Reconstruction of cardiovascular black-blood T2-weighted image by deep learning algorithm: A comparison with intensity filter

Ryo Ogawa[1] , Tomoyuki Kido[1], Masashi Nakamura[1], Atsushi Nozaki[2], R Marc Lebel[3], Teruhito Mochizuki[1,4] and Teruhito Kido[1]

Abstract

Background: Deep learning–based methods have been used to denoise magnetic resonance imaging.

Purpose: The purpose of this study was to evaluate a deep learning reconstruction (DL Recon) in cardiovascular black-blood T2-weighted images and compare with intensity filtered images.

Material and Methods: Forty-five DL Recon images were compared with intensity filtered and the original images. For quantitative image analysis, the signal to noise ratio (SNR) of the septum, contrast ratio (CR) of the septum to lumen, and sharpness of the endocardial border were calculated in each image. For qualitative image quality assessment, a 4-point subjective scale was assigned to each image (1 = poor, 2 = fair, 3 = good, 4 = excellent).

Results: The SNR and CR were significantly higher in the DL Recon images than in the intensity filtered and the original images ($p < .05$ in each). Sharpness of the endocardial border was significantly higher in the DL Recon and intensity filtered images than in the original images ($p < .05$ in each). The image quality of the DL Recon images was significantly better than that of intensity filtered and original images ($p < .001$ in each).

Conclusions: DL Recon reduced image noise while improving image contrast and sharpness in the cardiovascular black-blood T2-weight sequence.

Introduction

Cardiovascular black-blood T2-weighted imaging is widely used to examine acute myocardial infarction,[1] often using a short-inversion-time inversion-recovery (STIR) sequence. The technique is sensitive to cardiovascular motion, which along with stagnant blood in the lumen in turn can lead to myocardial signal loss and reduced contrast. Further, parallel imaging can be used to decrease scanning time, which

[1]Department of Radiology, Ehime University Graduate School of Medicine, Toon, Japan
[2]MR Collaboration and Development, GE Healthcare, Tokyo, Japan
[3]MR Collaboration and Development, GE Healthcare, Calgary, Canada
[4]Department of Radiology, I.M. Sechenov First Moscow State Medical University, Russia

Corresponding author:
Ryo Ogawa, Department of Radiology, Ehime University Graduate School of Medicine, Shitsukawa, Toon 791-0295, Japan.
Email: qq8y7cvd@tiara.ocn.ne.jp

often decreases the signal to noise ratio (SNR) in these images. The myocardial SNR in STIR images may be improved by increasing the slice thickness. However, this approach increases the unnecessary signal from slow blood flow, and partial volume effects reduce the sensitivity to edema.[2] The effect of adjusting the sequence on SNR and contrast in cardiovascular T2-weighted imaging has been verified in several studies.[3,4] In addition to changing parameters and sequences, using an intensity filter in post-processing is another option. This technique can remove noise, but may occasionally lead to unclear edges and does not improve the image quality. Recent application of deep learning in radiology has allowed for advances in lesion detection and evaluation and image segmentation.[5,6] Deep learning–based methods, particularly those derived from convolutional neural networks (CNN), have also been used to denoise magnetic resonance imaging (MRI).[7,8] Recently, the dedicated deep learning reconstruction (DL Recon) for 2-dimensional (2D) MRI has also been developed, with the aim to reduce noise and refine image quality. This technology can improve image quality in post-processing without changing the parameters or sequence. Therefore, we needed to verify the reliability of DL Recon images in volunteers before clinical application. The objective of this preliminary study was to validate a deep learning algorithm in cardiovascular T2-weighted images and compare with conventional intensity filtered images.

Material and methods

Volunteers

Fifteen healthy volunteers (15 men; median age 31 (29–35) years) underwent cardiovascular black-blood T2-weighted MRI. The median body mass index of volunteers was 23.8 (21.7–25.3) kg/m2. This study was approved by institutional review board. All participants gave informed consent.

MRI protocol

All MRI examinations were performed with a clinical 3T MR scanner (SIGNA Architect, GE healthcare, WI, USA). A breath-hold black-blood T2-weight image with STIR in 3 short-axis slices (basal, mid, apical) was acquired. Imaging parameters were as follows: repetition time, 2 R to R interval; echo time, 70 ms; slice thickness, 6 mm; field of view, 340×340 mm^2; matrix, 320×320 (reconstruction matrix, 512×512); flip angle, 107; echo train length, 32; acceleration factor, 3. The same raw data were used to reconstruct a set of the following images: the original image, a DL Recon image with the noise reduction level of 25%, 50%, 75%, and 100%, and the original image with an intensity filter.

Deep learning reconstruction

DL images were obtained using a prototype of Recon DL pipeline. The DL Recon receives raw k-space data and outputs high quality images. The prototype software uses a deep CNN, which reconstructs the MR images with higher SNR, higher edge sharpness, and the reduction of truncation artifacts. The CNN was integrated into the standard reconstruction process, allowing the adjustable parameter of the noise reduction level between 0—100%, so that the noise variance of the DL image is reduced to the noise reduction level. The network also recognizes the truncation artifact and works for de-ringing to improve image sharpness. The deep CNN contains over 4.4 million trainable parameters in more than 10,000 kernels and was trained using a dataset of high-resolution near-perfect image and low-resolution image with more truncation artifact or higher noise level, accounting for a database of 4 million images. Image augmentations, such as intensity gradients, rotations and flips, phase manipulations, and additional Gaussian noise were conducted, and 4 million training datasets were created. A single epoch of training involving 4 million iterations was conducted. The Rectified Linear Unit activation was used for the activation function of this CNN. The ADAM optimizer was used to minimize the loss.[9] The reconstruction was conducted using a central processing unit system on a computer with an Intel Xeon E5-2680 v3 at 2.5 GHz. Efficacy of the DL Recon was evaluated by retrospectively reconstructing images with noise reduction factors of 25%, 50%, 75%, and 100%.

Image intensity filter

The acquired images were also applied to the intensity filter and then compared with their DL Recon and the original counterparts. This filter, which was used routinely in clinical practice, separated the image into homogenous and heterogenous regions. The homogenous regions were uniformly smoothed in every direction with a simple box-like filter. The heterogenous regions were smoothed only along gradient perpendicular direction to preserve edges. The edge filter, defined by a blend ratio, aimed to integrate the structured region with the original image. In contrast, the smoothing filter, also defined by the blend ratio, was used to coalesce the homogenous region with the original image. In routine clinical practice, the combination of high sharpening and high smoothing was generally used.

Quantitative image analysis

Quantitative image analysis was performed for all 270 images using the workstation (SYNAPSE VINCENT, Fujifilm Corp., Ltd., Tokyo, Japan). The SNR of septal measured as the ratio of the signal intensity to the standard

deviation[10] was averaged across 3 regions of interests of septal. The contrast ratio (CR) was calculated using the following equation: CR = (Cm - Cl)/Cl (note: Cm, signal intensity of septal myocardium; Cl, signal intensity of lumen).[11] The CR was calculated as the mean of 3 regions of interests of septal and lumen. To evaluate the sharpness of the endocardial border, the signal intensity profile of the septal endocardial border was calculated. Image sharpness was defined as a gradient value (signal intensity/ millimeter) between 20% and 80% of the total intensity range. The gradient value was averaged across the 3 profiles.

Of several DL Recon strengths (25, 50, 75, and 100%), the highest performance one, determined based on the quantitative analysis, was used to compare with intensity filtered and original images.

Intra- and inter-observer reproducibility of SNR, CR, and sharpness were assessed across 3 imaging approaches (original, DL Recon [strength 100%], and intensity filtered images). Intra-observer variability was determined by repeating quantitative image analysis 4 weeks after initial evaluation. Inter-observer variability was determined by 2 radiologists with 7 and 13 years of experience with cardiovascular MRI.

Qualitative image analysis

Image quality was independently evaluated by 3 radiologists (readers 1, 2, and 3), with 7, 9, and 13 years of experience with cardiovascular MRI, using a 4-point subjective scale (1 = poor, 2 = fair, 3 = good, 4 = excellent). The image quality was defined as follows: 1: non-diagnostic due to severely blurred, noisy, and inhomogeneous signal; 2: markedly blurred and noisy; 3: mildly blurred and little noise; 4: well-defined myocardial borders, and almost no noise. The inter-observer reproducibility of image quality was assessed in each imaging method (original, DL Recon [strength 100%], and intensity filtered images).

Statistical analysis

Data are expressed as mean ± standard deviation or median (interquartile range). The differences in SNR, CR, sharpness, and image quality values between 3 examined groups were analyzed using Wilcoxon matched-pairs signed-rank tests and Bonferroni correction was used for multiple comparisons.

Intra- and inter-observer reproducibility of SNR, CNR, and sharpness values were assessed using the Bland– Altman method and Spearman correlation coefficients. Inter-observer reproducibility of qualitative image quality was determined by Kappa value. All statistical analyses were conducted using the JMP software (Version 11.2; SAS Institute, Inc., Cary, NC, USA).

Results

All scans were successfully completed. The median scan time of black-blood T2-weight images was 11 (10–11) s for each slice. The median heart rate was 59 (53–64) beats/min. The reconstruction time of DL Recon images in each noise reduction factor was approximately 50 s in 3 short-axis slices. The reconstruction time did not change according to the noise reduction factor. Figure 1 shows DL Recon images with noise reduction factors of 25%, 50%, 75%, and 100%. Figure 2 illustrates a case with the original image, its DL Recon (strength 100%), and intensity filtered images.

Quantitative image analysis

Table 1 shows the quantitative values of all 270 images. Forty-five DL Recon images with 100% strength were compared with intensity filtered and original images Figure 3. The SNR and CR were significantly higher in the DL Recon than that of intensity filtered and original images ($p < .05$ in each). The sharpness of the endocardial border was significantly higher in the DL Recon and intensity filtered images than in the original images ($p < .05$ in each).

Qualitative image quality analysis

Forty-five DL Recon images with 100% strength were compared with intensity filtered and original images. The image quality of the DL Recon (3.5 ± 0.7) was better than that of intensity filtered (3.0 ± 0.5) and original images (2.8 ± 0.5) ($p < .001$ in each). The image quality of the intensity filtered images was better than that of their original counterparts ($p < .05$ in each).

Analysis of reproducibility

There was significant correlation in SNR, CR, and sharpness values (Intra-observer: r = 0.92, $p < .001$; r = 0.93, $p < .001$; r = 0.93, $p < .001$; Inter-observer: r = 0.74, $p < .001$; r = 0.89, $p < .001$; r = 0.84, $p < .001$). Mean intra-observer differences (bias) were 0.2 (95% confidence interval [CI] -2.7 to 3.1) for SNR, 0 (95% CI -2.9 to 2.9) for CR, −3.7 (95% CI -118.7 to 111.3) for sharpness (Figure 4). Mean inter-observer differences (bias) were −1.7 (95% CI -6.9 to 3.5) for SNR, −0.8 (95% CI -7.1 to 5.5) for CR, −2.9 (95% CI -148.5 to 142.7) for sharpness (Figure 4). The kappa values of inter-observer agreement (readers 1 and 2) for qualitative image quality were 0.81 for the original, 0.79 for the DL Recon, and 0.80 for the intensity filtered images. The kappa values of inter-observer agreement (readers 1 and 3) for qualitative image quality were 0.81 for the original, 0.74 for the DL Recon, and 0.73 for the intensity filtered images.
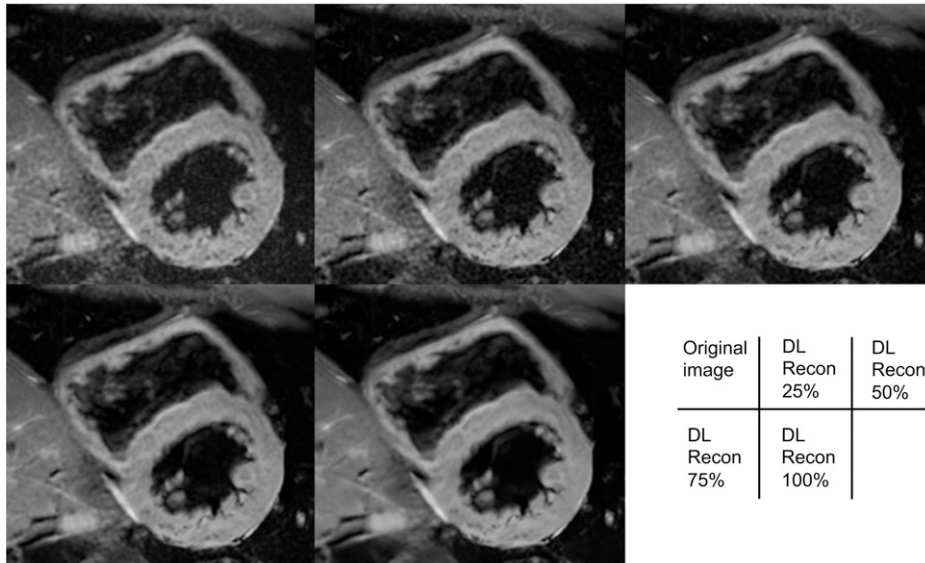
Figure 1. Images from different strengths of deep learning reconstruction. The noise is reduced as the strength changes. DL Recon: deep learning reconstruction.
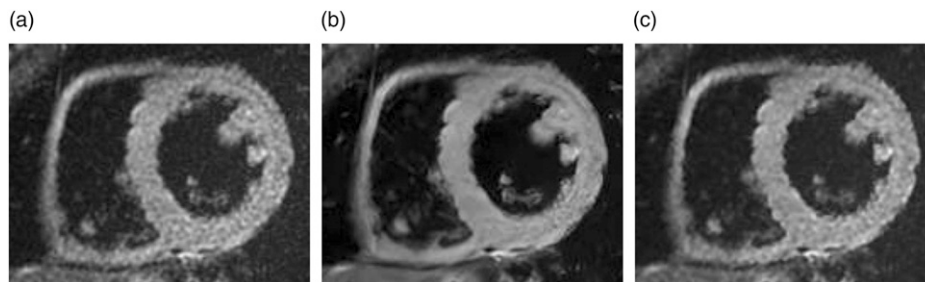
Figure 2. A case with (A) original image, (B) deep learning reconstruction, and (C) intensity filter. The deep learning reconstruction (strength 100%) showed better noise reduction than that of intensity filter.

Table 1. Signal to noise ratio, contrast ratio, and sharpness values in the original image, deep learning reconstruction, and intensity filter.

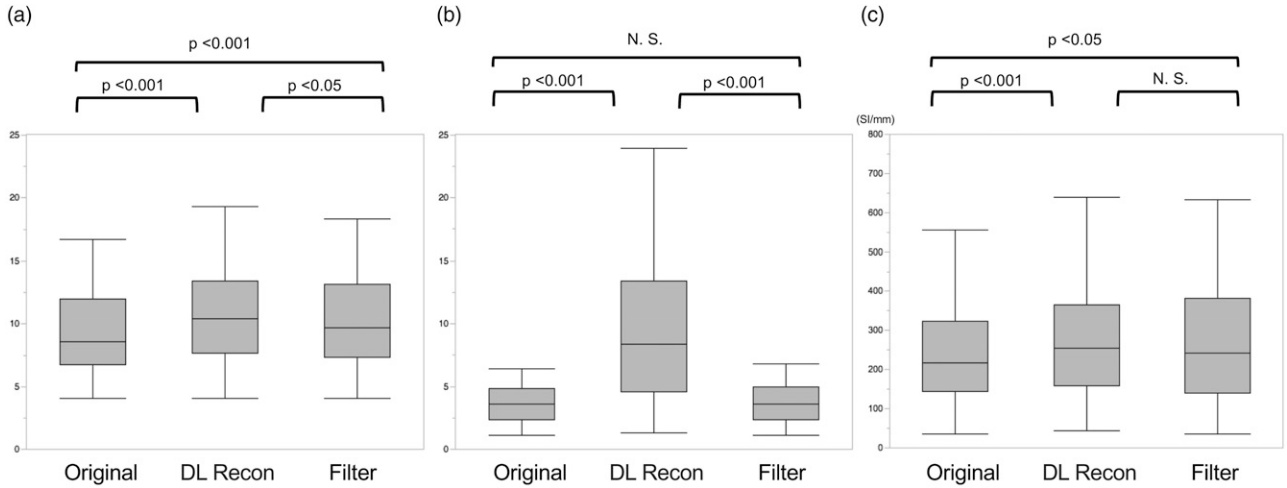|	Signal to noise ratio	Contrast ratio	Sharpness (SI/mm)
Original image	8.6 (6.8–12.0)	3.6 (2.4–4.9)	218.0 (144.1–322.9)
Deep learning reconstruction 25%	8.7 (12.1–6.9)	5.3 (3.6–7.3)	232.8 (156.2–365.6)
Deep learning reconstruction 50%	9.3 (12.5–7.2)	6.1 (4.0–8.8)	247.9 (158.6–355.0)
Deep learning reconstruction 75%	10.0 (13.1–7.5)	6.9 (4.3–11.1)	250.0 (154.9–358.4)
Deep learning reconstruction 100%	10.4 (7.7–13.5)	8.4 (4.6–13.4)	254.7 (157.8–366.0)
Intensity filter	9.7 (7.4–13.2)	3.6 (2.4–5.0)	243.0 (139.7–382.8)

Data are presented as median (first quartile, third quartile).
SI/mm: signal intensity/millimeter.

Discussion

This preliminary study showed the application of deep learning methods to cardiovascular T2-weight sequence can improve image quality. Regarding the resolution and acceleration factor of this study, the DL Recon with 100% strength showed a good performance, but the appropriate strength may change depending on the resolution or amount of noise.

The intensity filter used in this study appeared to affect not only noise but also structural details. Among similar smoothing filter, a popular technique utilizing a Gaussian

Figure 3. Quantitative values in the original image, deep learning reconstruction, and intensity filter. (A) signal to noise ratio, (B) contrast ratio, and (C) sharpness. DL Recon: deep learning reconstruction, N.S.: not significant, SI/mm: signal intensity/millimeter.
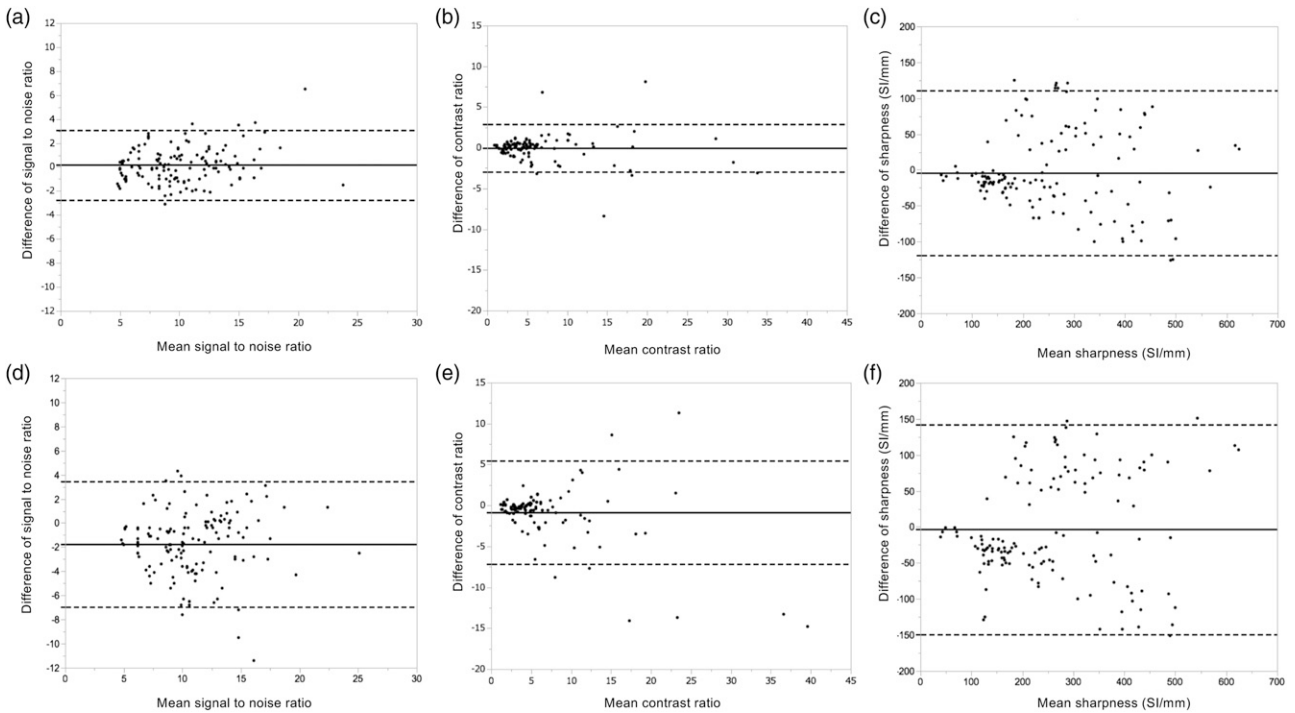
Figure 4. Bland–Altman plots for intra- and inter-observer agreement of the signal to noise ratio, contrast ratio, and sharpness. The mean bias (solid line) and 95% confidence intervals (dotted line) are shown. (A) Intra-observer agreement for signal to noise ratio, (B) intra-observer agreement for contrast ratio, (C) intra-observer agreement for sharpness, (D) inter-observer agreement for signal to noise ratio, (E) inter-observer agreement for contrast ratio, (F) inter-observer agreement for sharpness.

filter has shown by averaging local values to likely remove both noise and anatomical boundaries.[12] In contrast, the DL Recon could distinguish isolate noise by learning noise thresholds provided by high frequency components extracted from images.[12] In addition to noise reduction, this study showed DL Recon could improve contrast between myocardium and lumen better than the intensity filter. Since the cardiovascular T2-weighted image often causes poor contrast between myocardium and lumen due to noise or stagnant blood in the lumen,[13] it was possible that by effectively removing noise in the lumen, the DL Recon technique was able to enhance image.

The qualitative image quality of the DL Recon images was significantly better than that of intensity filtered and original images. Given that no noise was detected by eye in DL Recon images, these high quality and visually sharper edges could be attributed to increased contrast between myocardium and lumen.

In general, the novel image acquisition and reconstruction technique such as compressed sensing takes long time due to the need of a graphics processing unit. In this study, the reconstruction duration was approximately 50 s without using a graphics processing unit. Indeed, this would not only boost the clinical workflow but also lift any restrictions on hardware. Another study of the DL Recon also showed fast reconstruction speed, it was an advantage of DL Recon.[14]

This study has some limitations. First, the sample size was relatively small and only healthy individuals were included in this preliminary study. Since we used a prototype of DL Recon, we needed to verify the reliability of DL Recon images in healthy cases first. Improvement in myocardial noise, contrast, and qualitative evaluation were confirmed, but a comparison between normal and abnormal myocardium was not possible in healthy individuals. To reveal the clinical benefit of DL Recon, further patient studies are needed. Second, the peak SNR and structure similarity index were not evaluated in this study.[15] These metrics are often used to determine the noise reduction networks. However, there was difficulty in our study design to prepare additional noise images to compared with ground truth images. Third, the acceleration factor was set to 3. By incorporating parallel imaging with DL Recon,[16] we could take more accelerate and high-resolution images with finer image quality in the future.

In conclusion, the DL Recon reduced image noise and improved contrast and sharpness in the cardiovascular T2-weighted image. Compared with the conventional method of intensity filter, the deep learning–based technique proved superior in improving image quality.

Declarations of conflicting interest

Funding

ORCID iD

Ryo Ogawa ⓘ https://orcid.org/0000-0002-3261-2752

References

1. Eitel I, Friedrich MG. T2-weighted cardiovascular magnetic resonance in acute cardiac disease. J Cardiovasc Magn Reson 2011; 13: 13.

2. Keegan J, Gatehouse PD, Prasad SK, et al. Improved turbo spin-echo imaging of the heart with motion-tracking. J Magn Reson Imaging 2006; 24: 563–570.

3. Zhu Y, Yang D, Zou L, et al. T2STIR preparation for single-shot cardiovascular magnetic resonance myocardial edema imaging. J Cardiovasc Magn Reson 2019; 21: 72.

4. Cocker MS, Shea SM, Strohm O, et al. A new approach towards improved visualization of myocardial edema using T2-weighted imaging: a cardiovascular magnetic resonance (CMR) study. J Magn Reson Imaging 2011; 34: 286–292.

5. Yasaka K, Akai H, Kunimatsu A, et al. Deep learning with convolutional neural network in radiology. Jpn J Radiol 2018; 36: 257–272.

6. Leiner T, Rueckert D, Suinesiaputra A, et al. Machine learning in cardiovascular magnetic resonance: basic concepts and applications. J Cardiovasc Magn Reson 2019; 21: 61.

7. Jiang D, Dou W, Vosters L, et al. Denoising of 3D magnetic resonance images with multi-channel residual learning of convolutional neural network. Jpn J Radiol 2018; 36: 566–574.

8. Zhang K, Zuo W, Chen Y, et al. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. IEEE Trans Image Process 2017; 26: 3142–3155.

9. Kingma DP and Ba J. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980

10. Constantinides CD, Atalar E, McVeigh ER. Signal-to-noise measurements in magnitude images from NMR phased arrays. Magn Reson Med 1997; 38: 852–857.

11. Downs RK, Bashir MH, Ng CK, et al. Quantitative contrast ratio comparison between T1 (TSE at 1.5T, FLAIR at 3T), magnetization prepared rapid gradient echo and subtraction imaging at 1.5T and 3T. Quantitative Imaging Med Surg 2013; 3: 141–146.

12. Kidoh M, Shinoda K, Kitajima M, et al. Deep learning based noise reduction for brain MR imaging: tests on phantoms and healthy volunteers. Magn Reson Med Sci 2020; 19(3): 195–206.

13. Payne AR, Casey M, McClure J, et al. Bright-blood T2-weighted MRI has higher diagnostic accuracy than dark-blood short tau inversion recovery MRI for detection of acute myocardial infarction and for assessment of the ischemic area at risk and myocardial salvage. Circ Cardiovasc Imaging 2011; 4: 210–219.

14. Hammernik K, Klatzer T, Kobler E, et al. Learning a variational network for reconstruction of accelerated MRI data. Magn Reson Med 2018; 79: 3055–3071.

15. Ramzi Z, Ciuciu P and Starck J-L. Benchmarking MRI reconstruction neural networks on large public datasets. Appl Sci 2020; 10: 1816.

16. Sriram A, Zbontar J, Murrell T, et al. Grappanet: combining parallel imaging with deep learning for multi-coil MRI reconstruction. In: IEEE/CVF Conference on computer vision and pattern recognition (CVPR), Seattle, WA, 13-19 June 2020, pp. 14303–14310, 2020. IEEE.