

Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging

Mehmet Akçakaya^{1,2} | Steen Moeller² | Sebastian Weingärtner^{1,2,3} | Kâmil Uğurbil²

¹Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, Minnesota

²Center for Magnetic Resonance Research, University of Minnesota, Minneapolis, Minnesota

³Computer Assisted Clinical Medicine, University Medical Center Mannheim, Heidelberg University, Mannheim, Germany

Correspondence

Mehmet Akçakaya, University of Minnesota, 200 Union Street S.E., Minneapolis, MN, 55455.
Email: akcakaya@umn.edu

Funding information

National Institutes of Health, Grant/Award Numbers: R00HL111410, P41EB015894, U01EB025144, R01NS085188, P30NS076408, and P50NS098573; National Science Foundation, Grant/Award Number: CAREER CCF-1651825.

Purpose: To develop an improved k-space reconstruction method using scan-specific deep learning that is trained on autocalibration signal (ACS) data.

Theory: Robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction trains convolutional neural networks on ACS data. This enables nonlinear estimation of missing k-space lines from acquired k-space data with improved noise resilience, as opposed to conventional linear k-space interpolation-based methods, such as GRAPPA, which are based on linear convolutional kernels.

Methods: The training algorithm is implemented using a mean square error loss function over the target points in the ACS region, using a gradient descent algorithm. The neural network contains 3 layers of convolutional operators, with 2 of these including nonlinear activation functions. The noise performance and reconstruction quality of the RAKI method was compared with GRAPPA in phantom, as well as in neurological and cardiac in vivo data sets.

Results: Phantom imaging shows that the proposed RAKI method outperforms GRAPPA at high (≥ 4) acceleration rates, both visually and quantitatively. Quantitative cardiac imaging shows improved noise resilience at high acceleration rates (rate 4:23% and rate 5:48%) over GRAPPA. The same trend of improved noise resilience is also observed in high-resolution brain imaging at high acceleration rates.

Conclusion: The RAKI method offers a training database-free deep learning approach for MRI reconstruction, with the potential to improve many existing reconstruction approaches, and is compatible with conventional data acquisition protocols.

KEYWORDS

accelerated imaging, convolutional neural networks, deep learning, image reconstruction, k-space interpolation, nonlinear estimation, parallel imaging

1 | INTRODUCTION

Long scan times remain a limiting factor in MRI, often necessitating trade-offs with spatial and temporal resolution or coverage. In most applications, some form of accelerated imaging is required. Most clinical scans rely on parallel imaging for acceleration,^{1–3} where the local sensitivities of phased array coil elements are used for reconstruction to enable undersampled data acquisitions. The two most common parallel imaging approaches are sensitivity encoding, (SENSE),² and generalized autocalibrating partially parallel acquisition (GRAPPA).³ SENSE is a reconstruction approach that is formulated in the image domain as a linear inverse problem with a priori knowledge of the receiver sensitivities. In the GRAPPA algorithm, the reconstruction is an interpolation in k -space using shift-invariant convolutional kernels to estimate missing k -space lines from acquired lines, which constitutes a linear reconstruction method. For GRAPPA, these convolutional kernels are determined from a small amount of fully sampled reference data, referred to as autocalibration signal (ACS). More recently, an alternative approach called iterative self-consistent parallel imaging reconstruction (SPIRiT),⁴ which iteratively performs local interpolation and enforces data consistency has also been proposed, facilitating the use of regularization in the reconstruction process.

An alternative line of work relies on non-uniform sampling of k -space.^{5–10} Other than non-Cartesian acquisitions,^{5–8} randomly undersampled Cartesian acquisitions have recently gained attention with the advent of compressed sensing.^{9,10} Although such undersampled acquisitions can be reconstructed using the linear conjugate gradient SENSE⁸ or SPIRiT method, they are more traditionally processed with nonlinear iterative algorithms that enforce some regularization based on the compressibility of images in a transform domain. Popular regularizations include l_1 norm of wavelet coefficients,¹⁰ total variation,^{9,11} low-rank schemes,^{12,13} and methods based on an adaptive generation of the sparsity transform domain.^{14–16} Regularized compressed sensing has been shown to have favorable noise properties in high-dimensional data sets,¹⁷ however, linear parallel imaging has been shown to outperform it in lower-dimensional, especially 2D acquisitions.¹⁸ Although the inherent denoising properties of CS-based techniques are desirable, the need for incoherent undersampled acquisitions, as well as its residual artifacts that are not readily accepted by clinical radiologists,¹⁸ have hindered its deployment. Therefore, even though noise amplification remains a challenge for the linear reconstruction algorithms used in parallel imaging, these approaches continue to be the prominent technique for accelerated imaging in clinical MRI.

Recently, machine learning based techniques have gathered interest as a possible means to improve reconstruction quality with different undersampling patterns. So far, the focus of the MRI reconstruction community has been on

generating more advanced regularizers by training on large amounts of data sets, along with both random and uniform undersampling patterns.^{19–23} Initial results indicate the promise of this approach in improving the reconstruction quality or providing a simple mapping from undersampled k -space to the desired image. Commonly, this training is performed in a manner that is not patient- or scan-specific, and therefore requires large databases of MR images.

In this study, we develop a k -space reconstruction method that uses deep learning on a small amount of scan-specific ACS data. We focus on uniformly undersampled acquisitions, to improve the quality of parallel imaging reconstructions and for integration with existing data acquisition approaches. However, extensions to other sampling patterns such as random and non-Cartesian are possible within the described k -space method. Unlike most recent deep learning approaches, the proposed scheme does not require any training database of images. Instead, the neural networks are trained for each specific scan, or set of scans, with a limited amount of ACS data, similar to GRAPPA. The proposed method, called robust artificial-neural-networks for k -space Interpolation (RAKI) reconstruction has an intuitive premise: instead of generating linear convolutional kernels from the ACS data as in k -space interpolation-based reconstruction methods such as GRAPPA, we propose to learn nonlinear convolutional neural networks from the ACS data. Using scan- and subject-specific deep learning, and learning all the necessary neural network parameters from the ACS data, any dependence on training databases or assumptions about image compressibility are avoided. After motivating our nonlinear reconstruction approach, a brief review of convolutional neural networks is provided. We then introduce our reconstruction and training methods. Phantom experiments and in vivo imaging are performed, comparing the proposed reconstruction to the GRAPPA parallel imaging approach.

2 | THEORY

2.1 | Nonlinear estimation of missing k -space lines

Reconstruction methods based on interpolation kernels in k -space aim to synthesize missing k -space lines as a linear combination of acquired lines across all coils.^{3,4,24,25} Specifically, for uniformly undersampled k -space data, GRAPPA uses linear convolutional kernels to estimate the missing data.³ Therefore, for the j th coil k -space data S_j , we have:

$$S_j(k_x, k_y - m\Delta k_y) = \sum_{c=1}^{n_c} \sum_{b_x=-B_x}^{B_x} \sum_{b_y=-B_y}^{B_y} g_{j,m}(b_x, b_y, c) S_c(k_x - b_x\Delta k_x, k_y - Rb_y\Delta k_y), \quad (1)$$

where R is the acceleration rate; $S_j(k_x, k_y - m\Delta k_y)$ are the unacquired k-space lines, with $m = 1, \dots, R-1$; $g_{j,m}(b_x, b_y, c)$ are the linear convolution kernels for estimating the data in the j^{th} coil specified by the location m as above; n_c is the number of coils; and B_x, B_y are specified by the kernel size. The convolutional kernels, $g_{j,m}(b_x, b_y, c)$, need to be estimated before the reconstruction. This is typically done by acquiring an ACS region, either integrated into the undersampled acquisition as central k-space lines or as a separate acquisition.³ Subsequently, a sliding window approach is used in the ACS region to identify the fully sampled acquisition locations specified by the kernel size and the corresponding missing entries. The former, taken across all coils, is used as rows of a calibration matrix, \mathbf{A} ; while the latter, for a specific coil, yields a single entry in the target vector, \mathbf{b} .²⁶ Therefore, for each coil j and missing location $m = 1, \dots, R-1$, a set of linear equations are formed, from which the vectorized kernel weights $g_{j,m}(b_x, b_y, c)$, denoted $\mathbf{g}_{j,m}$, are estimated via least squares, as $\mathbf{g}_{j,m} = \arg \min_{\mathbf{g}} \|\mathbf{b} - \mathbf{A}\mathbf{g}\|_2$.

Because the encoding process in a multi-coil MRI acquisition is linear, the reconstruction for sub-sampled data is also expected to be linear. This is the premise of linear parallel imaging methods, which aim to estimate the underlying reconstruction process that exploits redundancies in the multi-coil acquisition, using a linear system with a few degrees of freedom. These degrees of freedom are captured in the small convolutional kernel sizes for GRAPPA or the smooth coil sensitivity maps for SENSE. In essence, linear functions with such limited degrees of freedoms form a subset of all linear functions. In linear parallel imaging, the underlying reconstruction is approximated with a function from this restricted subset. Our hypothesis is that the underlying reconstruction function, although linear in nature, can be better approximated from a restricted subset of nonlinear functions with similarly few degrees of freedom. Examples of this phenomenon (i.e., using non-linear functions with few degrees of freedom to better approximate higher-dimensional linear functions) can be observed even in very low dimensions. For instance, assume a 2D hyperplane, which is the range space of a 2D linear mapping, is to be approximated. On average, using a 1D non-linear mapping, parametrized by 1 independent variable, such as an Archimedean spiral, will be a better approximation compared to a 1D linear mapping, which is a line through the origin, also parametrized by 1 independent variable. We hypothesize that similar approximations can be made in MRI reconstruction of multi-coil acquisitions. For instance, if the true underlying linear mapping is n -dimensional, an m -dimensional nonlinear mapping may outperform an m -dimensional linear mapping, where $m \ll n$. Therefore, we seek to estimate missing k-space lines from acquired ones using such a non-linear mapping.

We confine the modeling of this non-linear process to few degrees of freedom using convolutional neural networks (CNNs). Therefore, the degrees of freedom in the previous discussion refers to the CNN parameters. Note, similarly in the linear case for GRAPPA, the degrees of freedom refers to the coefficients of the linear convolutional kernels. We hypothesize that this restricted, yet non-linear, function space defined through CNNs enables to non-linearly learn redundancies among coils, without learning specific k-space characteristics, which may lead to overfitting. This is similar to the GRAPPA reconstruction, which successfully extracts such redundancies with a linear procedure, without adapting to k-space signal characteristics. The idea of using a larger class of nonlinear functions in modeling redundancies among coils for GRAPPA has also been explored previously,²⁶ although a fixed nonlinear kernel mapping was used instead of the deep learning strategies explored here. Therefore, the modeling of the multi-coil system with a non-linear approximation with few degrees of freedom, instead of a linear approximation with similarly few degrees of freedom, is hypothesized to extract comparable coil information without overfitting, while offering improved noise resilience.

Another reason for using non-linear estimation relates to the presence of noise in the calibration data. For the GRAPPA formulation above, there is noise in both the target data, \mathbf{b} , and the calibration matrix, \mathbf{A} . When both sources of noise are present, it has been shown that linear convolutional kernels incur a bias when estimated via least squares, and this bias leads to non-linear effects on the estimation of missing k-space lines from acquired ones.²⁶ In the presence of such data imperfections, which are in addition to the model mismatches related to the degrees of freedom described earlier, a non-linear approximation procedure may improve reconstruction quality and reduce noise amplification.

2.2 | Convolutional neural networks for MRI reconstruction in k-space

We propose to replace the linear estimation in GRAPPA that uses convolutional kernels with a non-linear estimation that uses CNNs.²⁷ The proposed method, RAKI, is designed to calibrate the CNN from ACS data without necessitating use of any external training database for learning.

CNNs are a special type of artificial neural network, whose use dates back to the late 1980 s.²⁸ They have recently gained interest because of the availability of efficient training implementations along with substantial gains in computing power,²⁹ use of fast-converging nonlinear layers,³⁰ and increased access to large training data sets.³¹ Therefore, CNNs have been successfully applied to multiple fields, including image classification,²⁹ face recognition,³² image super-resolution,³³ and image denoising.³⁴ CNNs are simple mathematical models that

estimate a (potentially nonlinear) function $f: X \rightarrow Y$, for two arbitrary sets X and Y . The fundamental idea is to combine multiple layers of simple modules, such as linear convolutions and certain point-wise nonlinear operators, to effectively represent complex functions of interest.²⁷

For the problem of nonlinearly estimating missing k-space data from acquired k-space data, a set of functions, $f_{j,m}$ are desired such that

$$S_j(k_x, k_y - m\Delta k_y) \approx f_{j,m} \left(\{S_c(k_x - b_x\Delta k_x, k_y - Rb_y\Delta k_y)\}_{b_x \in [-B_x, B_x], b_y \in [-B_y, B_y], c \in [1, n_c]} \right), \quad (2)$$

where $[a, b]$ denotes the set of integers from a to b inclusive. In RAKI, these functions $f_{j,m}$ are approximated using CNNs, and their parameters are learned from the ACS data. CNN literature and the associated deep learning procedures are commonly based on mappings over the real field.²⁷ Therefore, before any processing, we map all the complex k-space data to real-valued numbers. For a complex-valued k-space data set, \mathbf{s} of size $n_x \times n_y \times n_c$, where n_c is the number of coils, we embed this into a real-valued space as a data set of size $n_x \times n_y \times 2n_c$, where the real part of \mathbf{s} is concatenated with the imaginary part of \mathbf{s} along the third (channel) dimension. Therefore, effectively $2n_c$ input channels are processed. We note that other embeddings, which will increase dimensionality in the other dimensions, are possible, but were not explored because the ease of implementation of the channel dimension.

Before a CNN can be trained, its structure (i.e., the number of layers and layer operations) needs to be determined. There is

no optimization procedure for choosing the structure, and this is done in a heuristic manner.²⁷ In this study, we use a 3-layer structure, as depicted in Figure 1. Each layer, except the last, is a combination of linear convolutional kernels and a nonlinear operation called the rectified linear unit (ReLU), which has desirable convergence properties.³⁰ The linear convolution in CNNs, denoted by $*$, e.g., $\mathbf{z} = \mathbf{w} * \mathbf{s}$, and is defined as follows:

$$z(x, y, j) = \sum_{m=1}^{n_c} \sum_{p=1}^{b_x} \sum_{q=1}^{b_y} w(p, q, m, j) \cdot s(x+p-1, y+q-1, m) \text{ for } j = 1 \dots, n_{out}, \quad (3)$$

where \mathbf{s} is an input data of size $n_x \times n_y \times n_c$, \mathbf{w} is a kernel of size $b_x \times b_y \times n_c \times n_{out}$, and \mathbf{z} is the output data of size $(n_x - b_x + 1) \times (n_y - b_y + 1) \times n_{out}$. The ReLU is defined as

$$\text{ReLU}(x) = \max(x, 0). \quad (4)$$

The last layer performs the output estimation without any nonlinear operations. The details of the network are given next.

The first layer of our network takes in the sub-sampled zero-filled k-space of size $n_x \times n_y \times 2n_c$, which has been embedded into the real field. The convolutional filters in this layer, denoted by \mathbf{w}_1 , are of size $b_1^x \times b_1^y \times 2n_c \times n_1$. The operation of this layer is given as

$$F_1(\mathbf{s}) = \text{ReLU}(\mathbf{w}_1 * \mathbf{s}). \quad (5)$$

The second layer of our network takes in the output of the first layer and applies convolutional filters, denoted by

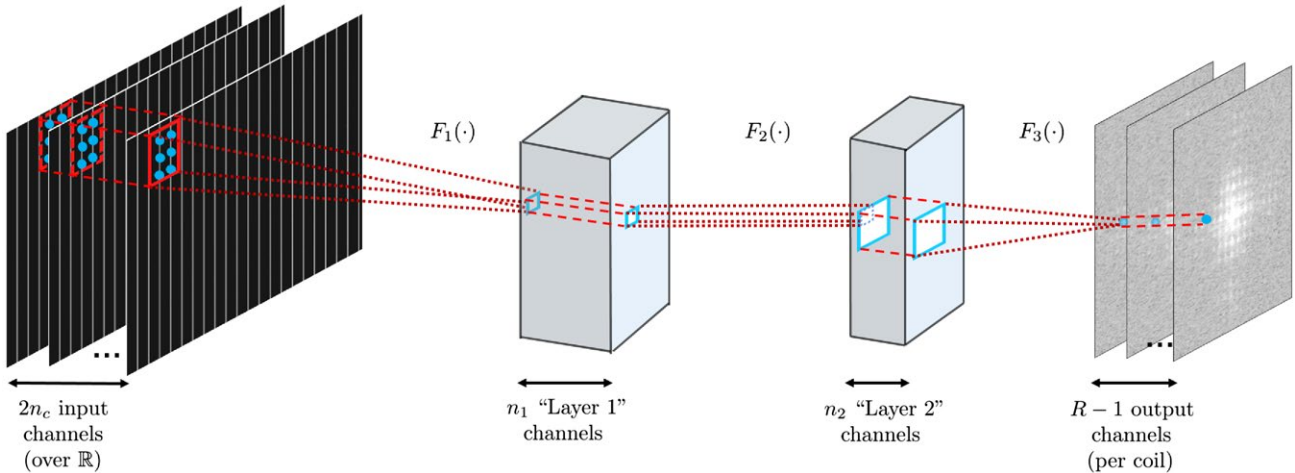


FIGURE 1 The 3-layer network structure used in this study for the convolutional neural network. The first layer, $F_1(\cdot)$, takes in the sub-sampled zero-filled k-space of size $n_x \times n_y \times 2n_c$, as embedded into the real field. The convolutional filters in this layer, \mathbf{w}_1 are of size $b_1^x \times b_1^y \times 2n_c \times n_1$. This is followed by a rectified linear unit (ReLU) operation. The second layer of our network, $F_2(\cdot)$, takes in the output of the first layer and applies convolutional filters, denoted by \mathbf{w}_2 , are of size $b_2^x \times b_2^y \times n_1 \times n_2$. This layer also includes a ReLU operation. These two layers nonlinearly combine the acquired k-space lines. The final layer of the network, $F_3(\cdot)$, produces the desired reconstruction output by applying convolutional filters, \mathbf{w}_3 , of size $b_3^x \times b_3^y \times n_2 \times n_{out}$. In our implementation all missing lines per coil (over reals) is estimated simultaneously, therefore $n_{out} = R - 1$

\mathbf{w}_2 , of size $b_2^x \times b_2^y \times n_1 \times n_2$. The operation of this layer is given as

$$F_2(\mathbf{s}) = \text{ReLU}(\mathbf{w}_2 * \mathbf{s}). \quad (6)$$

Intuitively, the first two layers are non-linearly combining the acquired k-space lines, essentially integrating the ideas of creating “virtual” channels and coil compression with non-linear processing. The final layer of the network produces the desired reconstruction output by applying convolutional filters, \mathbf{w}_3 , of size $b_3^x \times b_3^y \times n_2 \times n_{out}$, yielding

$$F_3(\mathbf{s}) = \mathbf{w}_3 * \mathbf{s}. \quad (7)$$

Therefore, the overall mapping is given by

$$F(\mathbf{s}) = F_3(F_2(F_1(\mathbf{s}))). \quad (8)$$

We note that a bias term is typically used in all layers in CNN applications. However, these were not included in our setup, as they change if the k-space is linearly scaled by a constant factor, for instance because of changes in the receiver gain, ultimately limiting robustness. In our implementations, all layers use kernel dilation³⁵ of size R in the k_y direction to only process the acquired k-space lines. We implement a CNN for each output channel and perform the estimation of all the missing k-space lines, within the specified kernel size, leading to $n_{out} = R - 1$. Because there are $2n_c$ output channels over the real field, this process yields a set, $\{F^j(\cdot)\}_{j=1}^{2n_c}$, of non-linear reconstruction functions, where each $F^j(\cdot)$ is an approximation of $\{f_{j,m}\}_m$ in Equation (2) using CNNs.

2.3 | Database-free learning of scan-specific neural networks

To learn a non-linear reconstruction function, $F^j(\cdot)$, the unknown network parameters, $\theta_j = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ for that input channel need to be estimated. Note we omit the subscript j for the kernels for ease of notation. Let \mathbf{s} denote the zero-filled acquired k-space lines, as the input to the network. The goal is to minimize a loss function between the reconstructed k-space lines using $F^j(\cdot)$, and the known ground truth for the target missing k-space lines in the ACS region, denoted by \mathbf{Y}_j , which is formatted to match the output structure with the same number of n_{out} channels in the third dimension. The loss function we use is the mean squared error (MSE), matching the original GRAPPA³ formulation:

$$L(\theta_j) = \|\mathbf{Y}_j - F^j(\mathbf{s}; \theta_j)\|_F^2, \quad (9)$$

where $\|\cdot\|_F$ denotes the Frobenius norm. The main difference with GRAPPA is that instead of solving a linear least squares problem to calculate one set of convolutional kernels; we need to solve a non-linear least squares problem to calculate three sets of convolutional kernels. Unlike the linear least squares problem, a closed-form minimization of the objective function is not possible. Therefore, in RAKI, gradient descent with backpropagation³⁶ and momentum³⁷ is used to minimize Equation (9). To update the convolutional kernels, at iteration t , we have

$$\bar{\mathbf{w}}_i^{(t)} = \mu \bar{\mathbf{w}}_i^{(t-1)} + \eta \frac{\partial L}{\partial \mathbf{w}_i^{(t-1)}}, \quad (10)$$

$$\mathbf{w}_i^{(t)} = \mathbf{w}_i^{(t-1)} - \bar{\mathbf{w}}_i^{(t)}, \quad (11)$$

where μ is the momentum rate, η is the learning rate, $i \in \{1, 2, 3\}$, and backpropagation is used to calculate the derivative $\partial L / \partial \mathbf{w}_i^{(t-1)}$. The same approach is used to calculate all convolutional filters. We note that although stochastic gradient descent is popular in most deep learning applications because of the immense size of the training data sets,³¹ we can use simple gradient descent because of the limited size of the ACS region.

3 | METHODS

3.1 | Reconstruction algorithm: Implementation details

The parameters used for the network are as follows: $b_1^x = 5$, $b_1^y = 2$, $n_1 = 32$, $b_2^x = 1$, $b_2^y = 1$, $n_2 = 8$, $b_3^x = 3$, $b_3^y = 2$, $n_o = R - 1$. Note that with this choice of parameters, at most $2R k_y$ lines are not estimated, which is the same as GRAPPA with a kernel size of $[5, 4]$.³ Three convolutional kernels $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ are trained for each output channel. The input is the zero-filled sub-sampled k-space, on which all the necessary Fourier transform shifts in k-space have been performed using phase correction. Kernel dilation³⁵ is used at each layer to process only the acquired data. We note that there is no effective dilation in the second layer because of our choice of kernel size. The momentum rate is set to $\mu = 0.9$. The learning rate is k-space scaling-dependent. We scale the k-space such that the maximum absolute value across all input channels is set to 0.015, and use $\eta = 100$ for the first layer and $\eta = 10$ for the next two layers. A slower learning rate for the later layers is important for convergence, as has been reported in image processing applications.^{33,34} The $n_{out} = R - 1$ outputs per processed k-space location is placed back to the corresponding missing k-space locations, consistent with standard GRAPPA practice.

The gradient descent for learning and the reconstruction was implemented in MATLAB (v9.0, The MathWorks, Natick, MA), with the CNN convolution and backpropagation operations implemented using the MatConvNet toolbox.³⁸ An implementation of the gradient descent, the CNN architecture, and the reconstruction algorithm, will be provided online (<https://people.ece.umn.edu/~akcakaya/RAKI.html>).

3.2 | Phantom imaging

Phantom imaging was performed at a 3 T Siemens Magnetom Prisma (Siemens Healthcare, Erlangen, Germany) system using a 32-channel receiver head coil-array. A head-shaped resolution phantom was imaged using a spoiled gradient recalled echo (GRE) sequence. The imaging parameters were: 2D multi-slice, FOV = $220 \times 220 \text{ mm}^2$, in-plane resolution = $0.7 \times 0.7 \text{ mm}^2$, matrix size = 320×320 , slice thickness = 4 mm, TR/TE = 500 ms/15 ms, flip angle = 70° , 27 slices, bandwidth = 360 Hz/pixel.

The fully sampled acquisition had an SNR of ~ 200 . To test the noise robustness of RAKI, complex Gaussian noise was added to the raw k-space to bring the image domain SNR measured within an ROI in the phantom down to 20. This noisy k-space was retrospectively undersampled using uniform undersampling of rates, $R \in \{1, 2, 3, 4, 5, 6\}$. ACS data of size 40 was taken from the central k-space lines, and this was used for the training of the CNN in RAKI reconstruction and for the determination of the convolutional kernels for GRAPPA reconstruction. GRAPPA reconstruction was implemented using a [5,4] kernel. For both GRAPPA and RAKI, no boundary extension was used. Therefore, edges of k-space that do not fully lie in the edges of the convolutional windows were not evaluated. For both methods, this meant the outermost $2(R-1) k_y$ lines were not calculated. The normalized mean square error (NMSE) of each reconstructed k-space was calculated with respect to the acquired k-space data. Additional experiments using higher SNR phantom data are provided in Supporting Information Figure S1.

3.3 | In vivo imaging

The imaging protocols were approved by the local institutional review board, and written informed consent was obtained from all participants before each examination for this HIPAA-compliant study.

Cardiac imaging was performed at a 3 T Siemens Magnetom Prisma (Siemens Healthcare, Erlangen, Germany) system using a 30-channel receiver body coil-array. SAPPHIRE myocardial T_1 mapping³⁹ was acquired in a mid-ventricular short-axis slice on 4 healthy volunteers (3 men, 40 ± 17 y). A balanced steady-state free precession image acquisition was used with the following parameters: FOV = 300

$\times 300 \text{ mm}^2$, TR/TE = 2.6 ms/1.0 ms, flip angle = 35° , bandwidth = 1085 Hz/pixel. 11 images with different T_1 weights were acquired in a single breath-hold. One acquisition was performed with resolution = $1.7 \times 1.7 \text{ mm}^2$, matrix size = 176×176 , slice thickness = 8 mm, in-plane acceleration (iPAT) = 2, ACS lines = 32. Another high-resolution acquisition was performed with resolution = $1.1 \times 1.1 \text{ mm}^2$, matrix size = 272×272 , slice thickness = 6 mm, iPAT = 5. For this high-resolution acquisition, a separate calibration scan with resolution = $1.1 \times 4.7 \text{ mm}^2$, corresponding to 64 ACS k_y lines, were performed in the same breath-hold with a 3 seconds delay to allow for T_1 recovery.

These data sets were used to test the robustness of RAKI across different contrasts. Specifically, the CNNs in RAKI and the convolutional kernels in GRAPPA were estimated using ACS data from only one T_1 -weighted image. These CNNs and convolutional kernels were then used to reconstruct all 11 images with varying T_1 weights. Therefore, this tested whether RAKI captured the coil information, as in GRAPPA, or performed over-fitting to specific image or k-space features. Three experiments were performed with these data sets: (1) the standard resolution acquisitions were reconstructed with RAKI and GRAPPA, where the CNN parameters and convolutional kernels were estimated using the integrated ACS region of the image with no T_1 preparation; (2) the standard resolution acquisitions were retrospectively undersampled to $R = 4$, and RAKI and GRAPPA reconstructions were performed on the same datasets and using the same ACS region of the image with no T_1 preparation; and (3) the high resolution acquisitions were reconstructed with RAKI and GRAPPA, where the training and convolutional kernels were estimated using the ACS data from the additional low-resolution calibration scan. Non-linear reconstructions change the statistical distribution of the noise and introduce dependencies between the reconstruction noise and reconstructed signal beyond additivity, hindering the definition of the equivalent of a g-factor.^{40,41} For these quantitative MRI data sets, the spatial variability of the T_1 times in the homogeneous muscle tissue of the myocardium were used as a surrogate for noise performance⁴² for both reconstruction methods.

Brain imaging was performed at a 7 T Siemens Magnex Scientific (Siemens Healthcare, Erlangen, Germany) system using a 32-channel receiver head coil-array. Anatomical T_1 -weighted imaging was acquired in a patient (man, 76 y) and in a healthy volunteer (man, 43 y) using a standard Siemens 3D-MPRAGE sequence with the following parameters: FOV = $230 \times 230 \times 154 \text{ mm}^3$, resolution = $0.6 \times 0.6 \times 0.6 \text{ mm}^3$, TR/TE = 3100 ms/3.5 ms, inversion time = 1500 ms, flip angle = 6° , bandwidth = 140 Hz/pixel, ACS lines = 40. The k-space data was inverse Fourier transformed along the slice direction, and a central slice, matrix size = 384×384 , was processed for the rest of the study. A single slice of a 3D data set

was chosen over 2D imaging to enable a 0.6 mm isotropic acquisition, while overcoming inherent SNR limitations of the latter. The patient data was acquired with $iPAT = 3$ and was also retrospectively undersampled to $R = 6$. RAKI and GRAPPA reconstructions were performed on both these data sets. The healthy volunteer data was acquired with $iPAT = 3, 4, 5, 6$. To further help the visualization of artifacts and avoid SNR losses, $iPAT = 5$ and 6 data were acquired with 2 averages. Therefore, the only SNR penalty between $iPAT = 3$ and 6 acquisitions is because of differences in coil encoding. All data sets were reconstructed using RAKI and GRAPPA. These data sets were used to test RAKI at very high resolutions for robustness and potential blurring artifacts. Additional experiments on the effect of calibration region on the high-SNR high-resolution volunteer data are provided in Supporting Information Figure S2.

Additional brain imaging was performed at a 3 T Siemens Magnetom Prisma (Siemens Healthcare, Erlangen, Germany) system using a 32-channel receiver head coil-array. T_1 -weighted imaging was acquired in a healthy subject (man, 41 y) using a standard Siemens 3D-MPRAGE sequence with the following parameters: $FOV = 224 \times 224 \times 179 \text{ mm}^3$, resolution = $0.7 \times 0.7 \times 0.7 \text{ mm}^3$, $TR/TE = 2400 \text{ ms}/2.2 \text{ ms}$, inversion time = 1000 ms, flip angle = 8° , bandwidth = 210 Hz/pixel, ACS lines = 40. Two separate acquisitions were performed with $iPAT = 2$ and 5. The k-space data was inverse Fourier transformed along the slice direction, and a central slice, matrix size = 320×320 , was processed for the rest of the study. The $R = 2$ acquisition was retrospectively undersampled to $R = 4$ and 6. Reconstructions were performed using both RAKI and

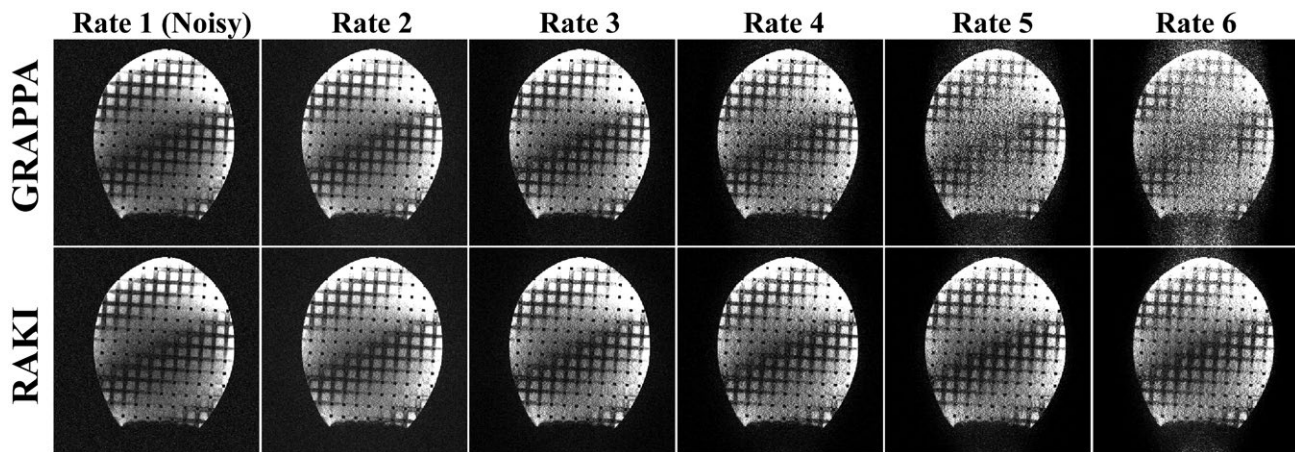


FIGURE 2 Results of phantom experiments. Images reconstructed using GRAPPA (top row) and the proposed RAKI (bottom row) reconstructions from noisy data sets for various acceleration rates, as well as a fully sampled noisy reference image. Starting with rate 4, the noise performance of RAKI improves on that of GRAPPA. RAKI reconstructions at acceleration rates 4 and 5 show visually improved noise properties over GRAPPA, which has high levels of noise amplification

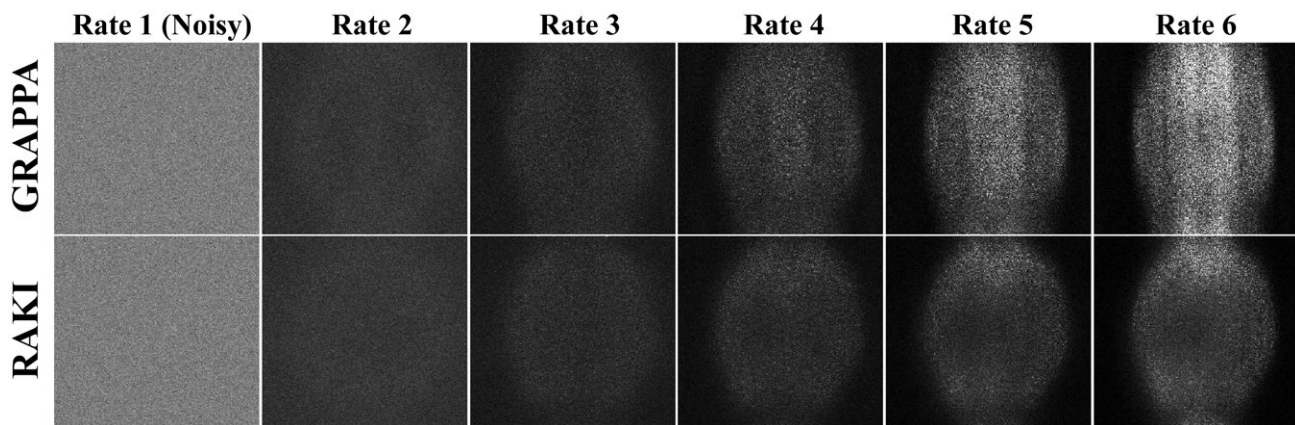


FIGURE 3 The difference images for the reconstructions in Figure 2, as compared to the original acquisition without additional noise, confirm the observations regarding the improved noise resilience of RAKI. NMSE quantification is also consistent with these observations: 0.65, 0.51, 0.51, 0.66, and 0.89 for GRAPPA versus 0.65, 0.51, 0.46, 0.47, and 0.52 for RAKI from rates 2 to 6

GRAPPA for $R = 2, 4, 5, 6$ to test the robustness of RAKI at very high resolutions for potential blurring artifacts.

4 | RESULTS

4.1 | Phantom imaging

Figure 2 depicts the results from phantom imaging. GRAPPA and the proposed RAKI reconstructions from noisy data for various acceleration rates. Although virtually no difference is observed at acceleration rates of 2 and 3, starting with rate 4, the noise performance of RAKI improves on that of GRAPPA. RAKI reconstructions at acceleration rates 4 and 5 show visually improved noise properties over GRAPPA, which has high levels of noise amplification, especially in the center part of the phantom. The difference images depicted in Figure 3 are also in agreement with these observations, where the errors in GRAPPA reconstruction become apparent and structured from rate 4 onward. The noise amplification behavior of the 2 reconstructions is further reflected with the quantitative NMSE performance. GRAPPA reconstruction exhibits NMSE 0.65, 0.51, 0.51, 0.66, and 0.89 from rates 2 to 6, respectively, whereas RAKI reconstruction has NMSE 0.65, 0.51, 0.46, 0.47, and 0.52 for the same rates. This corresponds to 0%, 0%, 11%, 28%, and 41% decrease in NMSE, respectively, using RAKI as compared to GRAPPA. Note the NMSE values reported here are not monotonically increasing because of several reasons, including (1) the averaging nature of NMSE in capturing the non-uniform noise characteristics, (2) noise-suppression outside the object in the reconstructions, (3) a noiseless reference in the NMSE calculation, and (4) the tendency of k-space interpolation techniques to exhibit g-factors < 1 in certain areas.

4.2 | In vivo imaging

Figure 4 shows representative T_1 -weighted images from SAPHIRE T_1 mapping acquired at the conventional resolution of $1.7 \times 1.7 \text{ mm}^2$. The images corresponding to the shortest inversion time, with a low SNR, are shown in Figure 2A for the acquisition acceleration rate of 2, as well as a retrospective acceleration rate of 4. There is visually no discernible difference between GRAPPA and RAKI at rate 2. However, at rate 4, the improved noise performance of RAKI over GRAPPA becomes visually apparent. Figure 2B shows T_1 -weighted images acquired without any preparation, corresponding to the highest SNR images in the acquisition. At this higher SNR, the differences between RAKI and GRAPPA are visually minor even at the acceleration rate of 4, with a slight improvement in the left ventricle blood pool for the proposed approach. These observations are reflected in the quantitative T_1 measurements, with virtually no difference

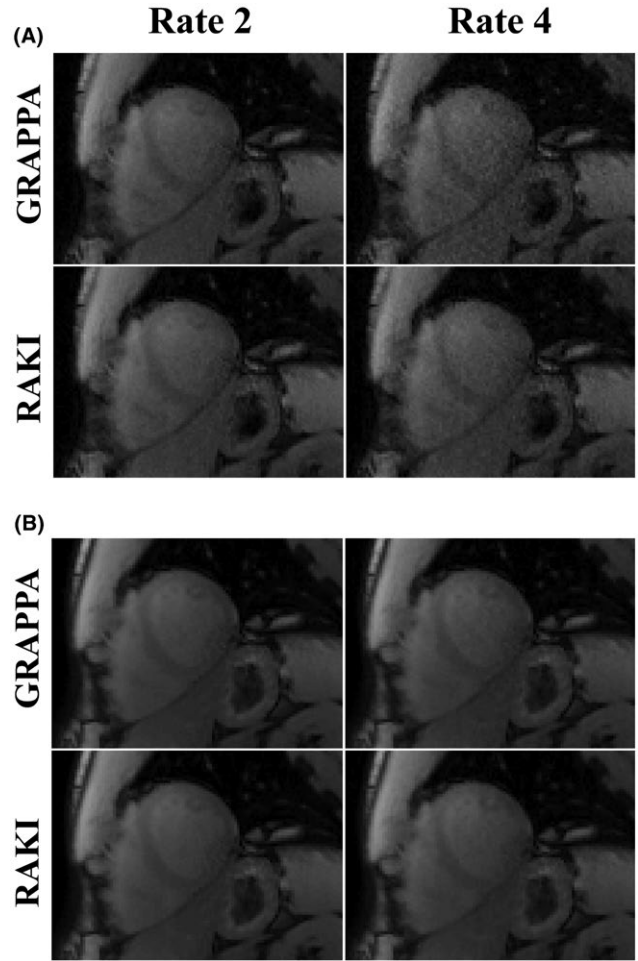


FIGURE 4 Representative T_1 -weighted images from SAPHIRE T_1 mapping acquired at the conventional resolution of $1.7 \times 1.7 \text{ mm}^2$. (A) Images with low acquisition SNR, corresponding to the shortest inversion time, at the acquisition acceleration rate of 2, and a retrospective acceleration rate of 4. At rate 4, the noise performance improvement of RAKI over GRAPPA becomes apparent. (B) Images with high acquisition SNR, where the differences between RAKI and GRAPPA are visually minor even at the acceleration rate of 4, with a slight improvement in the left ventricle blood pool for RAKI. A quantitative analysis of the resulting T_1 maps indicate that comparable quantification is observed for the 2 reconstructions, except for a 10% reduced noise variability with RAKI at rate 4

($<0.5\%$) in mean myocardial T_1 value among the 2 methods at either rate, or $<1\%$ difference in spatial variability at rate 2. However, RAKI showed an improvement of 10% in spatial variability over GRAPPA at rate 4. These trends are consistent across all subjects, where comparable quantification is observed with RAKI and GRAPPA reconstructions, with the exception of a 23% reduced noise variability with RAKI at rate 4 (average T_1 time \pm average spatial variability: RAKI: $1562 \pm 197.0 \text{ ms}$ vs. GRAPPA: $1542 \pm 258.8 \text{ ms}$). These experiments show the improved noise resilience of RAKI, as well as its robustness in reconstructing varying contrast weightings.

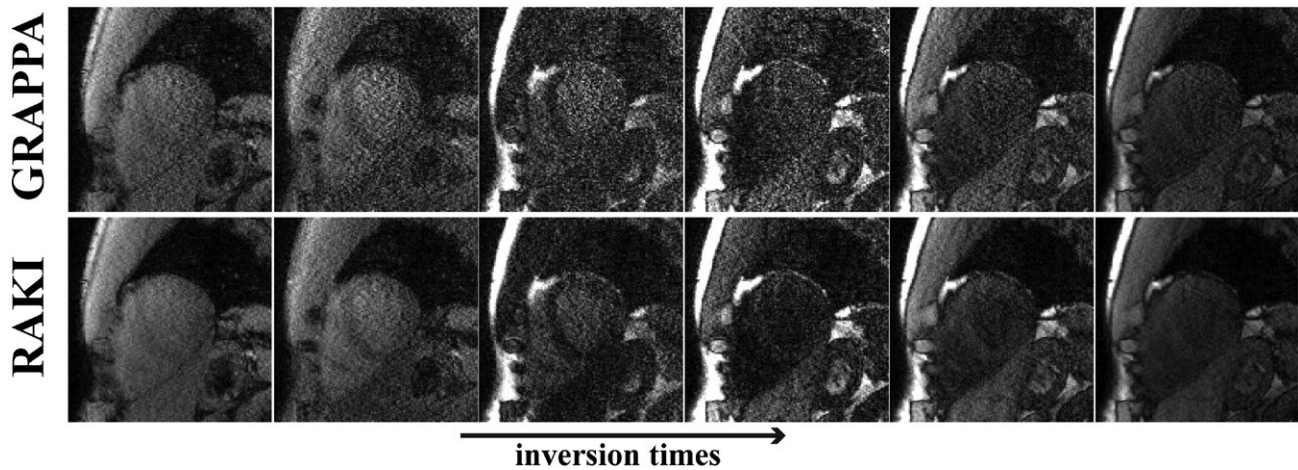


FIGURE 5 Six T_1 -weighted images, corresponding to the shortest inversion times, acquired using the SAPHIRE sequence at a high spatial resolution of $1.1 \times 1.1 \text{ mm}^2$, and acceleration rate 5. At this high acceleration rate, the noise improvement of RAKI over GRAPPA is observable for all images. A quantitative analysis of the resulting T_1 maps indicate that the spatial variability of the myocardial T_1 is improved by 37% using RAKI versus GRAPPA

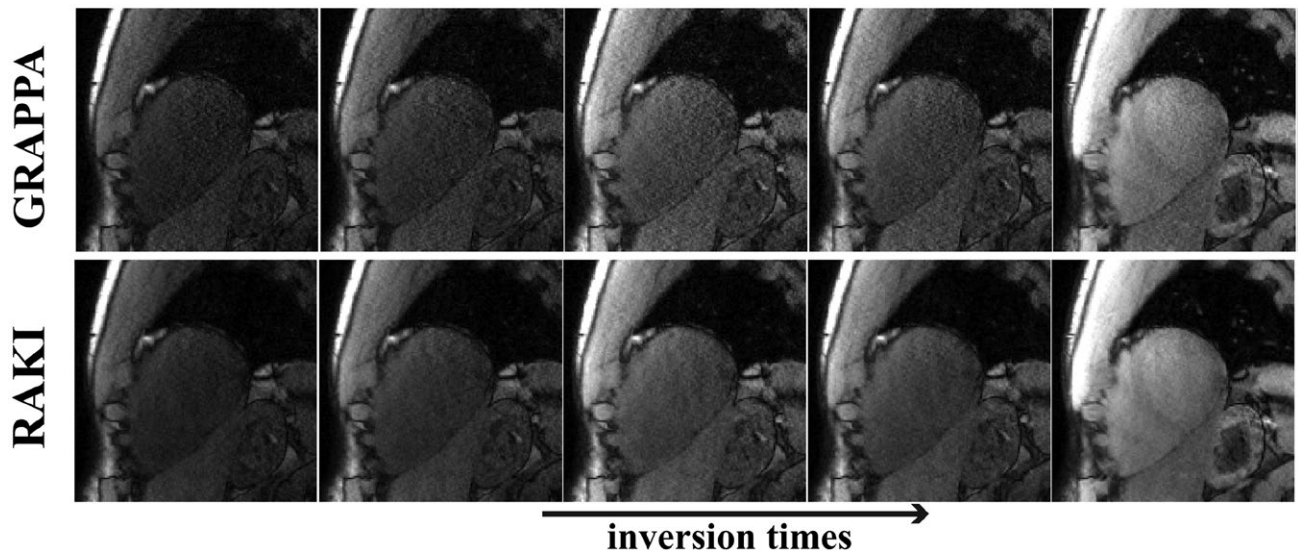


FIGURE 6 Five T_1 -weighted images, corresponding to the longest inversion times, from the same data set depicted in Figure 5. At this high acceleration rate, the noise improvement of RAKI is observable for all images. Quantitative analysis of T_1 map spatial variability shows 37% improvement using RAKI versus GRAPPA

Figure 5 and 6 show all of the 11 T_1 -weighted images acquired using the SAPHIRE sequence at the improved spatial resolution of $1.1 \times 1.1 \text{ mm}^2$ and acceleration rate 5. In this substantially lower SNR regime, because of the higher acceleration rate and improved resolution, the noise improvement of RAKI over GRAPPA is visually apparent for all images. A quantitative analysis of the resulting T_1 maps indicate that the spatial variability of the myocardial T_1 is 253.8 ms using the RAKI-reconstructed images versus 403.5 ms using the GRAPPA-reconstructed images, indicating a 37% improvement. The improvement is observable across all subjects, with an overall 48% improvement in spatial variability with RAKI. We note that in 2 cases, GRAPPA reconstruction do

not provide image quality that allows for reasonable quantitative evaluation, as residual artifacts and excessive noise amplification prevent reliable delineation of the myocardium against the blood-pools. For the remaining 2 cases, the improvements in spatial variability are 37% and 31% for RAKI.

Figure 7 depicts a slice of the high-resolution 7 T MPAGE acquisition on the patient, both at the acquisition acceleration rate of 3, as well as at a retrospective acceleration rate of 6. At rate 3, RAKI and GRAPPA methods both successfully reconstruct the image with little residual artifacts. At rate 6, GRAPPA reconstruction suffers from pronounced noise amplification, whereas the proposed RAKI reconstruction exhibits better noise tolerance. However, some residual

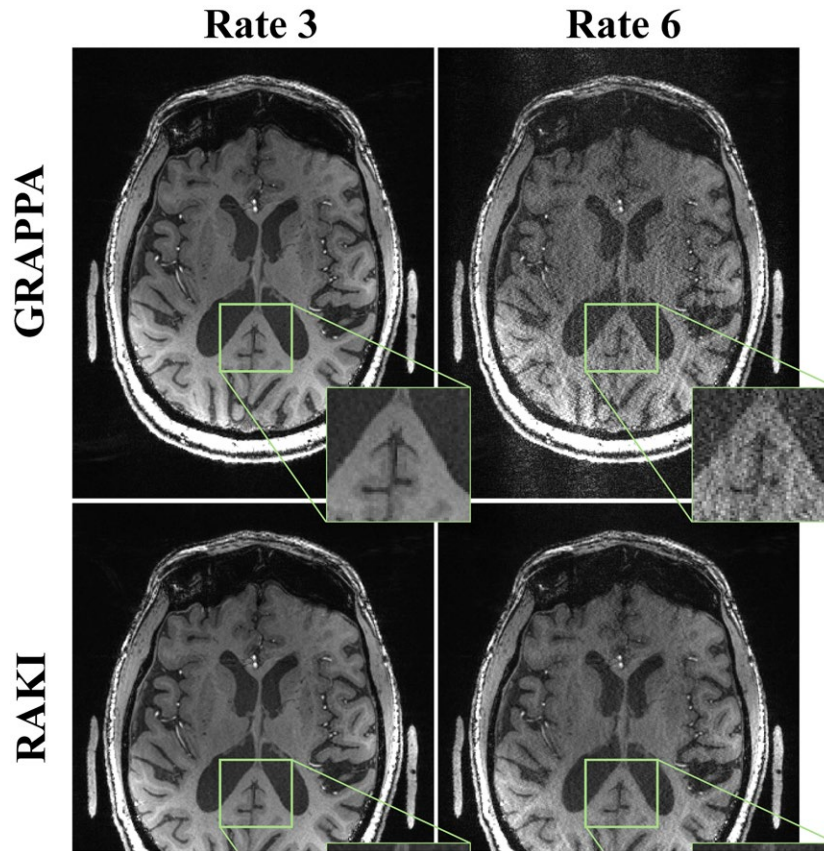


FIGURE 7 A central slice of the high-resolution (0.6 mm isotropic) 7 T MPRAGE acquisition, both at the acquisition acceleration rate of 3, as well as at a retrospective acceleration rate of 6. At rate 3, GRAPPA (top) and RAKI (bottom) methods both successfully reconstruct the image with little residual artifacts. At rate 6, GRAPPA reconstruction suffers heavily from noise amplification, whereas the proposed RAKI reconstruction exhibits improved noise tolerance without blurring artifacts

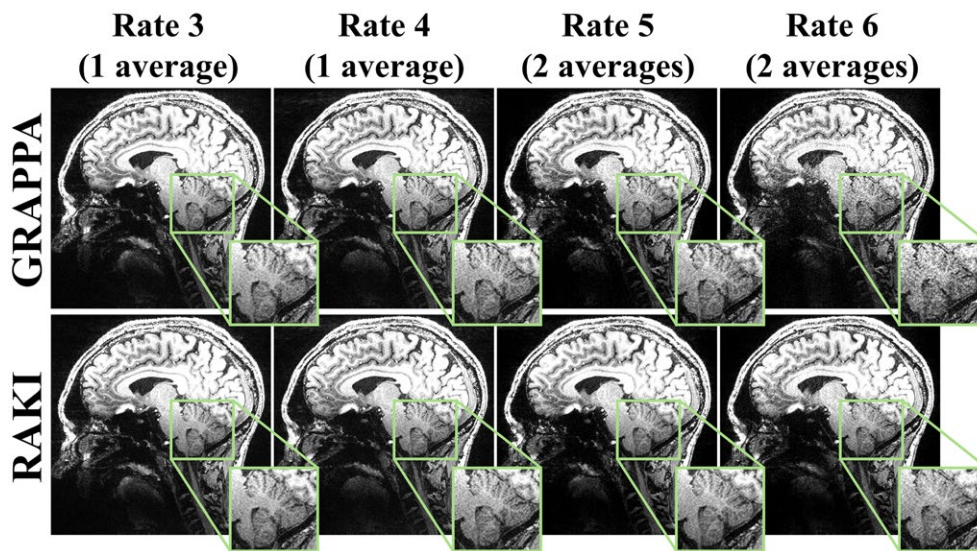


FIGURE 8 A central slice of the high-resolution (0.6 mm isotropic) MPRAGE acquisition at 7 T, acquired at $R = 3, 4, 5$, and 6 , where the first 2 rates were acquired with 1 average, and the latter 2 were acquired with 2 averages. Therefore, the only SNR penalty between $R = 3$ and 6 is because of differences in coil encoding. The results show that at this high SNR, GRAPPA (top) and RAKI (bottom) reconstruct the images with little residual artifacts up to $R = 4$. At $R = 5$, slight differences between GRAPPA and RAKI can be observed, with the latter showing better noise performance. At $R = 6$, the difference is further pronounced. At the higher rates, RAKI has improved noise tolerance while exhibiting no blurring artifacts

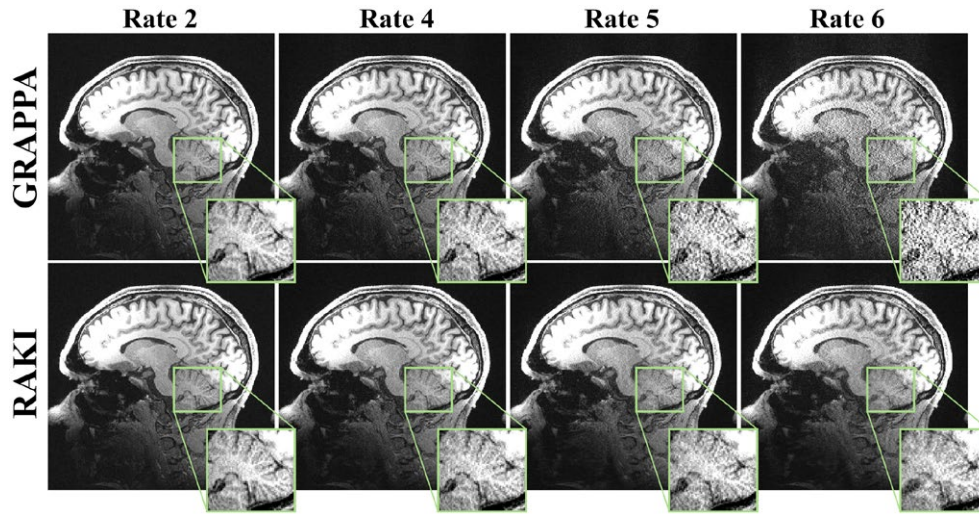


FIGURE 9 A central slice of the high-resolution (0.7 mm isotropic) MPRAGE acquisition at 3 T, acquired at $R = 2$ and 5, as well as retrospective $R = 4$ and 6. At rates 2 and 4, GRAPPA (top) and RAKI (bottom) methods both successfully reconstruct the image with little residual artifacts. At rates 5 and 6, noise amplification becomes visible for GRAPPA reconstruction. At these rates, RAKI has improved noise tolerance and exhibits no blurring artifacts

aliasing artifacts remain because of the high-level of in-plane acceleration. No blurring artifacts are observed with either method, even though the acquisition is performed at a very high resolution.

Figure 8 depicts a slice of high-SNR 0.6 mm isotropic 7 T MPRAGE acquisition, at acceleration rates 3, 4, 5, and 6, where the latter 2 acceleration rates were acquired with 2 averages. At rates 3 and 4, both RAKI and GRAPPA successfully reconstructs the images with little artifacts. At $R = 5$, GRAPPA and RAKI starts exhibiting slight differences, with the latter showing better noise performance. The differences become more pronounced at $R = 6$, while the quality also degrades for both reconstructions. At the higher rates, RAKI has improved noise amplification with no apparent blurring.

Figure 9 depicts a slice of the 0.7 mm isotropic 3 T MPRAGE acquisition, at the acquisition acceleration rates of 2 and 6, as well as at a retrospective acceleration rates of 4 and 6. At rates 2 and 4, RAKI and GRAPPA methods both successfully reconstruct the image with little residual artifacts. At rate 5, RAKI visually outperforms GRAPPA in terms of noise amplification, while incurring no visible reconstruction artifacts or blurring at this high-resolution. $R = 6$ reconstructions also show RAKI has visible noise advantages over GRAPPA, although at this high acceleration rate, both reconstructions have visible residual artifacts.

5 | DISCUSSION

In this study, we developed a k-space method, called RAKI, enabling deep learning for reconstruction of undersampled data from ACS data, as an alternative to linear k-space

interpolation techniques. This enables us to generate a non-linear function to estimate the missing k-space data from the acquired k-space lines without additional training on databases containing many subjects. In effect, our approach creates subject-specific neural networks for a non-linear reconstruction, which extends on the subject-specific linear convolution kernels used successfully in many parallel imaging reconstruction techniques, including GRAPPA. Our results indicate that RAKI offers improved noise performance compared to GRAPPA in phantom and in vivo imaging.

The use of machine learning for MRI reconstruction has recently received increasing interest. These recent works aim to build on the immense success of deep learning in image processing applications and have focused on training reconstruction algorithms on large amounts of data sets. Big data training has led to different regularization approaches for MR images, most often as regularizers without an explicit formula and based on CNNs,^{19,22,23,43} or as unsupervised learning of parameters for variational networks.²¹ In this work, we took an MRI-focused approach and considered the problem as a k-space estimation problem instead of an image domain regularization problem. This approach avoids dependency on large amounts of training data, and instead uses a small amount of ACS data, similar to existing parallel imaging reconstructions that are used in clinical practice. We also note that almost a decade ago, the use of neural networks was proposed for improving SENSE reconstruction in a scan-specific manner.⁴⁴ However, this earlier work did not use CNNs and was not further explored. Furthermore, the work in Sinha et al⁴⁴ uses an image domain approach based on SENSE formulation. Because of the use of non-linear reconstructions, this is not necessarily equivalent to our

k-space formulation, so its relative efficacy to our approach is hard to compare.

The nonlinearity in our convolutional network architecture improves the noise performance of the reconstructions, as presented in our experiments. This nonlinearity was introduced by the use of ReLU functions in all layers, except the last one. Other nonlinear activation functions, such as sigmoid functions,²⁷ have also been explored in neural network community and can also be applied as part of our method. However, ReLU has several advantages. Because it has a gradient that is either 0 or 1—unlike the sigmoid function, which can take arbitrarily small non-zero derivatives—its derivative can be back-propagated across multiple layers without convergence issues, enabling deeper network architectures. The nature of the ReLU function also enables sparsity in representations.⁴⁵ With a random initialization of weights, half of the output values are zeroes, facilitating an inherent data sparsification.

The use of nonlinear methods to improve GRAPPA reconstruction has been explored before^{26,46} but without the use of concepts from neural networks. A fixed set of virtual channels for a higher-dimensional feature space²⁶ or transform-domain sparsity⁴⁶ has also been proposed before. RAKI differs from these techniques, by refraining from any explicit assumptions on k-space or transform domains. Instead neural networks were shown to effectively represent a nonlinear function for estimating missing k-space data, successfully facilitating model-free reconstruction.

The advantage of our method over a linear reconstruction approach like GRAPPA becomes apparent at high acceleration rates or at low SNRs. In the other regimes, GRAPPA is known to perform very well, therefore differences are not noticeable. At the high acceleration rate regime, both RAKI and GRAPPA approaches require larger ACS regions for better estimation of the reconstruction filters. Because of the multi-layer network architecture, RAKI has more unknowns. In most applications, where one is interested in these higher rates (e.g., diffusion, perfusion, or quantitative MRI), it is not difficult to obtain one set of high-quality ACS data of sufficient size, which can be shared across multiple scans. The success of training a set of CNNs from one set of ACS data for varying contrast acquisitions was shown in our cardiac imaging experiments. Nonetheless, there may be applications where ACS region cannot be increased easily. In these cases, several modifications to our method are possible. One approach would be to reduce the number of layers. Another would be to reduce the output sizes for each layer. An alternative approach would be to perform a hybrid reconstruction, by first performing linear GRAPPA reconstruction in the central region to grow the ACS region and then use this larger area for training the neural network in RAKI reconstruction. All

these approaches warrant further investigation in specific applications, but these were not pursued in this study.

Unfortunately there are no existing methods for optimally determining the network structure in deep learning applications.²⁷ Therefore, several parameters for the network architecture were determined heuristically. These include the number of layers, the kernel sizes for each layer, the output sizes for the first and second layer, as well as the activation functions. The kernel sizes were chosen to not increase the number of unestimated k_y boundary lines beyond a [5,4] GRAPPA kernel, while enabling contributions from neighboring locations in the first and last layers. The number of layers and output sizes of each layer were chosen to limit the number of unknowns in the CNN, in order not to increase the required ACS data for training. The output size for the last layer was also a design choice, and our networks were trained to output all missing lines per coil. Other alternatives are possible. For instance, a different network can be trained per coil per each missing k-space locations relative to the acquisition location. However, this increased the number of unknowns by roughly a factor of $R - 1$ and provided virtually no benefit in preliminary observations. An alternative approach would be to train the network to output all the missing lines for all coils. This would require deeper networks and bigger output sizes per layer. We did not explore this alternative because we wanted to avoid interactions across coils in the output layer to provide a fair comparison to GRAPPA. However, alternative network architectures with such interactions may be beneficial and will be explored in future work.

Another design decision for the CNNs used in this work was the exclusion of bias terms for all three layers. In standard deep learning practice, both a set of convolutional filters and a set of biases are trained at each layer. For instance, in the first layer, this would correspond to convolutional filters \mathbf{w}_1 of size $b_1^x \times b_1^y \times 2n_c \times n_l$, and a bias term \mathbf{b}_1 of size $1 \times 1 \times n_l$, and the output of this layer would be calculated as $\text{ReLU}(\mathbf{w}_1 * \mathbf{x} + \mathbf{b}_1)$. However, because of their additive nature, inclusion of biases increases dependence on how the l_2 (or l_∞) norm of the k-space is scaled. This creates major limitations in processing data with multiple-varying contrast, as in quantitative MR parameter mapping, when training is done on one set of ACS data with a specific contrast weighting, which was the case in our experiments. Therefore, biases were not used in our CNNs, although they may find use in other settings.

The application of the CNN to missing lines is not significantly more time-consuming than the application of GRAPPA kernels. For the CNN architecture and GRAPPA kernels used in this manuscript, for acquired $n_x \times n_y / R$ k-space points, the complexity of the CNN approach per k-space point is $(1280n_c^2 + 512n_c + 96(R - 1)n_c)$ real additions and multiplications, and 40 logical comparisons. For GRAPPA

this is approximately $20(R-1)n_c^2$ complex additions and multiplications per acquired k-space point. Therefore, in terms of the number of real multiplications, for $n_c = 32$, RAKI has a ~ 7 -fold overhead at $R = 4$ and a ~ 4 -fold overhead at $R = 6$, compared with GRAPPA. However, the calibration of the GRAPPA kernel can be done with a simple linear least squares problem, whereas the CNN needs to be trained with a gradient descent. Our MATLAB implementation for the gradient descent has an approximate run-time of 5–6 s for 250 iterations, which can likely be improved using optimized implementations or parallelization with graphics processing units (GPU).

There are limitations to the use of the current proposed approach. As with other parallel imaging approaches, the coil geometry in the acquisition has to be able to support the desired acceleration rates. Aliasing artifacts may persist if the rate is increased further (see Supporting Information Figure S1). As discussed earlier, the CNN architecture was heuristically selected. The performance of the approach may change based on different network parameters, similar to how the performance of linear methods is affected by kernel sizes. The optimization problem in the deep learning portion of the approach is based on minimizing a nonlinear least squares objective function depicted in Equation (9). Unlike the linear least squares problem in GRAPPA, this objective function is non-convex. Therefore, there is a potential of converging to local minima. This is typically handled using momentum in gradient descent approaches, which has been shown to work well in practice.³⁷ Nonetheless, the solution to Equation (9) is not exact, but approximate. A further limitation in our current implementation is the use of fixed learning rates in the gradient descent algorithm. Although these rates were optimized heuristically for in vivo data sets, they may not be optimal in all applications (see Supporting Information Figure S1 for phantom data). To overcome this limitation, further work on implementations and optimizers that can adaptively update the learning rate is warranted and is one of our current topics of research.

Because our proposed method works with uniformly undersampled acquisitions, it has the potential to be applied to existing highly accelerated and low-SNR data sets to improve reconstruction quality, such as those generated by Human Connectome projects.⁴⁷ It could also easily be used in most protocols, as it does not require modifications to data acquisition.

Our RAKI reconstruction naturally extends to other approaches that utilize k-space interpolation. Although we concentrated on GRAPPA and uniform undersampling for 2D acquisitions in this study because of their high use, RAKI reconstruction is applicable to both 2D and 3D acquisitions with uniform undersampling,^{3,48,49} as well as to simultaneous multi-slice imaging.^{24,50} It is also applicable to random undersampling patterns as an alternative to SPIRiT convolution kernels,⁴ and can be combined with image regularization in

this setting.¹⁰ It can also be extended to non-Cartesian acquisitions, as an alternative to linear convolutional approaches, such as radial²⁵ or spiral⁴ acquisitions. These extensions are beyond the scope of the current study and will be explored in future work.

6 | CONCLUSION

The proposed reconstruction uses a scan-specific deep learning approach with convolutional neural networks trained on limited ACS data to improve on the reconstruction quality of linear k-space interpolation-based methods.

REFERENCES

1. Meyer CH, Hu BS, Nishimura DG, Macovski A. Fast spiral coronary artery imaging. *Magn Reson Med*. 1992;28:202–213.
2. Lustig M, Pauly JM. SPIRiT: iterative self-consistent parallel imaging reconstruction from arbitrary k-space. *Magn Reson Med*. 2010;64:457–471.
3. Sodickson DK, Manning WJ. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. *Magn Reson Med*. 1997;38:591–603.
4. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: sensitivity encoding for fast MRI. *Magn Reson Med*. 1999;42:952–962.
5. Griswold MA, Jakob PM, Heidemann RM, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med*. 2002;47:1202–1210.
6. Peters DC, Korosec FR, Grist TM, et al. Undersampled projection reconstruction applied to MR angiography. *Magn Reson Med*. 2000;43:91–101.
7. Mistretta CA, Wieben O, Velikina J, et al. Highly constrained backprojection for time-resolved MRI. *Magn Reson Med*. 2006;55:30–40.
8. Pruessmann KP, Weiger M, Bornert P, Boesiger P. Advances in sensitivity encoding with arbitrary k-space trajectories. *Magn Reson Med*. 2001;46:638–651.
9. Block KT, Uecker M, Frahm J. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint. *Magn Reson Med*. 2007;57:1086–1098.
10. Lustig M, Donoho D, Pauly JM. Sparse MRI: the application of compressed sensing for rapid MR imaging. *Magn Reson Med*. 2007;58:1182–1195.
11. Knoll F, Bredies K, Pock T, Stollberger R. Second order total generalized variation (TGV) for MRI. *Magn Reson Med*. 2011;65:480–491.
12. Lingala SG, Hu Y, DiBella E, Jacob M. Accelerated dynamic MRI exploiting sparsity and low-rank structure: k-t SLR. *IEEE Trans Med Imaging*. 2011;30:1042–1054.
13. Haldar JP, Liang ZP. Imaging: From nano to macro. In Proceedings of the IEEE International Symposium on Biomedical, Rotterdam, Netherlands, 2010:716–719.
14. Akcakaya M, Basha TA, Goddu B, et al. Low-dimensional-structure self-learning and thresholding: regularization beyond compressed sensing for MRI Reconstruction. *Magn Reson Med*. 2011;66:756–767.

15. Ravishanker S, Bresler Y. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans Med Imaging*. 2011;30:1028-1041.
16. Doneva M, Bornert P, Eggers H, Stehning C, Senegas J, Mertins A. Compressed sensing reconstruction for magnetic resonance parameter mapping. *Magn Reson Med*. 2010;64:1114-1120.
17. Akcakaya M, Basha TA, Chan RH, Manning WJ, Nezafat R. Accelerated isotropic sub-millimeter whole-heart coronary MRI: compressed sensing versus parallel imaging. *Magn Reson Med*. 2014;71:815-822.
18. Sharma SD, Fong CL, Tzung BS, Law M, Nayak KS. Clinical image quality assessment of accelerated magnetic resonance neuroimaging using compressed sensing. *Invest Radiol*. 2013;48:638-645.
19. Wang S, Su Z, Ying L, et al. *Accelerating magnetic resonance imaging via deep learning*. In Proceedings of the 13th IEEE International Symposium on Biomedical Imaging Prague, Czech Republic, 2016:514-517.
20. Yang Y, Sun J, Li H, Xu Z. *Deep ADMM-Net for compressive sensing MRI*. In Proceedings of the 30th International Conference on Neural Information Processing Systems, Pages, Barcelona, Spain, 2016:10-18.
21. Hammernik K, Knoll F, Sodickson D, Pock T. Learning a variational model for compressed sensing MRI reconstruction. In Proceedings of the 24th Annual Meeting of ISMRM, Singapore, 2016. p. 1088. doi: 10.1002/mrm.26977.
22. Han YS, Yoo J, Ye JC. *Deep learning with domain adaptation for accelerated projection reconstruction MR*. *arXiv preprint arXiv*. 2017;1703.01135. doi: 10.1002/mrm.27106.
23. Lee D, Yoo J, Ye JC. *Deep artifact learning for compressed sensing and parallel MRI*. *arXiv preprint arXiv*. 2017;1703.01120.
24. Setsompop K, Gagoski BA, Polimeni JR, Witzel T, Wedeen VJ, Wald LL. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magn Reson Med*. 2012;67:1210-1224.
25. Seiberlich N, Ehse P, Duerk J, Gilkeson R, Griswold M. Improved radial GRAPPA calibration for real-time free-breathing cardiac imaging. *Magn Reson Med*. 2011;65:492-505.
26. Chang Y, Liang D, Ying L. Nonlinear GRAPPA: a kernel approach to parallel MRI reconstruction. *Magn Reson Med*. 2012;68:730-740.
27. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436-444.
28. LeCun Y, Boser B, Denker JS, et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput*. 1989;541-551.
29. Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, NV, 2012. p. 1097-1105.
30. Nair V, Hinton G. Rectified linear units improve restricted Boltzmann machines. In Proceedings of the 27th International Conference on International Conference on Machine, Haifa, Israel: Learning; 2010:807-814.
31. 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 Conference on Computer Vision and Pattern Recognition, Miami, FL, 2009:248-255.
32. Sun Y, Chen Y, Wang, X, Tang X. Deep learning face representation by joint identification-verification. In Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, Canada, 2014:1988-1996.
33. Dong C, Loy CC, He K, Tang X. Image super-resolution using deep convolutional networks. *IEEE Trans Pattern Anal Mach Intell*. 2016;38:295-307.
34. Jain V, Seungh S. Natural image denoising with convolutional networks. In Proceedings of the 21st International Conference on Neural Information Processing, Systems, Vancouver, Canada, 2008:769-776.
35. Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. In Proceedings of the International Conference on Learning Representations, San Juan, Puerto Rico, 2016. arXiv:1511.07122v3.
36. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proc IEEE*. 1998;86:2278-2324.
37. Sutskever I, Martens J, Dahl GE, Hinton G. On the importance of initialization and momentum in deep learning. In Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, 2013. p. 1138-1147.
38. Vedaldi A, Lenc K. *MatConvNet – convolutional neural networks for MATLAB*. *Arxiv preprint arXiv*. 2014;1412.4564.
39. Weingartner S, Akcakaya M, Basha T, et al. Combined saturation/inversion recovery sequences for improved evaluation of scar and diffuse fibrosis in patients with arrhythmia or heart rate variability. *Magn Reson Med*. 2014;71:1024-1034.
40. Jaspan ON, Fleisher R, Lipton ML. Compressed sensing MRI: a review of the clinical literature. *Br J Radiol*. 2015;88:20150487.
41. Robson PM, Grant AK, Madhuranthakam AJ, Lattanzi R, Sodickson DK, McKenzie CA. Comprehensive quantification of signal-to-noise ratio and g-factor for image-based and k-space-based parallel imaging reconstructions. *Magn Reson Med*. 2008;60:895-907.
42. Kellman P, Hansen MS. T1-mapping in the heart: accuracy and precision. *J Cardiovasc Magn Reson*. 2014;16:2.
43. Kwon K, Kim D, Park H. A parallel MR imaging method using multilayer perceptron. *Med Phys*. 2017;44:6209-6224.
44. Sinha N, Saranathan M, Ramakrishnan KR, Suresh S. *Parallel magnetic resonance imaging using neural networks*. doi: 10.1002/mrm.27106. San Antonio, TX; 2007:p. III-149-III-152.
45. Glorot X, Bordes A, Deep BY. *sparse rectifier neural networks*. In Proceedings of the 14th International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, FL; 2011:315-323.
46. Weller DS, Polimeni JR, Grady L, Wald LL, Adalsteinsson E, Goyal VK. Sparsity-promoting calibration for GRAPPA accelerated parallel MRI reconstruction. *IEEE Trans Med Imaging*. 2013;32:1325-1335.
47. Van Essen DC, Smith SM, Barch DM, et al. The WU-Minn human connectome project: an overview. *Neuroimage*. 2013;80:62-79.
48. Blaimer M, Breuer FA, Mueller M, et al. 2D-GRAPPA-operator for faster 3D parallel MRI. *Magn Reson Med*. 2006;56:1359-1364.
49. Breuer FA, Blaimer M, Mueller MF, et al. Controlled aliasing in volumetric parallel imaging (2D CAIPRINHA). *Magn Reson Med*. 2006;55:549-556.
50. Moeller S, Yacoub E, Olman CA, et al. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn Reson Med*. 2010;63:1144-1153.

SUPPLEMENTAL INFORMATION

Additional Supporting Information may be found in the online version of this article.

FIGURE S1: Results from a higher SNR version of the experiment in the Phantom Imaging subsection. These highlight the effect of the heuristic choice of learning rate on the current implementation of the gradient-descent (g.d.) in the RAKI algorithm. RAKI reconstructions using in-vivo optimized (opt.) learning rates (bottom row) exhibits more artifacts than those using phantom optimized learning rates (middle row) during the training stage. All reconstructions perform similarly for $R = 2, 3, 4$, with RAKI using the in-vivo optimized learning rates showing higher noise amplification at $R = 4$. At $R = 5, 6$, all reconstructions suffer from residual aliasing and noise artifacts (arrows). Here, RAKI with phantom optimized learning rate has a similar degree of residual aliasing with GRAPPA (top row), and slightly lower noise amplification. RAKI with the in vivo optimized learning rate displays higher residual aliasing compared to the other two methods. These results indicate the use of fixed learning rates as in the current implementation, is a

limitation, and the use of alternative optimizers for RAKI is warranted during the learning stage.

FIGURE S2: RAKI reconstructions on a 0.6mm isotropic 7T MPRAGE dataset, where the CNN was trained from the full ACS region (top row), and half the ACS region in the readout direction (middle and bottom rows). The reconstructions are almost identical, with minor differences in noise due to the number of points available in the training datasets, showing the training is robust to changes in the calibration region, if the SNR of the calibration region is not substantially altered.

How to cite this article: Akçakaya M, Moeller S, Weingärtner S, Uğurbil K. Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging. *Magn Reson Med*. 2018;00:1–15. <https://doi.org/10.1002/mrm.27420>