PHYSICS-DRIVEN DEEP TRAINING OF DICTIONARY-BASED ALGORITHMS FOR MR IMAGE RECONSTRUCTION

Saiprasad Ravishankar, Il Yong Chun, and Jeffrey A. Fessler

Dept. of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA

ABSTRACT

Techniques involving learned dictionaries can outperform conventional approaches involving (nontrained) analytical sparsifying models for MR image reconstruction. Inspired by iterative dictionary learning-based reconstruction methods, we propose a novel efficient image reconstruction framework involving multiple iterations (or layers). Each layer involves applying a transformation to image patches, thresholding, and then reconstructing the patches in a dictionary, followed by an update of the image using observed k-space measurements. We train the transforms, thresholds, and dictionaries within the multi-layer algorithm to minimize reconstruction errors. Our experiments demonstrate that for highly undersampled k-space data, such trained reconstruction algorithms provide high quality results.

Index Terms— Sparse representations, Dictionary learning, Transform learning, Fast algorithms, Thresholding, Inverse problems, Magnetic resonance imaging, Machine learning.

1. INTRODUCTION

Image reconstruction in imaging modalities such as magnetic resonance imaging (MRI) or X-ray computed tomography, etc., often involves solving ill-posed inverse problems. MRI is a popular but relatively slow imaging modality, where measurements are samples in the Fourier space (or k-space) of the object, and are acquired only sequentially. Subsampling of k-space data can significantly accelerate MRI.

While conventional image reconstruction methods for limited k-space data [1–5] exploit the sparsity of images in analytical dictionaries or sparsifying transforms or alternative properties such as low-rank properties [6], recent research has demonstrated benefits of learned image models including synthesis dictionaries [7] and sparsifying transforms [8] for image reconstruction. The dictionaries or transforms could be learned from training data [9, 10] and used for image reconstruction, or they could be learned simultaneously while reconstructing images [7, 11] from limited data. The latter approach is referred to as blind compressed sensing. For example, given k-space data $y \in \mathbb{C}^m$, the following problem can be solved for jointly reconstructing the image $x \in \mathbb{C}^p$ and learning a dictionary (often much smaller than the image) $D \in \mathbb{C}^{n \times L}$ [7]:

$$\min_{x, D \in \mathcal{D}, Z} \sum_{j=1}^{N} \|P_j x - D z_j\|_2^2 + \nu \|Ax - y\|_2^2 \quad \text{s.t.} \quad \|z_j\|_0 \leq s \; \forall \, j. \tag{1}$$

Here, A denotes the sensing or measurement operator for the imaging modality (for MRI with undersampled k-space data, this is the undersampled Fourier encoding), $\nu > 0$ is a parameter set according to noise level in k-space, P_j is an operator that extracts a patch of image x as a vector $P_j x$, and z_j denotes the sparse coefficient vector for the patch, with $P_j x \approx D z_j$. Matrix $Z \in \mathbb{C}^{L \times N}$ has the vectors z_j as its columns, and the ℓ_0 "norm" counts the number of non-zero entries in a vector. While Problem (1) uses an ℓ_0 sparsity constraint with parameter s, alternative versions can be readily constructed using the ℓ_1 norm or using sparsity penalties. The set $\mathcal{D}$ denotes the set of feasible dictionaries, e.g., matrices with unit ℓ_2 norm columns [12].

Methods for Problem (1) (cf. [7]) start with an initial image and alternate between learning the dictionary and coefficients (D, Z) (*dictionary learning step*), and updating the image (a least squares *image update step*) to account for the imaging model $y \approx Ax$. These methods tend to be computationally expensive as they involve learning dictionaries and sparse coefficients for many image patches in each iteration.

In this work, we propose a novel efficient physics-driven MR image reconstruction framework involving multiple iterations or layers. In each layer, the method first decorrupts (removes artifacts) image patches by applying a sparsifying transform, nonlinearity (shrinkage), and dictionary, and then performs a least squares update of the image accounting for the imaging model or physics. The unique aspect of the proposed approach is that we *train* the parameters (transforms, dictionaries, and soft-shrinkage functions) of each layer of the image reconstruction algorithm to explicitly minimize reconstruction errors using training data synthesized by retrospective undersampling. The learned algorithm can then be efficiently applied to reconstruct other test data with finite number of layers. We present preliminary results for MR image reconstruction from limited data illustrating the performance of our trained sparsity-driven algorithms.

2. PROPOSED IMAGE RECONSTRUCTION METHOD AND TRAINING

We propose an image reconstruction framework that involves multiple iterations called layers [13], where each layer involves a *decorruption step* for image patches, followed by an *image update step* accounting for the imaging model. The decorruption step in the kth layer involves applying a transformation $W^k \in \mathbb{C}^{L\times n}$ to image patches, thresholding the result by soft-thresholding with $\gamma^k \in \mathbb{R}^L$ denoting a vector of thresholds (a different threshold corresponding to each transform row), and multiplying the (sparse) output by a dictionary $D^k \in \mathbb{C}^{n\times L}$. A complex-valued scalar is soft-thresholded as $S_\tau(c) = \max(|c| - \tau, 0)e^{i\angle c}$ for $\tau \geq 0$.

We train the parameters of the decorruption step in each layer (Fig. 1) to minimize reconstruction errors. The training data consists of a collection of reference images reconstructed from densely (or fully) sampled k-space measurements, and we retrospectively under-sample these k-space data to generate subsampled measurements. The first layer requires initial reconstructions (e.g., $A^\dagger y$, where $(\cdot)^\dagger$ denotes pseudoinverse of a matrix) as input. We train the dictionary, transform, and thresholds one layer at a time. Once the parameters for a specific layer are trained, the training image reconstructions are passed through that layer and the resulting images are used for training the subsequent layer. The training for the kth layer minimizes the error between the decorrupted versions of (estimated) patches $\{b_l^{k-1}\}$ (extracted from training image reconstructions), and the corresponding training patches $\{b_l^{\text{train}}\}$ (extracted from reference images) by solving

$$\min_{D^k\in\mathcal{D},W^k,\gamma^k} \sum_{l=1}^{N'} \left\| b_l^{\text{train}} - D^k S_{\gamma^k}(W^k b_l^{k-1}) \right\|_2^2, \quad (2)$$

where the set $\mathcal{D}$ here contains matrices with unit norm columns. Formulation (2) also appears in prior work [14], but with the discontinuous hard-thresholding.

For training, we optimize Problem (2) using an efficient block coordinate descent (BCD) approach. Let B^{train} and B^{k-1} denote matrices with patches b_l^{train} and b_l^{k-1} as columns. With $D^k S_{\Gamma^k}(W^k B^{k-1}) = \sum_{l=1}^{L} d_l^k S_{\gamma_l^k}(g_l^{k^T} B^{k-1})$, where G^k denotes the transpose of W^k and subscript l for lower-case letters denotes the lth column of a matrix (i.e., d_l^k is a column of D^k) or the lth entry of a vector, the proposed BCD training algorithm updates each variable in the triplet $(d_l^k, \gamma_l^k, g_l^k)$, and cycles over all such triplets.

The dictionary column d_l^k is updated by solving the following sub-problem, where the residual $E_l^k \triangleq B^{\text{train}} - \sum_{j\neq l} d_j^k S_{\gamma_j^k}(g_j^{k^T} B^{k-1})$ is fixed based on the most recent estimates of the other variables:

$$\min_{d_l^k} \left\| E_l^k - d_l^k S_{\gamma_l^k}(g_l^{k^T} B^{k-1}) \right\|_F^2 \quad \text{s.t.} \quad \left\| d_l^k \right\|_2 = 1. \quad (3)$$

The dictionary column is updated in closed-form [15] (using sparse multiplications) as follows:

$$\hat{d}_l^k = \frac{E_l^k S_{\gamma_l^k}(g_l^{k^T} B^{k-1})}{\left\| E_l^k S_{\gamma_l^k}(g_l^{k^T} B^{k-1}) \right\|_2} \quad (4)$$

The (non-negative) soft-threshold γ_l^k and the transform row $g_l^{k^T}$ are updated by minimizing the same objective as in (3), and using sub-derivative descent with backtracking line search (for step sizes). This approach ensures monotonic decrease of the cost in (2). The gradients with respect to γ_l^k and $g_l^{k^T}$ are as follows:

$$\frac{\partial \phi}{\partial \gamma_l^k} = -2\text{Re}\Big(\big(S_{\gamma_l^k}(f_l^k) - d_l^{k^T} E_l^k \big) \odot 1_{|f_l^k|>\gamma_l^k} \odot \text{sign}^*\left(f_l^k\right) \Big) \mathbf{1}$$

$$\frac{\partial \phi}{\partial g_l^{k^T}} = \left(2h_l^k - 2i\gamma_l^k \text{Im}(h_l^k) \oslash \left|f_l^k\right| + 2i\gamma_l^k h_l^{k^*} \odot c_l^k\right) \odot 1_{|f_l^k|>\gamma_l^k} \times \left(B^{k-1}\right)^H,$$

where ϕ denotes the objective in (3); $f_l^k \triangleq g_l^{k^T} B^{k-1}$, $h_l^k \triangleq S_{\gamma_l^k}(f_l^k) - d_l^{k^T} E_l^k$, and $c_l^k \triangleq \text{Im}(f_l^k) \odot \text{sign}(f_l^k) \oslash (|f_l^k| \odot |f_l^k|)$; $\mathbf{1}$ denotes a column vector of ones; $1_{|f_l^k|>\gamma_l^k}$ and $\text{sign}(\cdot)$ denote the indicator function (value 0 when condition is violated and 1 otherwise) and complex phase computed element-wise; $\odot$ and $\oslash$ denote element-wise multiplication and element-wise division; $(\cdot)^*$ denotes complex conjugate; $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ denote real and imaginary parts, respectively; and i denotes the imaginary number.

Once the multi-layer algorithm is trained, the learned operations can be applied to reconstruct test images from (undersampled) k-space data. This reconstruction approach repeatedly solves the following least-squares image update problem using the trained models and the measured data y:

$$\min_x \sum_{j=1}^{N} \|P_j x - D^k S_{\gamma^k}\left(W^k P_j x^{k-1}\right)\|_2^2 + \nu\|Ax - y\|_2^2. \quad (5)$$

For example, in the case of single coil Cartesian sampled MRI, assuming all maximally overlapping image patches (including patches that wrap around image boundaries) are used, the least squares solutions can be computed cheaply using FFTs [7]. The proposed simple multi-layer procedure is somewhat reminiscent of layers in a deep neural network, but the connections here are guided by both MR physics in the system matrix A and data y as well as the trained transforms and dictionaries, and trained soft-thresholding operators. Our framework differs from recent works [16, 17] that learn a denoised reference image for use with MRI reconstruction.

3. RESULTS AND DISCUSSION

3.1. Experimental Setup

We work with a multi-slice dataset with 32 slices, each 512×512 provided by Michael Lustig, UC Berkeley, and apply

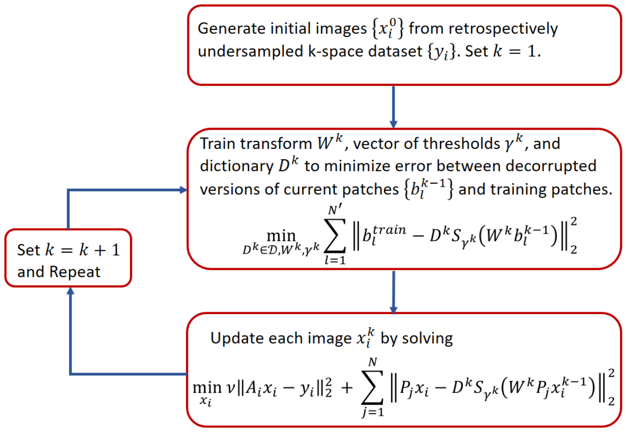

Fig. 1. Schematic for *training* an image reconstruction algorithm with $k_{\max}$ layers. Fully-sampled k-space data of reference or training MR images are assumed given. The k-space data are then retrospectively undersampled to generate a dataset $\{y_i\}_{i=1}^{M}$ of k-space measurements. The image reconstruction algorithm involves multiple iterations or layers, with each layer involving a decorruption step and an image update step. We train the transform, vector of thresholds for transform rows, and dictionary in each decorruption step to minimize the error between the decorrupted and reference or training patches. The image update step updates the image using information from the measured k-space.

the optimal sampling strategy in compressed sensing, called multi-level sampling [18, 19]. We trained the reconstruction algorithm in Fig. 1 with 20 layers based on four slices. We simulated 10%, 20%, and 30% undersampled (single coil) k-space imaging data during training. Fig. 3 shows the tenfold multi-level sampling pattern (feasible for 3D imaging). We used the following settings during training: 64×256 dictionaries and 256×64 transforms; 30 BCD iterations to minimize (2) in each layer; 4 inner descent iterations for γ_l^k and g_l^k updates during each BCD iteration; and 50,000 randomly selected image patches to train the dictionary, transform, and thresholds in each layer.

3.2. Results and Discussion

Fig. 2(a) shows that the average peak-signal-to-noise ratio (PSNR) of the images reconstructed while training the algorithm improves as the number of layers (iterations) is increased. The first layer itself markedly improved the reconstruction quality compared to the initial zero-filling image estimates, and subsequent layers further substantially improved image quality for the training data. Fig. 2 shows the trained dictionary, transform, and soft-thresholds for the first layer, with the dictionary columns and transform rows shown as 8×8 patches. The soft-thresholds for different rows of the transform are quite different, as are the atoms or rows of the transform and the dictionary columns that display many interesting geometric features.

Table 1. PSNR (dB) values (computed using complex images) corresponding to zero-filling and Sparse MRI reconstructions, and for the proposed method, shown for different sampling ratios (SR).

Test image	SR	Zero-filled	Sparse MRI	Trained
#1	30 %	26.0	27.3	**29.0**
	20 %	24.9	26.3	**28.3**
	10 %	23.7	25.0	**27.2**
#2	30 %	25.2	26.4	**27.9**
	20 %	24.2	25.4	**27.1**
	10 %	23.1	24.3	**26.3**

We applied the trained algorithm to reconstruct different test slices. Fig. 3 shows the reconstruction results and reconstruction errors for the proposed method and for the conventional Sparse MRI technique [1, 20] with wavelets and total variation sparsity (using built-in parameters, which worked well). The trained algorithm achieves better reconstruction accuracy and image sharpness than Sparse MRI for highly undersampled (10x) measurements. Table 1 shows the PSNR values for two test slices using different methods at multiple undersampling factors. The proposed trained method achieves the best PSNR values in all cases.

4. CONCLUSIONS

We presented a new image reconstruction framework involving multiple iterations or layers, with trained models in each layer. The transforms, thresholds (for nonlinearities), and dictionaries in each layer were learned using an efficient block coordinate descent algorithm to minimize reconstruction errors for training data. The learned algorithm is then applied to reconstruct test data with finite number of layers (or iterations). Our approach works well for MR image reconstruction from highly undersampled k-space measurements. More detailed experiments and exploration of other useful properties (e.g., orthogonality, etc.) for the dictionary (and transform) in the proposed algorithm will be pursued in future work. We also plan to explore other applications such as image reconstruction in low-dose X-ray computed tomography in future work.

5. ACKNOWLEDGMENTS

This work was supported in part by the following grants: ONR grant N00014-15-1-2141, DARPA Young Faculty Award D14AP00086, ARO MURI grants W911NF-11-1-0391 and 2015-05174-05, NIH grants R01 EB 023618 and P01 CA 059827, and a UM-SJTU seed grant.

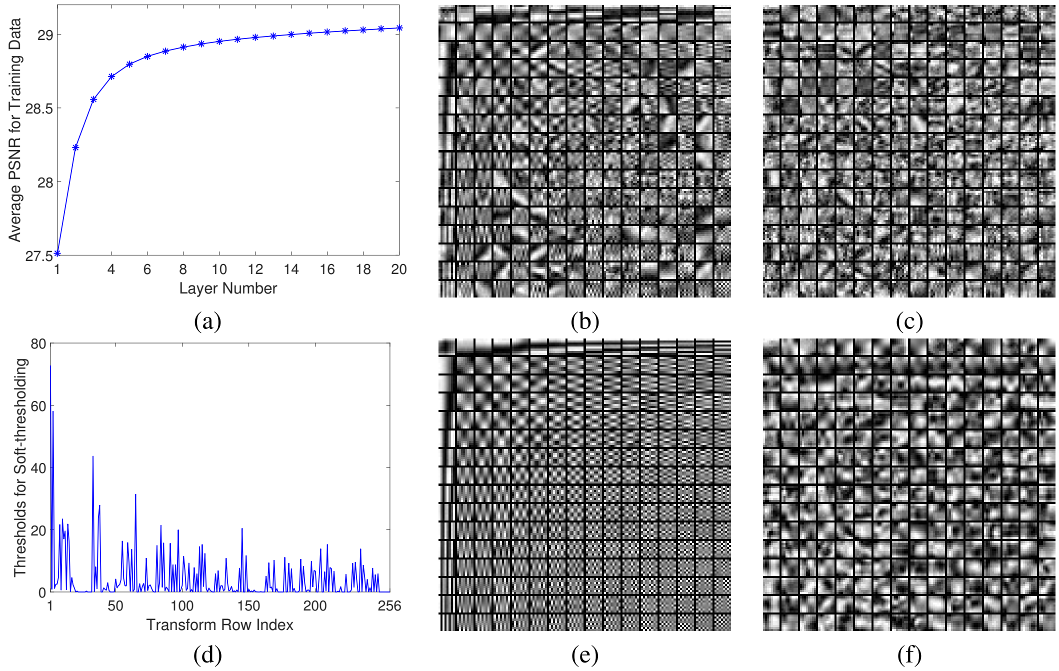

Fig. 2. Results from training the reconstruction algorithm with 20 layers at 10 fold undersampling of k-space. (a) Average PSNR (computed between magnitudes of reconstructed and reference images) in decibels (dB) for the training images for each iteration/layer (average PSNR of the initial zero-filling reconstructions is 26.0 dB). Trained models in the first layer: (b) real and (c) imaginary parts of the 64×256 dictionary; (d) thresholds for soft-thresholding for transform rows; and (e) real and (f) imaginary parts of the 256×64 transform. Dictionary and transform atoms are shown as 8×8 patches.

6. REFERENCES

[1] M. Lustig, D.L. Donoho, and J.M. Pauly, “Sparse MRI: The application of compressed sensing for rapid MR imaging,” *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[2] M. Lustig, J. M. Santos, D. L. Donoho, and J. M. Pauly, “k-t SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity,” in *Proc. ISMRM*, 2006, p. 2420.

[3] J. Trzasko and A. Manduca, “Highly undersampled magnetic resonance image reconstruction via homotopic l_0-minimization,” *IEEE Trans. Med. Imaging*, vol. 28, no. 1, pp. 106–121, 2009.

[4] Y. Kim, M. S. Nadar, and A. Bilgin, “Wavelet-based compressed sensing using gaussian scale mixtures,” in *Proc. ISMRM*, 2010, p. 4856.

[5] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, “k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI,” *Magnetic Resonance in Medicine*, vol. 61, no. 1, pp. 103–116, 2009.

[6] J. P. Haldar and Z. P. Liang, “Spatiotemporal imaging with partially separable functions: A matrix recovery approach,” in *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2010, pp. 716–719.

[7] S. Ravishankar and Y. Bresler, “MR image reconstruction from highly undersampled k-space data by dictionary learning,” *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1028–1041, 2011.

[8] S. Ravishankar and Y. Bresler, “Data-driven learning of a union of sparsifying transforms model for blind compressed sensing,” *IEEE Transactions on Computational Imaging*, vol. 2, no. 3, pp. 294–309, 2016.

[9] X. Zheng, S. Ravishankar, Y. Long, and J. A. Fessier, “Union of learned sparsifying transforms based low-dose 3D CT image reconstruction,” in *International Conference on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, 2017, pp. 69–72.

[10] X. Zheng, I. Y. Chun, Z. Li, Y. Long, and J. A. Fessler, “Sparse-view X-ray CT reconstruction using ℓ_1 prior with learned transform,” Oct. 2017, submitted. Online: `http://arxiv.org/abs/1711.00905`.

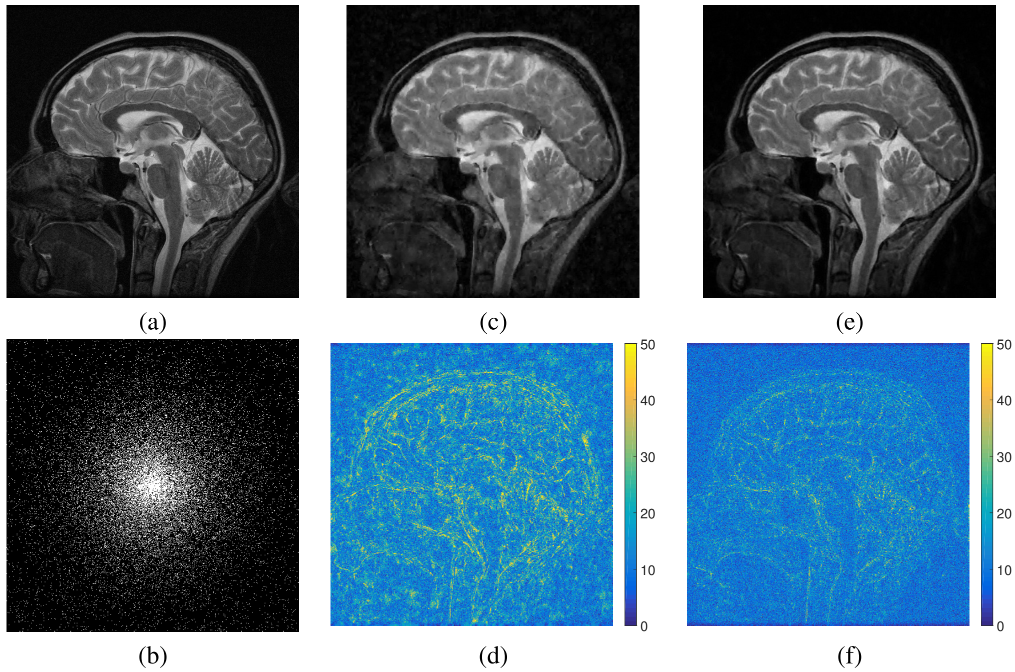

Fig. 3. Results with the trained MRI reconstruction algorithm (only magnitudes of complex-valued images are displayed): (a) Reference test image (peak intensity of 316.2); (b) k-space sampling mask with 10 fold undersampling [18]; (c) reconstruction with Sparse MRI [1] using Wavelets and Total variation sparsity (PSNR = 25.0 dB); (d) reconstruction error map for Sparse MRI; (e) reconstruction with trained algorithm (PSNR = 27.2 dB); and (f) reconstruction error map for the proposed trained method. The error map is obtained as the magnitude of the complex difference between the reconstructed and reference images. PSNR here is computed from the norm of the error maps.

[11] S. Ravishankar and Y. Bresler, "Efficient blind compressed sensing using sparsifying transforms with convergence guarantees and application to magnetic resonance imaging," *SIAM Journal on Imaging Sciences*, vol. 8, no. 4, pp. 2519–2557, 2015.

[12] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on signal processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[13] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. ICML*, Haifa, Israel, Jun. 2010, pp. 399–406.

[14] R. Rubinstein and M. Elad, "Dictionary learning for analysis-synthesis thresholding," *IEEE Transactions on Signal Processing*, vol. 62, no. 22, pp. 5962–5972, 2014.

[15] S. Ravishankar, B. E. Moore, R. R. Nadakuditi, and J. A. Fessler, "Low-rank and adaptive sparse signal (LASSI) models for highly accelerated dynamic imaging," *IEEE Transactions on Medical Imaging*, vol. 36, no. 5, pp. 1116–1128, 2017.

[16] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng, and D. Liang, "Accelerating magnetic resonance imaging via deep learning," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 2016, pp. 514–517.

[17] S. Wang, Z. Su, L. Ying, X. Peng, and D. Liang, "Exploiting deep convolutional neural network for fast magnetic resonance imaging," in *Proc. ISMRM*, 2016, p. 1778.

[18] B. Roman, A. Hansen, and B. Adcock, "On asymptotic structure in compressed sensing," *arXiv preprint arXiv:1406.4178*, 2014.

[19] I. Y. Chun and B. Adcock, "Compressed sensing and parallel acquisition," *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 1–23, May 2017.

[20] M. Lustig, "Michael Lustig home page," `http://www.eecs.berkeley.edu/~mlustig/Software.html`, 2014, [Online; accessed October, 2014].