Physikalisch-Technische Bundesanstalt
Braunschweig and Berlin
National Metrology Institute

Data-Driven Regularization Methods in Imaging

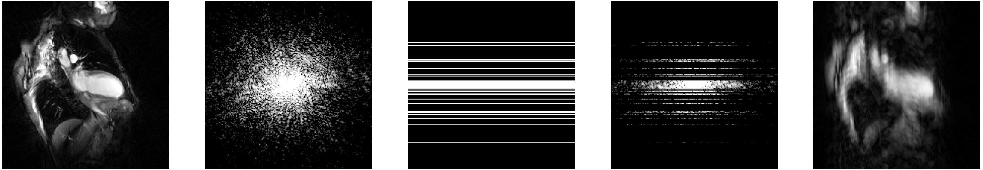
Thematic Einstein Semester on

Mathematics of Imaging in Real-World Challenges

Andreas Kofler, PhD
Physikalisch-Technische Bundesanstalt, Berlin and Braunschweig, Germany
Division of Medical Physics and Metrological Information Technology

October 29th, 2021

Problem: Measurements are **incomplete**.



$\mathbf{x}$	$\mathbf{y} := \mathbf{F}\mathbf{x}$	$\mathbf{S}_I$	$\mathbf{y}_I := (\mathbf{S}_I \circ \mathbf{F})\mathbf{x}$	$\mathbf{x}_I := \mathbf{F}_I^{\mathsf{H}}\mathbf{y}_I$
unknown image	fully-sampled k-space data (Cartesian)	binary mask	undersampled k-space data	initial reconstruction, $R \approx 8$

$\rightarrow$ scanning time $\sim \frac{1}{R}$

Machine Learning/Data-Driven Regularization

Regularization in the form of penalty terms:

$$F_{\mathcal{R}}(\mathbf{x}) = D(\mathbf{A}\mathbf{x}, \mathbf{y}) + \mathcal{R}(\mathbf{x}) \to \min_{\mathbf{x}}!$$

Machine Learning/Data-Driven Regularization

Regularization in the form of penalty terms:

$$F_{\mathcal{R}}(\mathbf{x}) = D(\mathbf{A}\mathbf{x}, \mathbf{y}) + \mathcal{R}(\mathbf{x}) \rightarrow \min_{\mathbf{x}}!$$

$\rightarrow$ What is a good choice for the regularization $\mathcal{R}(\mathbf{x})$?

Machine Learning/Data-Driven Regularization

Regularization in the form of penalty terms:

$$F_{\mathcal{R}}(\mathbf{x}) = D(\mathbf{A}\mathbf{x}, \mathbf{y}) + \mathcal{R}(\mathbf{x}) \rightarrow \min_{\mathbf{x}}!$$

$\rightarrow$ What is a good choice for the regularization $\mathcal{R}(\mathbf{x})$?

Main Idea:
Let $\mathcal{R}(\mathbf{x})$ be learned from data, i.e. $\mathcal{R}(\mathbf{x}) = \mathcal{R}_{\Theta}(\mathbf{x})$, where Θ is a set of trainable parameters.

Conventions, Dataset, Code, . . .

In the following:

$$\mathbf{F}_I\mathbf{x} + \mathbf{e} = \mathbf{y}_I,$$

- $\mathbf{x} \in \mathbb{C}^N$, $N = N_x \cdot N_y \cdot N_t$, $\mathbf{y} \in \mathbb{C}^M$ with $M < N$ (underdetermined system)
- $\mathbf{F}_I = \mathbf{S}_I\mathbf{F}$, where $\mathbf{S}_I$ binary mask, $I \subset J = \{1, \ldots, N\}$
- $\mathbf{x}_I := \mathbf{F}_I^{\mathsf{H}}$ - zero-filled reconstruction
- dynamic cardiac MRI
- Cartesian sampling grid
- single-coil

Neural Networks in a Nutshell - Neural Networks

A neural network with L layers is typically a function f_Θ of the form

$$f_\Theta = (\sigma_L \circ f^L_{\Theta_L}) \circ \ldots \circ (\sigma_1 \circ f^1_{\Theta_1}),$$

where $\Theta := \cup_{l=1}^L \Theta_l$ are called the weights.

Neural Networks in a Nutshell - Neural Networks

A neural network with L layers is typically a function f_Θ of the form

$$f_\Theta = (\sigma_L \circ f^L_{\Theta_L}) \circ \ldots \circ (\sigma_1 \circ f^1_{\Theta_1}),$$

where $\Theta := \cup_{l=1}^{L} \Theta_l$ are called the weights.

Training a network on a dataset of pairs $\mathcal{D} := \{(x_i, y_i)_{i=1}^{M}\}$ refers to finding a set of weights such that

$$\Theta^* \in \arg\min_{\Theta} \mathcal{L}(\Theta) \quad \text{with} \quad \mathcal{L}(\Theta) := \frac{1}{M} \sum_{i=1}^{M} l(f_\Theta(x_i), y_i),$$

where $l(\cdot,\cdot)$ is an appropriate **Loss-Function**, e.g. the L_2-error.

Neural Networks in a Nutshell - Training/Validation/Test Sets

Neural Networks in a Nutshell - Training/Validation/Test Sets

- Training Set: optimize Θ on this

Neural Networks in a Nutshell - Training/Validation/Test Sets

- Training Set: optimize Θ on this
- Validation Set: tune hyper-parameters, choose when to stop training, ...

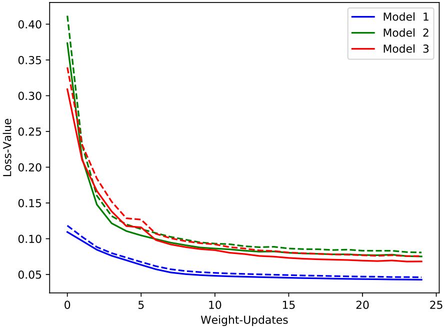
Neural Networks in a Nutshell - Training/Validation/Test Sets

- Training Set: optimize Θ on this
- Validation Set: tune hyper-parameters, choose when to stop training, ...
- Test Set: apply the model and report results

Neural Networks in a Nutshell - Training/Validation/Test Sets

- Training Set: optimize Θ on this
- Validation Set: tune hyper-parameters, choose when to stop training, ...
- Test Set: apply the model and report results

Remark: Never (EVER!) make decisions about the model on the test set!!

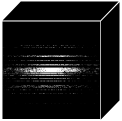
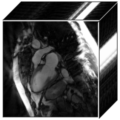


Types of Neural Networks for Image Reconstruction

Main different types of (supervised) NNs for image reconstruction

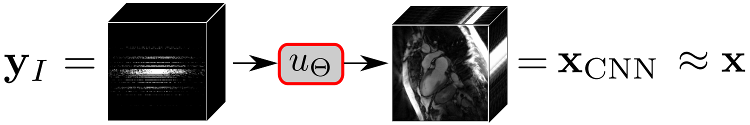
- Full inversion
- Post-Processing methods
- Decoupled methods
- Iterative/Cascaded networks

Learn the Full Inversion

$\mathbf{y}_I =$  $\longrightarrow \boxed{u_\Theta} \longrightarrow$  $= \mathbf{x}_{\mathrm{CNN}} \approx \mathbf{x}$

Idea: Learn to invert the forward model [13].

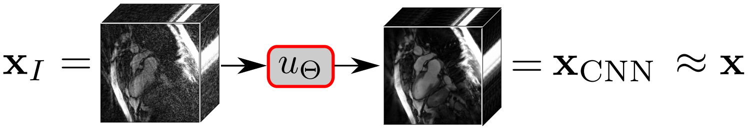
Learn the Full Inversion



Idea: Learn to invert the forward model [13].

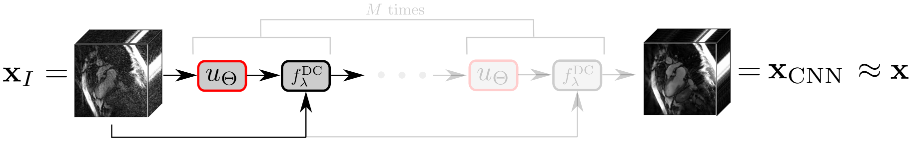
Remark: Typically, not recommended, as the physical model is (at least partially) known!

Post Processing



Idea: Train a network to reduce noise/artefacts from an initial image, see e.g. [6], [11], [3], [7].

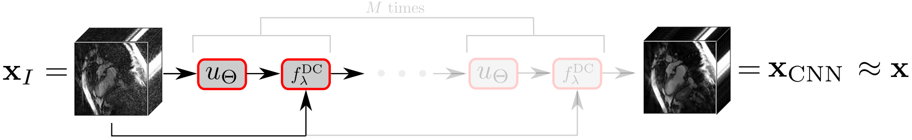
Decoupled Methods



Part of the network NOT Part of the network

u_Θ - Network f_λ^{DC} - Data-consistency block which contains $\mathbf{F}_I$, $\mathbf{F}_I^{\mathsf{H}}$

Idea: First learn, then reconstruct, e.g. [5], [4], [8].



Part of the network   NOT Part of the network

u_Θ - Network f_λ^{DC} - Data-consistency block which contains $\mathbf{F}_I$, $\mathbf{F}_I^{\mathsf{H}}$

Idea: Learn to reconstruct [1], [2], [12], [10], [9].

Outline

In this tutorial: **Iterative Neural Networks**

Consider the functional

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{F}_I\mathbf{x} - \mathbf{y}_I\|_2^2 + \frac{\lambda}{2}\|\mathbf{x}\|_2^2.$$

Tikhonov-Regularization

Consider the functional

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{F}_I\mathbf{x} - \mathbf{y}_I\|_2^2 + \frac{\lambda}{2}\|\mathbf{x}\|_2^2.$$

Calculate the derivative, set it to zero, re-arrange and obtain the system:

$$(\mathbf{F}_I^{\mathsf{H}}\mathbf{F}_I + \lambda\,\mathbf{I})\,\mathbf{x} = \mathbf{F}_I^{\mathsf{H}}\mathbf{y}_I,$$

which has a closed-form solution

$$\mathbf{x}^* = \mathbf{F}^{\mathsf{H}}\mathbf{M}\mathbf{F}\mathbf{x}_I,$$

where

$$\mathbf{M}_{k,k} = \begin{cases} 1/(1+\lambda) & \text{if } k \in I, \\ 1/\lambda & \text{if } k \notin I. \end{cases}$$

Generalized Tikhonov-Regularization

For some prior $\mathbf{x}_{\text{prior}}$, consider the functional

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{F}_I\mathbf{x} - \mathbf{y}_I\|_2^2 + \frac{\lambda}{2}\|\mathbf{x} - \mathbf{x}_{\text{prior}}\|_2^2.$$

Generalized Tikhonov-Regularization

For some prior $\color{red}{\mathbf{x}_{\text{prior}}}$, consider the functional

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{F}_I\mathbf{x} - \mathbf{y}_I\|_2^2 + \frac{\lambda}{2}\|\mathbf{x} - \color{red}{\mathbf{x}_{\text{prior}}}\|_2^2.$$

Calculate the derivative, set it to zero, re-arrange and obtain the system:

$$\left(\mathbf{F}_I^{\mathsf{H}}\mathbf{F}_I + \lambda\,\mathbf{I}\right)\mathbf{x} = \mathbf{F}_I^{\mathsf{H}}\mathbf{y}_I + \lambda\,\color{red}{\mathbf{x}_{\text{prior}}}$$

which has a closed-form solution

$$\mathbf{x}^* = \mathbf{F}^{\mathsf{H}}\left(\mathbf{\Lambda}\mathbf{F}\color{red}{\mathbf{x}_{\text{prior}}} + \frac{1}{1+\lambda}\mathbf{y}_I\right)$$

where

$$\mathbf{\Lambda}_{k,k} = \begin{cases} \lambda/(1+\lambda) & \text{if } k \in I, \\ 1 & \text{if } k \notin I. \end{cases}$$

Network Architecture

The network in [1] alternates between the following two steps for $1 \leq k \leq M$:

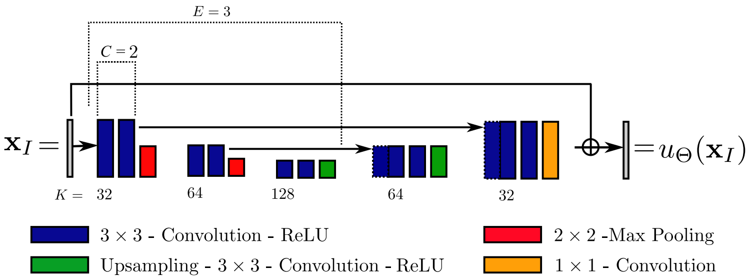
$$\begin{aligned} \mathbf{x}_{\mathrm{CNN}}^{k} &= u_{\Theta}(\mathbf{x}^{k}) \\ \mathbf{x}^{k+1} &= \arg\min_{\mathbf{x}} \|\mathbf{F}_{I}\mathbf{x} - \mathbf{y}_{I}\|_2^2 + \lambda \|\mathbf{x} - \mathbf{x}_{\mathrm{CNN}}^{k}\|_2^2 \end{aligned}$$

Network Architecture

The network in [1] alternates between the following two steps for $1 \leq k \leq M$:

$$
\begin{aligned}
\mathbf{x}_{\mathrm{CNN}}^{k} &= u_{\Theta}(\mathbf{x}^{k}) \\
\mathbf{x}^{k+1} &= \arg\min_{\mathbf{x}} \|\mathbf{F}_I \mathbf{x} - \mathbf{y}_I\|_2^2 + \lambda \|\mathbf{x} - \mathbf{x}_{\mathrm{CNN}}^{k}\|_2^2 \\
&= \mathbf{F}^{\mathsf{H}}(\boldsymbol{\Lambda}\mathbf{F}\mathbf{x}_{\mathrm{CNN}}^{k} + \frac{1}{1+\lambda}\mathbf{y}_I)
\end{aligned}
$$

with $\mathbf{x}^0 := \mathbf{F}_I^{\mathsf{H}}\mathbf{y}_I$ and $\boldsymbol{\Lambda}$ as before.



Hyper-parameters:

- E - number of encoding stages
- C - number of convolutional layers per stage
- K - initial number of filters

Let's begin! ;)

https://github.com/ckolbPTB/TES_21_22_Tutorials/

Questions:

- In the code `tutorial1_data_driven_reg_methods_apply_cnn_modl.ipynp`, try to vary the regularization parameter λ, the CNN-block for the MODL architecture, the length of the network, etc... What are your observations?
- Do the networks work for all different λ and different acceleration factors?
- Which of thee provided pre-trained networks seems to be better?
- Following the recipe on page 13 in this presentation, what would the data-consistency block look like for the following denoising problem

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{x} - \mathbf{y}_I\|_2^2 + \frac{\lambda}{2}\|\mathbf{x} - \mathbf{x}_{\text{prior}}\|_2^2\,?$$

Thank you for your attention.

Physikalisch-Technische Bundesanstalt
Braunschweig and Berlin
Abbestraße 2
10587 Berlin

Andreas Kofler
Phone: +49 (0)30 3481-7749
eMail: andreas.kofler@ptb.de
GitHub: `https://github.com/koflera/`

www.ptb.de

References I

[1] Hemant K Aggarwal, Merry P Mani, and Mathews Jacob.
Modl: Model-based deep learning architecture for inverse problems.
IEEE Transactions on Medical Imaging, 38(2):394–405, 2018.

[2] Kerstin Hammernik, Teresa Klatzer, Erich Kobler, Michael P Recht, Daniel K Sodickson, Thomas Pock, and Florian Knoll.
Learning a variational network for reconstruction of accelerated MRI data.
Magnetic Resonance in Medicine, 79(6):3055–3071, 2018.

[3] Andreas Hauptmann, Simon Arridge, Felix Lucka, Vivek Muthurangu, and Jennifer A Steeden.
Real-time cardiovascular mr with spatio-temporal artifact suppression using deep learning–proof of concept in congenital heart disease.
Magnetic Resonance in Medicine, 81(2):1143–1156, 2019.

References II

[4] Andreas Hauptmann, Felix Lucka, Marta Betcke, Nam Huynh, Jonas Adler, Ben Cox, Paul Beard, Sebastien Ourselin, and Simon Arridge.
Model-based learning for accelerated, limited-view 3-d photoacoustic tomography.
IEEE Transactions on Medical Imaging, 37(6):1382–1393, 2018.

[5] Chang Min Hyun, Hwa Pyung Kim, Sung Min Lee, Sungchul Lee, and Jin Keun Seo.
Deep learning for undersampled MRI reconstruction.
Physics in Medicine & Biology, 63(13):135007, 2018.

[6] Kyong Hwan Jin, Michael T McCann, Emmanuel Froustey, and Michael Unser.
Deep convolutional neural network for inverse problems in imaging.
IEEE Transactions on Image Processing, 26(9):4509–4522, 2017.

References III

[7] Andreas Kofler, Marc Dewey, Tobias Schaeffter, Christian Wald, and Christoph Kolbitsch.
Spatio-temporal deep learning-based undersampling artefact reduction for 2d radial cine MRI with limited training data.
IEEE Transactions on Medical Imaging, 39(3):703–717, 2020.

[8] Andreas Kofler, Markus Haltmeier, Tobias Schaeffter, Marc Kachelriess, Marc Dewey, Christian Wald, and Christoph Kolbitsch.
Neural networks-based regularization for large-scale medical image reconstruction.
Physics in Medicine & Biology, 65(13):135003, 2020.

[9] Andreas Kofler, Markus Haltmeier, Tobias Schaeffter, and Christoph Kolbitsch.
An end-to-end-trainable iterative network architecture for accelerated radial multi-coil 2d cine mr image reconstruction.
Medical Physics, 48(5):2412–2425, 2021.

References IV

[10] Chen Qin, Jo Schlemper, Jose Caballero, Anthony N Price, Joseph V Hajnal, and Daniel Rueckert.
Convolutional recurrent neural networks for dynamic MR image reconstruction.
IEEE Transactions on Medical Imaging, 38(1):280–290, 2018.

[11] Christopher M. Sandino, Neerav Dixit, Joseph Y. Cheng, and Shreyas S. Vasanawala.
Deep convolutional neural networks for accelerated dynamic magnetic resonance imaging.
Proceedings of 31st Conference of Neural Information Processing Systems (NIPS), Medical Imaigng meets NIPS Workshop., 2017. Online. Available at `www.doc.ic.ac.uk/bglocker/public/mednips2017/mednips_2017_paper_19.pdf`.

References V

[12] Jo Schlemper, Jose Caballero, Joseph V Hajnal, Anthony N Price, and Daniel Rueckert.
A deep cascade of convolutional neural networks for dynamic MR image reconstruction.
IEEE Transactions on Medical Imaging, 37(2):491–503, 2018.

[13] Bo Zhu, Jeremiah Z Liu, Stephen F Cauley, Bruce R Rosen, and Matthew S Rosen.
Image reconstruction by domain-transform manifold learning.
Nature, 555(7697):487–492, 2018.