


# Deep Learning Enables Reduced Gadolinium Dose for Contrast-Enhanced Brain MRI

Enhao Gong, MS <sup>1,2\*</sup> John M. Pauly, PhD,<sup>1</sup> Max Wintermark, MD,<sup>2</sup> and Greg Zaharchuk, PhD, MD<sup>2</sup>

**Background:** There are concerns over gadolinium deposition from gadolinium-based contrast agents (GBCA) administration.

**Purpose:** To reduce gadolinium dose in contrast-enhanced brain MRI using a deep learning method.

**Study type:** Retrospective, crossover.

**Population:** Sixty patients receiving clinically indicated contrast-enhanced brain MRI.

**Sequence:** 3D T<sub>1</sub>-weighted inversion-recovery prepped fast-spoiled-gradient-echo (IR-FSPGR) imaging was acquired at both 1.5T and 3T. In 60 brain MRI exams, the IR-FSPGR sequence was obtained under three conditions: precontrast, postcontrast images with 10% low-dose (0.01mmol/kg) and 100% full-dose (0.1 mmol/kg) of gadobenate dimeglumine. We trained a deep learning model using the first 10 cases (with mixed indications) to approximate full-dose images from the precontrast and low-dose images. Synthesized full-dose images were created using the trained model in two test sets: 20 patients with mixed indications and 30 patients with glioma.

**Assessment:** For both test sets, low-dose, true full-dose, and the synthesized full-dose postcontrast image sets were compared quantitatively using peak-signal-to-noise-ratios (PSNR) and structural-similarity-index (SSIM). For the test set comprised of 20 patients with mixed indications, two neuroradiologists scored blindly and independently for the three postcontrast image sets, evaluating image quality, motion-artifact suppression, and contrast enhancement compared with precontrast images.

**Statistical Analysis:** Results were assessed using paired t-tests and noninferiority tests.

**Results:** The proposed deep learning method yielded significant ( $n = 50$ ,  $P < 0.001$ ) improvements over the low-dose images ( $>5$  dB PSNR gains and  $>11.0\%$  SSIM). Ratings on image quality ( $n = 20$ ,  $P = 0.003$ ) and contrast enhancement ( $n = 20$ ,  $P < 0.001$ ) were significantly increased. Compared to true full-dose images, the synthesized full-dose images have a slight but not significant reduction in image quality ( $n = 20$ ,  $P = 0.083$ ) and contrast enhancement ( $n = 20$ ,  $P = 0.068$ ). Slightly better ( $n = 20$ ,  $P = 0.039$ ) motion-artifact suppression was noted in the synthesized images. The non-inferiority test rejects the inferiority of the synthesized to true full-dose images for image quality (95% CI:  $-14\%$ – $9\%$ ), artifacts suppression (95% CI:  $-5\%$ – $20\%$ ), and contrast enhancement (95% CI:  $-13\%$ – $6\%$ ).

**Data Conclusion:** With the proposed deep learning method, gadolinium dose can be reduced 10-fold while preserving contrast information and avoiding significant image quality degradation.

**Level of Evidence:** 3

**Technical Efficacy:** Stage 5

J. MAGN. RESON. IMAGING 2018;00:000–000.

Magnetic resonance imaging (MRI) is a powerful imaging technique providing unique information to distinguish different tissue properties and disorders. Additionally, MRI contrast agents with unique relaxation parameters are applied to further boost the visibility of pathology and delineation of lesions. Gadolinium-based contrast agents (GBCAs) are

widely used in MRI exams<sup>1,2</sup> and have been indispensable for monitoring treatment and investigating pathology, for applications such as angiography,<sup>3</sup> neuroimaging for multiple sclerosis<sup>4–7</sup> and tumors,<sup>8</sup> and liver imaging.<sup>9</sup>

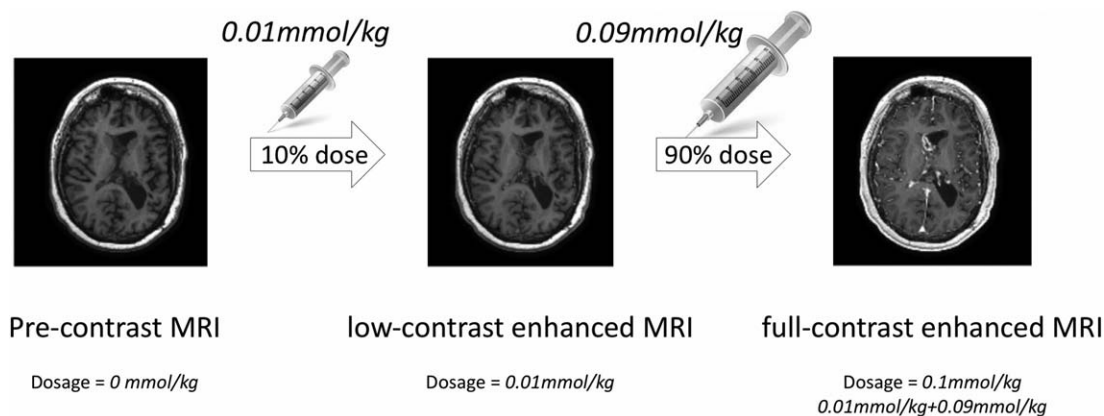
Recently, the identification of gadolinium deposition within the brain and body has raised safety concerns about

View this article online at [wileyonlinelibrary.com](http://wileyonlinelibrary.com). DOI: 10.1002/jmri.25970

Received Sep 13, 2017, Accepted for publication Jan 25, 2018.

\*Address reprint requests to: E.G., Department of Electrical Engineering, Stanford University, 1267 Lakeside Dr., Apt. 2082, Sunnyvale, CA, 94085.  
E-mail: [enhaog@stanford.edu](mailto:enhaog@stanford.edu)

From the <sup>1</sup>Department of Electrical Engineering, Stanford University, Stanford, California, USA; and <sup>2</sup>Department of Radiology, Stanford University, Stanford, California, USA



**FIGURE 1:** Detailed illustration of the protocol and procedure.  $T_1$ -weighted IR-FSPGR images are acquired under three different conditions: precontrast, post 10% contrast (low-dose), and post 100% contrast (full-dose).

GBCAs in the medical community.<sup>10–16</sup> Studies have demonstrated increased signal intensity on the unenhanced  $T_1$ -weighted MR images in the dentate nucleus (DN) and globus pallidus (GP) that is correlated with previous GBCA exposure<sup>17–20</sup> and the deposition is independent of renal function.<sup>21</sup> While initial reports focused on linear GBCAs, more recently it has been shown that deposition also occurs with macrocyclic GBCAs as well, albeit at lower levels.<sup>12,22</sup> Reducing the GBCA dose reduces the degree of deposition, which is especially relevant for patients (eg, multiple sclerosis patients) who need repeated contrast administration and are at the highest risk of gadolinium deposition. Studies have also shown that GBCA doses should be as low as possible, since a higher dose has been linked to nephrogenic systemic fibrosis (NSF) development.<sup>23</sup> In addition, by reducing the amount of GBCA administration, it may be possible to reduce the costs of contrast-enhanced MRI studies. Therefore, there are several reasons why it would be beneficial to reduce the dosage of GBCA while preserving the information provided by the contrast.

Recently, machine learning (ML) and deep learning (DL) methods have shown great potential in medical imaging applications such as reconstruction,<sup>24–27</sup> denoising,<sup>28</sup> and contrast synthesis.<sup>29</sup> The purpose of this work was to develop and validate a new DL-based technology to reduce GBCA dose levels while maintaining the image quality and contrast information of full-dose contrast images.

## Materials and Methods

### Clinical Dataset Acquisition With Different Contrast Enhancements

Studies were performed on six different 1.5T and 3T MRI scanners (MR750, MR750w, Signa HDx; GE Healthcare, Waukesha, WI) as shown in Fig. 1. At this hospital, patients evaluated for suspected or known enhancing brain abnormalities receive perfusion imaging with bolus contrast-enhanced dynamic susceptibility contrast (DSC). Before perfusion imaging, a 10% low-dose preload (0.01 mmol/kg, gadobenate dimeglumine; Multihance, Bracco,

Princeton, NJ) is used to reduce the effects of leakage on cerebral blood volume measurements.<sup>30</sup> A waiver of informed consent was granted by our Institutional Review Board to review postcontrast imaging of this preload dose.

In this study, all patients underwent 3D noncontrast axial  $T_1$ -weighted imaging using an inversion-recovery-prepped fast spoiled gradient echo (IR-FSPGR, also known as  $T_1$  BRAVO) with the following parameters: TR/TE/TI 9.5/3.8/400 msec; flip angle 13°; matrix  $512 \times 512$ , field-of-view 240 mm, ~300–350 slices reconstructed with 1 mm thickness and 0.5 mm offset. As part of a quality control study, we acquired the same sequence immediately following the low-dose preload to confirm compliance with the protocol. Following this, the patients received a standard DSC study with the remaining 90% of the dose (0.09 mmol/kg) followed by postcontrast 3D  $T_1$  IR-FSPGR with the same parameters. The patients received no additional gadolinium dosage or risk compared with a standard full-dose (0.1 mmol/kg) protocol. Detailed illustration of the protocol and procedure is shown in Figs. 1 and 2a.

### Subject Information

Sixty patients (49.7 age  $\pm$  17.7 years, 36 males, 24 females) with a wide range of clinical indications and brain abnormalities were included in this study. The detailed breakdown of age, genders, and broad classes of indication (extracted from clinical reports) is shown in Table 1.

### Data Preprocessing

The detailed workflow of the method is shown in Fig. 2b, which includes preprocessing steps applied to precontrast (0 mmol/kg), low-dose postcontrast (0.01 mmol/kg), and true full-dose postcontrast (0.1 mmol/kg) IR-FSPGR sequences for each subject. To remove the systematic differences between signal intensity levels in nonenhancing regions (such as scalp fat), efficient rigid coregistration using SimpleElastic software,<sup>31</sup> and signal normalization based on average voxel value within a mask was conducted. This step is required since the transmit and receive gains used for the sequences were determined from separate prescans, and thus were not guaranteed to be the same. The model training and prediction was conducted after coregistration and normalization, based on the assumption as shown in Fig. 3 that the contrast signal (remaining

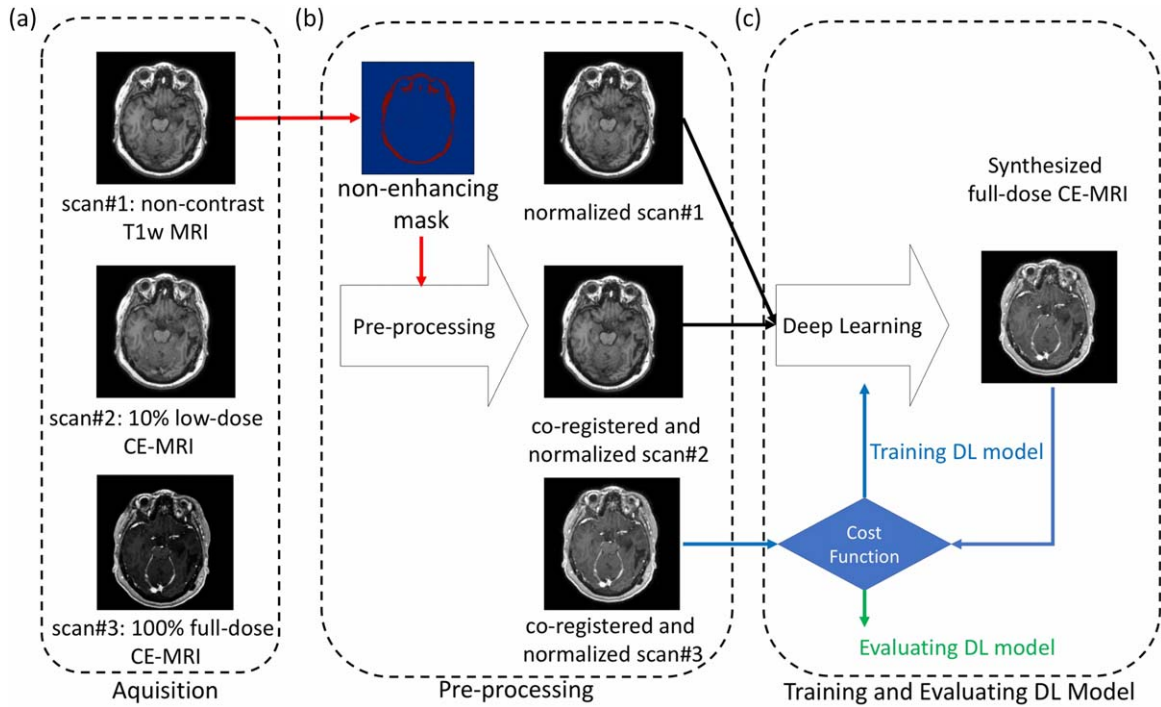


FIGURE 2: Detailed workflow of the process includes acquisition, preprocessing, and deep learning. (a) For acquisition, three different image sets are acquired. (b) Preprocessing steps include image coregistration and signal normalization. Normalization is required to remove bias in the images obtained at different dosage levels, which is performed using a mask of the nonenhancing tissues. This step is only required because of different transmit and receive gains for the different sequences. (c) Deep learning model is trained and evaluated.

TABLE 1. Clinical Subjects' Information of the Datasets With Mixed Indications, Grouped by Broad Classes of Clinical Indications

Clinical indications	Age	Patient number (male/female)	Patient number (train/test)	Scanner model (1.5T/3T)
Glioma	42.2 ± 16.9	8/5	4/9	4/9
Glioma <sup>a</sup>	50.8 ± 16.2	19/11	0/30	1/29
Lymphoma	62.0 ± 16.0	2/1	1/2	0/3
Meningioma	56.3 ± 17.0	1/1	2/0	0/2
Metastases	69.3 ± 6.2	1/2	1/2	0/3
Tumor, nonglioma	54.0 ± 16.8	1/2	2/1	1/2
Unknown	33.0 ± 26.0	1/1	1/1	0/1
Arachnoid cyst	37.0	1/0	0/1	1/0
Hydrocephalus	50.0	1/0	0/1	0/1
Infection	56.0	0/1	1/0	0/1
Vascular malformation	34.0	1/0	0/1	0/1
Total	49.7 ± 17.7	36/24	10/50	7/53

Training cases and testing cases cover a wide range of indications and pathologies, as well as different patient age, genders, and scanner models, to validate the generalization of the proposed technique. There were 10 patients with mixed indications involved in training, a different group of 20 patients with mixed indications involved in testing (both quantitative and qualitative), and a separate group of 30 glioma patients involved in additional testing (quantitative only).

<sup>a</sup>Separate from the first set of glioma patients. These 30 glioma patients were selected for additional quantitative evaluation.

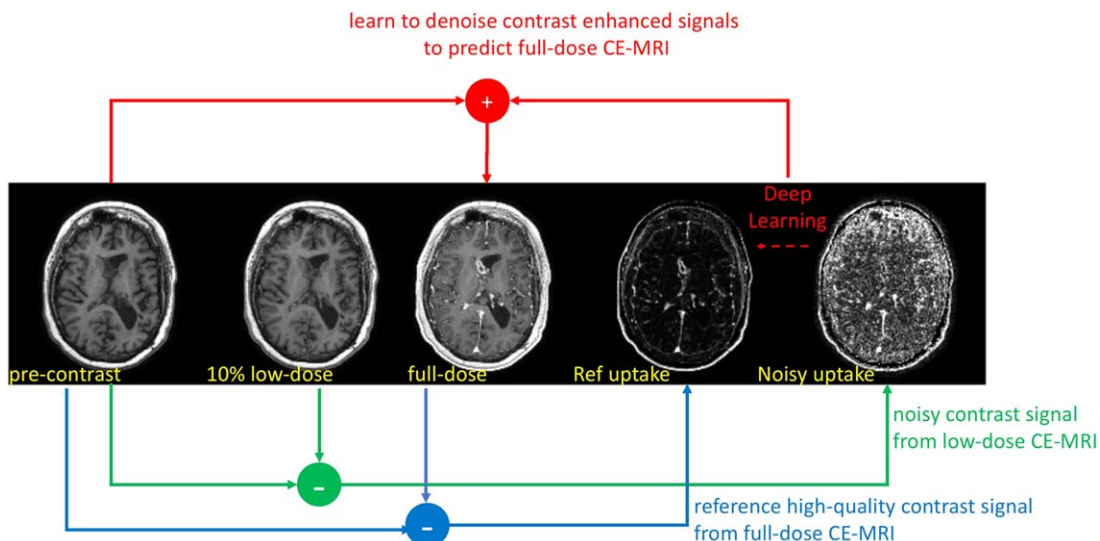


FIGURE 3: Illustration of the signal model to synthesize full-dose contrast-enhanced MRI. Compared with the reference high-quality contrast uptake between precontrast and full-dose CE-MRI, the contrast uptake in the low-dose CE-MRI is noisy but does include contrast information. By using a deep learning method, we learn the denoising to generate high-quality contrast uptake and then combine this with the precontrast scan to synthesize a full-dose CE-MRI.

signal differences) between low-dose CE-MRI and precontrast MRI is a well aligned and scaled, but a noisier version of the contrast signal between full-dose CE-MRI and precontrast MRI.

### Training the DL Model

Using the true 100% full-dose CE-MRI as the ground truth, we trained a deep network to take the postprocessed zero-dose precontrast MRI and the 10% low-dose postcontrast CE-MRI as inputs. The output of the model approximates the full-dose postcontrast CE-MRI. For patients not included in training, the final trained model can be directly applied to synthesize the full-dose images for evaluation. The basic idea of how the DL model works is shown in Figs. 2c and 3.

During training, this network implicitly learns the guided denoising of the noisy contrast uptake extracted from the difference signal between low-dose and precontrast (zero-dose) images, which can be scaled to generate the contrast enhancement of full-dose images.

The detailed DL model architecture is shown as in Fig. 4. It is an encoder-decoder convolutional neural network (CNN) with bypass connections and residual connections. This type of deep network architecture has demonstrated superior performance in other medical imaging applications such as segmentation<sup>32</sup> and MRI reconstruction.<sup>33,34</sup>

We trained the DL model on the first 10 patients with mixed indications (training set), and then tested its performance

### Deep learning model architecture

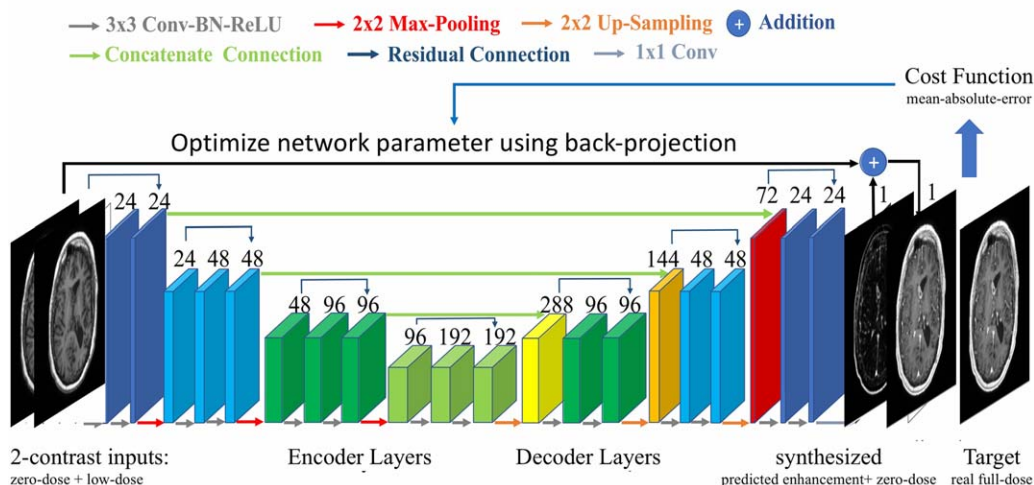


FIGURE 4: Deep learning network architecture used in this work. This model is an encoder-decoder convolutional neural network with three encoder steps and three decoder steps. In each step, there are three convolutional layers. Bypass concatenate connections combine symmetric layers to avoid resolution loss. The residual connections enable the model to synthesize a full-dose image by predicting the enhancement signal. The cost function compares the prediction and the reference ground-truth full-dose image, which enables the optimization of the network parameters.



on 20 patients with mixed indications. In addition, we conducted quantitative evaluation of the model on a separate dataset with 30 patients with known gliomas, none of whom were part of the training set, to see if performance varies for a specific clinical indication. Table 1 shows further details about the cases included in each set. For each patient in the training set, we used the 300–350 2D slices of the coregistered 3D volumes, excluding the five slices at the base and five slices at the top of the acquired volume due to their low signal-to-noise ratio (SNR). Standard DL data augmentation techniques, consisting of eight types of transformation (incremental 90° rotations and flips), were applied to each case in the training set to avoid overfitting and ensure the robustness of the model.

During training of the DL model, stochastic gradient descent (SGD) and back-projection were used to optimize the network parameters with respect to a cost function comparing predicted and true full-dose MR images. In this work, the mean-absolute-error (MAE) cost function, also known as the  $L_1$  loss, was used. Training was performed with 200 epochs of SGD and ADAM<sup>35</sup> for optimization. 10% of the training set with random permutations were used for validation to optimize hyperparameters and choose the best model among all iterations.

### Applying the Trained DL Model

For the patients in the test sets, we applied the trained model to the coregistered and normalized precontrast (zero-dose) and low-dose postcontrast images as inputs. The forward-pass of the model applied to this input data was used to produce the synthesized full-dose images. Creation of each synthesized full-dose 512-by-512 image took around 0.1 sec to calculate. This process was then repeated for each slice to generate the entire 3D volume and was stored in a new DICOM folder for further evaluation. For both training and testing the neural network, we used the Keras framework with Tensorflow backend, CUDA8 and CUDNN5.1, on a Linux server with 2 NVIDIA GTX 1080-TI GPUs.

### Performance Evaluation and Statistical Analysis

To evaluate the performance of the proposed method, we conducted and analyzed several assessments, including both quantitative metrics (for all 50 patients in two test sets) and qualitative ratings (only for the test sets comprised of the 20 patients with mixed indications) from neuroradiologists. Additionally, we conducted subjective visual assessment of the cases to characterize possible failure mechanisms.

For quantitative and nonsubjective measurement of similarity, the peak SNR (PSNR) and structural similarity index (SSIM) were used, where PSNR measures voxelwise differences (errors) while SSIM measures nonlocal structural similarity. The metrics were computed for both low-dose images and the synthesized full-dose images on 50 patients compared to the ground-truth (the true full-dose CE-MRI) to show the performance of the proposed DL method.

To assess subjective quality, two board-certified neuroradiologists (G.Z. and M.W. with 11 years' and 13 years' experience, respectively) separately and blindly rated three different postcontrast CE-MRI volumes (low-dose, synthesized full-dose and true full-dose) for the 20 patients with various clinical indications.

Specifically, the readers were independently presented the precontrast MRI and a randomized "version" of postcontrast CE-MRI and asked to rate their quality, which mimics the clinical practice of reading CE-MRI studies. For each case evaluated, three detailed ratings were performed: general image quality, suppression of aliasing/motion artifacts, and the degree of the enhancement compared against precontrast MRI, using a five-point Likert scale ranging from 1 (poor) to 5 (excellent). To assess differences,  $t$ -tests based on the average ratings from both neuroradiologists were used. The image ratings were dichotomized into low (1–3) and high (4–5) ratings, and the percentage of high ratings was calculated for each image type and each criteria. Noninferiority tests of the synthesized and true full-dose images were performed, using Stata 15.1 (Stata-Corp, College Station, TX) and R v. 3.3.1 (r-project.org) with v. 1.3 of the ExactCidiff package. The 95% confidence interval (CI) for the proportion of high ratings was calculated and compared to a lower bound interval with a noninferiority margin of –15%. This test evaluates (with a significance level of  $P = 0.05$ ) whether the proportion of high ratings for synthesized images was no more than 15 percentage points lower than that for true full-dose images.

In addition, subjective visual and free-text assessments of the images were reviewed to identify any common failure modes, to better characterize and understand the limitations of the proposed technique.

## Results

### Quantitative Comparison Evaluated With Similarity Metrics

As shown by the detailed metrics summarized in Table 2, the proposed method achieves significantly superior performance in approximating full-dose contrast. For the 50 testing cases, the synthesized CE-MRI images were improved significantly compared with the low-dose image on both quantitative metrics, with greater than 11.0% improvements in SSIM ( $P < 0.001$ ) and greater than 5.0 dB gain in PSNR ( $P < 0.001$ ). These improvements in PSNR and SSIM were similar regardless of whether the test dataset was comprised of the 20 patients with mixed indications or of the 30 glioma patients.

### Diagnostic Quality Evaluated From Neuroradiologists' Ratings

Detailed qualitative metrics rated by the two neuroradiologists are also shown in Table 3. As shown in Fig. 5a, the synthesized images led to significant improved scores compared with true low-contrast images ( $P < 0.001$ ). There were no significant differences in overall image quality or perceived contrast enhancement for the synthesized images compared with the true full-dose images, although there was a trend towards slightly lower (but not significant) scores for the synthesized images on overall image quality ( $P = 0.083$ ) and enhancement ( $P = 0.068$ ). The synthesized full-dose images were rated better (significant,  $P = 0.039$ ) for the suppression of motion artifacts, and the magnitude of this

**TABLE 2. Subject Information and Evaluations Using Quantitative Nonsubjective Similarity Metrics (Compared Against Full-Dose Ground-Truth) on Testing Datasets**

Testing dataset	Patient number (male/female)	Ages	Contrast	PSNR(dB)	SSIM
#1: Mixed indications	20 (12/8)	45.3 ± 18.2	Low-dose	22.97 ± 1.44	0.75 ± 0.07
			Synthesized	<b>27.99</b> ± 2.35	<b>0.84</b> ± 0.06
#2: Glioma patients	30 (19/11)	50.8 ± 16.2	Low-dose	23.03 ± 1.23	0.76 ± 0.08
			Synthesized	<b>28.12</b> ± 2.18	<b>0.86</b> ± 0.08
Total	50 (31/19)	48.6 ± 17.0	Low-dose	23.01 ± 1.30	0.76 ± 0.08
			Synthesized	<b>28.07</b> ± 2.26	<b>0.85</b> ± 0.07

Testing Dataset #1: 20 patients test set with mixed indications.

Testing Dataset #2: 30 patients test set with glioma for additional evaluation

Quantitative nonsubjective similarity metrics: PSNR, peak signal-to-noise-ratio; SSIM, structural similarity index.

improvement was moderate (0.46 out of 5-point scale). This was most noticeable in the nonenhancing regions.

For the noninferiority tests, the 95% CI for the difference was calculated with a noninferiority margin for the difference of 15% (ie, the percentage of high ratings (4-5) for the synthesized images being no worse than 15% points lower than the percentage of high ratings for the true full-dose). As shown in Fig. 5b, the lower bounds of the CIs were all above -15%; therefore, the null hypothesis of inferiority of the synthesized to true full-dose images was rejected at the  $P=0.05$  level for image quality (95% CI: -14-9%), artifacts suppression (95% CI -5-20%) and contrast enhancement (95% CI: -13-6%).

### Visual Assessment and Characterization of Failure Mechanisms

Detailed examples in several representative cases are shown in Figs. 6-8. When images are shown side-by-side, restoration of the contrast signal is clearly noted on the synthesized full-dose images, although the enhancement is slightly less conspicuous and enhancement in small vessels is underestimated compared with the true full-dose images.

We did identify several artifacts related to the proposed DL method. We observed large differences between the true full-dose images and the synthesized full-dose images within the arteries near the base of the brain, as shown in Fig. 9a. In addition, Fig. 9b shows a case where suboptimal coregistration causes the spurious appearance of contrast enhancement on the synthesized full-dose images. In total, we identified four minor local artifacts related to suboptimal coregistration in four cases of the 20 test cases.

### Discussion

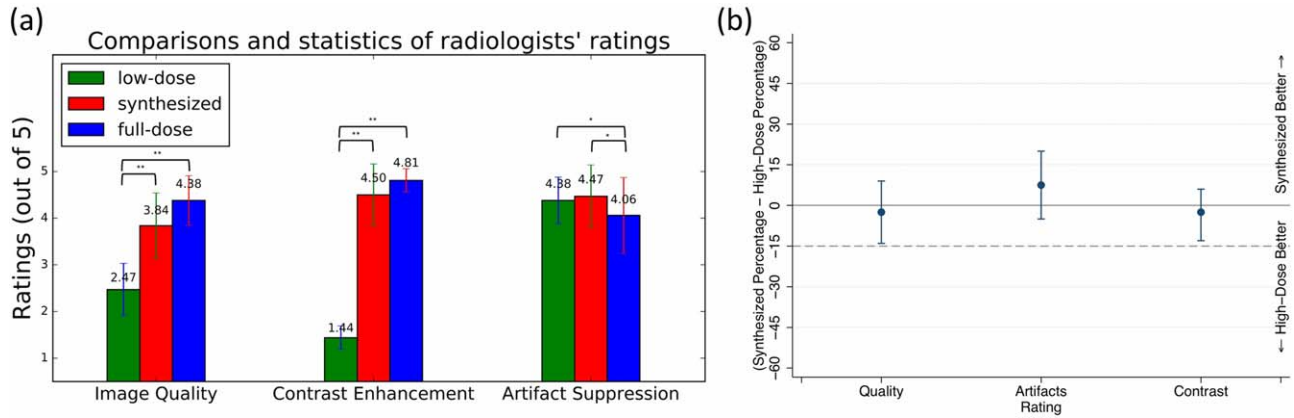
We have described, implemented, and validated a DL framework for improving the quality of CE-MRI using gadolinium doses 10-fold lower than typically used. In this study, we demonstrate that the DL method can successfully extract the contrast signal from low-dose gadolinium-enhanced images and significantly improve the image quality and contrast enhancement from such images. To be clear, simply amplifying 10× the contrast enhancement of the 10% low-dose CE-MRI without advanced algorithms is not a viable approach, as it results in severely degraded image quality.

**TABLE 3. Comparison and Statistics of the Quantitative Ratings (Averaged Scores Among Two Radiologists) on Three Types of CE-MRI (Acquired Low-Dose Images, Synthesized Full-Dose Images, and True Full-Dose Images)**

Contrast	Dose	IQ	CE	AS
Low-dose	0.01 mmol/kg	2.47 ± 0.56	1.44 ± 0.25	<b>4.38</b> ± 0.50
Synthesized	0.01 mmol/kg	<b>3.84</b> ± 0.70	<b>4.50</b> ± 0.66	<b>4.53</b> ± 0.64
Full-dose	0.1 mmol/kg	<b>4.38</b> ± 0.53	<b>4.81</b> ± 0.25	4.06 ± 0.81

The acquired or synthesized scans that are rated the best among the other alternatives, or not significantly different ( $P>0.05$ ) from the best one, are shown with red and bold highlights.

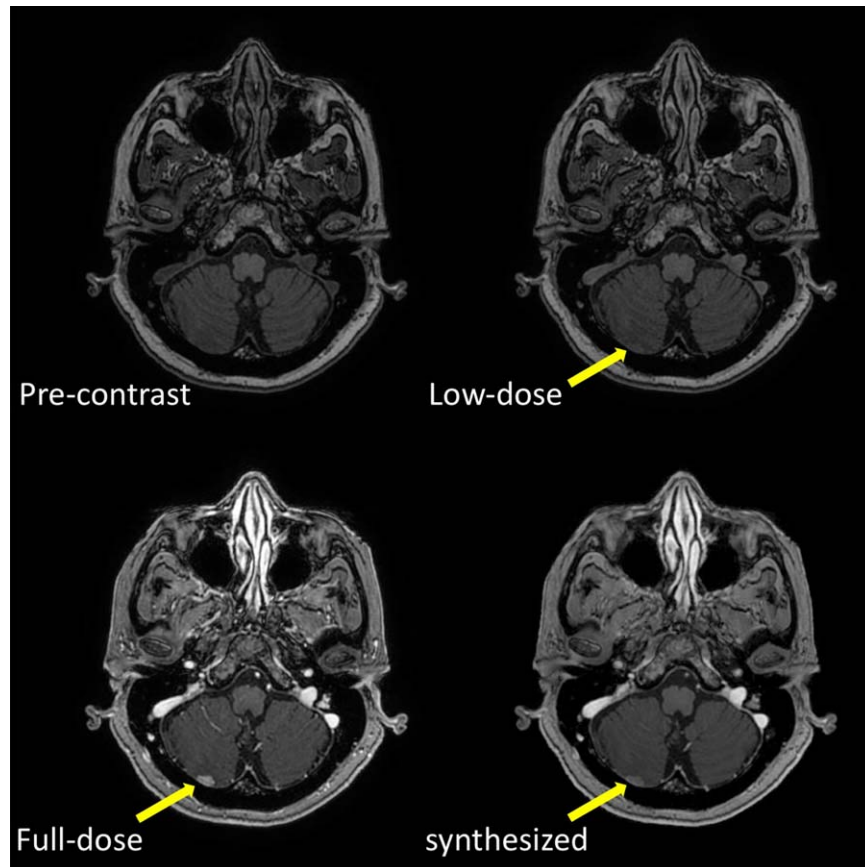
Qualitative ratings: IQ, overall image quality; CE, clarity of contrast enhancement; AS, artifact suppression.



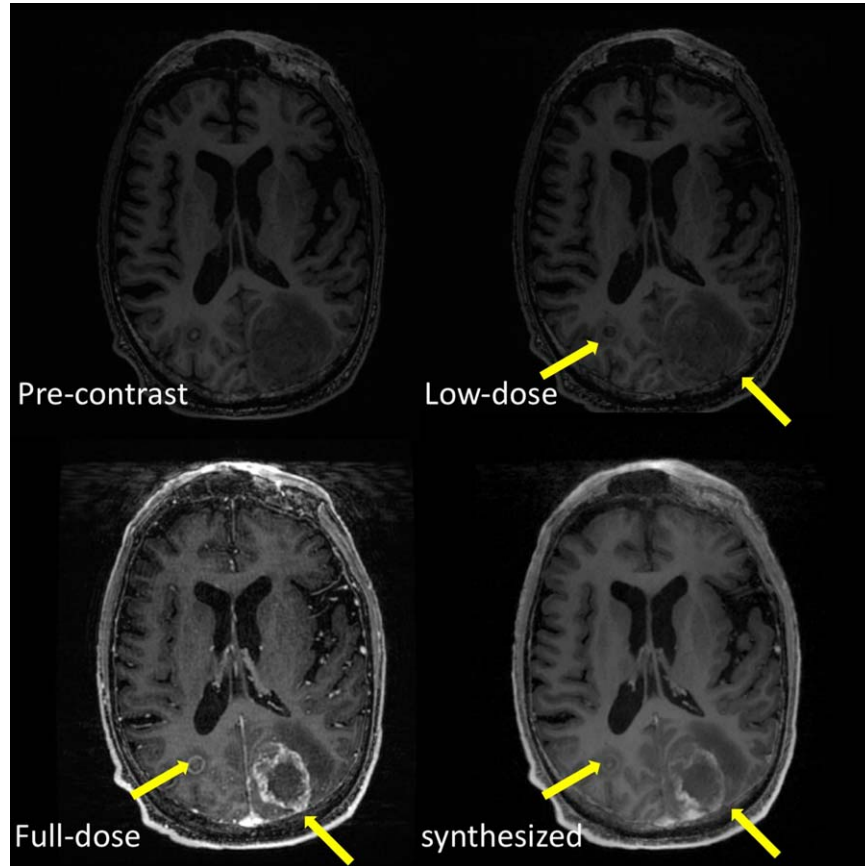
**FIGURE 5: (a)** Comparison and statistics of two radiologists' average ratings on different CE-MRI series (low-dose, synthesized, and true full-dose). The synthesized images show significant improvement compared with the acquired low-dose images, while there is no significant difference compared with the true full-dose images for diagnostic contrast enhancement. The synthesized full-dose images also show slight but significantly better artifact suppression. Significant differences are marked with \* ( $P < 0.05$ ) or \*\* ( $P < 0.005$ ). **(b)** The statistics from noninferiority test is shown. The lower bounds of 95% CI's of the difference in high rating percentages, between the synthesized and true full-dose images, are above the -15% margin across all three criteria.

From the evaluation results based on both nonsubjective quantitative metrics and subjective qualitative assessment from two neuroradiologists, the proposed method leads to significant improvement over the acquired low-dose images. The noninferiority test rejects

the inferiority of the synthesized images to the true full-dose images by a noninferiority margin of -15%, suggesting the synthesized full-dose imaging does not lead to significant quality degradation. The model was evaluated on patients with various age, genders, and indications,



**FIGURE 6:** For a patient with intracranial metastatic disease, the synthesized images result in similar highlighting of contrast enhancement in the lesions. The lesions show improved visibility in the synthesized full-contrast version while they cannot be reliably appreciated in low-dose CE-MRI. Moreover, the synthesized CE-MRI shows a similar outline of a metastatic lesion in the right posterior cerebellum (yellow arrow) compared with the true full-dose CE-MRI.



**FIGURE 7:** Results for a patient with a glioma, in which the synthesized and true full-dose CE-MRI shows similar contrast uptake. Additionally, the synthesized image demonstrates improved suppression of motion artifacts and aliasing compared with the true full-dose CE-MRI.

demonstrating that the model generalizes across different conditions.

Both the qualitative reader study and visual assessment show the proposed synthesized full-dose CE-MRI results in small but significant improvements in motion artifact/aliasing suppression. This surprising result may arise as the synthesized full-dose image is estimated from images with lower peak intensities and fewer motion artifacts. Further evaluation of the causes of the suppressed motion artifacts will be included in future studies with more controlled prospective experiments and protocols.

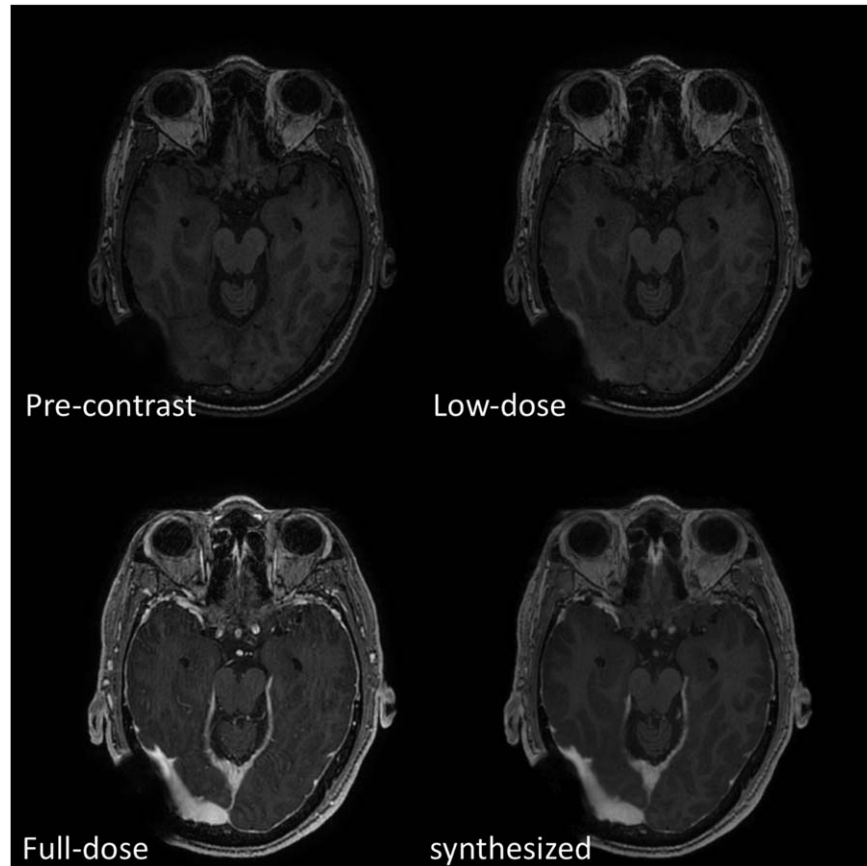
The impact of reducing the contrast dose by 90% while retaining diagnostic information could have a large impact on patient well-being and imaging costs. The proposed technique is based on the assumption that the contrast uptake in low-dose CE-MRI is a well-aligned, scaled, but noisier version of the contrast uptake in full-dose CE-MRI. This assumption is valid in most cases; however, we did see several artifacts related to the DL method, specifically caused by inflow effects and imperfect coregistration.

First, artificial enhancement within arteries may be generated due to inflow effects. We identified diminished contrast levels in arteries at the base of the brain on the

synthesized full-dose images. This difference is due to the nonlinear signal intensity changes generated by inflow effects in the arteries, which is more obvious for the  $T_1$  imaging in the lower slices and at a lower contrast level. At the 10% low-contrast level, the inflow effect competes with the contrast enhancement, leading to the underestimation of enhancement in the synthesized images. Since this effect only occurs in vascular structures with a clear spatial dependency and does not affect the identification of enhancing parenchymal lesions, it would likely have minimal impact on clinical evaluation. However, we will continue to investigate ways to mitigate or eliminate this for future clinical applications. For example, this artifact could be possibly reduced with improved masking or modified CE-MRI sequences that are less sensitive to inflow.

Second, imperfect coregistration can cause local artifacts. We witnessed potential errors resulting from both in-slice and through-slice misregistration of the three high-resolution  $T_1$  series. While we did perform coregistration as part of the study, there were still four instances in which incorrect contrast synthesis related to imperfect coregistration still occurred in a local region. This issue could be solved or mitigated using more sophisticated coregistration techniques, including potentially nonrigid coregistration.



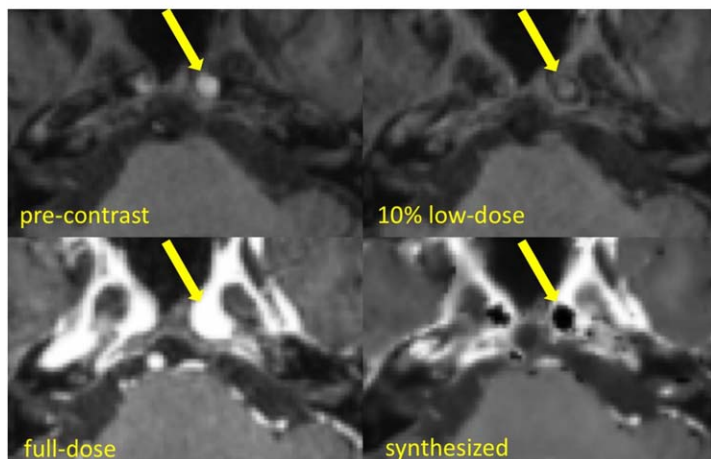


**FIGURE 8:** In a patient with a programmable shunt and intracranial hypotension, the proposed method generates contrast with diagnostic quality from the low-dose acquisition, also demonstrating improved noise suppression and contrast enhancement. The engorgement of the dural sinuses and pachymeningeal enhancement are clearly seen on both synthesized and true full-dose images, while there is less noise in nonenhancing regions on the synthesized image.

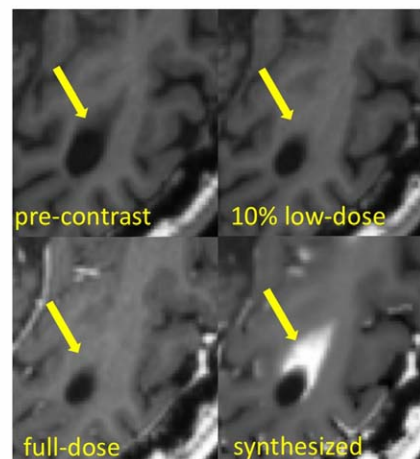
This artifact mechanism, on the other hand, reflects how the model generates the synthesized contrast, confirming that the network model does not hallucinate arbitrary

anatomical structures seen in training datasets but relies on augmenting the subtle contrast differences between the pre-contrast and low-dose postcontrast scans.

**(a) in-flow artifacts**



**(b) suboptimal co-registration**



**FIGURE 9:** (a) A zoomed-in example of in-flow artifacts causing reduced contrast enhancement on the low-dose images in arteries at the base of the brain, which results in underestimation of true contrast enhancement in the synthesized full-contrast images. (b) A zoomed-in example of an artifact related to suboptimal coregistration. The method mistakes the coregistration error as contrast change and incorrectly highlights it on the synthesized full-dose images. This issue can be addressed with slower but more accurate coregistration and methods for estimating different resolutions.

There are several limitations of this study. First, a relatively small cohort was included for training and testing the model. In comparison, recent studies<sup>36,37</sup> used around 300 exams to train models for predicting image quality and detecting aneurysms. However, there are intrinsic differences between the models generating images and the models generating binary classification, since the former trains on more samples per patient. In the presented work, each case presented provides more than 300 thin slices per volume, tens-of-thousands of local samples per image, and eight different versions of each using data augmentation for training. Thus, the number of training samples used is similar to existing DL studies on improving MRI reconstruction<sup>25,26</sup> and denoising.<sup>28</sup> The best evidence that the amount of training data is sufficient to train the model is its performance on the test set of new patients not included in training, as was done in the current study. However, to evaluate performance in a wider range of pathologies, it would be helpful to include a larger number of patients for both training and evaluation in future studies. With larger datasets, we can potentially train more complicated models to further improve the performance. In addition, separate models can potentially be fine-tuned for different clinical applications from a model trained on a broad population of cases, a method known as transfer-learning. Second, the work is based on a retrospective study. A prospective trial would avoid any possible bias in case selection, although we do not believe that such a bias exists in our series. In addition, the amount of time for contrast uptake was not controlled in our study. The reference 100% contrast dose images in this study are a combination of a delayed 10% predose and standard-delay imaging with the remaining 90% of the dose, which may result in slightly different enhancement patterns compared with a single standard-delay 100% dose. This could be improved by performing a 2-day protocol, with low and normal dose images acquired with well-controlled postcontrast timing, although this would require giving the patients a cumulative higher dose (by 10% more than the standard dose). However, such a trial would be valuable for definitive model training and evaluation.

The proposed method demonstrates positive results for approximating full-dose CE-MRI from low-dose CE-MRI, and several additional improvements are possible. First, other MRI sequences can be included in the model, which can then use these complementary contrasts to improve the prediction of the contrast enhancement. Second, in this work we evaluated the performance of the algorithm using a 2D CNN model. As has been shown in other DL applications, further performance gains are achievable with 3D models.<sup>38</sup> For example, there is still imperfect enhancement synthesized in small vessels. Although this does not likely affect the identification of enhancing parenchymal lesions, improved 3D models may better distinguish small vessels

from noise to further refine the proposed technique. Finally, here we used the MAE loss as cost function to train the DL model, which can be further improved by using multiple loss functions<sup>39</sup> or using Generative Adversarial Network (GAN) as a data-driven cost function for reconstruction.<sup>26</sup>

In this work we used a fixed level of gadolinium contrast with 90% reduction from the original clinical usage level. More validation is helpful to determine and justify the choice of the level of dose reduction recommended for different applications and for imaging different parts of body. It is also reasonable to explore the feasibility of further dose reduction and performance improvement, by incorporating improved DL models, using larger training datasets with various pathologies as well as other complementary information such as multicontrast MRI, multiple imaging modalities, and clinical history information.

In conclusion, while larger datasets, better experimental design, and more advanced models may lead to further improvements, the current study suggests that DL is a promising solution to reduce gadolinium dose used for brain MRI exams while preserving image quality and avoiding significant degradation in contrast enhancement, which could be of great benefit to patients.

## References

1. Brasch RC, Weinmann HJ, Wesbey GE. Contrast-enhanced NMR imaging: animal studies using gadolinium-DTPA complex. *Am J Roentgenol* 1984;142:625–630.
2. Carr DH, Brown J, Bydder GM, et al. Gadolinium-DTPA as a contrast agent in MRI: initial clinical experience in 20 patients. *Am J Roentgenol* 1984;143:215–224.
3. Burrell M, Llovet JM, Ayuso C, et al. MRI angiography is superior to helical CT for detection of HCC prior to liver transplantation: an explant correlation. *Hepatology* 2003;38:1034–1042.
4. Kappos L, Moeri D, Radue EW, et al. Predictive value of gadolinium-enhanced magnetic resonance imaging for relapse rate and changes in disability or impairment in multiple sclerosis: a meta-analysis. *Lancet* 1999;353:964–969.
5. Silver NC, Good CD, Barker GJ, et al. Sensitivity of contrast enhanced MRI in multiple sclerosis. Effects of gadolinium dose, magnetization transfer contrast and delayed imaging. *Brain* 1997;120:1149–1161.
6. McFarland HF, Frank JA, Albert PS, et al. Using gadolinium-enhanced magnetic resonance imaging lesions to monitor disease activity in multiple sclerosis. *Ann Neurol* 1992;32:758–766.
7. Miller DH, Barkhof F, Nauta JJ. Gadolinium enhancement increases the sensitivity of MRI in detecting disease activity in multiple sclerosis. *Brain* 1993;116:1077–1094.
8. Zahra MA, Hollingsworth KG, Sala E, Lomas DJ, Tan LT. Dynamic contrast-enhanced MRI as a predictor of tumour response to radiotherapy. *Lancet Oncol* 2007;8:63–74.
9. Semelka RC, Martin DR, Balci C, Lance T. Focal liver lesions: comparison of dual-phase CT and multisequence multiplanar MR imaging including dynamic gadolinium enhancement. *J Magn Reson Imaging* 2001;13:397–401.
10. Boyd AS, Zic JA, Abraham JL. Gadolinium deposition in nephrogenic fibrosing dermopathy. *J Am Acad Dermatol* 2007;56:27–30.
11. McDonald RJ, McDonald JS, Kallmes DF, et al. Intracranial gadolinium deposition after contrast-enhanced MR imaging. *Radiology* 2015;275:772–782.

12. Murata N, Gonzalez-Cuyar LF, Murata K, et al. Macrocyclic and other non-group 1 gadolinium contrast agents deposit low levels of gadolinium in brain and bone tissue: preliminary results from 9 patients with normal renal function. *Invest Radiol* 2016;51:447–453.
13. McDonald RJ, McDonald JS, Kallmes DF, et al. Gadolinium deposition in human brain tissues after contrast-enhanced MR imaging in adult patients without intracranial abnormalities. *Radiology* 2017: 161595.
14. Tibussek D, Rademacher C, Caspers J, et al. Gadolinium brain deposition after macrocyclic gadolinium administration: a pediatric case-control study. *Radiology* 2017:161151.
15. Gulani V, Calamante F, Shellock FG, Kanal E, Reeder SB. Gadolinium deposition in the brain: summary of evidence and recommendations. *Lancet Neurol* 2017 31;16:564–570.
16. Malayeri AA, Brooks KM, Bryant LH, et al. National Institutes of Health perspective on reports of gadolinium deposition in the brain. *J Am Coll Radiol* 2016;13:237–241.
17. Kanda T, Ishii K, Kawaguchi H, Kitajima K, Takenaka D. High signal intensity in the dentate nucleus and globus pallidus on unenhanced T1-weighted MR images: relationship with increasing cumulative dose of a gadolinium-based contrast material. *Radiology* 2013;270:834–841.
18. Errante Y, Cirimele V, Mallio CA, Di Lazzaro V, Zobel BB, Quattrocchi CC. Progressive increase of T1 signal intensity of the dentate nucleus on unenhanced magnetic resonance images is associated with cumulative doses of intravenously administered gadodiamide in patients with normal renal function, suggesting dechelation. *Invest Radiol* 2014;49:685–690.
19. Ramalho J, Castillo M, AIObaidy M, et al. High signal intensity in globus pallidus and dentate nucleus on unenhanced T1-weighted MR images: evaluation of two linear gadolinium-based contrast agents. *Radiology* 2015;276:836–844.
20. Robert P, Lehericy S, Grand S, et al. T1-weighted hypersignal in the deep cerebellar nuclei after repeated administrations of gadolinium-based contrast agents in healthy rats: difference between linear and macrocyclic agents. *Invest Radiol* 2015;50:473.
21. Kanda T, Fukusato T, Matsuda M, et al. Gadolinium-based contrast agent accumulates in the brain even in subjects without severe renal dysfunction: evaluation of autopsy brain specimens with inductively coupled plasma mass spectroscopy. *Radiology* 2015;276:228–232.
22. Bjørnerud A, Vatnehol SA, Larsson C, Due-Tønnessen P, Hol PK, Groote IR. Signal enhancement of the dentate nucleus at unenhanced MR imaging after very high cumulative doses of the macrocyclic gadolinium-based contrast agent gadobutrol: an observational study. *Radiology* 2017;285:434–444.
23. Khawaja AZ, Cassidy DB, Al Shakarchi J, et al. Revisiting the risks of MRI with Gadolinium based contrast agents—review of literature and guidelines. *Insights Imaging* 2015;6:553–558.
24. Golkov V, Dosovitskiy A, Sperl JI, et al. q-Space deep learning: twelve-fold shorter and model-free diffusion MRI scans. *IEEE Trans Med Imaging* 2016;35:1344–1351.
25. Hammernik K, Klatzer T, Kobler E, et al. Learning a variational network for reconstruction of accelerated MRI data. *arXiv preprint arXiv:1704.00447*. 2017.
26. Mardani M, Gong E, Cheng JY, et al. Deep generative adversarial networks for compressed sensing automates MRI. *arXiv preprint arXiv:1706.00051*. 2017.
27. Zhu B, Liu JZ, Rosen BR, Rosen MS. Image reconstruction by domain transform manifold learning. *arXiv preprint arXiv:1704.08841*. 2017.
28. Gong E, Pauly JM, Zaharchuk G. Boosting SNR and/or resolution of arterial spin label (ASL) imaging using multi-contrast approaches with multi-lateral guided filter and deep networks. In: *Proc 25th Meeting ISMRM, Honolulu*; 2017. 3983.
29. Bahrami K, Shi F, Rekik I, Shen D. Convolutional neural network for reconstruction of 7T-like images from 3T MRI using appearance and anatomical features. In: *International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*. Berlin: Springer International Publishing. 2016. 39–47.
30. Boxerman JL, Prah DE, Paulson ES, et al. The role of preload and leakage correction in gadolinium-based cerebral blood volume estimation determined by comparison with MION as a criterion standard. *Am J Neuroradiol* 2012;33:1081–1087.
31. Marstal K, Berendsen F, Staring M, Klein S. 2016. SimpleElastix: A user-friendly, multi-lingual library for medical image registration. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 134–142.
32. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham, Switzerland: Springer; 2015. 234–241.
33. Lee D, Yoo J, Ye JC. Deep artifact learning for compressed sensing and parallel MRI. *arXiv preprint arXiv:1703.01120*. 2017.
34. Gong E, Pauly JM, Zaharchuk G. Improving the PI+CS reconstruction for highly undersampled multi-contrast MRI using local deep network. In: *Proc 25th Scientific Meeting ISMRM, Honolulu*; 2017. 5663.
35. Kingma D, Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. 2014.
36. Esses SJ, Lu X, Zhao T, et al. Automated image quality evaluation of T2-weighted liver MRI utilizing deep learning architecture. *J Magn Reson Imaging* 2017 [Epub ahead of print].
37. Nakao T, Hanaoka S, Nomura Y, et al. Deep neural network-based computer-assisted detection of cerebral aneurysms in MR angiography. *J Magn Reson Imaging* 2017 [Epub ahead of print].
38. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Berlin: Springer International Publishing; 2016. 424–432.
39. Zhao H, Gallo O, Frosio I, Kautz J. Loss functions for image restoration with neural networks. *IEEE Trans Comput Imaging* 2017;3:47–57.