Limitations & caveats of deep learning

Jeffrey A. Fessler

EECS Department, BME Department, Dept. of Radiology
University of Michigan

ISMRM course on Deep Learning:
"Everything" you want to know

2018-09-16

Declaration: No relevant financial interests or relationships to disclose

Outline

Outline

Medical imaging overview

Overview of medical imaging:

$$\boxed{\begin{array}{c}\text{Acquire}\\ \text{Data}\end{array}} \xrightarrow{\substack{\text{raw data}\\ \boldsymbol{y}}} \boxed{\begin{array}{c}\text{Reconstruct}\\ \text{Images}\end{array}} \xrightarrow{\substack{\text{images}\\ \hat{\boldsymbol{x}}}} \boxed{\begin{array}{c}\text{Process}\\ \text{Images}\end{array}} \rightarrow \boxed{\begin{array}{c}\text{Analyze}\\ \text{Diagnose}\\ \text{Quantify}\\ \text{Interpret}\\ \text{Intervene}\end{array}}$$

Most obvious place for machine learning is post-processing:

reconstructed images $\hat{\boldsymbol{x}}$ $\rightarrow$ ML-based image processing (classification or regression) $\rightarrow$ interpretation

...

Most obvious place for machine learning is post-processing:

reconstructed images $\hat{\boldsymbol{x}}$ $\rightarrow$ $\boxed{\text{ML-based image processing (classification or regression)}}$ $\rightarrow$ interpretation

(Several ISMRM sessions; special issue of IEEE Trans. on Med. Imaging in May 2016 [1].)

Another (initially less obvious?) place for machine learning (this course, Tue 16:15 session):

$$\begin{array}{c} \text{raw data} \\ \boldsymbol{y} \end{array} \rightarrow \boxed{\begin{array}{c} \text{ML-based} \\ \text{image reconstruction} \end{array}} \rightarrow \begin{array}{c} \text{images} \\ \hat{\boldsymbol{x}} \end{array}$$

...

Another (initially less obvious?) place for machine learning (this course, Tue 16:15 session):

raw data $\boldsymbol{y}$ $\rightarrow$ $\boxed{\text{ML-based image reconstruction}}$ $\rightarrow$ images $\hat{\boldsymbol{x}}$

Possibly easier (than diagnosis) due to lower bar:

- current reconstruction methods based on simplistic image models;
- human eyes are better at detection than inverse problems.

Another (initially less obvious?) place for machine learning (this course, Tue 16:15 session):

raw data $\boldsymbol{y}$ $\rightarrow$ $\boxed{\text{ML-based image reconstruction}}$ $\rightarrow$ images $\hat{\boldsymbol{x}}$

Possibly easier (than diagnosis) due to lower bar:

- current reconstruction methods based on simplistic image models;
- human eyes are better at detection than inverse problems.

June 2018 special issue of IEEE Trans. on Medical Imaging [2]:

IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 37, NO. 6, JUNE 2018 1289

Image Reconstruction Is a New Frontier of Machine Learning

Ge Wang, *Fellow, IEEE*, Jong Chu Ye, *Senior Member, IEEE*, Klaus Mueller, *Senior Member, IEEE*, and Jeffrey A. Fessler, *Fellow, IEEE*

A more speculative opportunity for machine learning:

raw data $\boldsymbol{y}$ $\rightarrow$ $\boxed{\text{ML-based "magic"}}$ $\rightarrow$ interpretation

...

A more speculative opportunity for machine learning:

raw data $\boldsymbol{y}$	$\rightarrow$	ML-based "magic"	$\rightarrow$	interpretation

- CT sinogram to vessel diameter [3]
- k-space to ???

Caveat: seeing is believing...

One more opportunity for ML in medical imaging:

$$\boxed{\text{Acquire Data}} \xrightarrow{\text{raw data } \boldsymbol{y}} \boxed{\text{Reconstruct Images}} \xrightarrow{\text{images } \hat{\boldsymbol{x}}} \boxed{\begin{array}{c}\text{Analyze}\\ \text{Diagnose}\\ \text{Quantify}\\ \text{Interpret}\\ \text{Intervene}\end{array}}$$

...

Machine learning in medical imaging: scan design

One more opportunity for ML in medical imaging:

Acquire Data $\xrightarrow[\boldsymbol{y}]{\text{raw data}}$ Reconstruct Images $\xrightarrow[\hat{\boldsymbol{x}}]{\text{images}}$ Analyze / Diagnose / Quantify / Interpret / Intervene

k-space sampling design using ML methods:
"Learning-based compressive MRI" [4, 5]
(Volkan Cevher group, June 2018 IEEE T-MI)
Caveat: single coil only so far; hard to generalize to parallel MRI?

Outline

Simpler methods for ML in image reconstruction

Many possible ways to use ML ideas in image reconstruction

Basic “fast” methods:

- Enhance raw data (k-space, sinogram, . . .)
- Enhance poorly reconstructed image
  - patch-based
  - image-based

Caveat: computation / quality trade-offs

. . .

Simpler methods for ML in image reconstruction

Many possible ways to use ML ideas in image reconstruction

Basic "fast" methods:

- Enhance raw data (k-space, sinogram, ...)
- Enhance poorly reconstructed image
  - patch-based
  - image-based

Caveat: computation / quality trade-offs

Basic "slow" methods:

- Auto-tune regularization parameter(s)
- Provide an initial image for "conventional" iterative recon

Caveat: may not fully exploit the potential of ML

Advanced "fast" methods for ML-based IR

- ML-based "prior" image for iterative reconstruction [6]:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \beta \|\boldsymbol{x} - \boldsymbol{x}_{\text{prior}}\|_p^p$$

Caveat: fast for $p = 2$, but $p = 1$ more robust to errors in prior image

- ML-based "prior" image for iterative reconstruction [6]:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \beta \|\boldsymbol{x} - \boldsymbol{x}_{\text{prior}}\|_p^p$$

  Caveat: fast for $p = 2$, but $p = 1$ more robust to errors in prior image

- Unrolled loop (recurrent NN) with learned components [7–10]
  (See talks by Thomas Pock and others)

  Caveat: all the issues with CNN methods forthcoming

Nonlinear encoder methods for ML-based IR

- ML-based nonlinear encoder, *e.g.*, autoencoder or generative adversarial network (GAN) [11, 12]: nonlinear generalizations of subspace models
- learn G: maps low-dimensional latent parameter $\boldsymbol{z}$ into high-dimensional image $\boldsymbol{x}$

▶ Synthesis form [13]:

$$\hat{\boldsymbol{x}} = G(\hat{\boldsymbol{z}}), \qquad \hat{\boldsymbol{z}} = \arg\min_{\boldsymbol{z}} \|\boldsymbol{A} G(\boldsymbol{z}) - \boldsymbol{y}\|_2^2$$

Caveat: $\hat{\boldsymbol{x}} \in \text{Range}(G)$, non-convex minimization

Nonlinear encoder methods for ML-based IR

- ML-based nonlinear encoder, *e.g.*, autoencoder or generative adversarial network (GAN) [11, 12]: nonlinear generalizations of subspace models
- learn G: maps low-dimensional latent parameter $\boldsymbol{z}$ into high-dimensional image $\boldsymbol{x}$

▸ Synthesis form [13]:

$$\hat{\boldsymbol{x}} = G(\hat{\boldsymbol{z}}), \qquad \hat{\boldsymbol{z}} = \arg\min_{\boldsymbol{z}} \|\boldsymbol{A}G(\boldsymbol{z}) - \boldsymbol{y}\|_2^2$$

Caveat: $\hat{\boldsymbol{x}} \in \text{Range}(G)$, non-convex minimization

▸ Regularizer form:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \beta R_{\text{encoder}}(\boldsymbol{x})$$

$$R_{\text{encoder}}(\boldsymbol{x}) = \min_{\boldsymbol{z}} \|\boldsymbol{x} - G(\boldsymbol{z})\|_p^p$$

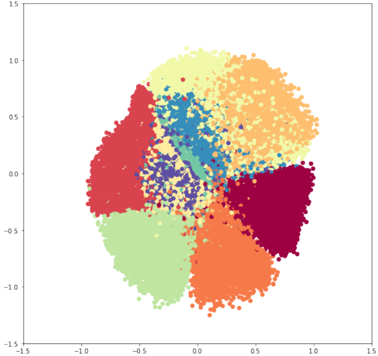
Caveat: expensive non-convex double minimization, but more robust to encoder

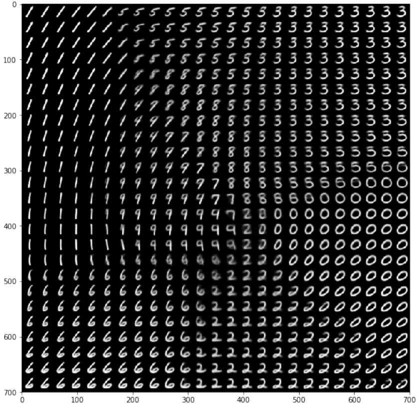
Nonlinear encoder illustration

From jupyter notebook for [14] (13 layer CNN with $\approx$ 300K learned parameters) at

`https://github.com/skolouri/swae/blob/master/MNIST_SlicedWassersteinAutoEncoder_Circle.ipynb`

$\boldsymbol{z} \in \mathbb{R}^2 \quad \mapsto \quad \boldsymbol{x} = G(\boldsymbol{z}) \in \mathbb{R}^{28 \times 28}$





Caveat: Where is 4?

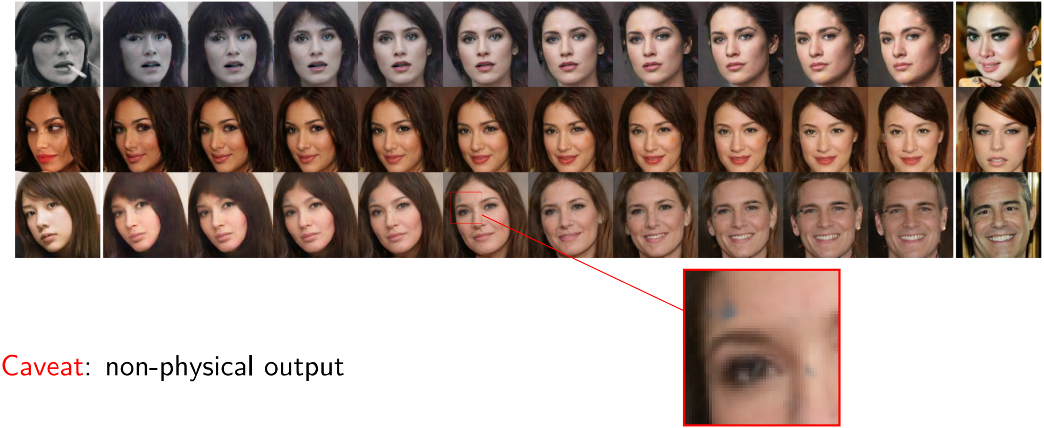
Generative Adversarial Networks (GAN) example

From Google's [15]:

Much more realistic than linear interpolation (averaging)
"setting a new milestone in visual quality" [15]

Generative Adversarial Networks (GAN) example

From Google's [15]:

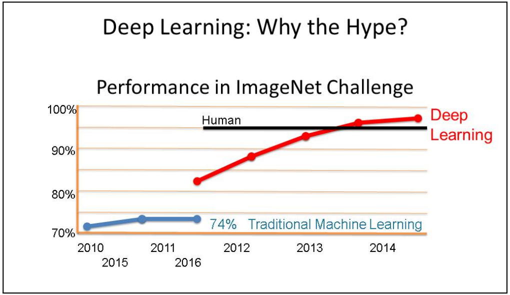


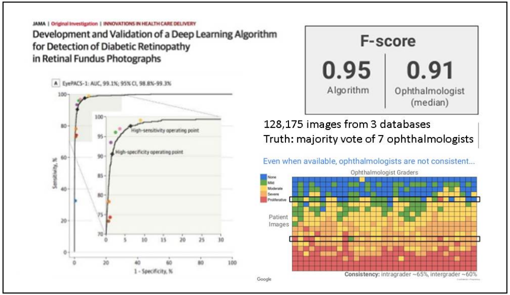
Caveat: non-physical output

Outline

Caveat: data size

From Dr. Bradley Erikson's synopsis (`http://cds.ismrm.org/protected/18MPresentations/abstracts/E1347.html`):



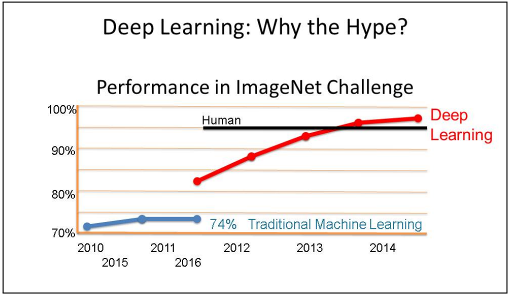


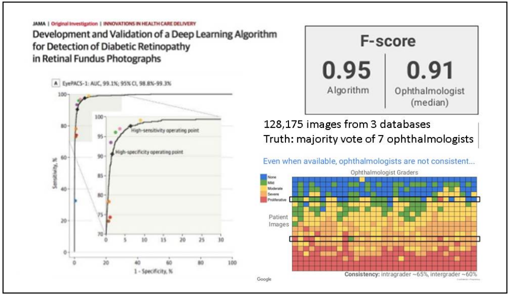
- ImageNet (`http://image-net.org/about-stats`): 14,197,122 images
- 2017 Nature paper on skin cancer classification [16]: 129,450 clinical images
- Chest X-ray study [17]: 100,000 images

. . .

Caveat: data size

From Dr. Bradley Erikson's synopsis (`http://cds.ismrm.org/protected/18MPresentations/abstracts/E1347.html`):





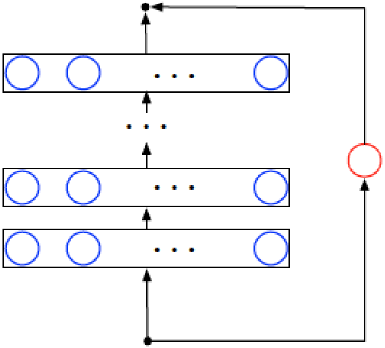
- ImageNet (`http://image-net.org/about-stats`): 14,197,122 images
- 2017 Nature paper on skin cancer classification [16]: 129,450 clinical images
- Chest X-ray study [17]: 100,000 images

Mitigation:

- Train and process image patches (*e.g.*, local operations like denoising)
- Transfer learning (pre-trained networks) [18]

Caveat: local minimizers

NN training highly nonconvex
$\Longrightarrow$ many local minimizers
...

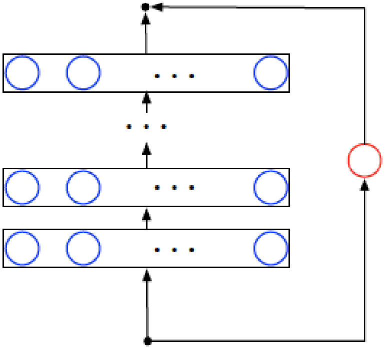


Caveat: local minimizers

NN training highly nonconvex
$\Longrightarrow$ many local minimizers

"Adding One Neuron Can Eliminate All Bad Local Minima"
exponential activation function [19]

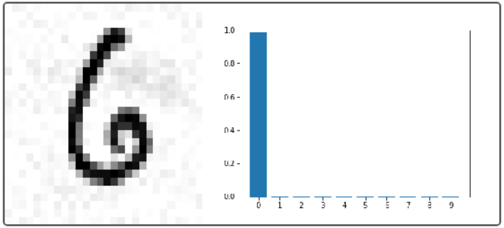
Caveat: binary classification only

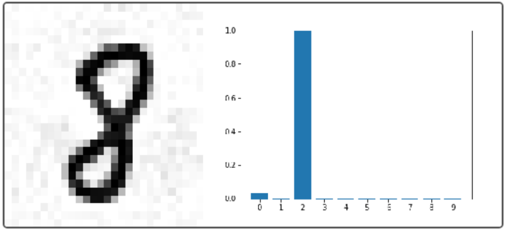


Caveat: NN methods can be fooled

Daniel Geng and Rishi Veerapaneni `https://ml.berkeley.edu/blog/2018/01/10/adversarial-examples/`

MNIST NN trained with 50000 images

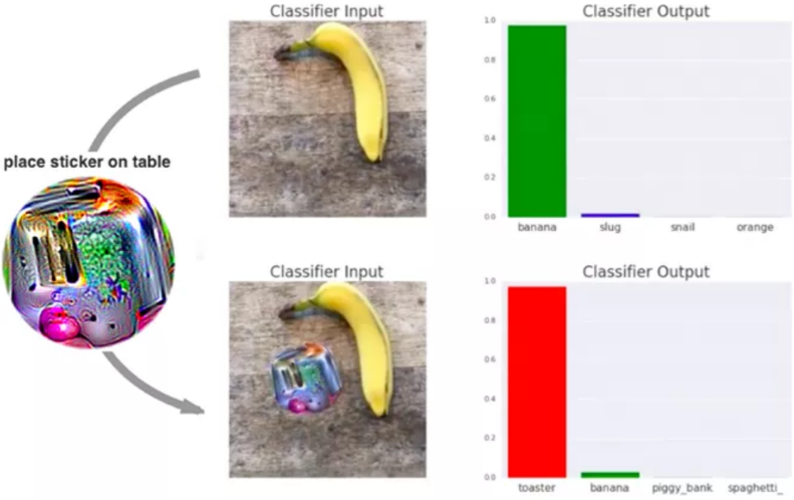




Caveat: adversarial concoctions

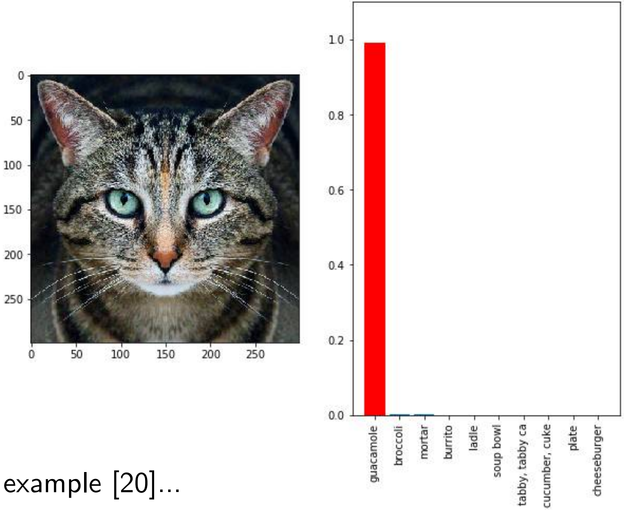
More fooling around

https://gizmodo.com/this-simple-sticker-can-trick-neural-networks-into-thin-1821735479



Obligatory cat picture

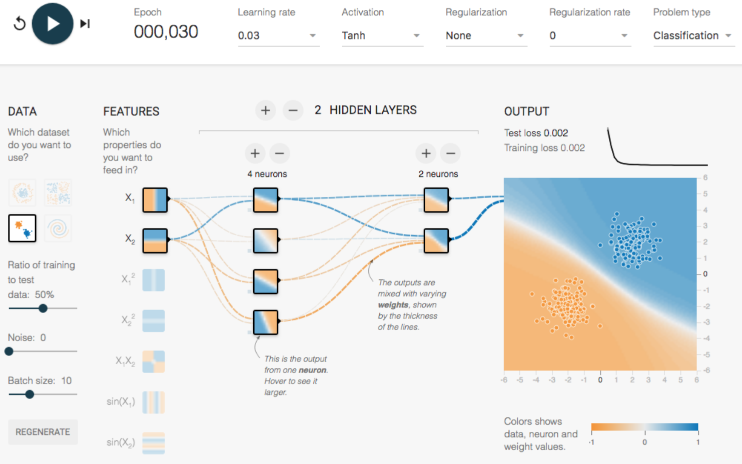
https://www.theregister.co.uk/2017/11/06/mit_fooling_ai/



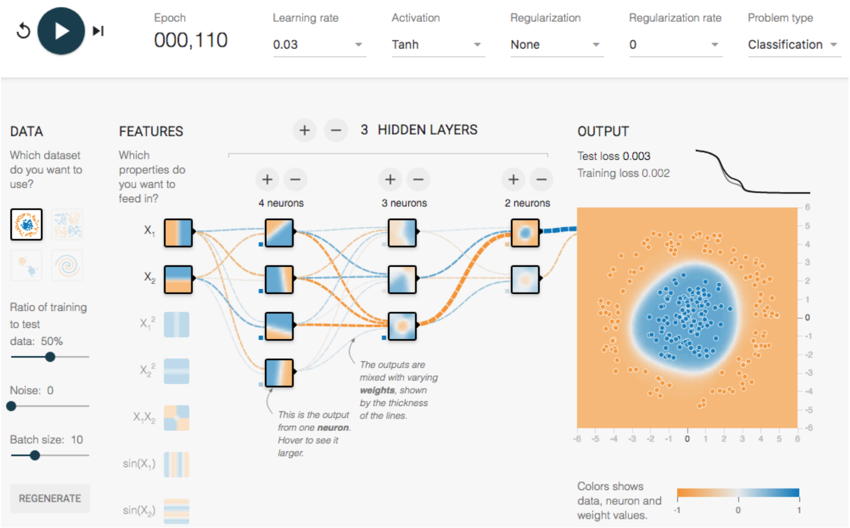
Or panda/gibbon example [20]...

TensorFlow playground I

http://playground.tensorflow.org



TensorFlow playground II



Caveat: myriads of choices

- Architecture choices
  - Input properties (just data, or functions of data, e.g., powers?)
  - # of layers
  - # of neurons per layer
  - Activation function: ReLU / Tanh / Sigmoid / Linear

Caveat: myriads of choices

- Architecture choices
  - Input properties (just data, or functions of data, e.g., powers?)
  - # of layers
  - # of neurons per layer
  - Activation function: ReLU / Tanh / Sigmoid / Linear
- Cost function (loss function) choices
  - Regularization (None or L1 or L2)
  - Regularization parameter
  - Regression / classification

Caveat: myriads of choices

- Architecture choices
  - Input properties (just data, or functions of data, e.g., powers?)
  - # of layers
  - # of neurons per layer
  - Activation function: ReLU / Tanh / Sigmoid / Linear
- Cost function (loss function) choices
  - Regularization (None or L1 or L2)
  - Regularization parameter
  - Regression / classification
- Algorithm choices
  - Learning rate (or schedule)
  - Batch size
  - # of epochs

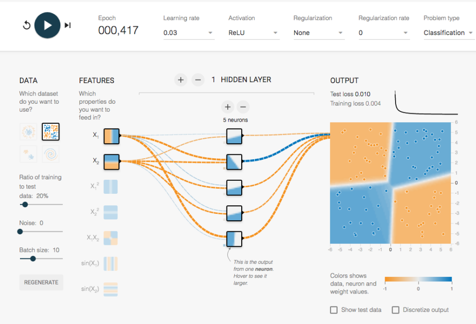
Caveat: even more choices

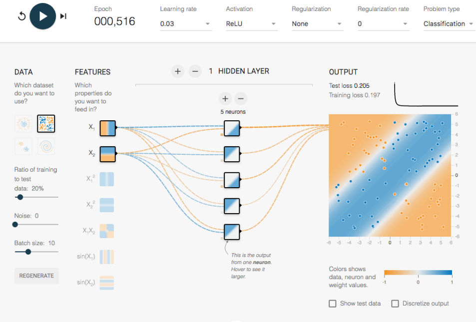
- Other NN design choices
  - Data scaling
  - Batch normalization
  - Dropout
  - Other training loss functions
    - MSE, MAE, VGG (perceptual), VGG+MSE, VGG+MAE, ...
    - MSE-based training can over-smooth details [21–23]
    - Training loss may matter more than network structure [24]
    - Task-based training of NNs (Kyle Myers et al., 1995 [25])
  - Evaluation metrics: NRMSE, SSIM, NPWE, CHO, ...
  - . . .

(Hand crafting...)

Caveat: local minimizers

Two realizations of stochastic gradient descent reach drastically different answers for a binary classification problem

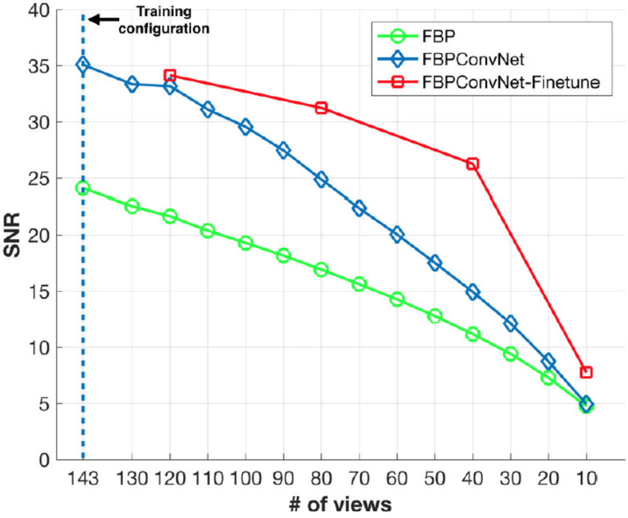




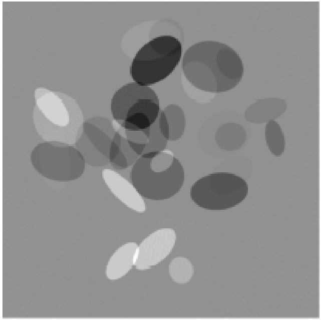
Caveat: others

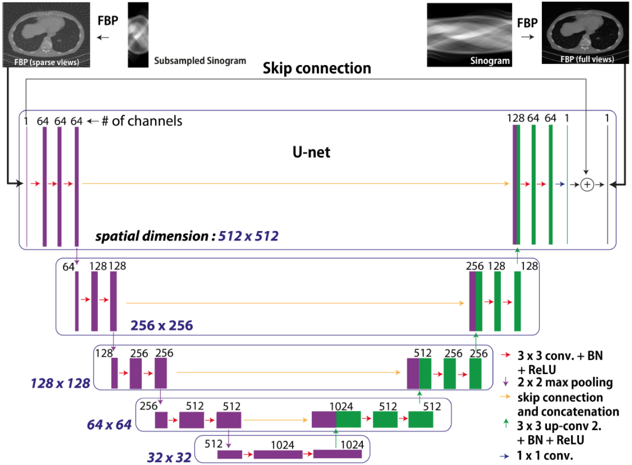
- Math skills may atrophy
  IEEE SP Magazine paper on GAN methods [26]: 1 equation!
- complex data in MRI
- Generalizability
  - Different noise levels
  - Different receive coil sensitivities
  - Different k-space sampling patterns
    [27] trained with various sampling patterns
- Fair comparisons
  - spend day(s) training a CNN
  - use default parameters for comparison methods!?
- Poor reproducibility due to unclear descriptions (Survey: [28])
- Problem size: 2D vs 3D vs 3D+time (dynamic), GPU memory constraints

Generalize CNN to different sampling patterns?
(CT views, cf radial MRI)
(Jin et al., IEEE T-IP, Sep. 2017 [29])



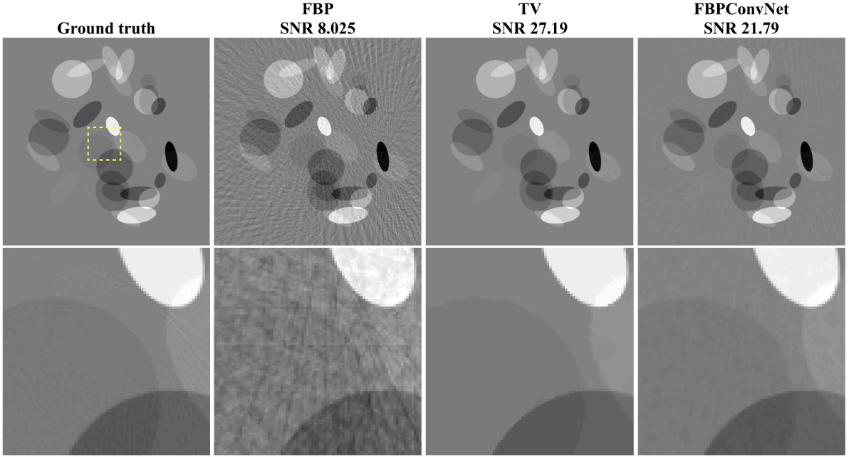
DL trained using 500 ellipse images (Jin et al., IEEE T-IP, Sep. 2017 [29])





Caveat: deep not necessarily better II

Sparse-view CT with 50 views: TV beats deep CNN [29]



Outline

Non-deep ML-based IR methods

Use training data to learn:

- dictionary $\boldsymbol{D}$ (for patches)
- sparsifying transform(s) $\boldsymbol{\Omega}$ (for patches)
- or convolutional versions thereof [30, 31]

ML-based regularized optimization problem using M image patches:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \beta R_{\mathrm{ML}}(\boldsymbol{x})$$

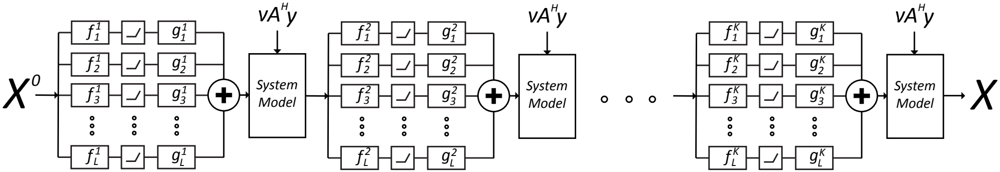
$$R_{\mathrm{ML-DL}}(\boldsymbol{x}) = \min_{\{\boldsymbol{z}_m\}} \sum_{m=1}^{M} \|\boldsymbol{P}_m \boldsymbol{x} - \boldsymbol{D}\boldsymbol{z}_m\|_2^2 + \alpha \|\boldsymbol{z}_m\|_0$$

$$R_{\mathrm{ML-ST}}(\boldsymbol{x}) = \min_{\{\boldsymbol{z}_m\}} \sum_{m=1}^{M} \|\boldsymbol{\Omega}\boldsymbol{P}_m \boldsymbol{x} - \boldsymbol{z}_m\|_2^2 + \alpha \|\boldsymbol{z}_m\|_0$$

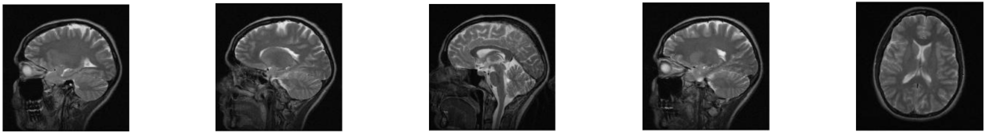
Alternative: blind adaptive learned dictionary [32] or learned sparsifying transform [33].
Caveat: double minimization, not “deep,” but more interpretable

Caveat: careful comparisons needed I

Unrolled loop method with 20 layers trained with $1.3 \cdot 10^6$ MR image 8×8 patches [9]



Tested with 5 different images:



Caveat: careful comparisons needed II

Results:

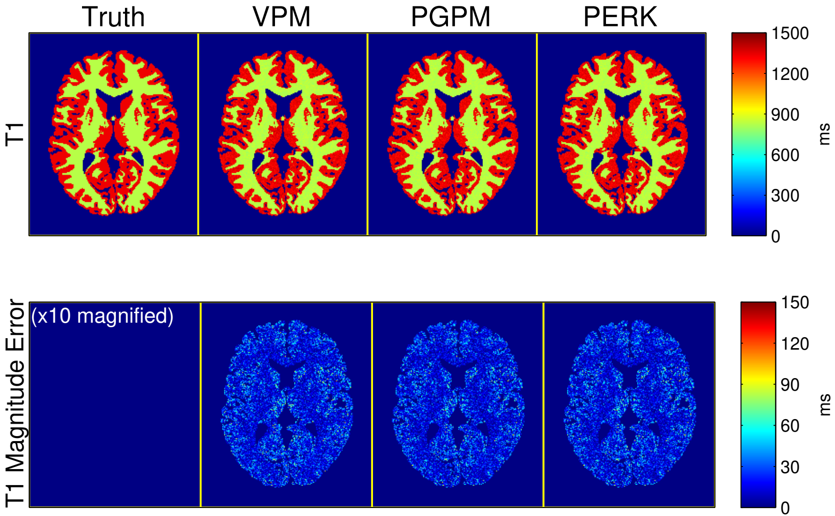
UF	Image	Zero-filled	Sparse MRI	UTMRI	Proposed
3.3×	1	25.6	26.7	28.3	28.2
	2	25.2	26.6	27.9	27.8
	3	26.0	27.3	29.3	28.9
	4	25.4	26.7	28.2	28.1
	5	27.2	28.9	30.6	30.3
Avg. PSNR change	-	-	**1.36**	**2.98**	**2.78**
5×	1	24.7	25.9	27.6	27.5
	2	24.2	25.5	27.2	27.0
	3	24.9	26.3	28.5	28.0
	4	24.4	25.7	27.6	27.4
	5	26.2	27.9	29.8	29.5
Avg. PSNR change	-	-	**1.38**	**3.26**	**3.0**
Approx recon time	-	-	**100s**	**240s**	**50s**

Sparse MRI [34] total variation and wavelets
UTMRI [35] (union of learned sparsifying transforms): adaptive, not “deep”

Caveat: DL is not always needed

Quantitative MRI: images $\rightarrow$ $\boxed{\text{estimation}}$ $\rightarrow$ parameters (T1,T2,...)

- Traditional nonlinear estimation methods:
  - nonlinear least squares
  - dictionary matching (quantized maximum likelihood via variable projection
- Machine-learning methods
  - deep neural network regression [36–39]
    Caveat: long training times
  - parameter estimation via kernel regression (PERK)
    Gopal Nataraj et al., ISBI 2017, IEEE T-MI 2018 [40, 41]


Truth
VPM
PGPM
PERK
T1
1500
1200
900
600
300
0
ms
T1 Magnitude Error
(x10 magnified)
150
120
90
60
30
0
ms

Outline

Summary

- Much excitement but many challenges
- Artificial intelligence vs artificial features

Summary

- Much excitement but many challenges
- Artificial intelligence vs artificial features
- 2017 Med. Phys. point/counter-point on ML in radiology [42]:

  "AI fares pretty well on "low hanging" targets of sharply defined skin cancers in colorful 2D photographs [16] but will face challenges from 3D gray scale, fuzzy radiology images where lesions are often subtle or diffuse, differentials are wider, and artifacts masquerade."
- Ali Rahimi of Google likens some ML methods to alchemy (trial and error)
  `http://www.sciencemag.org/news/2018/05/ai-researchers-allege-machine-learning-alchemy`
  `https://openreview.net/pdf?id=rJWF0Fywf`

  "Researchers, he said, do not know why some algorithms work and others don't, nor do they have rigorous criteria for choosing one AI architecture over another."

Summary

- Much excitement but many challenges
- Artificial intelligence vs artificial features
- 2017 Med. Phys. point/counter-point on ML in radiology [42]:

  "AI fares pretty well on "low hanging" targets of sharply defined skin cancers in colorful 2D photographs [16] but will face challenges from 3D gray scale, fuzzy radiology images where lesions are often subtle or diffuse, differentials are wider, and artifacts masquerade."
- Ali Rahimi of Google likens some ML methods to alchemy (trial and error)
  `http://www.sciencemag.org/news/2018/05/ai-researchers-allege-machine-learning-alchemy`
  `https://openreview.net/pdf?id=rJWF0Fywf`

  "Researchers, he said, do not know why some algorithms work and others don't, nor do they have rigorous criteria for choosing one AI architecture over another."
- DL is just one tool in the ML toolbox

Recommended reading (incomplete lists)

- Overviews: [28, 42, 43]
- Generative models: [14, 44]:
- Deep learning myths [45]
- NN complexity analysis / function approximation [46–48] [49]
- Application to MR fingerprinting [36, 39]
- MR reconstruction / enhancement using CNN [10, 50–57]
- Dynamic MR reconstruction using CNN [58]
- ...

Talk and code available online at
`http://web.eecs.umich.edu/~fessler`

Bibliography I

[1] H. Greenspan, B. van Ginneken, and R. M. Summers. "Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique." In: *IEEE Trans. Med. Imag.* 35.5 (May 2016), 1153–9.

[2] G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler. "Image reconstruction is a new frontier of machine learning." In: *IEEE Trans. Med. Imag.* 37.6 (June 2018), 1289–96.

[3] E. Haneda, B. Claus, P. FitzGerald, G. Wang, and B. De Man. "CT sinogram analysis using deep learning." In: *Proc. 5th Intl. Mtg. on Image Formation in X-ray CT.* 2018, 419–22.

[4] L. Baldassarre, Y-H. Li, J. Scarlett, B. Gozcu, I. Bogunovic, and V. Cevher. "Learning-based compressive subsampling." In: *IEEE J. Sel. Top. Sig. Proc.* 10.4 (June 2016), 809–22.

[5] B. Gozcu, R. K. Mahabadi, Y-H. Li, E. Ilicak, T. Cukur, J. Scarlett, and V. Cevher. "Learning-based compressive MRI." In: *IEEE Trans. Med. Imag.* 37.6 (June 2018), 1394–406.

[6] G. Yang, S. Yu, H. Dong, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, Y. Guo, and D. Firmin. "DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction." In: *IEEE Trans. Med. Imag.* 37.6 (June 2018), 1310–21.

[7] K. Gregor and Y. LeCun. "Learning fast approximations of sparse coding." In: *Proc. Intl. Conf. Mach. Learn.* 2010.

[8] Y. Chen and T. Pock. "Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration." In: *IEEE Trans. Patt. Anal. Mach. Int.* 39.6 (June 2017), 1256–72.

[9] S. Ravishankar, A. Lahiri, C. Blocker, and J. A. Fessler. "Deep dictionary-transform learning for image reconstruction." In: *Proc. IEEE Intl. Symp. Biomed. Imag.* 2018, 1208–12.

[10] K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll. "Learning a variational network for reconstruction of accelerated MRI data." In: *Mag. Res. Med.* 79.6 (June 2018), 3055–71.

[11] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. *Generative adversarial networks*. 2014.

[12] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever, and P. Abbeel. "InfoGAN: interpretable representation learning by information maximizing generative adversarial nets." In: *Neural Info. Proc. Sys.* 2016, 2172–80.

[13] A. Bora, A. Jalal, E. Price, and A. G. Dimakis. "Compressed sensing using generative models." In: *Proc. Intl. Conf. Mach. Learn.* Vol. 70. 2017, 537–46.

[14] S. Kolouri, C. E. Martin, and G. K. Rohde. *Sliced-Wasserstein autoencoder: an embarrassingly simple generative model*. 2018.

[15] D. Berthelot, T. Schumm, and L. Metz. *BEGAN: boundary equilibrium generative adversarial networks*. 2017.

[16] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun. "Dermatologist-level classification of skin cancer with deep neural networks." In: *Nature* 542.7639 (2017), 115–8.

[17] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng. *CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning*. 2017.

[18] H. Shan, Y. Zhang, Q. Yang, U. Kruger, W. Cong, and G. Wang. "3D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2D trained network." In: *IEEE Trans. Med. Imag.* 37.6 (June 2018), 1522–34.

[19] S. Liang, R. Sun, J. D. Lee, and R. Srikant. *Adding one neuron can eliminate all bad local minima*. 2018.

[20] I. J. Goodfellow, J. Shlens, and C. Szegedy. "Explaining and harnessing adversarial examples." In: *Proc. Intl. Conf. on Learning Representations*. 2015.

[21] J. Johnson, A. Alahi, and L. Fei-Fei. *Perceptual losses for realtime style transfer and super-resolution*. 2016.

[22] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi. *Photo-realistic single image super-resolution using a generative adversarial network*. 2016.

Bibliography III

[23] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang. "Low dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss." In: *IEEE Trans. Med. Imag.* 37.6 (June 2018), 1348–57.

[24] B. Kim, H. Shim, and J. Baek. "Performance comparison of deep learning based denoising techniques in low-dose CT images." In: *Proc. 5th Intl. Mtg. on Image Formation in X-ray CT.* 2018, 426–9.

[25] K. J. Myers, M. P. Anderson, D. G. Brown, R. F. Wagner, and K. M. Hanson. "Neural network performance for binary discrimination tasks, Part II: effect of task training, and feature pre-selection." In: *Proc. SPIE 2434 Med. Im. 1995: Im. Proc.* 1995, 828–37.

[26] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath. "Generative adversarial networks: an overview." In: *IEEE Sig. Proc. Mag.* 35.1 (Jan. 2018), 53–65.

[27] H. K. Aggarwal, M. P. Mani, and M. Jacob. "Model based image reconstruction using deep learned priors (MODL)." In: *Proc. IEEE Intl. Symp. Biomed. Imag.* 2018, 671–4.

[28] M. T. McCann, K. H. Jin, and M. Unser. "Convolutional neural networks for inverse problems in imaging: A review." In: *IEEE Sig. Proc. Mag.* 34.6 (Nov. 2017), 85–95.

[29] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser. "Deep convolutional neural network for inverse problems in imaging." In: *IEEE Trans. Im. Proc.* 26.9 (Sept. 2017), 4509–22.

[30] I. Y. Chun and J. A. Fessler. "Convolutional dictionary learning: acceleration and convergence." In: *IEEE Trans. Im. Proc.* 27.4 (Apr. 2018), 1697–712.

[31] I. Y. Chun and J. A. Fessler. *Convolutional analysis operator learning: acceleration, convergence, application, and neural networks*. 2018.

[32] S. Ravishankar and Y. Bresler. "MR image reconstruction from highly undersampled k-space data by dictionary learning." In: *IEEE Trans. Med. Imag.* 30.5 (May 2011), 1028–41.

[33] S. Ravishankar and Y. Bresler. "Efficient blind compressed sensing using sparsifying transforms with convergence guarantees and application to MRI." In: *SIAM J. Imaging Sci.* 8.4 (2015), 2519–57.

[34] M. Lustig, D. Donoho, and J. M. Pauly. "Sparse MRI: The application of compressed sensing for rapid MR imaging." In: *Mag. Res. Med.* 58.6 (Dec. 2007), 1182–95.

[35] S. Ravishankar and Y. Bresler. "Data-driven learning of a union of sparsifying transforms model for blind compressed sensing." In: *IEEE Trans. Computational Imaging* 2.3 (Sept. 2016), 294–309.

[36] P. Virtue, S. X. Yu, and M. Lustig. "Better than real: Complex-valued neural nets for MRI fingerprinting." In: *Proc. IEEE Intl. Conf. on Image Processing*. 2017, 3953–7.

[37] A. Lahiri, J. A. Fessler, and L. Hernandez-Garcia. "Optimized design of MRF scan parameters for ASL signal acquisition." In: *ISMRM Workshop on MR Fingerprinting*. 2017.

[38] A. Lahiri, J. A. Fessler, and L. Hernandez-Garcia. "Optimized scan design for ASL fingerprinting and multiparametric estimation using neural network regression." In: *Proc. Intl. Soc. Mag. Res. Med.* 2018, p. 309.

[39] O. Cohen, B. Zhu, and M. S. Rosen. "MR fingerprinting Deep RecOnstruction NEtwork (DRONE)." In: *Mag. Res. Med.* 80.3 (Sept. 2018), 885–94.

[40] G. Nataraj, J-F. Nielsen, and J. A. Fessler. "Dictionary-free MRI parameter estimation via kernel ridge regression." In: *Proc. IEEE Intl. Symp. Biomed. Imag.* 2017, 5–9.

[41] G. Nataraj, J-F. Nielsen, C. D. Scott, and J. A. Fessler. "Dictionary-free MRI PERK: Parameter estimation via regression with kernels." In: *IEEE Trans. Med. Imag.* (2018). To appear.

[42] G. Wang, M. Kalra, and C. G. Orton. "Machine learning will transform radiology significantly within the next five years." In: *Med. Phys.* 44.6 (June 2017), 2041–4.

[43] G. Wang. "A perspective on deep imaging." In: *IEEE Access* 4 (Nov. 2016), 8914–24.

[44] I. Deshpande, Z. Zhang, and A. Schwing. "Generative Modeling using the sliced Wasserstein distance." In: *Proc. IEEE Conf. on Comp. Vision and Pattern Recognition*. 2018.

Bibliography V

[45] S. Rakhlin. *MythBusters: A Deep Learning Edition*. Slides dated Jan 18-19, 2018. 2018.

[46] N. Golowich, A. Rakhlin, and O. Shamir. *Size-independent sample complexity of neural networks*. 2017.

[47] T. Liang, T. Poggio, A. Rakhlin, and J. Stokes. *Fisher-Rao metric, geometry, and complexity of neural networks*. 2017.

[48] M. Raghu, B. Poole, J. Kleinberg, S. Ganguli, and J. Sohl-Dickstein. "On the expressive power of deep neural networks." In: *Proc. Intl. Conf. Mach. Learn.* Vol. 70. 2017, 2847–54.

[49] S. Liang and R. Srikant. "Why deep neural networks for function approximation?." In: *Proc. Intl. Conf. on Learning Representations*. 2017.

[50] S. Ravishankar, I. Y. Chun, and J. A. Fessler. "Physics-driven deep training of dictionary-based algorithms for MR image reconstruction." In: *Proc., IEEE Asilomar Conf. on Signals, Systems, and Comp.* Invited. 2017, 1859–63.

[51] M. Mardani, E. Gong, J. Y. Cheng, S. Vasanawala, G. Zaharchuk, M. Alley, N. Thakur, S. Han, W. Dally, J. M. Pauly, and L. Xing. *Deep generative adversarial networks for compressed sensing automates MRI*. 2017.

[52] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen. "Image reconstruction by domain-transform manifold learning." In: *Nature* 555 (Mar. 2018), 487–92.

[53] Y. Han, J. Yoo, H. H. Kim, H. J. Shin, K. Sung, and J. C. Ye. "Deep learning with domain adaptation for accelerated projection-reconstruction MR." In: *Mag. Res. Med.* 80.3 (Sept. 2018), 1189–205.

[54] K. H. Jin and M. Unser. "3D BPConvNet to reconstruct parallel MRI." In: *Proc. IEEE Intl. Symp. Biomed. Imag.* 2018, 361–4.

[55] H. Jeelani, J. Martin, F. Vasquez, M. Salerno, and D. S. Weller. "Image quality affects deep learning reconstruction of MRI." In: *Proc. IEEE Intl. Symp. Biomed. Imag.* 2018, 357–60.

[56] T. M. Quan, T. Nguyen-Duc, and W-K. Jeong. "Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss." In: *IEEE Trans. Med. Imag.* 37.6 (June 2018), 1488–97.

[57] T. Eo, Y. Jun, T. Kim, J. Jang, H-J. Lee, and D. Hwang. "KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images." In: *Mag. Res. Med.* (2018).

[58] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert. "A deep cascade of convolutional neural networks for dynamic MR image reconstruction." In: *IEEE Trans. Med. Imag.* 37.2 (Feb. 2018), 491–503.