Training a neural network for Gibbs and noise removal in diffusion MRI

Matthew J. Muckley[1], **Benjamin Ades-Aron**[1,2], **Antonios Papaioannou**[1], **Gregory Lemberskiy**[1], **Eddy Solomon**[1], **Yvonne W. Lui**[1], **Daniel K. Sodickson**[1], **Els Fieremans**[1], **Dmitry S. Novikov**[1], **Florian Knoll**[1]

[1]Center for Advanced Imaging Innovation and Research (CAI2R), Department of Radiology, New York University School of Medicine, New York, NY, USA

[2]NYU Tandon School of Engineering, Brooklyn, NY, USA

Abstract

Purpose—To develop and evaluate a neural network–based method for Gibbs artifact and noise removal.

Methods—A convolutional neural network (CNN) was designed for artifact removal in diffusion-weighted imaging data. Two implementations were considered: one for magnitude images and one for complex images. Both models were based on the same encoder-decoder structure and were trained by simulating MRI acquisitions on synthetic non-MRI images.

Results—Both machine learning methods were able to mitigate artifacts in diffusion-weighted images and diffusion parameter maps. The CNN for complex images was also able to reduce artifacts in partial Fourier acquisitions.

Conclusions—The proposed CNNs extend the ability of artifact correction in diffusion MRI. The machine learning method described here can be applied on each imaging slice independently, allowing it to be used flexibly in clinical applications.

Keywords

denoising; diffusion MRI; Gibbs ringing; neural network

1 | INTRODUCTION

The Gibbs phenomenon[1,2] and noise are artifacts that affect all magnetic resonance imaging scans. The traditional solution for both low signal-to-noise ratio (SNR) and Gibbs artifacts has been to apply a smoothing filter to the image. Although smoothing can reduce the

variance of noise, it cannot directly correct for the Rician-biased signal present in low-SNR regimes, nor can it accurately reproduce images that are Gibbs-artifact-free. Methods using Gegenbauer polynomials have been shown to be able to detect edges and accurately estimate images up to the edge for the purpose of Gibbs removal.[3] Subvoxel shifting is another method that can substantially reduce Gibbs ringing in acquisitions without partial Fourier encoding.[4] Model-based reconstruction methods[5–8] can improve SNR and/or reduce Gibbs ringing. In parametric protocols with repetitive acquisitions such as diffusion MRI, correlations across the diffusion directions can be used for denoising. Recently, a random matrix theory-based marchenko-pastur principle components analysis (MP-PCA) method has been developed that takes advantage of these correlations for noise estimation[9] and removal.[10] Also for diffusion, the DESIGNER pipeline[11] based on MP-PCA and subvoxel shifting has been demonstrated and validated.

A drawback of many of these methods is that they rely on models that are not fully general. MP-PCA requires the distribution of the noise to be unaltered by processing, as well as many diffusion directions in order to take advantage of the corresponding correlations,[9,10] but many clinical protocols only use a small number of diffusion directions. Meanwhile, subvoxel shifting does not perform well on partial Fourier acquisitions.[4] Partial Fourier imaging is a complex case due to the interaction between blurring effects from the undersampling and ringing effects from the Fourier series approximation of the image. The combination of these effects is difficult to model explicitly due to the presence of ringing artifacts of varying frequency and phase.[4] As a result, some groups refrain from applying partial Fourier in diffusion imaging.

Machine learning offers a framework for building models that can be applied to simultaneously remove a variety of artifacts in an acquisition-flexible manner. One potential class of models suitable for these tasks would be convolutional neural networks (CNNs).[12,13] CNNs are models composed of simple building blocks—often convolution kernels only 3 pixels wide followed by an activation function. When many convolution kernels are used together, they can approximate highly nonlinear functions of many variables. The use of simple components makes training such models ideally suited for parallelization and efficient implementation on modern computer architectures. Several recent papers have demonstrated the applicability of CNNs for medical imaging in areas including image reconstruction,[14–18] image quality transfer,[19] super-resolution,[20] segmentation,[21] and artifact correction.[22]

Diffusion MRI is an intriguing test application for Gibbs and noise removal models. Both effects can have drastic effects on diffusion parameter maps. The calculation of higher order diffusion parameters can require the application of b-values up to $b = 2500$ s/mm^2,[23] which, coupled with strong diffusion gradients and long echo times, leads to low SNR.[24] The well-known 9% intensity variation due to Gibbs ringing[1,2] may translate into ~100% errors in the estimated signal moments or cumulants, in particular leading to the so-called "black voxels" (masked outliers) in kurtosis maps.[6] Thus, systemic defects—such as Gibbs or effects from machine learning-based processing—are substantially amplified in the corresponding parameter maps.

Here, our aim is to develop and evaluate a machine learning approach for Gibbs and noise removal. We show that the framework of CNNs allows the simultaneous removal of these artifact sources in an acquisition-flexible manner—specifically, we extend the ability for Gibbs artifact removal to the partial Fourier setting and noise removal to the setting with few diffusion directions. We accomplish this by training the model entirely on simulations of MRI acquisitions with images from the ImageNet data set.[25] Prior to the evaluation stage of our experiments, networks trained this way have never seen an MR image, mitigating overfitting risk. Our results suggest that, with CNN processing, high-quality parameter maps can be calculated from abbreviated acquisitions with a little more than half of the encoding time required for traditional full acquisitions, opening up new possibilities for diffusion MRI. The method can be applied independently to each imaging slice, allowing application to clinical protocols with few repetitive scans.

2 | METHODS

In MRI, images are reconstructed by sampling a Fourier series at a discrete set of locations where the Fourier series coefficients are the acquired k-space points. This suffers from the Fourier series approximation error that manifests as the Gibbs phenomenon.[1,2] In practice, the MRI scanner does not produce the true, noise-free Fourier series coefficients, but a noise-corrupted version of them. Here, we denote the true image at a given resolution as $\mathbf{m}_d$ and its noisy, Gibbs-corrupted version as $\mathbf{x}_d$, where d is an integer that indicates the dth image. Our goal is to learn a functional, $f_\theta(\cdot)$, that estimates $\mathbf{m}_d$ from $\mathbf{x}_d$ where θ are the parameters of the functional. We accomplish this by solving the following optimization problem:

$$\hat{\theta} = \operatorname{argmin}_\theta \sum_{d=1}^{D} \|\mathbf{m}_d - f_\theta(\mathbf{x}_d)\|_2^2, \tag{1}$$

where D is the size of the training data set.

Solving (1) requires a data set of image pairs $(\mathbf{m}_d, \mathbf{x}_d)$, $d \in [1, \ldots, D]$ or a means of generating them. We choose to simulate both $\mathbf{m}_d$ and $\mathbf{x}_d$ from a latent, high-resolution image that comes from the ImageNet data set.[25] This restricts the method to operating independently on 2D images. Figure 1 shows our simulation pipeline. We describe the details of the steps in Section 2.2.1. Section 2.1 describes the model architecture for $f_\theta(\cdot)$. In Section 2.3.2, we describe test experiments that were conducted on data sets completely separate from that used for training, including gold standard segmented EPI, prospective diffusion data on healthy and subjects with pathology, and cervical spine imaging. In these real-world diffusion validation experiments, we test the performance of two different networks and compare to state-of-the-art (SoA) methods across partial Fourier factors. We show results from these experiments in Section 3.

2.1 | Model architecture

There are many potential model architectures for accomplishing deGibbsing. Architectures such as variational[14] and cascaded[15] networks offer the most flexibility for modeling system physics. However, these models often require access to the raw data and system parameters,

for tasks such as sensitivity map estimation. Use of such models may require accessing proprietary information or signing research agreements with vendors. Furthermore, the use of models that rely on raw data can incur substantial data storage requirements. Finally, they preclude the use of the model on previously acquired image data sets where the raw data has been discarded.

Due to these considerations, we address the Gibbs phenomenon with an image-to-image architecture. Our model comes from the class of U-Net[26] architectures. Figure 2 shows a diagram of the architecture. At the midpoint of the U-Net (shown as the bottom of the U in Figure 2), convolution operations affect almost the entire image, enabling the network to learn global features such as conjugate k-space symmetry. We process complex images by inputting the real and imaginary components into the network as two separate channels.[15] After each convolution layer, we apply batch normalization.[27] After batch normalization, we apply rectified linear unit (ReLU) activation functions, as is standard practice.[28] Unlike the standard U-Net, our model includes wavelet transforms for downsampling and upsampling. The theoretical properties of these modifications were analyzed previously[16] and demonstrated to better-preserve high-resolution features. Based on the intuition that it is easier to learn residuals rather than signals,[29] we use a skipped connection over the entire architecture for the real and imaginary channels. For cases with complex inputs, we apply the magnitude operation as a final step after the skipped connection.

2.2 | Training

Code for training our models is available on Github at https://github.com/mmuckley/dldegibbs. All models were trained and tested using the PyTorch package. Links to trained models will be included in the GitHub repository. We trained the network by simulating MRI acquisitions on photographs from the ImageNet data base,[25] using resized, noise-free versions of the original photographs as target images $\mathbf{m}_d$. The data set contained 1 281 167 images in the training set and 50 000 images in the validation set. The "best" model was selected based on the minimal loss over the validation data set during the training epochs. Use of ImageNet for training requires specification of (a) an encoding simulation pipeline (Section 2.2.1) and (b) an optimization procedure (Section 2.2.2).

Although the images in ImageNet are not true MR images, we take this approach due to the fact that it is impossible to acquire large volumes of high-resolution diffusion images for the purpose of simulating Gibbs artifacts. Larger data sets are generally believed to be beneficial for learning,[30,31] and the benefits of larger data sets for denoising have been directly demonstrated.[32] An alternative would be to train from high-resolution anatomical data, but diffusion-weighted images (DWIs) can have substantially different features from their high-resolution anatomical counterparts. Thus, training a network on simulated Gibbs artifacts from high-resolution anatomical data runs the risk of having the network reconstruct signals that would not be present in the corresponding diffusion data. Furthermore, an anatomical data set would be highly homogeneous, limiting the generalizability of the resulting network. Conversely, ImageNet is a highly heterogeneous data set, enabling the network to learn more generic features and potentially enabling application of this approach in other MRI modalities that suffer from Gibbs artifacts. A key determinant of the feasibility of this

approach is the simulation of encoding operations as they occur in vivo, which we discuss below.

2.2.1 | Simulation—The simplest Gibbs simulation could be accomplished by cropping the k-space representation of a high-resolution image. However, in practice, MR acquisitions are affected by many factors other than the Gibbs phenomenon. We simulate a subset of these effects. Figure 1 shows a flowchart of the simulation pipeline. Figure S1 of the Supporting Information shows diagrams with the effects of key steps on images. The simulation begins with raw grayscale ImageNet image (converted via BT.601 color standards), resized to a standard size of 256×256 using bilinear interpolation.[33] The simulation then randomly flips and/or transposes the image, each with a 50% probability. After flipping and transposing, the simulation proceeds to the random phase step. The random phase simulation generates a random set of Gaussian radial basis functions for each simulated image—this is described in detail in Appendix A. An example of simulated phase compared to a true in vivo phase case is shown in Figure 1B,C. The random phase simulation requires a number of probabilistic parameters for specifying the distribution of the simulated phase—these are reported in the appendix.

Since most images have a no-signal background, following random phase simulation we apply random ellipsoid cropping to the images. We also include a 10% probability that no ellipsoid cropping occurs to handle cases that may not have a no-signal background. After image-domain cropping, the processing pipelines split. For the data, $\mathbf{x}_d$, we apply the FFT and crop the image from a 256×256 grid to a 100×100 grid in k-space to simulate the Fourier series truncation effect. The simulation adds noise at levels from 1 to 32 times the mean absolute k-space value, varying the noise level within this range on a uniform base-2 logarithmic scale (average noise ratio of 10).

After the inverse FFT, both processing pipelines are merged into an image normalization layer where both images are divided by the maximum absolute value in $\mathbf{x}_d$. The absolute value operation is then applied to $\mathbf{m}_d$. To simulate an artifact-free target, $\mathbf{m}_d$, we interpolate $\mathbf{m}_d$ from a 256×256 grid to a 100×100 grid using cubic splines. Finally, to provide the complex-valued CNN with more information, we concatenate a standard partial Fourier reconstruction method[34] to the complex-valued $\mathbf{x}_d$ output in the channel dimension prior to input to the neural network.

When magnitude images are used, the inputs have Rician noise rather than Gaussian. We maintain most of the above processing stream aside from three modifications. (a) The absolute value operation is applied after inverse FFT in the $\mathbf{x}_d$ stream, and (b) followed by dividing by the maximum value of the $\mathbf{x}_d$ image for both the $\mathbf{m}_d$ and $\mathbf{x}_d$ final images. This implicitly builds Rician bias correction into the network training.

2.2.2 | Optimization—Our optimization routine followed standard practice. We used the ImageNet data set[25] with standard training and validation splits. Our convolution weights were initialized with a uniform distribution.[35] ReLUs were initialized with zero parameters. We used the Adam algorithm[36] with a learning rate of 1×10^{-3} over 10 epochs to minimize the mean-squared error cost function in (1). At each epoch, we calculated the loss over the

entire validation data set and saved the model with the best validation loss for the in vivo test experiments. Each model was trained independently on each partial Fourier factor. We implemented our models in the PyTorch deep learning framework.[33] Training was performed on an IBM GPU computing cluster with Power9 8335-GTH compute nodes. Each node in the compute cluster had four Nvidia V100 GPUs with 16 GB of memory. Our specific trainings utilized one GPU with a batch size of 55. Training time took approximately 10 days.

2.3 | Testing

In the test phase, we evaluated the proposed CNN methods for their ability to process DWIs on different data sets. In all our test experiments, we define "SNR" to be the mean absolute image value divided by the complex Gaussian standard deviation. Note that this is different from the k-space ratio used during training. We conducted canonical signal processing experiments and in vivo experiments. To provide context for the performance of the CNN methods, we also processed the DWIs with a SoA method,[11] as well as a method designed for partial Fourier imaging.[34] In all cases with diffusion parameter maps, the raw DWIs were first processed by the designated method prior to parameter map estimation. To clearly show that our test data sets are distinct from our training data sets, we index test signals with t instead of d.

Raw—These parameter maps were calculated from the raw DWIs with no extra processing before the parameter estimation stage.

State-of-the-art[11]—Prior to parameter map estimation, the images were denoised using the random matrix theory-based MP-PCA method.[10] The MP-PCA method applies low-rank denoising using a cutoff based on the Marchenko-Pastur law.[37] We applied MP-PCA using a $5 \times 5 \times 5$ voxel kernel extent with Rician bias correction,[10] followed by Gibbs removal via subvoxel shifting[4] where each voxel was discretized into 20 subvoxel elements. The subvoxel shifting method searches in a subvoxel neighborhood for the minimum Gibbs distortion given by the sinc function and interpolates to that subvoxel location. Unlike the CNN methods, the SoA method is able to use information from low-b DWIs to denoise images from high-b diffusion gradients.

Magnitude-input CNN (MCNN)—The raw diffusion images after the magnitude operation, $\mathbf{x}_t$, are processed by a neural network trained with the simulation pipeline in Figure 1 with the target image $\mathbf{m}_t$ being a magnitude, deGibbsed image without PF. This particular network attempts to restore the missing partial Fourier information. This network has 31 646 338 parameters.

Standard Partial Fourier (Standard PF)—The raw diffusion images are processed by using a standard partial Fourier reconstruction method that utilizes the phase maps from the symmetric region of k-space[34] prior to parameter estimation.

Complex-input CNN (CCNN)—The raw complex-valued diffusion images, $\mathbf{x}_t$, are processed by a neural network trained with the simulation pipeline in Figure 1 with the

target image $\mathbf{m}_t$ being a magnitude, deGibbsed image without PF. This particular network attempts to restore the missing partial Fourier information. This network has 31 646 338 parameters.

2.3.1 | Canonical experiments—To characterize the deGibbsing and denoising performance of the network, we performed a set of canonical signal processing experiments with two data sets. The first data set consisted of a simple edge phantom, while the second data set consisted of the test split of the 2013 ImageNet challenge.[38] *Edge-spread function:* Edge-spread functions are routinely used to study the resolution properties of image processing methods. Toward the goal of identifying when the network blur small lesions, we conducted tests with a digital edge phantom across a variety of contrast-to-noise ratios. The edge phantom consisted of a 2D digital image with half of the image filled with 1s and the other half filled with 0s. This image was generated at a 1024×1024 matrix size (cropped to a 100×100 matrix size during simulation). The location of the edge was varied by rotating it. Figure 3 includes a composite image of the phantom generated from $f_{\hat{\theta}}(\mathbf{x}_t)$ at many different SNR levels. We performed two sets of experiments with the edge phantom: contrast-to-noise (CNR) and deGibbsing experiments. In a previous study,[4] it was noted that deGibbsing performance could vary depending on the angle of the edge. To assess the performance of the method on removing Gibbs artifacts at different angles, we generated the phantom at a set of angles beetween 0 and 45 degrees at a 1024 SNR level and computed the CCNN deGibbsed image.

We also performed CNR experiments. The goal of the CNR experiments was to quantify the full-width-half-maximum (FWHM) of the network's line-spread function across a variety of noise levels (analogous to a point-spread function in 1D). We accomplished this by applying the simulation pipeline in Figure 1 without random flipping and transposing at a variety of CNRs between 0 and 10, resulting in images at 100×100 matrix sizes. For each CNR, we applied both the CCNN model and the MCNN model (generating $f_{\hat{\theta}}(\mathbf{x}_t)$ realizations for each one). This resulted in a series of edge-spread function images. To compensate for noise in calculating the FWHM values, for each row of the resulting image, we fit a summation of logistic functions[39]:

$$l(s) = c + \sum_{i=1}^{3} \frac{\alpha_i}{1 + \exp\left(\frac{s - \beta_i}{\gamma_i}\right)}, \tag{2}$$

where c, α_i, β_i, and γ_i, $i \in [1, 2, 3]$ are fitting parameters. We computed the derivative of this function and used it to calculate a line-spread function for each row of the image. From the line-spread functions, we calculated the FWHMs and averaged across the rows. For each CNR level, we repeated and averaged this experiment 50 times.

ImageNet test data: The goal of the tests on the ImageNet data was to assess the network's performance across a broad variety of signals. For this, we used the test split from the 2013 ImageNet challenge.[38] This data was not used at all in training and is distinct from the validation data set. The test set included 40 152 images. We applied the simulation pipeline in Figure 1 and computed the power spectral ratio on the resulting images.[6] Let $S(\cdot)$ be the

operator that computes the power spectral density of the image. For each image-target pair in the validation data set, $(\mathbf{x}_t, \mathbf{m}_t)$, we computed

$$h_t(\boldsymbol{f}) = \sqrt{\frac{\mathcal{S}(f_{\hat{\theta}}(\mathbf{x}_t))}{\mathcal{S}(\mathbf{m}_t)}} \, . \qquad (3)$$

Then, we averaged over all $t \in [1, \ldots, n_{\text{test}}]$ to compute a mean $\bar{h}(f)$ for each partial Fourier factor and SNR level and computed profiles across f in the partial Fourier direction to assess resolution performance. As a last test, we calculated image metrics of MSE, SSIM, and PSNR over this simulated test set. The results are described in the Supporting Information and shown in Table S1.

2.3.2 | In vivo experiments—For our in vivo tests, we conducted four experiments:

1. A "gold standard" segmented EPI experiment. This was used to achieve a higher matrix size and produce reference diffusion images with almost no Gibbs artifact.

2. A prospective diffusion experiment. This tested the method vs the current SoA in a standard acquisition.

3. A pathology experiment. We acquired diffusion pathology data from a previous study and assessed stability of the MCNN model in the presence of MS lesions.

4. A cervical spine experiment. In addition to diffusion, Gibbs artifacts can confound clinical evaluations.[40,41] In this experiment, we evaluated the method in an alternative anatomy.

"Gold standard" segmented EPI—We acquired segmented EPI data for the purpose of simulating realistic diffusion data that is nearly Gibbs-free. In many cases, segmented EPI scans are not practical due to time limitations and difficulties with reconstruction, but here we use this sequence for the purpose of creating a diffusion reference at a 100×100 matrix size using a segmented EPI acquisition at a 128×128 matrix size. A volunteer was scanned with a diffusion-weighted imaging protocol under IRB approval. The scan was done after all model and training parameters were complete, so that these data qualify as a true prospective test set. The sequence was applied on a 3T Prisma scanner with a 16-channel head coil (Siemens Healthineers, Erlangen, Germany). Using a segmented EPI readout, we were able to acquire images at a 128×128 matrix size with a 2 mm isotropic resolution. The images were acquired at b-values of 500, 1000, and 2000 s/mm^2 in 39 isotropic directions. The acquisition had TE of 64 ms and a TR of 3 seconds. Reconstruction was accomplished with Siemens software via adaptive combine,[42] saving the resulting magnitude and phase images.

For the "gold standard" processing, we applied MP-PCA, apodized the 128×128 k-space data, applied the inverse FFT, and interpolated to a 100×100 image grid. This produced DWIs that were essentially noise and Gibbs-free. For the experimental data sets, we cropped the 128×128 k-space to 100×100 and added noise to simulate an SNR of 10 at $b = 0$ s/mm^2. Then we applied partial Fourier masks at PF factors of 5/8ths, 6/8ths, 7/8ths, and fully sampled, prior to inverse FFT.

To process the experimental data, we applied the CCNN, MCNN, and SoA processing methods to the images with simulated Gibbs. Parameter map estimation for all methods was performed on the processed image using a constrained weighted linear least squares approach.[43] Since the complex CNN can take into account smooth variations in the phase of the image to recover high-resolution information, we also included as a comparison a reconstruction using a standard partial Fourier method that estimates smooth phase from the symmetric region of k-space[34] applied with our own custom software to the raw data.

Prospective in vivo diffusion—The prospective in vivo diffusion experiments mirrors the in vivo segmented EPI experiments with the absence of the "gold standard" 100×100 parameter maps. DWIs were acquired with a spin-echo EPI sequence at a 100×100 matrix size with a 2-mm in-plane isotropic resolution over 38 slices and along 66 isotropically distributed directions at b-values of 0, 1000, and 1500 s/mm^2. We applied the diffusion-weighted acquisition four times: once for each PF factor of 5/8ths, 6/8ths, 7/8ths, and fully sampled, with the intention that the fully sampled DWIs with current SoA processing serve as a gold standard. The acquisition parameters were set based on the shortest values we could achieve for the fully sampled acquisition, leading to a TR of 9 seconds and a TE of 105 ms. The partial Fourier acquisitions used a standard protocol. For this data set, we did not apply any other corrections such as eddy current correction or motion correction, only using the adaptive combine method[42] outside of the vendor reconstruction software. As with the gold standard experiments, parameter map estimation was done with a weighted linear least squares approach.[43]

Diffusion pathology—To assess the method in the presence of pathology, we obtained data from a previous study of a patient with multiple sclerosis. For this, we had access to vendor-reconstructed DICOMs, so we were only able to process non-PF images with the MCNN and SoA models. The images were acquired using parallel imaging with an in-plane acceleration factor of 2 and a multiband factor of 2 with b-values of 0, 1000, and 2000 s/mm^2 over 89 isotropic directions with a TE of 100 ms and a TR of 3.5 seconds. The images were reconstructed at a 130×130 matrix size with a 1.69 mm isotropic resolution. This required retraining the MCNN model for the new target matrix size. After processing with the MCNN and SoA methods, we estimated parameter maps with a weighted linear least squares approach[43] after processing with the MCNN and SoA models.

Cervical spine experiments—To assess the performance of the method on other anatomies, we also acquired T2-weighted cervical spine data. The cervical spine data set was acquired with a standard clinical protocol that involved a 2D tubo spin-echo acquisition with 75% phase undersampling 240×320 acquisition matrix size with a TE of 103 ms and a TR of 3.8 seconds. The reconstructed images had a 0.625 isotropic resolution in the sagittal plane. The frequency-encoding direction was in the A/P direction. For performing deGibbsing at this resolution, we retrained the network, adjusting the MCNN model to produce a target image that included the 75% phase undersampling, but with Gibbs artifacts removed. This was accomplished by applying a phase undersampling mask in the simulation and altering the target resolution to 320×320. This adjustment discourages the network

from trying to fill in the missing high-resolution information. We applied mild dithering to the resulting image—a practice applied in previous works with CNN-based models.[44]

3 | RESULTS

3.1 | Canonical experimental results

Figure 3A,B shows the performance of the network in a canonical deGibbsing experiment with a hard edge with the edge oriented at 0 and 45 degrees, respectively. Both the CCNN and the MCNN methods performed similarly at both angles, with the main difference being the sharpness of the edges with the MCNN method both at the simulated edge and the image boundaries. These image boundary residual artifacts do not appear in the 45° plot since this profile does not reach the image boundaries.

Figure 3C,D shows the response of the CCNN in the edge experiments across a range of contrast-to-noise ratios. Above a CNR of 1, the CNN FWHM rapidly converges toward 1. Below a CNR ratio of 1, the network's FWHM increases as the CNR approaches 0. Figure 3 also shows a composite image of CNN image estimates, $f_{\hat{\theta}}(\mathbf{x}_t)$, across CNR factors. Figure 3 shows visually how as the CNR decreases, the edge begins to blur.

Figure 4 shows profiles of the mean spectral response, $\bar{h}(f)$, in the partial Fourier direction of the CCNN model across different partial Fourier factors and SNR levels over the entire test set. The spectral responses were calculated from the 2013 test split of ImageNet as described in Section 2.3.1. As the SNR decreases, the network applies more smoothing, resulting in a decrease of the spectral magnitude at high frequencies of about 30% at an SNR value of 1. The MCNN method exhibited similar profiles, but with more smoothing at low SNR—the MCNN smoothing without partial Fourier at an SNR of 1 led to about 50% attenuation of the high frequencies. In all cases, we observed features in the spectral responses around spectral cutoff points—this included the ripples at the edge of the sampled k-space in the No PF setting, with further cusps developing near the PF cutoff points in the PF settings.

Quantitative results on the ImageNet test set are shown in Table S1 in the Supporting Information. For MSE, the CCNN model performed better than the MCNN model at equivalent partial Fourier factors, and performance generally decreased with increased acceleration. This trend was repeated for PSNR, but MCNN models displayed higher SSIM values.

3.2 | In vivo results

"Gold standard" segmented EPI—Figure 5 shows $b = 0$ s/mm^2 images, average $b = 1000$ s/mm^2 images, and mean diffusivity (MD) images without partial Fourier (No PF) and with 5/8ths partial Fourier (5/8 PF) calculated from the segmented EPI experiments. The far right column shows the "gold standard" segmented EPI diffusion processing results. This data has minimal Gibbs artifacts. The other columns consist of experimental data processing methods that should approach the "gold standard" column.

Mean diffusivities calculated from the raw DWI data (Raw) show notable noise and Gibbs ringing artifacts, while this is removed with the SoA method. However, the SoA method begins to lose its ability to compensate for the Gibbs ringing and resolution loss when partial Fourier is utilized in the acquisition. Partial Fourier artifacts are reduced in the CNN methods. Other slices from this data are shown in the Supporting Information in Figures S7–S11.

Prospective in vivo diffusion—Figure 6 shows examples of images from the prospective data before and after processing at low ($b = 0$ s/mm^2) and high ($b = 1\,500$ s/mm^2) diffusion gradient strengths. The raw images (Raw) served as the input for the magnitude deep learning (MCNN) and complex deep learning (CCNN) methods. Figures 7 and 8 show parameter maps for prospective diffusion data. Figure 7 shows DWIs and mean diffusivity, whereas Figure 8 shows fractional anisotropy and mean kurtosis. Aside from the MCNN method, results are similar to those shown in Figure 6, with the SoA, MCNN, and CCNN methods all removing artifacts without partial Fourier, whereas the CCNN method is the only method able to remove artifacts at 5/8 partial Fourier. This is most obvious in the MD parameter maps. Similar trends are observed in the other diffusion parameter maps. The effects of denoising are primarily visible in the mean kurtosis images, where SoA, MCNN, and CCNN achieve substantial noise reduction.

Figure 9 compares MD maps across various partial Fourier factors between the MCNN, CCNN and SoA methods. The methods perform similarly without partial Fourier acceleration, but as partial Fourier acceleration increases, the image is continually degraded in the SoA method, with "black voxels" appearing around the lateral ventricles. The CCNN method mitigates the appearance of these artifacts in the parameter maps. Other slices from this data set are shown in the Supporting Information in Figures S12–S16.

Diffusion pathology—Figure 10 shows results from in vivo diffusion pathology. In this case, we only had access to magnitude DICOMs with iPaT acceleration, so only the MCNN method was used. We observed similar artifact characteristics as in the previous Figures 5, 7, and 8. In this case, all three methods provide similar diffusion parameter maps for MD, FA, and MK. Other slices from this data set, including zoomed images of pathology, are shown in the Supporting Information in Figures S17–S22.

Cervical spine experiments—We show cervical spine results in the Supporting Information in Figure S24.

4 | DISCUSSION

Our results indicate that with carefully crafted simulations, it is possible to train CNN models based solely on MR encoding operators that can be applied to in vivo MR imaging data. We characterized the performance of this method with canonical experiments, showing Gibbs removal capability with hard edges, as well as providing evidence that the network's denoising capability depends on signal-adaptive smoothing, and that this smoothing largely takes place at contrast-to-noise ratios below 1. In principle, there are no restrictions on applying the method in new applications other than retraining based on the application

image matrix size. When reconstruction and preprocessing steps failed to remove artifacts, the network also generally preserved them.

4.1 | Effects of in vivo reconstruction method

The results diverged depending on the reconstruction method between the segmented EPI and prospective diffusion experiments. This could be due to the ordering of partial Fourier versus adaptive coil combination. In the segmented EPI experiment, the partial Fourier masking was done after coil combination (as coil combination was done by the vendor reconstruction software). This exactly matched the partial Fourier masking of the simulation. In the prospective diffusion experiment, coil combination was done after partial Fourier masking, which did not match the simulation. This may be the reason for the failure of MCNN for prospective diffusion at extreme partial Fourier factors. The effect seems to have been benign for MCNN method at lower partial Fourier accelerations. The effect was also benign for the CCNN method, as it was trained to use smooth phase. As the prospective diffusion experiment is more representative, we would only recommend the CCNN method out of the methods we tested when using extreme partial Fourier acceleration. For other balanced accelerations like parallel imaging (eg, the pathology data), we did not observe effects on either model. The in vivo results here should be considered an initial demonstration of the potential of the method. Validation studies, particularly those studying the effects of reconstruction method, are necessary before adoption in target applications.

4.2 | Data set considerations

Practically, the success of the approach can be attributed to (a) having a sufficiently deep neural network able to adequately capture the difference between Gibbs-corrupted and ground-truth images and (b) to the heterogeneity of the synthetic training set and the richness of the encoding simulations providing ample training data. During training, the models saw over 10 million candidate phase maps simulated on over 1 million baseline images. Data sets of this size are difficult to acquire in medical imaging, and impossible to acquire for the purpose of simulating Gibbs artifacts in diffusion-weighted imaging. We ran one experiment on a smaller data set size, using 1% of ImageNet (about 12 000 images). Although the resulting models performed reasonably on ImageNet validation data, they failed on in vivo data, producing diffusion parameter maps outside of physiological ranges (shown in Figure S23 of the Supporting Information). This suggests that the proposed approach overfits on 1% of ImageNet, but this overfitting is mitigated when using the full ImageNet data set. Training such a model requires careful consideration of the simulation process and its relationship with the test data set. As we see in Figures 6 and 7, subtle or benign artifacts in the DWIs are substantially magnified in the corresponding parameter maps. To consider the effects of overfitting, we designed our experiments to have not only separate training, validation, and test data sets as different groups of images, but to have the test phase of the experiments consider images of a different class. Prior to our diffusion imaging experiments, we did not use MR images. In principle, this would allow detection of overfitting errors when testing on healthy subjects. As shown in Figure 10, we did not detect issues using the model on patients with pathology.

4.3 | Future directions

Our results were achieved by processing each DWI separately, which could be advantageous in that the processing of each image is independent of artifacts that may be present in other images. However, a disadvantage is that the proposed methods do not fully leverage the repetitive correlations present in applications such as diffusion MRI. It is difficult to denoise images at high *b*-values without access to information from low *b*-value images. Such relations are considered in the random matrix denoising method.[10] This suggests that one avenue to explore in the future for diffusion would be to combine the random matrix theory-based denoising[10] as a preprocessing step, able to incorporate the correlations between different DWIs, with the subsequently improved performance of CCNN-based removal of Gibbs and other artifacts.

Lastly, we note that an independent work recently appeared on the topic of Gibbs removal in MRI.[45] Although this and the present paper share target applications, they differ fundamentally in approach, with ours focusing on simulations and the effects of Gibbs on higher-order diffusion parameters such as mean kurtosis, while the cited work uses a k-space masking framework on in vivo MR data. Both these methods represent interesting approaches, with an item of future work being to explore relative benefits and tradeoffs in target applications.

5 | CONCLUSION

We have shown that Gibbs artifacts and noise can be substantially reduced by training a convolutional neural network with simulations of the image acquisitions. We demonstrated our method on canonical experiments and in vivo data that includes healthy subjects and those with pathology, and showed its potential for artifact removal. The method can be applied independently on each slice of the imaging data set, enabling its use in many clinical settings such as clinical diffusion MRI where few diffusion directions may be available. In the future, we will examine possible extensions, including consideration of 3D anatomical structures, correlations across repetitive acquisitions, explicit incorporation of physics into the model, and validation on other clinical applications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

APPENDIX A

PHASE SIMULATION

Inspired by results from functional approximation theory[46,47] and machine learning,[48] we approximate the phase as being a summation of Gaussian radial basis functions:

$$p_{\text{sim}}(\mathbf{r}) = \exp\left(i \sum_{s=1}^{S} \sum_{b=1}^{B_s} a_{b,s} e^{-\frac{1}{z_s}\left\|\mathbf{r} - \mathbf{r}_{0,b,s}\right\|_2^2}\right), \tag{A1}$$

where $p_{\text{sim}}(r)$ is the phase map, r is a 2-length vector for spatial position, $a_{b,s}$ is the amplitude of the (b,s)th basis function, $r_{0,b,s}$ is the center of the (b, s)th basis function, and z_s governs the width of the basis functions in the sth subset. In functional approximation theory, formulations like (A1) are used to build a candidate for $p_{\text{sim}}(r)$ given a set of points, $(a_{b,s}, r_{0,b,s})$, $b \in [1, \ldots, B_s]$ (z_s is usually chosen to be fixed). However, rather than estimate $p_{\text{sim}}(r)$ from a measured cloud of points, our goal is to randomly generate an example from a synthetic cloud of points. For each training example, we randomly generate $p_{\text{sim}}(r)$ according to the following model:

$$
\begin{aligned}
S &\sim \text{Poisson}(\lambda_S) \\
B_s &\sim \text{Poisson}(\lambda_B) \\
z_s &\sim |\mathcal{N}|(\mu_z, \sigma_z) \\
a_{b,s} &\sim \mathcal{N}(\mu_a, \sigma_a) \\
r_{0,b,s,j} &\sim \mathcal{V}(r_{\min,j}, r_{\max,j}),
\end{aligned}
$$

where $\mathcal{N}$ indicates a Gaussian distribution, $|\mathcal{N}|$ is a folded Gaussian distribution, and $\mathcal{V}$ indicates a uniform distribution. The parameters λ_S, λ_B, μ_z, σ_z, μ_a, and σ_a are set for training, and $r_{\min,j}$ and $r_{\max,j}$ are the minima and maxima for r in the jth imaging dimension. Contrary to practice in the approximation literature, we allow z_s to vary to avoid the degenerate training case where networks create filters to detect Gaussian functions of a fixed width.

When B_s and S are not too large, $p_{\text{sim}}(r)$ will typically be smooth as in real application settings; however, sharp changes can also occur during training when basis function centers are near each other. As a further form of data augmentation, we include a 1% probability that no phase at all is simulated. Figure 1B,C shows an example of simulated phase compared to real phase from in vivo data. For our experiments, we used $\lambda_S = 12$, $\lambda_B = 15$, $\mu_z = 64$ pixels^2 $\sigma_z = 100$ pixels^2, $\mu_a = 0°$, and $\sigma_a = 5°$. We set $r_{\min,j}$ and $r_{\max,j}$ to restrict the center of the Gaussian basis function to the image support in the jth dimension.

REFERENCES

1. Wilbraham H On a certain periodic function. Cambridge Dublin Math J. 1848;3:198–201.

2. Gibbs JW. Fourier's series. Nature. 1898;59:200.

3. Archibald R, Gelb A. A method to reduce the Gibbs ringing artifact in MRI scans while keeping tissue boundary integrity. IEEE Trans Med Imaging. 2002;21:305–319. [PubMed: 12022619]

4. Kellner E, Dhital B, Kiselev VG, Reisert M. Gibbs-ringing artifact removal based on local subvoxel-shifts. Magn Reson Med. 2016;76:1574–1581. [PubMed: 26745823]

5. Block KT, Uecker M, Frahm J. Suppression of MRI truncation artifacts using total variation constrained data extrapolation. Int J Biomed Imaging. 2008;2008.184123 [PubMed: 18784847]

6. Veraart J, Fieremans E, Jelescu IO, Knoll F, Novikov DS. Gibbs ringing in diffusion MRI. Magn Res Med. 2016;76:301–314.

7. Basu S, Fletcher T, Whitaker R. Rician noise removal in diffusion tensor MRI. In: International Conference on Medical Image Computing and Computer-Assisted Intervention Copenhagen, Denmark: Springer; 2006:117–125.

8. Varadarajan D, Haldar JP. A majorize-minimize framework for Rician and non-central chi MR images. IEEE Trans Med Imaging. 2015;34:2191–2202. [PubMed: 25935028]

9. Veraart J, Fieremans E, Novikov DS. Diffusion MRI noise mapping using random matrix theory. Magn Reson Med. 2015;76:1582–1593. 10.1002/mrm.26059. [PubMed: 26599599]

10. Veraart J, Novikov DS, Christiaens D, Ades-Aron B, Sijbers J, Fieremans E. Denoising of diffusion MRI using random matrix theory. NeuroImage. 2016;142:394–406. [PubMed: 27523449]

11. Ades-Aron B, Veraart J, Kochunov P, et al. Evaluation of the accuracy and precision of the diffusion parameter EStImation with Gibbs and NoisE removal pipeline. NeuroImage. 2018;183:532–543. [PubMed: 30077743]

12. LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. Proc IEEE. 1998;86:2278–2324.

13. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks In: Advances in Neural Information Processing Systems. Lake Tahoe, California, U.S.: Curran Associates Inc; 2012:1097–1105.

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. [PubMed: 29115689]

15. Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. IEEE Trans Med Imaging. 2018;37:491–503. [PubMed: 29035212]

16. Ye JC, Han Y, Cha E. Deep convolutional framelets: a general deep learning framework for inverse problems. SIAM J Imaging Sci. 2018;11:991–1048.

17. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. Nature. 2018;555:487. [PubMed: 29565357]

18. Knoll F, Hammernik K, Kobler E, Pock T, Recht MP, Sodickson DK. Assessment of the generalization of learned image reconstruction and the potential for transfer learning. Magn Reson Med. 2019;81:116–128. [PubMed: 29774597]

19. Tanno R, Worrall DE, Ghosh A, et al. Bayesian image quality transfer with CNNs: exploring uncertainty in dMRI super-resolution. In: International Conference on Medical Image Computing and Computer-Assisted Intervention Quebec City, Canada: Springer; 2017:611–619.

20. Chaudhari AS, Fang Z, Kogan F, et al. Super-resolution musculoskeletal MRI using deep learning. Magn Reson Med. 2018;80:2139–2154. [PubMed: 29582464]

21. Deniz CM, Xiang S, Hallyburton RS, et al. Segmentation of the proximal femur from MR images using deep convolutional neural networks. Sci Rep. 2018;8:16485. [PubMed: 30405145]

22. Schnurr AK, Chung K, Russ T, Schad LR, Zöllner FG. Simulation-based deep artifact correction with convolutional neural networks for limited angle artifacts. Z Med Phys. 2019;29:150–161. [PubMed: 30772110]

23. Fieremans E, Jensen JH, Helpern JA. White matter characterization with diffusional kurtosis imaging. NeuroImage. 2011;58:177–188. [PubMed: 21699989]

24. Basser PJ, Pajevic S. Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. Magn Reson Med. 2000;44:41–50. [PubMed: 10893520]

25. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: a large-scale hierarchical image database. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Miami, FL: IEEE; 2009:248–255.

26. 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 Boston, MA: Springer; 2015:234–241.

27. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:150203167 2015.

28. Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. In: Proceedings of International Conference on Artificial Intelligence and Statistics, Ft Lauderdale, FL; 2011:315–323.

29. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition; Las Vegas, NV, 2016:770–778.

30. Sun C, Shrivastava A, Singh S, Gupta A. Revisiting unreasonable effectiveness of data in deep learning era. In: Proceedings of the IEEE International Conference on Computer Vision; Honolulu, HI, 2017:843–852.

31. Zhang K, Zuo W, Gu S, Zhang L. Learning deep CNN denoiser prior for image restoration. In: Proceedings of the IEEE conference on computer vision and pattern recognition; Honolulu, HI, 2017:3929–3938.

32. Burger HC, Schuler CJ, Harmeling S. Image denoising: can plain neural networks compete with BM3D? In: 2012 IEEE conference on computer vision and pattern recognition, Providence, RI: IEEE; 2012:2392–2399.

33. Paszke A, Gross S, Chintala S, et al. Automatic differentiation in PyTorch; NIPS 2017 Autodiff Workshop:The future of gradient-based machine learning software and techniques, 2017.

34. Margosian P, Schmitt F, Purdy D. Faster MR imaging: imaging with half the data. Health Care Instrum. 1986;1:195.

35. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition; Boston, MA, 2015:1026–1034.

36. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv preprint arXiv:14126980 2014.

37. Marchenko VA, Pastur LA. Distribution of eigenvalues for some sets of random matrices. Matematicheskii Sbornik. 1967;114:507–536.

38. Russakovsky O, Deng J, Su H, et al. Imagenet large scale visual recognition challenge. Int J Comput Vis. 2015;115:211–252.

39. Li T, Feng H. Comparison of different analytical edge spread function models for MTF calculation using curve-fitting In: MIPPR 2009: Remote Sensing and GIS Data Processing and Other Applications International Society for Optics and Photonics. Bellingham, WA: SPIE; 2009;7498:74981H.

40. Vargas M, Delavelle J, Kohler R, Becker C, Lovblad K. Brain and spine MRI artifacts at 3 Tesla. J Neuroradiol. 2009;36:74–81. [PubMed: 18835643]

41. Morelli JN, Runge VM, Ai F, et al. An image-based approach to understanding the physics of MR artifacts. Radiographics. 2011;31:849–866. [PubMed: 21571661]

42. Walsh DO, Gmitro AF, Marcellin MW. Adaptive reconstruction of phased array MR imagery. Magn Reson Med. 2000;43:682–690. [PubMed: 10800033]

43. Veraart J, Sijbers J, Sunaert S, Leemans A, Jeurissen B. Weighted linear least squares estimation of diffusion MRI parameters: strengths, limitations, and pitfalls. NeuroImage. 2013;81:335–346. [PubMed: 23684865]

44. Sriram A, Zbontar J, Murrell T, Zitnick CL, Defazio A, Sodickson DK. GrappaNet: combining parallel imaging with deep learning for multi-coil MRI reconstruction. arXiv preprint arXiv:191012325 2019.

45. Zhang Q, Ruan G, Yang W, et al. MRI Gibbs-ringing artifact reduction by means of machine learning using convolutional neural networks. Magn Reson Med. 2019;82:2133–2145. [PubMed: 31373061]

46. Powell MJD. Radial basis functions for multivariable interpolation: a review Algorithms for Approximation. Oxford, UK: Clarendon Press; 1987:143–167

47. Carr JC, Beatson RK, Cherrie JB, et al. Reconstruction and representation of 3D objects with radial basis functions. Comput Graph ACM. 2001;67–76.
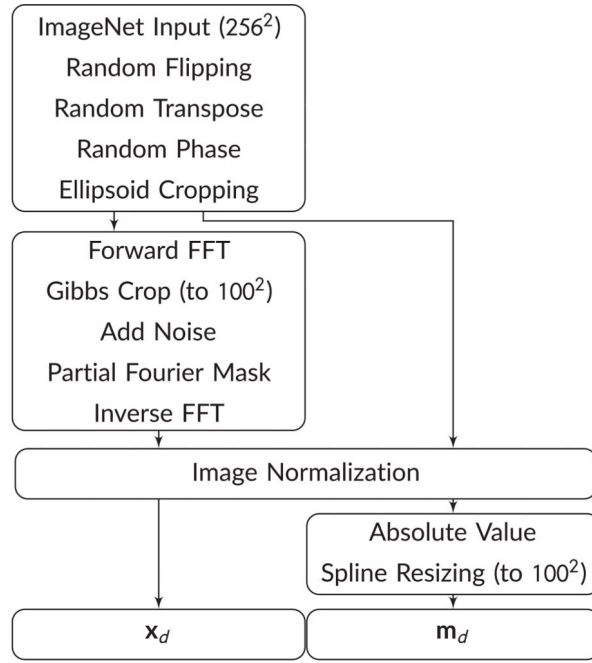
48. Broomhead DS, Lowe D. Radial basis functions, multi-variable functional interpolation and adaptive networks. Royal Signals Radar Establishment Malvern (United Kingdom). 1988.

(A) Simulation Pipeline

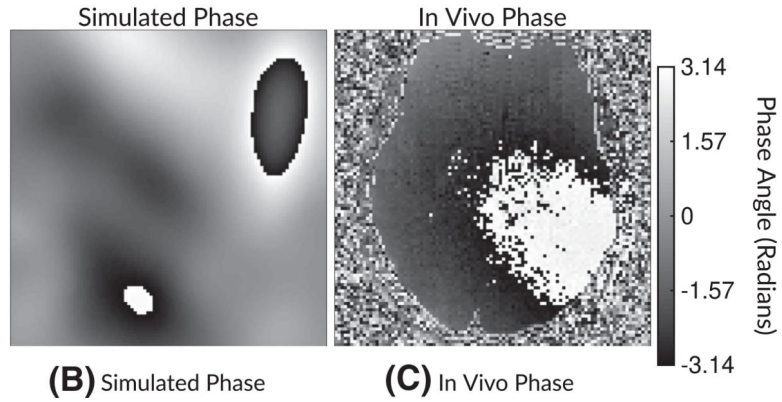**(B)** Simulated Phase   **(C)** In Vivo Phase

FIGURE 1.

A, Simulation pipeline and (B,C) example phase maps. The simulation pipeline begins with a 256×256 ImageNet image. Then, after random flipping and transposing, random phase and random ellipsoid cropping are applied. Then, the network input $\mathbf{x}_d$ has Gibbs ringing simulated along with noise and partial Fourier masking, followed by normalization. The target $\mathbf{m}_d$ images have the absolute value operation applied followed by spline resizing to the target image matrix size. An example of a phase map from this simulation is shown in (B), while a comparison in vivo phase map is shown in (C). Phase maps can have substantial variations that must be considered when applying partial Fourier imaging
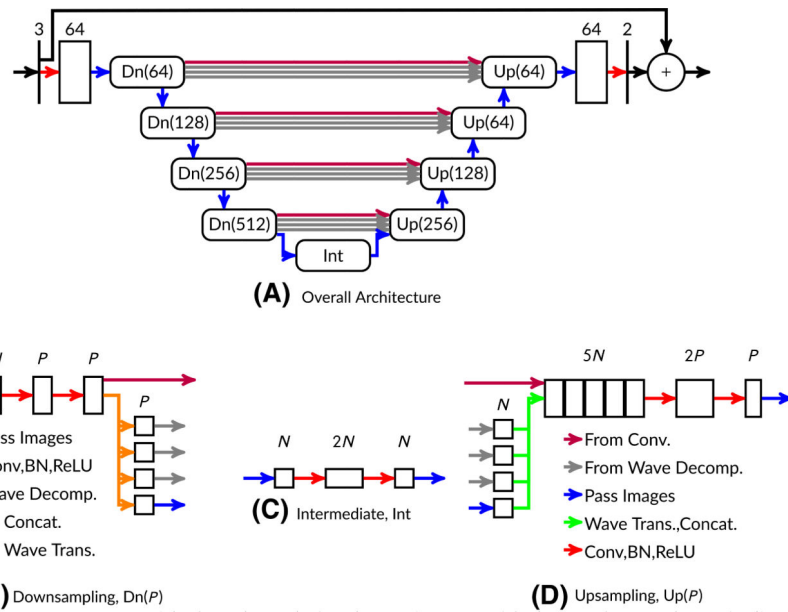
FIGURE 2.
Model architecture (A). Boxes with sharp boundaries denote images with the number of channels displayed above the box. Boxes with rounded boundaries are submodules. Also shown are example downsampling (B), upsampling (D), and intermediate (C) submodules at a constant network depth. The model as shown here has 31,646,338 parameters
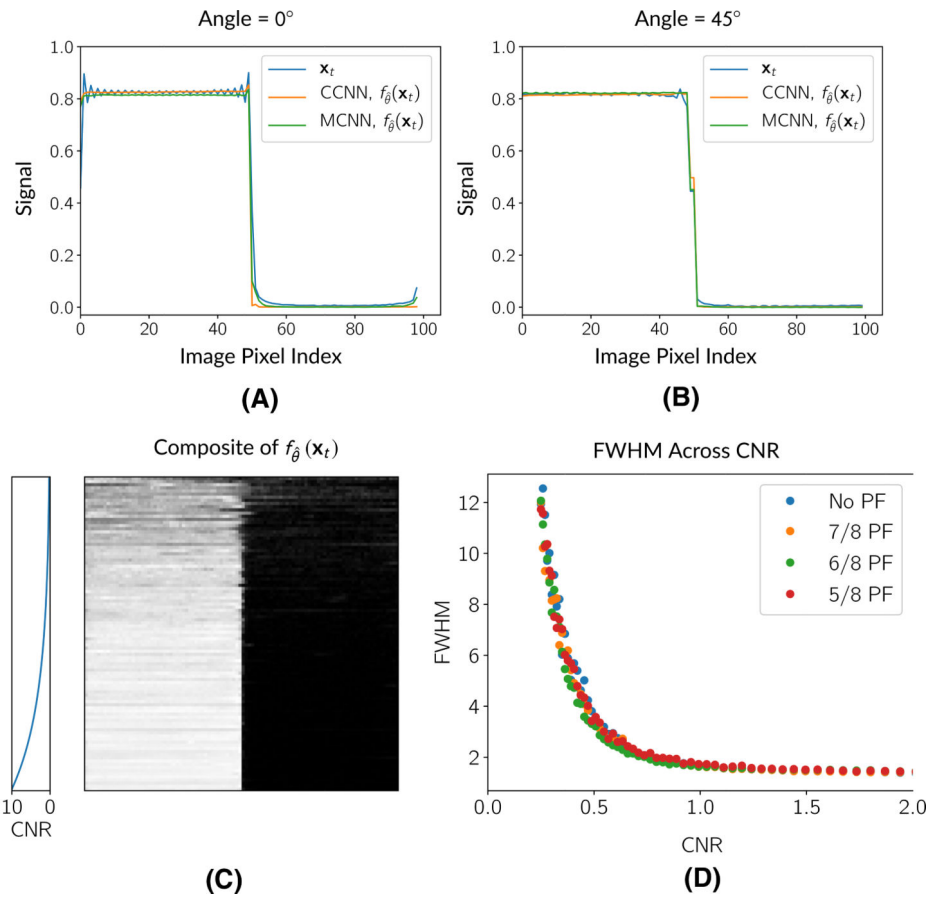
FIGURE 3.

Responses of CNN models to the edge phantom at different angles and across CNR levels. (A) and (B) show profiles across the edge phantom for the raw, Gibbs-corrupted data, $\mathbf{x}_t$, and the CCNN/MCNN models, $f_{\hat{\theta}}(\mathbf{x}_t)$, for edges at different angles. Both the CCNN and the MCNN method reduce the Gibbs artifacts. (C) shows a composite image of CCNN network estimates across a variety of CNRs from 0.25 to 10, while (D) shows the change in estimated FWHM as a function of CNR for the CCNN model. Above a CNR of 1, CNN performance is fairly stable. Below a CNR of 1, the FWHM increases as the CNR approaches 0
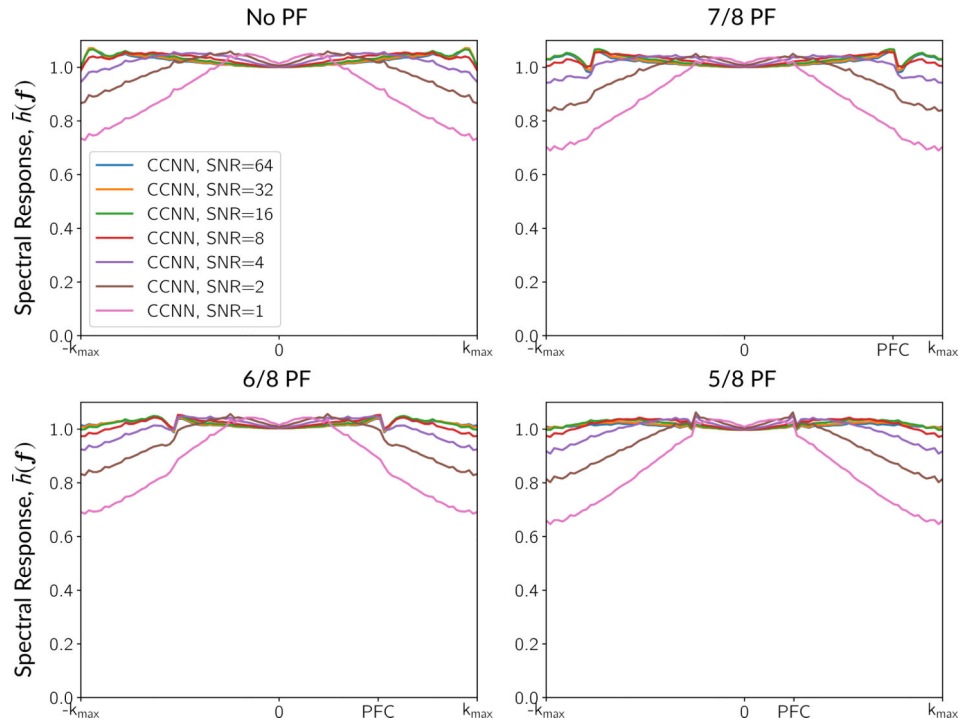
FIGURE 4.

Spectral responses in the partial Fourier direction across different partial Fourier factors and SNR levels for the CCNN model. "PFC" indicates the partial Fourier cutoff. Ripples can be seen that are related to the Gibbs effect and the PF cutoffs. In the partial Fourier cases, the CCNN response has a cusp in the frequency response at the level of the partial Fourier factor. The response beyond the PFC characterizes the ability to restore the missing Fourier harmonics. At lower SNRs, the PFC singularity gradually disappears, and the CCNN model becomes a low-pass filter, applying more smoothing to compensate for the increased noise. Profiles for the MCNN method were similar, but resulted in smoothing at the lower SNR levels
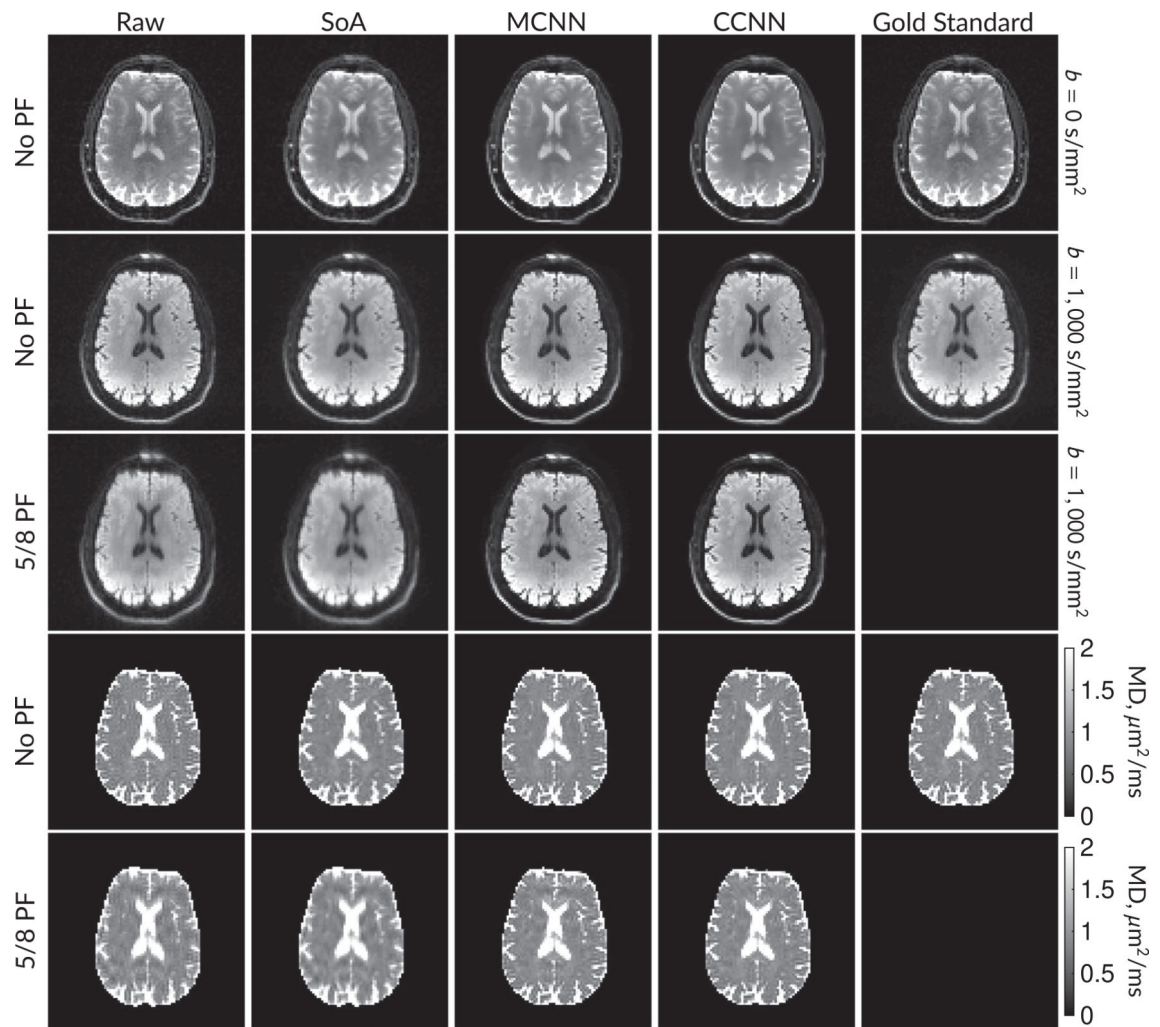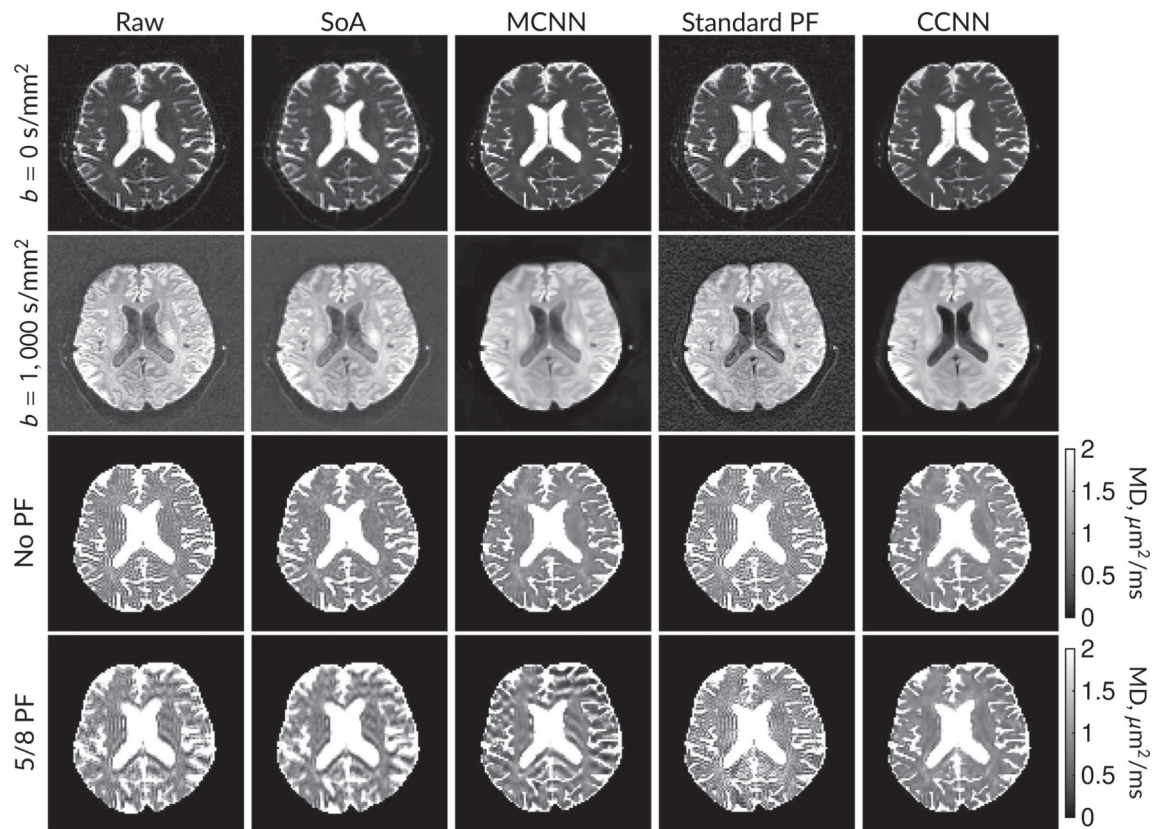
FIGURE 5.
DW and MD images from various methods performed on segmented EPI data. At the top are shown the $b = 0$ s/mm^2 and averaged $b = 1000$ s/mm^2 images from Raw, SoA, MCNN, and CCNN methods. The final column shows the "gold standard," results from apodizing the full segmented EPI data set and cropping. Comparisons with and without partial Fourier are shown for $b = 1000$ s/mm^2 images and mean diffusivity images. Partial Fourier artifacts are mitigated in the CNN methods
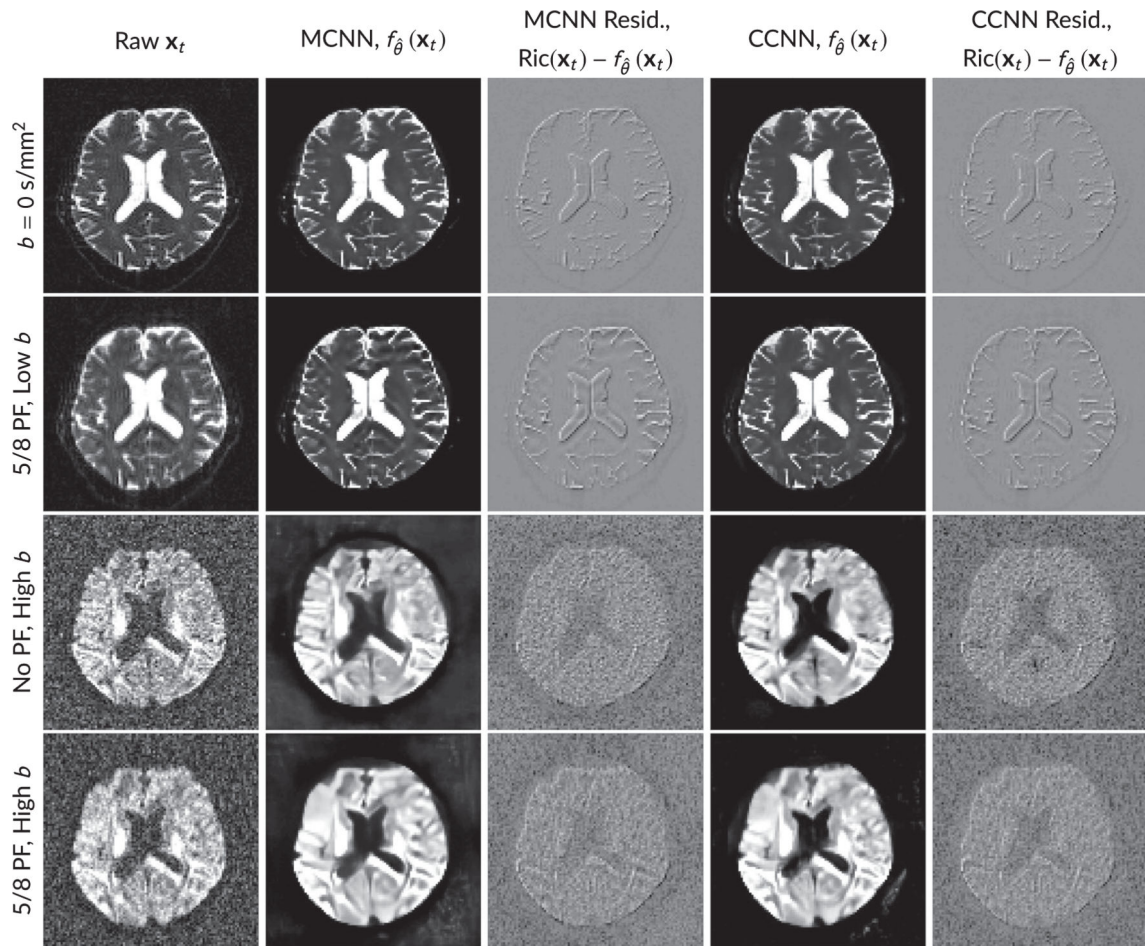
FIGURE 6.

Examples of images from in vivo data at low b ($b = 0$ s/mm^2) and high b ($b = 1500$ s/mm^2). Artifacts in the Raw image, $\mathbf{x}_t$, are corrected by the MCNN and CCNN models, $f_{\hat{\theta}}(\mathbf{x}_t)$. Also shown are the residuals between the CNN corrections and the original Raw image with Rician bias correction Ric($\mathbf{x}_t$).[10] The Gibbs and noise artifacts removed by the methods are observed in the residuals. Gibbs removal is prevalent at $b = 0$ s/mm^2, whereas noise removal is more prevalent at $b = 1500$ s/mm^2

FIGURE 7.

DW and MD images from various methods performed on prospective diffusion data. At the top are shown the $b = 0$ s/mm^2 and average $b = 1000$ s/mm^2 images from Raw, SoA, MCNN, standard partial Fourier, and CCNN methods. Rows 3 and 4 show results for mean diffusivity. The SoA method and both deep learning methods perform well without partial Fourier acceleration; however, at the 5/8ths partial Fourier factor, artifacts are present for all methods other than the CCNN method
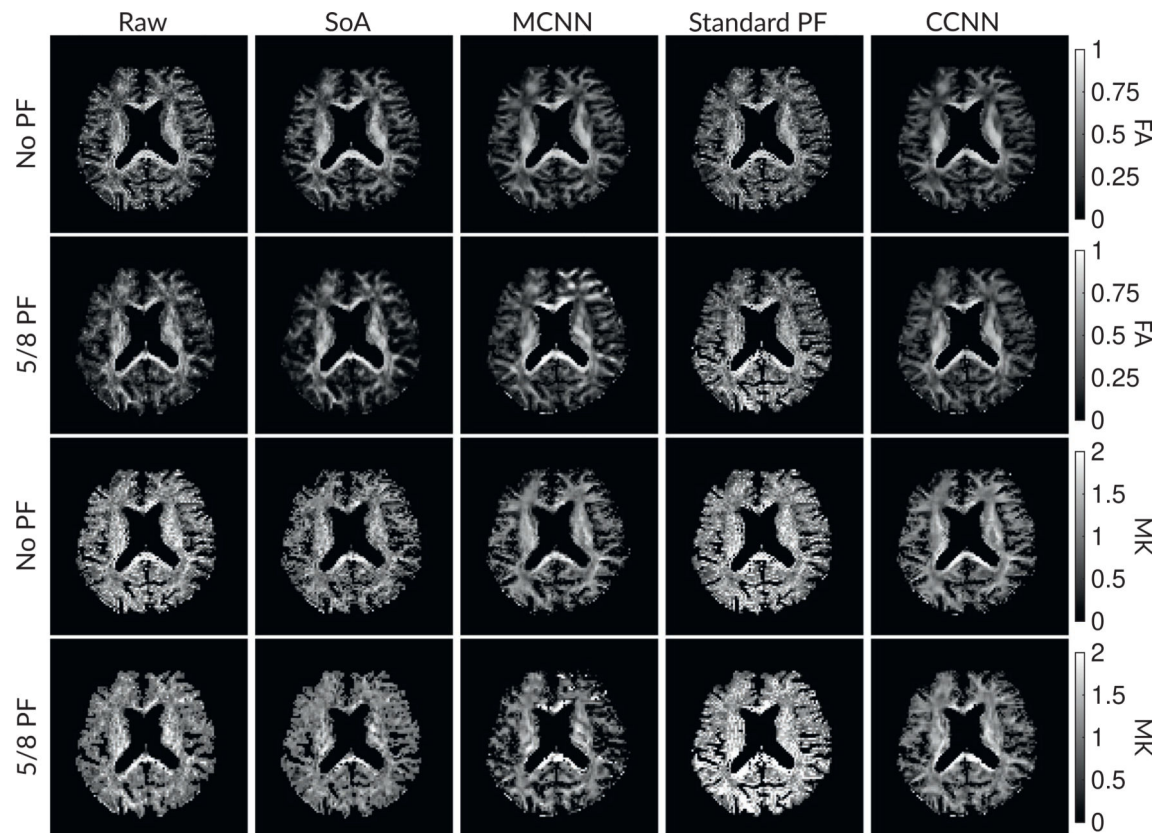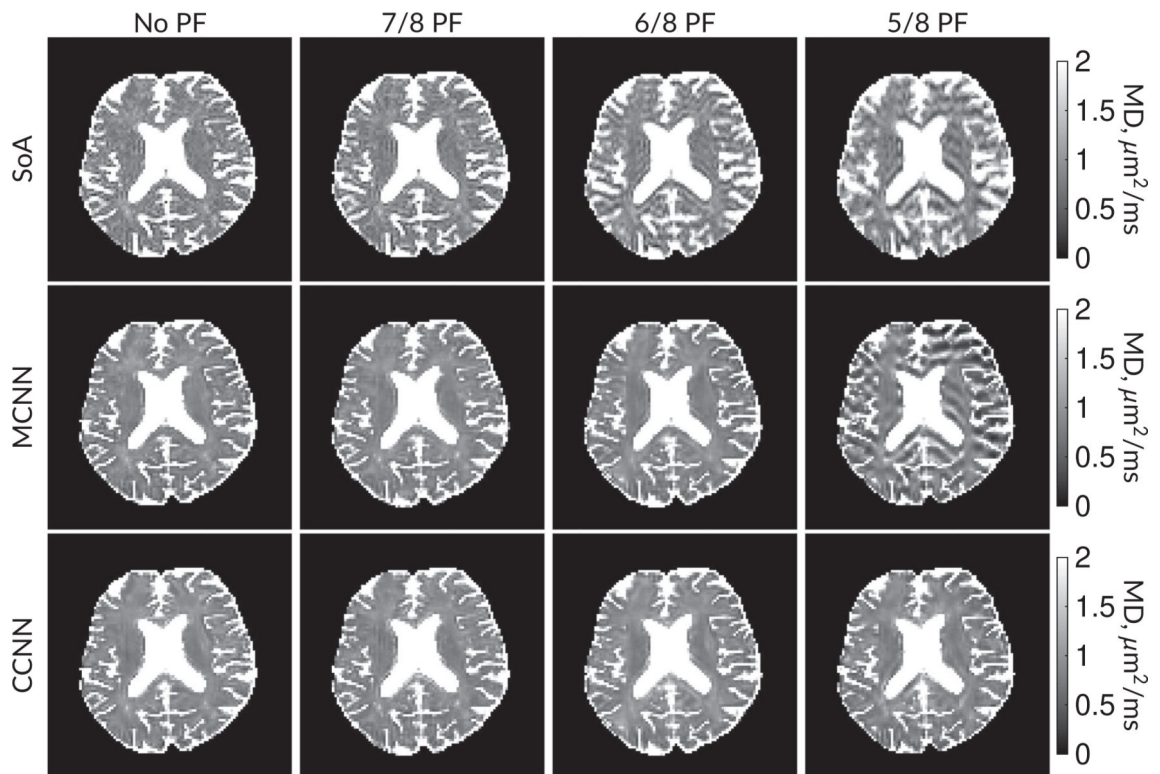
FIGURE 8.
Parameter maps from various methods with CSF masks. Rows 1 and 2 show the FA images from Raw, SoA, MCNN, standard partial Fourier, and CCNN methods, with MK images being in rows 3 and 4. Succeeding rows alternate the inclusion of partial Fourier encoding

FIGURE 9.

Comparison of mean diffusivity (MD) parameter maps across PF factors of 5/8ths, 6/8ths, 7/8ths, and without PF (No PF). The methods include SoA, MCNN, and CCNN methods. Without partial Fourier, all methods are similar; however, as the PF factor increases, substantial artifacts are introduced in the SoA and MCNN methods. The CCNN method is able to reduce the introduction of artifacts
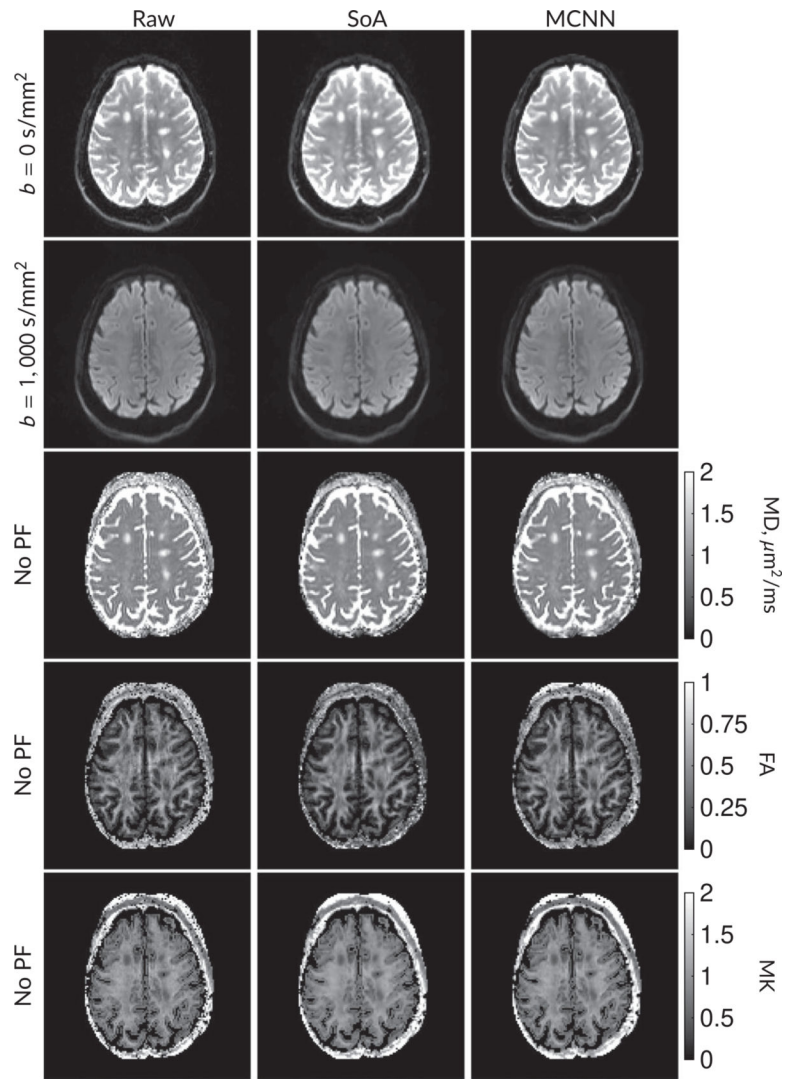
FIGURE 10.

Parameter maps calculated for a patient with multiple sclerosis. MS lesions are visible in all three parameter maps. Both the SoA and magnitude-input CNN (MCNN) methods are able to denoise parameter maps relative to the original (Raw) data. Images zoomed on the lesions are shown in the Supporting Information