EPI-Learning: Explicit Physics Induced Learning

Dr. Shira Nemirovsky-Rotman

CoMLab – Computational Medical imaging Laboratory

Department of Medical Imaging Sciences

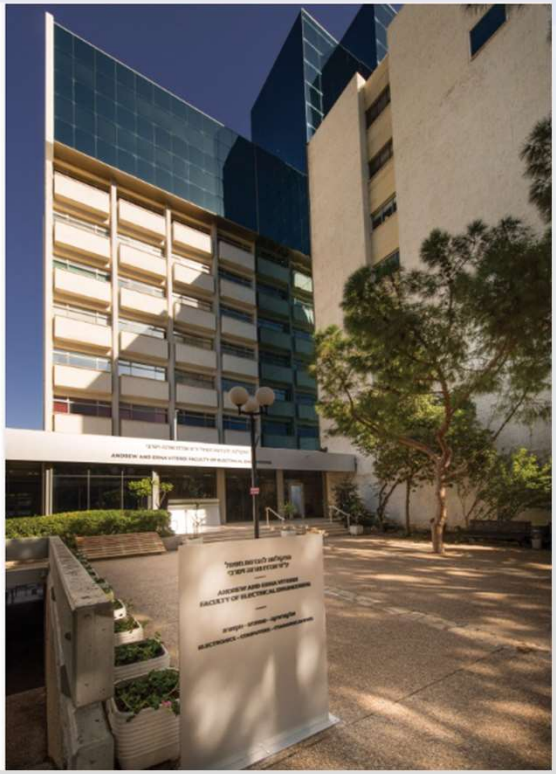
Faculty of Social Welfare and Health Sciences

University of Haifa, Israel

TECHNION
Israel Institute of Technology

TECHNION 100 EXPLORE | DISCOVER | INNOVATE

אוניברסיטת חיפה
UNIVERSITY OF HAIFA
جامعة حيفا



RAMBAM
Health Care Campus

אוניברסיטת חיפה
UNIVERSITY OF HAIFA
جامعة حيفا

The MRI Center Rambam-University of Haifa

New University of Haifa
and Rambam MRI Center (3 Tesla
Siemens Machine), October 2022

Rambam Healthcare
Campus, Haifa

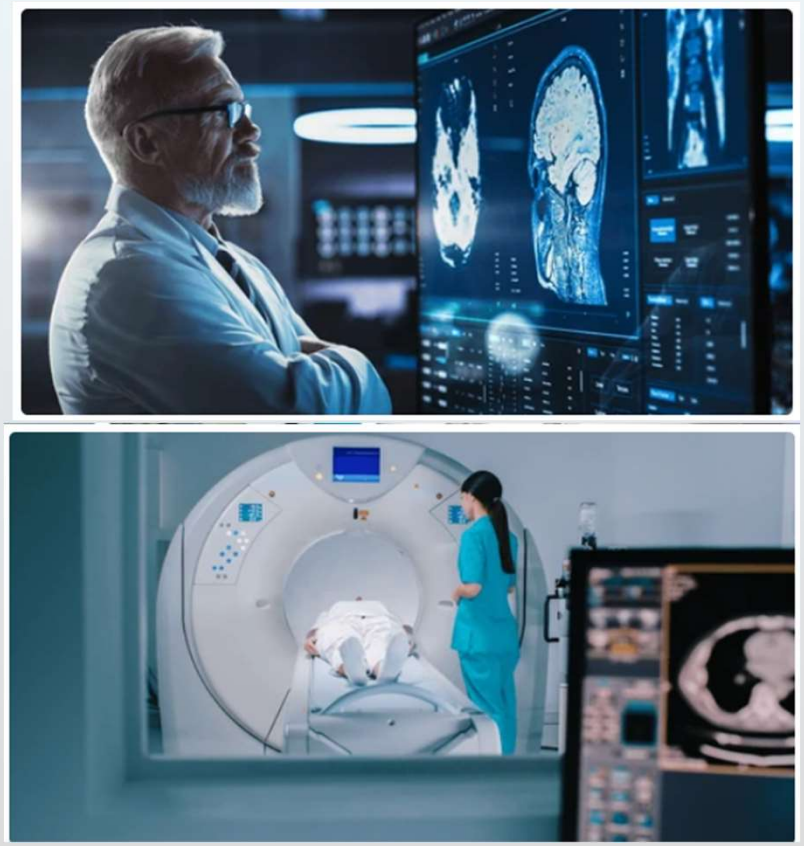
Discoveries Tower
CoMLab

Agenda

1. *Why medical imaging*
2. *Robustness and generalization*
3. *MRI background and basic principles*
4. *Physics-based Deep Learning for MRI reconstruction*
5. *EPI-Learning*
6. *EPI-Learning for Regression task: Diffusion-Weighted MRI*
7. *EPI-Learning for Reconstruction task: Low-Field MRI*
8. *Summary and Future Directions*

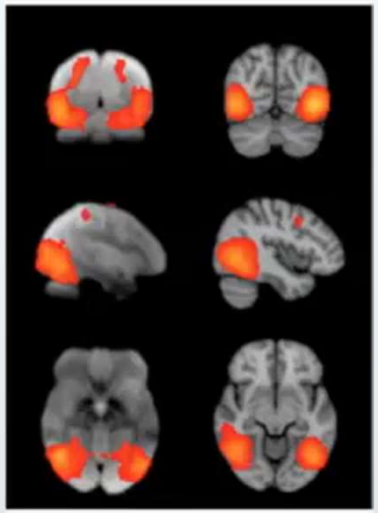
1. Why Medical Imaging?

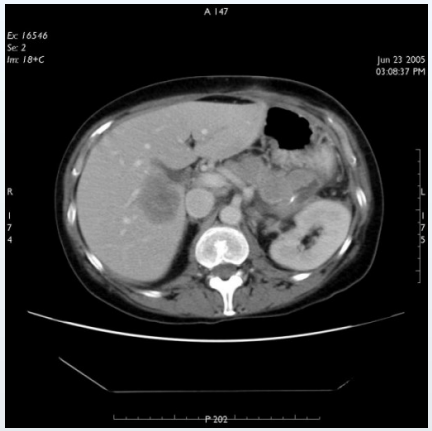
- Medical imaging has revolutionized medical diagnostics in the past 40-50 years.
- Basically – enables medical doctors to view what is happening inside the human body.
- Main advantages: non-invasive, improves medical diagnostics and treatment, imaging during surgical interventions.
- *Nearly every medical test today consists of medical imaging*.



MRI

fMRI

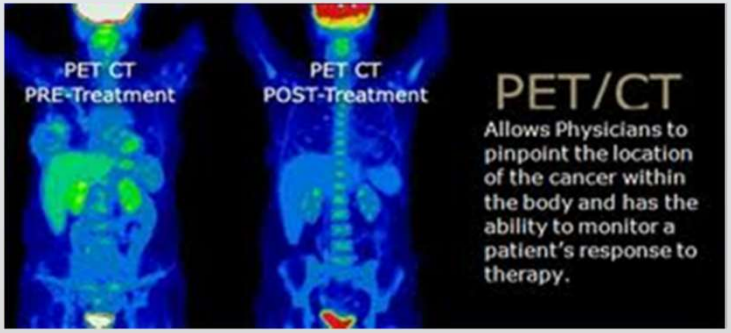




CT(Computed Tomography)

ULTRASOUND

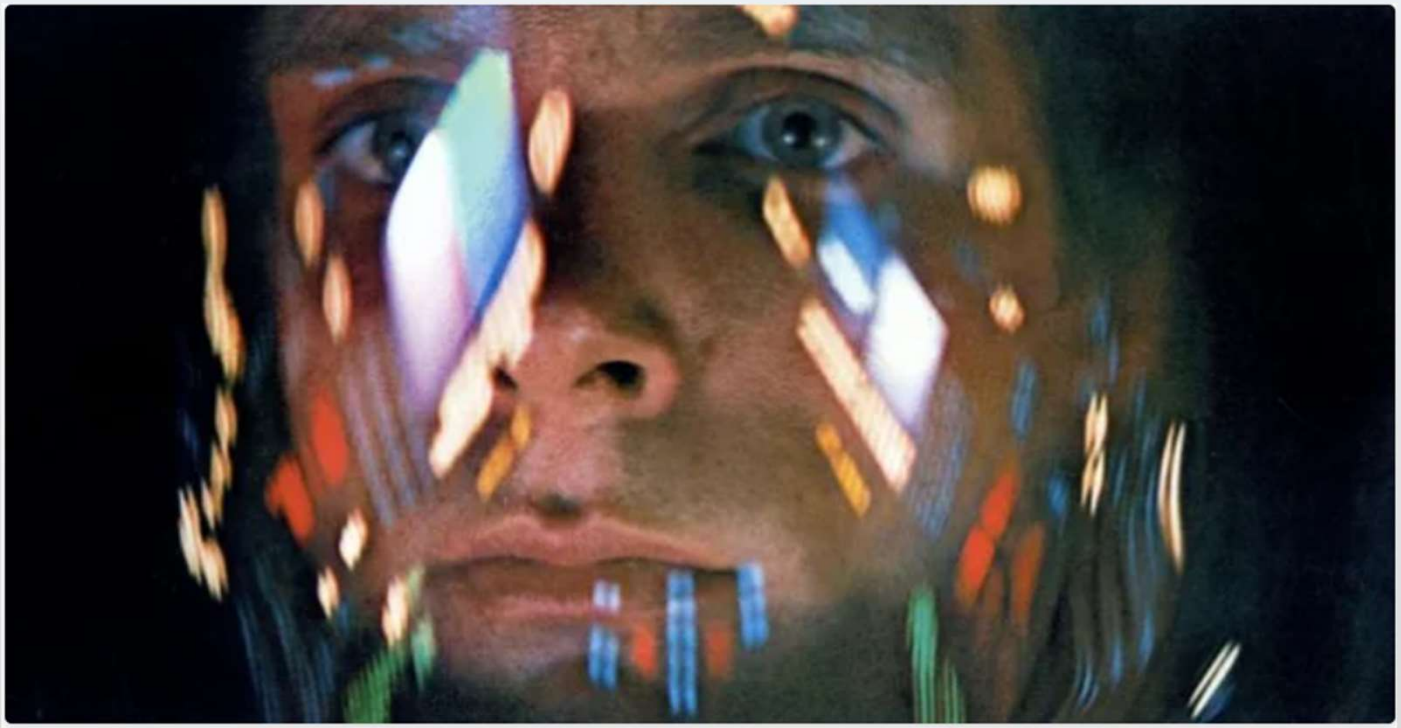
PET-CT



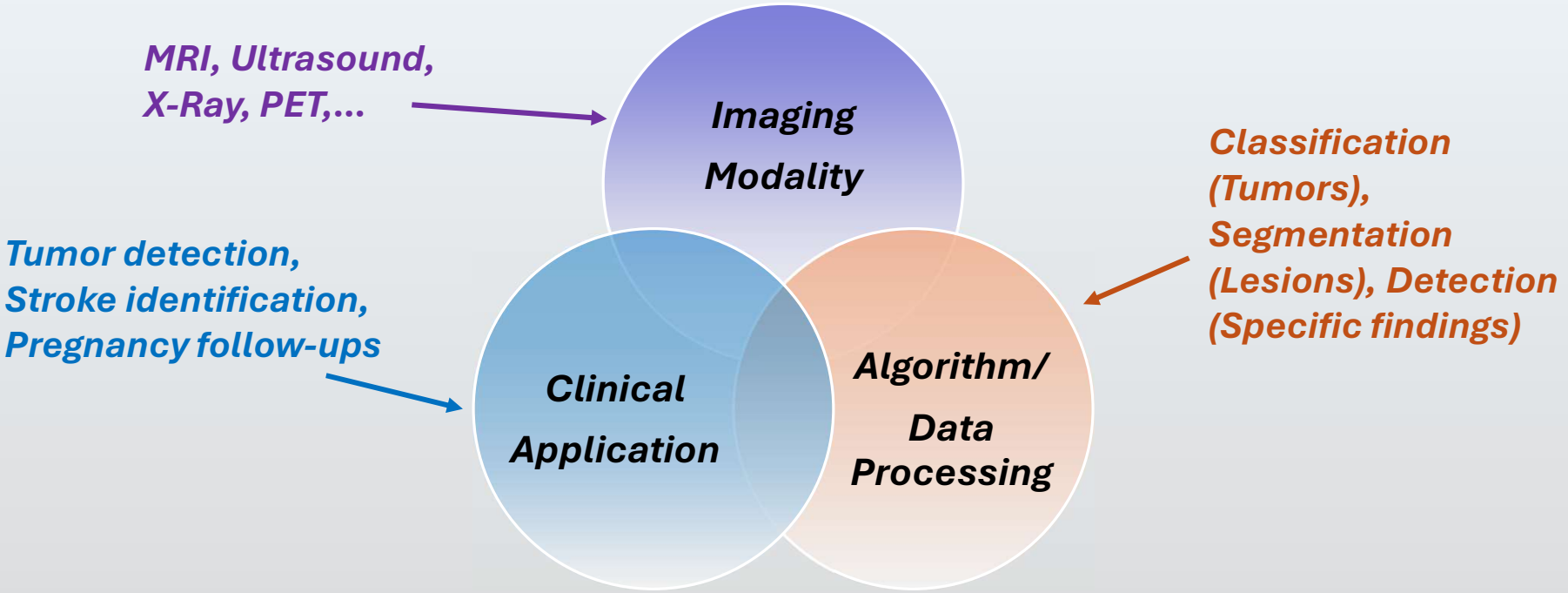
Just a Few Examples…

- Direct relation to reduced infant mortality and increased life expectancy.
- Emergency medicine
- Stroke detection
- Cancer detection and classification
- Detecting abnormalities in the fetus during pregnancy (ultrasound).
- And many more…

My God, it's full of stars...



The 3 Elements of (almost) every Medical Imaging Problem/Task



The Algorithm – ML, DL

- Deep learning - yields state-of-the-art performance in terms of accuracy and computational speed.
- However - challenges remain for deep diagnostic models:

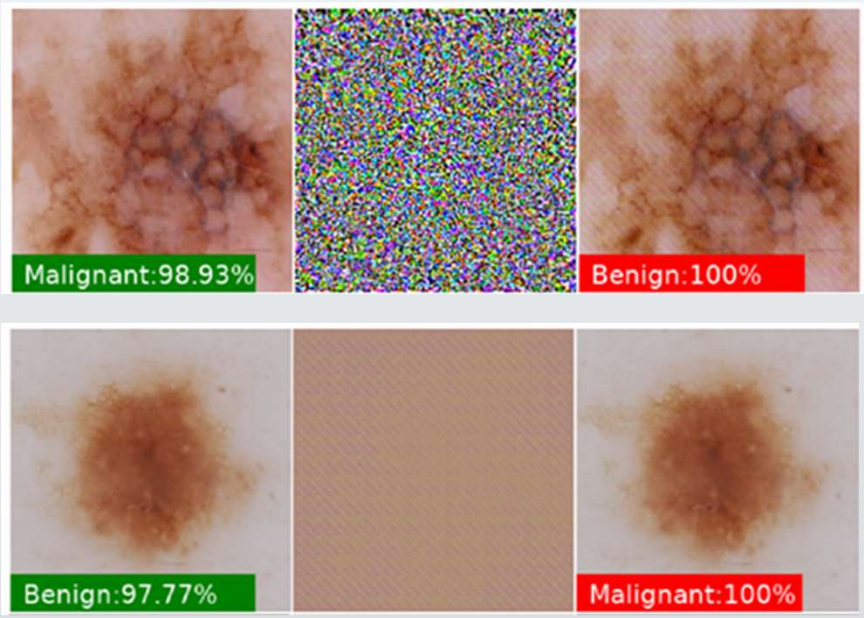
limited reliability and robustness to noise / other unknown perturbations in the data acquisition process

limited generalization to data sets characterized by distributions that differ from that of the training data (OOD)

These challenges stand at the core of the bottleneck that currently exists between theory and practice in terms of incorporating this technology into actual clinical settings.

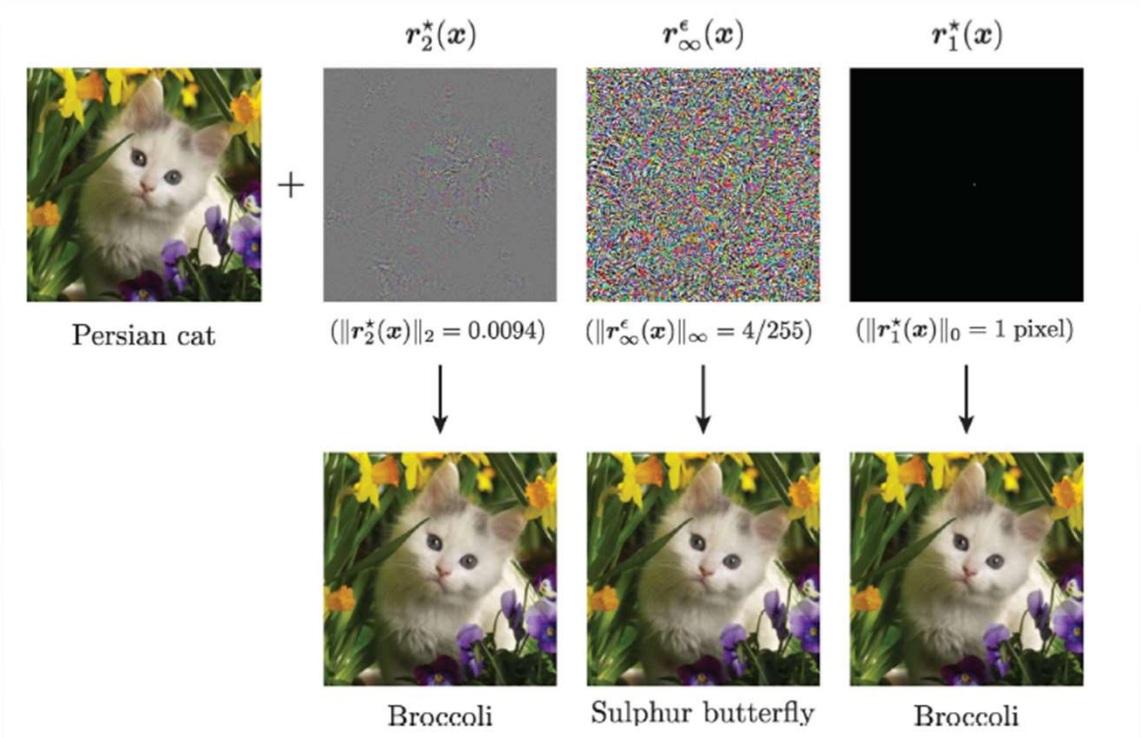
2. Robustness as a Performance Measure

- Typical performance measures - accuracy, specificity, sensitivity.
- However, especially when considering the clinical applicability of diagnostic algorithms, robustness is an additional important performance measure.
- ‘Robustness’ - a measure for a system’s stability, i.e. the extent of the variation that may be applied upon the input whilst obtaining valid predictions at the output.



Xu, M., Zhang, T., Li, Z., Liu, M. and Zhang, D., 2021. Towards evaluating the robustness of deep diagnostic models by adversarial attack. *Medical Image Analysis*, *69*, p.101977.

Adversarial Attacks in Computer Vision



Ortiz-Jiménez, G., Modas, A., Moosavi-Dezfooli, S.M. and Frossard, P., 2021. Optimism in the face of adversity: Understanding and improving deep learning through adversarial robustness. *Proceedings of the IEEE*, *109*(5), pp.635-659.

Adversarial Perturbations, Robustness Definition

- An Adversarial Perturbation $r(x) \in \mathbb{R}^D$ is the solution to

$$\min_{r \in \mathbb{R}^D} Q(r) \quad \text{such that } f_\theta(x+r) \neq f_\theta(x), r \in \mathcal{C}$$

$Q(r)$ - A general objective function

$\mathcal{C}$ – A general set of constraints that characterizes the perturbation

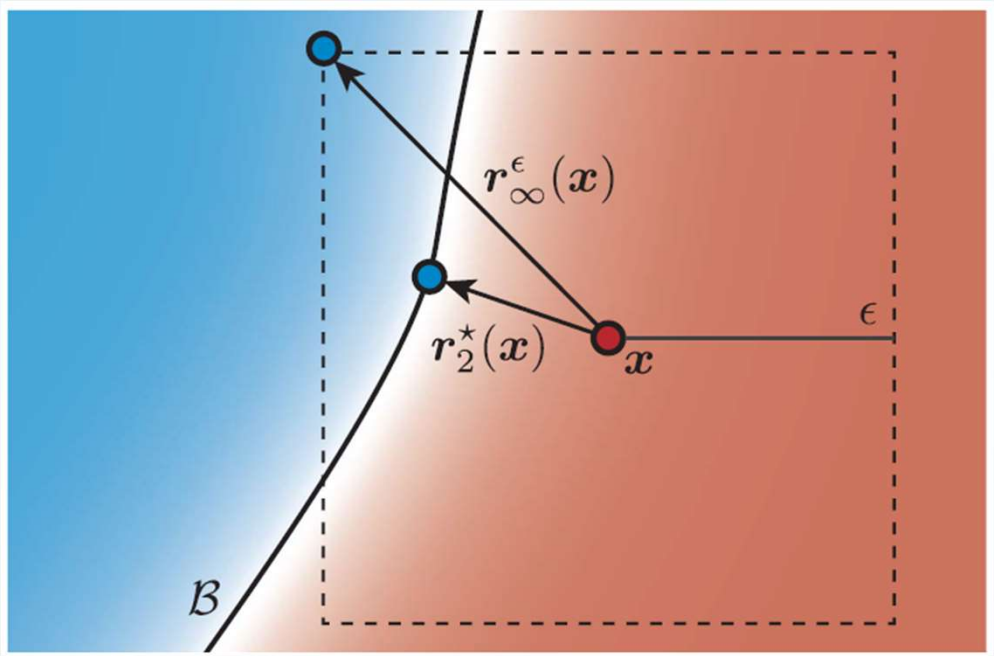
$x + r$ – Adversarial examples.

- Robustness of a DNN – the average distance of any data sample to the decision boundary of the network:

$$\rho_P = \mathbb{E}_{(x,y)\sim\mathcal{D}}[\|r_P(x)\|_P]$$

Ortiz-Jiménez, G., Modas, A., Moosavi-Dezfooli, S.M. and Frossard, P., 2021. Optimism in the face of adversity: Understanding and improving deep learning through adversarial robustness. *Proceedings of the IEEE*, *109*(5), pp.635-659.

Geometric Interpretation – Robustness



Ortiz-Jiménez, G., Modas, A., Moosavi-Dezfooli, S.M. and Frossard, P., 2021. Optimism in the face of adversity: Understanding and improving deep learning through adversarial robustness. *Proceedings of the IEEE*, *109*(5), pp.635-659.

Physics-Induced Deep Learning

- One approach to mitigate the limited robustness in DL is data augmentation.
- However, this may result with overfitting / un-interpretability.
- In recent years - the inclusion of physics information into the learning process and network design.
- Basic assumption: the inclusion of prior domain-related knowledge may shift the learning mechanism such as to produce improved stabilization of the network.

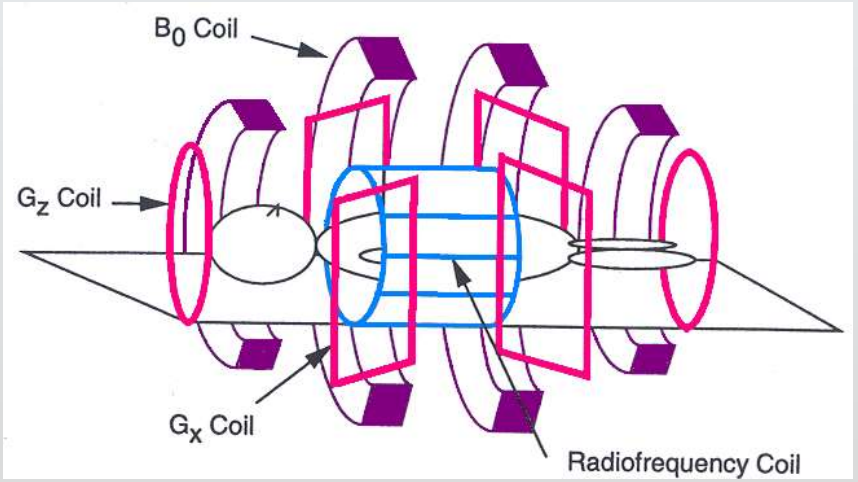
3. MRI Background

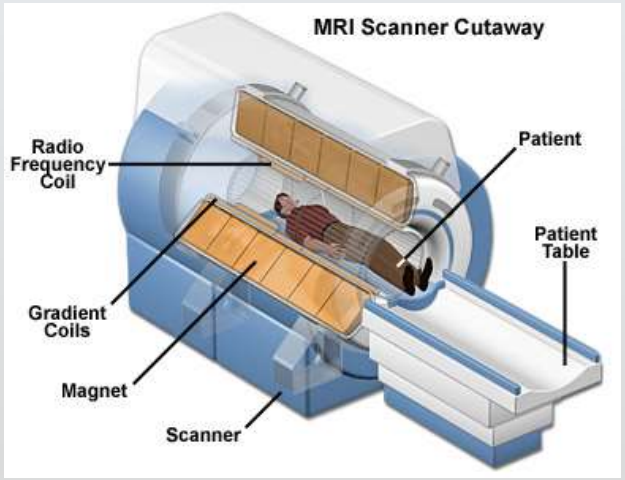
- Non-invasive
- High spatial resolution of tissues and organs
- Superior soft-tissue contrast
- Utilizes a strong magnetic field, radio-frequency pulses and computational algorithms to generate body images – brain, spine, organs, joints,…
- No ionizing radiation

MRI Background

The MRI is comprised of 3 main components:

- A superconducting primary magnet
- 3 magnetic field gradient coils
- RF transmitter and receiver





Taken from:
Lecture Notes by Prof. Guy Gilboa,
Faculty of Electrical and Computer Engineering,
Technion

MRI – T1 and T2 Relaxation Times

- Application of RF pulse creates non-equilibrium state (adding energy to the system).
- After the pulse is switched off, the system is relaxed back to equilibrium.
- There are 2 relaxation times which govern the return to equilibrium:
  - T1 (spin-lattice), equilibrium of z component.
  - T2 (spin-spin), x and y components.

Taken from:
Lecture Notes by Prof. Guy Gilboa,
Faculty of Electrical and
Computer Engineering,
Technion

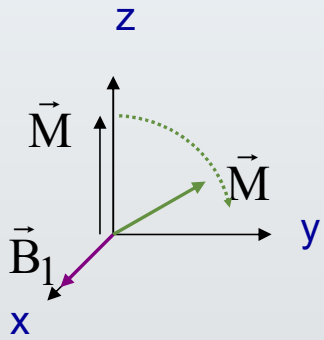
Tissue relaxation times for 1.5 Tesla

Tissue	T_1 (ms)	T_2 (ms)
White matter	790	90
Gray matter	920	100
Liver	500	50
Skeletal muscle	870	60
Lipid	290	160
Cartilage	1060	42

Taken from:
Lecture Notes by Prof. Guy Gilboa,
Faculty of Electrical and Computer Engineering,
Technion

Gradient Field Encoding

- Gradient coils are designed to create an additional B field that varies linearly across the scanner when current is driven into the coil. The slope of linear change is known as the gradient field and is directly proportional to the current driven into the coil.
- X-dimension is encoded by applying a frequency-encoding gradient G_{freq}.
- Protons precess at the frequency $\omega_x = \gamma G_x x$
- Y-dimension is encoded by applying a phase-encoding gradient G_{phase} for a period of τ_{pe} and then switched off.



Based on:
Lecture Notes by Prof. Guy Gilboa,
Faculty of Electrical and Computer Engineering,
Technion

Sensor Space and K-space Formalism

- The imaging is done through gradient coils which create linear spatial variation in the magnetic field in each axis:

$$G_x = \frac{\partial B_x}{\partial x}, \; G_y = \frac{\partial B_y}{\partial y}.$$

- *x* and *y* direction are encoded via frequency and phase.
- For the case of frequency encoded by the *x* direction and phase by the *y* direction the total measurements for a slice can be expressed by the following expression:

$$s(G_y, G_x) \propto \int_{slice} \int_{slice} \rho(x,y) e^{-j\gamma G_y y \tau_p e} e^{-j\gamma G_x x t} dx dy,$$

Based on:
Lecture Notes by Prof. Guy Gilboa,
Faculty of Electrical and Computer Engineering, Technion

K-space formalism – cont.

- Where $\rho(x, y)$ is the number of protons at position (x, y), τ_pe is the phase encoding period and t is the frequency encoding period.
- In the k-space formalism we assign:

$k_x = \frac{\gamma}{2\pi} G_x t, k_y = \frac{\gamma}{2\pi} G_y \tau_{pe}$ to get:

$$S(k_x, k_y) \propto \int_{slice} \rho(x, y) e^{-j2\pi k_x x} e^{-j2\pi k_y y} dx dy$$

Which is the Fourier transform of $\rho(x, y)$.

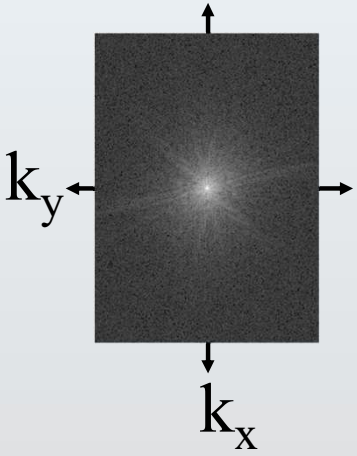
Based on:
Lecture Notes by Prof. Guy Gilboa, Faculty of Electrical and Computer Engineering, Technion

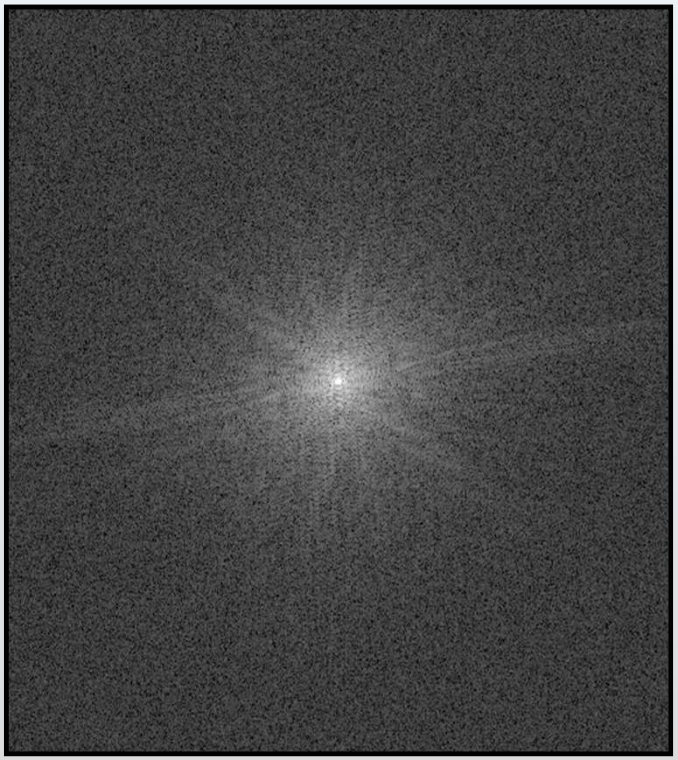
Image recovery

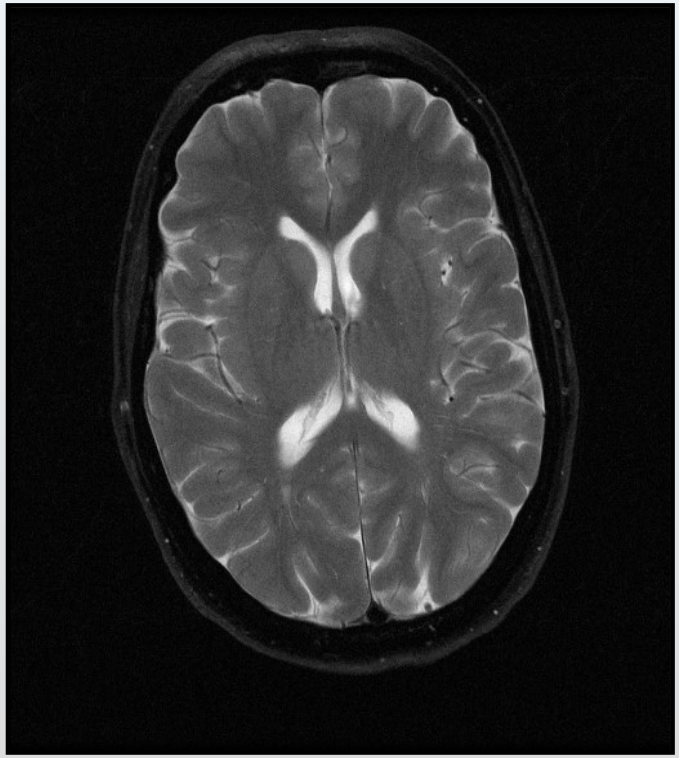
To recover the proton density $\rho(x, y)$, we simply use the inverse Fourier transform:

$$\rho(x,y) \propto \int_{-\infty}^{\infty} S(k_x, k_y) e^{+j2\pi(k_x x + k_y y)} k_x k_y$$

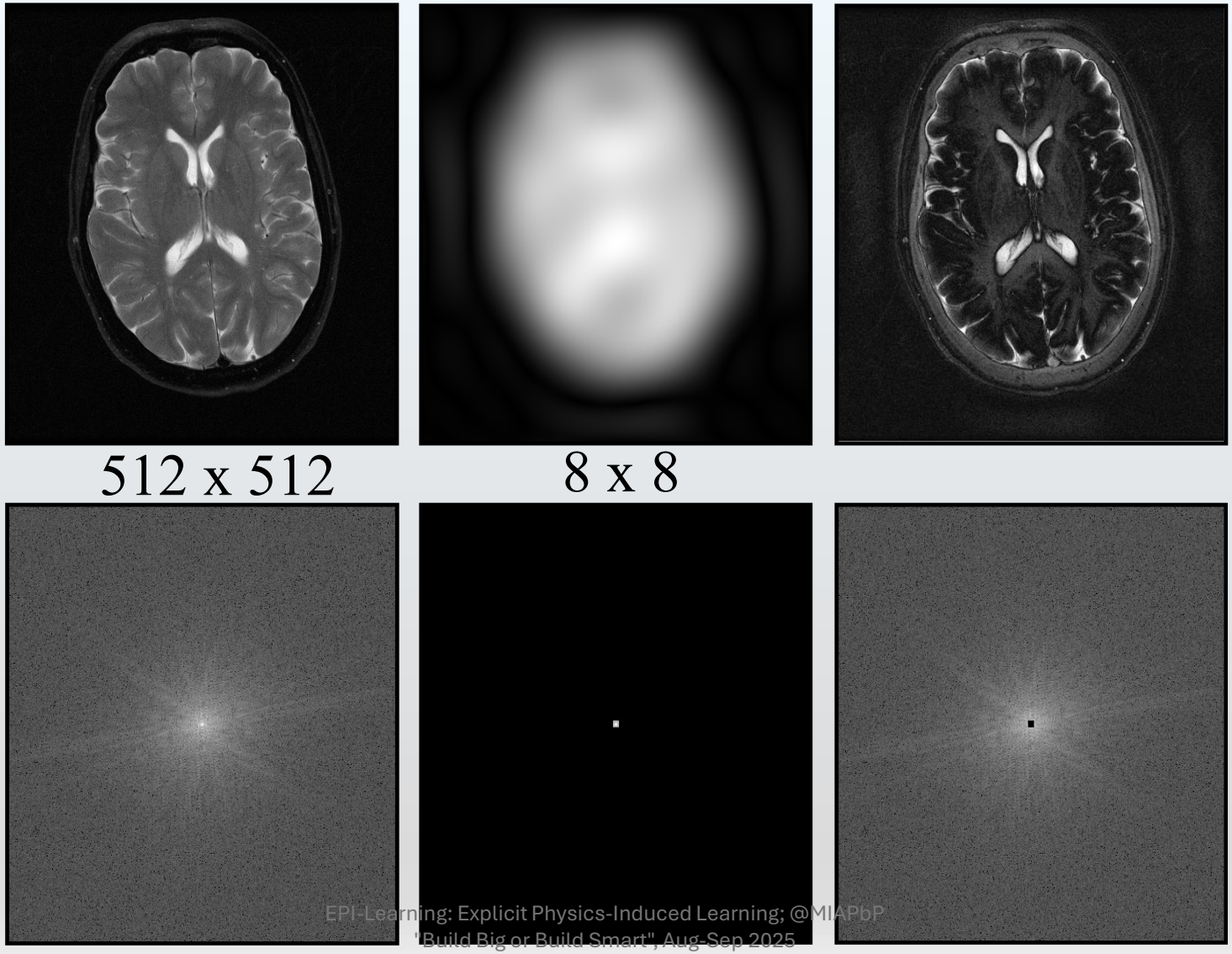
k-Space Representation, Image Space



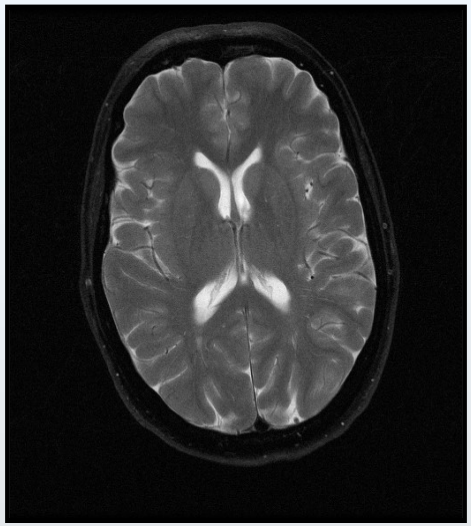
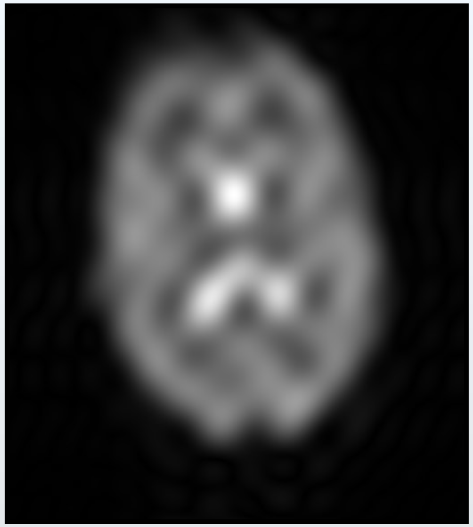
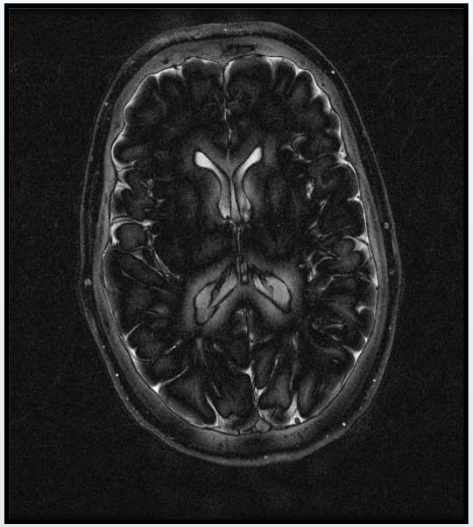




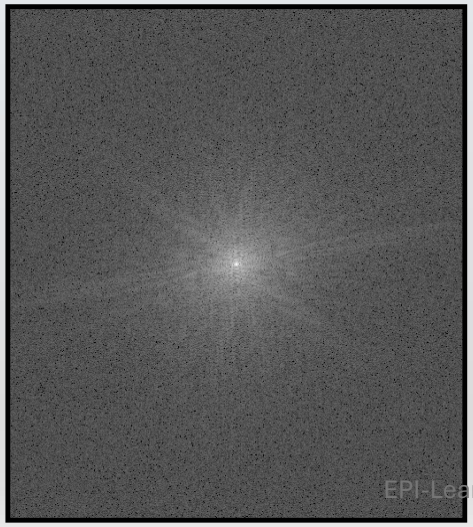
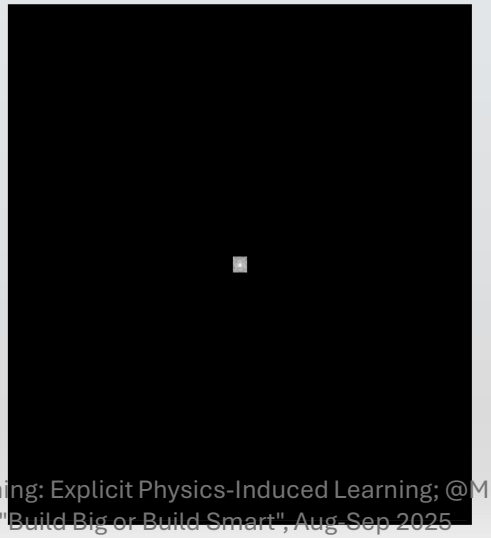
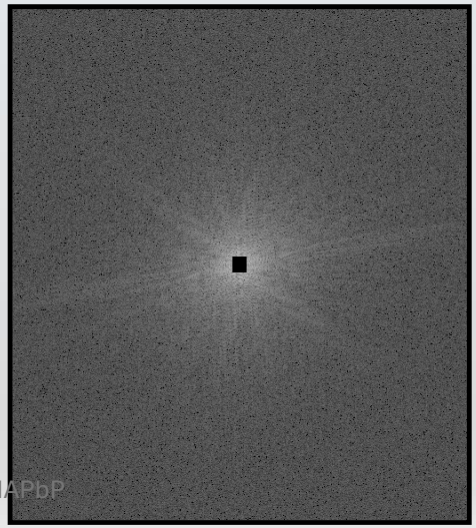
Taken from: Lecture Notes by Prof. Guy Gilboa, Faculty of Electrical and Computer Engineering, Technion



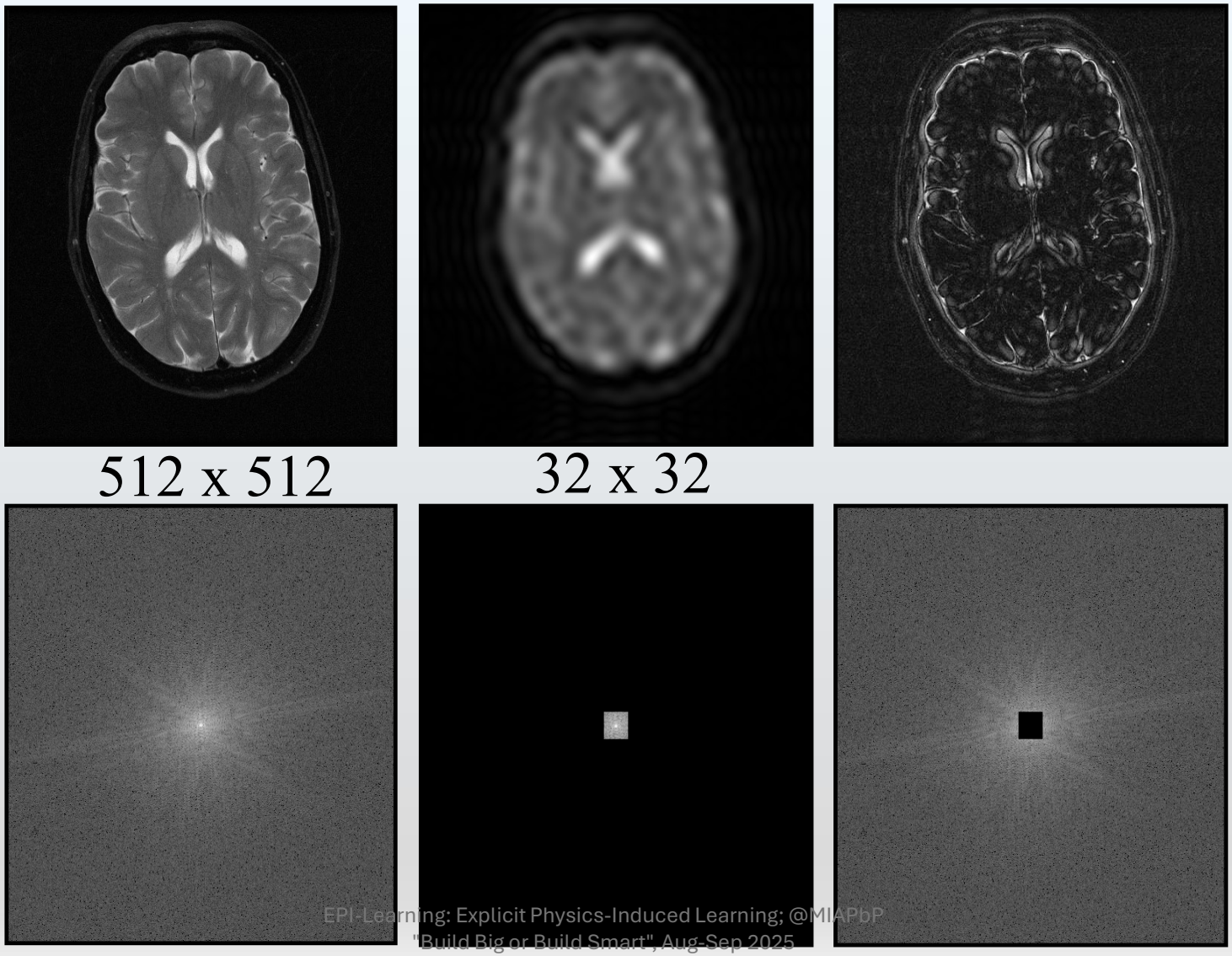
Taken from: Lecture Notes by Prof. Guy Gilboa, Faculty of Electrical and Computer Engineering, Technion

512 x 512 16 x 16

Taken from: Lecture Notes by Prof. Guy Gilboa, Faculty of Electrical and Computer Engineering, Technion



Taken from: Lecture Notes by Prof. Guy Gilboa, Faculty of Electrical and Computer Engineering, Technion

Problem Formulation of MRI Reconstruction

- MRI reconstruction may be viewed as a sensor-to-image translation task.
- Goal: recover an image x from the k-space signal y, which is corrupted by measurement noise n:

$$y = Ax + n$$

where: $A \in \mathbb{C}^{K \times N}$ a linear forward operator describing the MR acquisition model;

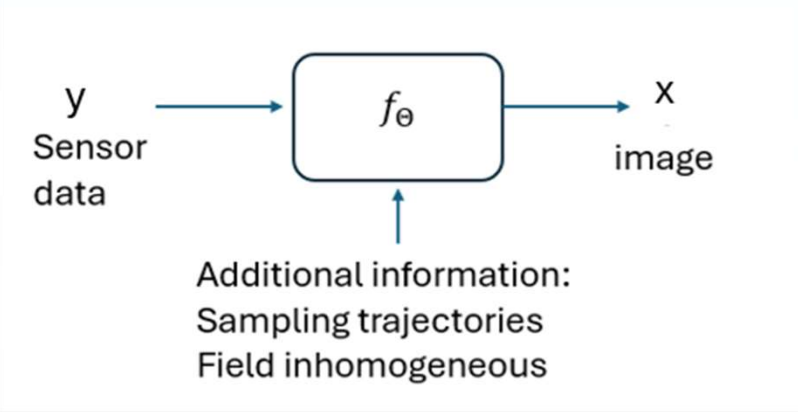
K - number of measurements, i.e. the dimensionality of the underlying k-space data;

N – number of image voxels.

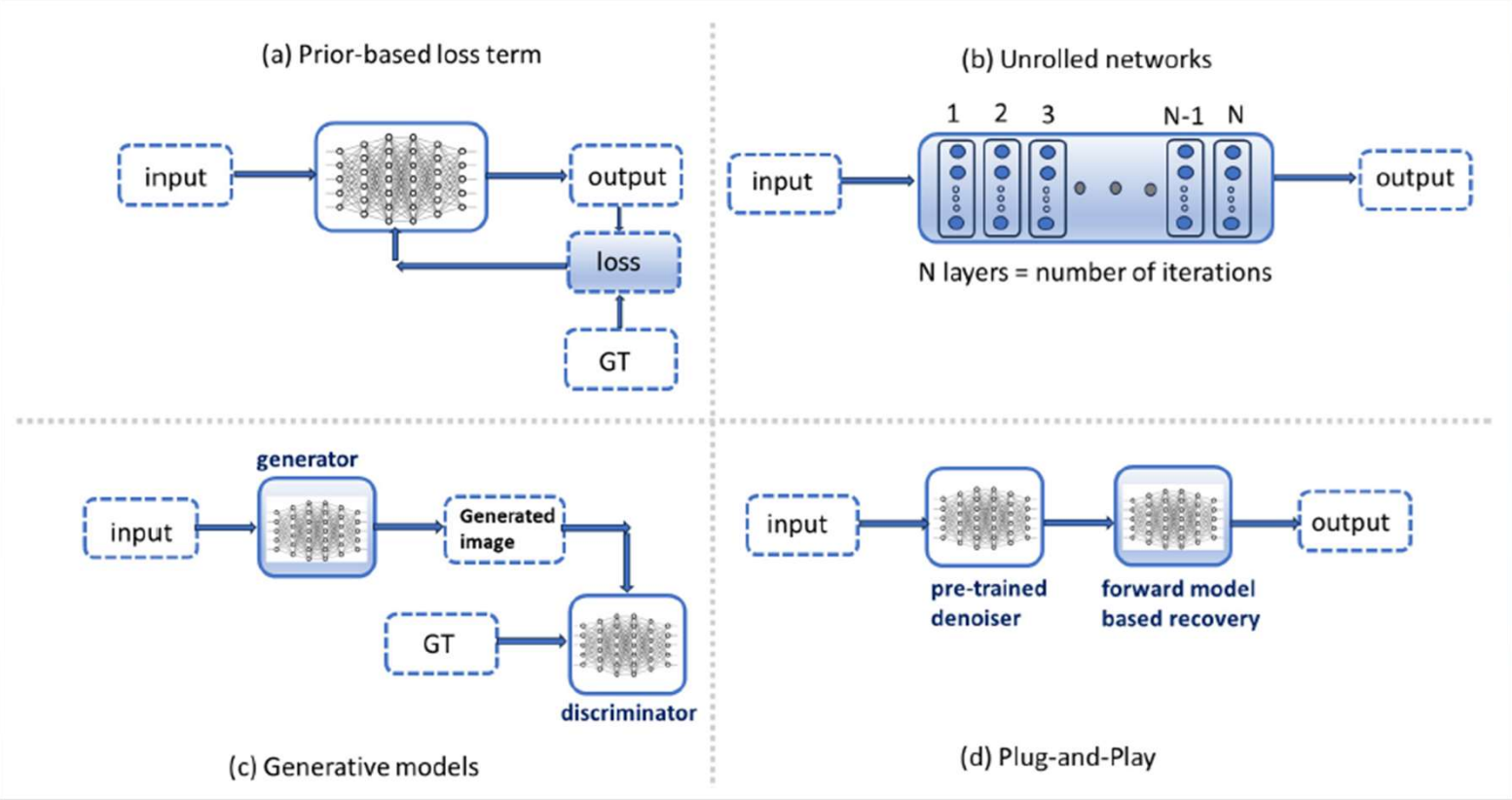
Problem Formulation of MRI Reconst. – cont.

- Depending on the imaging application and signal modelling, the operator A involves Fourier transformations, sampling trajectories, field inhomogeneities, etc.
- Objective: Learning the sensor-to-image mapping function

$f_\Theta: y \rightarrow x$, with learning parameters Θ.



4. Physics-based Approaches



Nemirovsky-Rotman, S. and Bercovich, E., 2024. Explicit physics-informed deep learning for computer-aided diagnostic tasks in medical imaging. *Machine Learning and Knowledge Extraction*, *6*(1), pp.385-401.

PI-DL: From Implicit to Explicit

- For many physics-induced learning methods, a known (or partially known) forward model is integrated into the learning process.
- Typically, for such methods, the forward model is ***implicitly*** integrated into the learning mechanism, either by applying its expression as part of the loss term, or by encapsulating it as part of the architecture design process.
- What if we ***explicitly*** incorporated this prior knowledge into the network architecture?

5. EPI-Learning – Problem Formulation

- Assume a known physical model of the form: input data $\boldsymbol{X}$, output data $\boldsymbol{Y}$, the forward operator $\boldsymbol{f}$, and noise $\boldsymbol{N}$ may be represented as:

$$\boldsymbol{Y} = \boldsymbol{f}(\boldsymbol{X}; \boldsymbol{I}) + \boldsymbol{N}$$

$\boldsymbol{I}$ is the prior of the physical coordinates that determine the form of $\boldsymbol{f}$

- We propose a solution based on a DNN architecture $\boldsymbol{g}$, whose weights are denoted by the parametrization $\boldsymbol{\Theta}$:

$$\widehat{\boldsymbol{X}} = \boldsymbol{g}_{\boldsymbol{\Theta}}^{\boldsymbol{I}}(\boldsymbol{Y}), \boldsymbol{\Theta} = \underset{\boldsymbol{\Theta}}{\mathbf{argmin}} \left\| \boldsymbol{X} - \widehat{\boldsymbol{X}} \right\| = \underset{\boldsymbol{\Theta}}{\mathbf{argmin}} \left\| \boldsymbol{X} - \boldsymbol{g}_{\boldsymbol{\Theta}}^{\boldsymbol{I}}(\boldsymbol{Y}) \right\|$$

Nemirovsky-Rotman, S. and Bercovich, E., 2024. Explicit physics-informed deep learning for computer-aided diagnostic tasks in medical imaging. *Machine Learning and Knowledge Extraction*, *6*(1), pp.385-401.

EPI-Learning – Novelty 1: Extended Architecture

- We extend the network's architecture $\boldsymbol{g}$ by *explicitly* feeding it with the physics prior coordinates $\boldsymbol{I}$ as input. We denote this new architecture by $\boldsymbol{g}_{\boldsymbol{\Theta}^{\boldsymbol{I_E}}}^{\boldsymbol{I_E}}$, where the index $\boldsymbol{E}$ stands for 'explicit'. The solution based on this architecture is:

$$\widehat{\boldsymbol{X}} = \boldsymbol{g}_{\boldsymbol{\Theta}^{\boldsymbol{I_E}}}^{\boldsymbol{I_E}}(\boldsymbol{Y};\boldsymbol{I})$$

$$\boldsymbol{\Theta}^{\boldsymbol{I_E}} = \underset{\boldsymbol{\Theta}^{\boldsymbol{I_E}}}{\mathbf{argmin}} \left\| \boldsymbol{X} - \widehat{\boldsymbol{X}} \right\| = \underset{\boldsymbol{\Theta}^{\boldsymbol{I_E}}}{\mathbf{argmin}} \left\| \boldsymbol{X} - \boldsymbol{g}_{\boldsymbol{\Theta}^{\boldsymbol{I_E}}}^{\boldsymbol{I_E}}(\boldsymbol{Y};\boldsymbol{I}) \right\|.$$

Nemirovsky-Rotman, S. and Bercovich, E., 2024. Explicit physics-informed deep learning for computer-aided diagnostic tasks in medical imaging. *Machine Learning and Knowledge Extraction*, *6*(1), pp.385-401.

EPI-Learning – Novelty 2: Modified Training Set

- We modify the data set for training the extended network architecture $\boldsymbol{g}_{\boldsymbol{\Theta}^{\boldsymbol{I_E}}}^{\boldsymbol{I_E}}$: for each input $\boldsymbol{X}$, we provide the network with the explicit prior information of the forward model $\boldsymbol{I_E}$, for which we add a (known) perturbation $\boldsymbol{\delta I_E}$.
- The perturbation is chosen from a set of perturbations that are distributed according to some chosen probability distribution: $\boldsymbol{\delta I_E \sim \mathcal{D}_{\delta I_E}}$.

Nemirovsky-Rotman, S. and Bercovich, E., 2024. Explicit physics-informed deep learning for computer-aided diagnostic tasks in medical imaging. *Machine Learning and Knowledge Extraction*, *6*(1), pp.385-401.

EPI-Learning – Novelty 2: Modified Training Set

- Therefore, the data set for training the extended network is $\left(\boldsymbol{X}_{\boldsymbol{I_E}+\boldsymbol{\delta I_E}}, \boldsymbol{Y}_{\boldsymbol{I_E}+\boldsymbol{\delta I_E}}\right)$, where:

$$\boldsymbol{Y}_{\boldsymbol{I_E}+\boldsymbol{\delta I_E}} = \boldsymbol{f}\left(\boldsymbol{X}_{\boldsymbol{I_E}+\boldsymbol{\delta I_E}}; \boldsymbol{I_E} + \boldsymbol{\delta I_E}\right) + \boldsymbol{n}\ .$$

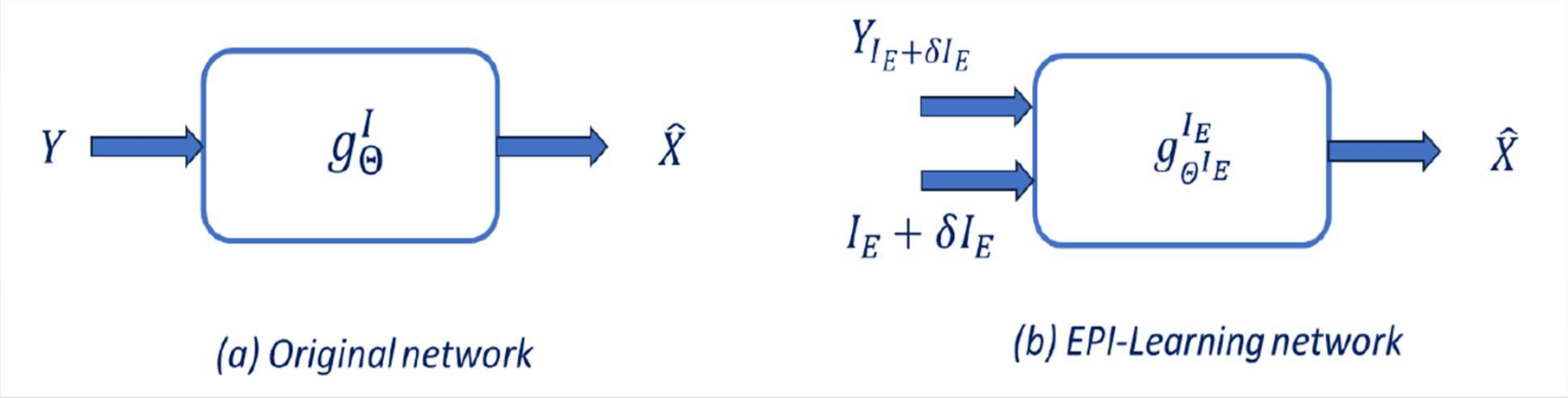
- The loss term for the extended network (namely: the *generalization loss*) is thus given by:

$$\mathcal{L}_{\boldsymbol{I_E}}^{\mathcal{D}_{\boldsymbol{\delta I_E}}} = \left\| \boldsymbol{Y}_{\boldsymbol{I_E}+\boldsymbol{\delta I_E}} - \boldsymbol{f}\left(\widehat{\boldsymbol{X}}; \boldsymbol{I_E} + \boldsymbol{\delta I_E}\right) \right\|$$

Nemirovsky-Rotman, S. and Bercovich, E., 2024. Explicit physics-informed deep learning for computer-aided diagnostic tasks in medical imaging. *Machine Learning and Knowledge Extraction*, *6*(1), pp.385-401.

EPI-Learning

- Training of the extended network $g_{\Theta^{I_E}}^{I_E}$ consists of this new data set $\left(X_{I_E+\delta I_E}, Y_{I_E+\delta I_E}\right)$.
- The input to the extended network consists of *both* $Y_{I_E+\delta I_E}$ *and its associated perturbed explicit prior information* $I_E + \delta I_E$.



Nemirovsky-Rotman, S. and Bercovich, E., 2024. Explicit physics-informed deep learning for computer-aided diagnostic tasks in medical imaging. *Machine Learning and Knowledge Extraction*, *6*(1), pp.385-401.

6. EPI-Learning for a Non-Linear Regression Task

- We define the following regression task: Let

$$\boldsymbol{Y} = \boldsymbol{f}(\boldsymbol{X}; \boldsymbol{I}) + \boldsymbol{N}$$

where: $\boldsymbol{Y} = \{\boldsymbol{y_i}\}_{\boldsymbol{i=1}}^{\boldsymbol{M}}, \boldsymbol{Y} \in \mathbb{R}^{\boldsymbol{M}}$ a set of $\boldsymbol{M}$ measurements,

- $\boldsymbol{I} = \{\boldsymbol{I_i}\}_{\boldsymbol{i=1}}^{\boldsymbol{K}}$, $\boldsymbol{I} \in \mathbb{R}^{\boldsymbol{K}}$ a set of explicit given physics prior coordinates,
- $\boldsymbol{X} = \{\boldsymbol{x_i}\}_{\boldsymbol{i=1}}^{\boldsymbol{K}}, \boldsymbol{X} \in \mathbb{R}^{\boldsymbol{K}}$ a set of un-known model parameters,
- $\boldsymbol{f}$: $\mathbb{R}^{\boldsymbol{K}} \rightarrow \mathbb{R}^{\boldsymbol{M}}$ a presumably known non-linear function.
- $\boldsymbol{N} = \{\boldsymbol{n}\}_{\boldsymbol{i=1}}^{\boldsymbol{M}}$, $\boldsymbol{n} \in \mathbb{R}^{\boldsymbol{M}}$ is an additive random noise vector of length $\boldsymbol{M}$.

Nemirovsky-Rotman, S. and Bercovich, E., 2024. Explicit physics-informed deep learning for computer-aided diagnostic tasks in medical imaging. *Machine Learning and Knowledge Extraction*, *6*(1), pp.385-401.

6. EPI-Learning for a Non-Linear Regression Task

- Goal: find an optimal estimation $\widehat{\boldsymbol{X}}$ of the model parameters $\boldsymbol{X}$, given the input measurements $\boldsymbol{Y}$ and the prior physics coordinates $\boldsymbol{I}$.
- We focus our interest on the highly non-linear model of a sum of exponentials:

$$\boldsymbol{s_m(a_1, a_2, \dots, a_L, c_1, c_2, \dots, c_L) = \sum_{l=1}^{L} a_l exp(-b_m c_l) + n} \quad ,$$

$$\boldsymbol{m = 1, \dots, M}\ , \quad \boldsymbol{M \in \mathbb{N}}$$

where $\boldsymbol{\{a_l, c_l\}_{l=1}^{L}}$ a set of parameters to be estimated from the samples $\boldsymbol{s_m(a_1, a_2, \dots, a_L, c_1, c_2, \dots, c_L)}$;

$\boldsymbol{a_l} \in \mathbb{R}$ and $\boldsymbol{0 < a_l < 1, \sum_{l=1}^{L} a_l = 1}$;

$\boldsymbol{c_l} \in \mathbb{R}$ such that $\boldsymbol{0 < c_l}$.

$\boldsymbol{b_m} \in \mathbb{R}, \boldsymbol{b_m > 0}$ are pre-chosen and known constant parameters for which the samples $\boldsymbol{s_m(a_1, a_2, \dots, a_L, c_1, c_2, \dots, c_L)}$ are evaluated

Nemirovsky-Rotman, S. and Bercovich, E., 2024. Explicit physics-informed deep learning for computer-aided diagnostic tasks in medical imaging. *Machine Learning and Knowledge Extraction*, 6(1), pp.385-401.

EPI-Learning for a Non-Linear Regression Task

- A sum of exponentials is useful for many applications.
- Diffusion-Weighted MRI - the two-compartment IVIM model ($\boldsymbol{L = 2}$) or the three-compartment NODDI model ($\boldsymbol{L = 3}$).
- For the case of the IVIM model, the forward model may be written as

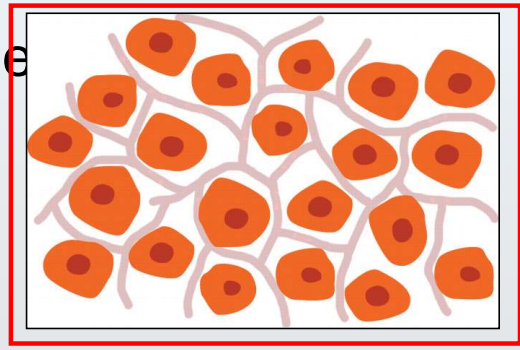
$$\boldsymbol{s_i = s_0\left[F_p \exp\left(-b_i D_p\right) + \left(-F_p\right)\exp\left(-b_i D\right)\right]}$$

Diffusion-Weighted MRI (DW-MRI) and the Intra-Voxel Incoherent Motion (IVIM) Model

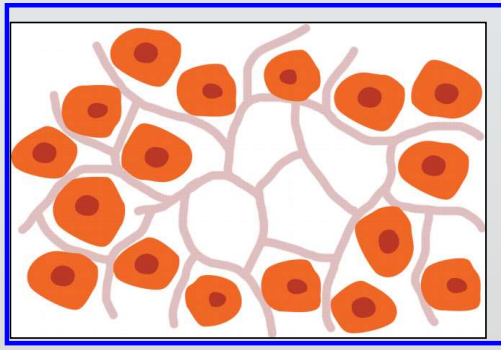
- A Bio-physical model that provides characterization of tissue cellular architecture by diffusion and perfusion parameters.
- ***The bi-exponential model****:

$$S_i = S_0\left(F_p\, e^{-b_i(D_p+D)} + (1-F_p)e^{-b_i D}\right)$$

- S_i — Attenuated signal
- S_0 — Reference signal
- F_p — Fraction of pseudo-diffusion contribution
- b_i — acquisition parameters
- D_p — Pseudo-diffusion coefficient – micro-circulation of blood in blood-vessels (intravascular perfusion)
- D — Diffusion coefficient – diffusion of water molecules in tissues



Restricted diffusion (low D valu



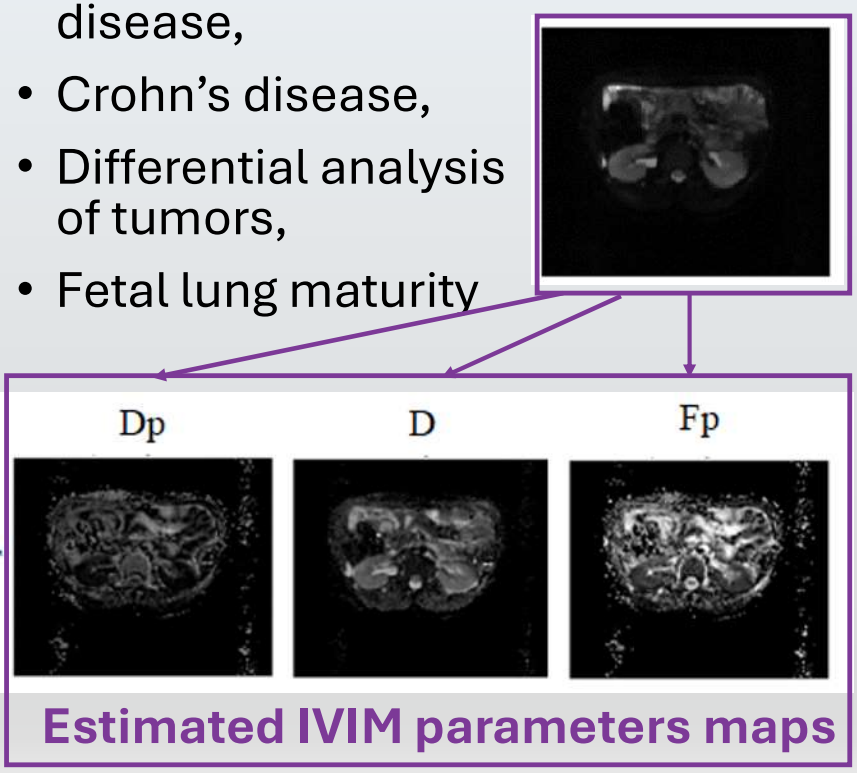
Unrestricted diffusion (high D valu

* Denis Le Bihan *el al,* "MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders," Radiology, vol. 161, no. 2, pp. 401–407,1986

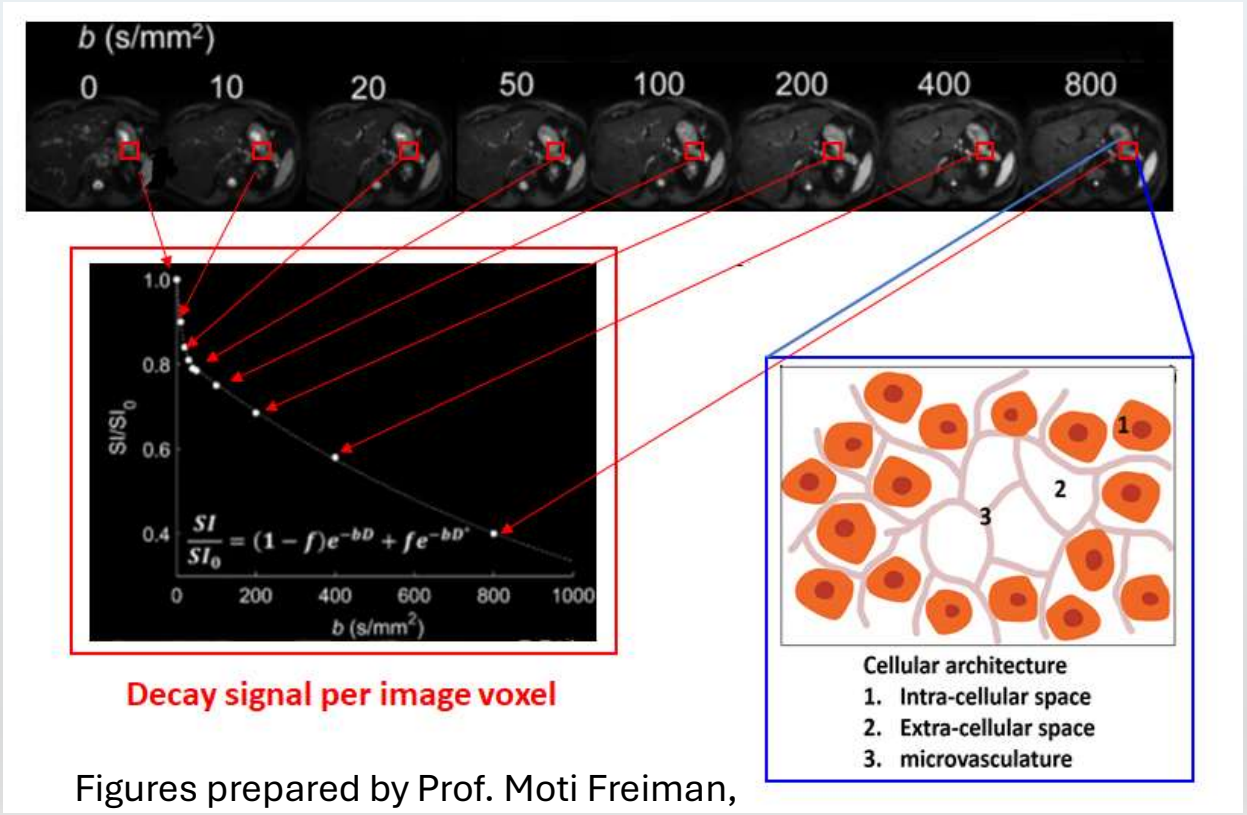
Figures prepared by Prof. Moti Freima
Faculty of Bio-Medical Engineering, T

Clinical Applications of IVIM Modelling

- Characterization of inflammation in chronic liver disease,
- Crohn's disease,
- Differential analysis of tumors,
- Fetal lung maturity



Estimated IVIM parameters maps

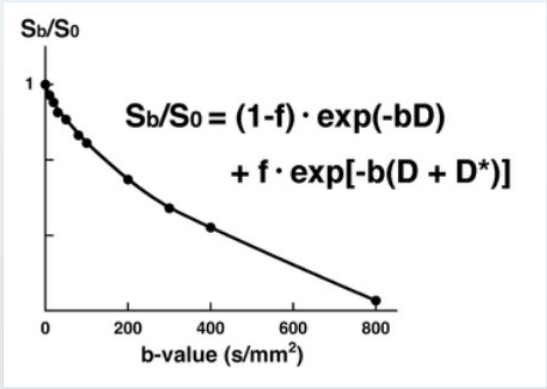


Figures prepared by Prof. Moti Freiman,
Faculty of Bio-Medical Engineering, Technion

Deep Neural Networks vs. Classical Solvers – IVIM Parameter Estimations

(I) Classical solvers: Non-Linear Least Squares (NLLS), Bayesian Shrinkage Prior (BSP).

Pros: Solver is agnostic to acquisition protocol.



$$\widehat{\Theta_v} = Arg\min_{\Theta_v} \sum_{i=1}^{N} \left(m_i^v - s_i^{\Theta_v}\right)^2$$

$$\Theta_v \equiv \left(D^v, D_p^v, F_p^v\right)$$

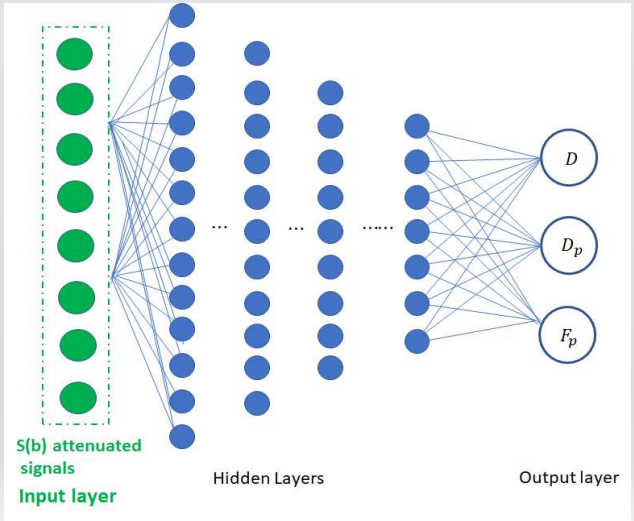
m_i^v Observed IVIM signal

$s_i^{\Theta_v}$ Forward-model IVIM signal

(II) Deep Neural Network (DNN) solvers (IVIM-NET; Barbieri, 2020):

Pros: more accurate and faster.

$$\hat{\Phi} = \underset{\Phi}{\text{Argmin}} \sum_{k=1}^{K} \sum_{i=1}^{N} \left\| f_{g_\Phi(\vec{s}^k)}\left(s_0^k, b_i\right) - s_i^k \right\|^2$$



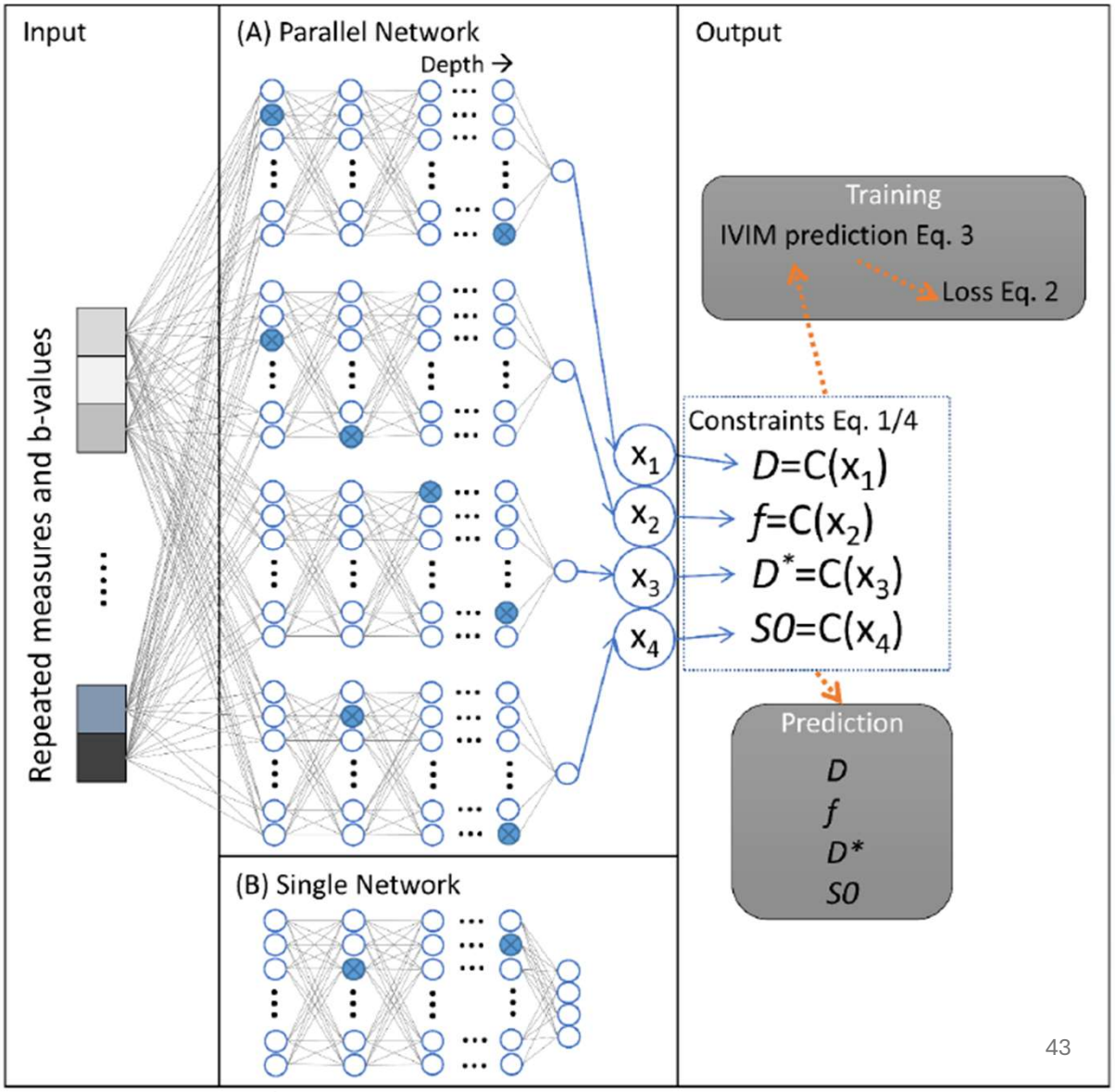
Barbieri et al. (2020). Deep learning how to fit an intravoxel

IVIM Model – Diffusion MRI

$$L = \frac{1}{|B|}\sum_{b\in B}\left(\frac{S(b)}{S(b=0)} - S_{net}(b)\right)^2$$

$$S_{net}(b) = fe^{-bD^*} + (1-f)e^{-bD}$$

Kaandorp, M. B.-C. (2021). Improved unsupervised physics-informed deep learning for intravoxel incoherent motion modeling and evaluation in pancreatic cancer patients. *Magnetic resonance in medicine*, 2250-2265



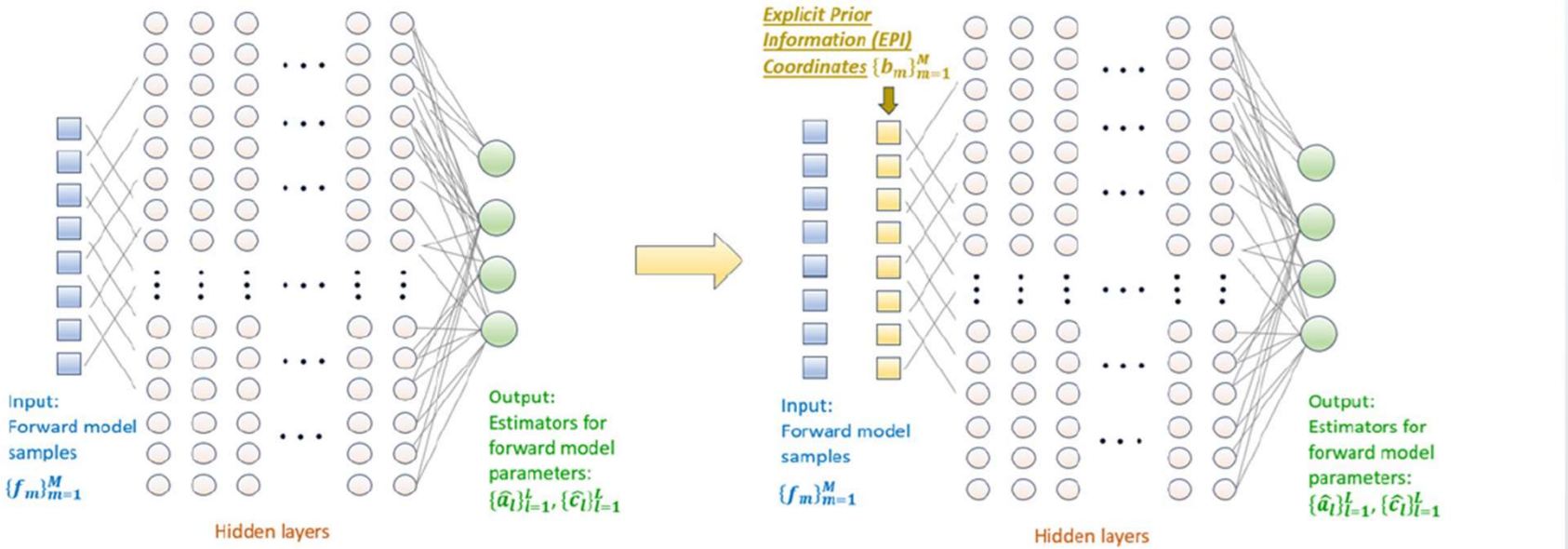
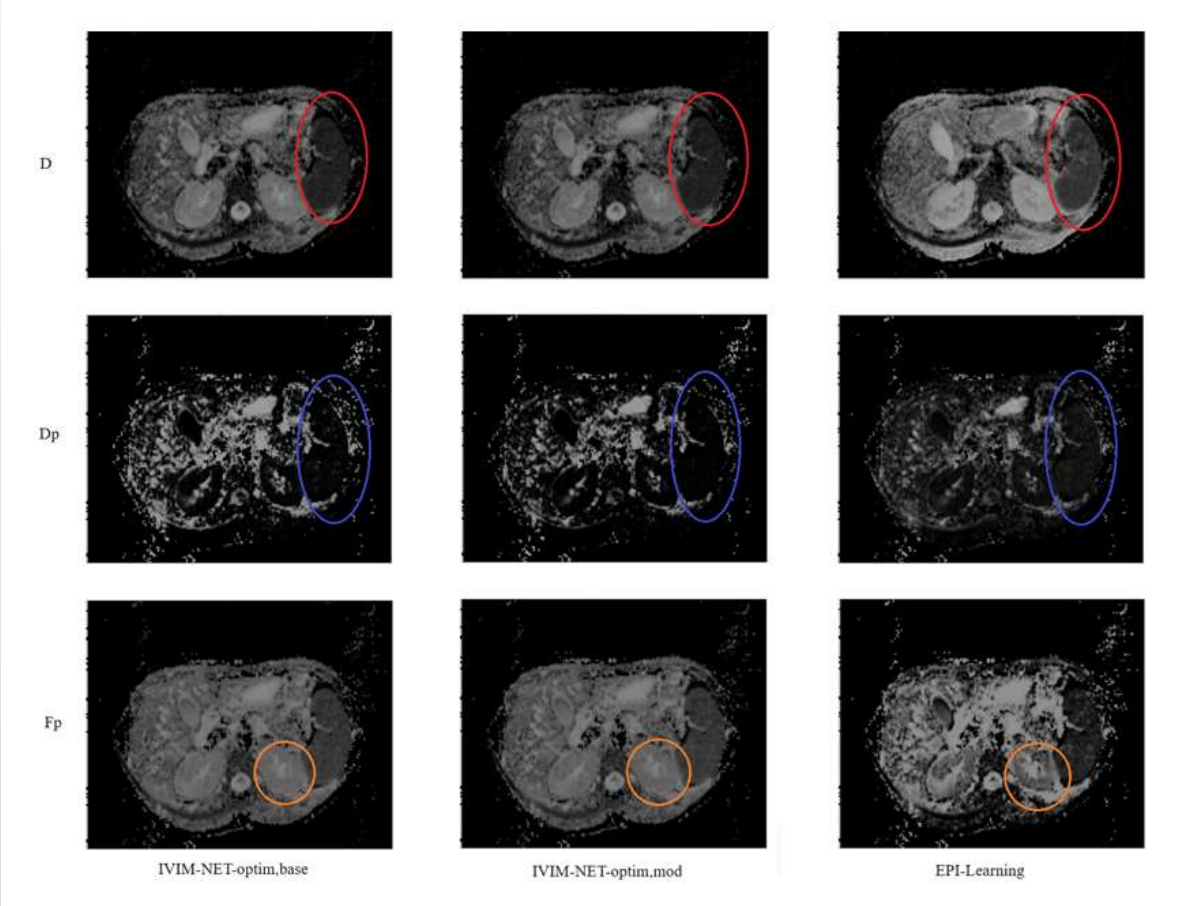


Figure 3. Extending a basic network architecture with the explicit prior information coordinates (EPI-Learning).

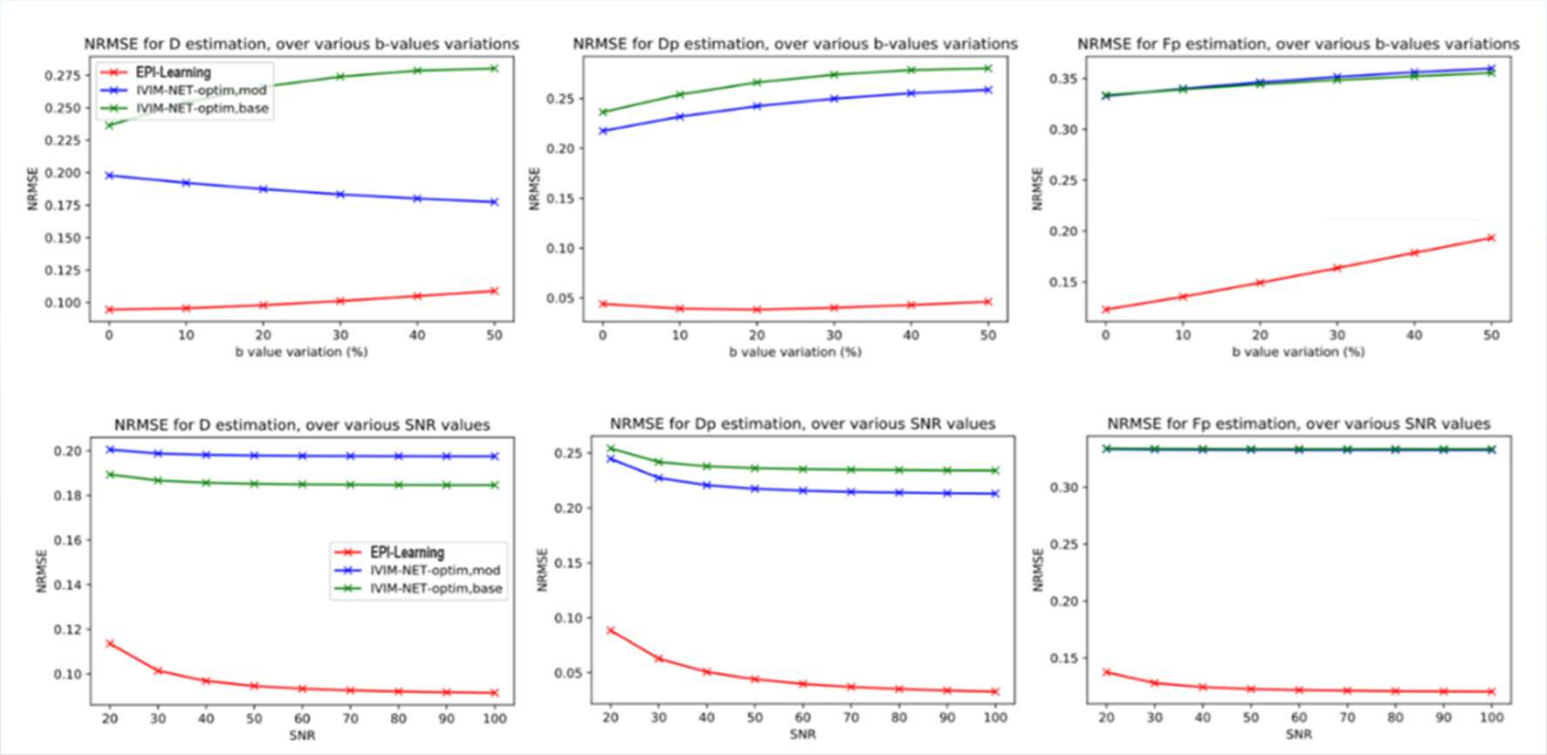
Nemirovsky-Rotman, S. and Bercovich, E., 2024. Explicit physics-informed deep learning for computer-aided diagnostic tasks in medical imaging. *Machine Learning and Knowledge Extraction*, *6*(1), pp.385-401.

EPI-Learning Results vs. Original DNN



Nemirovsky-Rotman, Shira, Elad Rotman, Onur Afacan, Sila Kurugol, Simon Warfield, and Moti Freiman. "Physically Motivated Deep-Neural Networks of the Intravoxel Incoherent Motion Signal Decay Model for Quantitative Diffusion-Weighted MRI." In *Proceedings of the Annual Conference of the ISMRM (International Society of Magnetic Resonance in Medicine, Virtual, 15–20 May 2021*.

EPI-Learning Results vs. Original DNN



GenLeAP: Generalized Learning based on Incorporation of the Acquisition Parameters for the Intra-Voxel Incoherent Motion Model in Diffusion-Weighted MRI, in prep. Joint work with Prof. Moti Freiman, Technion.

Working Hypotheses – EPI-Learning

Hypothesis 1: Improved *generalization* to input distribution:

- The network $\boldsymbol{g}_{\boldsymbol{\Theta}^{\boldsymbol{I_E}}}^{\boldsymbol{I_E}}$ is expected to have improved generalization compared to $\boldsymbol{g}_{\boldsymbol{\Theta}}$:

$$\mathbb{E}_{\boldsymbol{Y}\sim\boldsymbol{\mathcal{D}}_{\boldsymbol{Y}}^{*}}\left\{\left\|\boldsymbol{X}-\boldsymbol{g}_{\boldsymbol{\Theta}^{\boldsymbol{I_E}}}^{\boldsymbol{I_E}}(\boldsymbol{Y};\boldsymbol{I})\right\|\right\}\leq\mathbb{E}_{\boldsymbol{Y}\sim\boldsymbol{\mathcal{D}}_{\boldsymbol{Y}}^{*}}\left\{\left\|\boldsymbol{X}-\boldsymbol{g}_{\boldsymbol{\Theta}}^{\boldsymbol{I}}(\boldsymbol{Y})\right\|\right\}$$

- where the expectation operator $\mathbb{E}(\cdot)$ is taken over $\boldsymbol{Y}$ that is distributed according to $\boldsymbol{\mathcal{D}}_{\boldsymbol{Y}}^{*}$, a shifted distribution with respect to the original distribution $\boldsymbol{\mathcal{D}}_{\boldsymbol{Y}}$ (according to which the network $\boldsymbol{g}_{\boldsymbol{\Theta}}^{\boldsymbol{I}}$ was trained): $\boldsymbol{\mathcal{D}}_{\boldsymbol{Y}}^{*}\neq\boldsymbol{\mathcal{D}}_{\boldsymbol{Y}}$.

Nemirovsky-Rotman, S. and Bercovich, E., 2024. Explicit physics-informed deep learning for computer-aided diagnostic tasks in medical imaging. *Machine Learning and Knowledge Extraction*, *6*(1), pp.385-401.

Hypothesis 2: Improved *robustness* to random input variations: the network $\boldsymbol{g}_{\boldsymbol{\Theta}^{I_E}}^{\boldsymbol{I_E}}$ that is trained with the *generalization* loss $\boldsymbol{\mathcal{L}}_{\boldsymbol{I_E}}^{\boldsymbol{\mathcal{D}_{\delta I_E}}}$ is expected to have improved *robustness* to input variations with respect to $\boldsymbol{g}_{\boldsymbol{\Theta}}^{\boldsymbol{I}}$:

- $\mathbb{E}_{\boldsymbol{\delta Y}\sim\boldsymbol{\mathcal{D}}_{\boldsymbol{\delta Y}}}\left[\mathbb{E}_{\boldsymbol{Y}\sim\boldsymbol{\mathcal{D}}_{\boldsymbol{Y}}}\left\{\left\|\boldsymbol{g}_{\boldsymbol{\Theta}^{I_E}}^{\boldsymbol{I_E}}(\boldsymbol{Y}+\boldsymbol{\delta Y};\boldsymbol{I})-\boldsymbol{g}_{\boldsymbol{\Theta}^{I_E}}^{\boldsymbol{I_E}}(\boldsymbol{Y};\boldsymbol{I})\right\|\,|\boldsymbol{\delta Y}\right\}\right]\leq\mathbb{E}_{\boldsymbol{\delta Y}\sim\boldsymbol{\mathcal{D}}_{\boldsymbol{\delta Y}}}\left[\mathbb{E}_{\boldsymbol{Y}\sim\boldsymbol{\mathcal{D}}_{\boldsymbol{Y}}}\{\|\boldsymbol{g}_{\boldsymbol{\Theta}}^{\boldsymbol{I}}(\boldsymbol{Y}+\boldsymbol{\delta Y})-\boldsymbol{g}_{\boldsymbol{\Theta}}^{\boldsymbol{I}}(\boldsymbol{Y})\|\,|\boldsymbol{\delta Y}\}\right]$

where $\boldsymbol{\delta Y}$ is an input variation distributed according to $\boldsymbol{\delta Y}\sim\boldsymbol{\mathcal{D}}_{\boldsymbol{\delta Y}}$, and $\mathbb{E}_{\boldsymbol{\delta Y}\sim\boldsymbol{\mathcal{D}}_{\boldsymbol{\delta Y}}}(\cdot\,|\boldsymbol{\delta Y})$ denotes the conditional expectation value, conditioned on a given input variation of value $\boldsymbol{\delta Y}$.

Nemirovsky-Rotman, S. and Bercovich, E., 2024. Explicit physics-informed deep learning for computer-aided diagnostic tasks in medical imaging. *Machine Learning and Knowledge Extraction*, *6*(1), pp.385-401.

7. EPI-Learning for Image Reconstruction

- Define the following image reconstruction task:

$$\boldsymbol{Y} = \boldsymbol{A_I X} + \boldsymbol{N},$$

- where: $\boldsymbol{X} \in \mathbb{R}^{\boldsymbol{M^2}}$ is a column-wise vector representing an original, high-quality $\boldsymbol{M} \times \boldsymbol{M}$ image;
- $\boldsymbol{Y} \in \mathbb{R}^{\boldsymbol{M^2}}$ is a column-wise vector representing a degraded, low-quality $\boldsymbol{M} \times \boldsymbol{M}$ image;
- $\boldsymbol{A_I} \in \mathbb{R}^{\boldsymbol{M^2 X M^2}}$ is a known degradation operator matrix, whose entries are determined according to the given explicit physics prior coordinates $\boldsymbol{I}$;
- $\boldsymbol{N} \in \mathbb{R}^{\boldsymbol{M^2}}$ is additive random noise, independently and identically distributed (i.i.d.) according to some distribution $\boldsymbol{\mathcal{D}_N}$.

EPI-Learning in Image Reconstruction – cont.

- A decimation operator $\boldsymbol{D}$ may describe a degradation operator that is followed by sub-sampling (decimation operator):

$$\boldsymbol{Y} = \boldsymbol{D}\boldsymbol{A}_{\boldsymbol{I}}\boldsymbol{X} + \boldsymbol{N},$$

$\boldsymbol{Y} \in \mathbb{R}^{\boldsymbol{L}^2}, \boldsymbol{L} < \boldsymbol{M}$ is the down-sampled and corrupted image;

$\boldsymbol{N} \in \mathbb{R}^{\boldsymbol{L}^2}$ is the additive noise.

Re: Image-to-Sensor Equations in MRI

- The acquired k-space signal $\boldsymbol{y(t)}$ at time $\boldsymbol{t}$ may is expressed as:

$$\boldsymbol{y(t) = \int M(\rho(r), \vartheta, r, t) e^{-j2\pi k(t)\cdot r} dr + n(t),}$$

Where:

$\boldsymbol{r}$ is the spatial location

$\boldsymbol{\rho(r)}$ is the underlying magnetization

$\boldsymbol{\vartheta}$ is a set of parameters that model physiological or systemic changes

$\boldsymbol{k(t)}$ is the k-space location at time $\boldsymbol{t}$

$\boldsymbol{n(t)}$ is the measurement noise.

Hammernik, K., Küstner, T., Yaman, B., Huang, Z., Rueckert, D., Knoll, F. and Akçakaya, M., 2023. Physics-driven deep learning for computational magnetic resonance imaging: Combining physics and machine learning for improved medical imaging. *IEEE signal processing magazine*, *40*(1), pp.98-114.

Re: Image-to-Sensor Equations in MRI

A simplified and discretized acquisition model

$$\boldsymbol{y} = \boldsymbol{E}\boldsymbol{x} + \boldsymbol{n},$$

Where:

$\boldsymbol{x}$ is the image/magnetization,

$\boldsymbol{y}$ denotes the corresponding k-space measurements,

$\boldsymbol{E}$ is the forward encoding operator, which reflects the hardware characteristics of the scanner.

Image-to-Sensor Equations in LF-MRI

The discretized forward model:

$$\boldsymbol{S} = \boldsymbol{E}\boldsymbol{m} + \boldsymbol{n},$$

where:

$\boldsymbol{S}$ is the observed signal that corresponds to the LF scan,

$\boldsymbol{E}$ is the encoding matrix,

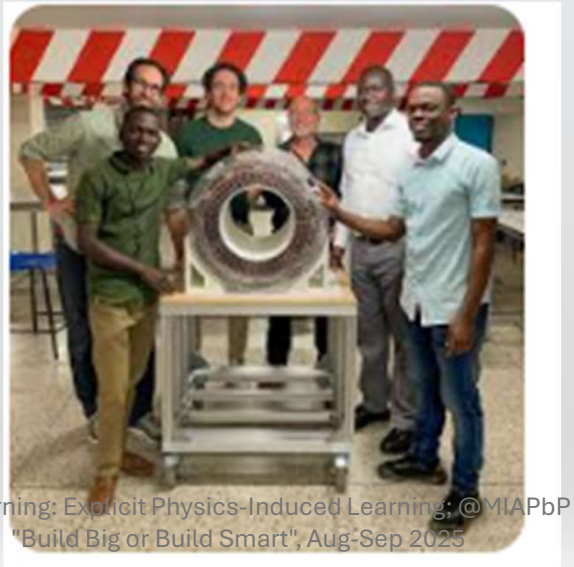
$\boldsymbol{m}$ is the corresponding HF image for which the inverse problem is solved.

The encoding matrix $\boldsymbol{E}$ encapsulates the entire system modeling. $\boldsymbol{E}$ depends on hardware, magnetic field and gradient maps, as well as image acquisition parameters: $\boldsymbol{E} = \boldsymbol{E}(\boldsymbol{B_0}, \boldsymbol{G_x}, \boldsymbol{G_y}, \dots)$,

where $\boldsymbol{B_0}$ is the magnetic field map,

$\boldsymbol{G_x}, \boldsymbol{G_y}$ represent the gradient magnetic field maps in the $\boldsymbol{x}, \boldsymbol{y}$ imaging dimensions.

Low-Field MRI Scanners- "MRI Democratization"



Why Low-Field (LF) MRI?

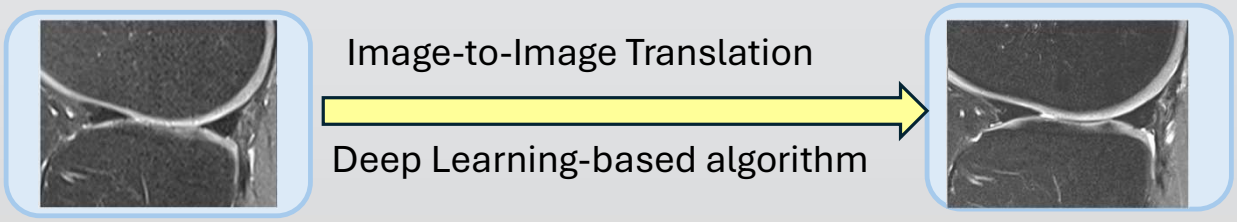
- Conventional High-Field (HF) scanners (1.5/3 Tesla) require special physical siting conditions, specialized technicians and high maintenance costs.
- High costs further contribute to lack of resources and increasing workloads on the national health care system.
- LF scanners: compact, portable, require significantly reduced installation and maintenance costs.
- Reduced sensitivity to metal implants.

LF-MRI's Limitations

- LF images suffer from reduced Signal-to-Noise Ratio (SNR), contrast and spatial resolution.

$$SNR \propto B_0^{7/4}$$

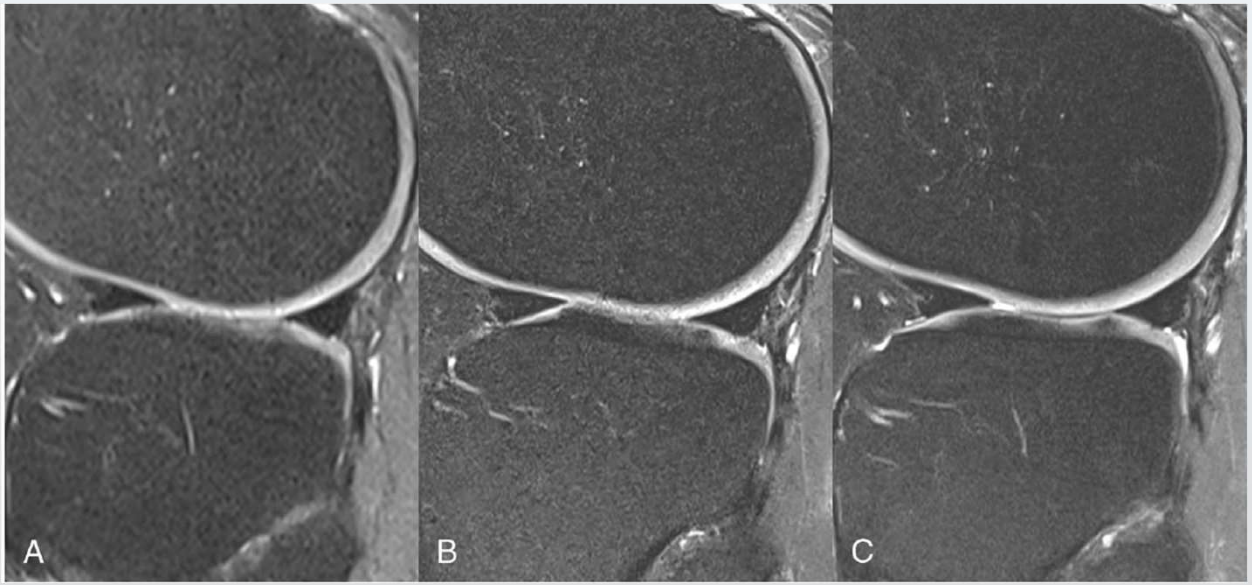
- Specific applications require high resolution and noise: degenerative disorders identification, injury diagnosis and more.
- The solution: State-of-the-Art DL, which can translate LF images to HF-comparable image quality: Image-to-mage Translation.



Input: Low-Field (LF) MRI scan

Output: High-Field (HF) quality produced image

The Goal of Image-to-Image Translation Algorithms

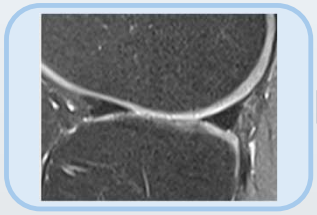


0.55 T 1.5 T 3 T

Sagittal MRI scans of the medial meniscus and articular cartilage of the knee obtained at (A) 0.55 T, (B) 1.5 T, (C) 3 T.

Figure taken from: *Runge, V., & Heverhagen, J. (2022). The clinical utility of magnetic resonance imaging according to field strength, specifically addressing the breadth of current state-of-the-art systems, which include 0.55 T, 1.5 T, 3 T, and 7 T. Investigative radiology, 1-12.*

3 Main DL Blocks for Low-Field to High-Field Image-to-Image Translation



LF-domain,
Low-resolution Input

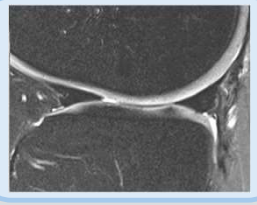
Block #1:
Domain
Adaptation

HF-domain, noisy
Low-resolution
image

Block #2:
Denoiser

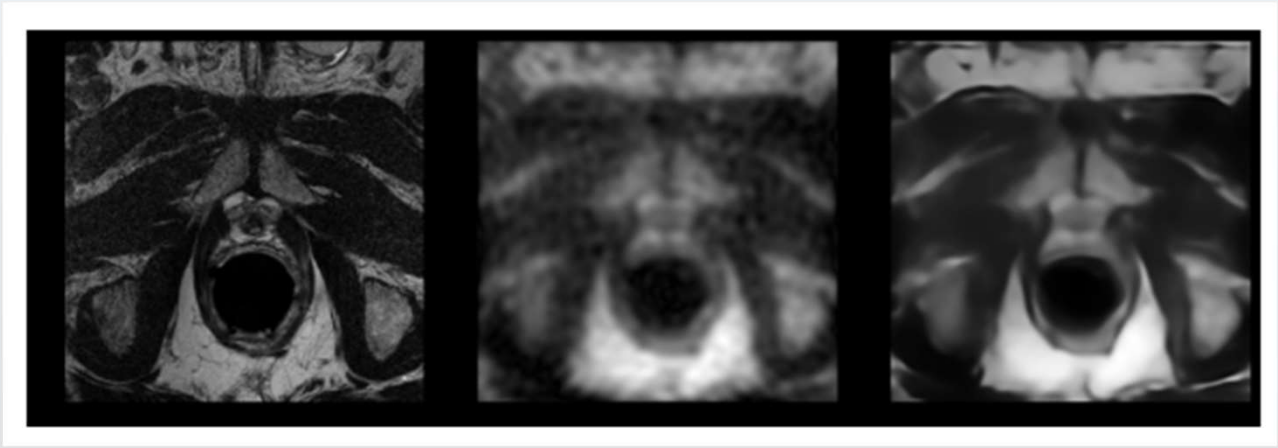
HF-domain,
Denoised,
Low-resolution
image

Block #3:
Super-
Resolution



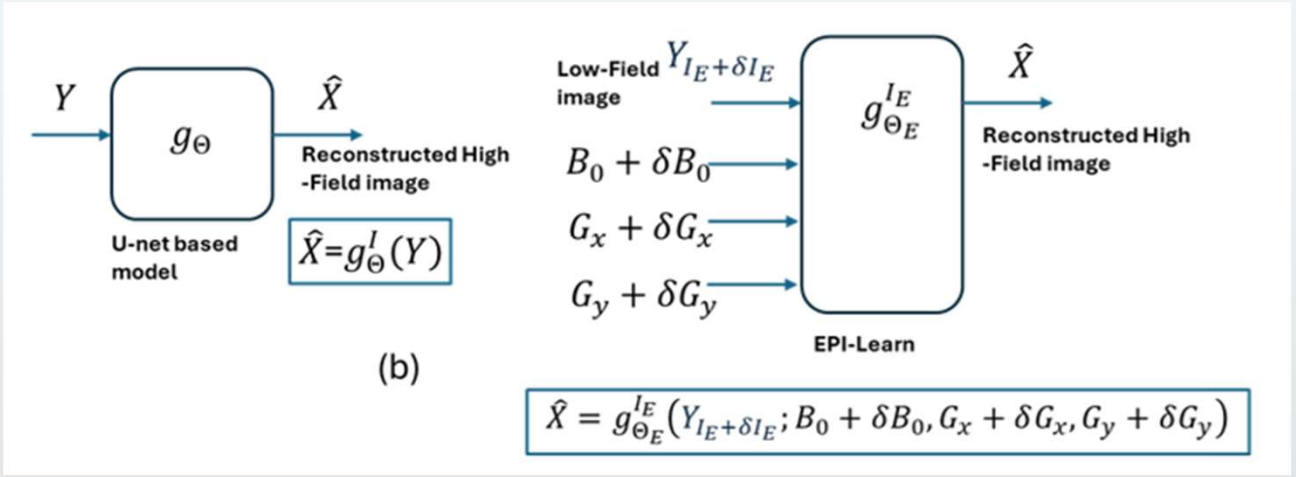
HF-domain,
High-resolution Output

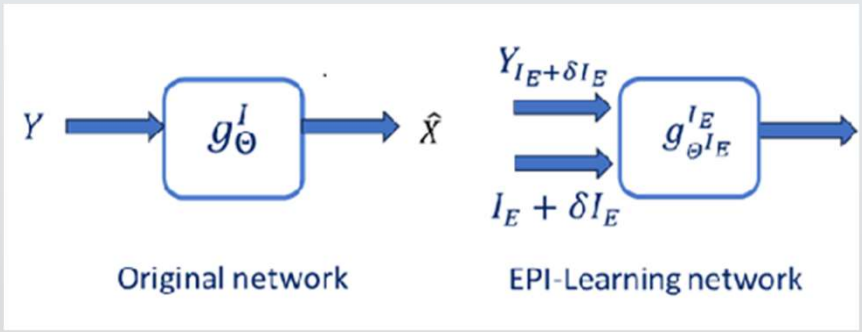
Promaxo LTD (California, USA)



Le, D.B.T., Sadinski, M., Nacev, A., Narayanan, R. and Kumar, D., 2021, August. Deep learning–based method for denoising and image enhancement in low-field MRI. In *2021 IEEE international conference on imaging systems and techniques (IST)* (pp. 1-6). IEEE.

Low-Field MRI – EPI-Learning





Ongoing joint collaboration of CoMLab (Dr. Shira Rotman) and Promaxo, LTD., California, USA.

Low-Field MRI – EPI-Learning

- The modified field maps will be generated as follows:
- Step (1): A set of $\boldsymbol{K = 30}$ matrices, representing the field maps of $\boldsymbol{B_0, G_x, G_y}$, will be provided by Promaxo, namely:
- $\left\{\boldsymbol{B_0^k}\right\}_{\boldsymbol{k=1}}^{\boldsymbol{K}}, \left\{\boldsymbol{G_x^k}\right\}_{\boldsymbol{k=1}}^{\boldsymbol{K}}, \left\{\boldsymbol{G_y^k}\right\}_{\boldsymbol{k=1}}^{\boldsymbol{K}}$.
- Step (2): For each of the three field maps, the empirical expectation value $\hat{\boldsymbol{\mu}}$ and variance $\hat{\boldsymbol{\sigma}}^{\boldsymbol{2}}$ (first- and second-order
- probability distribution moments) will be estimated entry-wise (for row, column indices $\boldsymbol{i, j}$), as follows:
- $\left(\hat{\boldsymbol{\mu}}_{\boldsymbol{B_0}}\right)_{\boldsymbol{i,j}} = \frac{\boldsymbol{1}}{\boldsymbol{K}}\sum_{\boldsymbol{k=1}}^{\boldsymbol{1}}\left(\boldsymbol{B_0^k}\right)_{\boldsymbol{i,j}}, \left(\hat{\boldsymbol{\mu}}_{\boldsymbol{G_x}}\right)_{\boldsymbol{i,j}} = \frac{\boldsymbol{1}}{\boldsymbol{K}}\sum_{\boldsymbol{k=1}}^{\boldsymbol{1}}\left(\boldsymbol{G_x^k}\right)_{\boldsymbol{i,j}}, \left(\hat{\boldsymbol{\mu}}_{\boldsymbol{G_y}}\right)_{\boldsymbol{i,j}} = \frac{\boldsymbol{1}}{\boldsymbol{K}}\sum_{\boldsymbol{k=1}}^{\boldsymbol{1}}\left(\boldsymbol{G_y^k}\right)_{\boldsymbol{i,j}}$.
- $\left(\hat{\boldsymbol{\sigma}}^{\boldsymbol{2}}_{\boldsymbol{B_0}}\right)_{\boldsymbol{i,j}} = \frac{\boldsymbol{1}}{\boldsymbol{K}}\sum_{\boldsymbol{k=1}}^{\boldsymbol{1}}\left[\left(\boldsymbol{B_0^k}\right)_{\boldsymbol{i,j}} - \left(\hat{\boldsymbol{\mu}}_{\boldsymbol{B_0}}\right)_{\boldsymbol{i,j}}\right]^{\boldsymbol{2}}, \left(\hat{\boldsymbol{\sigma}}^{\boldsymbol{2}}_{\boldsymbol{G_x}}\right)_{\boldsymbol{i,j}} = \frac{\boldsymbol{1}}{\boldsymbol{K}}\sum_{\boldsymbol{k=1}}^{\boldsymbol{1}}\left[\left(\boldsymbol{G_x^k}\right)_{\boldsymbol{i,j}} - \left(\hat{\boldsymbol{\mu}}_{\boldsymbol{G_x}}\right)_{\boldsymbol{i,j}}\right]^{\boldsymbol{2}}, \left(\hat{\boldsymbol{\sigma}}^{\boldsymbol{2}}_{\boldsymbol{G_y}}\right)_{\boldsymbol{i,j}}$
  $= \frac{\boldsymbol{1}}{\boldsymbol{K}}\sum_{\boldsymbol{k=1}}^{\boldsymbol{1}}\left[\left(\boldsymbol{G_y^k}\right)_{\boldsymbol{i,j}} - \left(\hat{\boldsymbol{\mu}}_{\boldsymbol{G_y}}\right)_{\boldsymbol{i,j}}\right]$
- Step (3): For each of the three field maps, their perturbed values will be generated by adding to each matrix entry a random variation that is uniformly distributed, according to the empirical second-order moments from Step (2):
- $\left(\boldsymbol{\delta B_0}\right)_{\boldsymbol{i,j}} \sim \boldsymbol{\mathcal{U}}\left(-\left(\hat{\boldsymbol{\sigma}}_{\boldsymbol{B_0}}\right)_{\boldsymbol{i,j}}, \left(\hat{\boldsymbol{\sigma}}_{\boldsymbol{B_0}}\right)_{\boldsymbol{i,j}}\right), \left(\boldsymbol{\delta G_x}\right)_{\boldsymbol{i,j}} \sim \boldsymbol{\mathcal{U}}\left(-\left(\hat{\boldsymbol{\sigma}}_{\boldsymbol{G_x}}\right)_{\boldsymbol{i,j}}, \left(\hat{\boldsymbol{\sigma}}_{\boldsymbol{G_x}}\right)_{\boldsymbol{i,j}}\right), \left(\boldsymbol{\delta G_y}\right)_{\boldsymbol{i,j}} \sim \boldsymbol{\mathcal{U}}\left(-\left(\hat{\boldsymbol{\sigma}}_{\boldsymbol{G_y}}\right)_{\boldsymbol{i,j}}, \left(\hat{\boldsymbol{\sigma}}_{\boldsymbol{G_y}}\right)_{\boldsymbol{i,j}}\right)$.

Ongoing joint collaboration of CoMLab (Dr. Shira Rotman) and Promaxo, LTD., California, USA.

8. Summary and Future Directions

- Explicit physics-prior, which encapsulates a presumably known, or partially-known, forward model
- Incorporation of such explicit prior, combined with extended training data set (created by applying perturbations over the prior)- → yields improved generalization to OODs and robustness to random input noise.
- The proposed approach may be translated into applications for which an inverse problem should be solved, based on a given forward model formulation.

Thank you!

Danke Schon!