OFF-THE-GRID MODEL BASED DEEP LEARNING (O-MODL)

Aniket Pramanik, Hemant Aggarwal, Mathews Jacob

University of Iowa, Iowa, USA

ABSTRACT

We introduce a model based off-the grid image reconstruction algorithm using deep learned priors. The main difference of the proposed scheme with current deep learning strategies is the learning of non-linear annihilation relations in Fourier space. We rely on a model based framework, which allows us to use a significantly smaller deep network, compared to direct approaches that also learn how to invert the forward model. Preliminary comparisons against image domain MoDL approach demonstrates the potential of the off-the-grid formulation. The main benefit of the proposed scheme compared to structured low-rank methods is the quite significant reduction in computational complexity.

Index Terms— off-the-grid, CNN, MRI

1. INTRODUCTION

The recovery of images from undersampled Fourier measurements is a classic problem in MRI and several other modalities. The popular approach tis to constrain the reconstructions using compactness priors including sparsity. Several researchers have recently introduced off-the-grid continuous domain priors that are robust to discretization errors [1, 2], which provide significantly improved image quality in a range of applications. However, the main challenge is the significant increase in computational complexity.

Recently, several researchers have introduced deep learning methods as fast and efficient alternatives to compressed sensing algorithms. Current approaches can be categorized into direct and model based strategies. The direct approaches directly estimate the images from the undersampled measurements or their transforms/features [3, 4]. These methods learn to invert the forward operator over the space/manifold of images. While this approach is more popular, a challenge with these schemes is the need to learn the inverse, which often requires large models (e.g. UNET); this often translates to the need for extensive training data. Model based strategies instead formulate the recovery as a penalized optimization problem using a forward model that mimics the acquisition scheme [5]; the penalty is designed to encourage the reconstructed image to be close to the space/manifold of images. Since they rely on an explicit forward model, they often only

require a much smaller network than direct methods, which significantly reduces the training data demand. Recent studies also show the benefit in sharing the network parameters across iterations, end-to-end training, as well as using optimization blocks within the network [5].

The main focus of this work is to extend deep learning to the off-the-grid setting, similar to the structured low-rank setting in [1, 2]. Specifically, we consider an iterative reweighted least-squares (IRLS) formulation of our off-the-grid algorithm, termed as GIRAF. We show that the structure of GI-RAF, which alternates between data consistency enforcement and denoising of the Fourier coefficients of the partial derivatives by projecting it to the constraint set, has extensive similarities to the MoDL framework. The main difference with MoDL is that the learning is performed in the Fourier domain, and the denoising network is linear; GIRAF relies on a residual convolution and deconvolution (flipped convolution) denoising filter-bank with shared filters, which is learned from the available k-space measurements in an iterative fashion. The residual filterbank projects the Fourier coefficients to the signal subspace, thus denoising them. To significantly reduce the computational complexity of these off-the-grid methods, we propose to use a deep non-linear convolutional neural network as the denoiser for the Fourier coefficients of the partial derivatives. Similar to GIRAF, we propose to use convolutional and deconvolutional blocks with shared parameters. The parameters of the deep network are learned in an end-to-end fashion as in MoDL. This work has connections with k-space deep learning approach [6], which follows a direct approach of estimating images from measurements, bypassing a forward model. By contrast, we follow a model based strategy, which has the benefits discussed aboe.

We determine the utility of the proposed scheme in the single coil image recovery setting. Thanks to the MoDL formulation, few training datasets were sufficient to reliably learn the parameters of the model. The comparisons of the proposed scheme against MoDL shows the potential of the proposed scheme. Our future work will focus on the extension of the framework to the multi-coil setting and more elaborate comparisons with state of the art algorithms.

2. PROBLEM SETTING AND BACKGROUND

We consider the recovery of the Fourier coefficients of a function $f(\mathbf{r})$ from its noisy undersampled Fourier measurements

$\mathbf{b} = \mathbf{A}\hat{f} + \boldsymbol{\eta}$, where $\hat{f}$ denotes the Fourier transform of $f(\mathbf{r})$, and $\mathbf{A}$ is an undersampling matrix, while η is additive noise.

2.1. Piecewise smooth images & annihilation relations

Recent work off-the-grid methods [1, 2] model the signal f to be piecewise smooth, when the partial derivatives of f vanish everywhere except on an edge set. Similar to [1], we model the edge set as the zero-set of a bandlimited function μ. With these assumptions, the first order partial derivatives of the signal satisfies $\nabla f.\mu = 0$, which translates to $\widehat{\nabla f} * h = 0$, where $h \overset{\mathcal{F}}{\leftrightarrow} \mu$ is the bandlimited Fourier transform of μ. The theory in [1] shows that if the assumed size of the filter h is greater than the true bandwidth, there exist several linearly independent finite impulse response (FIR) filters that satisfy $\widehat{\nabla f} * h_i = 0; i = 1, .., N$. The above relation can be expressed in a matrix form as

$$\underbrace{\left[\mathcal{G}(\mathbf{h}_1) \mid \mathcal{G}(\mathbf{h}_2) \mid \ldots \mid \mathcal{G}(\mathbf{h}_N) \right]}_{\mathcal{T}(\mathbf{H})} \widehat{\nabla f} = 0. \quad (1)$$

Here, $\mathcal{G}(\mathbf{h}_i)$ is a block Toeplitz matrix that corresponds to 2-D convolution. i.e, $\mathcal{G}(\mathbf{h}_i)\widehat{\nabla f} = h_i * \widehat{\nabla f}$. Using commutativity of convolutions, the above relation can also be expressed as

$$\underbrace{\left[\begin{array}{c} \mathcal{G}\left(\widehat{\mathbf{f}}_x\right) \\ \mathcal{G}\left(\widehat{\mathbf{f}}_y\right) \end{array} \right]}_{\mathcal{T}(\widehat{\nabla f})} \mathbf{H} = 0. \quad (2)$$

where $\mathcal{G}(\widehat{\mathbf{f}}_x)$ is the block Toeplitz matrix constructed from $\widehat{\mathbf{f}}_x$.

## 2.2. GIRAF algorithm for off-the-grid image recovery

In practice, the filters $h_i; i = 1, .., N$ are unknown. Two step methods [1, 2] estimate the filters $h_i; i = 1, .., N$ or equivalently $\mathcal{T}(\mathbf{H})$ from fully sampled regions of Fourier space, which is also called as calibration regions, using (2). When the Fourier samples are non-uniformly sampled, the above two step approach is not feasible. Note that (2) implies that the matrix $\mathcal{T}(\mathbf{H})$ is low-rank. The constrained version of GIRAF algorithm poses the recovery as
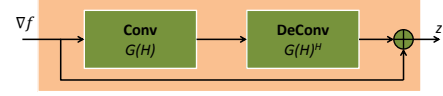
$$\hat{f} = \arg\min_{\hat{f}} \left\| \mathcal{T}(\widehat{\nabla f}) \right\|_* \text{ subject to } \mathbf{A}f = \mathbf{b}, \quad (3)$$

where $\| \cdot \|_*$ is the nuclear norm that encourages the matrix to be low-rank. GIRAF relies on an iterative least squares optimization problem, which majorizes the nuclear norm as $\|\mathbf{T}(\widehat{\nabla f})\|_* \leq \|\mathbf{T}(\widehat{\nabla f})\mathbf{H}\|_F^2$. This algorithm alternates between the estimation of $\mathbf{H}$ as $\mathbf{H} = \left[\mathbf{T}^*(\mathbf{T}(\widehat{\nabla f})) + \epsilon \mathbf{I}\right]^{-1/4}$ and updating f by
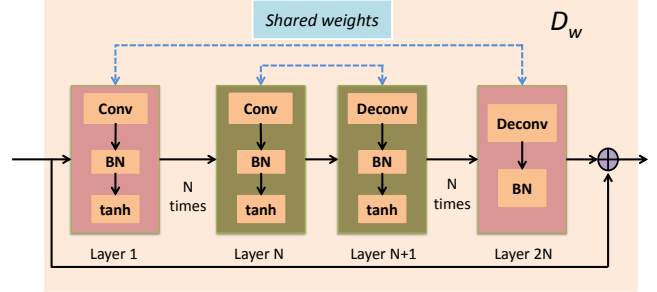
$$\hat{f} = \arg\min_{\hat{f}} \|\mathcal{T}(\mathbf{H}) \widehat{\nabla f}\|_F^2 \text{ subject to } \mathbf{A}f = \mathbf{b}. \quad (4)$$
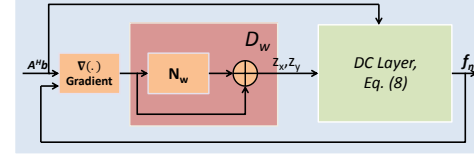
(a) Iterative algorithm

(b) GIRAF residual denoising block

(c) O-MoDL residual denoising block

Fig. 1. Outline of the recursive learning architectures. (a) Proposed iterative algorithm, which alternates between a residual denoiser block and data-consistency block. GIRAF and O-MoDL frameworks differs only in the structure of the residual denoiser blocks, shown in (b) and (c), respectively. (b). GIRAF residual denoising block, where the filters $\mathbf{H}$ are learned using structured low-rank matrix completion. (c) O-MoDL residual denoising block: the linear denoiser in GIRAF is replaced by a deep CNN architecture, which uses convolution and deconvolution blocks. The filter parameters of the corresponding convolution and deconvolution blocks are shared.

A challenge with the alternating minimization strategy is the high computational complexity of the algorithm. Motivated by the fast computation offered by our MoDL framework [5], we introduce a deep learning solution.

3. OFF-THE-GRID MODEL BASED DEEP LEARNING (O-MODL)

We first show that the GIRAF algorithm can be viewed as a recursive network, which has alternating blocks imposing data consistency and denoising of the data.

3.1. Network structure of GIRAF

We consider a penalized version of (4), using an auxiliary variable $\mathbf{z}$:

$$\mathcal{C}\{\mathbf{f}, \mathbf{z}\} = \arg\min_{\mathbf{f}, \mathbf{z}} \beta \|\widehat{\nabla f} - \mathbf{z}\|^2 + \lambda \|\mathcal{T}(\mathbf{H}) \mathbf{z}\|_F^2$$
$$\text{subject to } \mathbf{A}\mathbf{f} = \mathbf{b}. \quad (5)$$

(a) Original (b) MoDL (c) Proposed (d) SENSE

(e) $A^H b$ (f) MoDL (g) Proposed (h) SENSE

(i) Original (j) MoDL (k) Proposed (l) SENSE
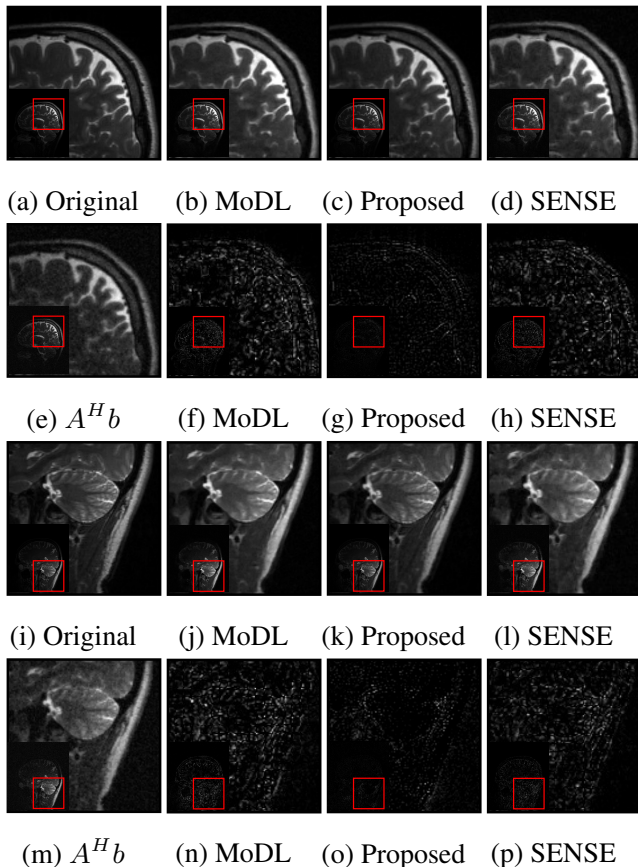
(m) $A^H b$ (n) MoDL (o) Proposed (p) SENSE

Fig. 2. Reconstruction results of two-fold accelerated data. Figures (a)-(h) correspond to Coil 1 images while (i)-(p) correspond to Coil 6 images. (b) MoDL:16.33 db (c) O-MoDL: 18.57 db, (d) SENSE: 12.72 db, (f)-(h) are the error images. (j) MoDL,16.74 db (k) O-MoDL,18.98 db, (l) SENSE, 12.88 db, (n)-(p) are the error images.

Note that the above formulation is equivalent to (4) when $\beta \to \infty$. We use an alternating algorithm to solve for $\mathbf{f}$ and $\mathbf{z}$:

$$\mathbf{f}_{n+1} = \arg\min_{\hat{f}} \|\widehat{\nabla f} - \mathbf{z}_n\|^2 \text{ subject to } \mathbf{Af} = \mathbf{b} \quad (6)$$

$$\mathbf{z}_{n+1} = \arg\min_{\mathbf{z}} \|\nabla \widehat{f_{n+1}} - \mathbf{z}\|^2 + \lambda \|\mathcal{T}(\mathbf{H})\,\mathbf{z}\|_F^2 \quad (7)$$

Problem (6) is analytically solved, when $\mathbf{A}$ is a sampling operator. Solving (7), we get $\mathbf{z} = \left[\mathbf{I} + \frac{\lambda}{\beta} \mathcal{T}(\mathbf{H})^H \mathcal{T}(\mathbf{H})\right]^{-1} \widehat{\nabla f}$. Using matrix inversion lemma and assuming $\lambda << \beta$, we approximate the solution as:

$$\mathbf{z} \approx \left[\mathbf{I} - \frac{\lambda}{\beta} \mathcal{T}(\mathbf{H})^H \mathcal{T}(\mathbf{H})\right] \widehat{\nabla f}. \quad (8)$$

Note from (2) that filters $\mathbf{H}$ are surrogates for the null space of $\mathcal{T}(\widehat{\nabla f})$. Hence, $\mathcal{T}(\mathbf{H})^H \mathcal{T}(\mathbf{H})\,\mathbf{z}$ can be viewed as

the projection of $\mathbf{z}$ onto the null-space of ; (8) can be viewed as a residual block, which removes the null-space components of $\widehat{\nabla f}$. The alternating algorithm specified by (8) and (6) can be unrolled to a deep architecture as in [5].

3.2. Deep learning in k-space

Motivated by the success of MoDL [5], we introduce a novel deep learning solution to reduce the computational complexity of GIRAF. Specifically, we replace the denoiser in GIRAF, specified by (8), by a CNN, whose parameters are learned from exemplar data. The structure of O-MoDL is similar to GIRAF, involving two alternating blocks, $\mathcal{D}_w = \mathcal{I} - \mathcal{N}_w$ (denoiser) and DC (data consistency). We rely on an n-layer O-MoDL $\mathcal{N}_w$ block consists of 'n' convolution layers followed by 'n' deconvolution (convolution with flipped filters) layers, followed by batch normalization (BN) and non-linear activation function. We chose the non-linear activation function as a hyperbolic tangent (tanh) activation function. Each convolution layer only retains the valid convolutions to eliminate boundary effects. The use of deconvolution blocks ensures the smaller image obtained after convolution grows back to its original size. The DC block refers to (6). The main difference of this framework with GIRAF is the depth of the denoiser network (see Fig. 2.1) and the presence of non-linearities and batch normalization steps within the network. The key difference of the proposed scheme with MoDL is that the learning is performed in k-space, unlike most deep learning image recovery strategies. The denoiser network uses convolution and deconvolution layers similar to GIRAF, which share weights. This is a another distinction with MoDL, which uses a series of convolutional blocks with no weight sharing. Similar to MoDL, we also share the same network across iterations.

4. RESULTS

Multi-channel brain MRI data was collected from five subjects at University of Iowa Hospitals using 3D T2 CUBE sequence with Cartesian readouts using a 12-channel head coil. The data from four subjects was used for training, while the data from the fifth subject was used for testing. We retrospectively undersampled the phase encodes to train and test the framework; we note that this approach is completely consistent with a future prospective acquisition, where a subset of phase encodes can be pre-selected and acquired. All the experiments were performed with variable-density Cartesian random sampling mask with different undersampling factors. We chose 10 most informative slices from each coil, yielding 120 from each subject to form a training dataset of 480 slices.

We trained a network with an $\mathcal{N}_w$ block consisting of five convolution and five deconvolution layers as described in section 3.2. Each layer consisted of 64 complex 3 x 3 filters. The iterative model was unrolled with five iterations, which was trained to minimize the mean square error between the output
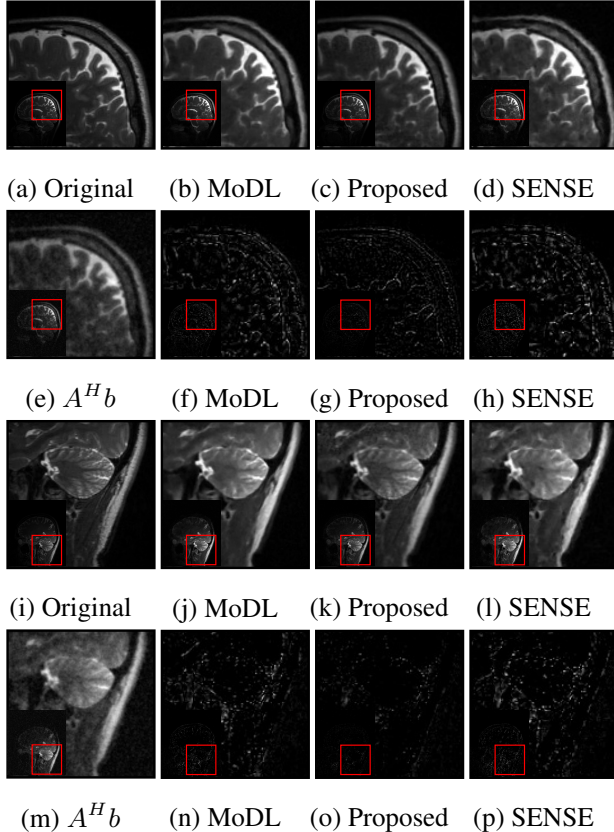
(a) Original   (b) MoDL   (c) Proposed   (d) SENSE

(e) $A^H b$ (f) MoDL   (g) Proposed   (h) SENSE

(i) Original   (j) MoDL   (k) Proposed   (l) SENSE

(m) $A^H b$ (n) MoDL   (o) Proposed   (p) SENSE

Fig. 3. Reconstruction results of four-fold accelerated data. Figures(a)-(h) correspond to Coil 1 images while (i)-(p) correspond to Coil 6 images. (b) MoDL:12.58 db (c) O-MoDL:13.84 db, (d) SENSE: 11.52 db, (f)-(h) are the error images. (j) MoDL,12.21 db (k) O-MoDL,13.54 db, (l) SENSE, 11.87 db, (n)-(p) are the error images.

and the non-undersampled images. The training took 7 hours on an NVIDIA Tesla p100 GPU on 480 slices. By contrast, the run time for reconstucting a single image is only 0.5 seconds. We compare the performance of O-MoDL against our image domain deep learning method (MoDL) and SENSE reconstructions for two-fold and four-fold accelerated data. For fair comparisons, we train both the networks with same number of iterations and trainable parameters. The SNR improvement in each case can be observed in table 1. Fig. 3 shows the reconstruction from two-fold accelerated noisy data. Similarly, Fig. 4 shows reconstruction from four-fold accelerated data. The reconstructions show that the proposed scheme can provide improved reconstructions compared to the MoDL scheme, which was trained exactly in the same fashion.

5. CONCLUSION

We proposed an off-the-grid deep learning architecture for model based MR image reconstruction. Unlike many of the current CNN architectures, the algorithm works in the Fourier domain. The algorithm is a non-linear extension of recent

2-fold/4-fold accelerated data			
Noise (σ)	MoDL	**Proposed**	SENSE
10	16.42/12.46	**18.30/13.71**	11.7/11.44
11	16.35/12.22	**18.29/13.68**	11.06/10.54
12	15.92/11.91	**18.27/13.68**	10.66/9.89
13	15.77/11.51	**18.25/13.66**	10.48/9.76

Table 1. Quantitative comparison of MoDL, Proposed, and SENSE reconstructions. The SNR is reported in dB and the two figures correspond to two fold and four fold acceleration.

structured low-rank off-the-grid methods, which rely on annihilation relations in the Fourier domain resulting from continuous domain image properties. Unlike the linear annihilation relations that are self learned in structured low-rank setting, the proposed framework learns non-linear annihilation relations in the Fourier domain from exemplar data; the non-linearities facilitate the the generalization of the annihilation properties to images unseen by the training algorithm, eliminating the need for self-learning the weights. The main benefit of the proposed scheme is the quite significant reduction in run time, compared to structured low-rank algorithms. The preliminary experimental comparisons demonstrate the improvements offered by the proposed scheme over image domain MoDL framework.

6. REFERENCES

[1] Greg Ongie and Mathews Jacob, "Off-the-Grid Recovery of Piecewise Constant Images from Few Fourier Samples," *SIAM J. Imag. Sci.*, vol. 9, no. 3, pp. 1004—-1041, 2016.

[2] Justin P Haldar, "Low-rank modeling of local k-space neighborhoods (loraks) for constrained mri," *IEEE trans. med. imag.*, vol. 33, no. 3, pp. 668–681, 2014.

[3] Jong Chul Ye, Yoseob Han, and Eunju Cha, "Deep convolutional framelets: A general deep learning framework for inverse problems," *SIAM J. Imag. Sci.*, vol. 11, no. 2, pp. 991–1048, 2018.

[4] Kyong Hwan Jin, Michael T. McCann, Emmanuel Froustey, and Michael Unser, "Deep Convolutional Neural Network for Inverse Problems in Imaging," vol. 29, pp. 4509–4522, 2017.

[5] Hemant Kumar Aggarwal, Merry P Mani, and Mathews Jacob, "Modl: Model based deep learning architecture for inverse problems," *IEEE trans. med. imag*, vol. 10.1109/TMI.2018.2865356., 2018,.

[6] Yoseob Han and Jong Chul Ye, "k-space deep learning for accelerated mri," *arXiv preprint arXiv:1805.03779*, 2018.