Deep Generative Adversarial Networks for Compressed Sensing (GANCS) Automates MRI

Morteza Mardani[1], *Enhao Gong*[1], *Joseph Y. Cheng*[1,2], *Shreyas Vasanawala*[2], *Lei Xing*[1,3], *John M. Pauly*[1] *

Abstract

Magnetic resonance imaging (MRI) suffers from aliasing artifacts when it is highly undersampled for real-time imaging. Conventional compressed sensing (CS) MRI analytics are not however cognizant of the image diagnostic quality, and substantially trade-off accuracy for speed in real-time imaging. To cope with these challenges we put forth a novel CS framework that permeates benefits from generative adversarial networks (GAN) to modeling a manifold of MR images from historical patients. Extensive evaluations on a large MRI dataset of pediatric patients by expert radiologists corroborate that GANCS retrieves images with improved quality and finer details relative to conventional CS and pixel-wise schemes. It also offers around 100 times faster inference than existing CS-MRI schemes.

1 Context and Motivation

Owing to its superb soft tissue contrast, (real-time) MRI visualization is of paramount importance for diagnostic and therapeutic guidance in next generation imaging platforms. MRI scans are however quite slow, taking several minutes to acquire clinically acceptable images, possibly deteriorated by motion especially for pediatric patients. As a result, the acquisition typically undergoes significant undersampling leading to a seriously ill-posed linear inverse problems for image reconstruction [1, 2]. The scanner collects the k-space data $\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \mathbf{v}$ with $\mathbf{\Phi} \in \mathbb{C}^{M \times N}$ capturing the Fourier transform and the coil maps where $M \ll N$, $\mathbf{v}$ accounts for modeling noise, and $\mathbf{x}$ is the image of interest.

In order to retrieve $\mathbf{x}$ from $\mathbf{y}$, conventional CS uses sparsity regularization in a proper transform domain such as Wavelet (WV) [1]. It however demands running several iterations of non-smooth optimization algorithms for convergence, and need manual hyper-parameter tunning to optimize the performance, which hinders *real-time* MRI. With the abundance of historical scans $\mathcal{X} := \{\mathbf{x}_k\}_{k=1}^K$, and the corresponding (possibly) noisy observations $\mathcal{Y} := \{\mathbf{y}_k\}_{k=1}^K$ we are motivated to pursue a deep learning based approach. Suppose the unknown and complex-valued image $\mathbf{x}$ lies in a *low-dimensional* manifold $\mathcal{M}$. No information is known about the manifold besides the training samples $\mathcal{X}$ and $\mathcal{Y}$. Given a new observation $\mathbf{y}$, the goal is then to recover $\mathbf{x}$.

Previous work. Recent attempts exist that *automate* medical image reconstruction by leveraging historical data; see e.g., [3–8]. By training a neural network to map out aliased images to the gold-standard ones, they gain speed up, but they suffer from blurring artifacts and possible hallucinations. This is mainly due to i) adopting pixel-wise costs that are oblivious to high-frequency texture details, which is crucial for drawing diagnostic decisions; and ii) lack of data consistency. GANs [9–11] are very powerful in modeling manifold of natural images that are perceptually appealing; see e.g., [12–16]. For natural image super-resolution, GANs retrieve realistic images under $4\times$ upscaling [14, 15].

2 Method

The inverse imaging solution amounts to finding a solution at the intersection of data-consistent images and the plausible high-quality image manifold. To effectively learn the manifold from training

*The authors are with the Stanford University, Departments of Electrical Engineering[1], Radiology[2], and Radiation Oncology[3].

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

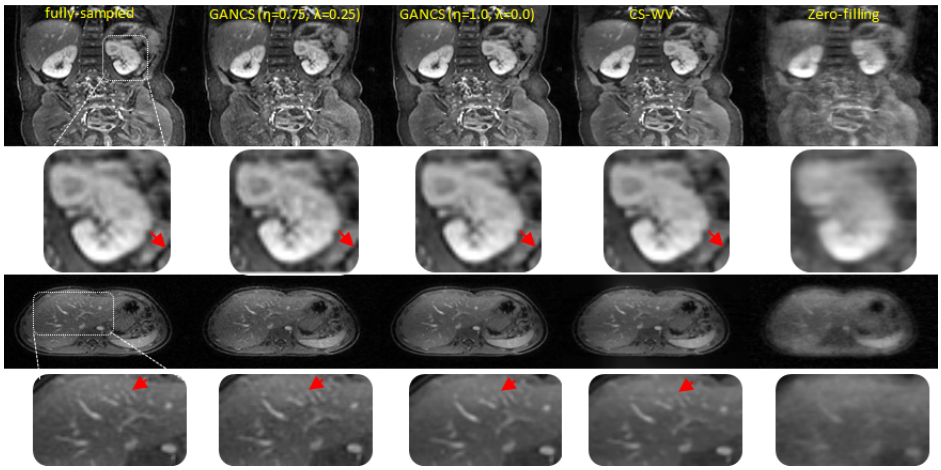

Figure 1: Representative coronal (1st row) and axial (3rd row) images for a test patient retrieved by fully-sampled (1st), GANCS ($\eta = 0.75, \lambda = 0.25$) (2nd), GANCS ($\eta = 1.0, \lambda = 0.0$) (3th), CS-WV (4th), zero-filling (input) (5th) for 5x undersampling.

Table 1: Average SNR (dB), SSIM, inference time (sec), and ROS comparison of various schemes.

scheme	ZF	CS-WV	GANCS ($\eta = 1, \lambda = 0$)	GANCS ($\eta = 0.75, \lambda = 0.25$)
SNR	13.66	18.22	21.4	18.03
SSIM	0.64	0.8	0.84	0.76
inference time	10^{-3}	12.28	0.03	0.03
overall quality	2	4.1	4.7	5
sharpness	2	3.5	4.1	5

samples, one needs to ensure *c1)* trained manifold contains plausible images; and *c2)* points on the manifold are data consistent, namely $\mathbf{y} \approx \mathbf{\Phi x}$, $\forall \mathbf{x} \in \mathcal{M}$? To address *c1)* we adopt GANs that are very powerful in estimating a prior distribution of sharp images that are visually plausible [9]. A tandem network of generator (G) and discriminator (D) nets is considered, with an initial estimate $\tilde{\mathbf{x}}$ as the input for G (this estimate can be simply obtained as the least-norm solution $\mathbf{\Phi}^\dagger \mathbf{y}$). The G net then projects $\tilde{\mathbf{x}}$ onto $\mathcal{M}$ to remove the aliasing artifacts, and subsequently G's output, say $\check{\mathbf{x}}$, passes through D net to score 1 if $\check{\mathbf{x}} \in \mathcal{X}$, and 0 otherwise. As per *c2*, G's output, i.e., $\check{\mathbf{x}}$, may be inconsistent with the data $\mathbf{y}$. Thus, $\check{\mathbf{x}}$ is projected onto the feasible set $\mathbf{y} = \mathbf{\Phi x}$ through the affine projection $\hat{\mathbf{x}} = \mathbf{\Phi}^\dagger \mathbf{y} + (\mathbf{I} - \mathbf{\Phi}^\dagger \mathbf{\Phi})\check{\mathbf{x}}$. This resembles a single iteration of projection onto intersection of convex sets (POCS) [17]. Allowing more iterations amounts to training for multiple copies of the G net, and it turns out to considerably improve the performance as will be reported in an extended version of this work.

One issue with GANs however is that it introduces high frequency noise that can hallucinate the images. To discard this noise we adopt a mixture of least-squares (LS) GAN [18] and pixel-wise ℓ_1 cost, weighted by λ and η, respectively, to train the generator [19]. A residual net with 5 residual blocks and skip connections is adopted for G, and D is a 8-layer CNN for LS classification.

3 Results

T1-weighted contrast-enhanced abdominal image volumes are acquired for 350 pediatric patients. With 330 training (50K slices) and 20 test patients (2K slices), the network is trained to map 5-fold undersampled (zero-filled) images (with artifacts) to fully sampled ones. For a random test patient, representative slices from axial and coronal orientations, retrieved by various methods, are shown in Fig. 1. Note, CS-WV is optimized for the best performance using the FISTA algorithm with BART toolbox [20]. It is apparent that GANCS ($\eta = 0.75$, $\lambda = 0.25$) better reveals the details of liver vessels with more realistic texture and contrast, while other schemes blurred out the images. More quantitative results such as the average SNR (dB), SSIM, inference time (sec), and radiologists opinion score (ROS), ranged from 0 (poor) to 5 (excellent), about quality and sharpness is also listed in Table I. Notice particularly 100 times speed up, and ROS improvement relative to CS-WV. Details

of this study will be reported in an extended version of this work, and in our ongoing research we are validating the performance with more radiologists and a better diagnostic quality assessment strategy.

References

[1] Michael Lustig, David Donoho, and John M Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007.

[2] Morteza Mardani, Georgios B Giannakis, and Kamil Ugurbil. Tracking tensor subspaces with informative random sampling for real-time MR imaging. *arXiv preprint, arXiv:1609.04104*, 2016.

[3] Hu Chen, Yi Zhang, Mannudeep K Kalra, Feng Lin, Peixi Liao, Jiliu Zhou, and Ge Wang. Low-dose CT with a residual encoder-decoder convolutional neural network (RED-CNN). *arXiv preprint, arXiv:1702.00288*, 2017.

[4] Bo Zhu, Jeremiah Liu, Bruce Rosen, and Matthew Rosen. Neural network MR image reconstruction with AUTOMAP: Automated transform by manifold approximation. In *Proceedings of the 25st Annual Meeting of ISMRM, Honolulu, HI, USA*, 2017.

[5] Shanshan Wang, Ningbo Huang, Tao Zhao, Yong Yang, Leslie Ying, and Dong Liang. 1D partial fourier parallel MR imaging with deep convolutional neural network. In *Proceedings of the 25st Annual Meeting of ISMRM, Honolulu, HI, USA*, 2017.

[6] Jo Schlemper, Jose Caballero, Joseph V. Hajnal, Anthony Price, and Daniel Rueckert. A deep cascade of convolutional neural networks for MR image reconstruction. In *Proceedings of the 25st Annual Meeting of ISMRM, Honolulu, HI, USA*, 2017.

[7] Angshul Majumdar. Real-time dynamic MRI reconstruction using stacked denoising autoencoder. *arXiv preprint, arXiv:1503.06383 [cs.CV]*, Mar. 2015.

[8] Ashish Bora, Ajil Jalal, Eric Price, and Alexandros G Dimakis. Compressed sensing using generative models. *arXiv preprint, arXiv:1703.03208*, 2017.

[9] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, pages 2672–2680, 2014.

[10] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint, arXiv:1511.06434*, 2015.

[11] Sanjeev Arora, Rong Ge, Yingyu Liang, Tengyu Ma, and Yi Zhang. Generalization and equilibrium in generative adversarial networks (GANs). *arXiv preprint, arXiv:1703.00573v2*, Mar. 2017.

[12] Jun-Yan Zhu, Philipp Krähenbühl, Eli Shechtman, and Alexei A Efros. Generative visual manipulation on the natural image manifold. In *European Conference on Computer Vision*, pages 597–613. Springer, 2016.

[13] Justin Johnson, Alexandre Alahi, and Fei-Fei Li. Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision*, pages 694–711. Springer, Mar. 2016.

[14] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. *arXiv preprint arXiv:1609.04802*, 2016.

[15] Casper Kaae Sonderby, Jose Caballero, Lucas Theis, Wenzhe Shi, and Ferenc Huszar. Amortised MAP inference for image super-resolution. *arXiv preprint, arXiv:1610.04490*, Oct. 2016.

[16] Raymond Yeh, Chen Chen, Teck Yian Lim, Mark Hasegawa-Johnson, and Minh N Do. Semantic image inpainting with perceptual and contextual losses. *arXiv preprint, arXiv:1607.07539*, 2016.

[17] Dimitri P Bertsekas. *Nonlinear programming*. Athena scientific Belmont, 1999.

[18] Xudong Mao, Qing Li, Haoran Xie, Raymond YK Lau, Zhen Wang, and Stephen Paul Smolley. Least squares generative adversarial networks. *arXiv preprint, ArXiv:1611.04076*, Apr. 2016.

[19] Hang Zhao, Orazio Gallo, Iuri Frosio, and Jan Kautz. Loss functions for image restoration with neural networks. *IEEE Transactions on Computational Imaging*, 3(1):47–57, 2017.

[20] Jonathan I. Tamir, Frank Ong, Joseph Y. Cheng, Martin Uecker, and Michael Lustig. Generalized magnetic resonance image reconstruction using the Berkeley Advanced Reconstruction Toolbox. In *ISMRM Workshop on Data Sampling and Image Reconstruction*, Sedona, 2016.