Deep learning using a biophysical model for Robust and Accelerated Reconstruction (RoAR) of quantitative and artifact-free $R_2^*$ images

Max Torop$^1$, Satya Kothapalli$^2$, Yu Sun$^1$, Jiaming Liu$^3$, Sayan Kahali$^2$, Dmitriy A. Yablonskiy$^2$, and Ulugbek S. Kamilov$^{1,3}$

$^1$Department of Computer Science and Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA
$^2$Department of Radiology, Washington University in St. Louis, St. Louis, MO 63110, USA
$^3$Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA

December 19, 2019

Running head: Deep learning for robust and fast $R_2^*$ reconstruction

Address correspondence to:
Ulugbek S. Kamilov, PhD
One Brookings Drive
Campus Box 1045
St. Louis, MO 63130, USA.
Email: kamilov@wustl.com

Dmitriy A. Yablonskiy, PhD
Mallinckrodt Institute of Radiology
4525 Scott Ave, Room 3216
St. Louis, MO 63110, USA.
Email: yablonskiyd@wustl.edu

This work was supported in part by NSF award CCF-1813910, NIH/NIA grant R01AG054513, Marilyn Hilton Award for Innovation in MS Research, Amazon Web Services Cloud Credits for Research program, and NVIDIA Corporation with the donation of the Titan Xp GPU for research.

Approximate word count: 250 (Abstract) 3,000 (body)

Submitted to Magnetic Resonance in Medicine.
Abstract

Purpose: To introduce a novel deep learning method for Robust and Accelerated Reconstruction (RoAR) of quantitative $R_2^*$ images from multi-gradient recalled echo (mGRE) MRI data.

Methods: RoAR trains a convolutional neural network (CNN) by adopting a self-supervised learning strategy to produce quantitative $R_2^*$ maps free from field inhomogeneity artifacts given (a) mGRE sequence input data, (b) the biophysical model describing $R_2^*$ signal decay, and (c) preliminary evaluated F-function accounting for contribution of macroscopic magnetic field inhomogeneities to the mGRE signal formation. Importantly, no ground-truth $R_2^*$ images are required and F-function is only needed during RoAR training but not application.

Results: We show that RoAR preserves all features of reconstructed images while offering significant improvements over existing methods in reconstruction speed (seconds vs. hours) and reduced sensitivity to noise. Even for data with SNR=5 RoAR produced $R_2^*$ maps with accuracy of 19% as compared with direct voxel-wise analysis accuracy of 44%. For SNR=10 the RoAR accuracy increased to 13% vs. 22% for direct reconstruction.

Conclusion: RoAR is trained to recognize the macroscopic magnetic field inhomogeneities directly from the input of magnitude-only mGRE data and does not require any ground-truth $R_2^*$ maps. Since RoAR utilizes information on the signal not just from individual voxels but also accounts for spatial patterns of the signals in the images, it reduces sensitivity of $R_2^*$ reconstruction to the noise in the data. These features plus high reconstruction speed provide significant advantages of using RoAR in clinical settings.

Keywords: MRI, gradient echo, image reconstruction, self-supervised learning
Introduction

Multi-Gradient-Recalled-Echo (mGRE) sequences are often used for different MRI applications with some even completing initial multi-center testings (e.g. [1,2]). Since the major goal of mGRE approaches is to produce quantitative metrics related to biological tissue microstructure, it is important to reduce the sensitivity of these metrics to the adverse effects of $B_0$ magnetic field inhomogeneities [3-6], and noise in the MRI signal. It is also important to develop metrics-generating reconstruction algorithms that are fast and robust for direct implementation on MRI scanners.

Recent image reconstruction methods in MRI are increasingly based on deep learning [7-10]. The current paradigm is based on training the weights of an artificial neural network (ANN) over a dataset in order for the network to produce an accurate estimate of the desired images. Generally, such training is done in a supervised fashion by collecting a large dataset of input signals and corresponding “ground-truth” images generated by processing data on a voxel-by-voxel analysis. The key benefit of using deep learning over the traditional voxel-by-voxel reconstruction lies in its excellent resilience to noise and superior computational speed. This was recently demonstrated on a related problem of reconstructing the oxygen extraction fraction (OEF) maps, where an ANN, with a single hidden layer, was trained and applied voxel-wise to produce the desired quantitative parameters given the mGRE signal input [11,12]. The ground-truth datasets for supervised training in these approaches were generated using simulated signals based on the analytical model [13,14].

In this paper, we present a novel reconstruction method called RoAR, based on self-supervised training of a deep convolutional neural network (CNN). RoAR does not need ground truth reconstructed images for training; instead, it is trained using only actual mGRE signals, the pre-calculated contribution of magnetic field inhomogeneities to the mGRE signal decay (described in terms of a factor $F(t)$ [3] in the mGRE signal model), and our knowledge of the analytical biophysical model connecting the mGRE signal with biological tissue microstructure. Herein we exemplify our analysis by demonstrating the efficiency and robustness of the RoAR method for generating quantitative maps of the mGRE signal transverse relaxation metric $R_2^* = 1/T_2^*$. The key advantage of RoAR is that it eliminates the need to explicitly characterize the macroscopic magnetic field inhomogeneities at test time, only requiring $F(t)$ during training. This means that the trained CNN can be directly applied on the mGRE magnitude images without the need
to produce phase images and compute $F(t)$. Additionally, by learning statistical dependencies in neighboring voxels, the CNN acts as an imaging prior that can regularize ill-posed reconstruction problems. Several recent studies have in fact shown that CNNs can serve as excellent priors for a wide range of image reconstruction problems (15–20).

Our results on experimental and synthetic data show that the proposed CNN-based $R_2^*$ reconstruction method not only reduces the reconstruction time by several orders of magnitude compared with the direct re-construction, but also significantly improves the reconstruction quality, thanks to the powerful regularization ability of our trained deep CNNs.

Methods

The mGRE signal at a single voxel can be expressed as (3):

$$S(t) = S_0 \cdot \exp(-R_2^* t - i\omega t) \cdot F(t),$$

where $t$ denotes the gradient echo time, $S_0 = S(0)$ is the signal intensity at $t = 0$ and $\omega$ is a local frequency of MRI signal. The complex valued function $F(t)$ in Eq. [1] describes the effect of macroscopic magnetic field inhomogeneities on the mGRE signal. The failure to account for such inhomogeneities is known to bias and corrupt the recovered $R_2^*$ maps. In this paper we use the Voxel Spread Function (VSF) approach (6) for calculating $F(t)$. The VSF method takes advantage of both the magnitude and phase of mGRE data from the same mGRE scan that is used for generating $R_2^*$ maps. In a standard approach, the $R_2^*$ maps, $\omega$ maps, and $S_0$ are jointly reconstructed from 3D mGRE signals acquired at different echo times $t$ by fitting Eq. [1] with pre-calculated $F(t)$ on a voxel-by-voxel basis to experimental data applying the non-linear least squares (NLLS) analysis. In this paper we propose a reconstruction algorithm that is based on training a convolutional neural network.

RoAR: Architecture and training

Figure 1 presents the details of our CNN model, which is based on the popular 2D U-Net architecture (21). The U-Net has been extensively used in medical image reconstruction and relies on
Figure 1: CNN model with a 10-channel input for the mGRE data $s = (s_1, \ldots, s_{10})$ and 2-channel output for reconstructing $p = (s_0, r^*_2)$ estimate maps of $S_0$ and $R^*_2$. Our model processes data from individual spatial slices extracted from 3D MRI data. The 3D image of the whole brain is obtained by concatenating the outputs of the CNN applied slice-by-slice.

A multi-scale decomposition, based on max-pooling, to make the effective size of its filters in the middle layers larger than that of the early and late layers [22]. Such multi-scale structure leads to a large receptive field of the CNN that has been shown to be effective for removing globally spread imaging artifacts typical in MRI [7].

Let $s = M(p; f)$ denote the magnitude of the signal model in Eq. [1], applied to a single 2D spatial slice extracted from the full 3D MRI data. Here $f$ represents the magnitude of pre-computed $F(t)$ values, stored in an array. We represent the magnitude of the measured mGRE signal of $N$ echo times as the vector

$$s = (s_1, \ldots, s_N), \quad [2]$$

and represent the corresponding absolute value of $S_0$ and true $R^*_2$ maps as another vector

$$p = (s_0, r^*_2). \quad [3]$$

Each vector $s_n$ in $s$ denotes a vectorized 2D image representing the magnitude of the data for one of the echo times.

Let $\hat{p} = I_\theta(s)$ denote our reconstruction model, implemented using our deep CNN architecture, that reconstructs an estimate $\hat{p}$ of the unknown true values of $p$ given the mGRE signal $s$. The vector $\theta$ denotes the trainable set of weights in the CNN. In order to increase the expressive power of the network [23], we rely on multichannel filters, which lead to multiple feature maps at each layer. In our data, the CNN takes $s$ as its 10-channel input and produces $\hat{p} = (\hat{s}_0, \hat{r}^*_2)$ as its 2-channel output. The volumetric image of the whole brain is obtained by applying the model...
Figure 2: Comparison of two approaches for training the reconstruction model. (a) In the standard supervised approach, the reconstruction model $I_\theta$ is optimized for the loss in the image domain, so that $\hat{p} = I_\theta(s)$ is close to the corresponding ground-truth images $p$. (b) In the proposed self-supervised approach, only access to the measurements $s$ and the biophysical model $M$ is assumed. The loss is formulated in the measurement domain, and the reconstruction model is trained so that $\hat{s} = M(I_\theta(s); f)$ is close to $s$

slice by slice. Note the difference between our deep convolutional architecture from that of recent approaches (11, 12) that apply a fully connected ANN voxel-by-voxel to map the mGRE signal to the desired quantitative parameters. These architectures consist of one input layer, one hidden layer with 10 nodes, and one output layer. The convolutional structure of our architecture allows it to process the whole slice of the volumetric data, thus taking into account complex statistical relationships between pixels and echo times.

The traditional supervised learning, illustrated in Figure 2(a), is carried out by minimizing the empirical loss over a training set consisting of $L$ slices $\{(s_\ell, p_\ell)\}_{\ell=1,...,L}$, as follows

$$\min_{\theta} \sum_{\ell=1}^{L} \mathcal{L}(I_\theta(s_\ell), p_\ell), \quad [4]$$

where $\mathcal{L}$ measures the discrepancy between the reconstruction $\hat{p}_\ell = I_\theta(s_\ell)$ generated by the CNN and the ground-truth $p_\ell$. Typical choices for $\mathcal{L}$ include the Euclidean and the $\ell_1$ distances. In practice, the minimization problem is solved by using stochastic gradient-based optimization algorithms such as Adam (24, 25). Once the optimal set of parameters $\theta^*$ has been learned on the training data, the operator $I_{\theta^*}$ is applied to solve the image reconstruction problem on previously unseen data. The major limitation of the traditional learning paradigm is that it requires large amounts of ground-truth data $\{p_\ell\}$ for training, which can be challenging to obtain in practice. One strategy, adopted by (11, 12), is to synthetically generate such data in individual voxels using the signal model (e.g. Eq. [1]). However, this strategy can be sensitive to the mismatch between the actual statistical distribution of entries in $p_\ell$ (which might have spatial statistical dependencies) and the assumed distribution used in simulation (often independent in space).
The key idea of RoAR is to use self-supervised learning, illustrated in Figure 2(b), to train the parameters $\theta$ of the reconstruction model $I_\theta$. The idea of using self-supervised learning has recently gained popularity in several distinct imaging applications for addressing the lack of ground-truth training data [26–30]. The self-supervised learning in RoAR is enabled through our knowledge of the analytical biophysical model connecting the mGRE signal with biological tissue microstructure.

Given a set of mGRE measurements $\{s_\ell\}$ and a set of arrays containing corresponding F-function values $\{f_\ell\}$ calculated from the same dataset using the VSF method, our approach can be formalized as the following optimization problem

$$\min_\theta \sum_{\ell=1}^L \mathcal{L}(M(I_\theta(s_\ell); f_\ell), s_\ell). \quad [5]$$

Note that the loss function $\mathcal{L}$ in Eq. [5] operates exclusively in the measurement space and does not require any ground-truth data $\{p_\ell\}$.

Solving the above optimization problem for the $\ell$th data element yields $\hat{p}_\ell = I_\theta(s_\ell)$, which acts as a latent image in the intermediate stage for our optimization, as shown in Figure 2(b). Intuitively, our method is searching for images $\hat{p}_\ell$ that are parameterized by $(\theta, \{s_\ell\})$ that best explain the measured mGRE signal dataset $\{s_\ell\}$. This strategy is called self-supervised because the measurements themselves provide the supervision to solve the reconstruction problem by exploiting the signal model $\mathcal{M}$ and the prior induced by the CNN.

**Denoising with RoAR**

In addition to RoAR’s ability to generate $R_2^*$ maps free of field inhomogeneity artifacts, RoAR can also be trained to decrease the influence of noise on RoAR-generated $R_2^*$ maps. This can be achieved by generating synthetic mGRE data $\{s_\ell\}$ from high SNR data and adding different levels of noise to produce datasets $\{\tilde{s}_\ell\}$, so that $\tilde{s}_\ell = s_\ell + e_\ell$, where $e_\ell$ denotes the noise and train RoAR by solving the following optimization problem

$$\min_\theta \sum_{\ell=1}^L \mathcal{L}(\mathcal{M}(I_\theta(\tilde{s}_\ell); f_\ell), s_\ell). \quad [6]$$

Here, it is important to emphasize that RoAR requires synthetic data to be generated from high SNR data only during training. As corroborated by our results, the trained model yields excellent
reconstruction results at test time, even on previously unseen data with high amounts of noise. As described below, the training is done by training the CNN with data containing varying amounts of noise.

**In vivo brain dataset**

For validating our method, we used previously published (31) brain image data collected from 26 healthy volunteers (age range 26-76) using a Siemens 3T Trio MRI scanner and a 32-channel phased-array head coil. Studies were conducted with approval of the local IRB of Washington University. All volunteers provided informed consent. The data was obtained using a 3D version of the mGRE sequence with 10 gradient echoes followed by a navigator echo (32) used to reduce artifacts induced by physiological fluctuations during the scan. Sequence parameters were flip angle $\text{FA} = 30^\circ$, voxel size of $1 \times 1 \times 2$ mm$^3$, first echo time $t_1 = 4$ ms, echo spacing $\Delta t = 4$ ms (monopolar readout gradients), repetition time $\text{TR} = 50$ ms, and the total imaging time for each acquisition was around 10 min. 9 of the 26 volunteers were scanned twice, at different times, making for a set of 35 MRI’s.

After applying the Fourier transform to the k-space, data from different channels were combined for each voxel to give a single mGRE signal $S(t)$ as described in (33). The effects of macroscopic magnetic field inhomogeneities were taken into account by including in CNN training the function $F(t)$, Eq [1], pre-estimated using the voxel spread function (VSF) approach (6).

**Performance evaluation**

We evaluated our method on both experimentally measured and synthetic data and compared the results to the traditional NLLS approach. First, we have directly trained RoAR on in vivo data to test its performance on actual experimental MRI data. The ability of RoAR to correct for noise is validated and quantified by training our model on noisy synthetic data with different noise levels. All neural networks were trained on a GeForce GTX 1080 Ti GPU (NVIDIA Corporation, Santa Clara, CA, USA), and implemented in TensorFlow (34), using the Adam optimizer to minimize the Euclidean distance
Before running NLLS on in vivo MRI’s, the VSF method was used to find an \( F(t) \) value at each voxel and each echo time \( t \). Here \( F(t) \) accounts for the adverse effect of macroscopic magnetic field inhomogeneities on a signal from a given voxel, at echo time \( t \) and is essential for true \( R^*_2 \) evaluation. At each iteration of the regression the signal model from Eq. [1] is parameterized by both the \( S_0 \) and \( R^*_2 \) reconstructions as well as given \( F(t) \). This accounts for the contribution of field inhomogeneities in the input and results in reconstructions without them. A brain extraction tool, implemented in the Functional Magnetic Resonance Imaging of the Brain Library (FMRIB), was used to mask out both skull and background voxels in all MRI data (35). NLLS was run over only the set of unmasked voxels, optimizing for 400 iterations at each spatial point. This method was implemented in MatLab R2018b (MathWorks, Natick, MA). The results of this method are then compared to those of a RoAR instance trained on in vivo data for 220 epochs (7 hours).

To further decrease the influence of noise on RoAR-generated \( R^*_2 \) maps and to test the efficacy of this method, we elected to generate synthetic 3D mGRE signals that are based on real mGRE data to serve as a “ground truth”. This approach ensures that synthetic data reflects the spatial statistical dependencies seen in real mGRE data. To do this we first run the NLLS method on a high SNR in vivo MRI to get \( S_0 \) and \( R^*_2 \) reconstructions. These reconstructions are used to parameterize the magnitude of the signal model shown in Eq. [1], generating signal on a voxel-by-voxel basis. Together these generated signals constitute full synthetic 3D mGREs. The 3D \( S_0 \) and \( R^*_2 \) maps used to generate a given synthetic MRI can be thought of as its ground truth, which we use only at test time to compare our methods. Furthermore these synthetic MRI’s inherit the realistic structure of the NLLS reconstructions used to generate them, which themselves get their structure from the in vivo MRI’s used for their reconstruction. Since at this stage we are only interested in the influence of noise on RoAR’s performance, in the synthetic mGRE data we excluded the \( F(t) \) factor accounting for the magnetic field inhomogeneities. The synthetic data was split into 23 MRI’s for training, 4 for validation and 8 for testing. MRI’s from patients who were scanned twice were put in the same sets, to avoid biasing our results.

In the synthetic training paradigm, we use copies of the simulated MRI’s with added noise. This noise is added before MRI’s are split into slices for training, and comes from the distribution \( \mathcal{N}(0, \bar{S}_0 \text{SNR}) \) where \( \bar{S}_0 \) represents the mean \( S_0 \) value over the entire 3D volume of a given MRI. The use of \( \bar{S}_0 \) ensures that noise level is standardized across MRI’s with different signal strength, and the use of SNR allows for control over noise level. For each copy SNR is randomly selected from
Figure 3: Examples of $R_2^*$ reconstructions from NLLS and RoAR methods from two in vivo slices. VSF was used to calculate NLLS results and train RoAR, but not during RoAR reconstruction. The top and bottom left images show Echo 1 of the 10 input images used to produce these reconstructions. The two rightmost images are maps of differences between the RoAR and NLLS results.

The interval $[5, 20]$ was used to make our models robust to a range of noise in input. Four copies of the synthetic data, and their noisy counterparts were used to train RoAR’s U-Net, which lasted for approximately 7 hours (60 epochs). Additionally, three sets were made, each consisting of copies of the 8 synthetic test MRI’s, with the noise level of all MRI’s in a set corresponding to either SNR = 5, 10 or 15. At test time, these sets were used to compare how robust both methods are to different levels of noise. We use the relative error (RE) metric as a means to quantitatively compare a reconstructed slice’s similarity to ground truth

$$RE = \frac{\|\tilde{r}_2^* - r_2^*\|}{\|r_2^*\|} \times 100\%,$$

where $\tilde{r}_2^*$ and $r_2^*$ represent a given methods $R_2^*$ slice output and the ground truth, respectively. We note that RE is computed inside brain mask.
Results

Figure 3 shows examples for two in vivo slices of the $R^*_2$ calculated by NLLS and RoAR trained on in vivo data, as described above. It can be seen that the reconstructions between the two methods are both of high quality and almost identical (SNR in original mGRE data is about 50). Importantly, in the RoAR approach, F-function was only used during training but not reconstruction, i.e. the final CNN model $I_\theta$ operates in the domain of magnitude images $s$ and does not need input from $F(t)$.

Table 1 provides a summary of the average relative error of results from the methods over all brain slices from the 8 synthetic test MRI’s, given separately for each of the three noisy test sets. Here brain slices correspond to slice 25 to 55 of a 72 slice MRI, 20 to 50 of a 60 slice MRI, and 30 to 60 when there are 88 slices. At every input noise level, RoAR has much lower RE then NLLS. The quality gap between RoAR and NLLS grows noticeably as the input gets noisier: from around 5% at SNR=15 to around 25% at SNR=5.

Table 1: Average Relative Errors of $R^*_2$ evaluation from NLLS and RoAR methods on the synthetic test data for three different noise levels. RE were computed inside brain masks that insure removing all skull voxels where the signal model in Eq. [1] is not applicable

| Method | SNR=5 | SNR=10 | SNR=15 |
|--------|-------|--------|--------|
| NLLS   | 43.9 %| 22.5 % | 15.2 % |
| RoAR   | 18.9 %| 13.0 % | 10.2 % |

Figure 4 shows examples of reconstruction results for the same two slices, from test sets with different noise levels (SNR 5 and 10). All NLLS reconstructions have higher noise levels when compared with the corresponding RoAR results. This is highlighted by the difference maps located under each reconstructed slice, visualizing the absolute value of their deviation from the ground truth. The NLLS difference maps are much brighter for the low SNR data and, although less extreme, are noticeably brighter than RoAR’s difference maps for the higher SNR data. The RE of each reconstructed image is also shown in the bottom left corner. In all images shown the RE of the RoAR results are lower than NLLS’s.
Discussion and Conclusions

In this manuscript we proposed a self-supervised Convolutional Neural Network approach for fast and robust reconstruction of $R_2^*$ maps from a multi-Gradient-Recalled Echo MRI data. The method is based on a deep learning that uses a biophysical model connecting MRI signal with underlying biological tissue microstructure. Figure 3 shows that RoAR has the ability to produce $R_2^*$ images from high SNR in vivo data of the same quality as NLLS-based voxel-by-voxel analysis. In this regard there are two important positives of using RoAR over NLLS. The first is RoAR’s faster runtime: RoAR takes 5 seconds to output $R_2^*$ and $S_0$ maps for a full brain (using a GeForce GTX 1080 Ti GPU) while NLLS takes 120 minutes on a modern PC (using 8 cores). The second improvement is that RoAR operates in the domain of magnitude mGRE images and is trained to recognize the contribution of macroscopic magnetic field inhomogeneities to the mGRE signal only from the magnitude data, thus providing $R_2^*$ maps free from macroscopic magnetic field inhomogeneity artifacts. At the same time, a standard NLLS analysis requires both magnitude and phase images to compute $F(t)$ values that are subsequently used during NLLS fitting. It is important to emphasize, that the pre-computed $F(t)$ is an essential input to the biophysical model that is used to train RoAR but $F(t)$ is not used by afterwards “trained” RoAR to produce $R_2^*$ maps. The input being transformed by the U-Net is just magnitude mGRE images and the output is $R_2^*$ maps free from macroscopic magnetic field inhomogeneity artifacts even though mGRE images are naturally affected by magnetic field inhomogeneities.

The synthetic MRI experiments show that RoAR is significantly more robust than NLLS with respect to noise in the data: as illustrated by Table 1 and Figure 4, the Relative Error in $R_2^*$ evaluation is significantly smaller in RoAR-reconstructed data than in the NLLS results. This is due to two limitations in NLLS. While RoAR reconstruction is based on the data from the entire image, the NLLS is a voxel-based reconstruction. Because of this NLLS cannot take advantage of the naturally occurring spatial patterns in the images that RoAR benefits from. The second limitation of NLLS is that it does not map $R_2^*$ from noisy signals from individual voxels to their cleaner $R_2^*$ counterparts. While RoAR was trained to take noisy signal and output $R_2^*$ corresponding to cleaner signal, NLLS is simply runs directly on the noisy data. This further limits NLLS output quality as it never gets to ”see” examples of what high SNR data looks like.
Figure 4: Two examples of the results obtained from the synthetic datasets with different noise levels (SNR 5 and 10). Images of Echo 1 ($s_1$) of the method input (before the addition of noise) and the ground truth $R^*_2$ are shown in the first column. Columns 2 and 3 correspond to reconstructions from input with SNR 10, while columns 4 and 5 correspond to those from input with SNR 5. The RE of each reconstruction is shown in its bottom left corner. Representative regions with 2X zoom are shown on the bottom right of each reconstruction. Underneath each $R^*_2$ reconstruction is its difference with the ground truth.

It is clear from the synthetic results that the relative improvement of $R^*_2$ estimates by RoAR
over NLLS grows as the input gets noisier. In Table 1 the difference between RoAR and NLLS’s RE changes from 4.9% to 9.6% when going from SNR=15 to SNR=10 and shoots up to 25.1% when going from SNR=10 to SNR=5. The results in Figures 4 qualitatively show this relative quality increase. While also strong with lower noise input, RoAR becomes an increasingly attractive options as noise increases.

In Conclusion, we introduced RoAR as a fast, self-supervised, method that can utilize only magnitude mGRE data to produce high quality $R_2^*$ maps free from artifacts resulting from the macroscopic magnetic field inhomogenieties. We also demonstrated RoAR’s ability to reconstruct $R_2^*$ maps with high accuracy even from noisy mGRE signals. This can be achieved because, as a CNN based approach, RoAR utilizes information on the signal not just from individual voxels (as in a voxel-based NLLS approach) but also spatial patterns of the signals in the images. The resilience of RoAR to the noise in the data, and high reconstruction speed provide significant advantages of using RoAR in clinical settings.

Acknowledgement: The authors are grateful to Vadim Omeltchenko, Sr. AWS Solutions Architect, for helpful discussion.

References

1. Ropele S, Wattjes MP, Langkammer C et al. Multicenter mapping in the healthy brain. Magn Reson Med 2014; 71:1103–1107.

2. Weiskopf N, Suckling J, Williams G et al. Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: a multi-center validation. Front Neurosci 2013; 7:95.

3. Yablonskiy DA. Quantitation of intrinsic magnetic susceptibility-related effects in a tissue matrix. Phantom study. Magn Reson Med 1998; 39:417–428.

4. Fernandez-Seara M, Wehrli F. Postprocessing technique to correct for background gradients in image-based R*(2) measurements. Magn Reson Med 2000; 44:358–366.

5. Hernando D, Vigen KK, Shimakawa AS, Reeder SB. R*(2) mapping in the presence of macroscopic B0 field variations. Magn Reson Med 2012; 68:3:830–40.
6. Yablonskiy DA, Sukstanskii AL, Luo J, Wang X. Voxel spread function method for correction of magnetic field inhomogeneity effects in quantitative gradient-echo-based MRI. Magn Reson Med 2013; 70:1283–1292.

7. Han YS, Yoo J, Ye JC. Deep learning with domain adaptation for accelerated projection-reconstruction MR. Magn Reson Med 2017; 80:1189–1205.

8. Lee D, Yoo J, Tak S, Ye JC. Deep Residual Learning for Accelerated MRI Using Magnitude and Phase Networks. IEEE Trans Biomed Eng 2018; 65:1985–1995.

9. Aggarwal HK, Mani MP, Jacob M. MoDL: Model Based Deep Learning Architecture for Inverse Problems. IEEE Trans Med Imag 2018; 38:394–405.

10. Lee J, Han Y, Ryu JK, Park JY, Ye JC. k-Space Deep Learning for reference-free EPI ghost correction. Magn Reson Med 2019; 82:2299–2313.

11. Domsch S, Mürle B, Weingärtner S, Zapp J, Wenz F, Schad LR. Oxygen extraction fraction mapping at 3 Tesla using an artificial neural network: A feasibility study. Magn Reson Med 2018; 79:890–899.

12. Hubertus S, Thomas S, Cho J, Zhang S, Wang Y, Schad LR. Using an artificial neural network for fast mapping of the oxygen extraction fraction with combined QSM and quantitative BOLD. Magn Reson Med 2019; 82:2199–2211.

13. Yablonskiy DA, Haacke EM. Theory of NMR signal behaviour in magnetically inhomogeneous tissues: the static dephasing regime. Magn Reson Med 1994; 32:749–763.

14. Ulrich X, Yablonskiy DA. Separation of cellular and BOLD contributions to T2* signal relaxation. Magn Reson Med 2016; 75:606–615.

15. Veen DV, Jalal A, Price E, Vishwanath S, Dimakis AG. Compressed Sensing with Deep Image Prior and Learned Regularization. ArXiv 2018; abs/1806.06438.

16. Sun Y, Xia Z, Kamilov US. Efficient and accurate inversion of multiple scattering with deep learning. Opt. Express 2018; 26:14678–14688.

17. Ulyanov D, Vedaldi A, Lempitsky V. Deep Image Prior. In: Proc IEEE Conf Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, June 2018 pp. 9446–9454.
18. Liu J, Sun Y, Xu X, Kamilov US. Image Restoration Using Total Variation Regularized Deep Image Prior. In: Proc IEEE Int Conf Acoustics, Speech and Signal Process, Brighton, UK, May 2019 pp. 7715–7719.

19. Sun Y, Liu J, Kamilov US. Block coordinate regularization by denoising. In: Proc. Advances in Neural Information Processing Systems 33, Vancouver, BC, Canada, December 8-14, 2019

20. Mataev G, Milanfar P, Elad M. DeepRED: Deep Image Prior Powered by RED. In: Proceedings of the IEEE International Conference on Computer Vision Workshops, Seoul, South Korea, October 2019 pp. 0–0.

21. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, Germany, November 2015 pp. 234–241.

22. Jin KH, McCann MT, Froustey E, Unser M. Deep Convolutional Neural Network for Inverse Problems in Imaging. IEEE Trans Image Process 2017; 26:4509–4522.

23. Krizhevsky A, Sutskevar I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks. In: Proc. Advances in Neural Information Processing Systems 25, Lake Tahoe, NV, USA, December 3-8, 2012 pp. 1097–1105.

24. Bottou L. “Neural Networks: Tricks of the Trade”, chapter Stochastic Gradient Descent Tricks, pp. 421–437. Springer, 2 ed., September 2012.

25. Kingma D, Ba J. Adam: A Method for Stochastic Optimization. In: International Conference on Learning Representations (ICLR), San Diego, USA, May 2015

26. Yaman B, Hosseini SAH, Moeller S, Ellermann J, Ugurbil K, Akakaya M. Self-Supervised Physics-Based Deep Learning MRI Reconstruction without Fully-Sampled Data. ArXiv 2019; abs/1910.09116.

27. Chen L, Bentley P, Mori K, Misawa K, Fujiwara M, Rueckert D. Self-supervised learning for medical image analysis using image context restoration. Med Image Anal 2019; 58:101539.

28. Kim B, Ye JC. Mumford–Shah loss functional for image segmentation with deep learning. IEEE Trans Image Process 2019; 29:1856–1866.
29. Senouf O, Vedula S, Weiss T, Bronstein A, Michailovich OV, Zibulevsky M. Self-supervised learning of inverse problem solvers in medical imaging. In: DART/MIL3ID@MICCAI, Shenzen, China, October 2019.

30. Lu AX, Kraus OZ, Cooper S, Moses AM. Learning unsupervised feature representations for single cell microscopy images with paired cell inpainting. PLOS Comput Biol 2019; 15:e1007348.

31. Wen J, Goyal MS, Astafiev SV, Raichle ME, Yablonskiy DA. Genetically defined cellular correlates of the baseline brain MRI signal. Proceedings of the National Academy of Sciences 2018; 115:E9727–E9736.

32. Wen J, Cross AH, Yablonskiy DA. On the role of physiological fluctuations in quantitative gradient echo MRI: implications for GEPCI, QSM, and SWI. Magn Reson Med 2015; 73:195–203.

33. Luo J, Jagadeesan BD, Cross AH, Yablonskiy DA. Gradient Echo Plural Contrast Imaging – Signal model and derived contrasts: T2*, T1, phase, SWI, T1f, FST2* and T2*-SWI. Neuroimage 2012; 60:1073–1082.

34. Abadi M, Agarwal A, Barham P et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. ArXiv 2015; abs/1603.04467.

35. Pechaud M, Jenkinson M, Smith S, BET2-MRI-Based Estimation of Brain, Skull and Scalp Surfaces, Technical report, FMRIB Technical Report TR06MP1. Oxford, England: Oxford University, 2006.
List of Figures

1 CNN model with a 10-channel input for the mGRE data $s = (s_1, \ldots, s_{10})$ and 2-channel output for reconstructing $p = (s_0, r_2^*)$ estimate maps of $S_0$ and $R_2^*$. Our model processes data from individual spatial slices extracted from 3D MRI data. The 3D image of the whole brain is obtained by concatenating the outputs of the CNN applied slice-by-slice.

2 Comparison of two approaches for training the reconstruction model. (a) In the standard supervised approach, the reconstruction model $I_\theta$ is optimized for the loss in the image domain, so that $\hat{p} = I_\theta(s)$ is close to the corresponding ground-truth images $p$. (b) In the proposed self-supervised approach, only access to the measurements $s$ and the biophysical model $M$ is assumed. The loss is formulated in the measurement domain, and the reconstruction model is trained so that $\hat{s} = M(I_\theta(s); f)$ is close to $s$.

3 Examples of $R_2^*$ reconstructions from NLLS and RoAR methods from two in vivo slices. VSF was used to calculate NLLS results and train RoAR, but not during RoAR reconstruction. The top and bottom left images show Echo 1 of the 10 input images used to produce these reconstructions. The two rightmost images are maps of differences between the RoAR and NLLS results.

4 Two examples of the results obtained from the synthetic datasets with different noise levels (SNR 5 and 10). Images of Echo 1 ($s_1$) of the method input (before the addition of noise) and the ground truth $R_2^*$ are shown in the first column. Columns 2 and 3 correspond to reconstructions from input with SNR 10, while columns 4 and 5 correspond to those from input with SNR 5. The RE of each reconstruction is shown in its bottom left corner. Representative regions with 2X zoom are shown on the bottom right of each reconstruction. Underneath each $R_2^*$ reconstruction is its difference with the ground truth.
List of Tables

1 Average Relative Errors of $R^*_2$ evaluation from NLLS and RoAR methods on the synthetic test data for three different noise levels. RE were computed inside brain masks that insure removing all skull voxels where the signal model in Eq. [1] is not applicable ........................................... 11