

Inverse problems and machine learning in medical physics

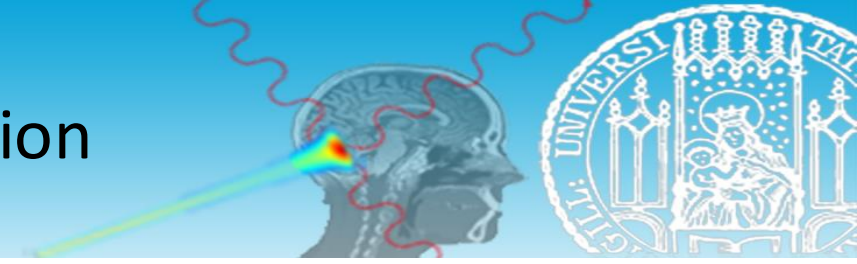
Machine learning for tomographic image
reconstruction or “deep reconstruction”

Dr. Chiara Gianoli

10/12/2024

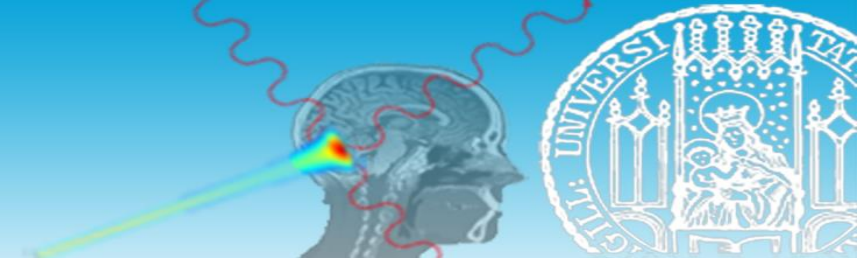
chiara.gianoli@physik.uni-muenchen.de

Tomographic image reconstruction

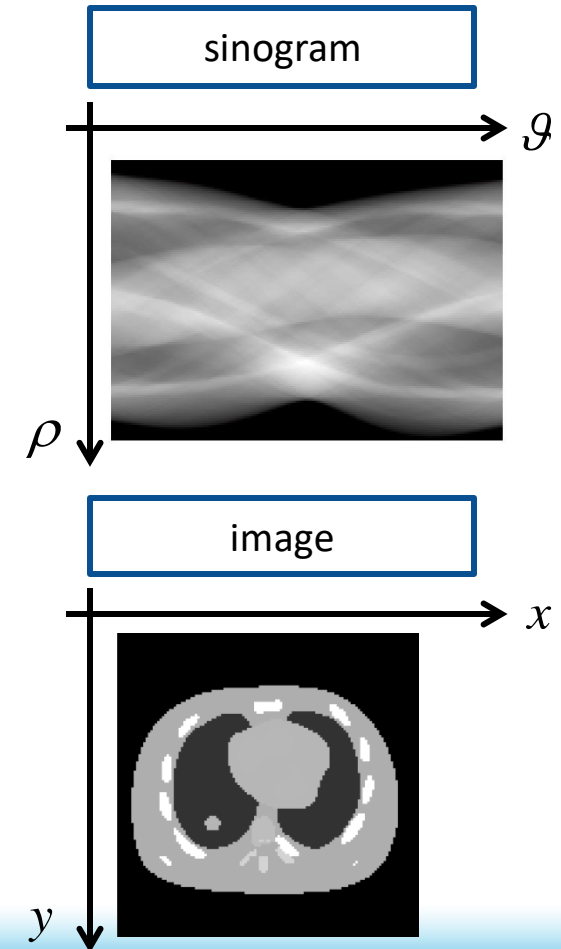


- Tomographic image reconstruction represents the building block of medical imaging
- Tomographic image reconstruction has been classified as [analytical reconstruction](#) or [numerical reconstruction](#)
- Very recently, data-driven, deep-learning-based tomographic image reconstruction has been introduced (i.e., [deep tomographic reconstruction](#))
 - Direct reconstruction methods
 - Unrolled iterative reconstruction methods
- The huge benefit of machine learning in reconstruction is the use of the [ground truth](#) (i.e., supervised learning), as obtained from high quality simulations or high quality measurements

Analytical reconstruction



- Tomographic image **acquisition** can be modelled as a **Radon transform**, or **sinogram**, of the variable describing the physical properties of the object of interest
 - The Radon Transform converts an image from **spatial domain** to **sinogram domain**, by integrating the variables along the integration lines, as a function of the projection angles
 - The analytical image reconstruction is based on the **Fourier slice theorem** that puts in correspondence the Radon Transform with the Fourier Transform of the image (i.e., the filtered back-projection)

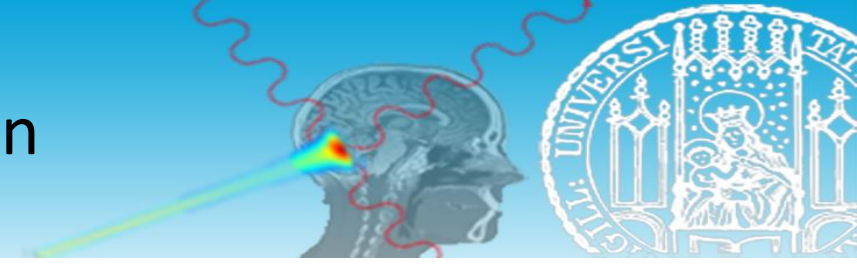


$$\hat{f}_\rho(w_x, w_y) = \int_{-\infty}^{+\infty} R(f) e^{-2\pi i(\rho w_\rho)} d\rho = \hat{R}(w_\rho)$$

Radon transform

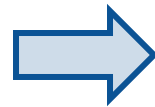
Fourier transform

Numerical image reconstruction

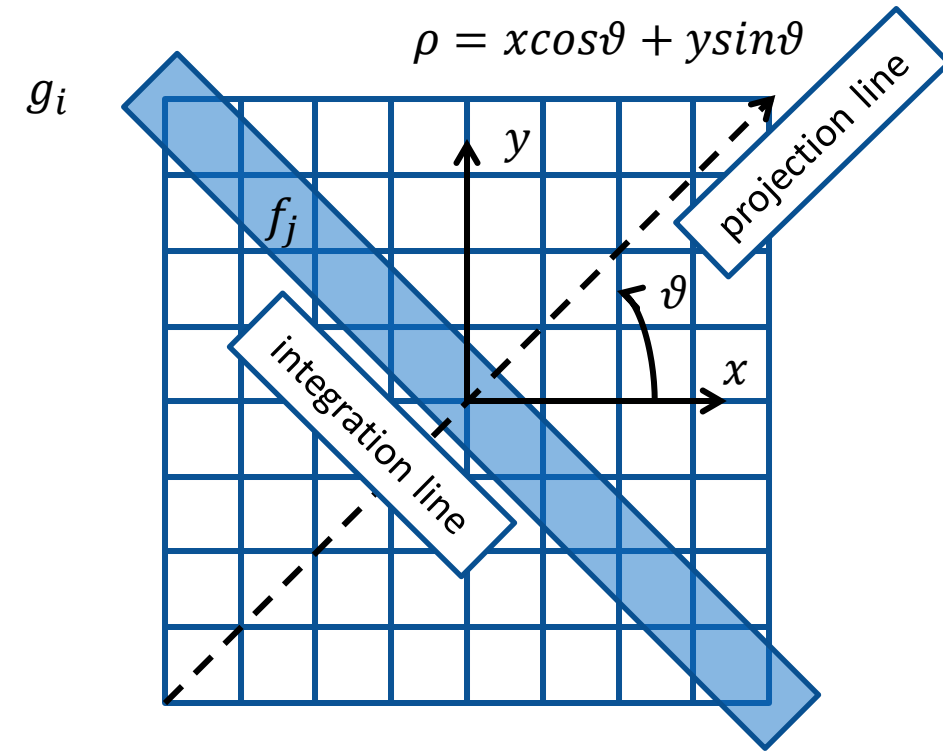


- The inverse problem of tomographic image reconstruction can be solved by means of numerical (iterative) algorithms
- Numerical algorithms can be considered as an iterative solver of a **system of linear equations**
 - I equations, one for each projection
 - J unknowns, one for each pixel

$$\begin{cases} a_{11}f_1 + a_{12}f_2 + \dots + a_{1J}f_J = g_1 \\ \dots \\ a_{I1}f_1 + a_{I2}f_2 + \dots + a_{IJ}f_J = g_I \end{cases}$$

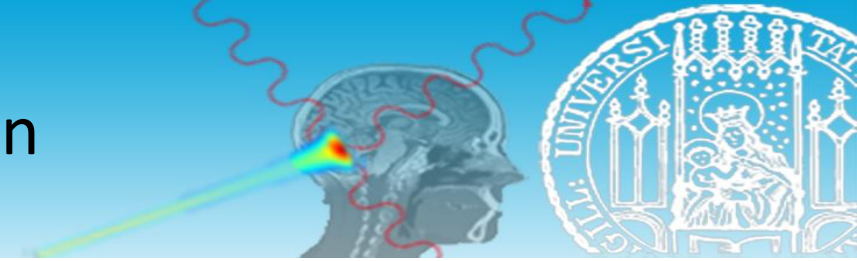


$$g_i = \sum_j a_{ij}f_j$$

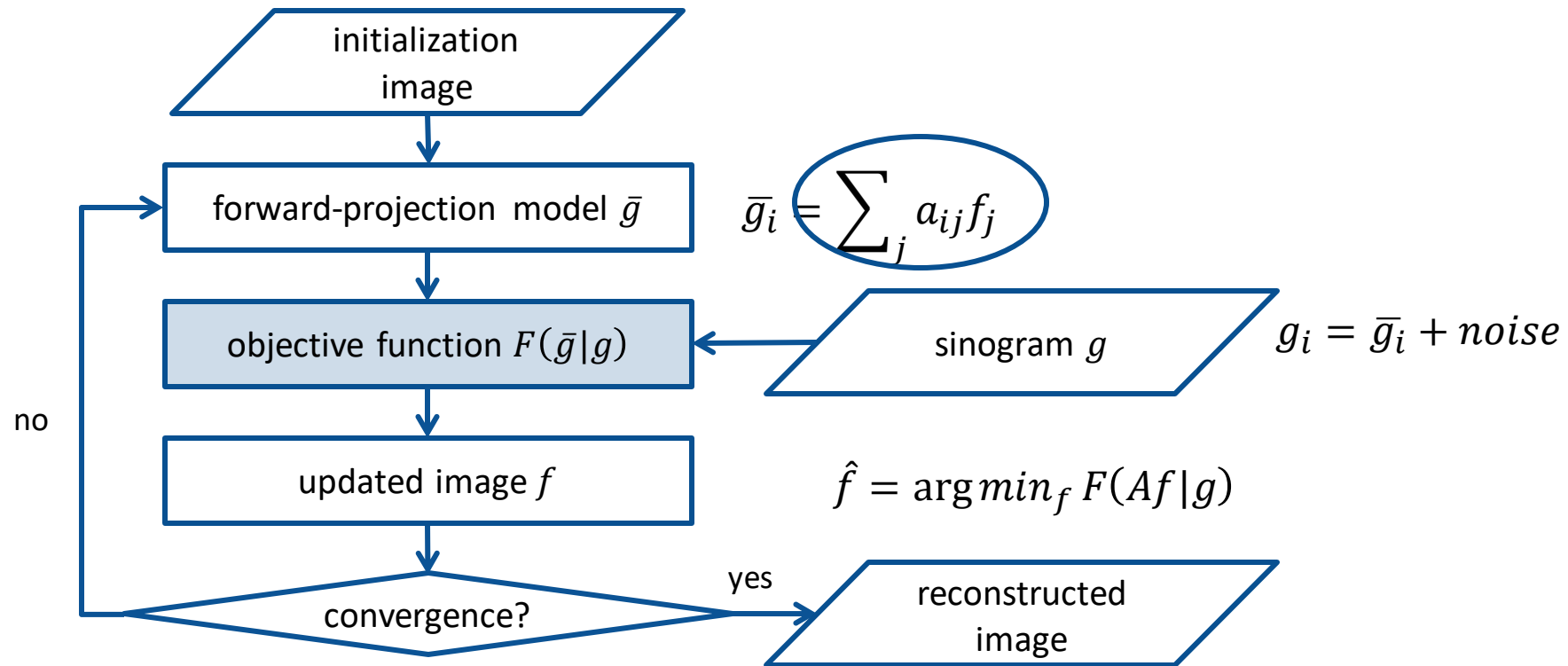


- The coefficients a_{ij} (i.e., the elements of the **system matrix**) express the intersection area/volume of the pixel/voxel j with the integration line of the projection i

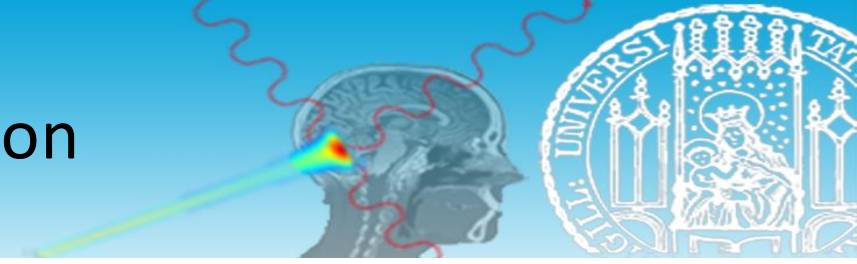
Numerical image reconstruction



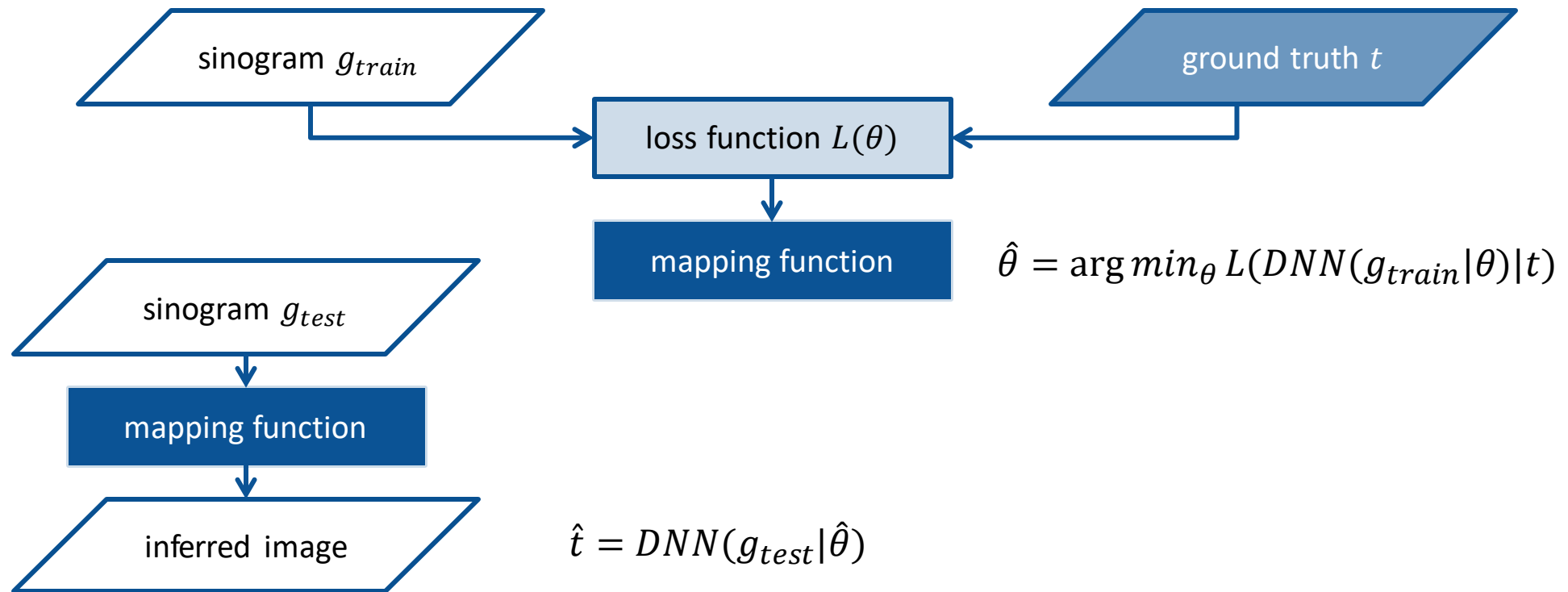
- The **iterative reconstruction paradigm** is to find the image that minimizes the “discrepancy” between the forward-projection of the image (i.e., the **model of the sinogram**) and the acquired sinogram



Deep tomographic reconstruction

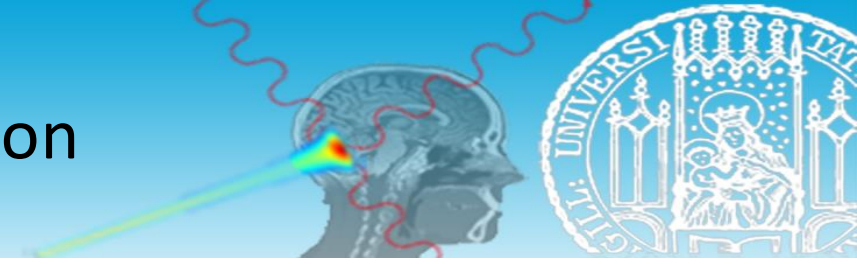


- The machine learning paradigm in tomographic image reconstruction is to find the parameters of the mapping function that infers the ground truth based on supervised prediction

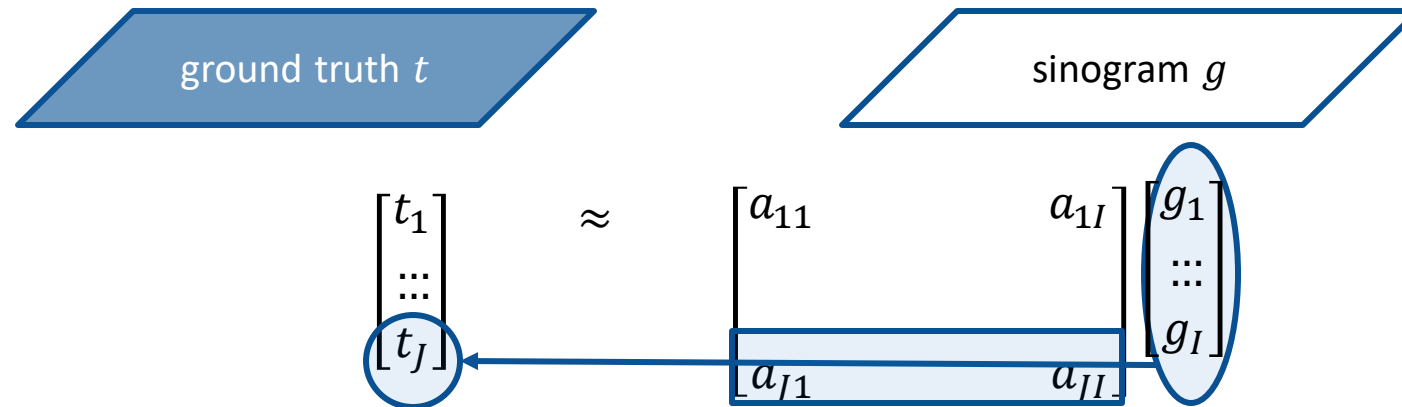


- Different from analytical and numerical approaches, deep learning deploys a method for reconstruction (i.e., an estimator), not a reconstruction (i.e., an estimate)

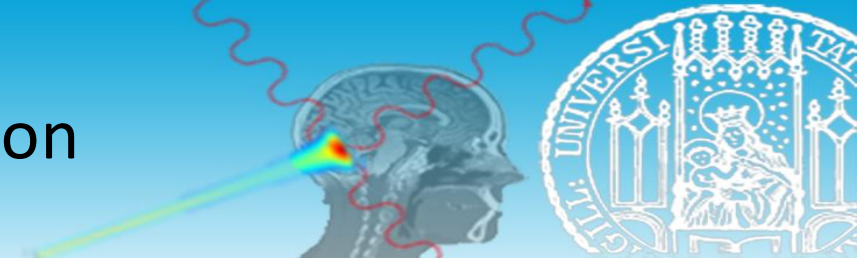
Deep tomographic reconstruction



- The back-projection is a linear mapping (i.e., matrix-vector multiplication) that can be described by a **fully connected layer** (i.e., linear layer) of an artificial neural network (ANN)

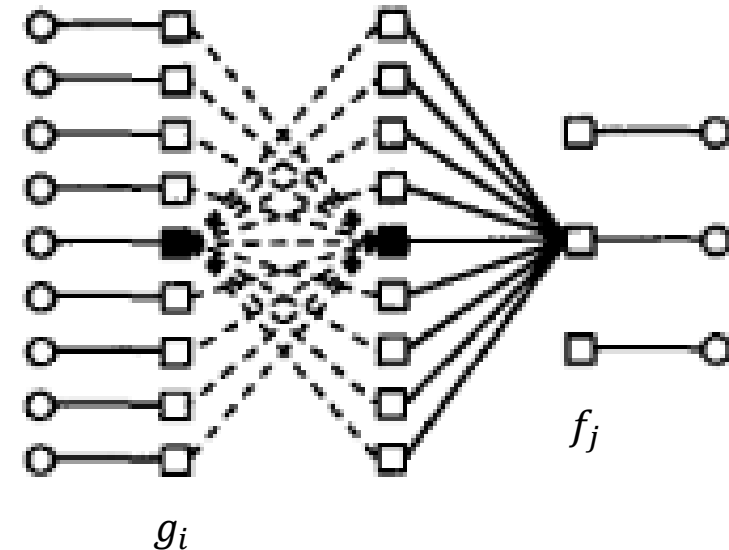


Deep tomographic reconstruction



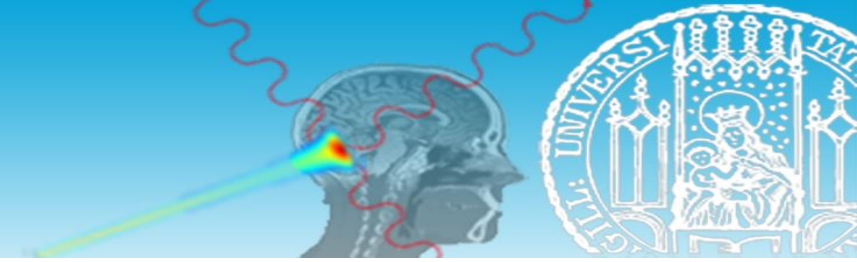
- One of the first ML attempt to deep tomographic reconstruction was based on the “pre-calculation” of the filters for the filtered back-projection, instead of being analytically calculated each time...
 - The learnable weights (learnt based on a point source) are applied along the projection lines of the sinogram
 - The back-projection is implemented for each projection lines of the sinogram as **fully connected layer** with **non-learnable weights** (rotational and shift-invariant approximation)

Floyd, C. E. (1991). An artificial neural network for SPECT image reconstruction. *IEEE transactions on medical imaging*, 10(3), 485-487.



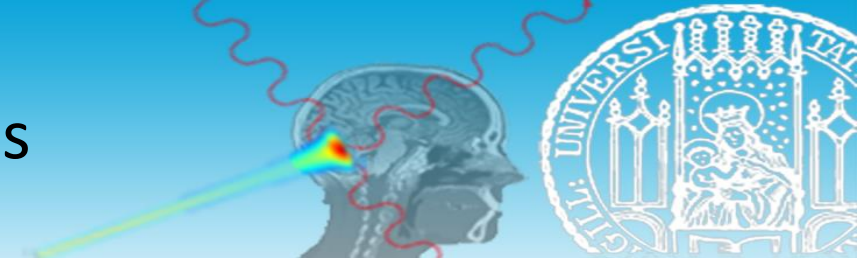
- In practice, this is suitable only for two-dimensional images

Exercise #2

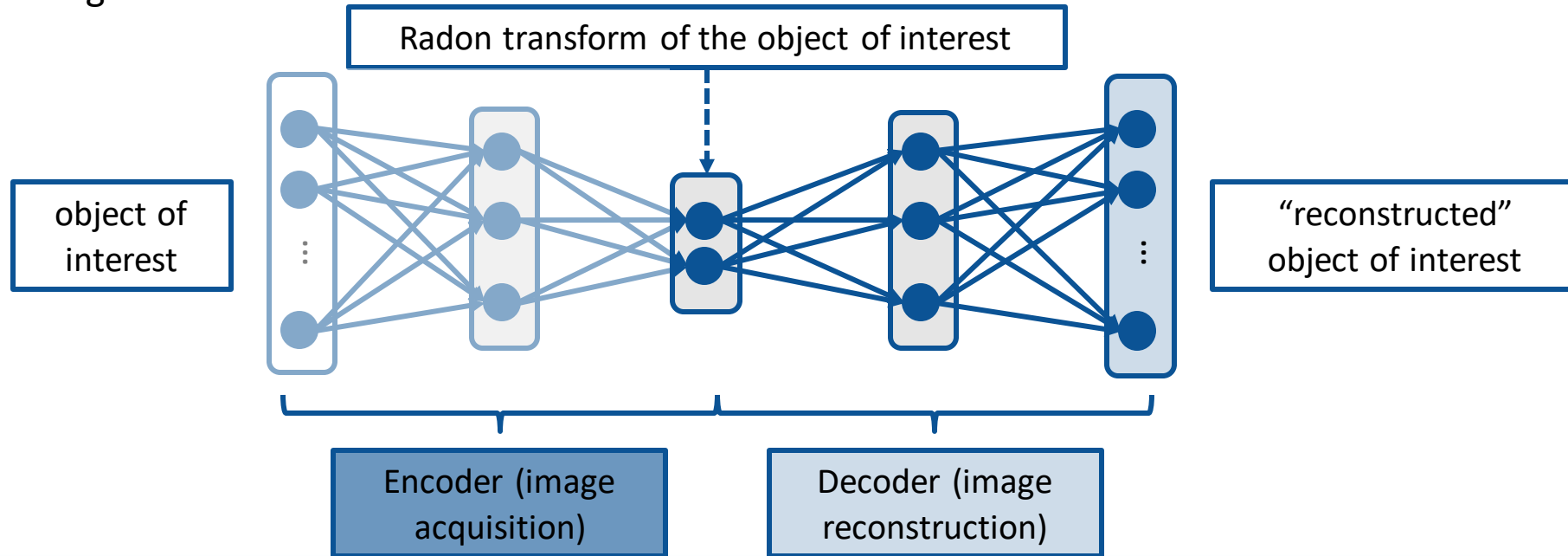


- Define the image of an ideal point source (i.e., **target data** of the network)
- Simulate the sinogram of the ideal point source, then add noise and blur along each projection line (i.e., **input data** of the network)
- Implement a **first fully connected network**, connecting the nodes of the **input layer** (i.e., the projection line) to all the nodes of the **hidden layer** (i.e., the filtered projection line)
 - The forward-pass function is a weighed average of the inputs with unknown weights (i.e., no bias, no activation function)
- Implement a **second fully connected network**, connecting the nodes of the hidden layer (i.e., the filtered projection line) to all the nodes of the output layer (i.e., the image)
 - The forward-pass function is a weighed average of the inputs with known weights from the system matrix
- Train the network based on input and target data (implement the backward-pass based on the gradient descend algorithms)

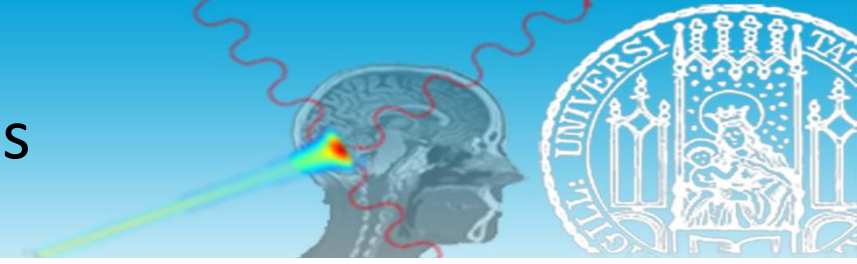
Direct reconstruction methods



- The purpose of domain transform is to map the sinogram (i.e., the projections) to the image
 - The measured sinogram encodes an intermediate representation of the object of interest in the projection domain (i.e., the Radon transform), similar to an encoding function
 - The measured sinogram is subsequently reconstructed into an image by an inversion of the encoding function, similar to a decoding function

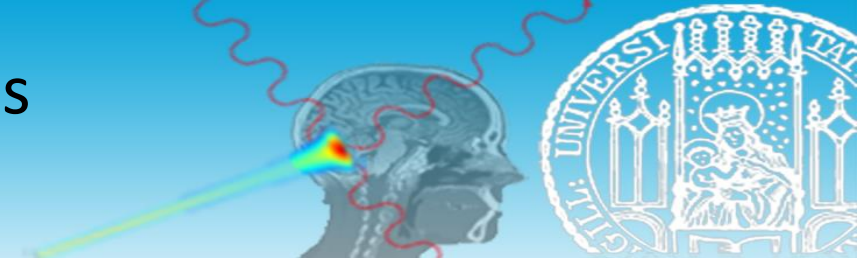


Direct reconstruction methods

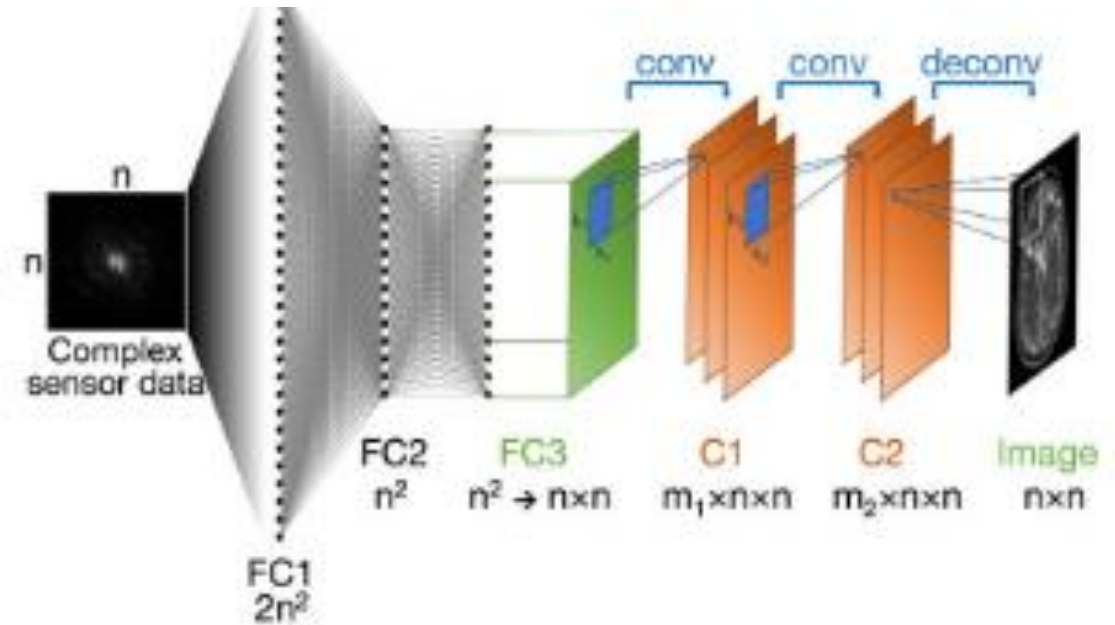


- In direct reconstruction methods, the domain transformation can be learnt from the network or explicitly given as input to the network
 - Direct reconstruction methods can entail the encoding of the Radon transform into a lower dimensional space (i.e., **compressed sensing**) and the decoding of the encoded Radon transform, typically by means of convolutional layers
 - The compressed sensing hypothesis is that a signal can be **represented by** and **perfectly recovered from** few non-zero coefficients in a suitable basis (i.e., dictionary)
 - *Wavelet* and *Shearlet* are common basis functions (<https://www.math.uh.edu/~dlabate/SHBookIntro.pdf>)
 - Transformation into a lower dimensional space can be based on manifolds (i.e., **manifold-based compressed sensing**)
 - The manifold hypothesis is that high dimensional data (i.e., a **continuous images**) lie on low-dimensional manifolds (i.e., a point) in a high-dimensional space (i.e., an **infinite dimensional vectoral space**)
 - Zero-dimensional manifolds are points, one-dimensional manifolds are lines, two-dimensional manifolds are surfaces...

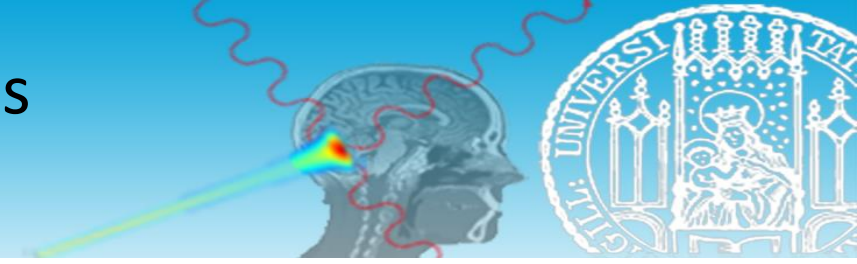
Direct reconstruction methods (AUTOMAP)



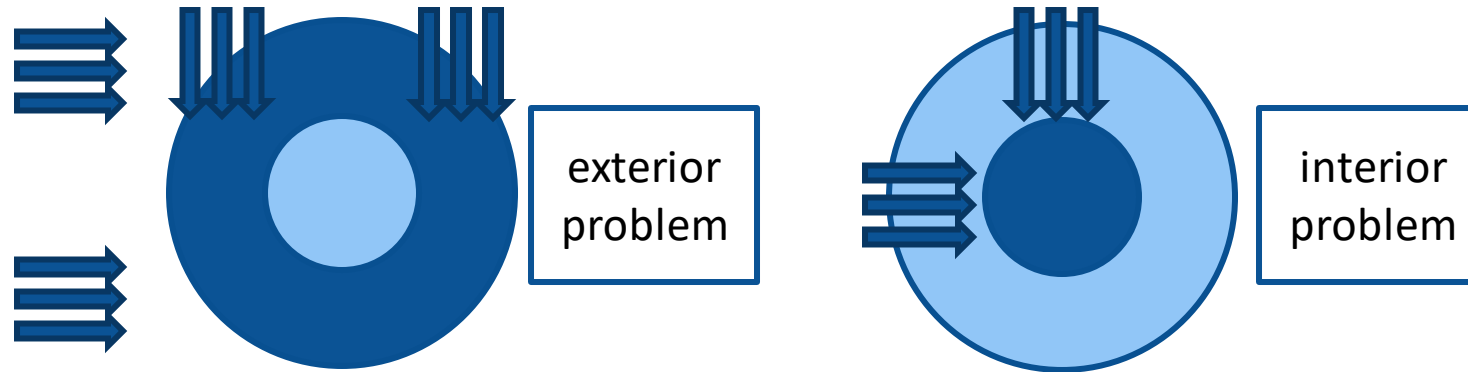
- The AUtomated TransfOrm by Manifold Approximation (AUTOMAP) is a deep neural network with feed-forward architecture, composed of **multiple fully-connected layers** followed by a **sparse convolutional auto-encoder** (i.e., encoder-decoder where the input and the output domains are the same)
- The network simultaneously learns an optimal convolutional domain (i.e., **manifold space**) and a sparse representations (i.e., **compressed sensing**) through a joint optimization (i.e., manifold encoding–decoding process)
 - Different from **compressed sensing**, the convolutional layers do not make hypothesis on the sparsifying transform (e.g., wavelet, shearlet...)
 - AUTOMAP is originally demonstrated for MRI but it is generally applicable to different image reconstruction problems



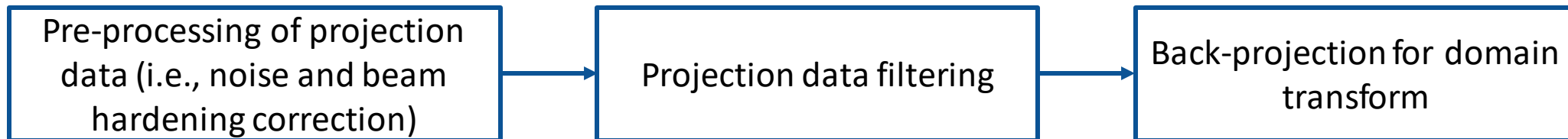
Direct reconstruction methods (iCT-Net)



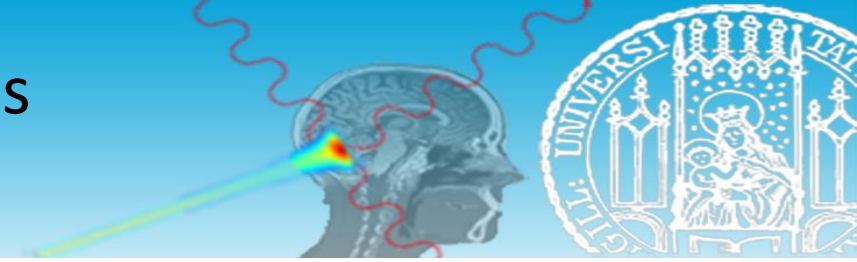
- The intelligent CT network (iCT-Net) is a deep neural network based on multi-channel convolutional layers intended for image reconstruction of truncated data (i.e., the limited and sparse angle problem, the exterior problem and the interior problem)



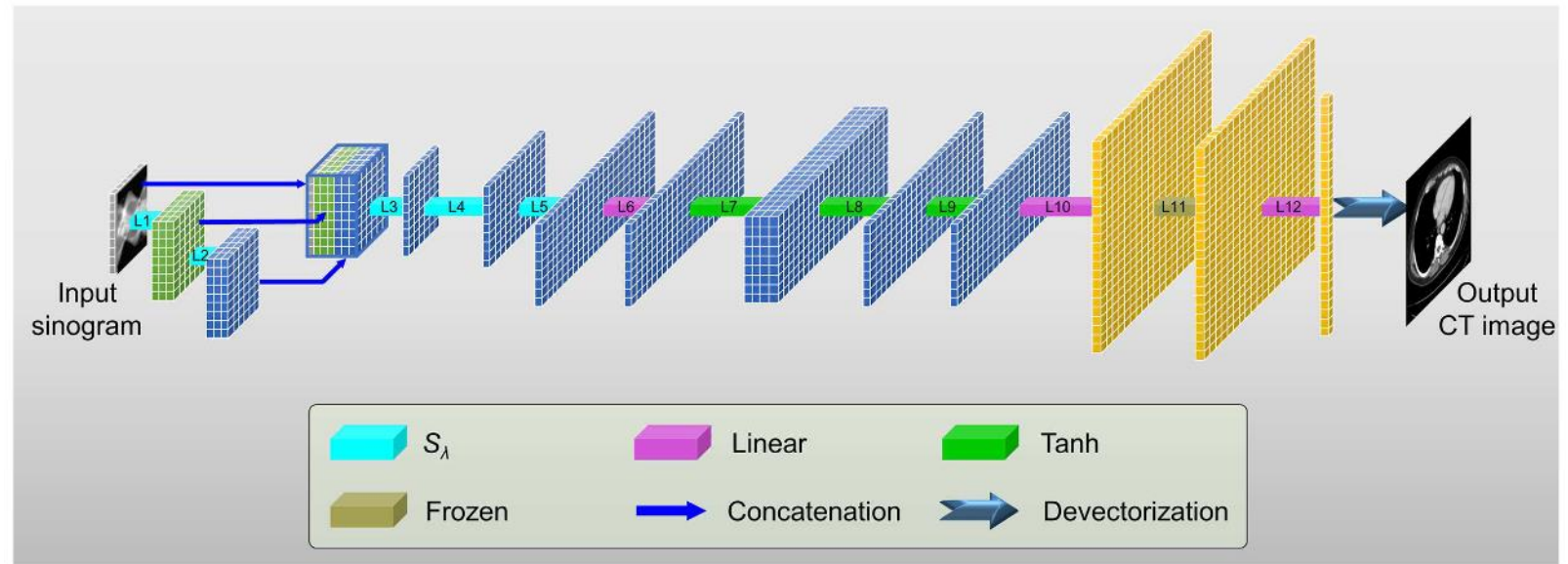
- The design of the iCT-Net is inspired by the **filtered back-projection pipeline** which consists of three major cascaded steps



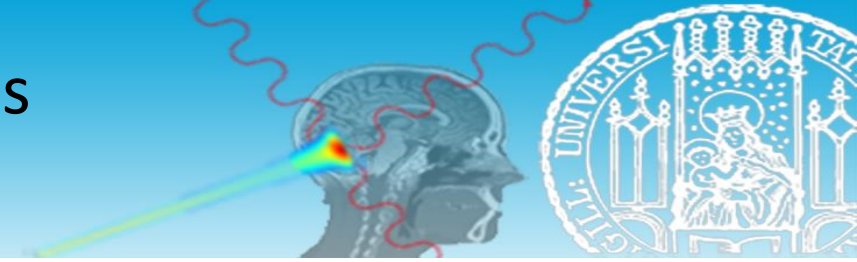
Direct reconstruction methods (iCT-Net)



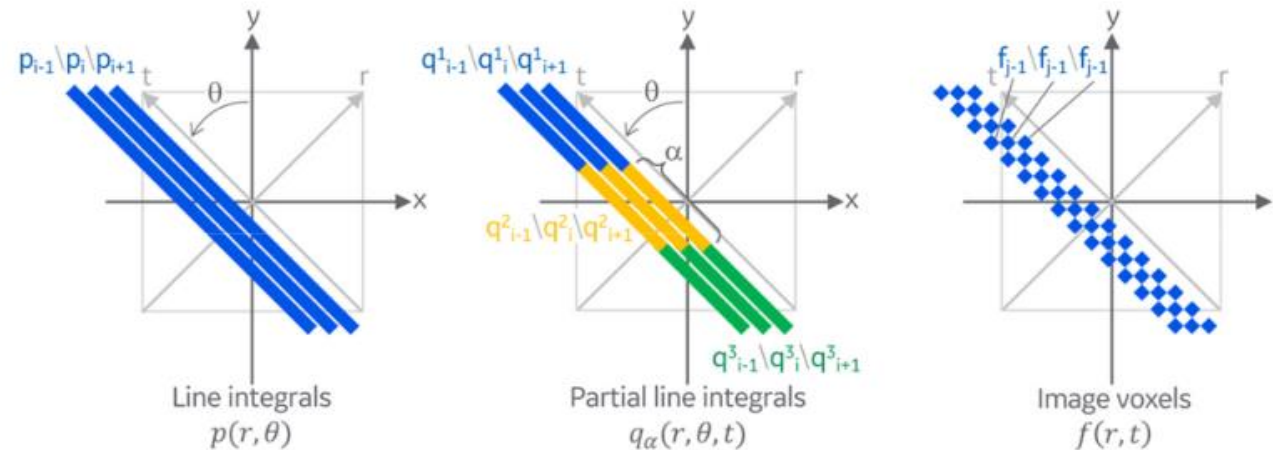
- Five **convolutional layers** (L1-L5 in figure) to suppress noise and convert a sparse-view sinogram into a dense-view sinogram (i.e., manifold learning or pre-processing step)
- Four **convolutional layers** (L6-L9 in figure) to extract features (i.e., filtering)
- A **fully connected layer** (L10 in figure) to perform the **domain transform** from the extracted feature space to image space (i.e., back-projection step)
- Two convolutional layers (L11-L12 in figure) to learn a combination of the image from each view (i.e., summation step but with learnable summation weights)
 - The rotational symmetry of the back-projection is explicitly implemented to reduce the dimensionality of the network



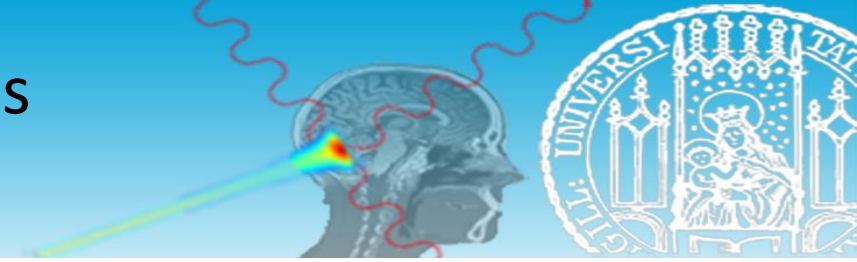
Direct reconstruction methods (hierarchical deep learning)



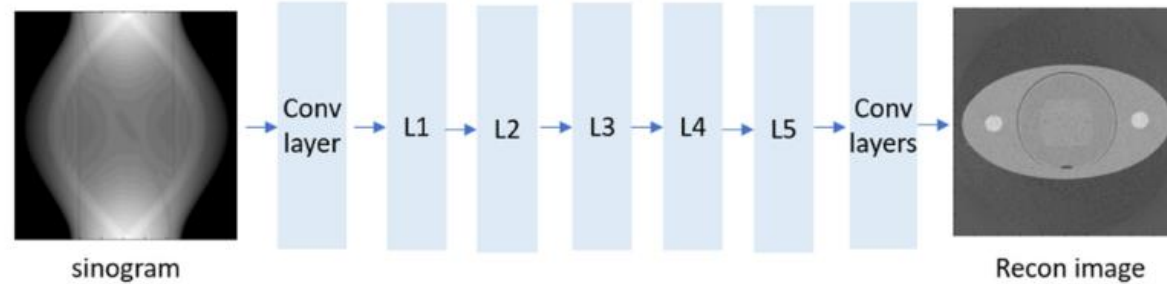
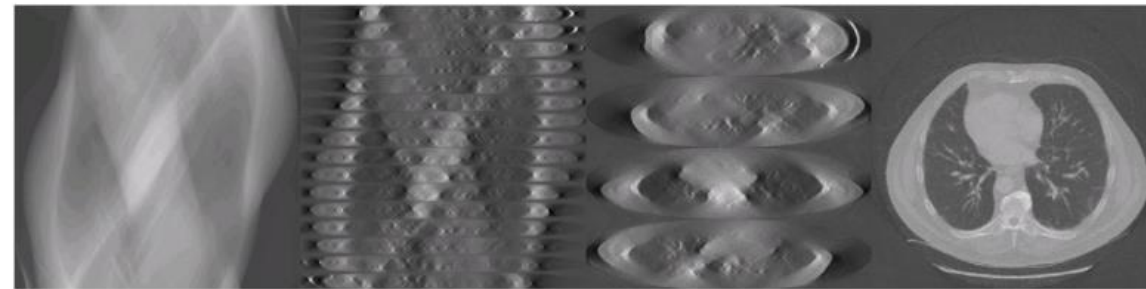
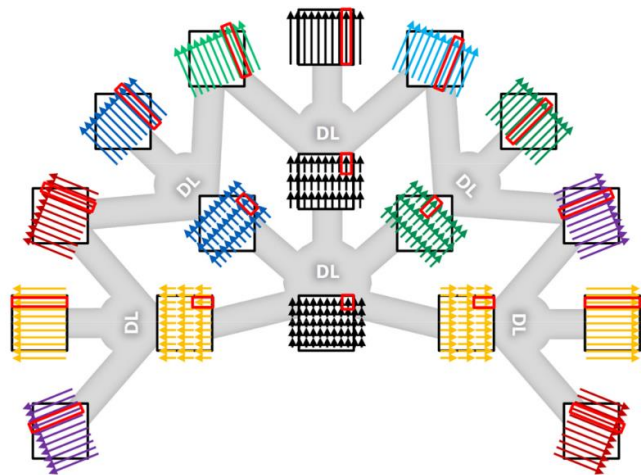
- In the hierarchical deep learning, the **image reconstruction** is fully learnt by interpreting the **domain transform** as a continuum of intermediate representations between the input and output data
 - A partial line integral is proposed as an **intermediate representation** between line integral and voxel according to a **hierarchical framework**
 - The reconstructed image is made by voxels, which are essentially line integrals over the “length” of the voxel



Direct reconstruction methods (hierarchical deep learning)

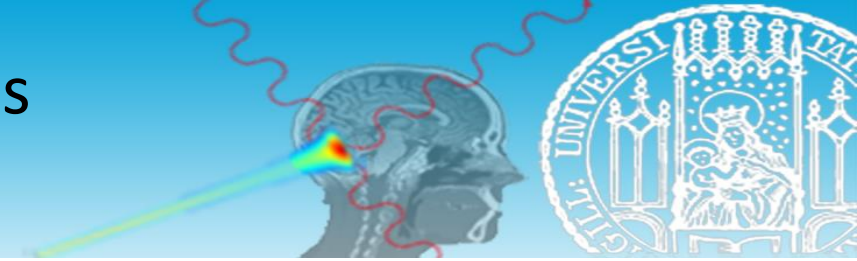


- The estimation of the partial line integrals only requires the line integrals that are at nearby angular positions and at nearby radial positions
- Similarly, the estimation of the voxel values requires as inputs only the partial line integrals that are at nearby radial and depth positions
- Sparse connections layers



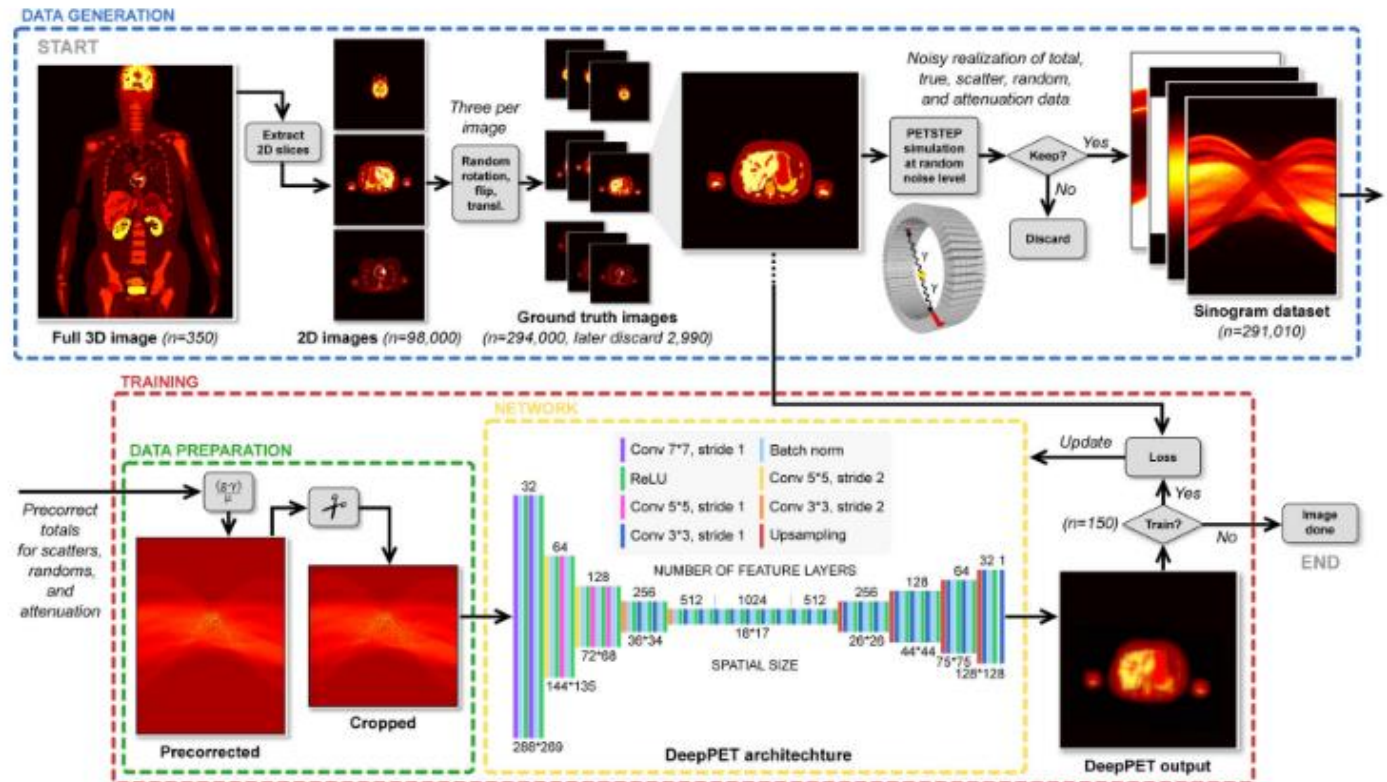
Fu, L., & De Man, B. (2019, May). A hierarchical approach to deep learning and its application to tomographic reconstruction. In *15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine* (Vol. 11072, p. 1107202). SPIE.

Direct reconstruction methods (Deep PET)



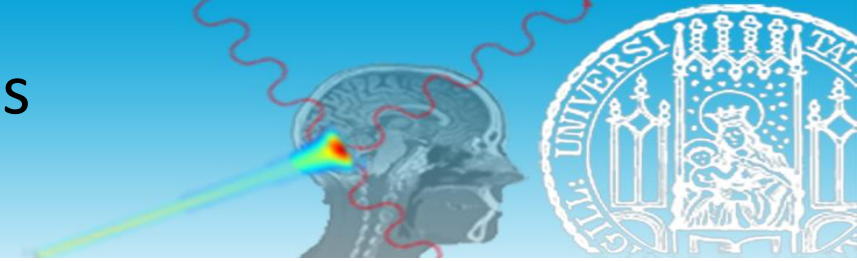
- The Deep PET is a **convolutional encoder–decoder network** without fully connected layers
- Shift-invariant mapping of the convolution to encode sinogram data into feature maps (convolutional encoding)

- Spatial down-sampling to introduce space variance (needed for **domain transform**)
- Convolutional decoding with spatial up-sampling

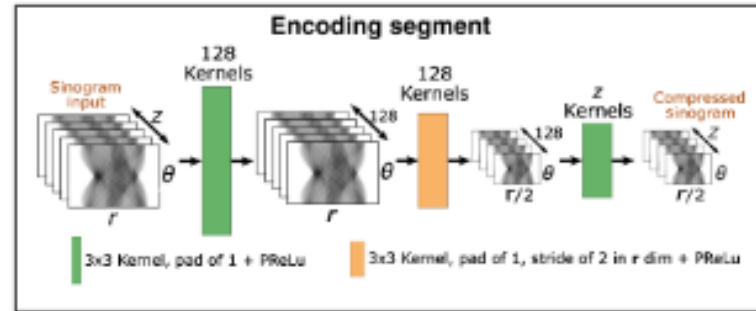


Hägström, I., Schmidlein, C. R., Campanella, G., & Fuchs, T. J. (2019). DeepPET: A deep encoder–decoder network for directly solving the PET image reconstruction inverse problem. *Medical image analysis*, 54, 253-262.

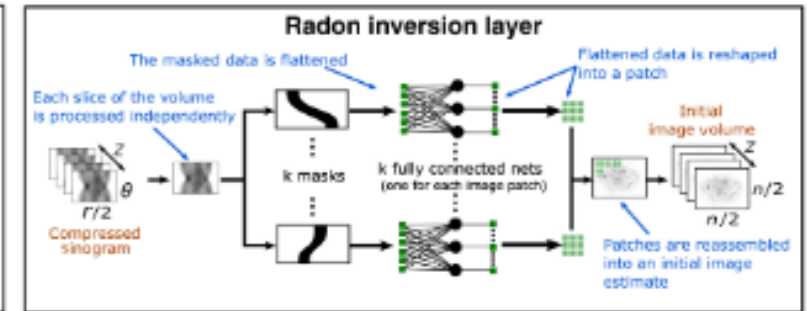
Direct reconstruction methods (DirectPET)



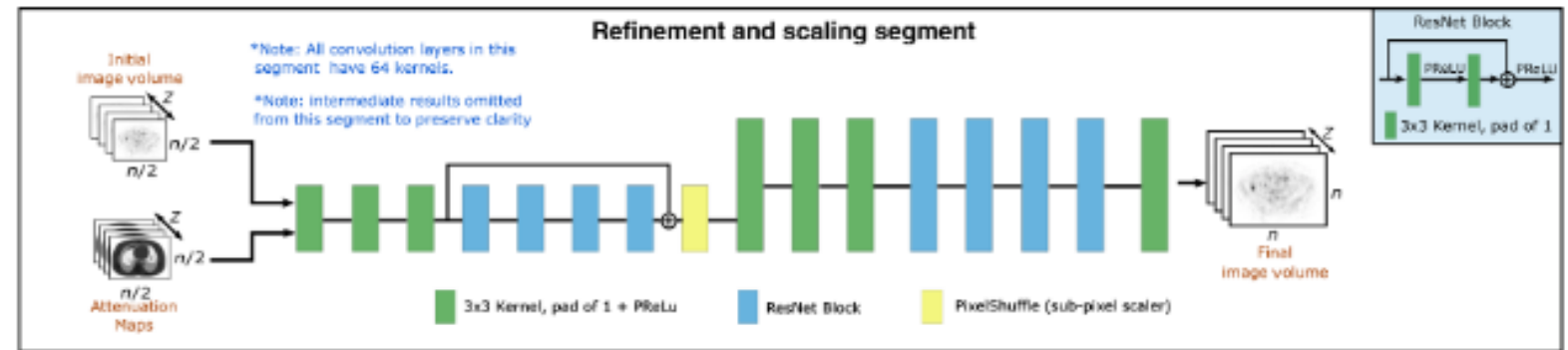
- The DirectPET is a large-scale direct neural network that performs **image reconstruction** by introducing a **Radon inversion layer**
 - An encoding segment compressing the sinogram data into a lower dimensional space
 - A **domain transformation** segment (i.e., Radon inversion) using sinogram data masking along with fully connected layers
 - A refinement and scaling segment enhancing and up-sampling the reconstructed image



(a)



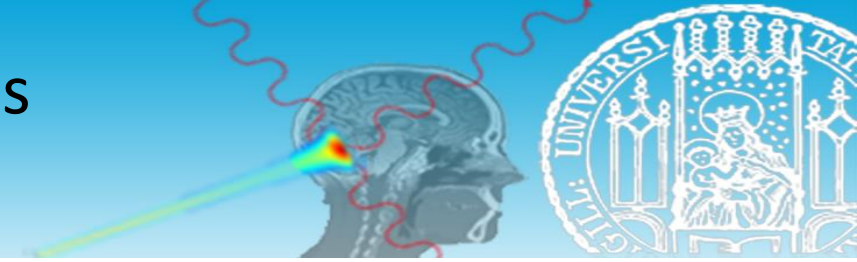
(b)



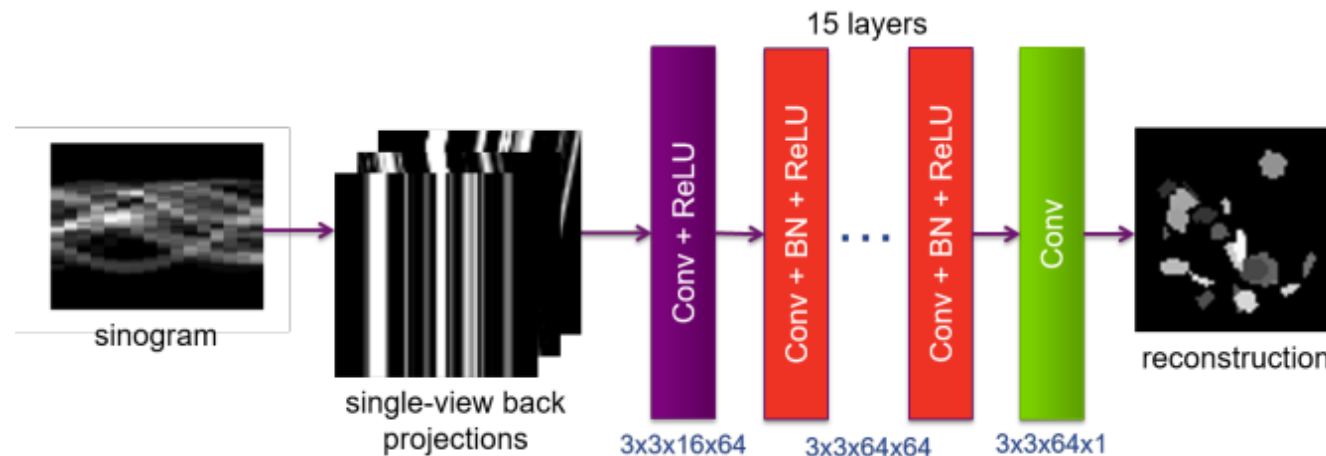
(c)

Whiteley, W., Luk, W. K., & Gregor, J. (2020). DirectPET: full-size neural network PET reconstruction from sinogram data. *Journal of Medical Imaging*, 7(3), 032503.

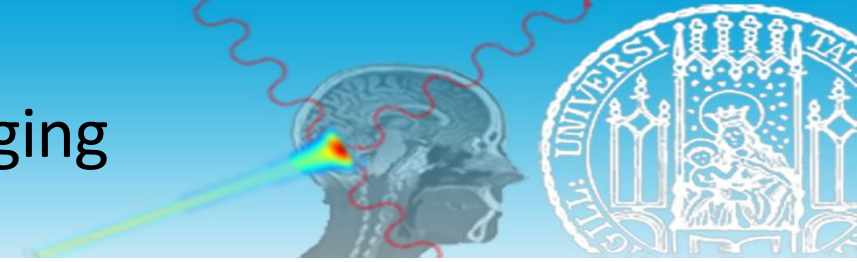
Direct reconstruction methods (DBP)



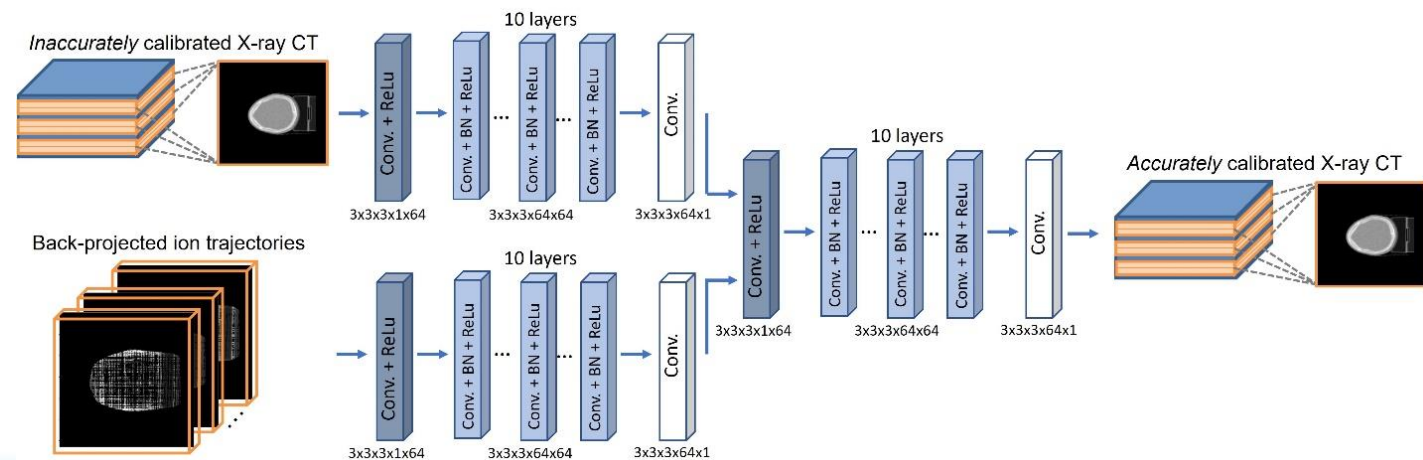
- With the Deep Back-Projection (DBP) the geometrical relationship between the **projection domain** and the **image domain** is encoded in **single-view back-projections** that are stacked and then fed as input into the convolutional neural network
 - Typically, image reconstruction needs non-linear and shift-variant mapping, as introduced by fully connected layers, down-sampling (encoding) and then up-sampling (decoding)
 - In this case, the spatial invariance of the **purely convolutional neural network** is retained



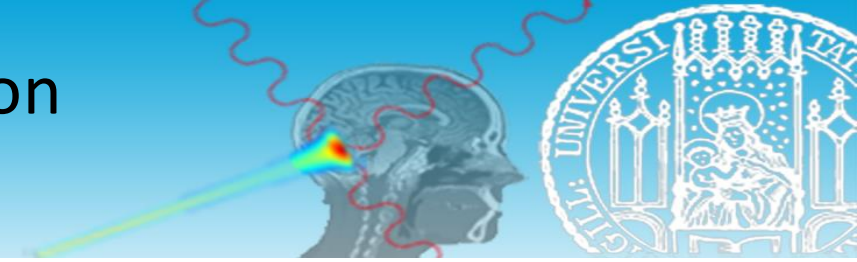
Deep reconstruction in ion imaging



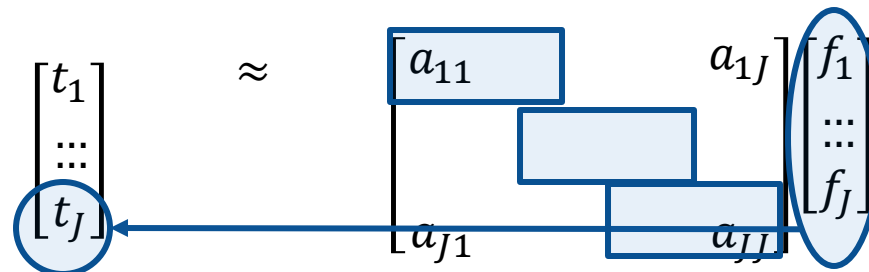
- Inspired by the ML literature about tomographic image reconstruction but extended toward hybrid X-ray and ion imaging
- With the “**deep back-projection**”, the geometrical relationship between the **projection domain** and the **image domain** is encoded in single-view back-projections that are stacked and fed as input to the **convolutional neural network**
- The network is designed as a two input branches (one for the **X-ray CT image** and the other for the **ion radiographies**) followed by an integration branch
- Two ion radiographies (two projection angles) are clustered according to the **scattering angles of the ion trajectories** (ten clusters for each projection angles)



Unrolled iterative reconstruction methods



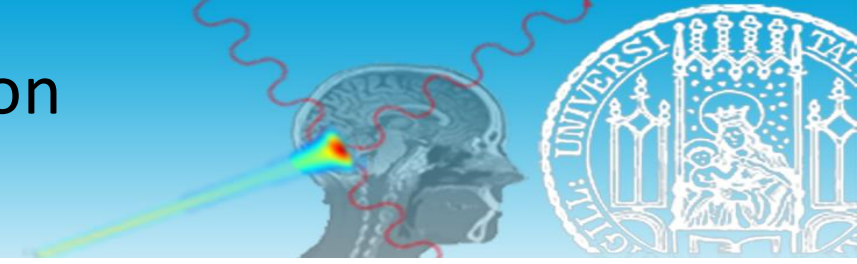
- **Interpretability** and **generalizability** in medical applications play fundamental roles but deep neural networks are usually difficult to interpret because of the huge number of parameters
- Algorithm unrolling or unfolding is proposed to improve the **interpretability** and the **generalizability** (i.e., overfitting) of the deep neural network
 - Explicitly connected to the iterative algorithms used in imaging (and signal processing)
 - Explicitly based on domain knowledge as in imaging (and signal processing)
- If the normal operator of the forward-projection model is a convolution (i.e., denoising and deblurring in the back-projection model), convolutional neural networks take part of **unrolled iterative reconstruction methods**



- Denoising and deblurring can be described by convolutional layers

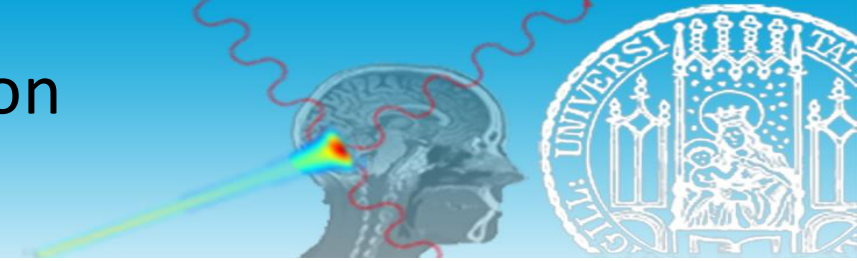


Unrolled iterative reconstruction methods

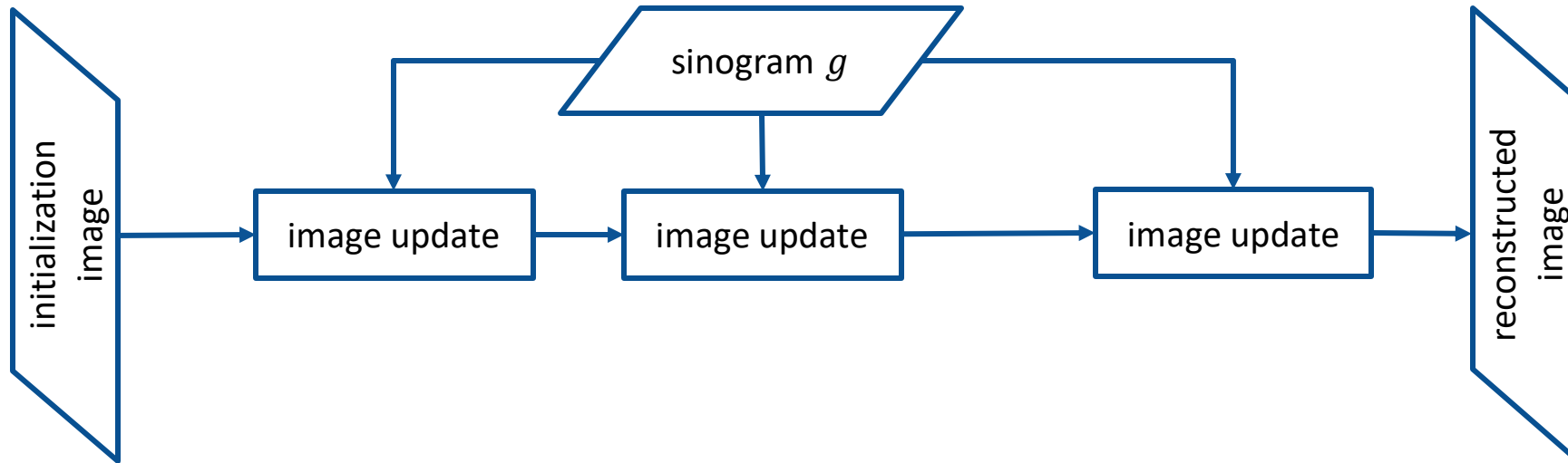


- Each iteration of the algorithm is represented as one **layer of the network**
 - Concatenating these layers forms a deep neural network
 - The number of layers in a deep network is typically much smaller than the number of iterations required in an iterative algorithm
- The network can be trained through all layers (i.e., end-to-end training) or layer per layer (i.e., sequential training)

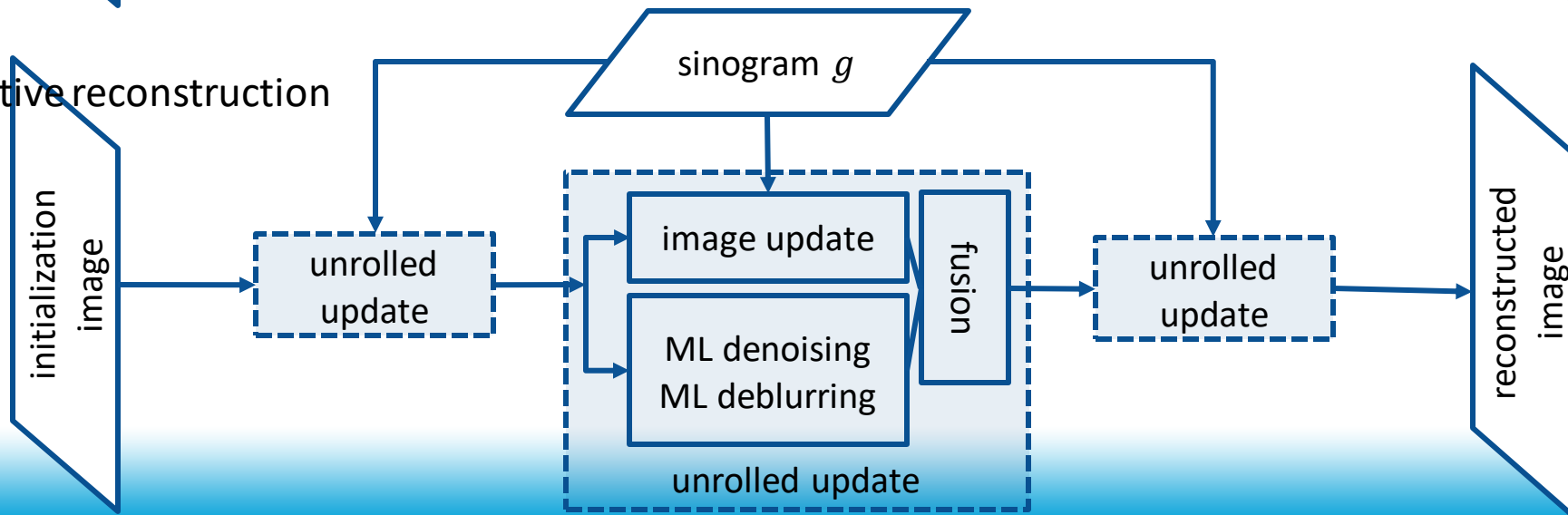
Unrolled iterative reconstruction methods



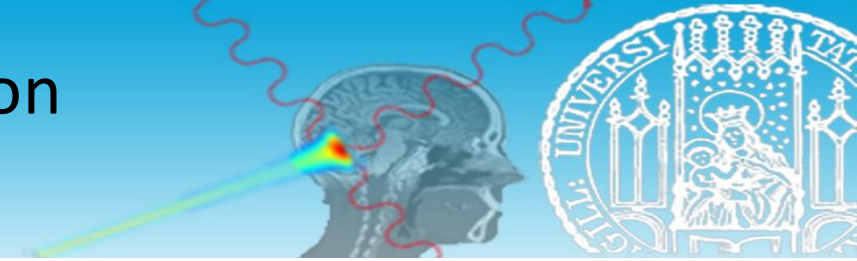
- “Unfolded” iterative reconstruction



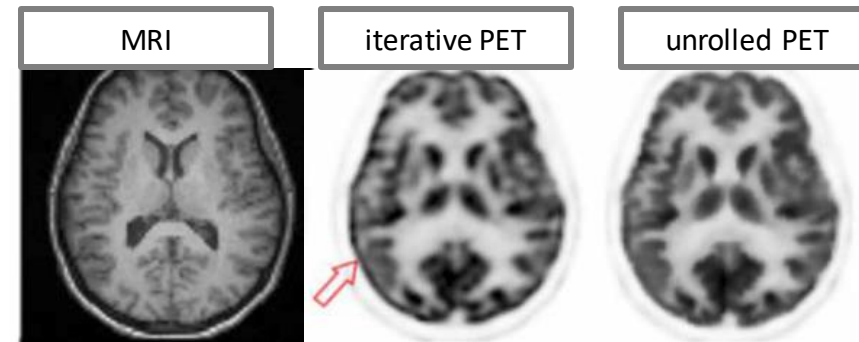
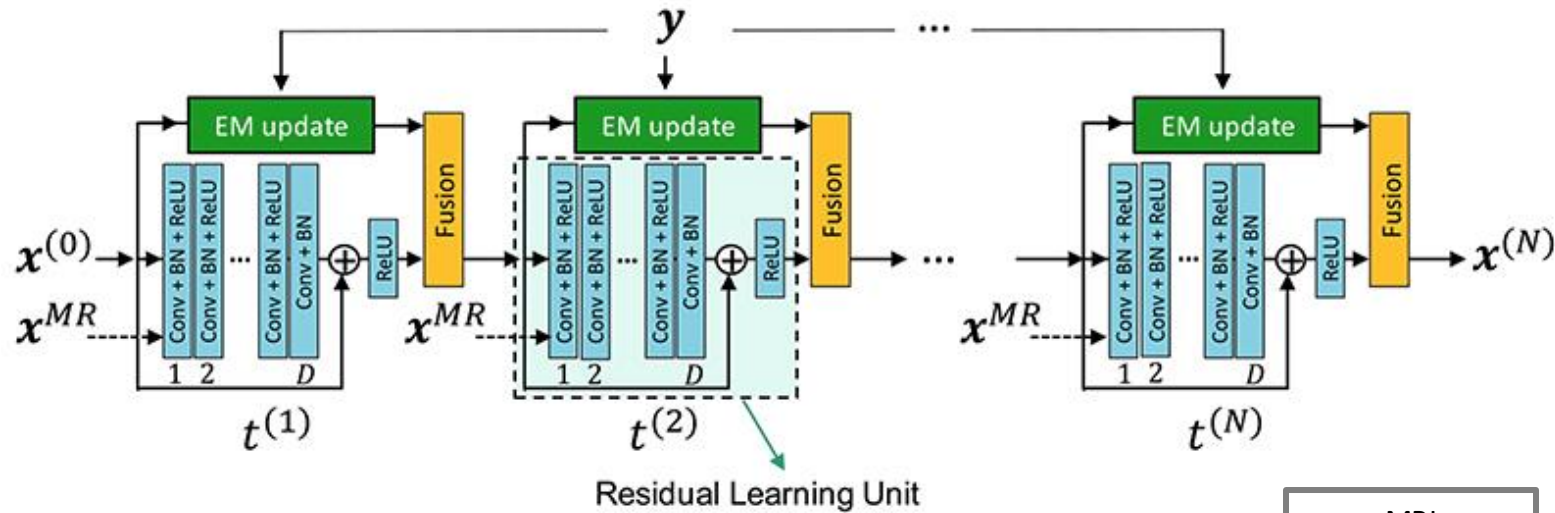
- Unrolled iterative reconstruction



Unrolled iterative reconstruction methods

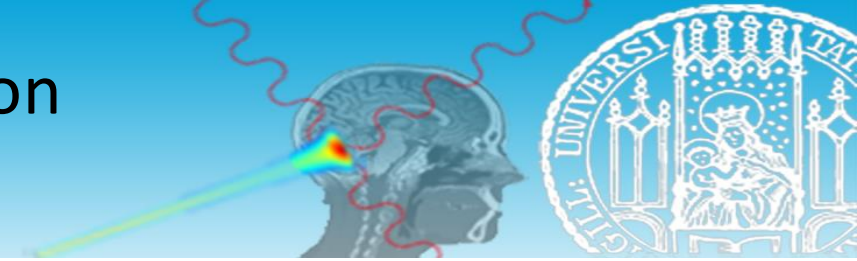


- Residual network based on convolutional layers trained during reconstruction

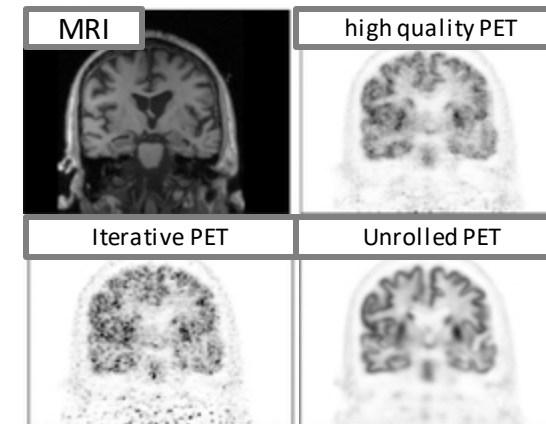
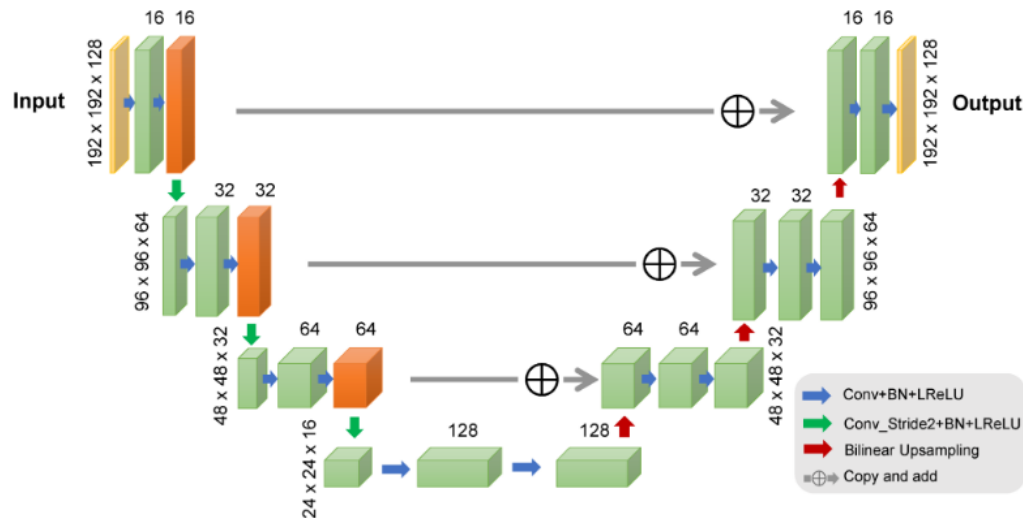


Mehranian, A., & Reader, A. J. (2020). Model-based deep learning PET image reconstruction using forward-backward splitting expectation-maximization. *IEEE transactions on radiation and plasma medical sciences*, 5(1), 54-64.

Unrolled iterative reconstruction methods

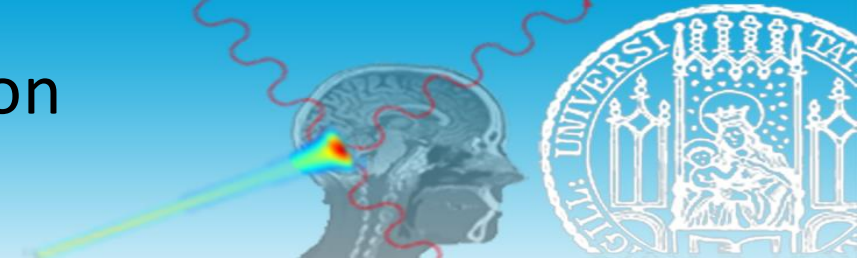


- Unrolled iterative reconstruction methods are typically **image domain networks** intended to reduce the **image noise** due to low count statistics PET sinograms based on prior information from high quality imaging
 - The network, a modified fully convolutional network (U-net), is trained using **patient-specific** prior information (i.e., MRI image) and the measured data (i.e., PET sinogram)
 - The training of the network is embedded in the tomographic image reconstruction algorithm

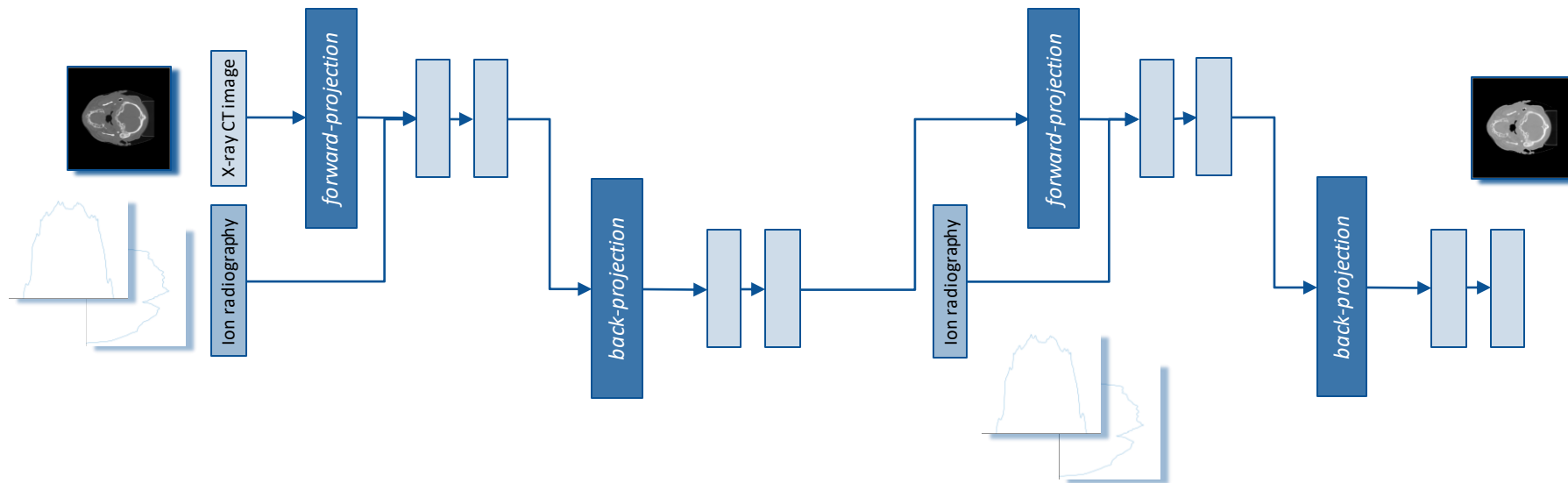


Gong, K., Catana, C., Qi, J., & Li, Q. (2018). PET image reconstruction using deep image prior. *IEEE transactions on medical imaging*, 38(7), 1655-1665.

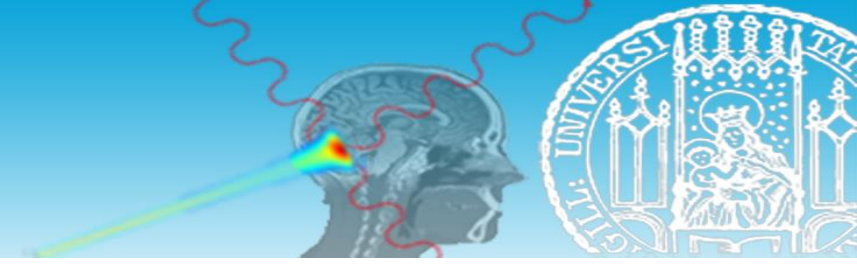
Unrolled iterative reconstruction methods



- The state-of-the-art in deep learning for the solution of inverse problems is the **Learned PrimalDual** architecture (i.e., **physics-informed design**),
- The Learned PrimalDual architecture explicitly contains the forward- and back-projections models, which are applied to the input data multiple times, as an **unrolled tomographic image reconstruction algorithm**



Outlook and conclusions



- Analytical image reconstruction is based on the [Fourier Slice Theorem](#) that mathematically defines the correspondence between the sinogram and the image
- Numerical image reconstruction is based on the forward-projection model that defines the correspondence between the measurements and the unknowns, so that the iterative algorithm matches the [forward-projection model](#) with the [measurements](#)
- Deep image reconstruction is instead based on the learning of the [ground truth image](#)
 - The network can include the [domain conversion](#) in different ways
 - As fully connected layer, thus informing the network about the space variance of the correspondence between image and sinogram
 - As a decoding branch, thus introducing space variance to the network
 - As input to the network in form of back-projection tensor