Robust Multi-shot EPI with Untrained Artificial Neural Networks: Unsupervised Scan-specific Deep Learning for Blip Up-Down Acquisition (BUDA)

Tae Hyung Kim1,2,3, Zijing Zhang1,2,4, Jaejin Cho1,2, Borjan Gagoski4,5, Justin Haider2, and Berkin Bilgic1,2
1Athens University of Technology, Athens, Greece, 2A. A. Martinos Center for Biomedical Imaging, Boston, MA, United States, 3Radiology, Harvard Medical School, Boston, MA, United States, 4Electrical Engineering, University of Southern California, Los Angeles, CA, United States, 5State Key Laboratory of Modern Optical Instrumentation, College of Optical Science and Engineering, Zhejiang University, Hangzhou, China, 6Boston Children’s Hospital, Boston, MA, United States

Synopsis

Blip Up-Down Acquisition (BUDA) has been successful in generating distortion-free multi-shot EPI (msEPI) without navigators, utilizing a fieldmap and structured low-rank constraints. Recently, a scan-specific artificial neural network (ANN) motivated by structured low-rank modeling, named LORAKI, has been proposed for refined MRI reconstruction, where its training employed fully-sampled autocalibrated signal (ACS). Although applying LORAKI framework to BUDA is beneficial, acquiring fully-sampled ACS for msEPI is not practical. We propose scan-specific unsupervised ANNs for improved BUDA msEPI without training data. Experiment results indicate that the proposed BUDA-LORAKI exhibits advantages, with up to 1.5x reduction in NRMSE compared to standard BUDA reconstruction.

Introduction

Multi-shot EPI has been extensively employed for fast imaging with reduced distortion and blurring over single-shot acquisition. However, resolving shot-to-shot phase variations remains as a challenge. Blip Up-Down Acquisition (BUDA) and related approaches14,5 have been proposed, utilizing interleaved blip-up and -down k-space data acquisition, a fieldmap in the forward model for distortion correction and the reconstruction using structured low-rank modeling6.

A recent work7 has shown that model-based structured low-rank matrix modeling8–15 can be generalized into artificial neural network (ANN) models. Its iterative structure enables the use of a small parametric space that can be trained by an autocalibration signal (ACS). It prevents potential “hallucinations” introduced from external training data which do not generalize to test acquisitions, and are advantageous when large-scale training data are unavailable. Therefore, incorporating LORAKI to BUDA will be impactful, but acquiring ACS with msEPI would necessitate uneven echo spacings, longer readouts, increased blurring, and later TE.

In this work, we propose BUDA-LORAKI, a scan-specific unsupervised ANN for BUDA msEPI. Recent studies16,17 have shown that the structure of ANNs can be powerful for inverse problems without training data (“untrained” neural networks16,17).

Theory

The following can be formulated for BUDA reconstruction using LORAKS low-rank framework8–10,

$$\min_\rho f(\rho) = \min_\rho \|A\rho - d\|^2 + \mu J(B\rho),$$

with the BUDA forward model A (coil sensitivity, fieldmap, Fourier transform, and undersampling), the distortion-free images \(\rho\), the acquired data \(d\), the SENSE forward model B (coil sensitivity and Fourier transform), and the LORAKS regularization \(J(\cdot)\). We will name this as BUDA-LORAKS.

It has been shown that the low-rank methods8–15 are composed of convolutional operations, so their iterative counterparts (e.g. Landweber iteration, Conjugate Gradient) can be represented as convolutional neural network models2. Specifically, the gradient of the above objective function is

$$\nabla f(\rho) = A^H A(\rho) - A^H d + \mu B^H \left( c_o (c_b (B\rho)) \right),$$

with \(c_o, c_b\) are convolution operations corresponding to the LORAKS regularization. The LORAKS framework was introduced from this observation, by adding nonlinear ReLU activation function. Let \(g(\rho)\) be a blind objective function for LORAKS, we define its gradient as,

$$\nabla g(\rho) = A^H A(\rho) - A^H d + \mu B^H \left( c_b (ReLU (c_o (B\rho))) \right),$$

providing improved reconstruction by added nonlinearities7.

The proposed network structure adopted conjugate gradient (CG) algorithm30. Unlike standard LORAKS requiring ACS for training, we propose to circumvent the disadvantages of incorporating ACS into msEPI by unsupervised reconstruction. We trained the network without training data, using the following loss function, let \(y = h_\theta(x)\) be the output of the proposed LORAKS network (reconstructed blip-up and -down images) with the network parameter \(\theta\),

$$\min_\theta \| A(y) - d \|_1 + \lambda_1 \sum_i |\text{y}_{di}|^2 + |\text{y}_{wi}|^2 + \lambda_2 TV(y),$$

with \(|\text{y}_{wi}|, |\text{y}_{di}|\) are magnitude of reconstructed blip-up and -down images at voxel \(i\), respectively. The first term enables training without the fully-sampled training data\(^1\), and second and third terms are specific to improve msEPI reconstruction, where the magnitude of blip-up and -down images should be consistent (taking magnitude prevents us from vulnerability to shot-to-shot phase variations), and the reconstructed images are spatially smooth through total variation.

Method

Three different dataset were acquired on 3T Siemens scanners using 32-channel reception.

1. msEPI diffusion: Using a Prisma system, 1x1x4 mm\(^3\), b=1000 s/mm\(^2\) diffusion data were acquired at Rep/TE=4ms/83ms, using 4 shots (2 blip-up, 2 -down) and TR/TE=83ms/3900ms. Using all 4 shots provided a reference BUDA reconstruction, after which only 2 shots were utilized to compare undersampled reconstructions.
References


Figures

Figure 1. The structure of the proposed network. “A” represents the BUDA forward model, “B” is SENSE-encoding (coil sensitivities and Fourier transform), “VC” represents augmentation of virtual conjugate coils for the phase constraint. Some parameter choices are: the number of iteration $K=7$, convolution kernel size=7, the number of output channels of the first convolutional layer=64, $\lambda_1 = 1$, $\lambda_2 = 0.5$.

Figure 2. The reconstruction results of the diffusion data with $b=1000\text{s/mm}^2$. Top two rows display reconstructed images, the bottom two rows show error images (10x-amplified). For each shot, 4x in-plane acceleration was applied. Two shots (one blip-up and one blip-down) were used for reconstruction. The reference images were generated by combining 4 shots (2 blip-up and 2 blip-down) through BUDA. Naive SENSE does not include the field map, resulting in distortion near frontal lobes.

Figure 3. The reconstruction results with GRE data from two shots (one blip-up and one blip-down) of 4x in-plane acceleration.
Figure 4. The reconstruction results with SAGE SMS data with four shots (two blip-up and two blip-down) of 8x in-plane acceleration with 2-slice SMS. The reference images were generated by combining 8 shots (4 blip-up and 4-down) through BUDA. Top two rows display reconstructed images, the bottom two rows show error images (10x-amplified).

Figure 5. T2 and $T2^*$ maps estimated from 9 echoes of the reconstructed SAGE SMS data. The figure shows only one slice.