Model-Based Image Reconstruction using Deep Learned Priors (MoDL)

Hemant K. Aggarwal, Merry P. Mani, and Mathews Jacob

The University of Iowa

April 6, 2018

Inverse Problems & Classical Solutions

Model-Based Problem Formulation

Given $y = Ax + n, \quad A \in \mathbb{C}^{M\times N}, \quad x \in \mathbb{C}^N$

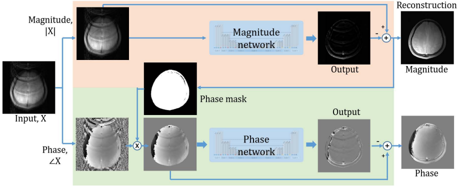
$$\hat{x} = \arg\min_x \|y - Ax\|_2^2 + \lambda\mathcal{R}(x)$$

$\mathcal{R}(x)$: regularization priors

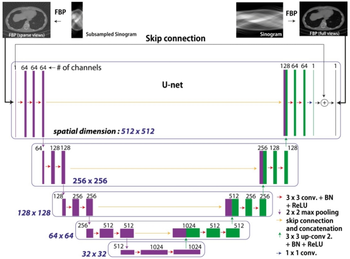
- Total Variation
- Wavelet-based sparsity
- Plug and Play denoisers
  - BM3D, non-local means

[1] Venkatakrishnan et al. Plug-and-Play priors for model based reconstruction GlobalSIP, 2013

Deep Artifact Learning[1]



U-NET architecture[2]



Joint Learning of Image manifold & Inverse: Challenges

- Large network: lots of training data
- Sensitive to acquisition setting: image matrix, undersampling pattern
  - Need several trained networks

[1] Lee et al. Deep artifact learning for compressed sensing and parallel MRI arXiv, 1703.01120

[2] Jin et al. Deep Convolutional Neural Network for Inverse Problems in Imaging MRI IEEE TIP, 2017

MoDL: model based recovery with DL priors

Model-Based Problem Formulation

$$\mathbf{x} = \arg\min_{\mathbf{x}} \underbrace{\|\mathbf{A}\,\mathbf{x} - \mathbf{b}\|_2^2}_{\text{data consistency}} + \lambda \underbrace{\|\mathcal{N}_{\mathbf{w}}(\mathbf{x})\|^2}_{\text{regularization}}$$

- $N_{\mathbf{w}}$: Predictor of noise and alias patterns in $\mathbf{x}$
- Determine $\mathbf{x}$ that is data-consistent and minimize aliasing

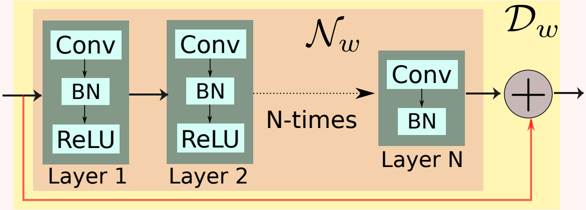
Denoising using noise predictor $\mathcal{N}_{\mathbf{w}}$

Model-Based Problem Formulation

$$\mathbf{x} = \arg\min_{\mathbf{x}} \underbrace{\|\mathbf{A}\,\mathbf{x} - \mathbf{b}\|_2^2}_{\text{data consistency}} + \lambda \underbrace{\|\mathcal{N}_{\mathbf{w}}(\mathbf{x})\|^2}_{\text{regularization}}$$

Denoiser

$$\mathcal{D}_{\mathbf{w}}(\mathbf{x}) = (\mathcal{I} - \mathcal{N}_{\mathrm{w}})\,(\mathbf{x}) = \mathbf{x} - \mathcal{N}_{\mathbf{w}}(\mathbf{x}).$$



Alternating minimization

Problem Formulation

$$\mathbf{x} = \arg\min_{\mathbf{x}} \|\mathbf{A}\,\mathbf{x} - \mathbf{b}\|_2^2 \; + \lambda \; \|\mathbf{x} - \mathcal{D}_{\mathbf{w}}(\mathbf{x})\|_2^2$$

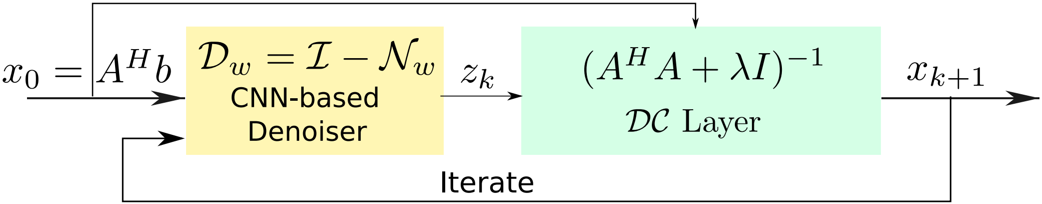
Algorithm

$$\begin{aligned} \mathbf{z}_k &= \mathcal{D}_{\mathbf{w}}\left(\mathbf{x}_k\right) \\ \mathbf{x}_{k+1} &= \left(\mathbf{A}^H\mathbf{A} + \lambda\,\mathbf{I}\right)^{-1}\left(\mathbf{A}^H\mathbf{b} + \lambda\,\mathbf{z}_k\right) \end{aligned}$$

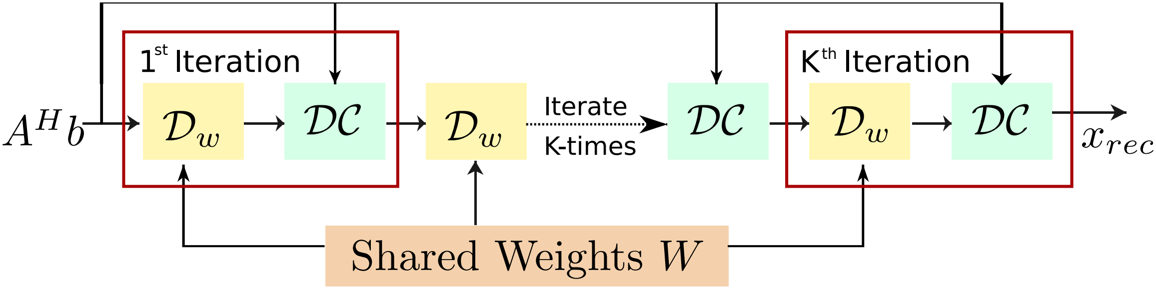
Recursive MoDL Architecture

$$\mathbf{z}_k = \mathcal{D}_{\mathbf{w}}\left(\mathbf{x}_k\right)$$

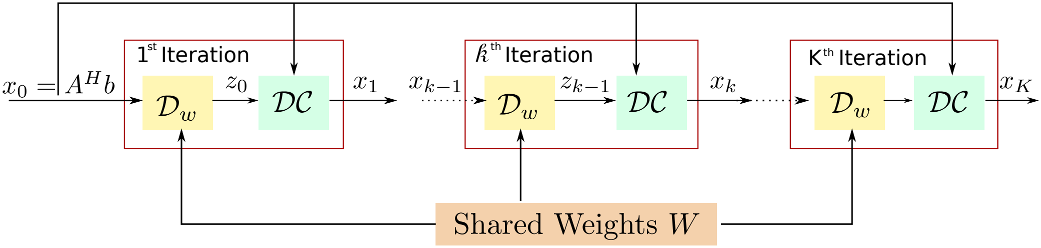
$$\mathbf{x}_{k+1} = \left(\mathbf{A}^H\mathbf{A} + \lambda\,\mathbf{I}\right)^{-1}\left(\mathbf{A}^H\mathbf{b} + \lambda\,\mathbf{z}_k\right)$$



Training: unroll the recursive network

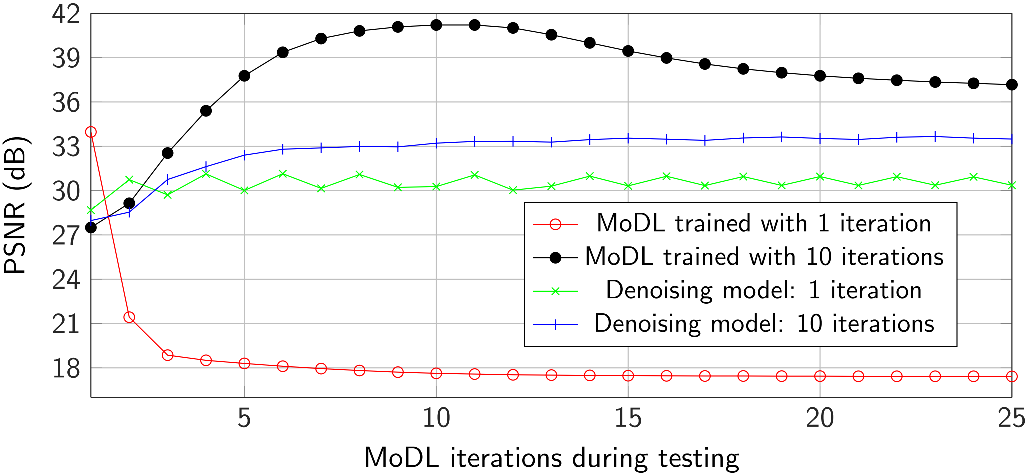


Shared weights: chain rule for gradients



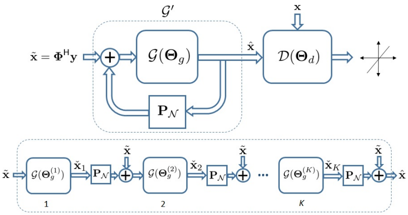
$$\partial_w \mathcal{C} = \sum_{k=1}^{K} \partial_{z_k} \mathcal{C} \ \partial_w z_k$$

Joint training is better than pre-trained denoisers

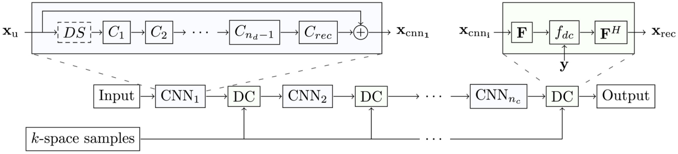


[1] Chang et al. One N/w to Solve Them All: Solving Linear Inverse Prob. using Deep Projection Models, ICCV, 2017

Differences with current iterative approaches



Recurrent GAN[3]



Cascade Network[4]

Challenges

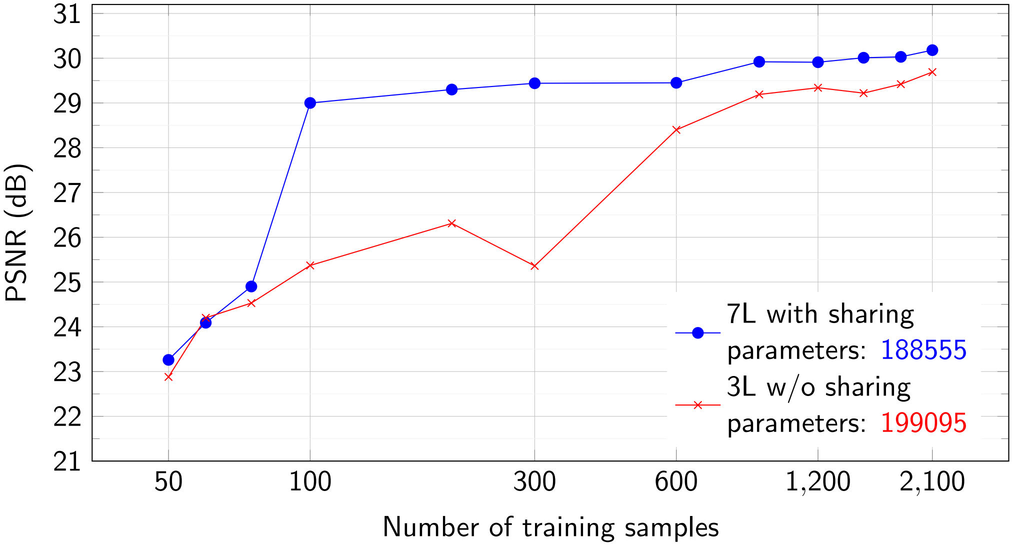
- Different networks at each iteration: not consistent with model based framework
  - Large capacity: Require significantly more training data
- More training data available: increase complexity of networks at each iteration

[3] Mardani et al. Recurrent GAN for Proximal Learning and Automated Compressive Image Recovery CVPR, 2018

[4] Schlemper et al. A Deep Cascade of CNN for Dynamic MR Image ReconstructionTMI, 2018

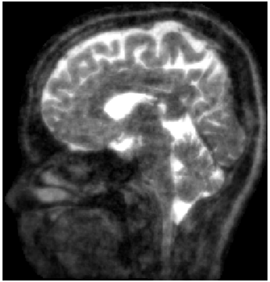
[5] Hammernik et al. Learning a Variational Network for Reconstruction of Accelerated MRI Data MRM, 2017

MoDL: Can be trained with less training data



MoDL: Weight sharing vs without weight sharing

Original slice

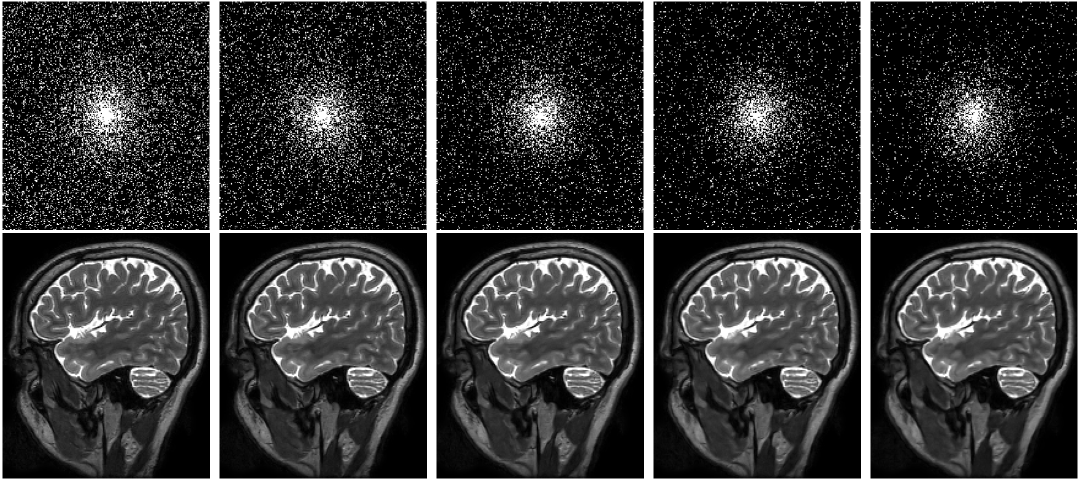


$A^H B$ at 6x, 24.97 dB

w/o sharing, 32.93 dB

with sharing, 38.67 dB

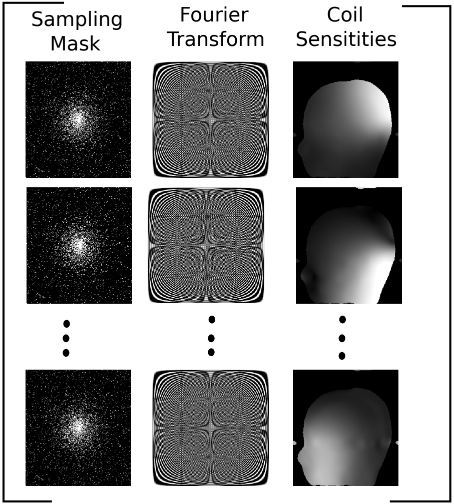
MoDL: Insensitivity to acquisition conditions



6-Fold, 39.43 dB | 8-Fold, 38.47 dB | 10-Fold, 37.75 dB | 12-Fold, 36.42 dB | 14-Fold, 35.87 dB

Parallel MRI: $(A^H A + \lambda I)$ not analytically invertible

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_M \end{bmatrix} = \begin{bmatrix} S\ FC_1 \\ S\ FC_2 \\ \vdots \\ S\ FC_M \end{bmatrix} =$$



Current approaches: Gradient descent (ISTA)

Alternating minimization

$$\mathbf{z}_k = \mathcal{D}_{\mathbf{w}}\left(\mathbf{x}_k\right)$$

$$\mathbf{x}_{k+1} = \arg\min_{x} \|Ax - b\|_2^2 + \lambda\|x - z_k\|_2^2$$

Gradient Descent to minimize DC subproblem[5]

$$\mathbf{x}_{k+1} = x_k - 2(A^H A + \lambda I)x_k + 2A^H b + 2\lambda z_k$$

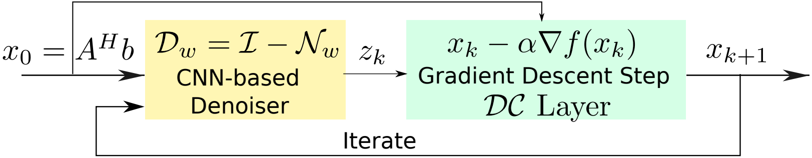
Shrinkage is cheap in CS setting: fast convergence

- Each DC block is in-expensive

[6] Hammernik et al. Learning a Variational Network for Reconstruction of Accelerated MRI Data. MRM, 2017

[6] Wang et al. Deep Networks for Image Super-Resolution with Sparse Prior. ICCV, 2015

Training difficulties with large unrolled network



Large number of iterations

- Large network with unrolling
  - Does not fit on GPUs
- Hammernik et al: does not use weight sharing

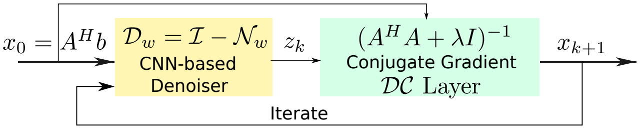
[3] Hammernik et al. Learning a Variational Network for Reconstruction of Accelerated MRI Data MRM, 2017

Solution: Numerical Optimization within DL Network

Sub-Problems

$$z_k = \mathcal{D}_{\mathbf{w}}(\mathbf{x}_k)$$

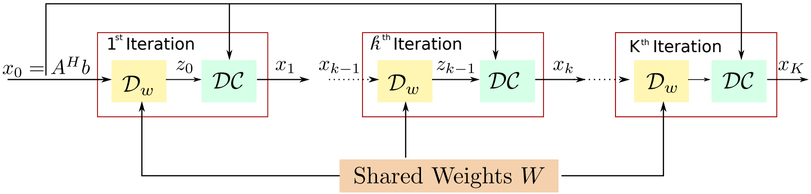
$$\mathbf{x}_{k+1} = \arg\min_x \|Ax - b\|_2^2 + \lambda \|x - z_k\|_2^2$$



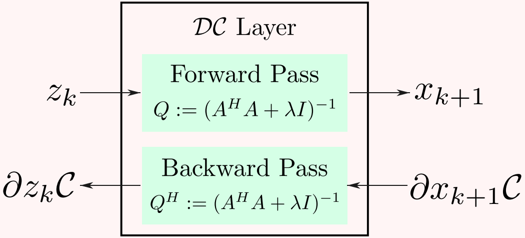
CG within network

- Faster convergence than ISTA
- Need fewer iterations: use larger network on GPU

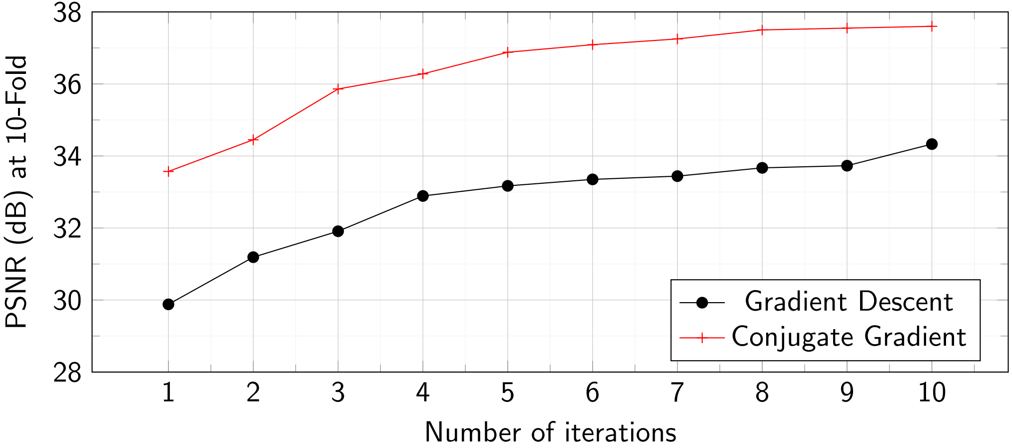
Backpropagation through CG Layer



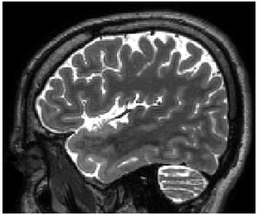
Gradient Computation



CG within network: improved performance

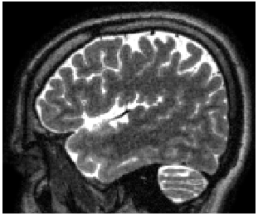


Parallel Imaging with DL (6x)

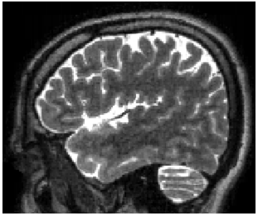


Original Image

$A^H B$, 22.93 dB



Tikhonov, 34.16 dB

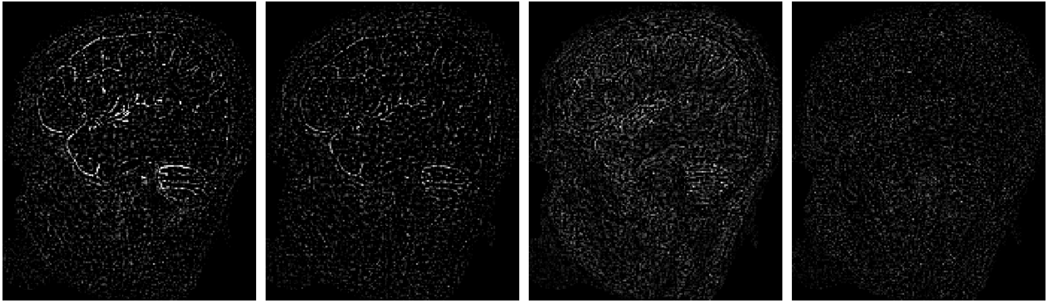


CSTV, 35.20 dB

Grad.Desc., 38.29 dB

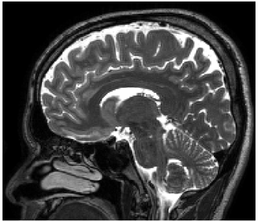
MoDL, 40.33 dB

Parallel Imaging with DL (6x)

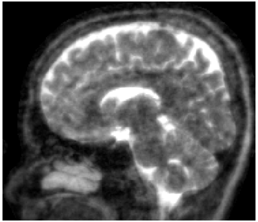


Tikhonov | Compressed Sensing | Gradient Descent | Proposed MoDL

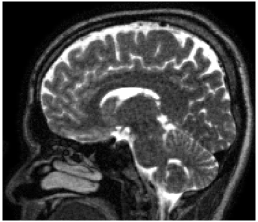
Parallel Imaging with DL (8x)



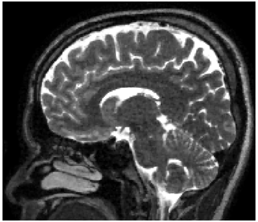
Original Image



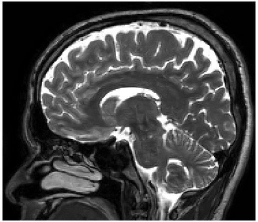
$A^H B$, 23.82 dB



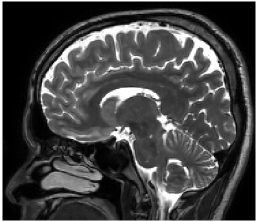
Tikhonov, 32.05 dB



CSTV, 34.43 dB

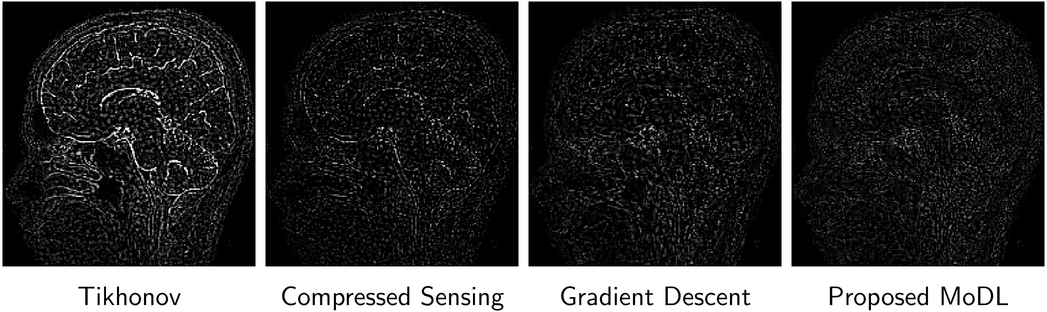


Grad.Desc., 35.22 dB



MoDL, 37.95 dB

Parallel Imaging with DL (8x)



MoDL-SToRM: Use patient-specific image priors

MoDL with SToRM prior

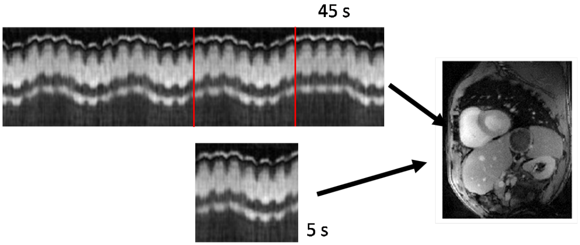
$$\mathbf{x} = \arg\min_{\mathbf{x}} \underbrace{\|\mathbf{A}\,\mathbf{x} - \mathbf{b}\|_2^2}_{\text{data consistency}} + \lambda_1 \underbrace{\|\mathcal{N}_{\mathbf{w}}(\mathbf{x})\|^2}_{\text{regularization}} + \lambda_2 \underbrace{\|tr(\mathbf{X}^{\mathbf{T}}\mathbf{W}\mathbf{X})\|^2}_{\text{SToRM Prior}}$$

[1] Biswas et al. Model-based Free Breathing Cardiac MRI Recon. using Deep Learned & SToRM priors ICASSP, 2018

MoDL-SToRM: Use patient-specific image priors

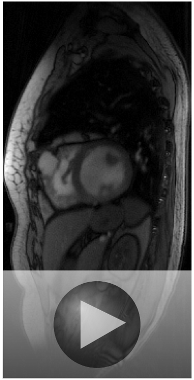
MoDL with SToRM prior

$$\mathbf{x} = \arg\min_{\mathbf{x}} \underbrace{\|\mathbf{A}\,\mathbf{x} - \mathbf{b}\|_2^2}_{\text{data consistency}} + \lambda_1 \underbrace{\|\mathcal{N}_{\mathbf{w}}(\mathbf{x})\|^2}_{\text{regularization}} + \lambda_2 \underbrace{\|tr(\mathbf{X}^{\mathbf{T}}\mathbf{W}\mathbf{X})\|^2}_{\text{SToRM Prior}}$$



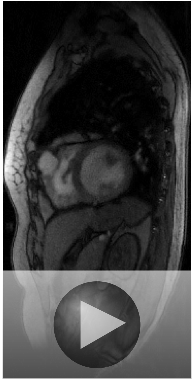
[1] Biswas et al. Model-based Free Breathing Cardiac MRI Recon. using Deep Learned & SToRM priors ICASSP, 2018

Dynamic image recovery using MoDL-SToRM



SToRM recon from 45s of acquisition

SToRM recon from 5s of acquisition



MoDL recon from 5s of acquisition

Conclusion

- Integrating DL priors with model based reconstruction: systematic approach
  - Weight sharing: reduced training data
  - More training data: better performance with iterating shared layers
- Decouple inversion of forward model from image manifold learning
  - Relatively insensitive to acquisition settings
  - Exploit fast algorithms for forward model evaluation
- Numerical optimization blocks within deep network
  - Complex forward model: parallel MRI
  - Faster convergence compared to LISTA
  - Add additional priors: (e.g. subject specific priors)
- Model based deep learning image recovery
  - Fast image recovery
  - Improved image quality

Conclusion

- Integrating DL priors with model based reconstruction: systematic approach
  - Weight sharing: reduced training data
  - More training data: better performance with iterating shared layers
- Decouple inversion of forward model from image manifold learning
  - Relatively insensitive to acquisition settings
  - Exploit fast algorithms for forward model evaluation
- Numerical optimization blocks within deep network
  - Complex forward model: parallel MRI
  - Faster convergence compared to LISTA
  - Add additional priors: (e.g. subject specific priors)
- Model based deep learning image recovery
  - Fast image recovery
  - Improved image quality

Conclusion

- Integrating DL priors with model based reconstruction: systematic approach
  - Weight sharing: reduced training data
  - More training data: better performance with iterating shared layers
- Decouple inversion of forward model from image manifold learning
  - Relatively insensitive to acquisition settings
  - Exploit fast algorithms for forward model evaluation
- Numerical optimization blocks within deep network
  - Complex forward model: parallel MRI
  - Faster convergence compared to LISTA
  - Add additional priors: (e.g. subject specific priors)
- Model based deep learning image recovery
  - Fast image recovery
  - Improved image quality

Conclusion

- Integrating DL priors with model based reconstruction: systematic approach
  - Weight sharing: reduced training data
  - More training data: better performance with iterating shared layers
- Decouple inversion of forward model from image manifold learning
  - Relatively insensitive to acquisition settings
  - Exploit fast algorithms for forward model evaluation
- Numerical optimization blocks within deep network
  - Complex forward model: parallel MRI
  - Faster convergence compared to LISTA
  - Add additional priors: (e.g. subject specific priors)
- Model based deep learning image recovery
  - Fast image recovery
  - Improved image quality

Thank You

Full paper: https://arxiv.org/abs/1712.02862

Computational Biomedical Imaging Group Laboratory

http://research.engineering.uiowa.edu/cbig/