

Mitigating transmit- B_1 artifacts by predicting parallel transmission images with deep learning: A feasibility study using high-resolution whole-brain diffusion at 7 Tesla

Xiaodong Ma  | Kâmil Uğurbil | Xiaoping Wu 

Center for Magnetic Resonance Research,
Radiology, Medical School, University of
Minnesota, Minneapolis, Minnesota, USA

Correspondence

Xiaodong Ma, Center for Magnetic
Resonance Research, University of
Minnesota, 2021 6th Street SE,
Minneapolis, MN 55455, USA.
Email: ma000311@umn.edu

Funding information

National Institutes of Health, Grant/Award
Numbers: U01 EB025144, P41 EB015894,
and P30 NS076408

Purpose: To propose a novel deep learning (DL) approach to transmit- B_1 (B_1^+)-artifact mitigation without direct use of parallel transmission (pTx), by predicting pTx images from single-channel transmission (sTx) images.

Methods: A deep encoder–decoder convolutional neural network was constructed and trained to learn the mapping from sTx to pTx images. The feasibility was demonstrated using 7 T Human-Connectome Project (HCP)-style diffusion MRI. The training dataset comprised images acquired on 5 healthy subjects using commercial Nova RF coils. Relevant hyperparameters were tuned with a nested cross-validation, and the generalization performance evaluated using a regular cross-validation.

Results: Our DL method effectively improved the image quality for sTx images by restoring the signal dropout, with quality measures (including normalized root-mean-square error, peak SNR, and structural similarity index measure) improved in most brain regions. The improved image quality was translated into improved performances for diffusion tensor imaging analysis; our method improved accuracy for fractional anisotropy and mean diffusivity estimations, reduced the angular errors of principal eigenvectors, and improved the fiber orientation delineation relative to sTx images. Moreover, the final DL model trained on data of all 5 subjects was successfully used to predict pTx images for unseen new subjects (randomly selected from the 7 T HCP database), effectively recovering the signal dropout and improving color-coded fractional anisotropy maps with largely reduced noise levels.

Conclusion: The proposed DL method has potential to provide images with reduced B_1^+ artifacts in healthy subjects even when pTx resources are inaccessible on the user side.

KEYWORDS

deep learning, diffusion MRI, Human Connectome Project, parallel transmission, ultrahigh field MRI

1 | INTRODUCTION

Ultrahigh field (UHF) MRI systems (operating at a field strength of 7 T and above) offer a practical solution for boosting image SNR and thereby, pushing the limit of image resolution, which have shown tremendous value in both clinical and neuroscience applications.¹ One example fMRI and diffusion MRI (dMRI) with high spatiotemporal resolutions achieved via the 7 T Human Connectome Project (HCP).²⁻⁴ However, a major challenge at UHF is the severe transmit B_1 (B_1^+) inhomogeneity encountered when using a conventional single-channel transmit RF volume coil.^{5,6} The B_1^+ inhomogeneity, if not corrected, can result in flip-angle variations across the brain, which in turn may yield variations in tissue contrast or even signal dropout, especially in lower brain regions such as the temporal lobe and cerebellum.

An effective way to address the B_1^+ inhomogeneity at UHF is by RF parallel transmission (pTx),⁷⁻¹³ a technique that uses a multi-channel RF transmit system and allows channel-specific RF pulse shapes to be applied. It is shown that pTx can substantially improve flip-angle uniformity and eliminate signal dropout across the brain when compared with conventional single-channel transmission (sTx). Additionally, pTx allows for control of RF power deposition in tissues (i.e., specific absorption rate [SAR]) by incorporating corresponding power constraints into the formulation of the pulse design problem.¹⁴⁻¹⁶ To date, pTx has been demonstrated in many UHF MRI applications including high-quality structural brain¹⁷⁻²⁰ and body²¹⁻²³ imaging, whole-brain high-resolution BOLD fMRI^{24,25} and dMRI.²⁶

However, the conventional pTx workflow is tedious and relies on special expertise. First, it requires on-fly calibration scans to be obtained in each subject to acquire prior knowledge of channel-specific B_1^+ maps and main field inhomogeneity (ΔB_0) maps for the subsequent pulse design. Second, it usually involves solving a non-convex optimization problem for pTx pulse waveforms,^{15,27} which can take up to several minutes or even longer especially when designing large-tip-angle pulses.²⁸ All these have been a hurdle that prevents pTx from being widely adopted in the UHF community.

Recently, great strides have been made to develop solutions for a user-friendly pTx workflow. One effective solution is the universal pulse (UP) method.²⁹ Instead of online pTx pulse design during the scan session, the UP method aims to pre-calculate a universal pTx pulse waveform using B_1^+ and ΔB_0 field maps acquired from a representative sample of adult population. It is shown that such UP can improve flip angle homogeneity across the human brain relative to the circularly polarized mode when used in a new human subject. The use of the UP

method is demonstrated for 7 T neuroimaging with various contrasts,^{25,29,30} and 7 T body imaging.³¹

Other solutions for a user-friendly pTx workflow include using machine learning in pulse design. SmartPulse has been introduced to improve the robustness of the UP method against the inter-subject variation and demonstrated useful for liver imaging at 3 T.³² In this method, a number of group-wise optimized UP candidates are optimized and 1 UP is selected during the scan depending on the features of the subject under scan. Another approach is machine-learning-based RF shimming proposed by Ianni et al.³³ Based on the assumption that RF field distribution is highly dependent on the shape of the subject, this method uses kernelized ridge regression to predict the RF magnitudes and phases of individual transmit channels (or RF shims) from several shape-related features, with reduced computational time and minimal requirement of B_1^+ mapping (only the k-space center line of B_1^+ maps is needed). It is shown that this method can be used to obtain SAR-efficient RF shims that can produce B_1^+ uniformities comparable to conventional RF shimming method.

Although effective in reducing the pTx workflow, the above-mentioned methods still require that the user has access to the pTx hardware, which is usually expensive. Here, we propose a novel deep-learning (DL) framework that aims to train a deep neural network to directly predict pTx-style images from those obtained with sTx. This is purely an image-to-image mapping method, which does not require any pTx expertise, pTx software or pTx hardware on the user side. Here, we demonstrate the feasibility of our DL method for creating 7 T HCP-style dMRI free of B_1^+ artifacts in healthy subjects. Our results show that the proposed method can substantially enhance image quality and improve downstream diffusion analysis relative to sTx acquisitions.

2 | METHODS

2.1 | Training dataset

Our training dataset was created based on our previous data acquisition aimed at demonstrating the use of pTx for high-resolution, whole-brain dMRI at 7 T.²⁶ Specifically, it consisted of data acquired on 5 healthy subjects using a 7 T MR scanner (Siemens, Erlangen, Germany). All subjects signed a consent form approved by the local Institutional Review Board. For each subject, the data comprised a pair of matched, 1.05-mm HCP-style dMRI datasets: one obtained with sTx using the commercial Nova single-channel transmit 32-channel receive (1Tx32Rx) coil and the other with pTx using the commercial Nova 8Tx32Rx coil. As in the 7 T HCP

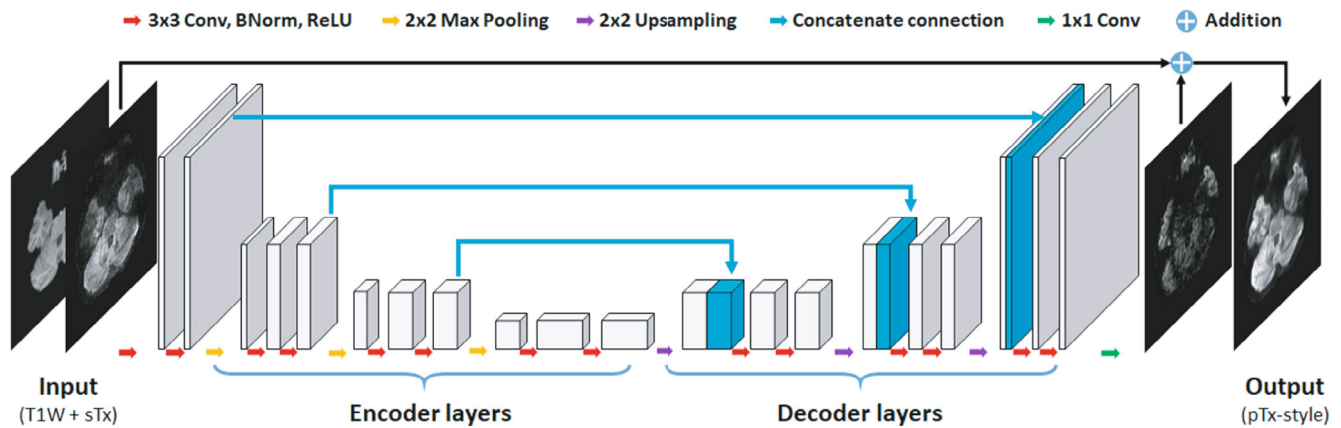


FIGURE 1 The architecture of the deep convolutional neural network developed in this study. The neural network takes single-transmission (sTx) diffusion and T_1 -weighted (T_1w) MPRAGE images as input and spits out the parallel-transmission (pTx) version of diffusion images. The core of the neural network is an encoder–decoder model with N encoder levels and N decoder levels (an example network with $N = 3$ is illustrated here). Each encoder level contains a 2×2 downsampling layer (via max pooling), followed by 2 repetitions of 3 consecutive layers: 3×3 convolution, batch-normalization and ReLU activation (3×3 Conv-BN-ReLU); each decoder level contains a 2×2 upsampling layer, followed by 2 repetitions of the 3×3 Conv-BN-ReLU operation. The number of channels are doubled after each downsampling layer, and are halved after each upsampling layer. Concatenated skip connections are added between corresponding encoder and decoder levels to reduce resolution loss, and a global addition connection is appended to promote the training performance by enabling residual learning. In this study, the number of encoder and decoder levels (i.e., N) and the number of output channels for the first layer were considered as hyperparameters and were tuned during the model selection process

dMRI protocol,³ the sTx acquisition used dielectric pads to improve the B_1^+ uniformities in the brain. Even with dielectric padding, the B_1^+ field of sTx was still inhomogeneous across the brain, with the coefficient of variation (i.e., SD/mean) of the whole-brain B_1^+ ranging from $\sim 22\%$ to $\sim 27\%$ for the 5 subjects scanned (Supporting Information Figure S1). Both sTx and pTx dMRI datasets consisted of 36 preprocessed image volumes (including 32 diffusion-weighted images with b-value of 1000 s mm^{-2} [b1000] and 4 b_0 images), with each volume having 100 slices covering the whole brain. The preprocessing was conducted following the HCP pipelines³⁴ for correction of head motion and geometric EPI distortions and for co-registration. As a result, the training dataset comprised 18 000 samples (5 subjects \times 100 slices \times 36 volumes), each sample being a pair of corresponding sTx and pTx image slices. The intensity of sTx and pTx images were independently normalized to the range of 0 to 1 to form the training dataset.

2.2 | Deep encoder–decoder convolutional neural network

We constructed a deep encoder–decoder convolutional neural network (CNN)^{35,36} for a specific realization of our DL method to predict pTx diffusion images given sTx ones. The reason we chose an encoder–decoder

CNN was its demonstrated use for various applications including image reconstruction,^{37–39} denoising,⁴⁰ and artifact correction.⁴¹ Specifically, our encoder–decoder CNN (Figure 1) comprised a same number, N , of encoder and decoder levels. Each encoder level contained a 2×2 downsampling layer (via max pooling) and 2 repetitions of 3×3 convolution, batch-normalization, and ReLU activation (3×3 Conv-BN-ReLU) operations; each decoder level contained a 2×2 upsampling layer and 2 repetitions of 3×3 Conv-BN-ReLU operations. The number of output channels was doubled after each downsampling layer, and was halved after each upsampling layer. Further, skip connections were added with concatenation operations to reduce resolution loss, and a global addition connection was appended to improve the training performance by enabling residual learning.

Our encoder–decoder CNN took T_1 -weighted (T_1w) MPRAGE images (acquired with sTx) as additional input because this was shown in our pilot study to result in better prediction accuracy than using sTx diffusion images alone as input. The T_1w images were acquired in the same subjects at 0.7-mm isotropic resolutions and were co-registered with the diffusion images. They were downsampled using the cubic interpolation and were 0 padded to have the same resolution and size as the diffusion images. To match the input and output dimensions of the neural network, the input images were 0

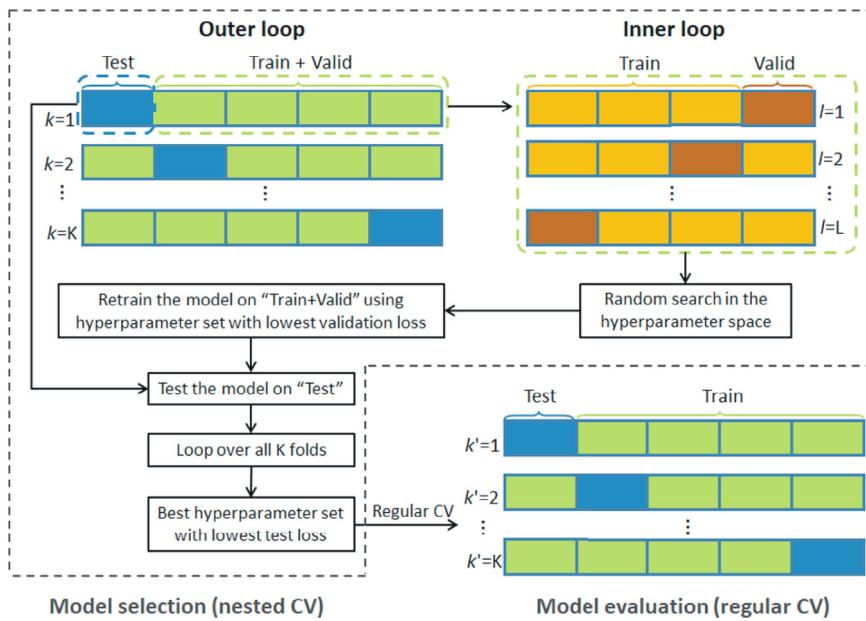


FIGURE 2 The diagram of model selection and evaluation. For model selection, a nested 5-fold (with dataset split into 3/1/1 for training/validation/testing and each fold comprising data of a single subject) cross-validation (CV) was used. Hyperparameter tuning was conducted in the inner loop by training the model on the training data and selecting the hyperparameter set with the minimum validation loss. In the outer loop, the model with tuned hyperparameters was trained on both training and validation data and the hyperparameter set with the minimum testing loss were selected. The generalization performance of the model with tuned hyperparameters was evaluated using a regular 5-fold (with dataset split into 4/1 for training/testing and each fold comprising data of a single subject) CV

padding from 173×207 to 224×224 , and the output images were cropped from 224×224 to 173×207 .

Our encoder–decoder CNN was implemented using the Flux package in Julia,⁴² and trained on a Linux workstation using a NVIDIA TITAN RTX GPU with 24GB memory. The loss function was formulated to measure the mean-squared-error (MSE) between the output of our encoder–decoder CNN and the pTx diffusion images. The minimization was conducted using the Adam algorithm.⁴³ For improved training performance, decaying learning rates were considered with the learning rate at the i -th epoch (LR_i) being:

$$LR_i = LR_0 e^{-DF \cdot \text{floor}\left(\frac{i-1}{DS}\right)}, \quad (1)$$

where LR_0 is the initial learning rate, DF the learning rate decay factor, and DS the learning rate decay step. The source code is publicly available at <https://github.com/XiaodongMa-MRI/ImageBasedDLForpTX.jl>.

2.3 | Model selection and evaluation

We conducted cross-validation (CV) (Figure 2) for model selection and evaluation. In model selection, a nested 5-fold CV (with dataset split into 3/1/1 for training/validation/testing and each fold comprising data of a single subject) was performed to tune relevant hyperparameters.⁴⁴ Six hyperparameters were considered: the number, N , of encoder or decoder levels, number of output channels for the first layer, mini-batch size, initial learning

rate, learning rate decay factor, and learning rate decay step. The hyperparameter tuning was carried out using a random search,⁴⁴ in which the optimal hyperparameter set (consisting of tuned hyperparameters) was chosen from a pool of 50 candidate hyperparameter sets created by randomly sampling the hyperparameter space (spanned by the 6 hyperparameters under consideration) (Supporting Information Table S1). More details about how the nested 5-fold CV was performed are provided in the Supporting Information.

In model evaluation, a regular 5-fold CV (with dataset split into 4/1 for training/testing and each fold comprising data of a single subject) was conducted to estimate the generalization performance of the model with the optimal hyperparameter set. The regular 5-fold CV involved 5 iterations. In each iteration, the model was trained on the training data and was tested on the testing data. The test loss averaged across all 5 iterations was calculated to evaluate the generalization performance.

In both model selection and evaluation, each training involved in the nested and regular CV used an early-stop strategy to reduce over fitting.⁴⁵ With the early-stop strategy, each training was conducted as follows: the model was trained using a total of 30 epochs; after every epoch (starting from the 11-th epoch) the trained model as well as the associated validation loss were recorded; from the 20 candidate epoch-specific trained models, the one with the lowest validation loss was selected as the final trained model. The training time was ~ 65 min when training on data of 4 subjects and was ~ 35 min when training on data of 3 subjects. The inference time was ~ 23 ms for prediction of a single slice.

2.4 | Evaluation of image quality

We further evaluated the quality of the predicted diffusion images by examining how close they would be to the pTx acquisitions in various brain regions. This was done by using image results from the regular 5-fold CV. Specifically, for each subject, the diffusion images predicted by the model with the best hyperparameter set (when trained on the other 4 subjects) were used to calculate region-specific metrics for image quality assessment, including normalized root-mean-square error (nRMSE), peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM),⁴⁶ and point spread function (PSF), all in reference to pTx acquisitions. In each case, 10 brain regions of interests (ROI) were considered, including 9 anatomic regions defined by the Montreal Neurological Institute (MNI) structural atlas⁴⁷ plus the whole brain. Note that the whole-brain ROI was defined as the brain mask calculated from the reference pTx images, where cerebrospinal fluid signals were excluded. The details of how each quality metric was calculated are provided in the Supporting Information. For comparison, region-specific nRMSE, PSNR, SSIM values, and PSF were also calculated using the sTx diffusion images. For each region-specific quality measure, a paired *t* test was used to determine whether there would be a difference between sTx and our DL method. A significant difference was inferred when the *P* value ≤ 0.05 .

2.5 | Diffusion analysis

We performed diffusion analysis to investigate how our DL method would improve DTI in comparison to sTx acquisitions. For this, the image results of the regular 5-fold CV were used. Specifically, for each subject, the diffusion images predicted by the model with the optimal hyperparameter set (when trained on the other 4 subjects) were used to fit the DTI model using FSL's dtifit routine⁴⁸ to derive fractional anisotropy (FA), mean diffusivity (MD), and fiber orientation vectors (i.e., the principal eigenvectors).

For both FA and MD, region-specific nRMSE, PSNR, SSIM, and PSF were calculated to measure the deviation from or similarity to the reference metric as obtained by fitting the DTI model to the pTx acquisitions. For fiber orientation vectors, region-specific angular errors of the principal eigenvectors were evaluated to quantify the angular difference from the reference principal eigenvectors derived from the pTx acquisitions. All region-specific quality measures for FA and MD maps and the region-specific angular errors of the principal eigenvectors were calculated for the same 10 brain ROIs as

in the aforementioned evaluation of image quality. For comparison, region-specific quality measures for FA and MD maps as well as the region-specific angular errors of the principal eigenvectors (all in reference to the pTx acquisitions) were also calculated based on the diffusion metrics derived by fitting the DTI model to the sTx diffusion images, and paired *t* tests were conducted with a significant difference being inferred when the *P* value ≤ 0.05 .

2.6 | Test on new data from the 7 T HCP database

The generalizability of our final DL model was further examined by training the final model on data of all our 5 subjects and using the trained model to predict pTx diffusion images for 5 new subjects (No. 102311, 102816, 221319, 525541, and 927359) randomly selected from the 7 T HCP database.⁴⁹ The training was performed using a total of 20 epochs. The training time was ~ 80 min, and the inference time for prediction of an entire image volume was ~ 26 s. For each subject, the predicted diffusion images were compared against the sTx acquisitions. The predicted diffusion images were also used to fit the DTI model and the results were compared with those obtained using sTx acquisitions.

3 | RESULTS

3.1 | Model selection and evaluation

Per the results from nested CV, our final DL model was created using the following hyperparameters: the number of encoder and decoder levels = 4, the number of output channels for the first layer = 24, mini-batch size = 24, and learning rate initialization/decay factor/decay steps = $5.87 \times 10^{-3}/0.28/4$. The generalizability of our model was high, with the mean test loss across folds being as low as $6.49 \times 10^{-5} \pm 2.16 \times 10^{-5}$ (mean \pm SD across 5 subjects or folds).

3.2 | Evaluation of image quality

The use of our DL method appeared to substantially improve the image quality for both b0 and b1000 acquisitions when compared with sTx (Figure 3), effectively restoring the signal dropout as observed in the lower brain region such as the temporal pole. The anatomic structure and image contrast of the restored signals were comparable with the reference pTx acquisitions, despite

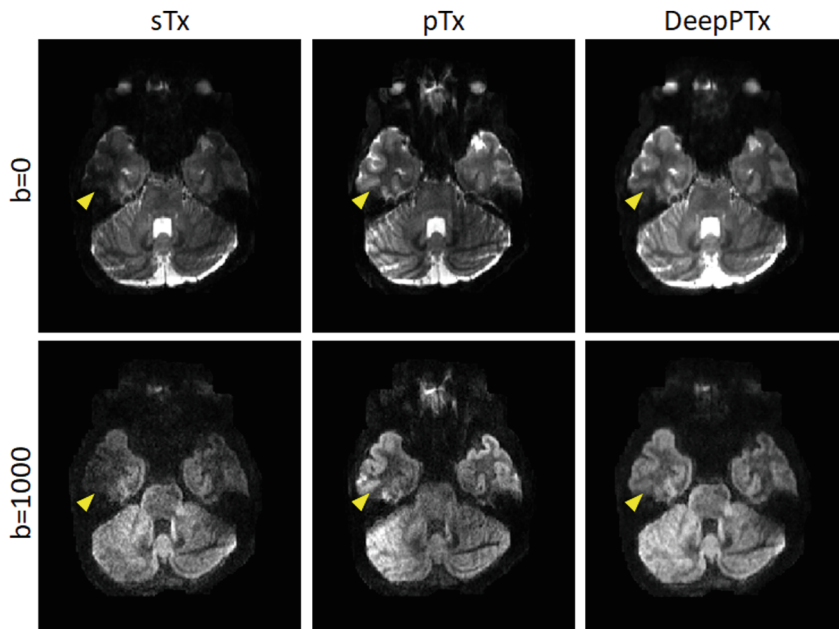


FIGURE 3 Comparing single transmission (sTx) diffusion images against their parallel transmission version as predicted by our deep learning (DL) method in reference to the images acquired using parallel transmission (pTx). Shown are example b_0 and b_{1000} diffusion images (1 diffusion direction) of a representative axial slice in lower brain from 1 subject, in which case the model with tuned hyperparameters was trained on data of the other 4 subjects. Note that the use of our DL method substantially improved the image quality by effectively recovering the signal dropout observed in the lower temporal lobe (as marked by the yellow arrowheads), producing images that were comparable to those obtained with parallel transmission

that slight image blurring was observed (e.g., in the cerebellum). Similar results were observed when inspecting the diffusion-weighted images for each individual subject (Supporting Information Figure S2).

Further quantitative analyses on quality measures (Figure 4) revealed that the use of our DL method decreased both region-specific nRMSE values and mean FWHM of PSF (averaged across both right–left and anterior–posterior directions) while increasing both region-specific PSNR and SSIM values in most brain regions relative to sTx acquisitions, with the improvement being significant for all quality measures when considering the following ROIs: whole brain, frontal lobe, temporal lobe, and insula. Quantitatively, the percentage changes of whole-brain nRMSE, PSNR, SSIM, and FWHM values (calculated as $|\text{val}_{\text{DL}} - \text{val}_{\text{sTx}}| / \text{val}_{\text{sTx}} * 100\%$, where val_{DL} is the quality measure of our DL method and val_{sTx} the quality measure of sTx, and “| |” denotes the absolute value) were 28%, 9%, 5%, and 0.8%, respectively. For the other 3 ROIs with a significant improvement in all quality measures, the percentage changes of nRMSE, PSNR, SSIM, and FWHM values ranged from 35% (temporal lobe) to 37% (insula), from 11% (frontal lobe) to 12% (temporal lobe), from 5% (insula) to 10% (temporal lobe), and from 0.6% (insula) to 1% (temporal lobe), respectively.

3.3 | Diffusion analysis

The use of our DL method improved DTI performances by substantially decreasing the fitting error, leading

to increased quality and accuracy for both FA and MD estimations especially in the lower temporal lobe when compared to the sTx acquisition (Figure 5). The sum-of-squared fitting error averaged across the whole brain decreased by as high as 68% (0.17 for the DL method vs. 0.53 for sTx acquisition) and appeared even lower than that of the reference pTx acquisition (0.17 for the DL method vs. 0.40 for pTx acquisition).

Further quantitative analyses on quality measures (Figure 6) showed that overall the results were in agreement with those of image quality evaluation. For both FA and MD, the use of our DL method decreased both region-specific nRMSE values and mean FWHM of PSF while increasing both region-specific PSNR and SSIM values in most brain regions relative to sTx acquisitions. For FA, the use of our method led to a significant improvement in all quality measures except for FWHM when considering the whole-brain ROI. The percentage changes of whole-brain nRMSE, PSNR, SSIM, and FWHM values were 10%, 5%, 2%, and 0.3%, respectively. The use of our method also significantly decreased nRMSE and significantly increased PSNR in the temporal lobe, with the percentage changes of nRMSE and PSNR being 23% and 14%, respectively. It also significantly increased SSIM and significantly decreased FWHM in 3 brain ROIs including caudate, cerebellum, and thalamus, with the percentage change of SSIM and FWHM ranging from 2% (cerebellum) to 5% (thalamus), and from 0.4% (cerebellum) to 0.7% (caudate), respectively. For MD, although decreasing the nRMSE value for most brain regions (7 of 10 ROIs), the use of our DL method did not bring a significant improvement to this quality measure in any brain

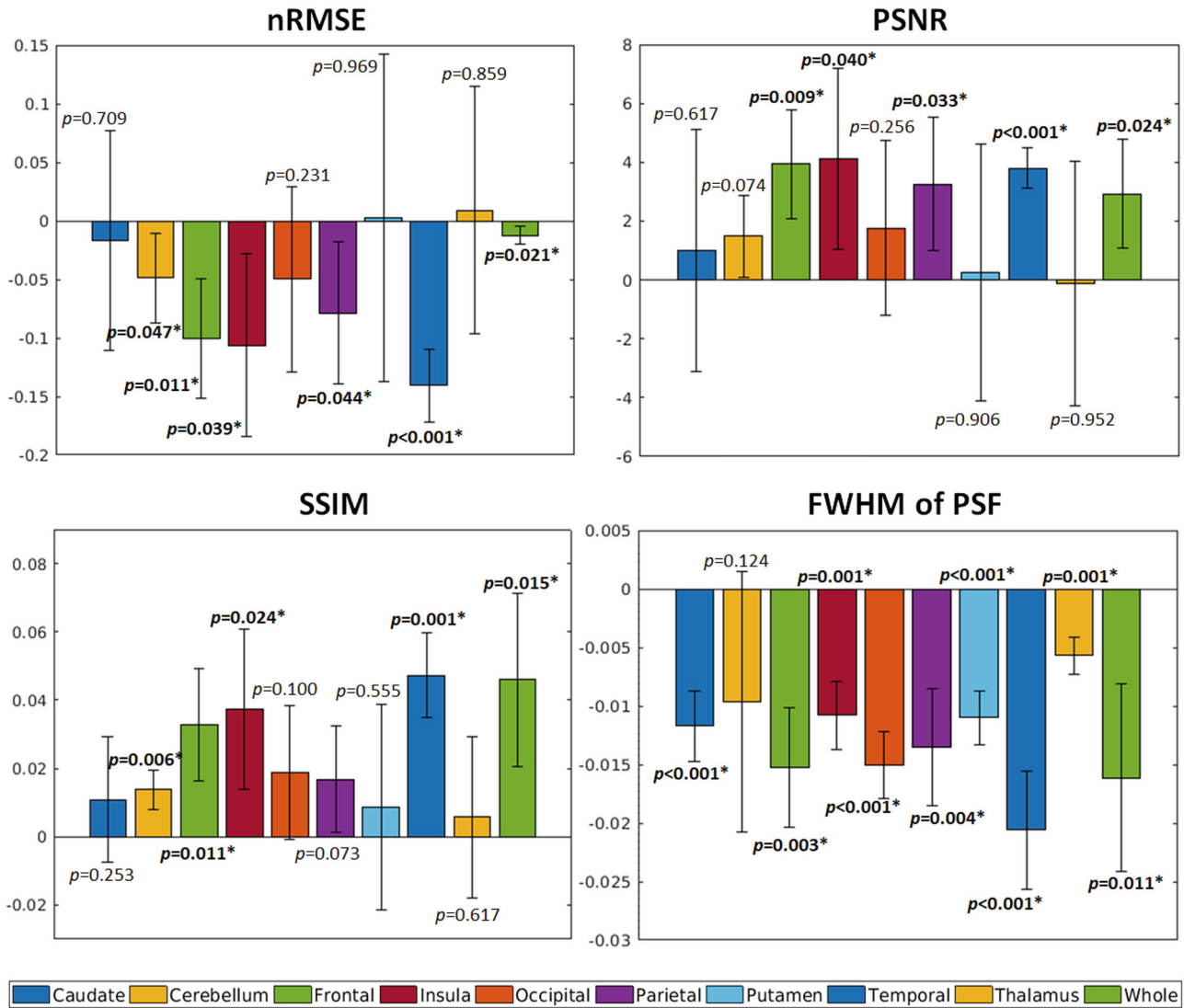


FIGURE 4 Comparing single transmission (sTx) and our deep neural network in terms of image quality. In each case, the image quality was evaluated using 4 quality measures: normalized root-mean-square error (nRMSE), peak signal-to-noise-ratio (PSNR), structural similarity index measure (SSIM), and FWHM of point spread function (PSF), all in reference to the acquisition with parallel transmission. Shown are mean and standard deviation (across 5 subjects) of the differences in region-specific nRMSE, PSNR, SSIM, and FWHM of PSF between sTx and our DL method (with the difference being the quality measure of the DL method minus that of sTx). For both sTx and the deep learning (DL) method, region-specific values of each quality measure were evaluated by considering a total of 10 brain regions of interest (including 9 brain regions as defined by the MNI152 standard-space structural atlas plus the whole brain). The numbers reported are the *P* values obtained from a paired *t* test, with significance being denoted by “*.” Note that the use of our DL method significantly decreased nRMSE values and FWHM of PSF, while increasing both PSNR and SSIM values in most brain regions including the whole brain

ROI. For those ROIs with an improvement, the percentage change of nRMSE ranged from 3% (insula) to 12% (temporal lobe). The use of our DL method also increased PSNR in all brain ROIs except for cerebellum and insula. For those 8 ROIs with an improvement, the improvement was found significant only in the parietal and temporal lobes, with the percentage change being 1% in both. However, the use of our method increased SSIM and decreased mean FWHM of PSF in all brain ROIs, with the improvement being significant in most brain ROIs including the

whole-brain ROI. For the ROIs with a significant improvement, the percentage change of SSIM and FWHM ranged from 2% (cerebellum) to 6% (temporal lobe) and from 0.2% (thalamus) to 2.3% (occipital lobe), respectively. For the whole-brain ROI, the percentage changes of nRMSE, PSNR, SSIM, and FWHM were 7%, 1%, 4%, and 1.2%, respectively.

Moreover, the use of our DL method improved the performances for the principal eigenvector estimation relative to sTx acquisition (Figure 7), producing better delineated

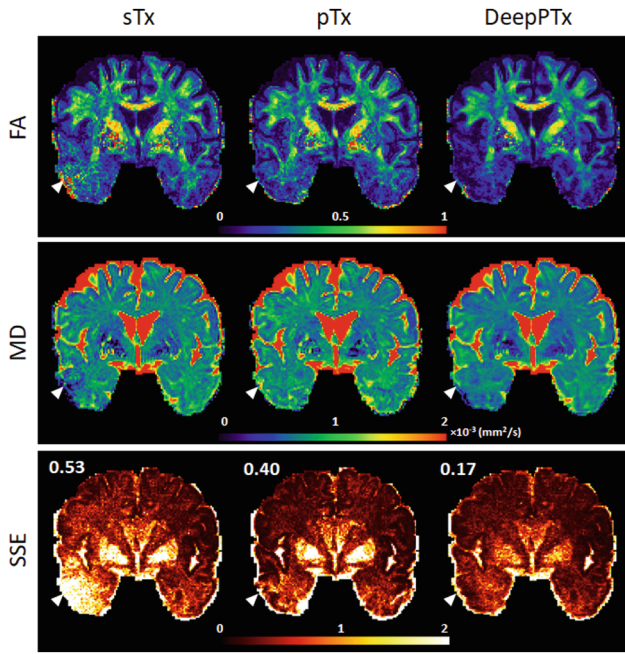


FIGURE 5 Comparing single transmission (sTx) against our deep learning (DL) method in reference to acquisition with parallel transmission (pTx) in terms of DTI performances. Shown are example fractional anisotropy (FA) and mean diffusivity (MD) maps of a representative coronal slice from 1 subject, in which case the neural network with tuned hyperparameters was trained on data of the other 4 subjects. Corresponding sum of squared fitting error maps are also shown, with numbers reported being the whole brain average sum-of-squared fitting error values. Note that the use of the DL method improved DTI performances by substantially decreasing the fitting error, leading to increased quality and accuracy for both FA and MD estimations especially in the lower temporal lobe (as indicated by arrowheads) when compared to sTx acquisition

fiber orientations that more closely resembled what was attainable with the pTx acquisition. This improvement was further verified by quantitative comparison of angular errors in reference to pTx acquisition (Figure 8)—the angular errors were significantly lower for our DL method than for sTx acquisition across all brain ROIs. The percentage change of angular errors ranged from 4% (occipital lobe) to 9% (caudate and thalamus), with its value being 6% for the whole-brain ROI. The percentage changes of all quality measures for raw images, FA, and MD, as well as of angular errors in each ROI are listed in Supporting Information Table S2.

3.4 | Test on new data from the HCP database

When used to test on unseen diffusion data of 5 new subjects from the 7 T HCP database, the final model

substantially enhanced the image quality for every subject by effectively restoring the signal dropout observed in the lower brain regions (including the cerebellum and the temporal pole), producing color-coded FA maps that presented largely reduced noise levels in those challenging regions (Figure 9).

4 | DISCUSSION

Here, we proposed and demonstrated a novel DL method that can be used in healthy subjects to create 7 T HCP-style diffusion images with reduced B_1^+ artifacts directly from images obtained using the sTx mode (which usually present strong B_1^+ artifacts). Essential to the efficacy of our method is the implementation of a deep encoder–decoder CNN with tuned hyperparameters (e.g., the number of encoder and decoder levels) and with facilitating structures (e.g., skip connections added between corresponding encoder and decoder levels to reduce resolution loss). The effectiveness of our method was demonstrated using 7 T HCP-style dMRI at ~ 1 -mm isotropic resolutions obtained in 5 healthy human subjects. The generalizability of our deep-learning model was estimated using 5-fold CV (i.e., training the deep encoder–decoder CNN with tuned hyperparameters on data of 4 subjects while testing the trained CNN on another subject), which shows that our method can be used to substantially improve the image quality relative to sTx acquisitions (Figures 3 and 4) and such improvement can translate into improved estimation of diffusion metrics (Figures 5–8) in the downstream diffusion analysis. The generalizability of our method was also examined by training the deep encoder–decoder CNN with tuned hyperparameters on data of all 5 subjects and testing the trained CNN on unseen data of 5 new subjects randomly chosen from the 7 T HCP database. The results (Figure 9) suggest that our method can enhance the image quality for 7 T HCP-style diffusion acquisition in healthy subjects by effectively restoring signal dropout present in the sTx images.

For the purpose of proving the principle, we conducted 7 T DTI to demonstrate the effectiveness of our method in improving the image quality for sTx acquisitions. The database that we acquired before following the 7 T HCP protocol actually has diffusion data obtained using a double-shell q-space sampling with b1000 and b2000. Here, we only used the diffusion images with b1000 and b0 images for DTI. Part of our future study is to include b2000 diffusion images and investigate how our method can be used to improve the quality for double-shell diffusion images.

The proposed DL method is based on the assumption that the transformation from sTx to pTx images can be

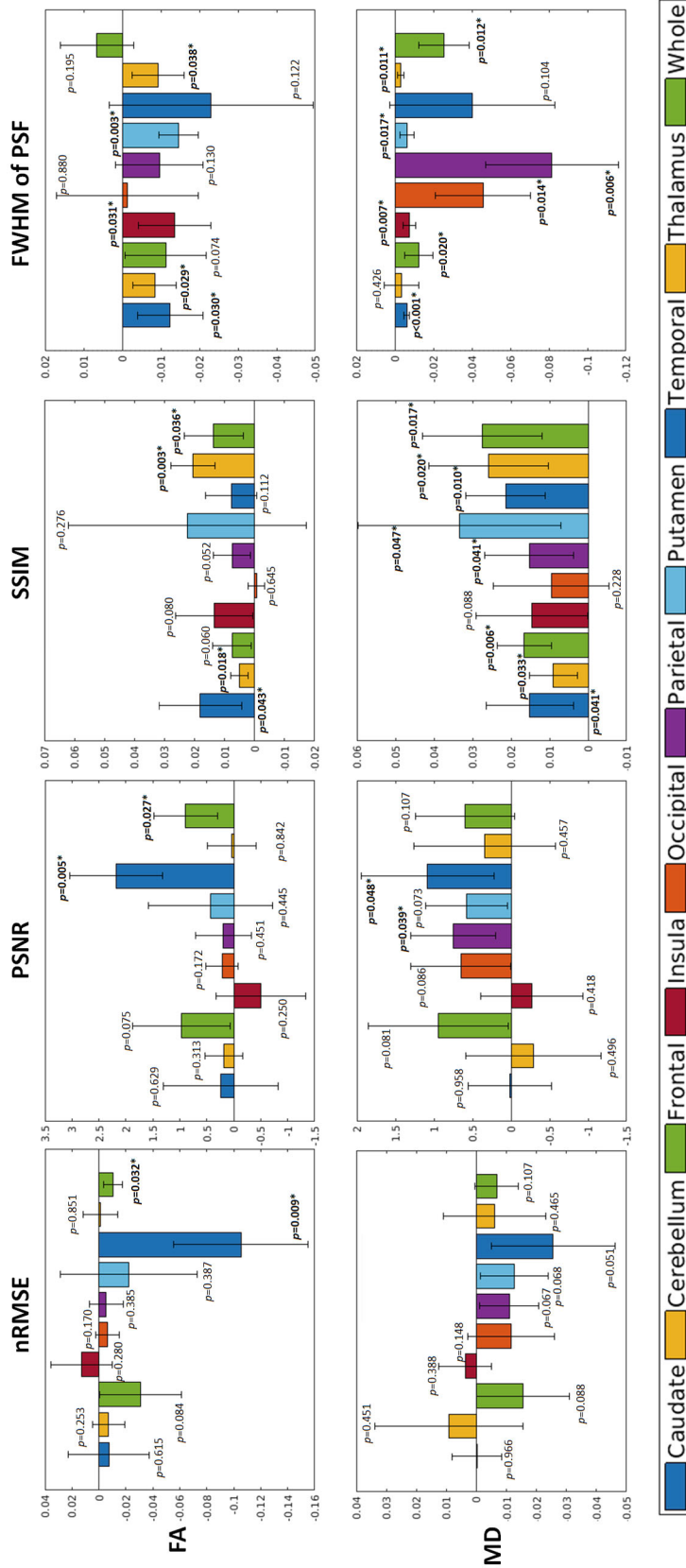


FIGURE 6 Comparing single transmission (sTx) and our deep learning (DL) method in terms of quality of diffusion tensor imaging metrics including fractional anisotropy (FA) and mean diffusivity (MD). The quality of each metric was evaluated using 4 quality measures: normalized root-mean-square error (nRMSE), peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), and FWHM of point spread function (PSF), all in reference to the acquisition with parallel transmission. For each metric, shown are mean and standard deviation (across 5 subjects) of the differences in region-specific nRMSE, PSNR, SSIM and FWHM of PSF between sTx and our DL method (with the difference being the quality measure of DL method minus that of sTx). For all cases, the region-specific values of each quality measure were evaluated by considering a total of 10 brain regions of interest (including 9 brain regions as defined by the MNI152 standard-space structural atlas plus the whole brain). The numbers reported are the P values obtained from a paired t test, with significance being denoted by “*.” For both FA and MD, note how the use of our DL method decreased nRMSE and FWHM of PSF, while increasing PSNR and SSIM in most brain regions

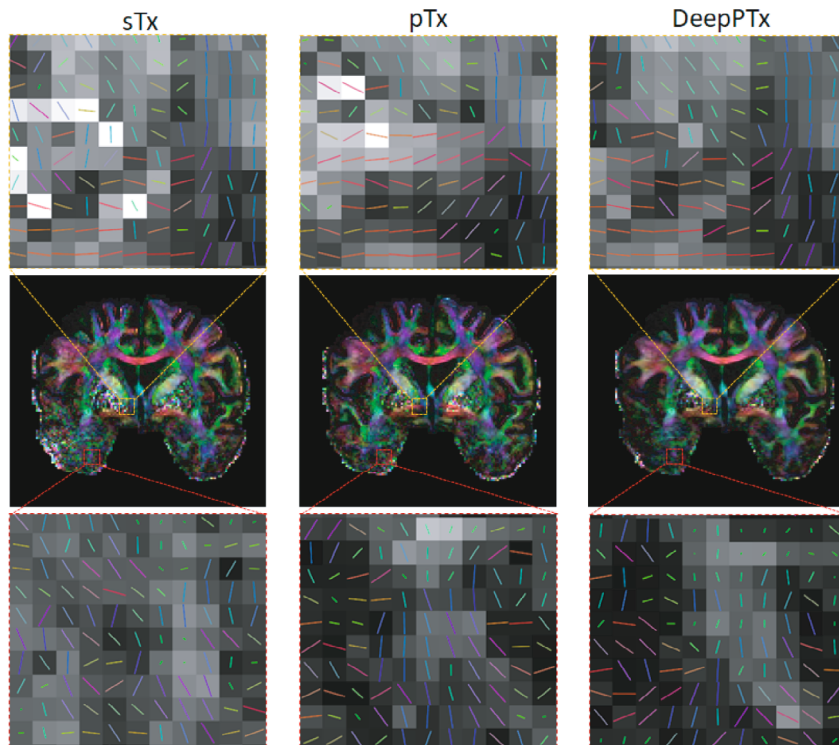


FIGURE 7 Comparing single transmission (sTx) against our deep learning (DL) method in reference to the acquisition with parallel transmission (pTx) in terms of performances for principal eigenvector estimation. Shown are example color-coded fractional anisotropy (FA) maps (middle row) of a representative coronal slice from 1 subject (in which case the neural network with tuned hyperparameters was trained on data of the other 4 subjects), along with 2 zoomed-in maps of principal eigenvector orientations overlaid on the respective FA maps: one in the putamen (top row) and the other in the temporal lobe (bottom row). Note that the use of our DL method improved principal eigenvector estimation performances by producing better delineated fiber orientations that more closely resembled what was attainable with the pTx acquisition

interpreted by a neural network (e.g., CNN). Intuitively, this is viable because the difference between sTx and pTx images mostly occurs in some local regions such as lower temporal lobe and cerebellum, and this local difference could be represented by a convolution operation with a relatively small kernel. Here, we choose the encoder–decoder CNN as the neural network, which has been proven effective in various image mapping applications including reconstruction and post-processing for MRI images. Our results show that the trained encoder–decoder CNN is able to restore signal dropouts in sTx images caused by B_1^+ inhomogeneity, while preserving local structure and overall image contrast.

Critical to the efficacy of our method is the implementation of a nested CV to tune relevant hyperparameters associated with the network structure and training configuration for our network. Here, 6 hyperparameters are tuned: (1) the number of encoder and decoder levels, (2) the number of output channels for the first layer, (3) mini-batch size, (4) learning rate initialization, (5) learning rate step size, and (6) learning rate decay factor. The first 2 hyperparameters determine the depth and width of the encoder–decoder CNN, thereby allowing for control over the model capacity and fitting errors. The other 4 hyperparameters determine how the neural network is trained and are commonly considered for tuning when training a deep learning model with the stochastic gradient descent optimization⁴³ to achieve improved training performances. Our results suggest that

tuning of these 6 hyperparameters is effective for optimizing our encoder–decoder CNN and can help produce a deep-learning model with high generalization performances.

In its current implementation, our method takes the T_1w images as an additional input. We found in our pilot study that this can help improve the prediction performance (mostly with sharp edges preserved) as compared to when using sTx images alone as input. This may be explained by the fact that the T_1w images bring in more features (e.g., tissue structures) for the neural network to learn and use. We note that this strategy has also been adopted in other studies aimed at diffusion image processing using DL with CNN.^{41,50} Part of our future work is to explore whether similar prediction performances can be achieved without using T_1w images by refining the neural network.

Consistent with the visual inspection of the raw images (Figure 3), the quantitative comparison of quality measures on raw images for our method versus sTx (Figure 4) shows that the images obtained from our method presented a large quality improvement in the temporal lobe, with the improvement being significant for all quality measures considered (i.e., nRMSE, PSNR, SSIM, and FWHM of PSF). Although the improvement in image quality overall translated into an improvement in DTI performances (Figure 5), quantitative comparison (Figure 6) shows that the improvements in FA and MD estimations in the temporal lobe were not significant for all quality measures,

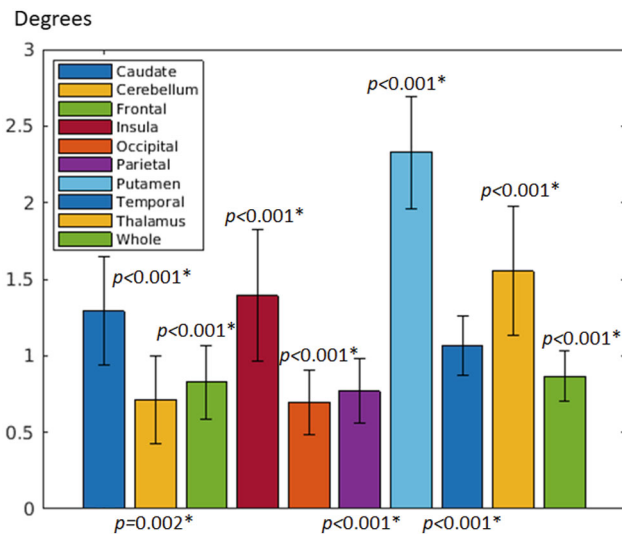


FIGURE 8 Comparing single transmission (sTx) and our deep learning (DL) method in terms of angular errors of principal eigenvector estimations. Shown are mean and standard deviations (across 5 subjects) of the differences in angular errors between sTx and our DL method (with the difference being the angular error of sTx minus that of the DL method). For both sTx and our DL method, region-specific angular errors were evaluated in reference to the acquisition with parallel transmission and by considering a total of 10 brain regions of interest (including 9 brain regions as defined by the MNI152 standard-space structural atlas plus the whole brain). The numbers reported are the P values obtained from a paired t test, with significance being denoted by “*.” Note that the use of our DL method substantially improved the performances for principal eigenvector estimations, significantly decreasing the angular errors across all the selected brain regions

with the improvement in SSIM on FA, the improvement in nRMSE on MD, and the improvement in FWHM on both FA and MD, all being found insignificant. Further, the improvements in both nRMSE and PSNR in the whole-brain ROI were not significant for MD estimation, whereas they were significant for raw images. We also found that for both FA and MD estimations, the percentage changes (relative to sTx) of the quality measures overall were lower than those for the raw images. These discrepancies can be explained by the fact that the ratio of b_0 and b_{1000} images is used in the DTI model fitting, thereby flattening the improvement with our method. The comparison of angular errors for principal eigenvectors (Figure 8), however, strongly suggests that our method can improve the delineation of fiber orientations, and the improvement is significant in all brain ROIs including the whole-brain ROI. This indicates that our method can improve the image quality while preserving signal variations among different diffusion directions, which in turn leads to more accurate fiber orientation depiction than the sTx acquisition.

The test of our final model on 5 new subjects from the 7T HCP database (Figure 9) suggests that our method can improve the quality of diffusion images obtained using conventional sTx methods, restoring signal dropouts and enhancing the quality of FA maps with higher SNR and clearer fiber orientation representation. This is encouraging as it shows that the neural network trained on data of only 5 subjects is already well generalizable and can be used to create diffusion images with mitigated B_1^+ artifacts given data of a new subject.

Our results (Figures 4 and 6) suggest that our deep-learning method yields little blurring (if any) in the raw images and FA and MD maps in terms of FWHM of PSF. A closer look at the PSF profiles (Supporting Information Figure S3) indeed revealed that the use of our method barely changed the shape of PSF profile relative to sTx acquisition. However, we notice that subtle blurring effects can be perceived especially in FA and MD maps, smoothing the image details in some local brain regions. To further investigate this blurring effect, we quantified the edges for raw images and FA and MD maps. The results (Supporting Information Figure S4) were compared to those of sTx acquisition showing that although enhancing the edges on raw images in all brain ROIs, the use of our method appeared to degrade the edges on both FA and MD maps especially in regions with strong B_1^+ artifacts (e.g., lower temporal lobe). Part of our future work is to investigate how best to reduce such blurring effects. A potential solution is to modify the loss function by introducing extra terms that can promote image edges (e.g., a term associated with image gradients),⁵¹ but likely at the cost of image fidelity in terms of nRMSE.

To investigate how our DL model would perform at different levels of signal dropout, we conducted a simulation study. A representative single-direction b_{1000} image slice obtained with sTx was chosen from 1 subject to show typical signal dropout in the temporal pole, and the region of signal dropout was manually defined as an ROI. The signal within the ROI was modulated by multiplying with 6 scaling factors (ranging from 0 to 1 in steps of 0.2) to generate 6 input images mimicking sTx acquisitions with different levels of signal dropout. Each of these input images was fed into our DL model to produce a predicted image. Our results (Supporting Information Figure S5) show that our DL model appeared capable of recovering the signal to some extent even for complete signal void; the prediction performances, however, decreased with increasing levels of signal dropout, with the improvement in nRMSE (evaluated within the ROI in reference to pTx acquisition) decreasing from $\sim 39\%$ (for the scaling factor of 1 corresponding to original signal dropout) to $\sim 15\%$ (for the scaling factor of 0 corresponding to complete signal void).

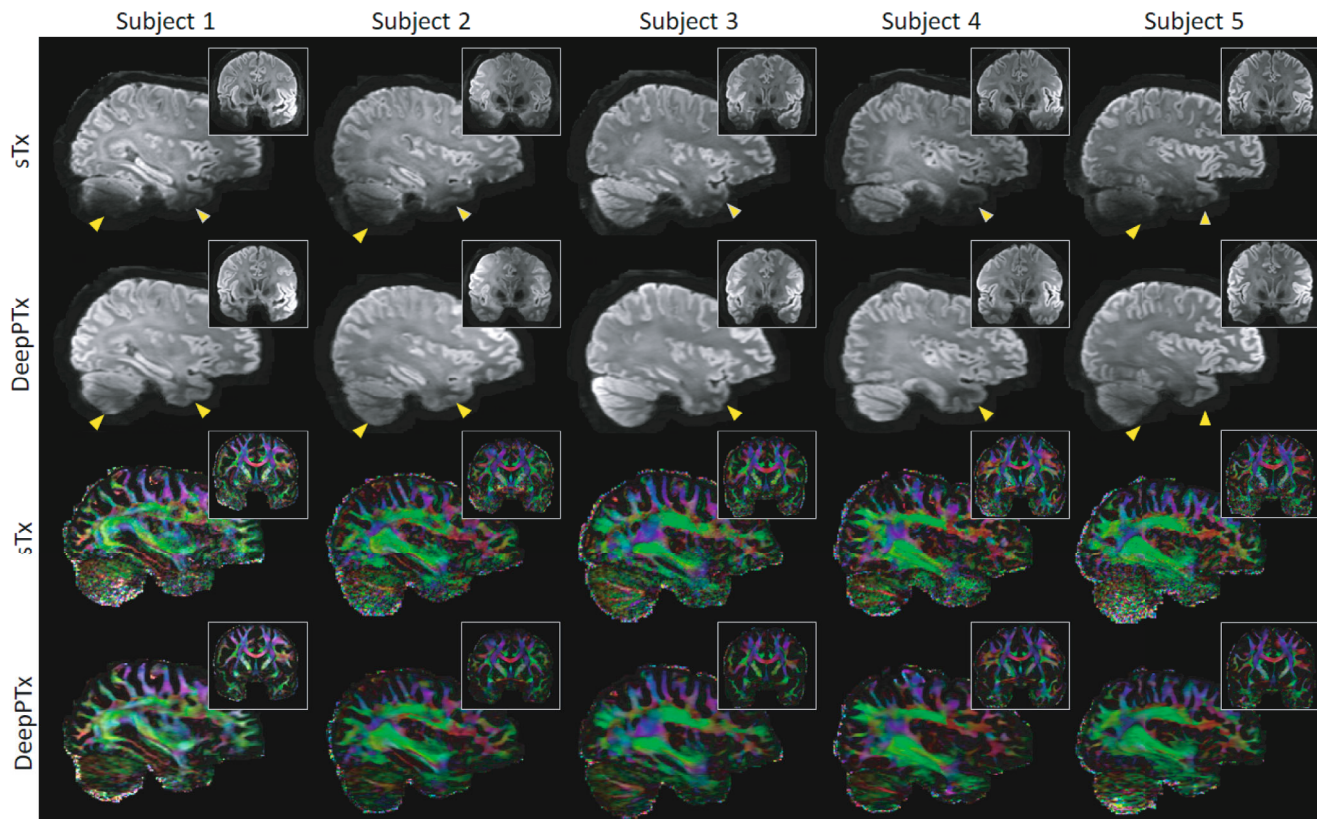


FIGURE 9 Testing of our final deep learning model on unseen data of 5 new subjects randomly chosen from the 7 T Human Connectome Project database. For each case, shown are mean diffusion-weighted images (with b1000) averaged across all diffusion directions (first and second rows) and color-coded fractional anisotropy (FA) maps (third and fourth rows), all in a representative sagittal slice with the insert showing a representative coronal slice. The final model with tuned hyperparameters was trained on our entire dataset of 5 subjects. Note that the use of our final model substantially enhanced the image quality by effectively restoring the signal dropout observed in the lower brain regions (as marked by yellow arrowheads), producing color-coded FA maps that presented largely reduced noise levels in those challenging regions

There are several aspects we can work on to improve our DL method. First, we can collect more data for improved training and prediction performance. Second, we may improve the prediction performance by exploring the use of other machine-learning frameworks. For example, we may consider the Generative Adversarial Networks (GANs),⁵² which has been shown effective in recovering lost signal in BOLD fMRI.⁵³ Moreover, it is valuable to investigate how our method would work in clinical applications where images have pathology.

5 | CONCLUSIONS


We have introduced and demonstrated a DL approach that can create 7 T HCP-style diffusion images free of B_1^+ artifacts for healthy subjects given images acquired with sTx. Our results show that our approach can substantially improve image quality by effectively restoring the signal dropout present in sTx images, thereby improving the

downstream diffusion analysis. As such, our approach has great potential to minimize transmit- B_1 -related artifacts even when pTx resources (including pTx expertise, pTx software or pTx hardware) are inaccessible on the user side.

ACKNOWLEDGMENTS

We thank John Strupp, Brian Hanna, and Jerahmie Radder for their assistance in setting up computation resources. This work was supported in part by National Institutes of Health (NIH) grants U01 EB025144, P41 EB015894, and P30 NS076408.

ORCID

Xiaodong Ma  <https://orcid.org/0000-0001-7158-8073>

Xiaoping Wu  <https://orcid.org/0000-0001-6021-9088>

REFERENCES

1. Ugurbil K. Imaging at ultrahigh magnetic fields: history, challenges, and solutions. *Neuroimage*. 2018;168:7-32.

2. Ugurbil K, Xu JQ, Auerbach EJ, et al. Pushing spatial and temporal resolution for functional and diffusion MRI in the human Connectome project. *Neuroimage*. 2013;80:80-104.
3. Vu AT, Auerbach E, Lenglet C, et al. High resolution whole brain diffusion imaging at 7 T for the human connectome project. *Neuroimage*. 2015;122:318-331.
4. Vu AT, Jamison K, Glasser MF, et al. Tradeoffs in pushing the spatial resolution of fMRI for the 7T human connectome project. *Neuroimage*. 2017;154:23-32.
5. Van De Moortele PF, Akgun C, Adriany G, et al. B-1 destructive interferences and spatial phase patterns at 7 T with a head transceiver array coil. *Magn Reson Med*. 2005;54(6):1503-1518.
6. Vaughan JT, Garwood M, Collins CM, et al. 7T vs. 4T: RF power, homogeneity, and signal-to-noise comparison in head images. *Magn Reson Med*. 2001;46:24-30.
7. Katscher U, Bornert P, Leussler C, van den Brink JS. Transmit SENSE. *Magn Reson Med*. 2003;49:144-150.
8. Adriany G, Van de Moortele PF, Wiesinger F, et al. Transmit and receive transmission line arrays for 7 Tesla parallel imaging. *Magn Reson Med*. 2005;53:434-445.
9. Zelinski AC, Wald LL, Setsompop K, et al. Fast slice-selective radio-frequency excitation pulses for mitigating B+1 inhomogeneity in the human brain at 7 Tesla. *Magn Reson Med*. 2008;59:1355-1364.
10. Cloos MA, Boulant N, Luong M, et al. kT-points: short three-dimensional tailored RF pulses for flip-angle homogenization over an extended volume. *Magn Reson Med*. 2012;67:72-80.
11. Zhu Y. Parallel excitation with an array of transmit coils. *Magn Reson Med*. 2004;51:775-784.
12. Grissom W, Yip CY, Zhang Z, Stenger VA, Fessler JA, Noll DC. Spatial domain method for the design of RF pulses in multicoil parallel excitation. *Magn Reson Med*. 2006;56:620-629.
13. Setsompop K, Alagappan V, Gagoski B, et al. Slice-selective RF pulses for in vivo B(1)(+) inhomogeneity mitigation at 7 Tesla using parallel RF excitation with a 16-element coil. *Magn Reson Med*. 2008;60:1422-1432.
14. Padormo F, Beqiri A, Hajnal JV, Malik SJ. Parallel transmission for ultrahigh-field imaging. *NMR Biomed*. 2016;29:1145-1161.
15. Hoyos-Idrobo A, Weiss P, Massire A, Amadon A, Boulant N. On variant strategies to solve the magnitude least squares optimization problem in parallel transmission pulse design and under strict SAR and power constraints. *IEEE Trans Med Imaging*. 2014;33:739-748.
16. Guerin B, Gebhardt M, Cauley S, Adalsteinsson E, Wald LL. Local specific absorption rate (SAR), global SAR, transmitter power, and excitation accuracy trade-offs in low flip-angle parallel transmit pulse design. *Magn Reson Med*. 2014;71:1446-1457.
17. Cloos MA, Boulant N, Luong M, et al. Parallel-transmission-enabled magnetization-prepared rapid gradient-echo T1-weighted imaging of the human brain at 7 T. *Neuroimage*. 2012;62:2140-2150.
18. Massire A, Vignaud A, Robert B, Le Bihan D, Boulant N, Amadon A. Parallel-transmission-enabled three-dimensional T2-weighted imaging of the human brain at 7 Tesla. *Magn Reson Med*. 2015;73:2195-2203.
19. Malik SJ, Padormo F, Price AN, Hajnal JV. Spatially resolved extended phase graphs: modeling and design of multipulse sequences with parallel transmission. *Magn Reson Med*. 2012;68:1481-1494.
20. Tse DHY, Wiggins CJ, Poser BA. High-resolution gradient-recalled echo imaging at 9.4T using 16-channel parallel transmit simultaneous multislice spokes excitations with slice-by-slice flip angle homogenization. *Magn Reson Med*. 2017;78:1050-1058.
21. Schmitter S, DelaBarre L, Wu XP, et al. Cardiac imaging at 7 Tesla: single- and two-spoke radiofrequency pulse design with 16-channel parallel excitation. *Magn Reson Med*. 2013;70:1210-1219.
22. Wu X, Schmitter S, Auerbach EJ, Ugurbil K, Van de Moortele PF. Mitigating transmit B 1 inhomogeneity in the liver at 7T using multi-spoke parallel transmit RF pulse design. *Quant Imaging Med Surg*. 2014;4:4-10.
23. Metzger GJ, Snyder C, Akgun C, Vaughan T, Ugurbil K, Van de Moortele PF. Local B-1(+) shimming for prostate imaging with transceiver arrays at 7T based on subject-dependent transmit phase measurements. *Magn Reson Med*. 2008;59:396-409.
24. Wu XP, Auerbach EJ, Vu AT, et al. Human Connectome project-style resting-state functional MRI at 7 Tesla using radiofrequency parallel transmission. *Neuroimage*. 2019;184:396-408.
25. Gras V, Poser BA, Wu X, Tomi-Tricot R, Boulant N. Optimizing BOLD sensitivity in the 7T human Connectome project resting-state fMRI protocol using plug-and-play parallel transmission. *Neuroimage*. 2019;195:1-10.
26. Wu XP, Auerbach EJ, Vu AT, et al. High-resolution whole-brain diffusion MRI at 7T using radiofrequency parallel transmission. *Magn Reson Med*. 2018;80:1857-1870.
27. Setsompop K, Wald LL, Alagappan V, Gagoski BA, Adalsteinsson E. Magnitude least squares optimization for parallel radio frequency excitation design demonstrated at 7 Tesla with eight channels. *Magn Reson Med*. 2008;59:908-915.
28. Xu D, King KF, Zhu Y, McKinnon GC, Liang ZP. Designing multichannel, multidimensional, arbitrary flip angle RF pulses using an optimal control approach. *Magn Reson Med*. 2008;59:547-560.
29. Gras V, Vignaud A, Amadon A, Le Bihan D, Boulant N. Universal pulses: a new concept for calibration-free parallel transmission. *Magn Reson Med*. 2017;77:635-643.
30. Gras V, Mauconduit F, Vignaud A, et al. Design of universal parallel-transmit refocusing kT -point pulses and application to 3D T2-weighted imaging at 7T. *Magn Reson Med*. 2018;80:53-65.
31. Aigner CS, Dietrich S, Schaeffter T, Schmitter S. Calibration-free pTx of the human heart at 7T via 3D universal pulses. *Magn Reson Med*. 2022;87:70-84.
32. Tomi-Tricot R, Gras V, Thirion B, et al. SmartPulse, a machine learning approach for calibration-free dynamic RF shimming: preliminary study in a clinical environment. *Magn Reson Med*. 2019;82:2016-2031.
33. Ianni JD, Cao Z, Grissom WA. Machine learning RF shimming: prediction by iteratively projected ridge regression. *Magn Reson Med*. 2018;80:1871-1881.
34. Glasser MF, Sotiropoulos SN, Wilson JA, et al. The minimal preprocessing pipelines for the human connectome project. *Neuroimage*. 2013;80:105-124.

35. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *Lect Notes Comput Sci*. 2015;9351:234-241.
36. Ye JC, Sung WK. Understanding geometry of encoder-decoder CNNs. *PMLR*. 2019;7064-7073.
37. Gong E, Zaharchuk G, Pauly J. Improving the PI+ CS reconstruction for highly undersampled multi-contrast MRI using local deep network. 2017. p 5663.
38. Lee D, Yoo J, Ye JC. Deep artifact learning for compressed sensing and parallel MRI. arXiv preprint arXiv:170301120 2017.
39. Zbontar J, Knoll F, Sriram A, et al. fastMRI: an open dataset and benchmarks for accelerated MRI. arXiv preprint arXiv:181108839 2018.
40. Tripathi PC, Bag S. CNN-DMRI: a convolutional neural network for denoising of magnetic resonance images. *Pattern Recogn Lett*. 2020;135:57-63.
41. Hu Z, Wang Y, Zhang Z, et al. Distortion correction of single-shot EPI enabled by deep-learning. *Neuroimage*. 2020;221:117170.
42. Innes M. Flux: elegant machine learning with Julia. *J Open Source Softw*. 2018;3:602.
43. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 2014.
44. Cawley GC, Talbot NLC. On over-fitting in model selection and subsequent selection bias in performance evaluation. *J Mach Learn Res*. 2010;11:2079-2107.
45. Goodfellow I, Bengio Y, Courville A. Deep learning. *Adapt Comput Mach Le*. 2016;1-775.
46. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process*. 2004;13:600-612.
47. Mazziotta J, Toga A, Evans A, et al. A probabilistic atlas and reference system for the human brain: international consortium for brain mapping (ICBM). *Philos Trans R Soc Lond B Biol Sci*. 2001;356:1293-1322.
48. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. *Neuroimage*. 2012;62:782-790.
49. Van Essen DC, Smith SM, Barch DM, et al. The WU-Minn human connectome project: an overview. *Neuroimage*. 2013;80:62-79.
50. Tian Q, Bilgic B, Fan Q, et al. DeepDTI: high-fidelity six-direction diffusion tensor imaging using deep learning. *Neuroimage*. 2020;219:117017.
51. Zhao H, Gallo O, Frosio I, Kautz J. Loss functions for image restoration with neural networks. *IEEE T Comput Imag*. 2017;3:47-57.
52. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks. *Commun AcM*. 2020;63:139-144.
53. Yan YX, Dahmani L, Ren JX, et al. Reconstructing lost BOLD signal in individual participants using deep machine learning. *Nat Commun*. 2020;11:5046.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

Figure S1 Transmit B_1 (B_1^+) inhomogeneity at the subject level when using the single transmission (sTx) setup. The commercial Nova sTx head coil was used in combination

with dielectric padding. Shown are magnitude B_1^+ maps in arbitrary units in 5 gapped representative axial slices, with the number reported being the coefficient of variation (i.e., SD/mean) of the B_1^+ maps across the whole brain. These B_1^+ maps were obtained using the actual flip angle imaging method and were co-registered to the MNI standard volume space. The brain mask resulting from the HCP structural preprocessing pipelines was used to mask the B_1^+ map for each subject. Note that strong B_1^+ inhomogeneity was present at 7 T using the sTx setup, with the coefficient of variation on average being as high as $\sim 27\%$.

Figure S2 Comparing single transmission (sTx) diffusion images against those predicted by our deep learning method (DL) in reference to the acquired parallel transmission (pTx) images. Shown are mean diffusion-weighted images with b1000 (averaged across all diffusion directions) of 1 representative coronal slice for each subject, in which case the model with tuned hyperparameters was trained on data of the other 4 subjects. Note that the use of DL substantially improved the image quality by effectively recovering the signal dropout observed in the lower temporal lobe (as marked by the yellow arrowheads), producing images that were more comparable to those obtained with pTx.

Figure S3 Comparing the point spread functions (PSFs) of single transmission (sTx) versus our deep learning (DL) method. Shown are mean (solid line) and range across 5 subjects (shaded area) of the center line of whole-brain PSFs in the anterior–posterior (AP) and (right–left) RL directions, for diffusion images (with b1000), fractional anisotropy (FA) and mean diffusivity (MD). In each case, the PSFs were estimated by comparing the sTx and DL results against those obtained with parallel transmission (serving as the reference). For both AP and RL directions, also reported are the mean and standard deviation (across subjects) of the FWHM of the PSFs in units of the number of pixels. Note that the use of our DL method brought little change to PSF in comparison to sTx, with the mean FWHM of the PSF (averaged over both RL and AP directions) found comparable to that of the sTx acquisition ($\sim 1\%$ difference).

Figure S4 Comparing the image edges of single transmission (sTx) versus our deep learning (DL) method. Shown in the left panel are diffusion images (b1k image), fractional anisotropy (FA), and mean diffusivity (MD), along with their respective edge images, all in a representative sagittal slice from a single subject; also shown for reference are the results obtained with parallel transmission. In each case, the edge images were estimated by applying a rotationally symmetric Laplacian of Gaussian filter (size, 15×15 pixels; SD, 1.5 pixels) to the corresponding images (with cerebrospinal fluid masked out to reduce the biases that it would otherwise induce in edge estimation because of its

extremely high or low signal intensity). Shown in the right panel are quantitative comparison showing region-specific mean and standard deviation values of the differences in edges between sTx and our DL method, along with the respective *P* values obtained from a paired *t* test (with significance being denoted by “*”). Note that although enhancing the edges of original diffusion weighted images, our DL method appeared to degrade the edges of both FA and MD maps, especially in the regions with strong transmit B_1 artifacts such as the lower temporal lobe (as indicated by red arrows), leading to a decrease in the region-specific edge values across the whole brain.

Figure S5 Testing the prediction performances of our deep learning (DL) model at different levels of signal dropout. A representative single-direction b1000 image slice obtained with single transmission (sTx) was chosen from 1 subject to show the typical signal dropout in the temporal pole. A region of interest (ROI) was manually drawn to define the region of signal dropout (red dashed contour). Different levels of signal dropout were simulated by decreasing the signal within the ROI by using various scaling factors ranging from 0 to 1 in steps of 0.2. Shown in the first 2 rows are the simulated sTx images with different levels of signal dropout (used as input to our DL model), and the predicted images (i.e., the output of our DL model), along with the images obtained using the parallel transmission (pTx) serving as the reference. The numbers reported are the normalized root mean squared errors within the ROI

in reference to pTx. Shown in the last 2 rows are the corresponding difference images between sTx and pTx and between DL and pTx at different levels of signal dropout. Note that our DL model appeared capable of recovering the signal to some extent even for complete signal dropout (i.e. no signal at all in the ROI in the input image) although the prediction performances were found to decrease with increasing levels of signal dropout.

Table S1 The hyperparameter space expanded by the 6 hyperparameters chosen in the current study for model selection. During the model selection, random search was carried out to tune the hyperparameters, for which a pool of 50 candidate hyperparameter sets was created by randomly sampling the space based on a uniform distribution.

Table S2 The percentage changes of quality measures between deep learning (DL) and single transmission (sTx) in each brain region. In each case, the percentage change was calculated as $|val_DL - val_sTx| / val_sTx * 100\%$ where *val_DL* and *val_sTx* are quality measures for DL and sTx, respectively; “| |” denotes the absolute value.

How to cite this article: Ma X, Uğurbil K, Wu X. Mitigating transmit- B_1 artifacts by predicting parallel transmission images with deep learning: A feasibility study using high-resolution whole-brain diffusion at 7 Tesla. *Magn Reson Med*. 2022;88:727-741. doi: 10.1002/mrm.29238