacts in random under-sampling contribute to the image quality improvement in CS iterative reconstruction. In contrast, low-incoherence artifacts are easier to learn by the proposed CNN, and therefore, higher PSNR and SSIM values were obtained in 1D-random under-sampling for the DRL-CNN. The characteristics of image reconstruction using the DRL-CNN were clarified by a comparison with conventional iterative reconstruction methods.

Acknowledgments

This work was partially supported by JSPS KAKENHI grants 16K06379, 19K04423 and the Kayamori Foundation of Informational Science Advancement. We acknowledge <http://brain-development.org/ixi-dataset/> as the source of the IXI data used in this paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

1. Donoho DL. Compressed sensing. *IEEE Trans Inform Theory* 2006; 52:1289–1306.
2. Candès EJ, Wakin MB. An introduction to compressive sampling. *IEEE Signal Process Mag* 2008; 25:21–30.
3. Lustig M, Donoho DL, Pauly JM. Sparse MRI: the application of compressed sensing for rapid MR imaging. *Magn Reson Med* 2007; 58:1182–1195.
4. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: sensitivity encoding for fast MRI. *Magn Reson Med* 1999; 42:952–962.
5. Griswold MA, Jakob PM, Heidemann RM, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med* 2002; 47:1202–1210.
6. Hinton G, Deng L, Yu D, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Process Mag* 2012; 29:82–97.
7. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems*, vol. 1, Lake Tahoe, Nevada, 2012; 1097–1105.
8. Roth HR, Lee CT, Shin HC, et al. Anatomy-specific classification of medical images using deep convolutional nets. *IEEE 12th International Symposium on Biomedical Imaging*, IEEE, NY, USA, 2015; 101–104.
9. Zhao X, Zhang H, Liu H, Qin Y, Zhang T, Zou X. Single MR image Super-resolution via channel splitting and serial fusion network. arXiv: 1901.06484v1.
10. Wang S, Su Z, Ying L, et al. Accelerating magnetic resonance imaging via deep learning. *IEEE 13th International Symposium on Biomedical Imaging*, IEEE, Prague, Czech Republic, 2016; 514–517.
11. Kwon K, Kim D, Park H. A parallel MR imaging method using multilayer perceptron. *Med Phys* 2017; 44:6209–6224.
12. Lee D, Yoo J, Tak S, Ye JC. Deep residual learning for accelerated MRI using magnitude and phase networks. *IEEE Trans Biomed Eng* 2018; 65:1985–1995.
13. Hyun CM, Kim HP, Lee SM, Lee S, Seo JK. Deep learning for undersampled MRI reconstruction. *Phys Med Biol* 2018; 63:135007.
14. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. *Nature* 2018; 555:487–492.
15. Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Trans Image Process* 2017; 26:3142–3155.
16. Daubechies I, Defrise M, De Mol C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun Pure Appl Math* 2004; 57:1413–1457.
17. Goldstein T, Osher S. The split Bregman method for L1-regularized problems. *SIAM J Imaging Sci* 2009; 2:323–343.
18. Liu Y, Cai JF, Zhan Z, et al. Balanced sparse model for tight frames in compressed sensing magnetic resonance imaging. *PLoS One* 2015; 10:e0119584.
19. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Las Vegas, USA, 2016; 770–778.
20. Maas AL, Hannun AY, Ng AY. Rectifier nonlinearities improve neural network acoustic models. *Proceedings of the 30th International Conference on Machine Learning*, vol. 30, Atlanta, GA, USA, 2013.
21. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *Proceedings of the 3rd International Conference on Learning Representations*, San Diego, CA, USA, 2015, arXiv:1409.1556.
22. IXI dataset. <https://brain-development.org/ixi-dataset/> (Accessed: July 30, 2018).
23. 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.
24. Huang J, Zhang S, Metaxas D. Efficient MR image reconstruction for compressed MR imaging. *Med Image Anal* 2011; 15:670–679.
25. Diederik PK, Jimmy B. Adam: a method for stochastic optimization. *Proceedings of the 2015 International Conference on Learning Representations*, San Diego, CA, USA, 2015.
26. Vedaldi A, Lenc K. MatConvNet: convolutional neural networks for MATLAB. *Proceedings of the 23rd ACM international conference on Multimedia*, Brisbane, Australia, 2015; 689–692.
27. Hammernik K, Knoll F, Sodickson DK, Pock T. On the influence of sampling pattern design on deep learning-based MRI reconstruction. *Proceeding of the 25th Annual Meeting of ISMRM*, Hawaii, 2017; 644.
28. Rudin LI, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. *Phys D Nonlinear Phenom* 1992; 60:259–268.