Deep Learning in CT Image Formation

Marc Kachelrieß

German Cancer Research Center (DKFZ)

Heidelberg, Germany

www.dkfz.de/ct

DEUTSCHES KREBSFORSCHUNGSZENTRUM IN DER HELMHOLTZ-GEMEINSCHAFT

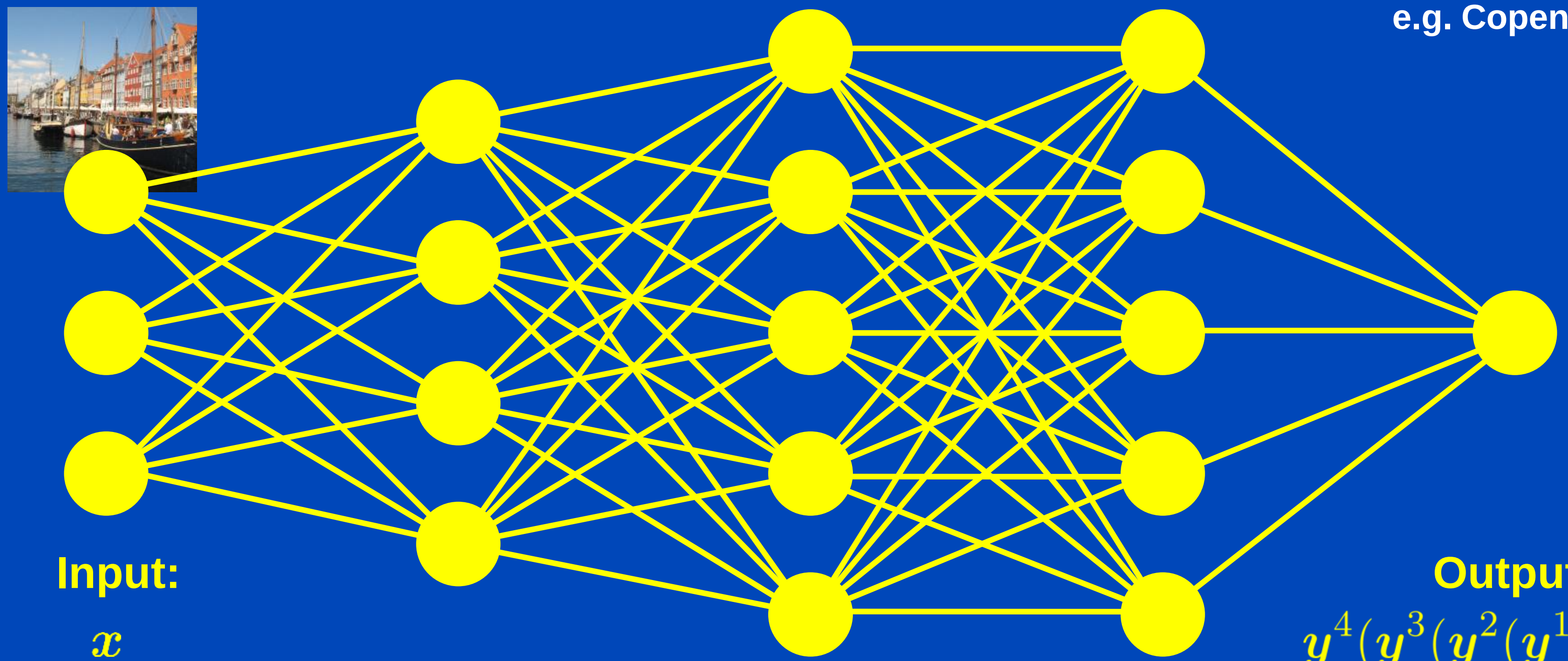
Fully Connected Neural Network

- Each layer fully connects to previous layer
- Difficult to train (many parameters in *W* and *b*)
- Spatial relations not necessarily preserved

Input | **Hidden** | **Hidden** | **Hidden** | **Output**

e.g. 512×512×3 pixels
e.g.

e.g. 1 label
e.g. Copenhagen

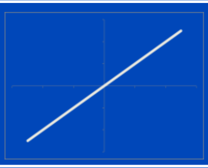
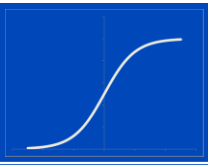
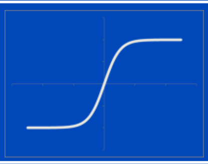
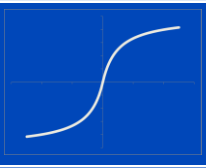
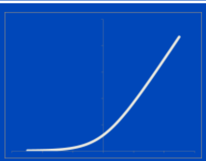


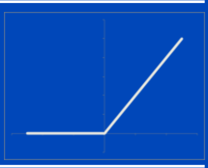
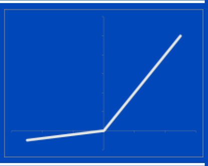
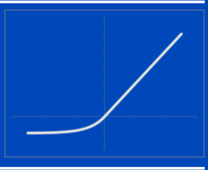
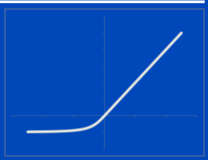
Input: x

Output: $y^4(y^3(y^2(y^1(x))))$

$\boldsymbol{y}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{W}\cdot\boldsymbol{x}+\boldsymbol{b})$ with $\boldsymbol{f}(\boldsymbol{x}) = \big(f(x_1), f(x_2), \ldots\big)$ point-wise scalar, e.g. $f(x) = x\vee 0 = \mathrm{ReLU}$

Activation Functions

Function	Equation	Plot
Identity	$f(x) = x$	
Sigmoid	$f(x) = \dfrac{1}{1 + e^{-x}}$	
Hard sigmoid	$f(x) = \begin{cases} 0 & \text{for } x < -\alpha \\ \frac{\alpha + x}{2\alpha} & \text{for } -\alpha \leq x < \alpha \\ 1 & \text{for } x \geq \alpha \end{cases}$	
Tanh	$f(x) = \dfrac{2}{1 + e^{-2x}} - 1$	
Softsign	$f(x) = \dfrac{x}{1 + \lvert x \rvert}$	
Softplus	$f(x) = \log(1 + \exp x)$	

Function	Equation	Plot
ReLU	$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	
Leaky ReLU	$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	
ELU	$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	
Inverse square root LU	$f(x) = \begin{cases} \frac{x}{\sqrt{1 + \alpha x^2}} & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	
…	…	…

Gradient Descent

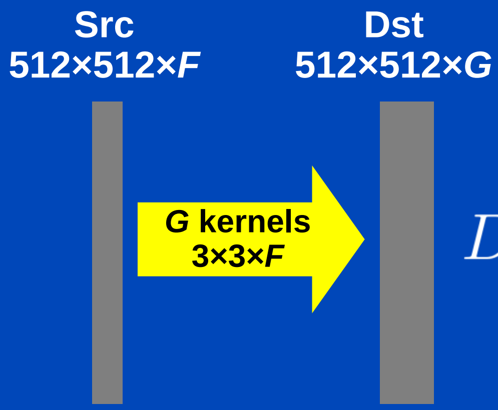
- Walk along the direction of the negative gradient
- Steepest descent
- Learning rate η

$$\boldsymbol{w}^{\text{new}} = \boldsymbol{w}^{\text{old}} - \eta \, \nabla_{\boldsymbol{w}} L(\boldsymbol{x}_n, \boldsymbol{y}_n, \boldsymbol{w})$$

- Easy to understand, but not optimal
- Methods in use
  - Batch gradient descent
  - Sochastic gradient descent
  - Mini-batch gradient descent
  - Conjugate gradient descent
  - Quasi Newton methods
  - Momentum methods
  - …

Convolutional Neural Network (CNN)

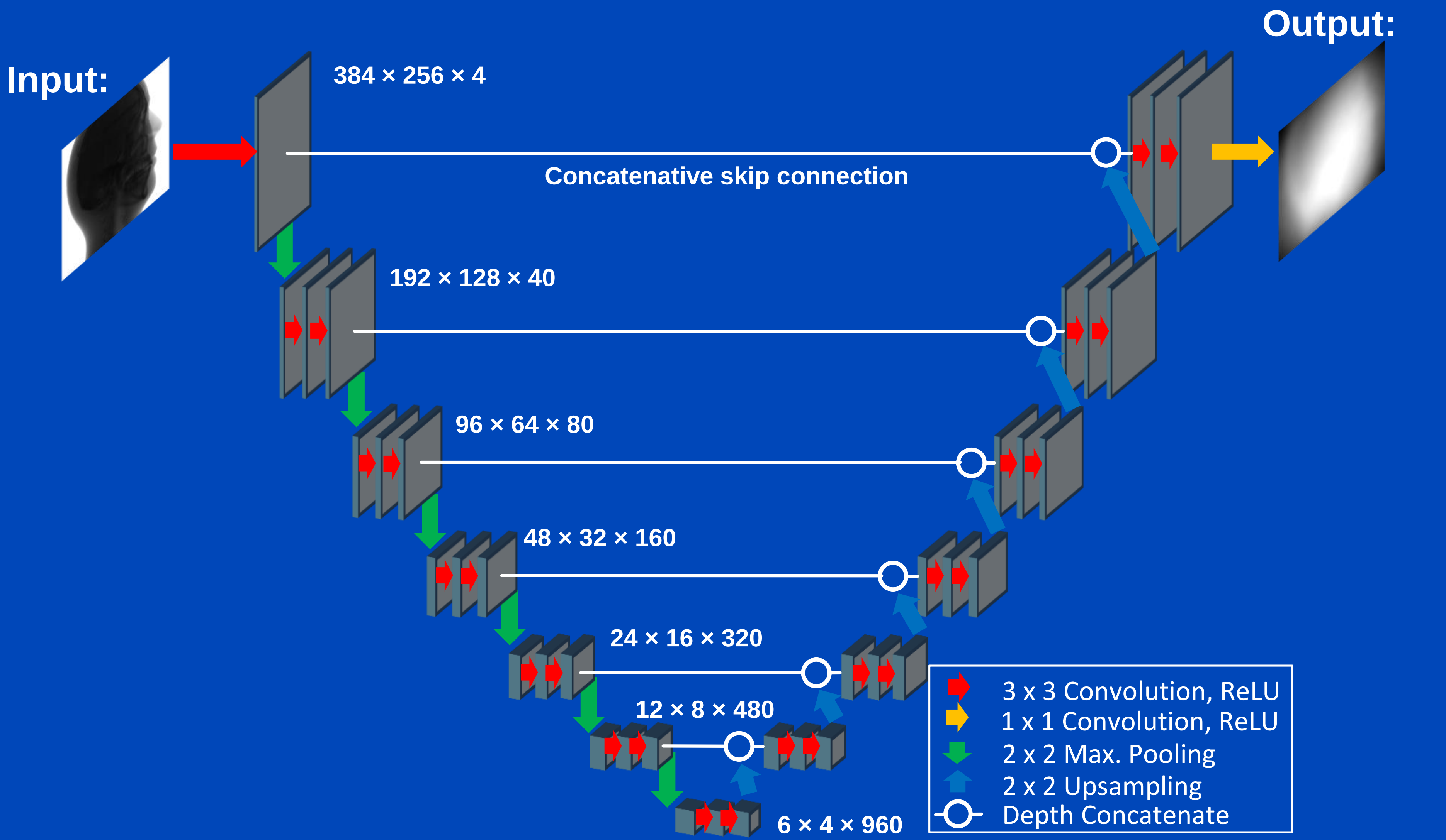
- Replace dense *W* in $y(x) = f(W \cdot x + b)$ by a sparse matrix *W* with sparsity being of convolutional type.
- CNNs consist (mainly) of convolutional layers.
- Convolutional layers are not fully connected.
- Convolutional layers are connected by small, say 3×3, convolution kernels whose entries need to be found by training.
- CNNs preserve spatial relations to some extent.



$$D_{i,j,g} = \sum_f S_{i,j,f} * K^g_{i,j,f} = \sum_{a,b,f} S_{i-a,j-b,f} K^g_{a,b,f}$$

Attention: No convolution in depth direction!

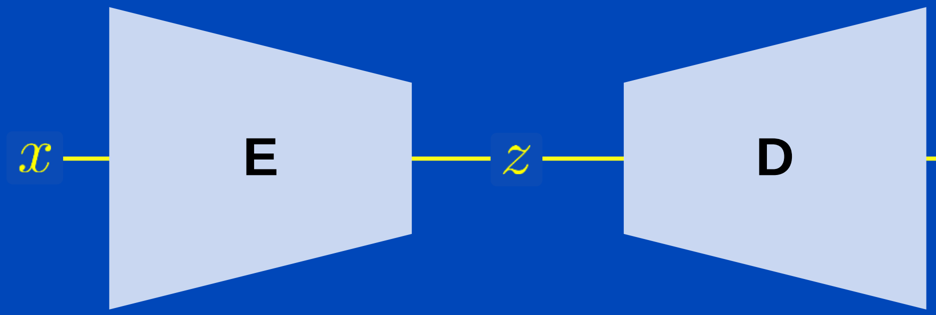
U-Net[1]



[1] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. Proc. MICCAI:234-241, 2015.

What is an Autoencoder?

- In and output domain are the same, here *x*.
- Bottleneck *z* enforces the encoder and decoder to do a good job.

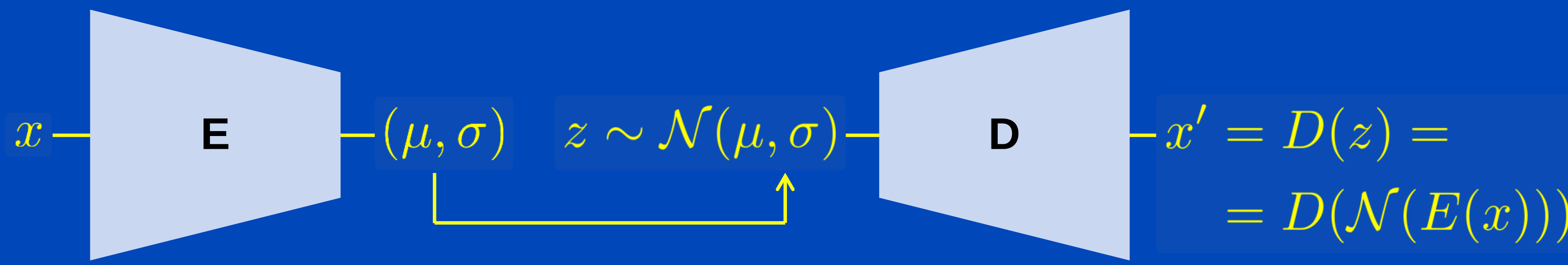


$$x' = D(z) = \\ = D(E(x))$$

- Examples:
  - Principal component analysis (linear autoencoder), lossless
  - PCA with dimensionality reduction (nonlinear due to clipping), lossy
  - Image compression and decoding, e.g. jpeg, lossy
- Latent space typically not interpretable.

What is a Variational Autoencoder?

- Make latent space regular.
- Allow to sample in latent space from a given distribution, here: normal distribution.



- The VAE is a generative model.
- It allows to generate new data by sampling new values from the normal distribution.

Loss Function

- The neural network parameters (weights and biases) *w* are chosen by minimizing a loss function (cost function)

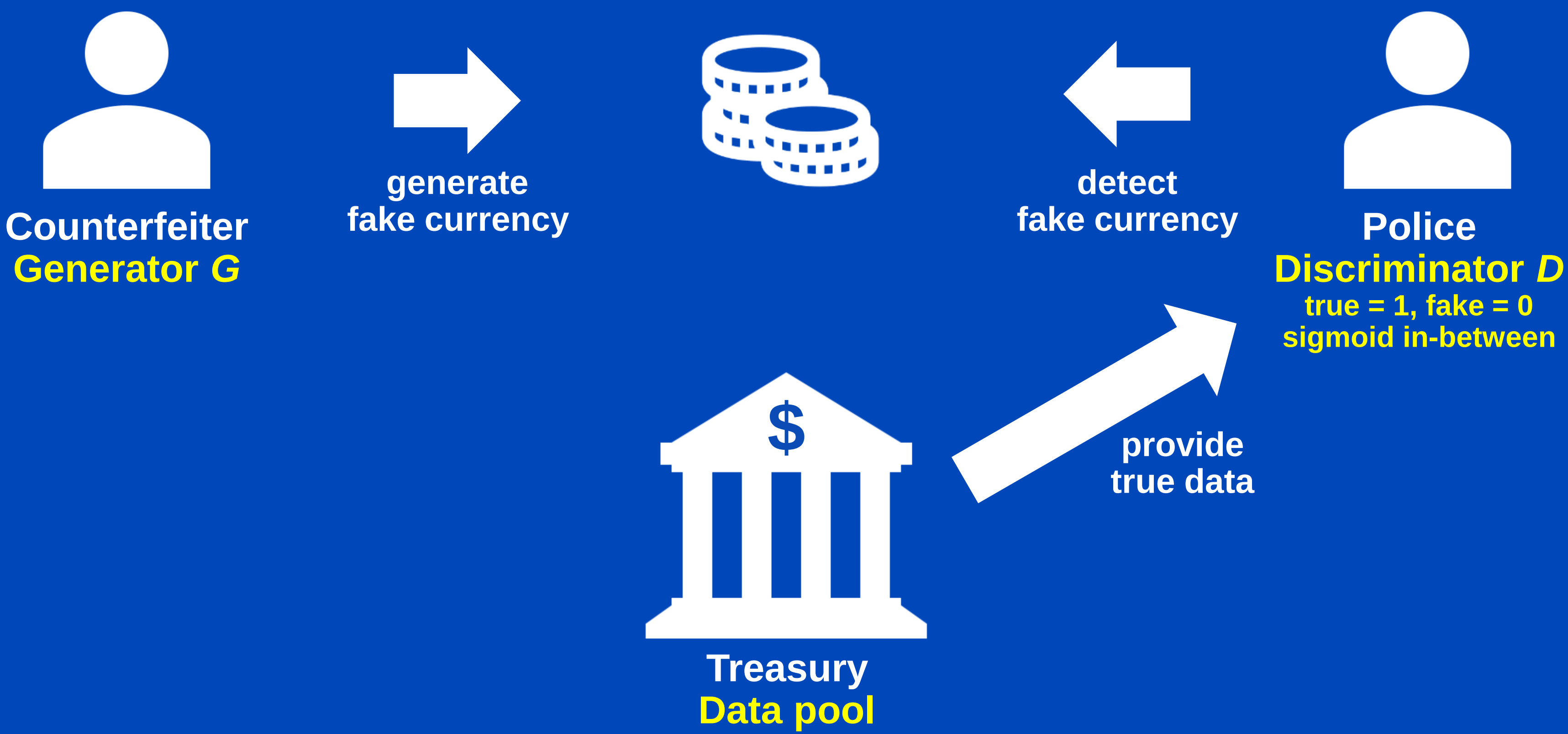
$$\boldsymbol{w} = \arg\min_{\boldsymbol{w}} \sum_{n=1}^{N} L(\boldsymbol{x}_n, \boldsymbol{y}_n, \boldsymbol{w})$$

  with x_n being the training data input, $y(x_n, w)$ being the network output, and y_n being the so-called labels, i.e. the training target, and N being the number of training samples.
- An example for such a loss function is the MSE loss

$$L(\boldsymbol{x}_n, \boldsymbol{y}_n, \boldsymbol{w}) = \big(\boldsymbol{y}(\boldsymbol{x}_n, \boldsymbol{w}) - \boldsymbol{y}_n\big)^2$$

Generative Adversarial Network[1] (GAN)

- Useful, if no direct ground truth (GT) is available, the training data are unpaired, unsupervised learning



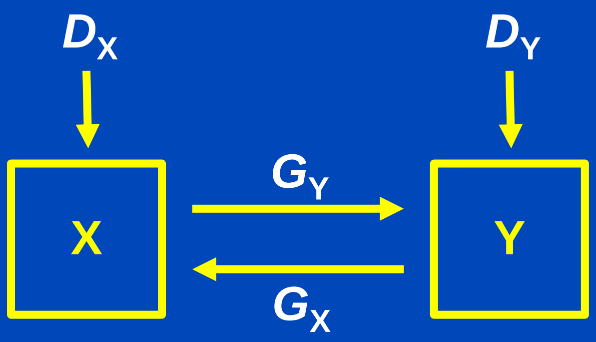
[1] I. Goodfellow et al. Generative Adversarial Nets, arXiv 2014

Generative Adversarial Network (GAN)

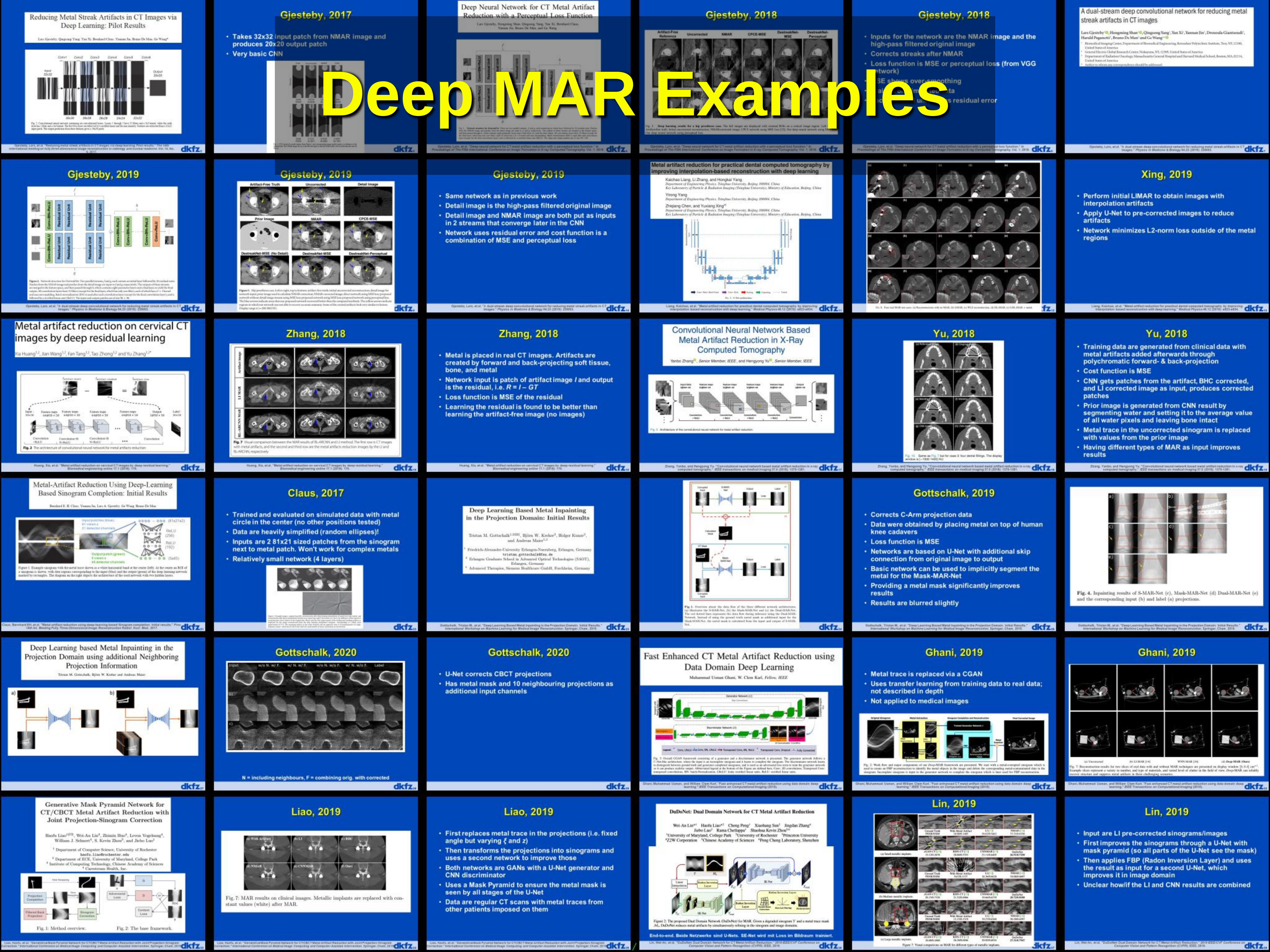
- **Typical loss function and minimax game:**

$$\min_G \max_D L(D,G) := \mathrm{E}_x \ln\big(1 - D(G(x))\big) + \mathrm{E}_y \ln D(y)$$

- **Conditional GAN[1]**
  - **Conditinal GANs sample the generator input *x* not from a uniform distribution but from a conditional distribution, e.g. noisy CT images.**
  - **Need some measure to ensure similarity to input distribution (e.g. pixelwise loss added to the minimax loss function)**
- **Cycle GAN[2]**
  - **Two GANs (X $\rightarrow$ Y and Y $\rightarrow$ X)**
  - **Demand cyclic consistency, i.e. $x = G_X(G_Y(x))$ and $y = G_Y(G_X(x))$**

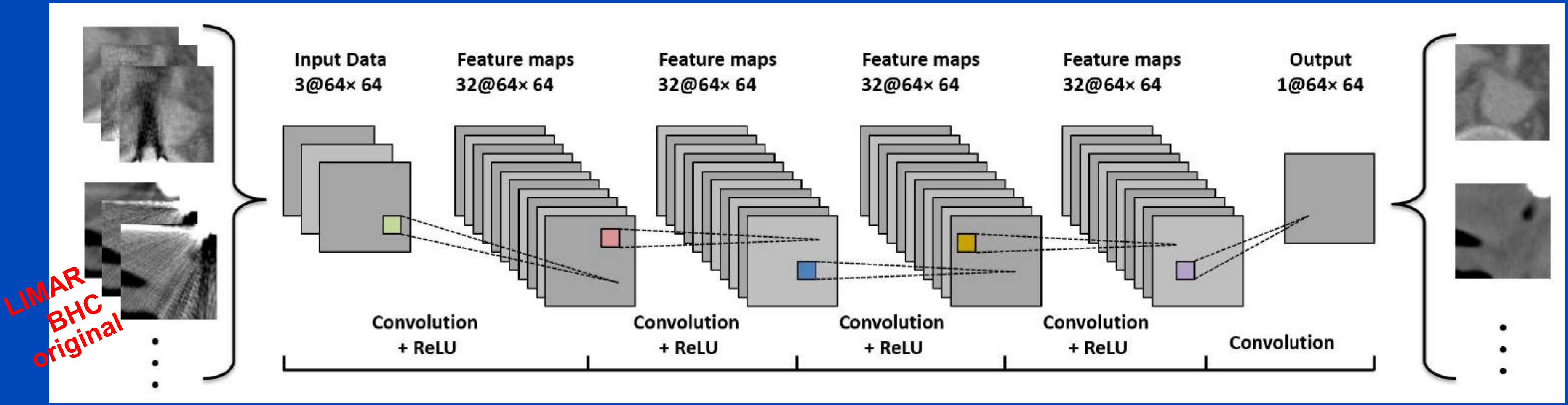


Deep MAR Examples

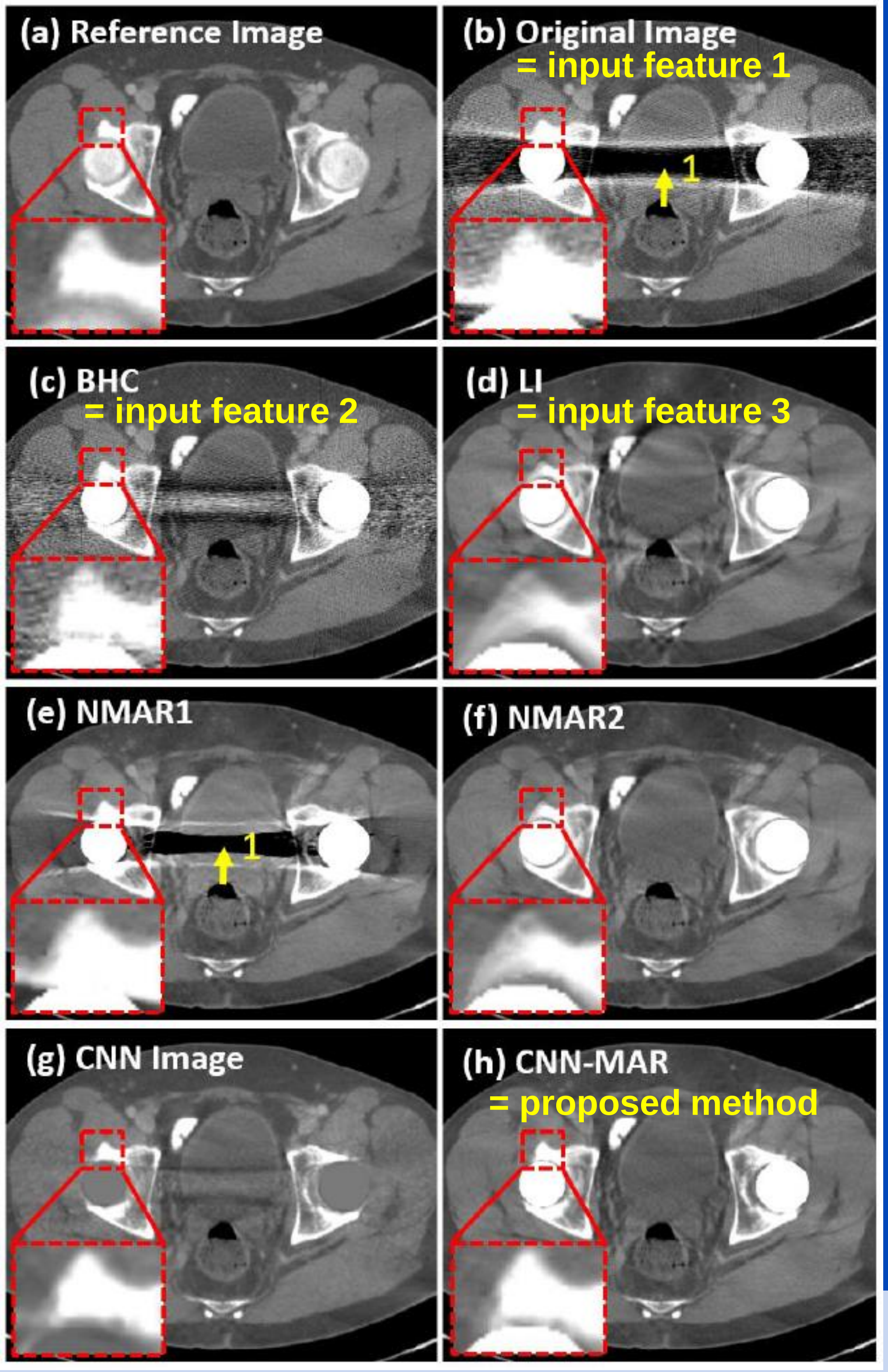


MAR Example

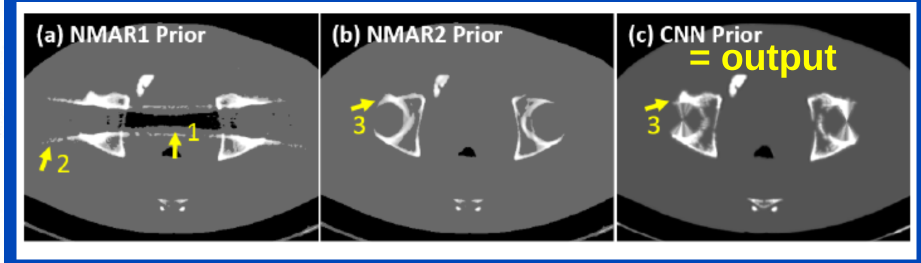
- Deep CNN-driven patch-based combination of the advantages of several MAR methods trained on simulated artifacts



- followed by segmentation into tissue classes
- followed by forward projection of the CNN prior and replacement of metal areas of the original sinogram
- followed by reconstruction

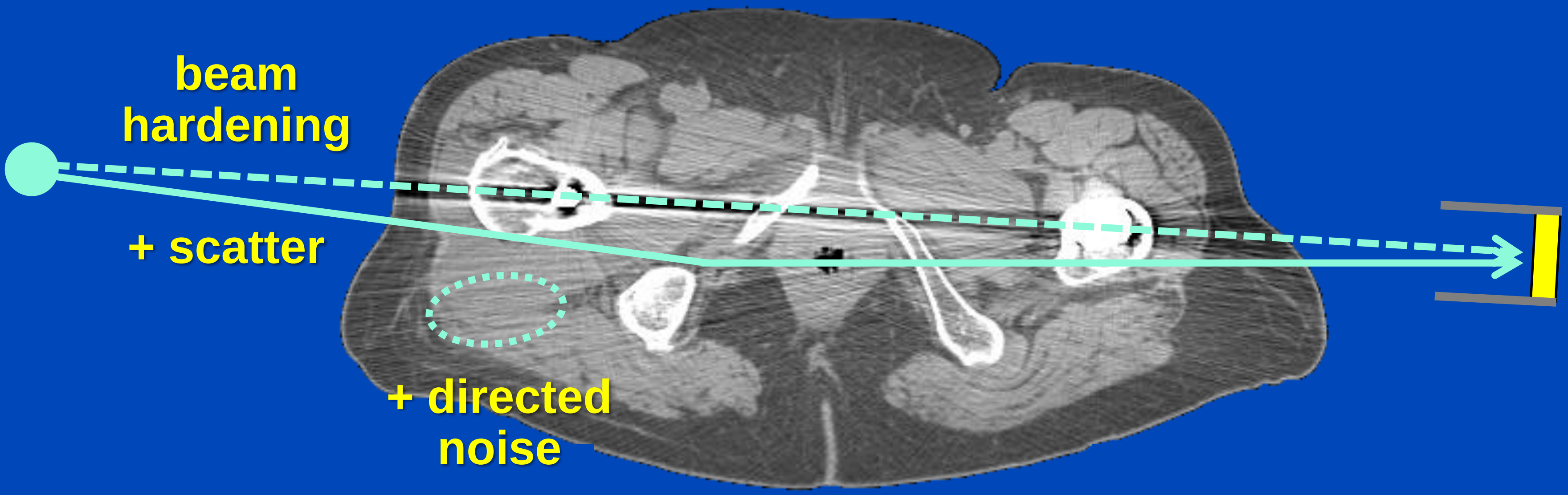


(a) Reference Image
(b) Original Image
= input feature 1
1
(c) BHC
= input feature 2
(d) LI
= input feature 3
(e) NMAR1
1
(f) NMAR2
(g) CNN Image
(h) CNN-MAR
= proposed method



(a) NMAR1 Prior
1
2
(b) NMAR2 Prior
3
(c) CNN Prior
= output
3

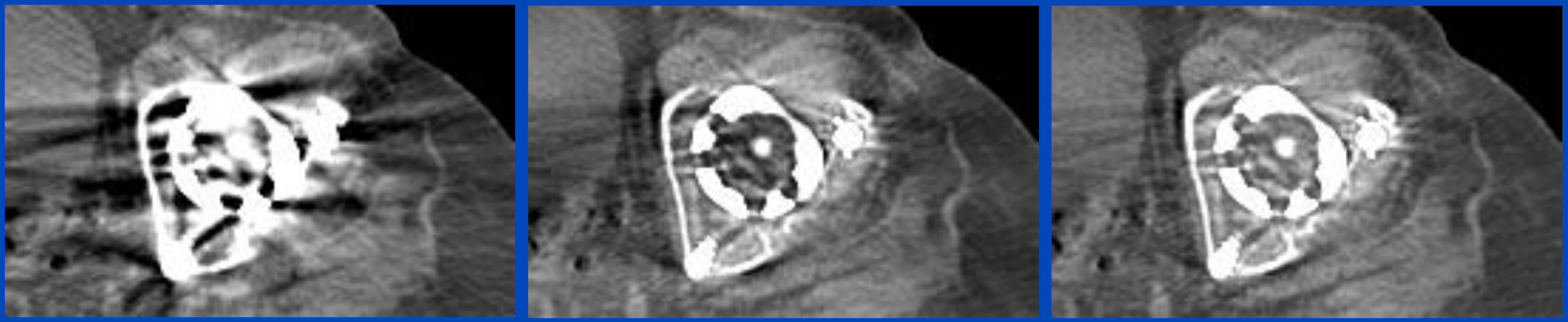
Metal artifacts are



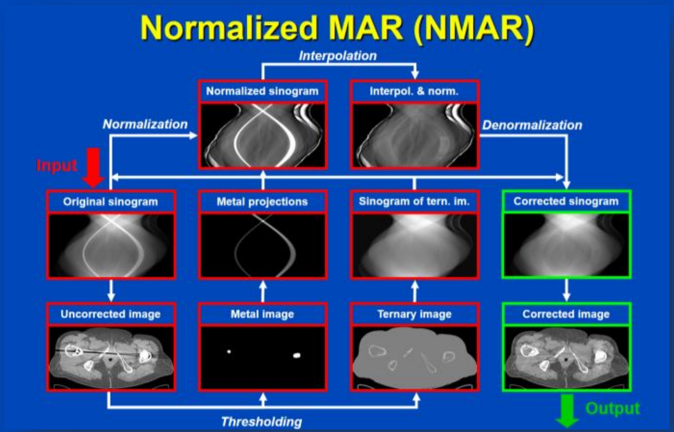
+ increased susceptibility to sampling artifacts and motion.

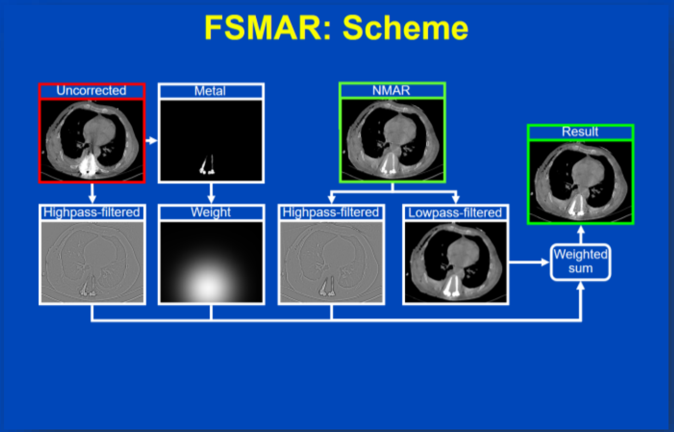
MAR without Machine Learning is a Good Alternative: Frequency Split Normalized MAR[1,2]

Uncorrected FSLIMAR FSNMAR



Patient with bilateral hip prosthesis, Somatom Definition Flash, (C=40/W=500).



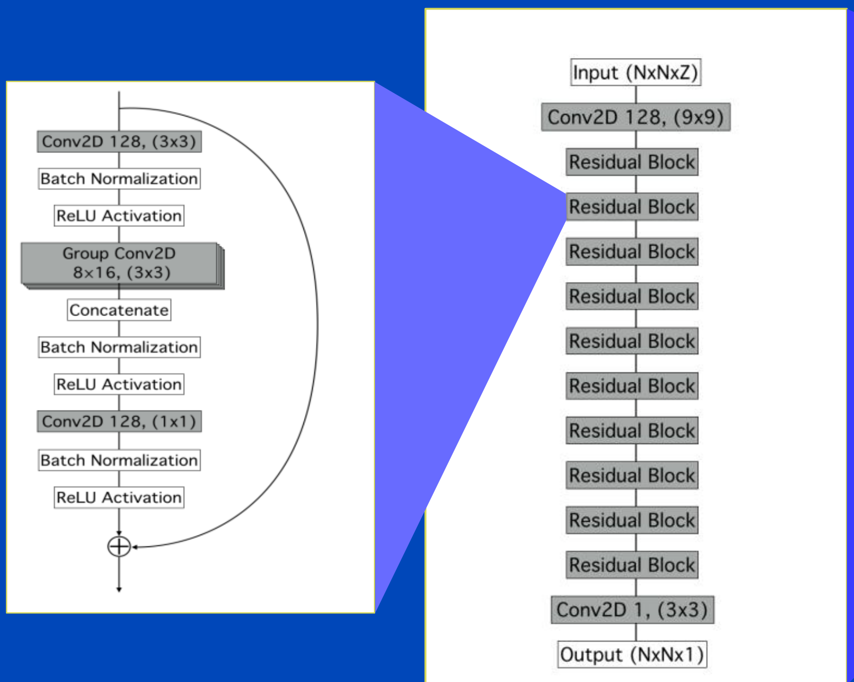


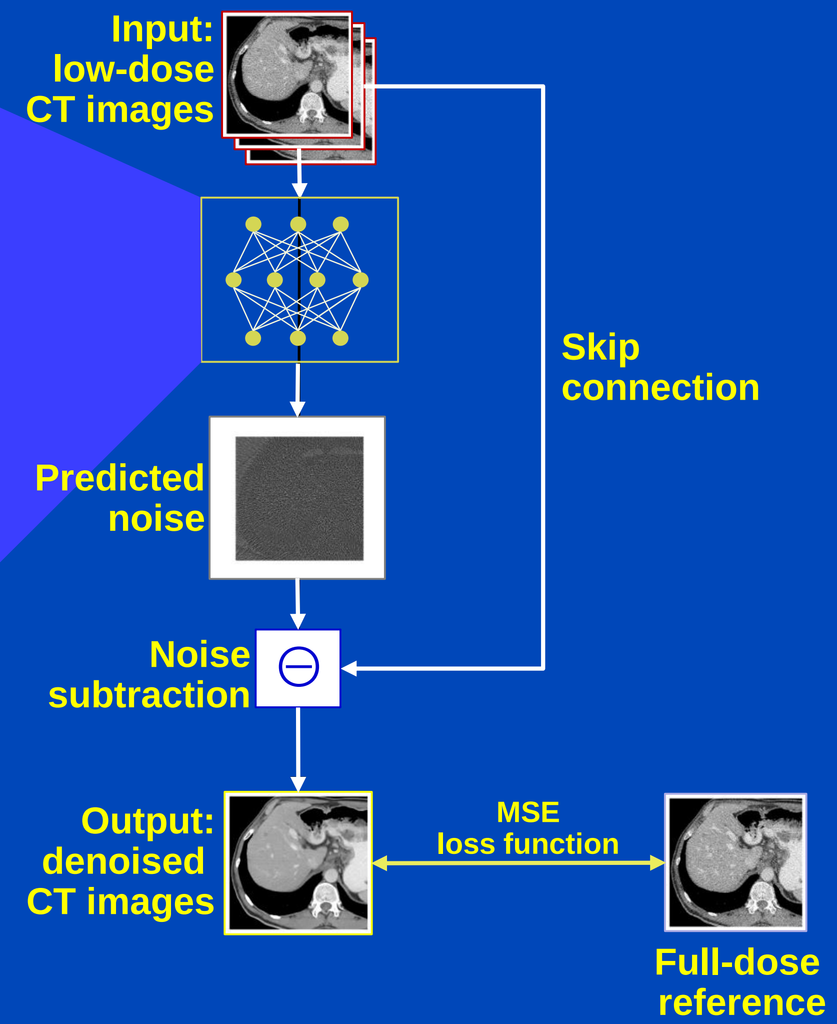
[1]E. Meyer, M. Kachelrieß. Normalized metal artifact reduction (NMAR) in computed tomography. Med. Phys. 37(10):5482-5493, Oct. 2010.
[2]E. Meyer, M. Kachelrieß. Frequency split metal artifact reduction (FSMAR) in CT. Med. Phys. 39(4):1904-1916, April 2012.

Summary on Deep MAR

- **Most common uses for networks:**
  - Improve image quality in image domain after MAR
  - Use network for the sinogram inpainting
  - Produce a prior image, e.g. for NMAR
- **Additional observations:**
  - Training data are often produced by segmenting an artifact-free CT image, adding metal and applying a polychromatic forward projection to different types of tissue separately.
  - As of today, it seems hard to outperform NMAR, or hard to give convincing clinical examples.

Noise Removal Example

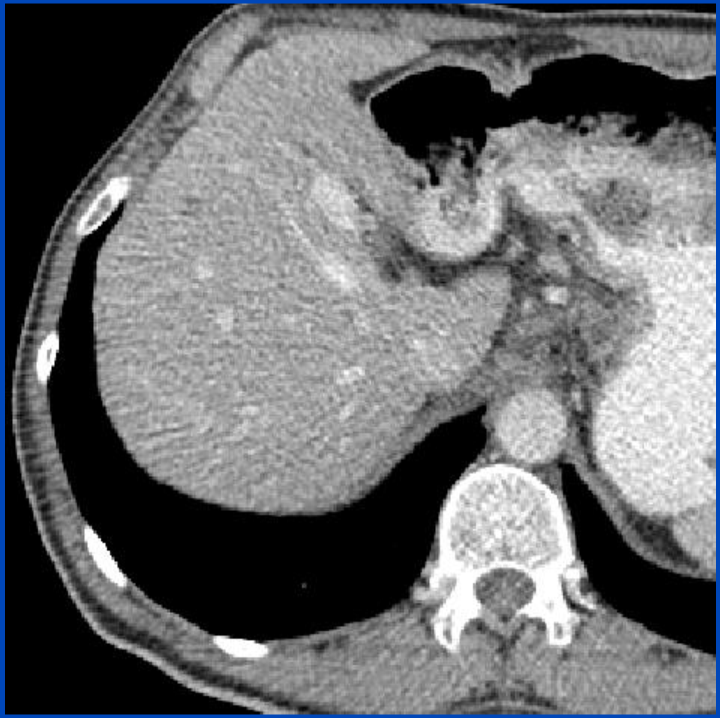


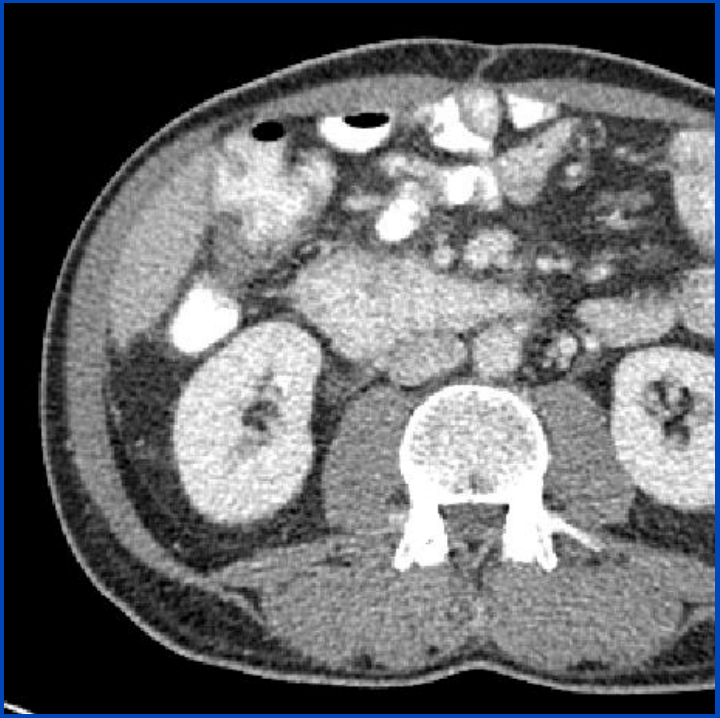


- Architecture based on state-of-the-art networks for image classification (ResNet).
- 32 conv layers with skip connections
- About 2 million tunable parameters in total
- Input is arbitrarily-size stack of images, with a fixed number of adjacent slices in the channel/feature dimension.

Andrew D. Missert, Shuai Leng, Lifeng Yu, and Cynthia H. McCollough. Noise Subtraction for Low-Dose CT Images Using a Deep Convolutional Neural Network. Proceedings of the 5th CT-Meeting: 399-402, 2018.

Noise Removal Example





Low dose images (1/4 of full dose)

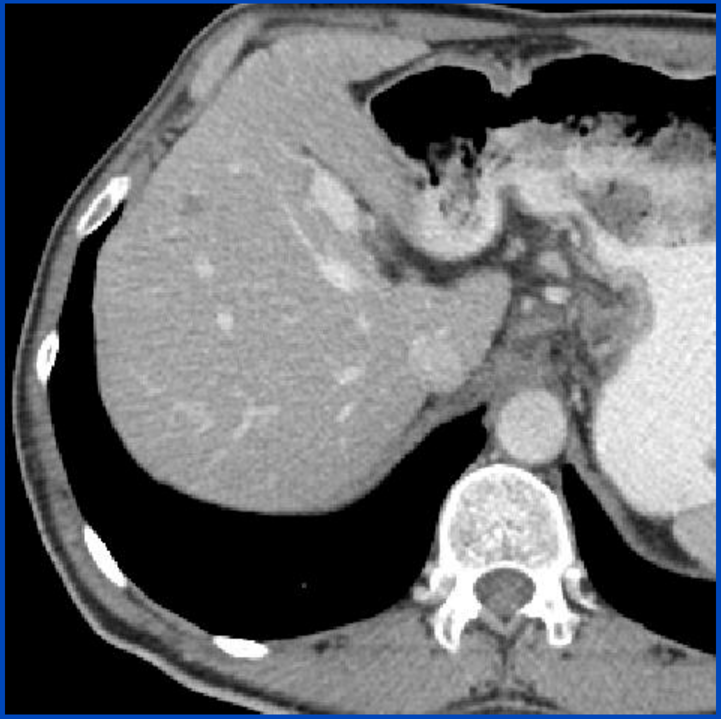
Andrew D. Missert, Shuai Leng, Lifeng Yu, and Cynthia H. McCollough. Noise Subtraction for Low-Dose CT Images Using a Deep Convolutional Neural Network. Proceedings of the 5th CT-Meeting: 399-402, 2018.

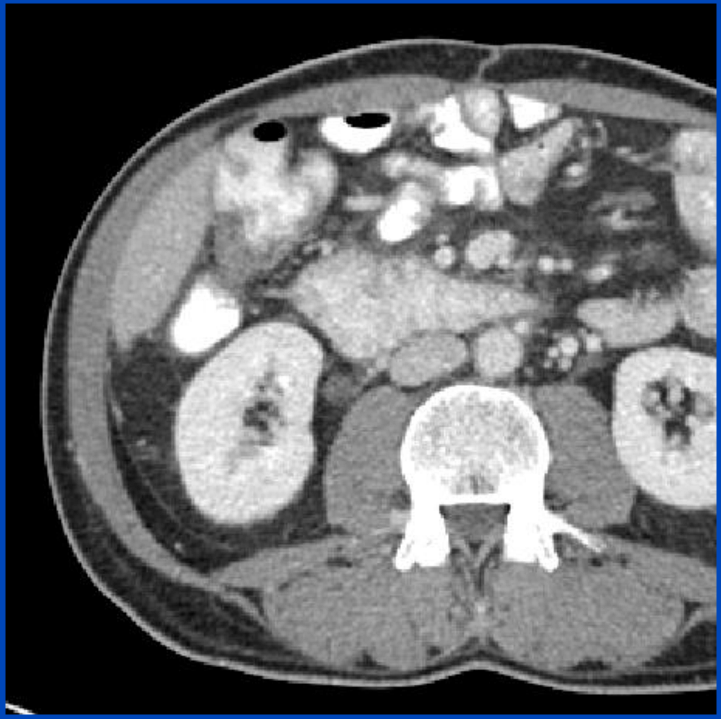
Noise Removal Example

Denoised low dose

Andrew D. Missert, Shuai Leng, Lifeng Yu, and Cynthia H. McCollough. Noise Subtraction for Low-Dose CT Images Using a Deep Convolutional Neural Network. Proceedings of the 5th CT-Meeting: 399-402, 2018.

Noise Removal Example





Full dose

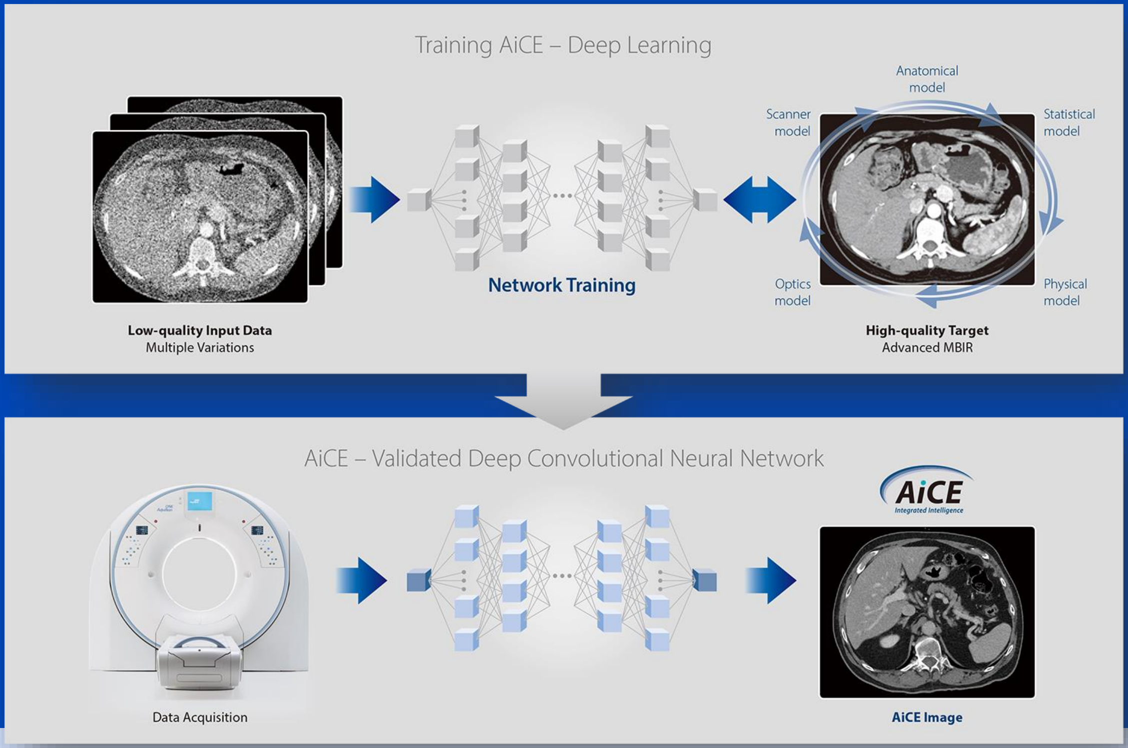
Noise Removal Example

Denoised full dose

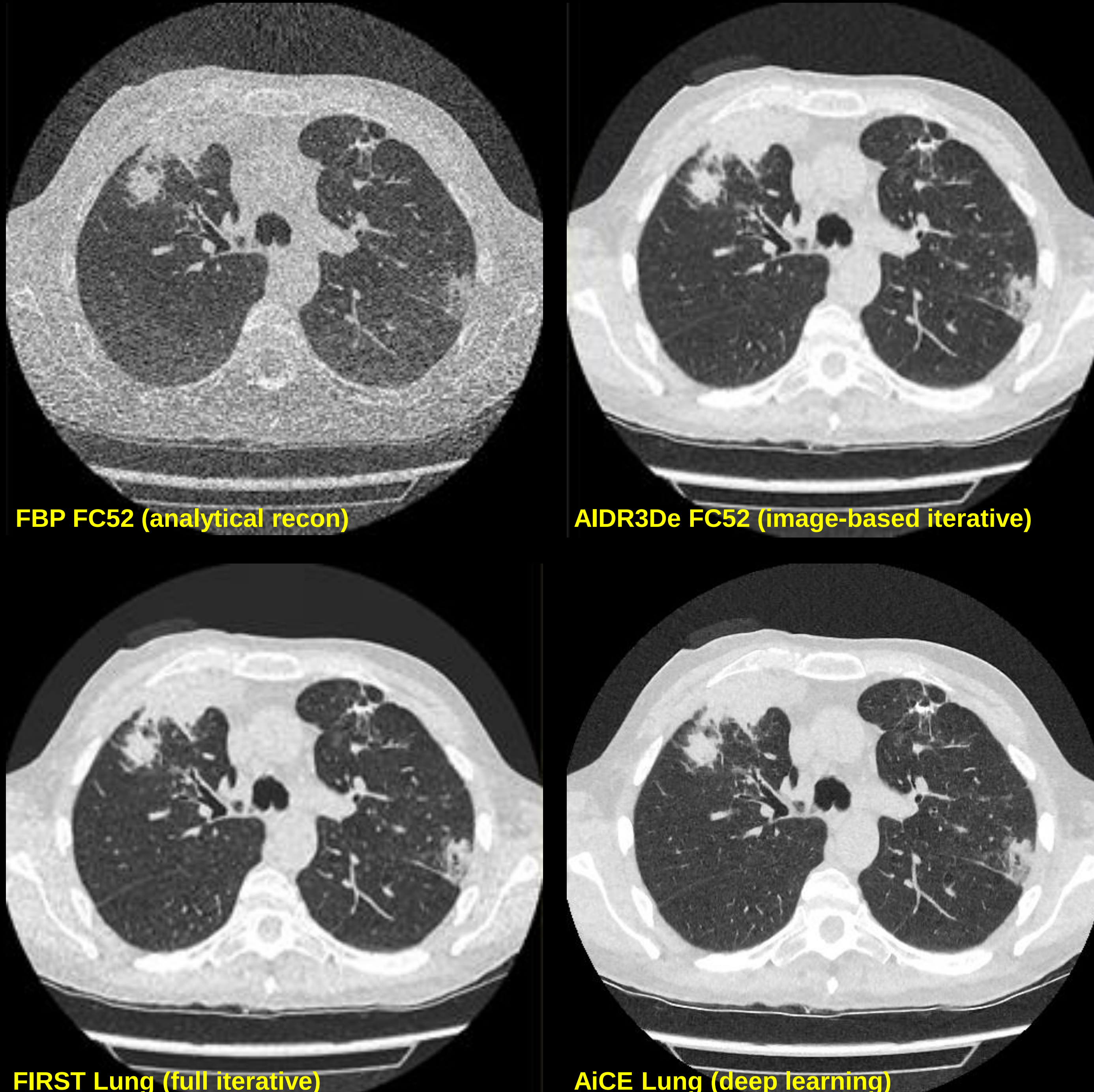
Andrew D. Missert, Shuai Leng, Lifeng Yu, and Cynthia H. McCollough. Noise Subtraction for Low-Dose CT Images Using a Deep Convolutional Neural Network. Proceedings of the 5th CT-Meeting: 399-402, 2018.

Canon's AiCE

- Advanced intelligent Clear-IQ Engine (AiCE)
- Trained to restore low-dose CT data to match the properties of FIRST, the model-based IR of Canon.
- FIRST is applied to high-dose CT images to obtain a high fidelity training target



U = 100 kV
CTDI = 0.6 mGy
DLP = 24.7 mGy·cm
D_{eff} = 0.35 mSv



Courtesy of
Radboudumc,
the Netherlands

GE's True Fidelity

- Based on a deep CNN
- Trained to restore low-dose CT data to match the properties of Veo, the model-based IR of GE.
- No information can be obtained in how the training is conducted for the product implementation.

2.5D DEEP LEARNING FOR CT IMAGE RECONSTRUCTION USING A MULTI-GPU IMPLEMENTATION

Amirkoushyar Ziabari[⋆], *Dong Hye Ye* [⋆ †], *Somesh Srivastava*[‡], *Ken D. Sauer* [⊕]
Jean-Baptiste Thibault [‡], *Charles A. Bouman*[⋆]

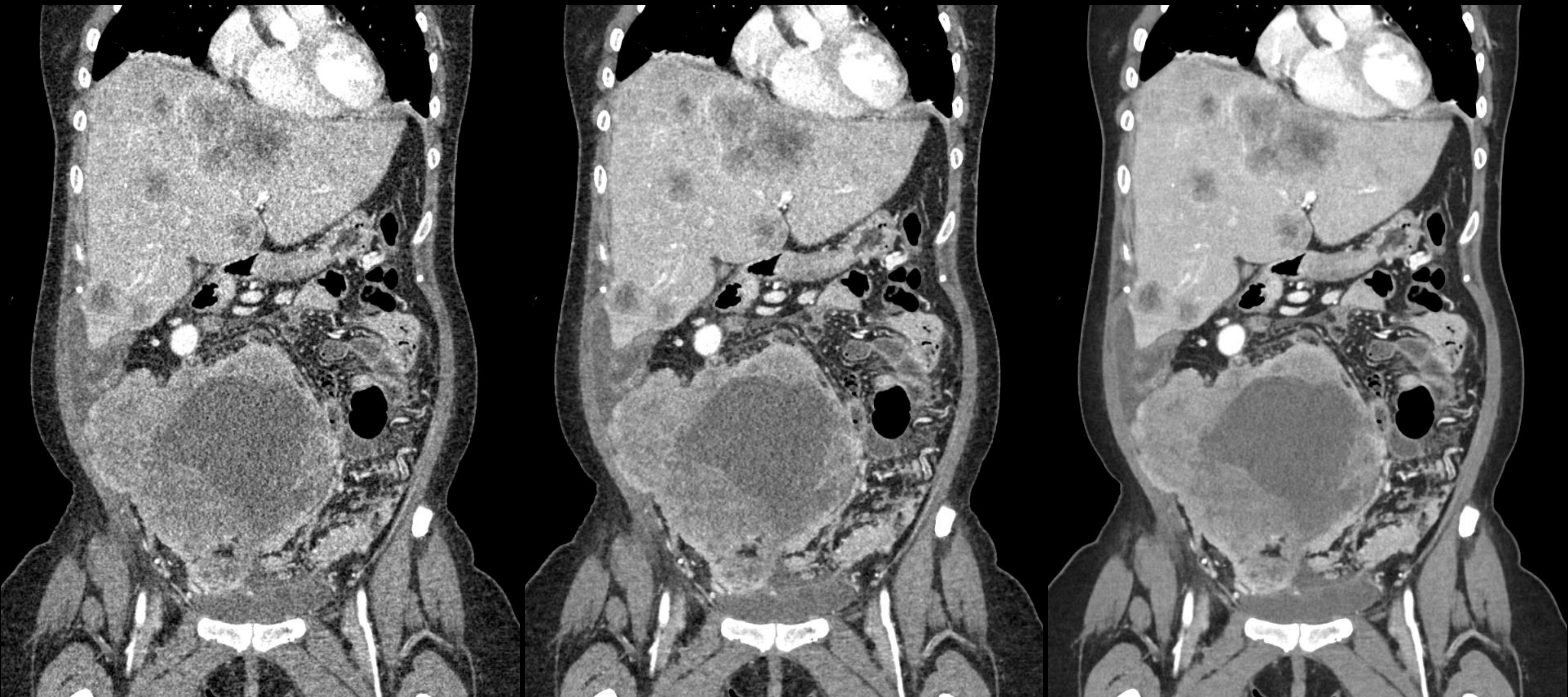
[⋆] Electrical and Computer Engineering at Purdue University
[†] Electrical and Computer Engineering at Marquett University
[‡] GE Healthcare
[⊕] Electrical Engineering at University of Notre Dame

ABSTRACT

While Model Based Iterative Reconstruction (MBIR) of CT scans has been shown to have better image quality than Filtered Back Projection (FBP), its use has been limited by its high computational cost. More recently, deep convolutional neural networks (CNN) have shown great promise in both denoising and reconstruction applications. In this research, we propose a fast reconstruction algorithm, which we call Deep

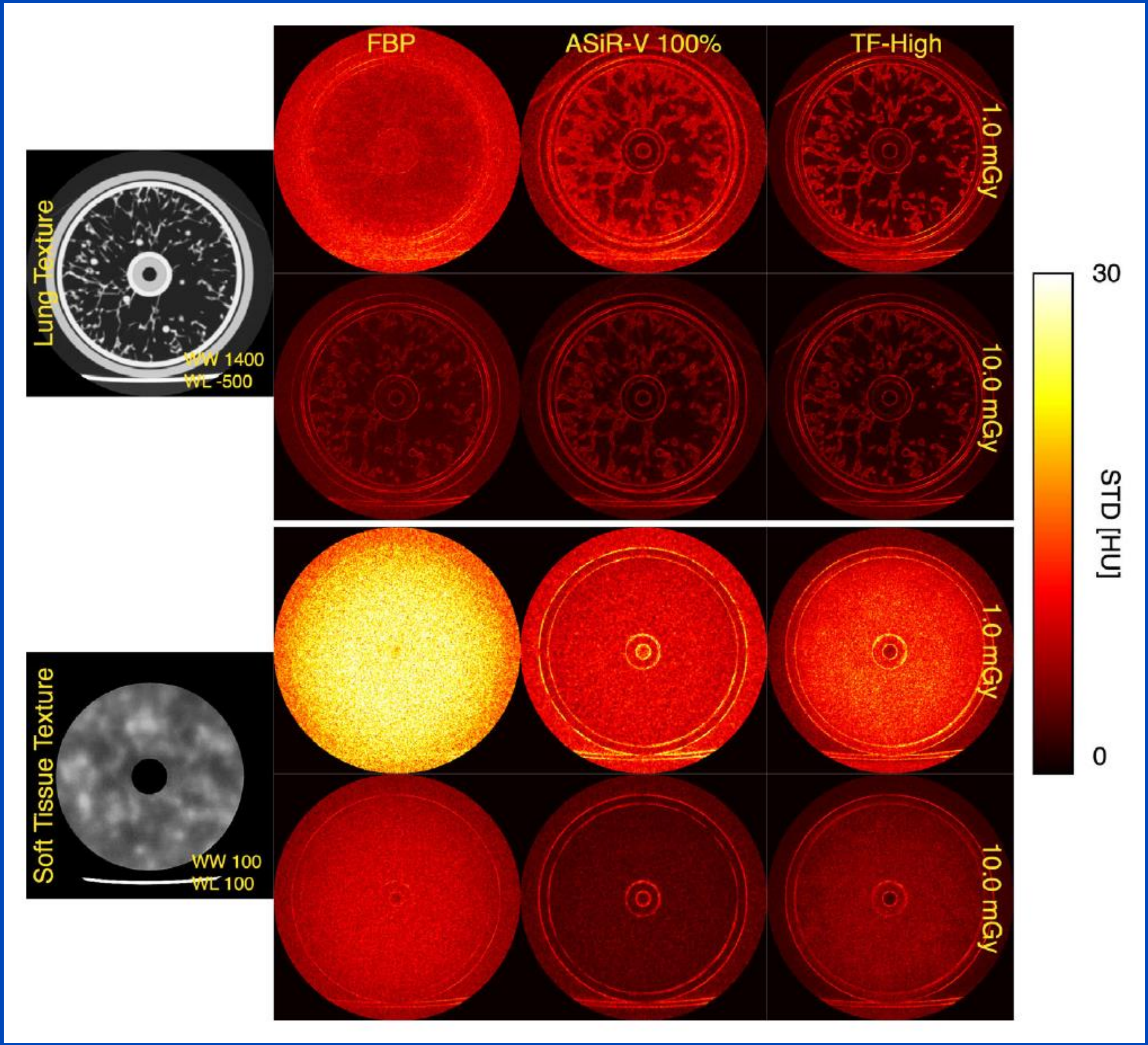
streaking artifacts caused by sparse projection views in CT images [8]. More recently, Ye, et al. [9] developed method for incorporating CNN denoisers into MBIR reconstruction as advanced prior models using the Plug-and-Play framework [10, 11].

In this paper, we propose a fast reconstruction algorithm, which we call Deep Learning MBIR (DL-MBIR), for approximately achieving the improved quality of MBIR using a deep residual neural network. The DL-MBIR method is trained to



FBP | ASIR V 50% | True Fidelity

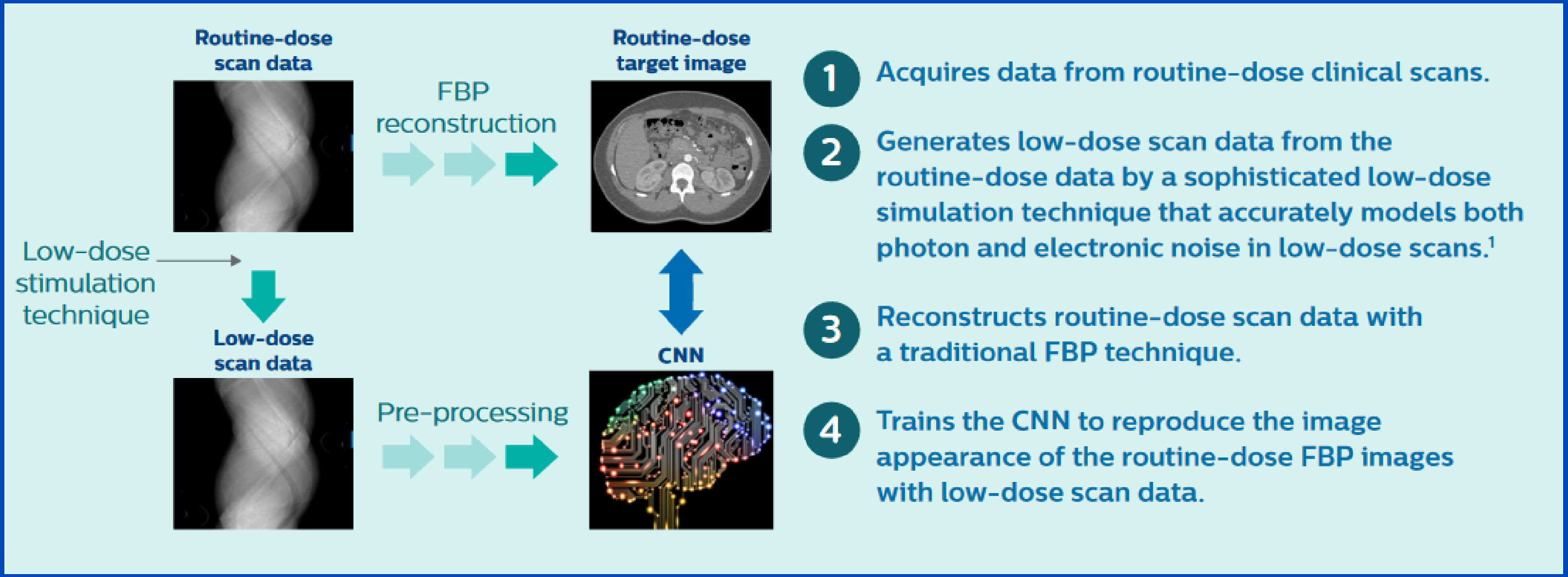
Courtesy of GE Healthcare

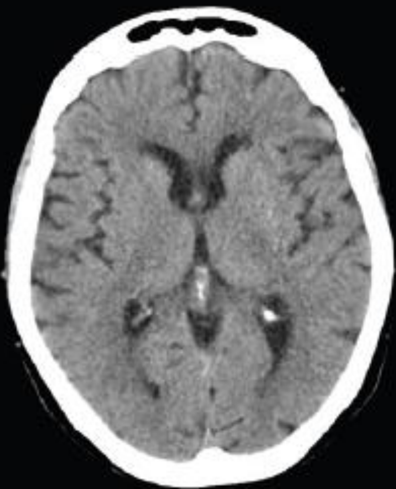


Solomon et al. Noise and spatial resolution properties of a commercially available deep learning-based CT reconstruction algorithm. Med. Phys. 47(9):3961-3971, Sept. 2020

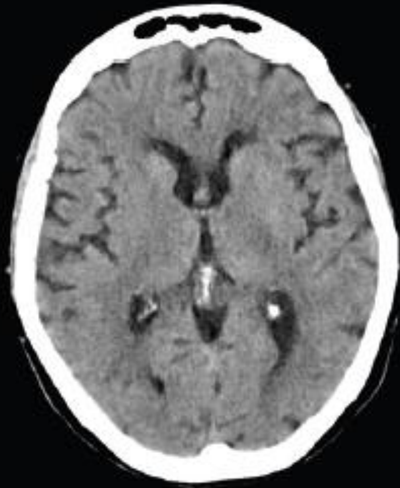
Philips' Precise Image

- Noise-injected data serve as low dose examples while their original reconstructions are the labels. A CNN learns how to denoise the low dose images.





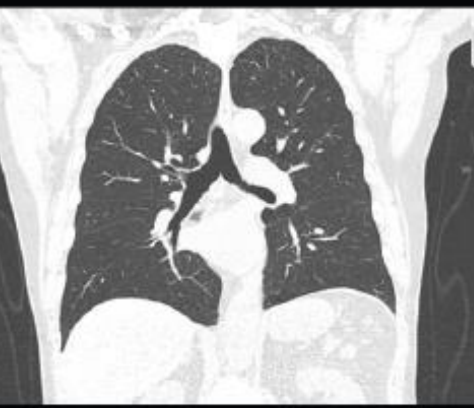
iDose⁴ 1.4 mSv



Precise Image 0.7 mSv

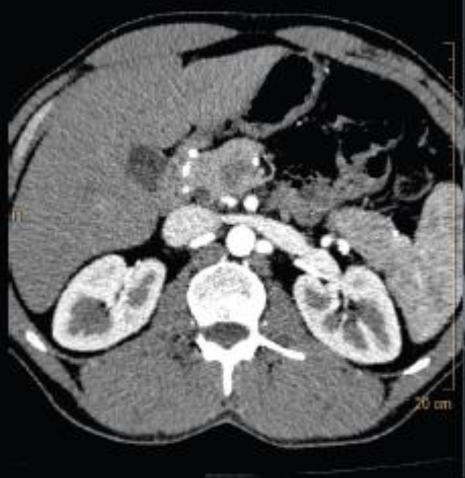
iDose⁴ 5.1 mSv

Precise Image 2.6 mSv



iDose⁴ 1.5 mSv

Precise Image 0.75 mSv



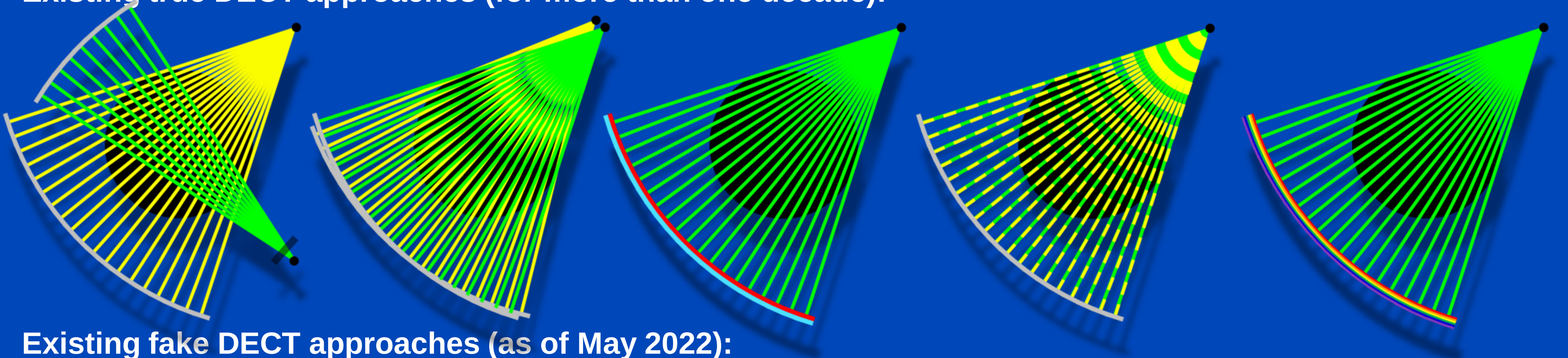
iDose⁴ 5.4 mSv

Precise Image 2.6 mSv

Taken from https://www.philips.com/c-dam/b2bhc/master/resource-catalog/landing/precise-suite/incisive_precise_image.pdf

True and Fake DECT

Existing true DECT approaches (for more than one decade):



Existing fake DECT approaches (as of May 2022):

[1] J. Ma, Y. Liao, Y. Wang, S. Li, J. He, D. Zeng, Z. Bian, "Pseudo dual energy CT imaging using deep learning-based framework: basic material estimation", *SPIE Medical Imaging 2018*.

[2] W. Zhao, T. Lv, P. Gao, L. Shen, X. Dai, K. Cheng, M. Jia, Y. Chen, L. Xing, "A deep learning approach for dual-energy CT imaging using a single-energy CT data", *Fully3D 2019.*

[3] D. Lee, H. Kim, B. Choi, H. J. Kim, "Development of a deep neural network for generating synthetic dual-energy chest x-ray images with single x-ray exposure", PMB 64(11), 2019.

[4] L. Yao, S. Li, D. Li, M. Zhu, Q. Gao, S. Zhang, Z. Bian, J. Huang, D. Zeng, J. Ma, "Leveraging deep generative model for direct energy-resolving CT imaging via existing energy-integrating CT images", *SPIE Medical Imaging 2020*.

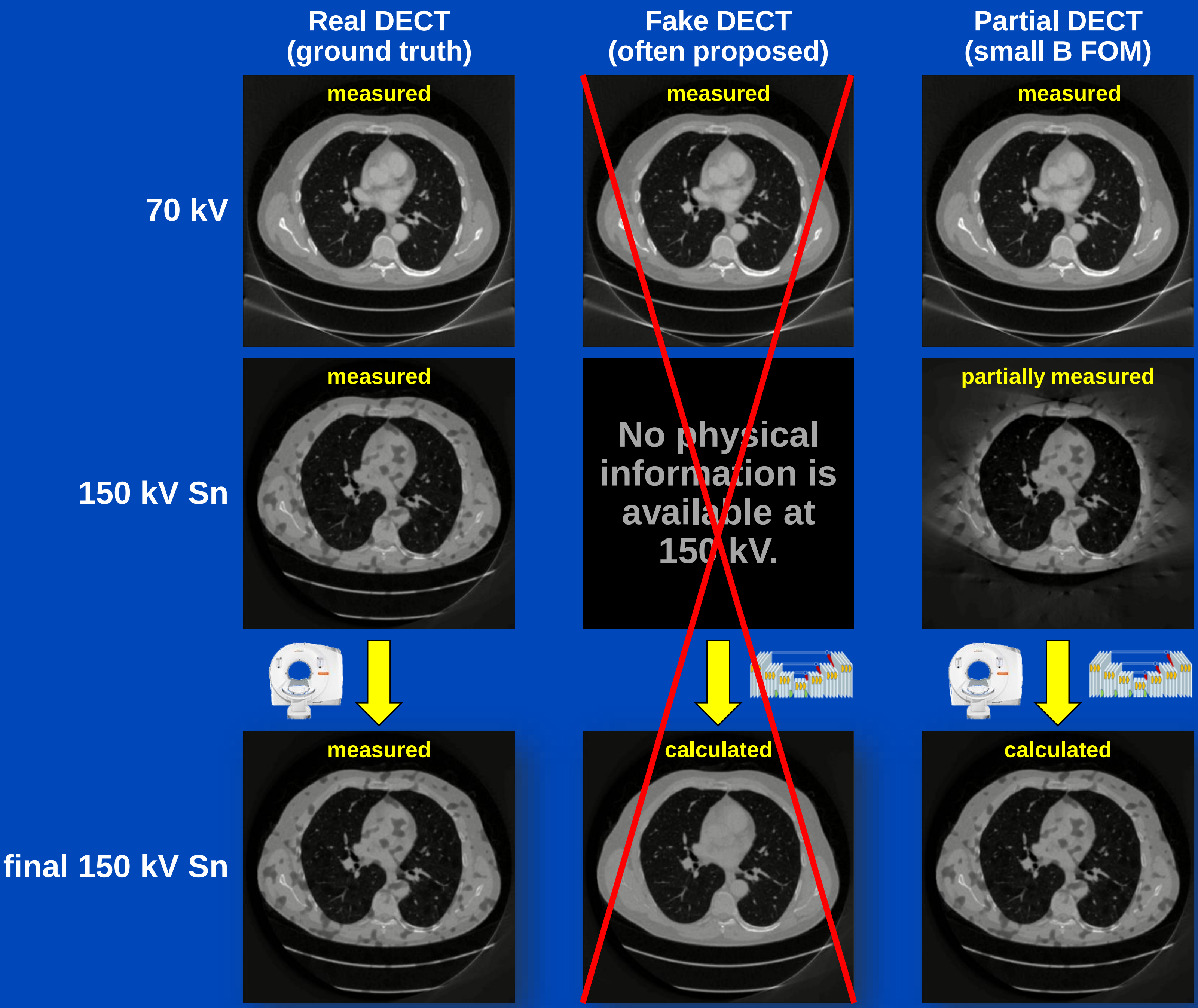
[5] D. P. Clark, F. R. Schwartz, D. Marin, J. C. Ramirez-Giraldo, C. T. Badea, "Deep learning based spectral extrapolation for dual-source, dual-energy x-ray CT", Med. Phys. 47 (9): 4150–4163, 2020.

[6] C. K. Liu, C. C. Liu, C. H. Yang, H. M. Huang, "Generation of brain dual-energy CT from single-energy CT using deep learning", Journal of Digital Imaging 34(1):149–161, 2021.

[7] T. Lyu, W. Zhao, Y. Zhu, Z. Wu, Y. Zhang, Y. Chen, L. Luo, S. Li, L. Xing, "Estimating dual-energy CT imaging from single-energy CT data with material decomposition convolutional neural network", Medical Image Analysis 70:1–10, 2021.

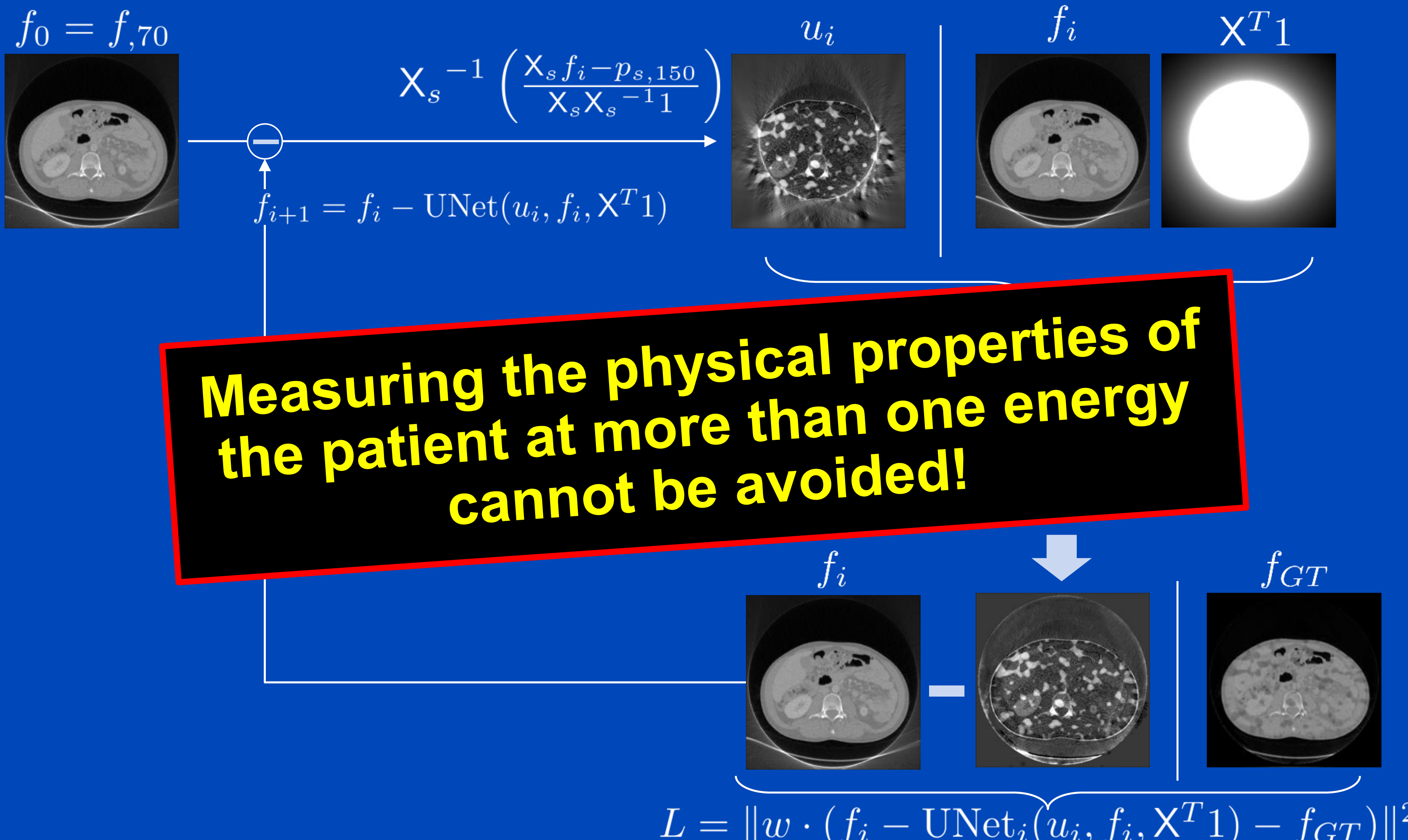
[8] F. R. Schwartz, D. P. Clark, Y. Ding, J. C. Ramirez-Giraldo, C. T. Badea, D. Marin, "Evaluating renal lesions using deep-learning based extension of dual-energy FoV in dual-source CT—A retrospective pilot study", European Journal of Radiology 139:109734, 2021.

[9] Y. Li, X. Tie, K. Li, J. W. Garrett, G.-H. Chen, "Deep-En-Chroma: mining the spectral fingerprints in single-kV CT acquisitions using energy integration detectors", *SPIE Medical Imaging 2022*.



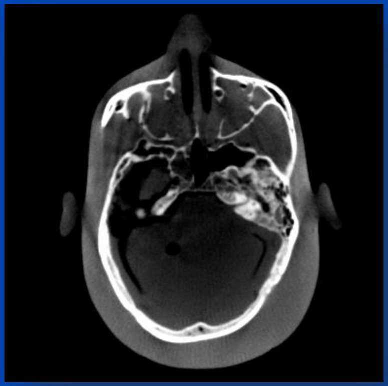
Real DECT
(ground truth)
Fake DECT
(often proposed)
Partial DECT
(small B FOM)
70 kV
measured
measured
measured
150 kV Sn
measured
No physical information is available at 150 kV.
partially measured
measured
calculated
calculated
final 150 kV Sn

Algorithm for Partial DECT



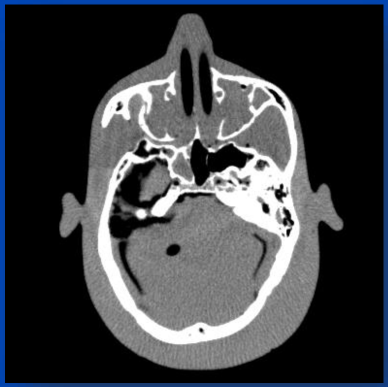
Measuring the physical properties of the patient at more than one energy cannot be avoided!

Deep Scatter Estimation



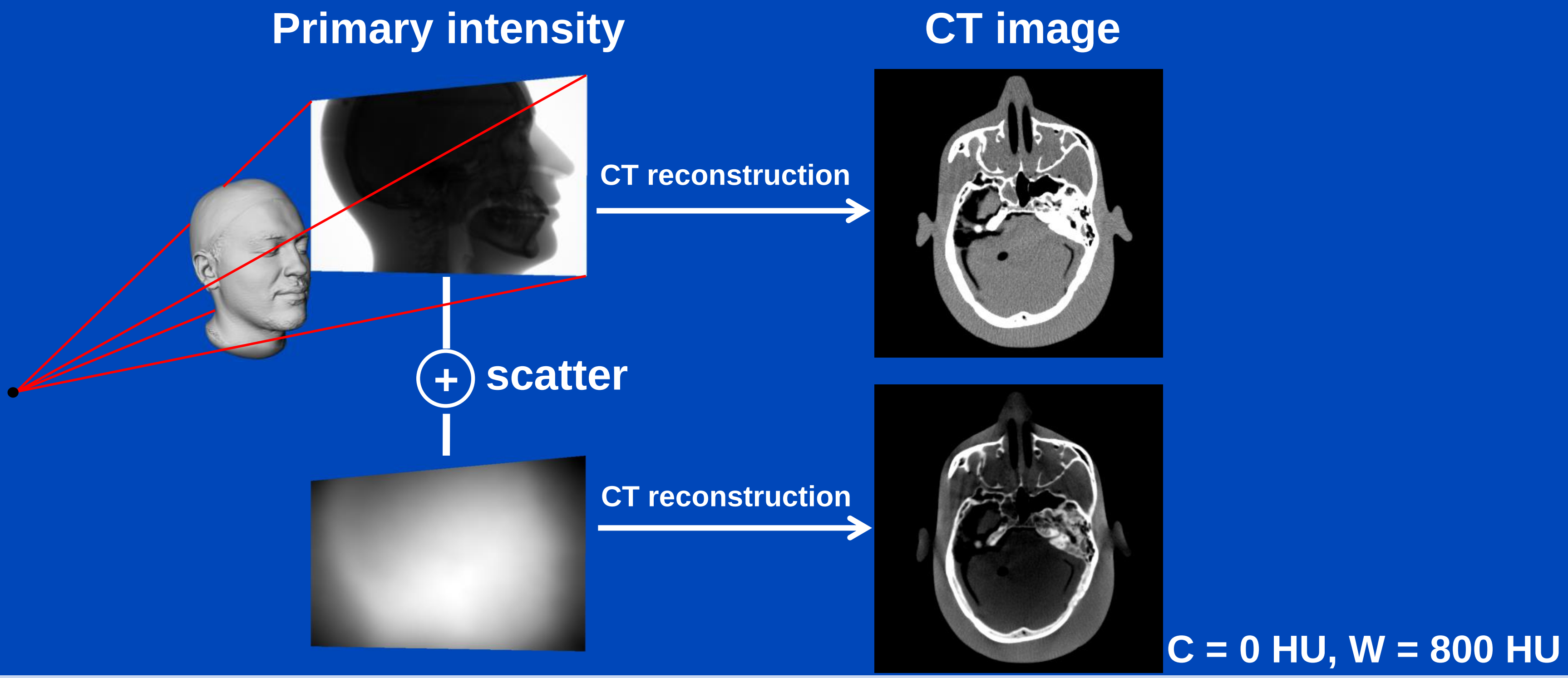
???

In real time?



Motivation

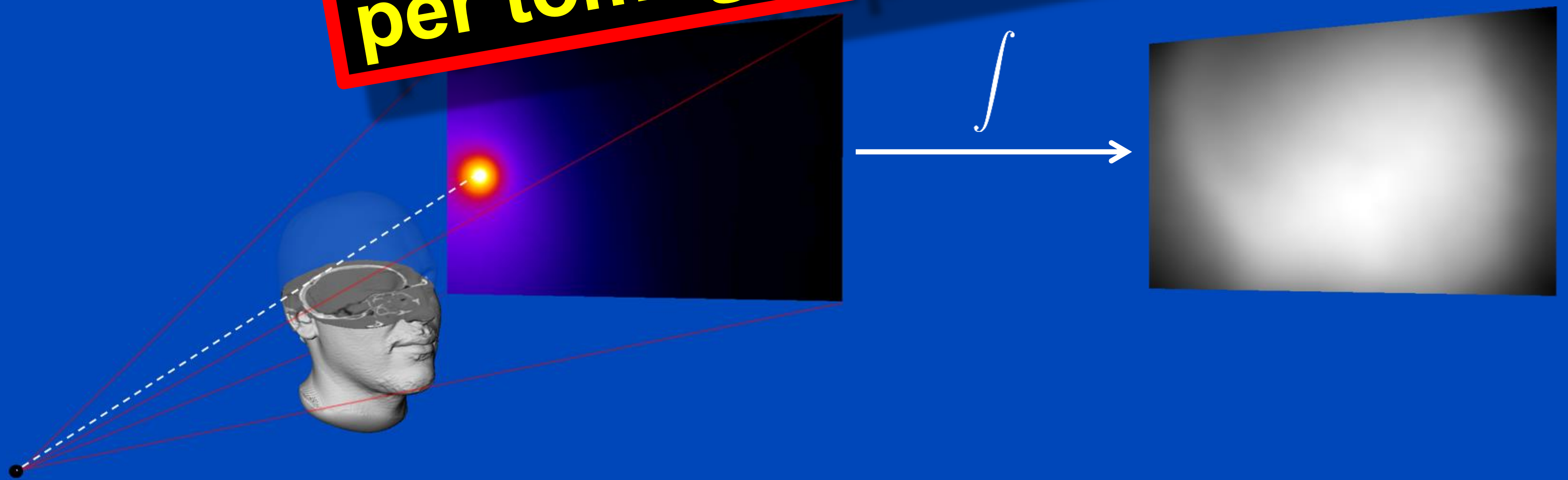
- X-ray scatter is a major cause of image quality degradation in CT and CBCT.
- Appropriate scatter correction is crucial to maintain the diagnostic value of the CT examination.



Monte Carlo Scatter Estimation

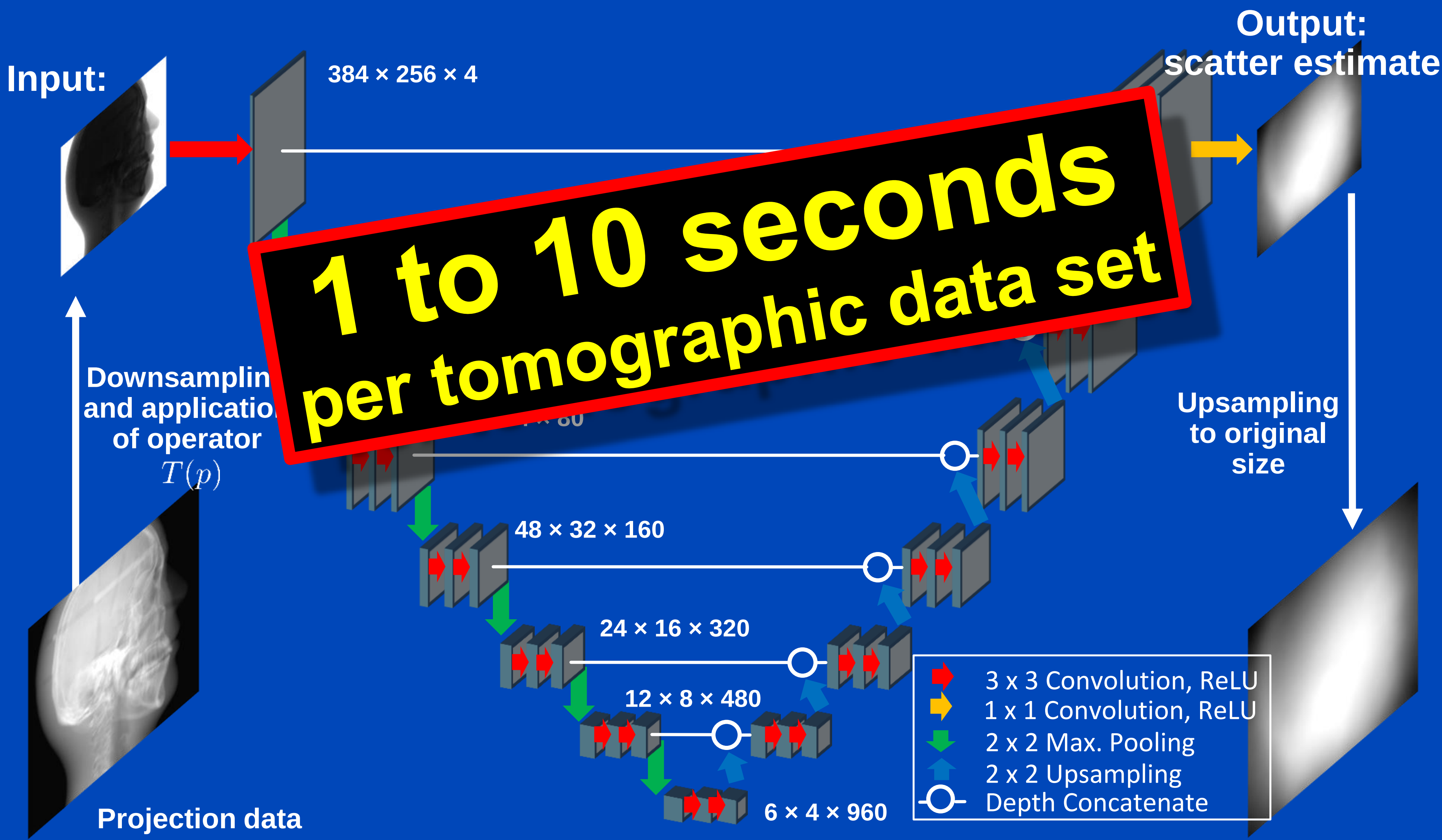
- Simulation of photon trajectories according to physical interaction probabilities.
- Simulating a large numb… …ories well approxima…

1 to 10 hours per tomographic data set



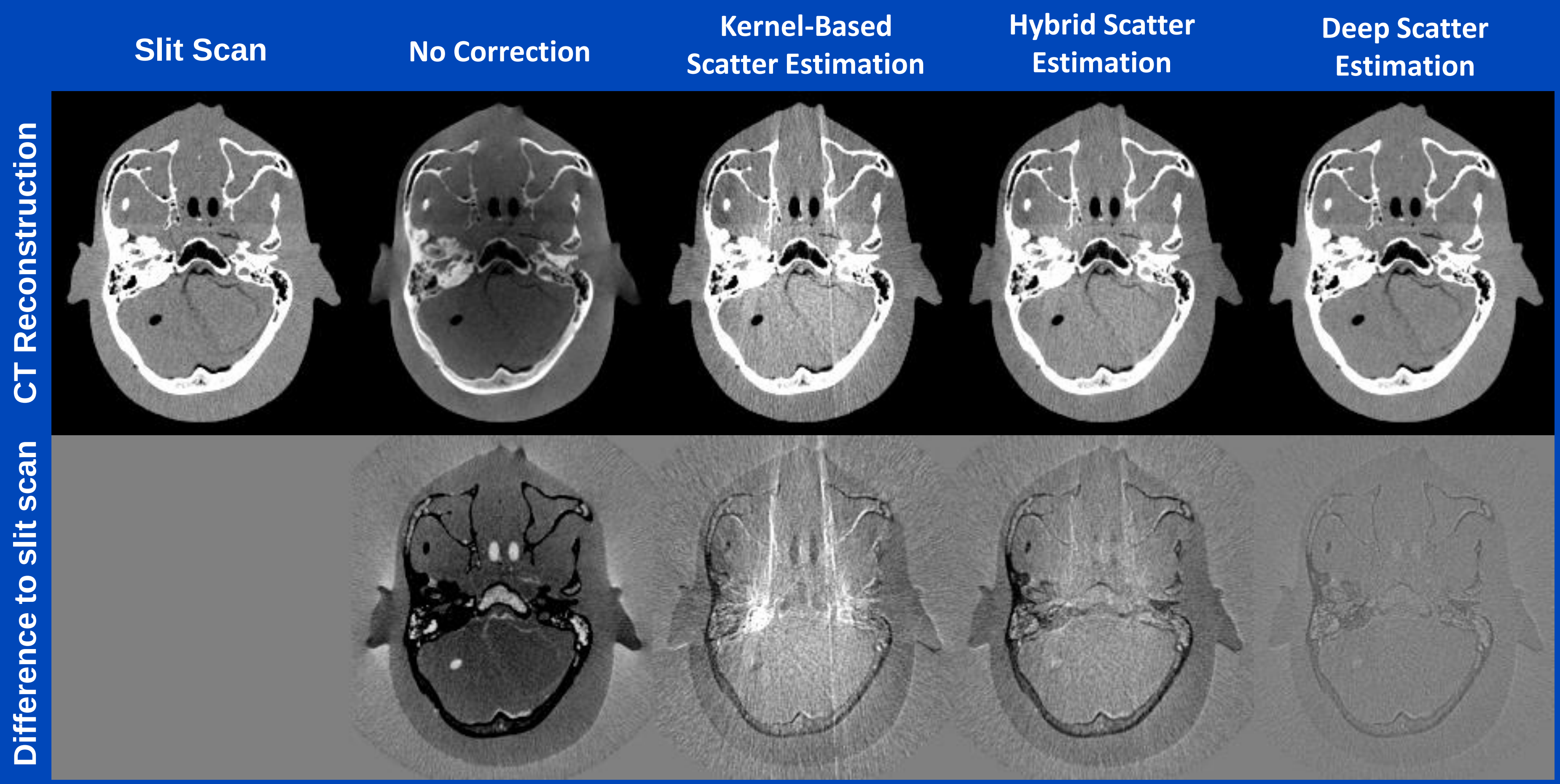
Deep Scatter Estimation

Network architecture & scatter estimation framework



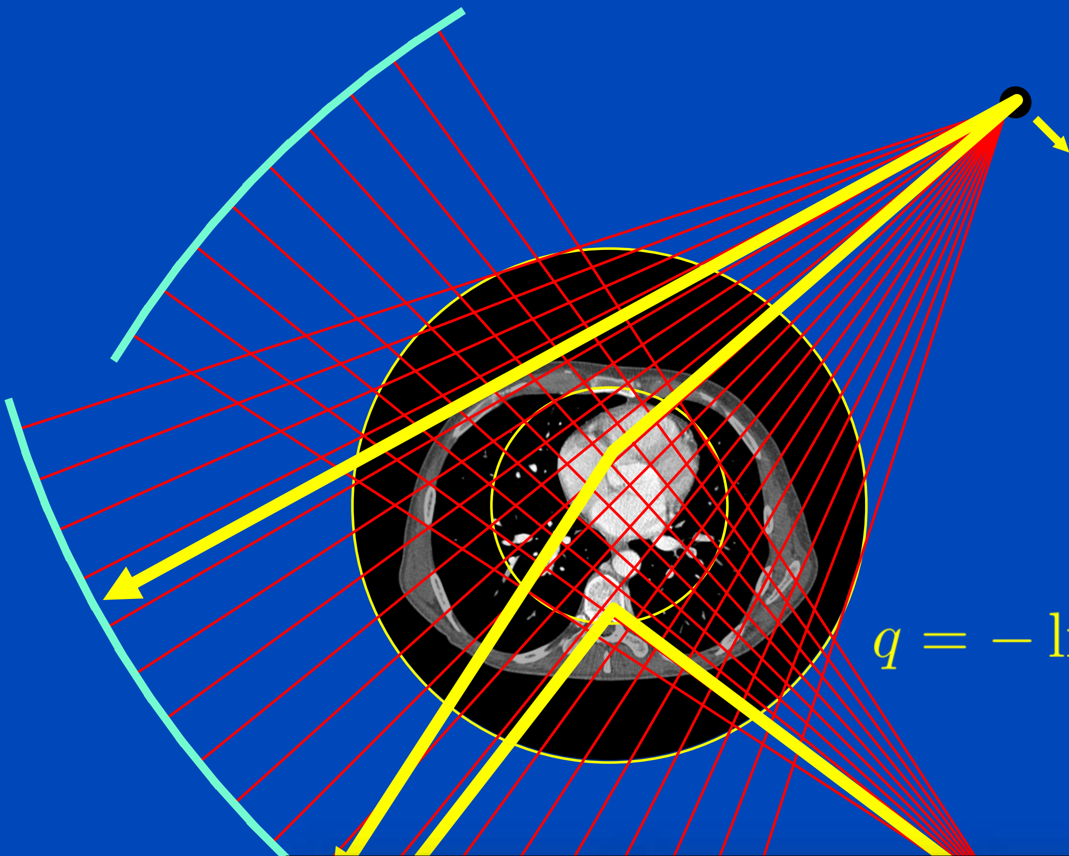
J. Maier, M. Kachelrieß et al. Deep scatter estimation (DSE). SPIE 2017 and Journal of Nondestructive Evaluation 37:57, July 2018.
J. Maier, M. Kachelrieß et al. Robustness of DSE. Med. Phys. 46(1):238-249, January 2019.

Reconstructions of Measured Data



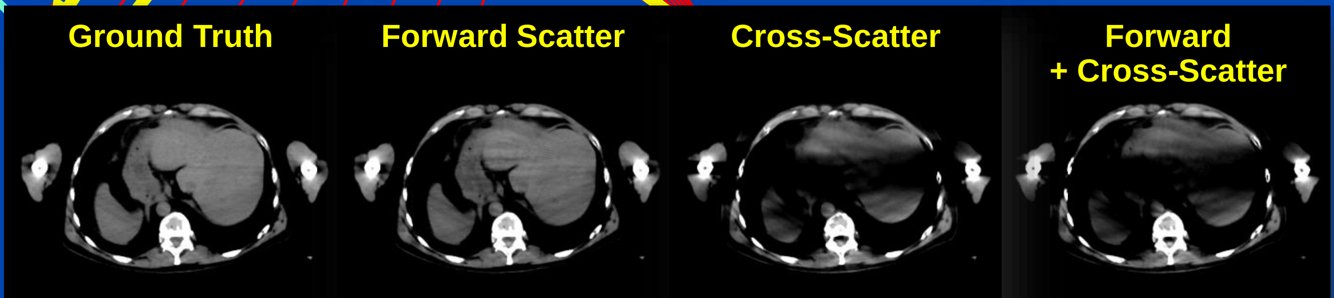
C = 0 HU, W = 1000 HU

Scatter in Dual Source CT (DSCT)



Siemens SOMATOM Force
dual source cone-beam spiral CT

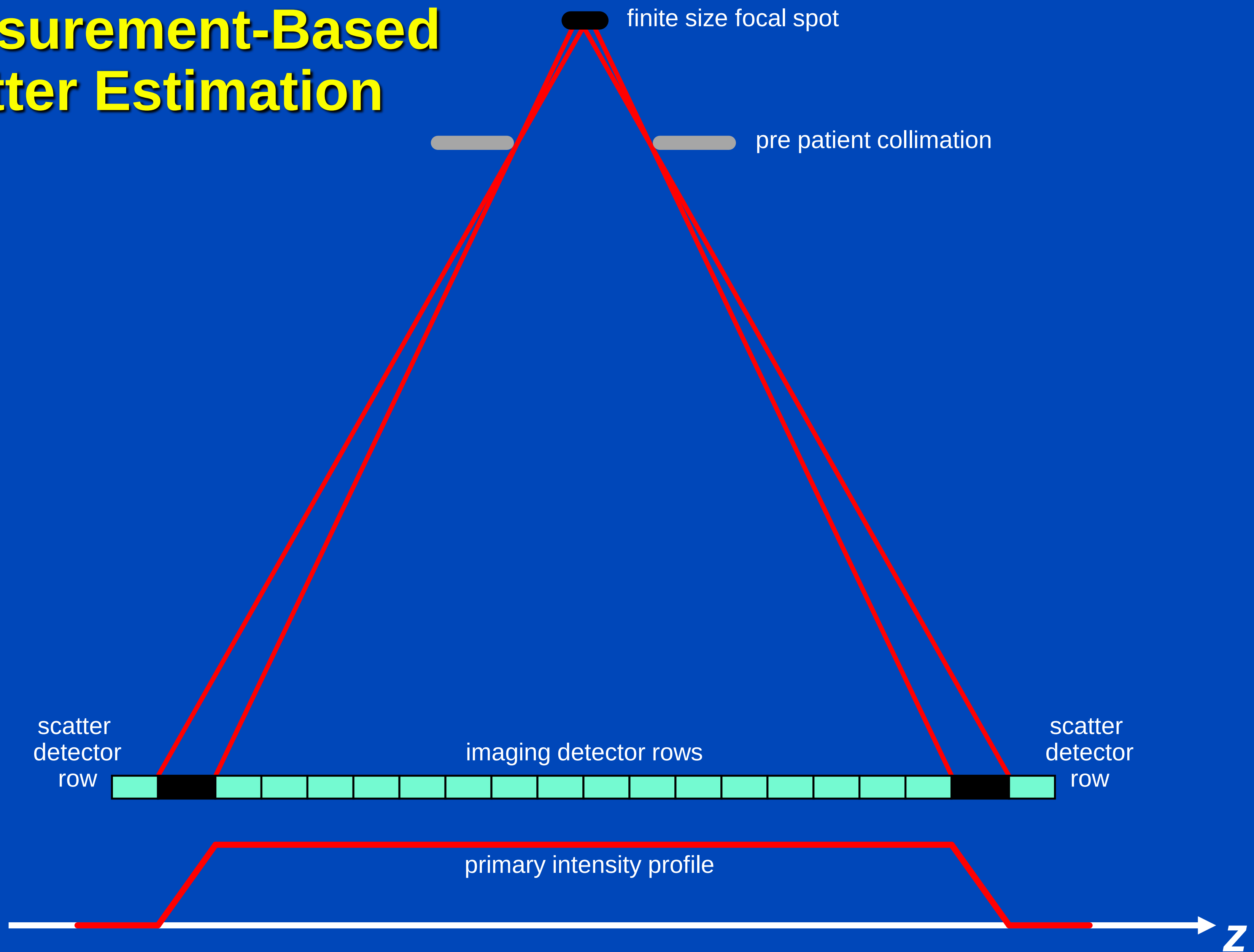
$$q = -\ln \frac{I_{\text{primary}} + S_{\text{forward}} + \rho\, S_{\text{cross}}}{I_0}$$



C = 40 HU, *W* = 300 HU, with 2D anti-scatter grid

J. Erath, T. Vöth, J. Maier, E. Fournié, M. Petersilka, K. Stierstorfer, and M. Kachelrieß. Deep learning-based forward and cross-scatter correction in dual source CT. Med. Phys. 48:4824–4842, July 2021.

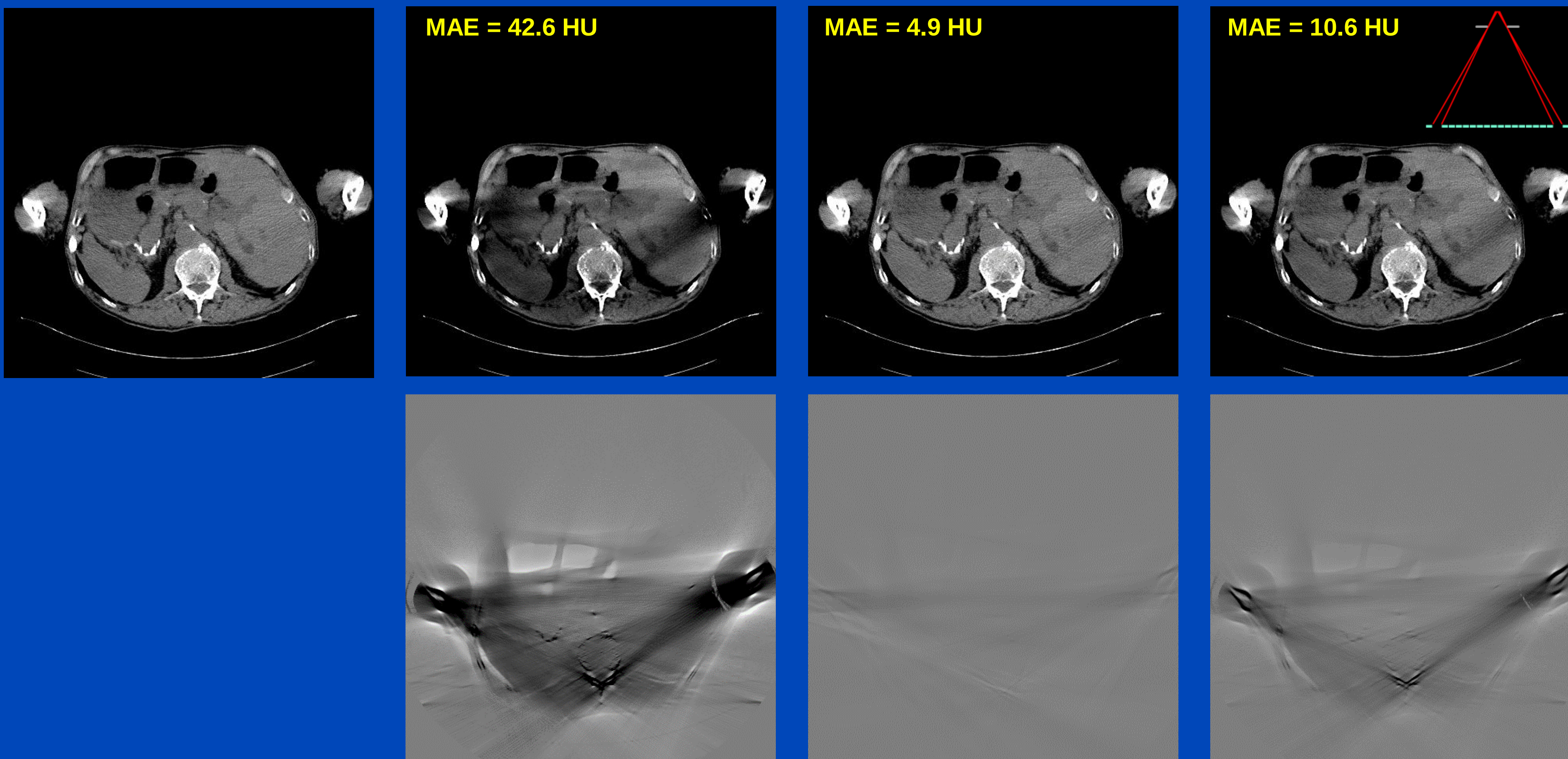
Measurement-Based Scatter Estimation



J. Erath, T. Vöth, J. Maier, E. Fournié, M. Petersilka, K. Stierstorfer, and M. Kachelrieß. Deep learning-based forward and cross-scatter correction in dual source CT. Med. Phys. 48:4824–4842, July 2021.

Cross-DSE

Ground Truth | **Uncorrected** | **xDSE (2D, xSSE)** | **Measurement-based**



xDSE (2D, xSSE) maps
primary + forward scatter + cross-scatter + cross-scatter approximation → cross-scatter

Images *C* = 40 HU, *W* = 300 HU, difference images *C* = 0 HU, *W* = 300 HU

J. Erath, T. Vöth, J. Maier, E. Fournié, M. Petersilka, K. Stierstorfer, and M. Kachelrieß. Deep learning-based forward and cross-scatter correction in dual source CT. Med. Phys. 48:4824–4842, July 2021.

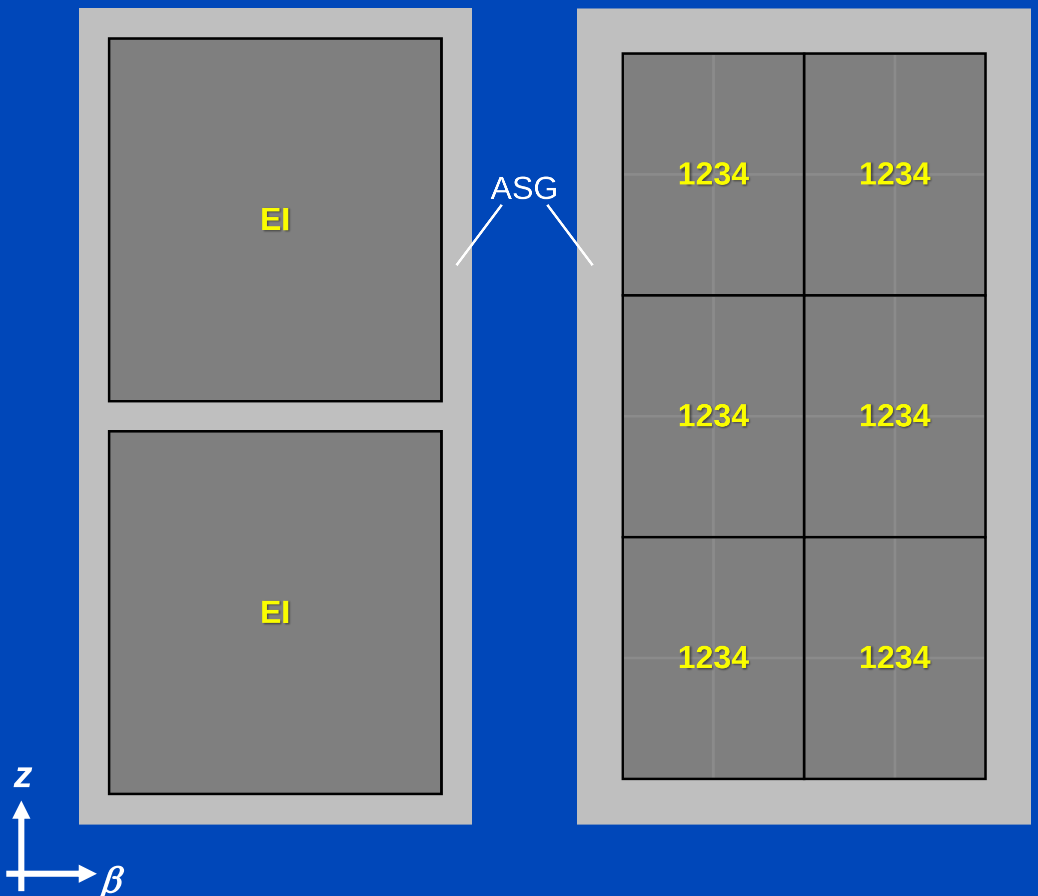
Conclusions on DSE

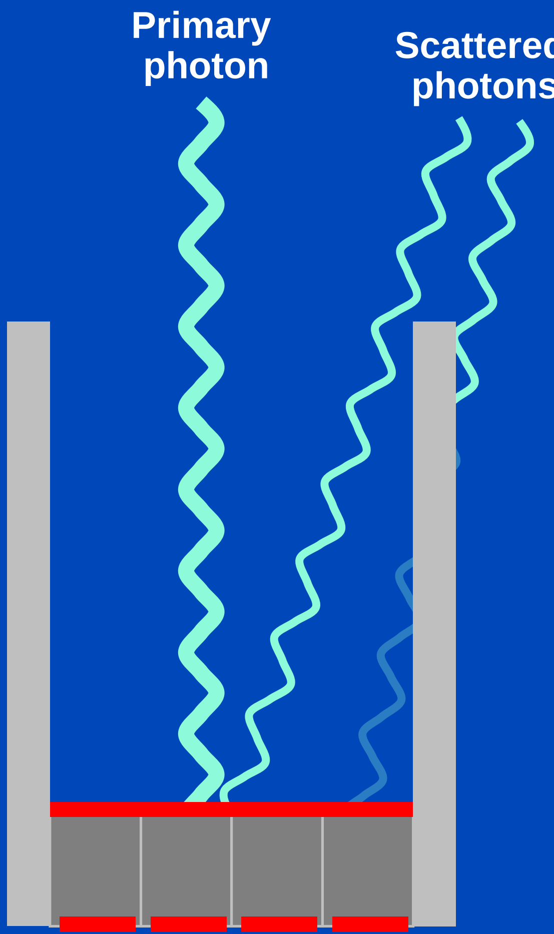
- DSE needs about 3 ms per CT and 10 ms per CBCT projection (as of 2020).
- DSE is a fast and accurate alternative to MC simulations.
- DSE outperforms kernel-based approaches in terms of accuracy and speed.
- Facts:
  - DSE can estimate scatter from a single (!) x-ray image.
  - DSE can accurately estimate scatter from a primary+scatter image.
  - DSE generalizes to all anatomical regions.
  - DSE works for geometries and beam qualities differing from training.
  - DSE may outperform MC even though DSE is trained with MC.
- DSE is not restricted to reproducing MC scatter estimates.
- DSE can rather be trained with any other scatter estimate, including those based on measurements.

Scatter of Coarse ASG

Conventional ASG
Somatom Force
920 × 96 detector pixels
pixel size 0.52 × 0.56 mm at iso

Coarse ASG
Naeotom Alpha
1376 × 144 macro pixels
pixel size 0.3 × 0.352 mm at iso

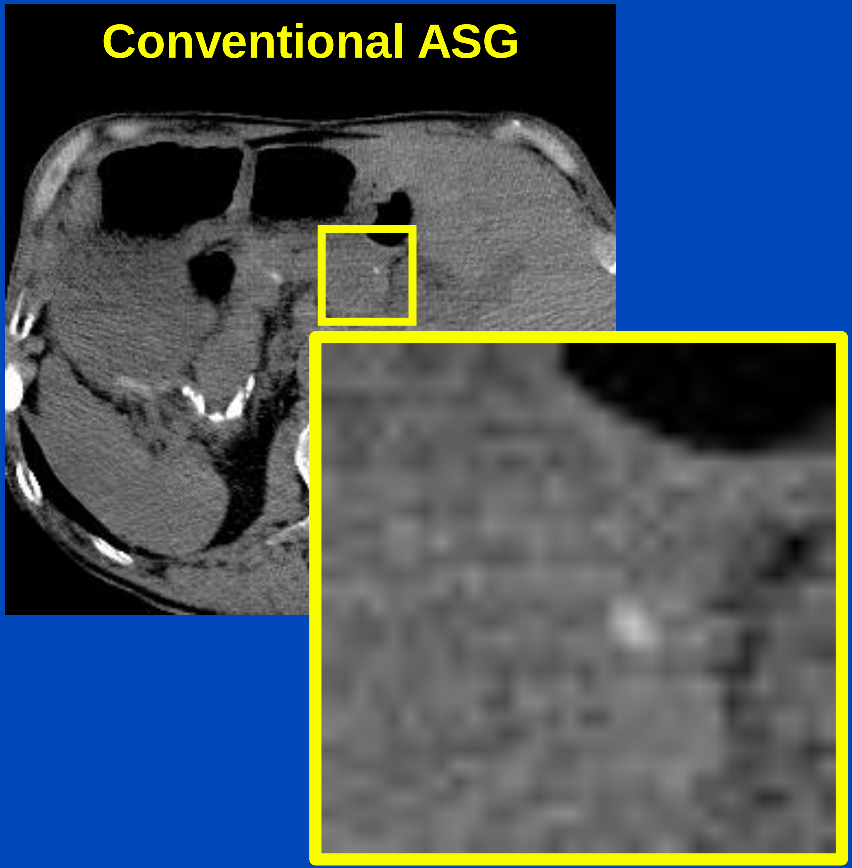


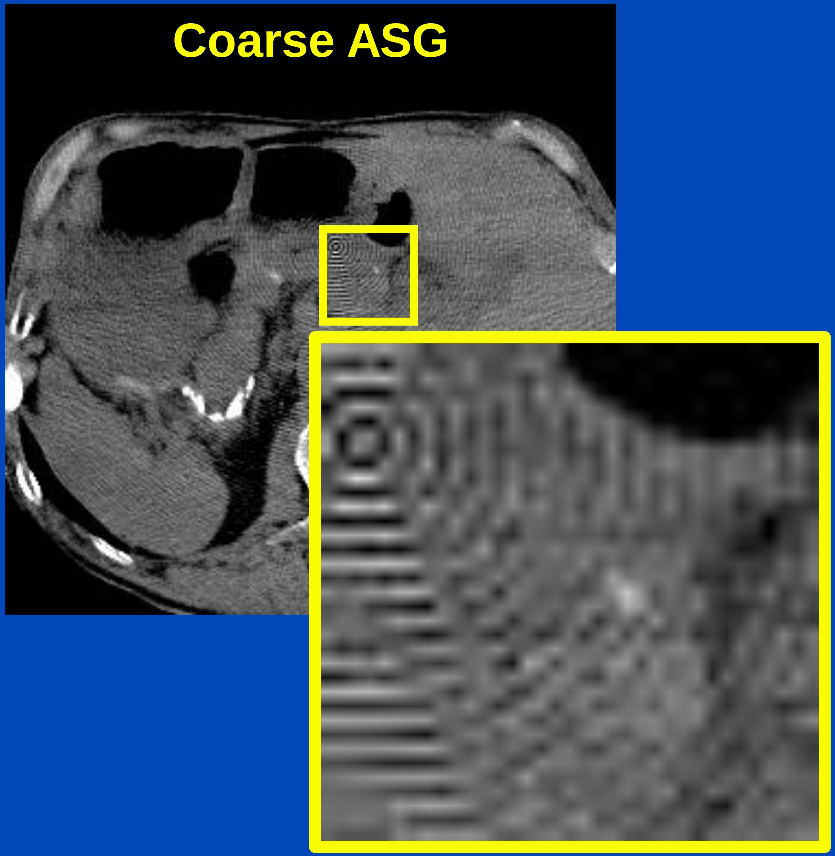


Coarse ASGs lead to changing scatter intensity between neighboring pixels.

This paper received the "Highest Impact Paper Award" for the highest impact score at the 7th International Conference on Image Formation in X-Ray Com-puted Tomography in June 2022

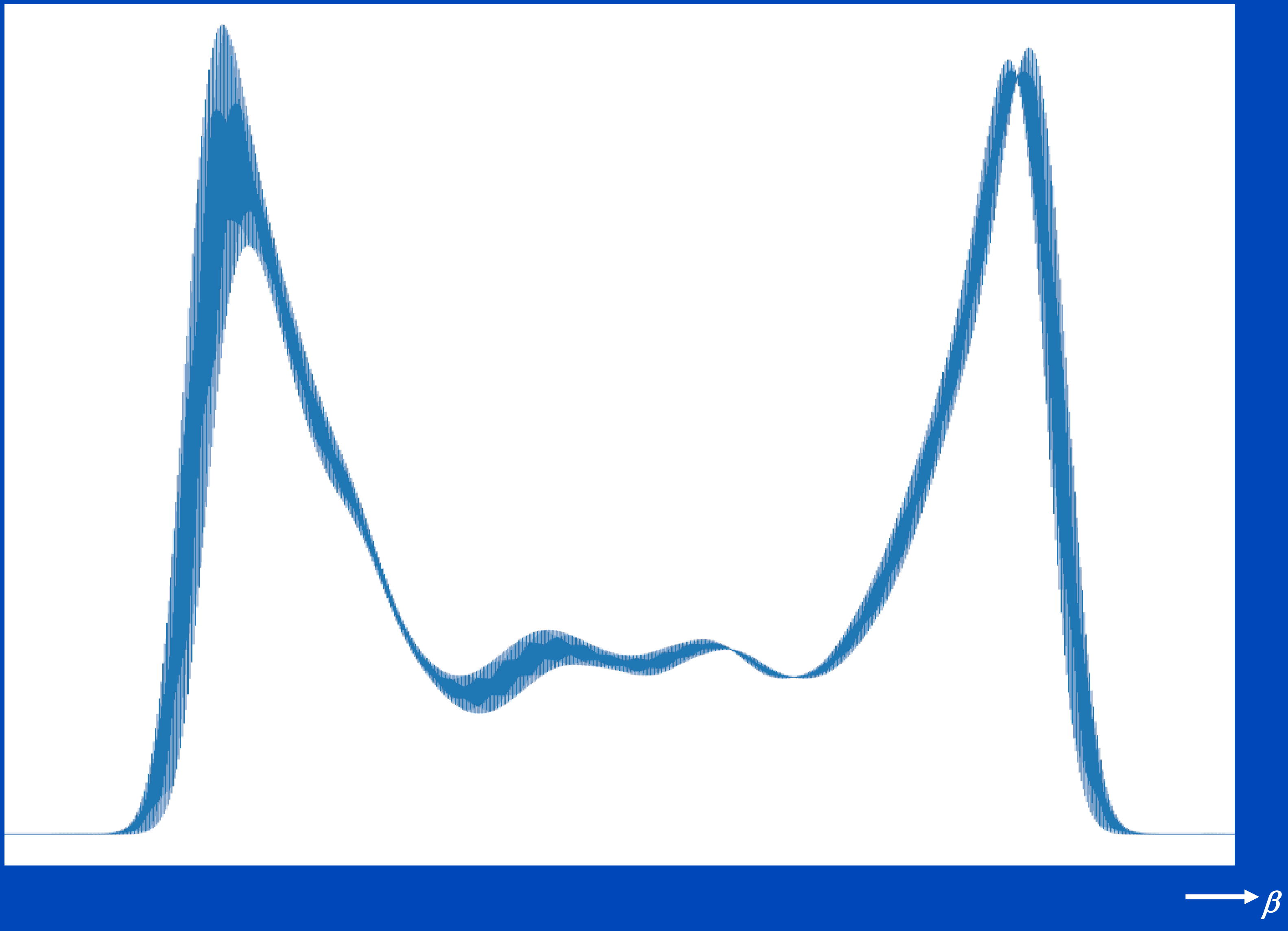
Scatter Artifacts of Coarse ASG



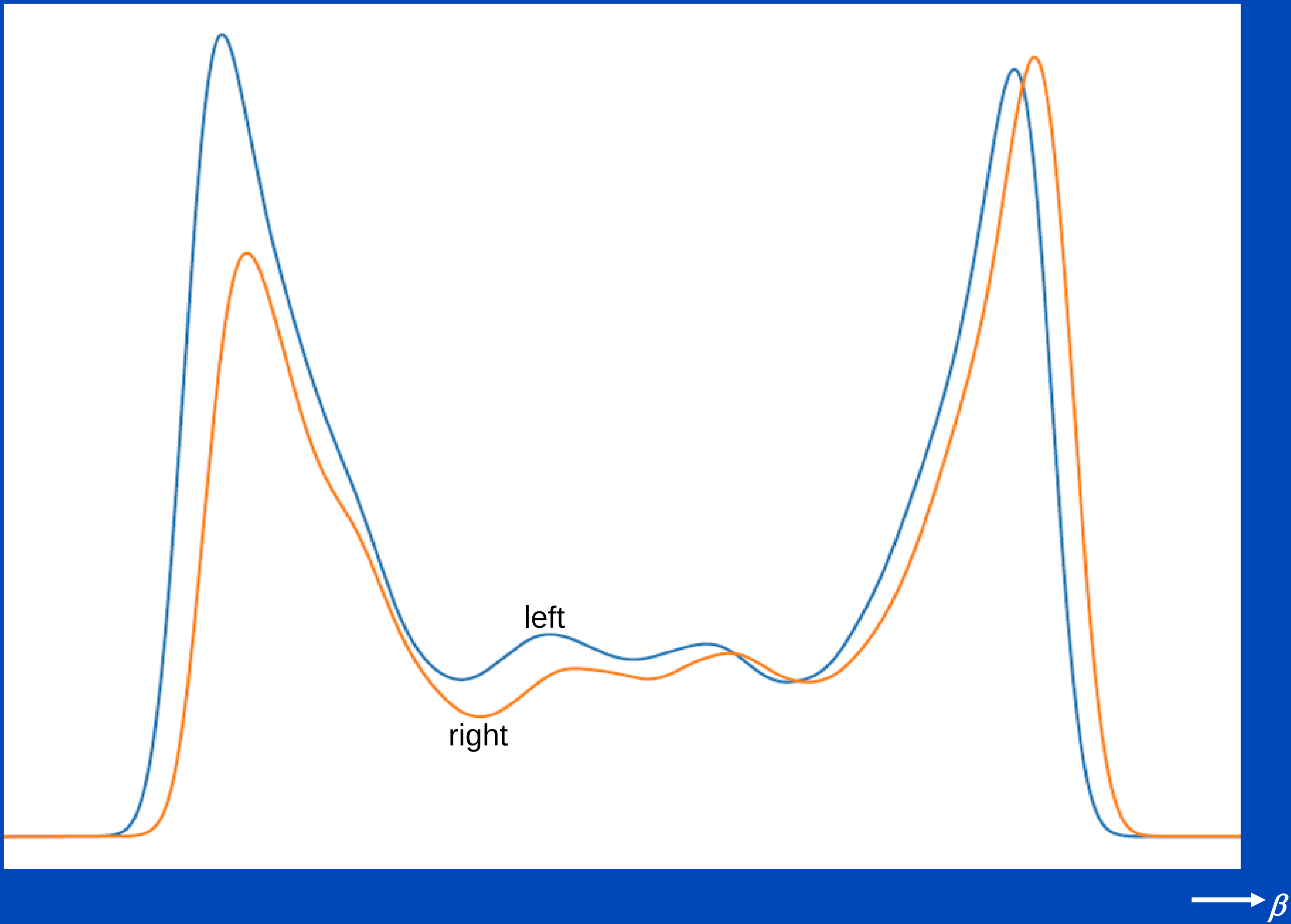


Coarse ASG can lead to scatter-induced moiré artifacts.

Reconstruction: C = 40 HU, W = 300 HU



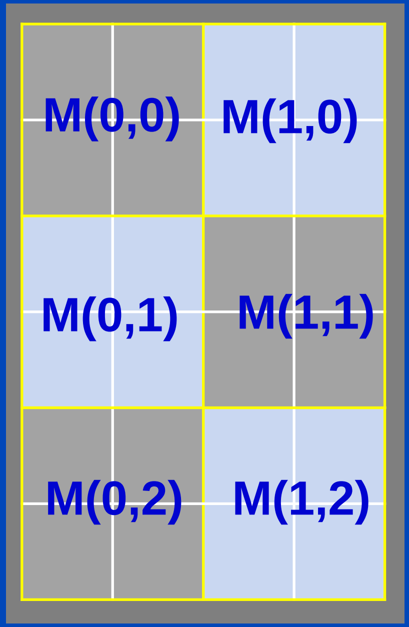
Scatter distribution averaged over all detector rows



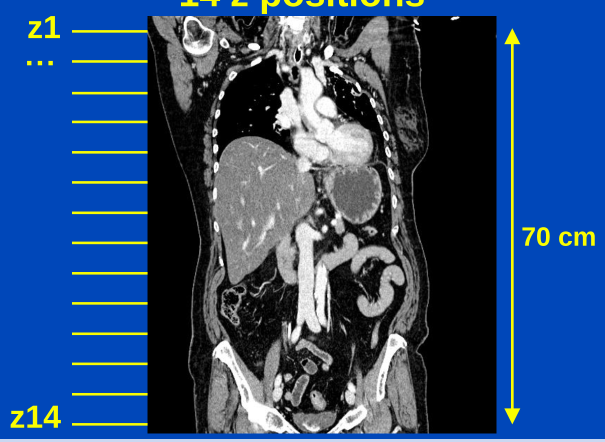
Scatter distribution averaged over all detector rows

Training and Validation Data

- **Monte Carlo simulation with the geometry of the photon counting CT scanner NAEOTOM Alpha (Siemens Healthineers)**
- **12 patients for training and 4 for validation**
- **14 z-positions with 36 projections each simulated for each patient**
- **8064 paired scatter and primary data pairs**
- **Simulation of coarse ASG with macro pixel with detector dimension of 1376 × 144 pixels**
- **6 different macro pixels locations**
- **Smooth only across same macro-pixel locations**



14 z positions



Training and validation patients with high variety and different clinical situations, important to consider scatter-to-primary ratio

Example of validation data set:

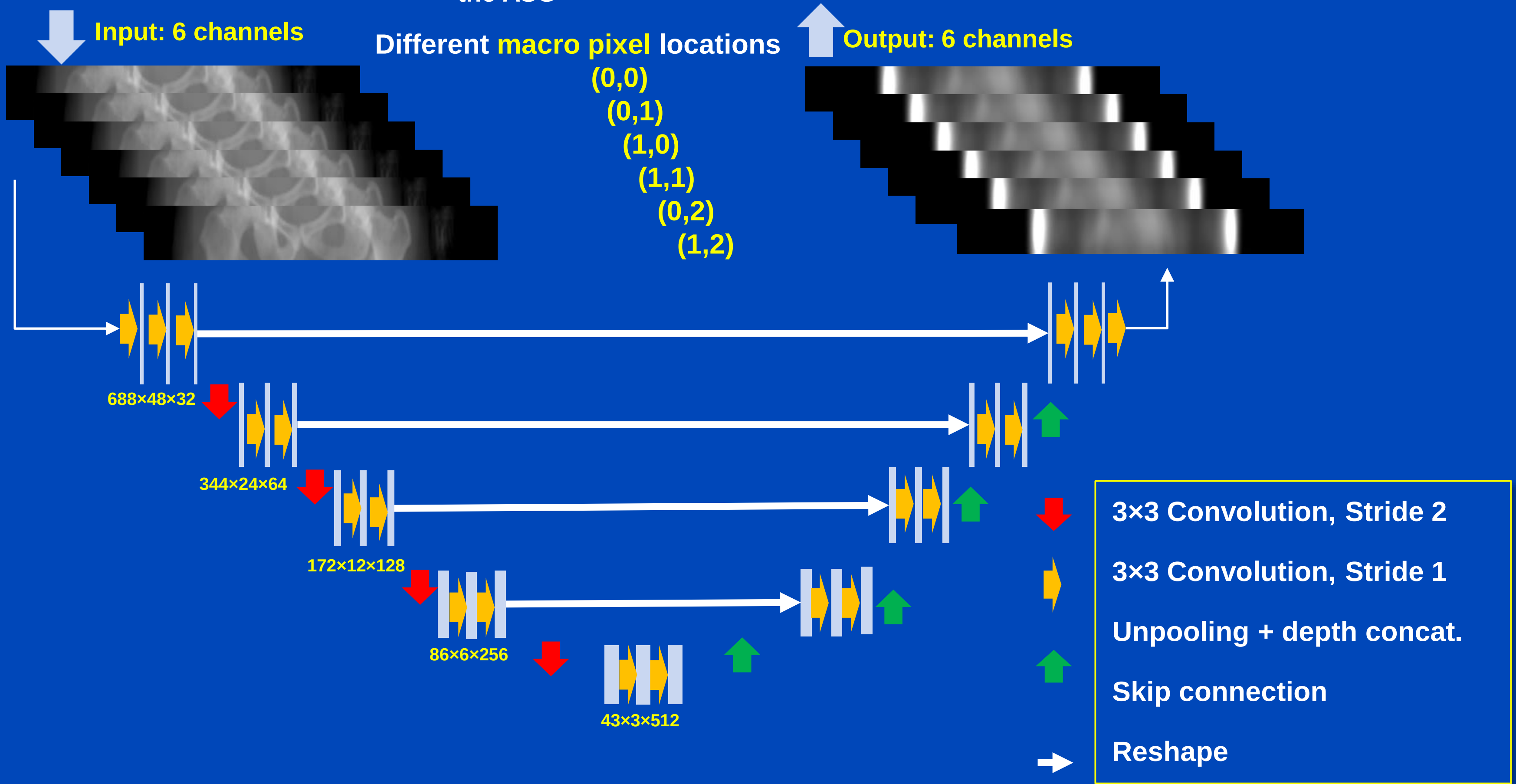
DSE for coarse ASG

Detector dimension
1376×144
Input mapping

$$p = -\ln\left(\frac{I_{\text{primary}}}{I_0} + \frac{I_{\text{scatter}}}{I_0}\right)$$

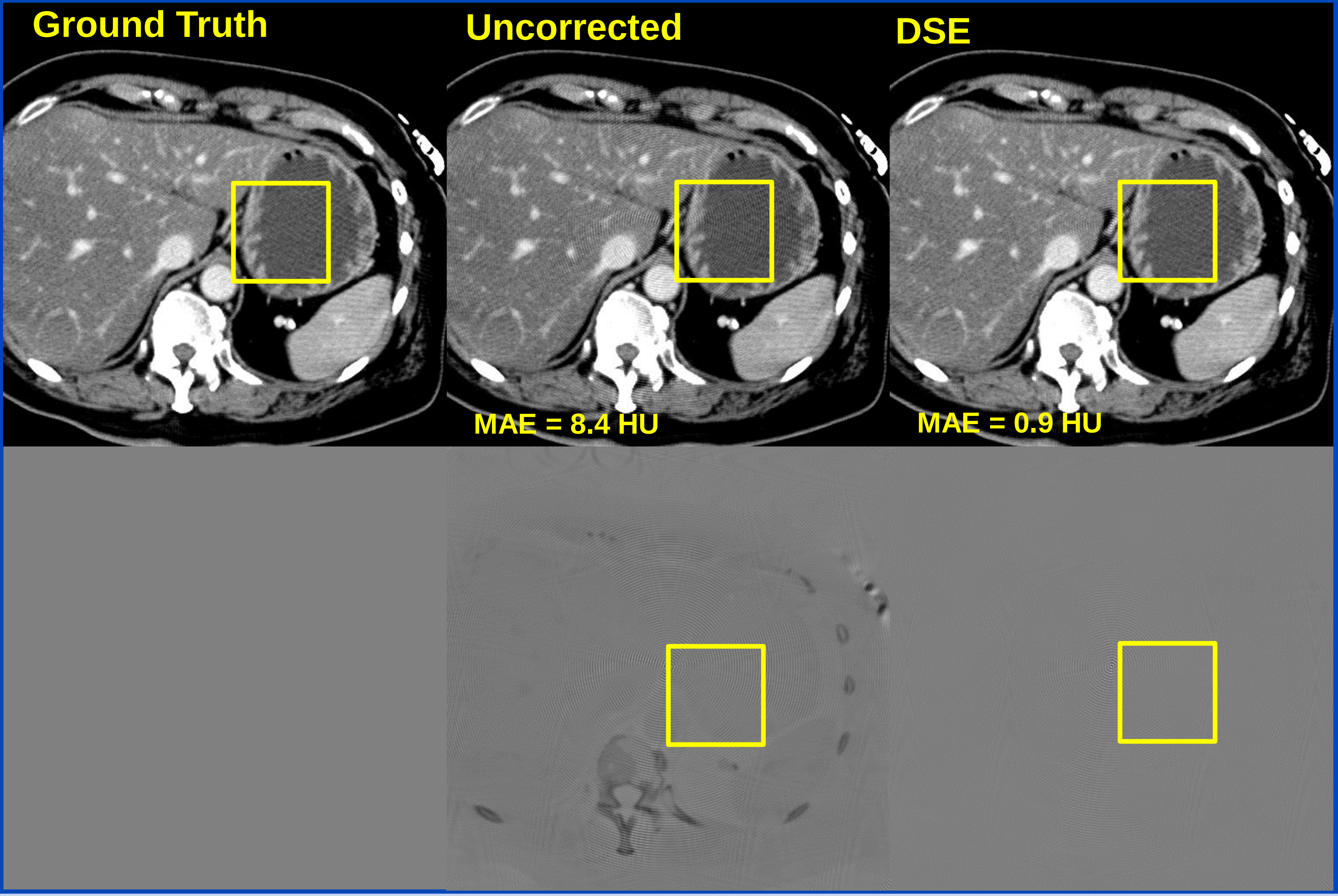
Each channel corresponds to a different pixel position between the lamallea of the ASG

Merging 6 different channels to obtain total scatter correction term



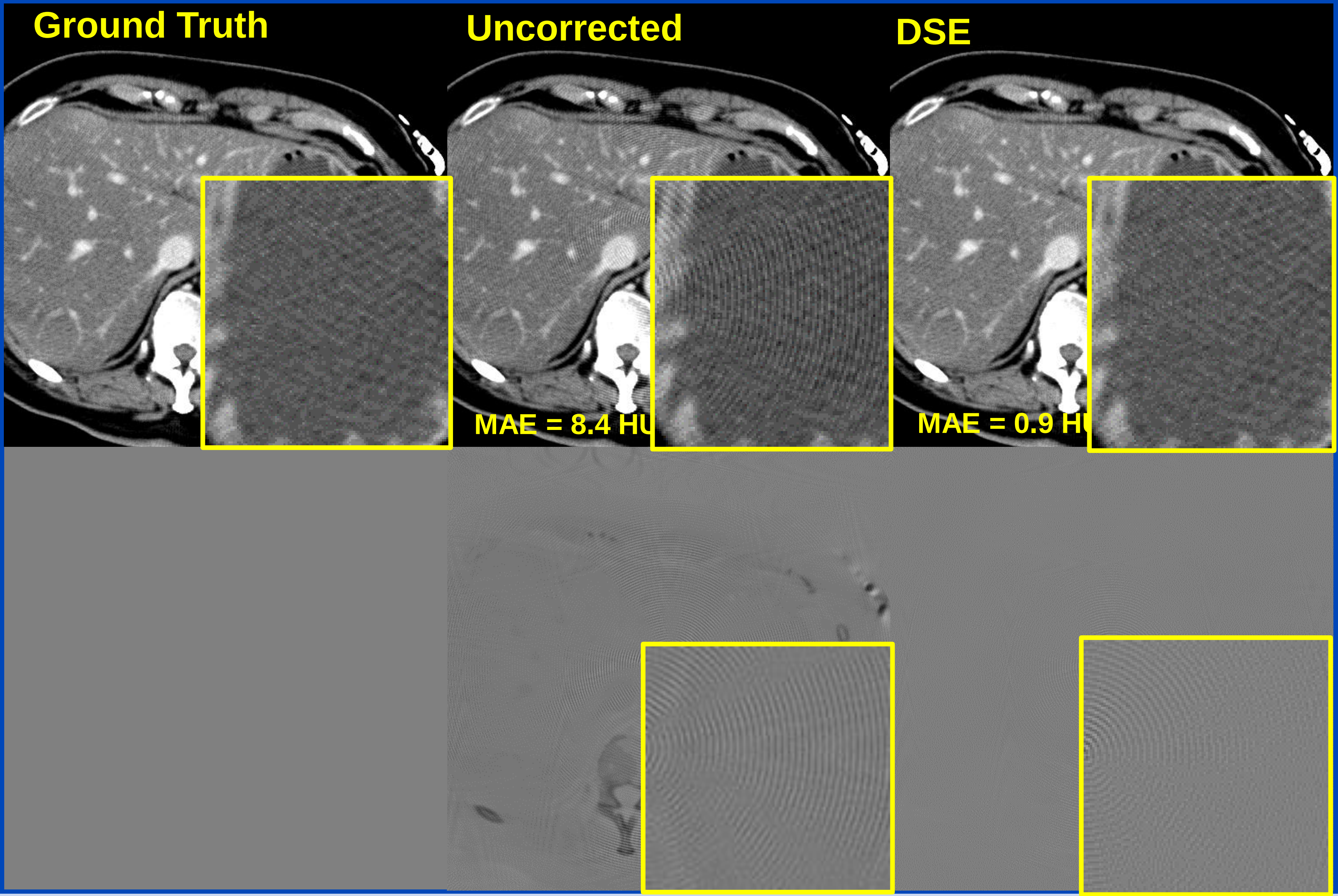
This paper received the "Highest Impact Paper Award" for the highest impact score at the 7th International Conference on Image Formation in X-Ray Com-puted Tomography in June 2022

Results in Reconstructed Images



Simulated Reconstruction *C* = 0 HU, *W* = 400 HU,
Difference to GT *C* = 0 HU, *W* = 50 HU

Results in Reconstructed Images

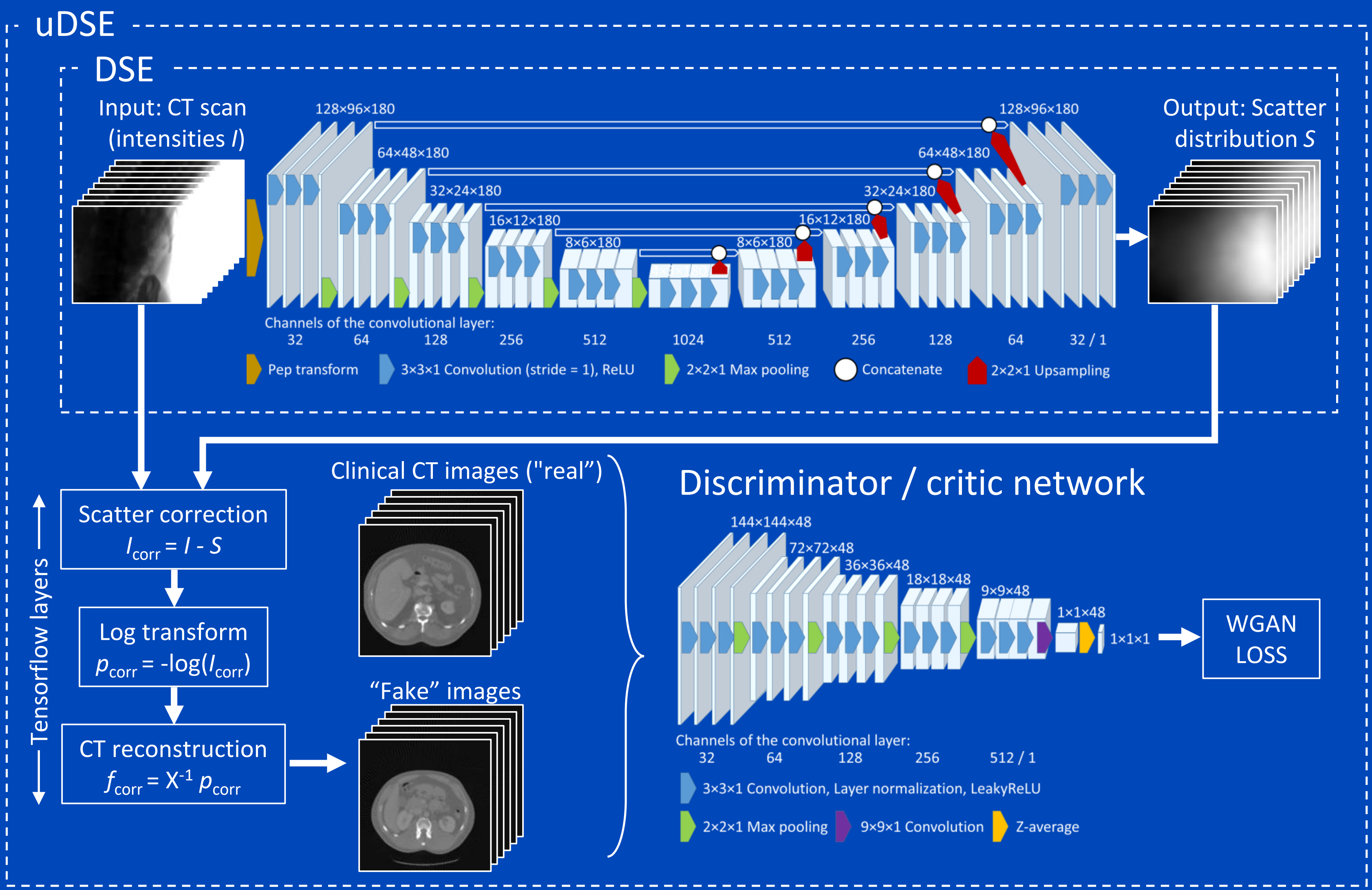


Simulated Reconstruction C = 0 HU, W = 400 HU,
Difference to GT C = 0 HU, W = 50 HU

Conclusions

- Coarse anti-scatter grid can lead to moiré artifacts due to scattered radiation.
- DSE reduces the mean absolute error (MAE) from about 9 HU to under 1 HU.
- The moiré pattern‘s amplitude can be reduced from 30 HU to less than 5 HU.

uDSE – Basis Principle



dkfz.

Datasets

- Training and testing data were generated using CT simulations based on 65 clinical CT reconstructions.
- Based on the corresponding voxel volumes, CBCT scans (120 kV, shifted detector, RFD = 1100 mm, RF = 700 mm, 360 views 360°) were simulated at five different z-positions within the abdomen region.
- Generation of one scatter corrupted dataset (30 patients) that was used as input to the generator network, one scatter-free dataset (30 patients) that was used to provide ideal reference for the critic network, and a scatter-corrupted dataset (remaining patients) for testing.

Training

- **Training of conventional DSE as reference using the following loss function:**

$$L_{\mathrm{DSE}}(\theta) = \sum_{n}^{B} \left| \frac{\mathrm{DSE}_{\theta}(I_n) - S_n}{S_n} \right|$$

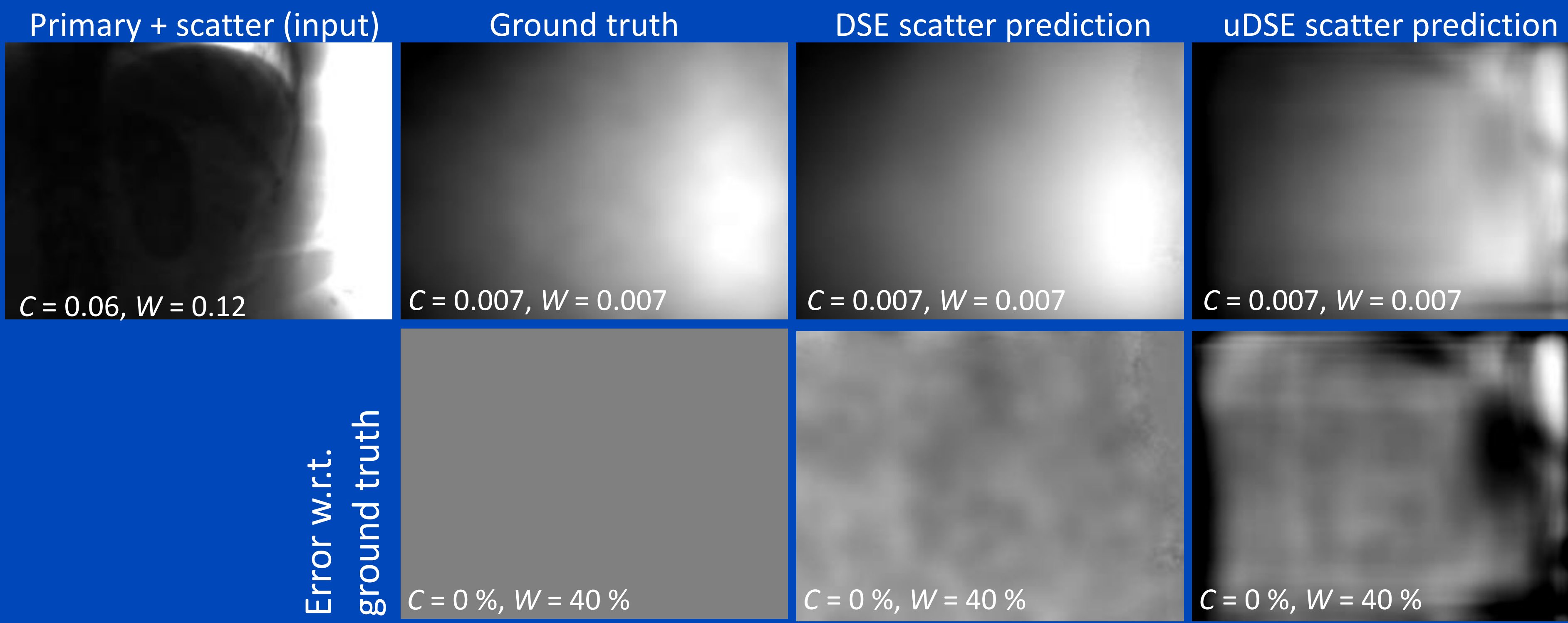
- **Training of uDSE using a WGAN setup:**

$$L_{\mathrm{critic}}(\theta_c) = \sum_{n}^{B} C_{\theta_c}(G_{\theta_g}(I_n)) - C_{\theta_c}(f_{\mathrm{real,\ n}}),$$

$$L_{\mathrm{gen}}(\theta_g) = -\sum_{n}^{B} C_{\theta_c}(G_{\theta_g}(I_n)),$$

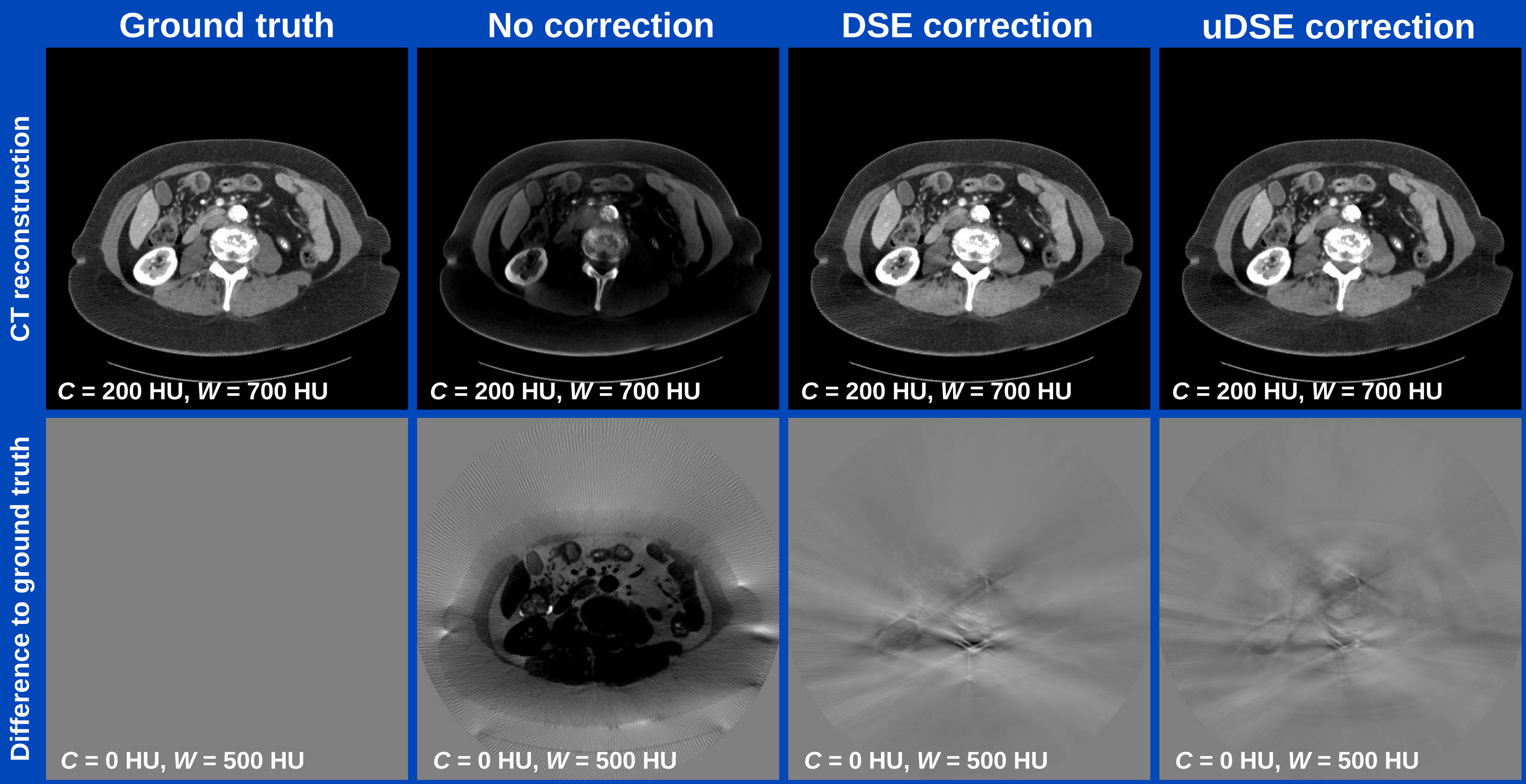
Results

Scatter Estimates



Results

CT Reconstructions



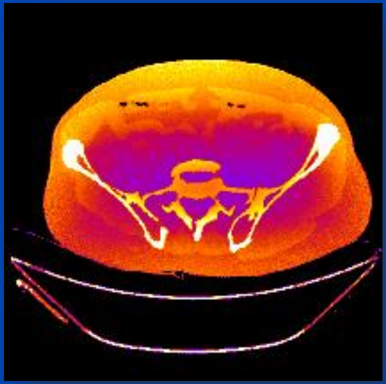
Conclusions on uDSE

- **This study demonstrates the feasibility of learning CT scatter estimation in absence of labeled data.**
- **uDSE is able to remove most of the present scatter artifacts and yields similar CT value accuracy (mean error of 27.9~HU vs. 24.7~HU) as a state-of-the-art supervised scatter estimation approach**
- **In general uDSE is not restricted to CBCT but can be trained with any tomographic input and any scatter-free reference as long as both distributions are sufficiently equal after scatter correction.**
- **Thus, uDSE has the potential to extend the concept of neural network-based scatter estimation and correction to scenarios where labels are not available or cannot be generated with sufficient accuracy.**

Deep Dose Estimation

???

In real time?



Estimation of Dose Distributions

- **Useful to study dose reduction techniques**
  - Tube current modulation
  - Prefiltration and shaped filtration
  - Tube voltage settings
  - …
- **Useful to estimate patient dose**
  - Risk assessment requires segmentation of the organs (difficult)
  - Often semiantropomorphic patient models take over
  - The infamous k-factors that convert DLP into D_{eff} are derived this way, e.g. k_{chest} = 0.014 mSv/mGy/cm
  - …
- **Could be useful for patient-specific CT scan protocol optimization**
- **However: Dose estimation does not work in real time!**

J. Maier, L. Klein, E. Eulig, S. Sawall, and M. Kachelrieß. Real-time estimation of patient-specific dose distributions for medical CT using the deep dose estimation. Med. Phys. 49(4):2259-2269, April 2022. *Best Paper within Machine Learning at ECR 2019!*

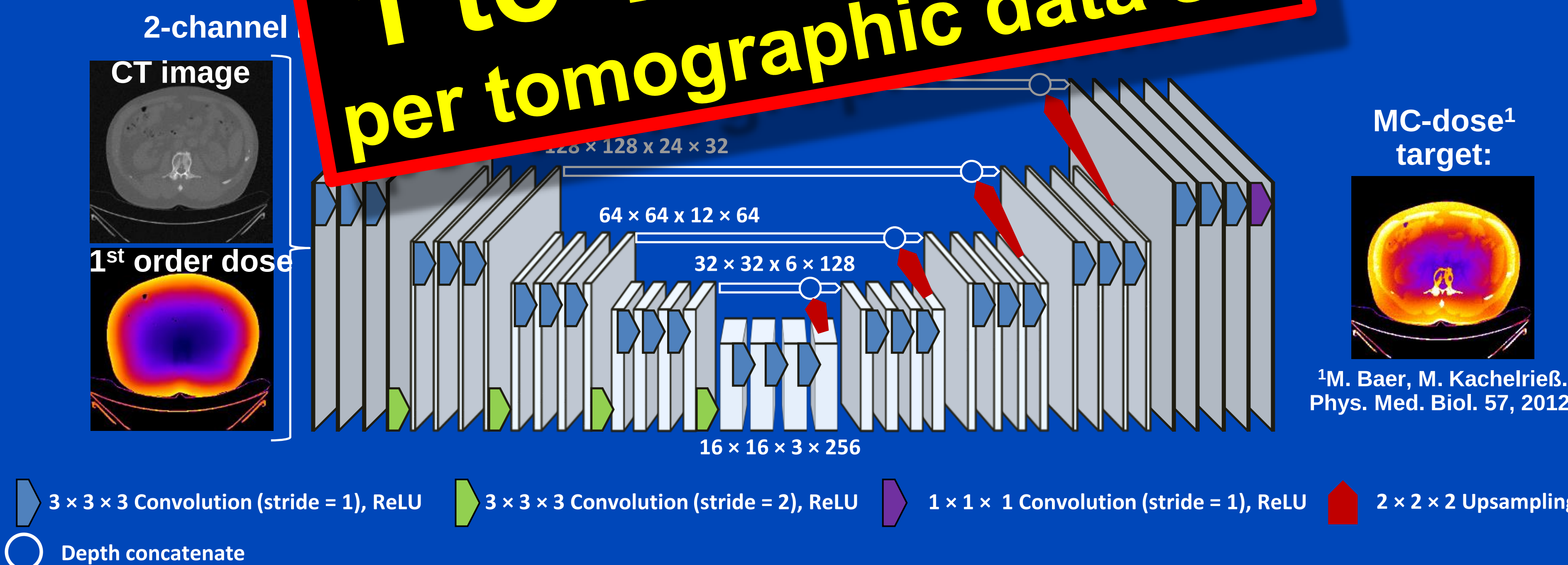
MC Dose Simulation for a 360° Scan

J. Maier, L. Klein, E. Eulig, S. Sawall, and M. Kachelrieß. Real-time estimation of patient-specific dose distributions for medical CT using the deep dose estimation. Med. Phys. 49(4):2259-2269, April 2022. *Best Paper within Machine Learning at ECR 2019!*

Deep Dose Estimation (DDE)

- Combine fast and accurate CT dose estimation using a deep convolutional neural network.
- Train the network to re… …mates given the … …mate.

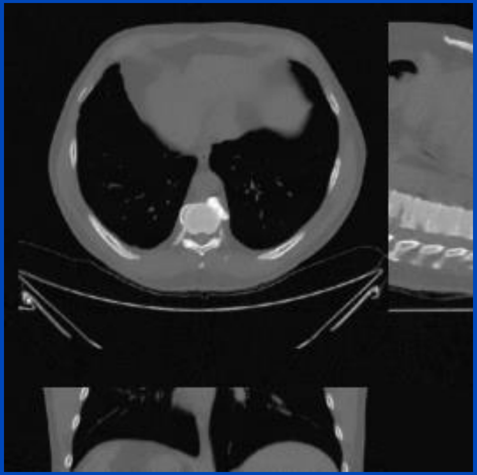
1 to 10 seconds per tomographic data set



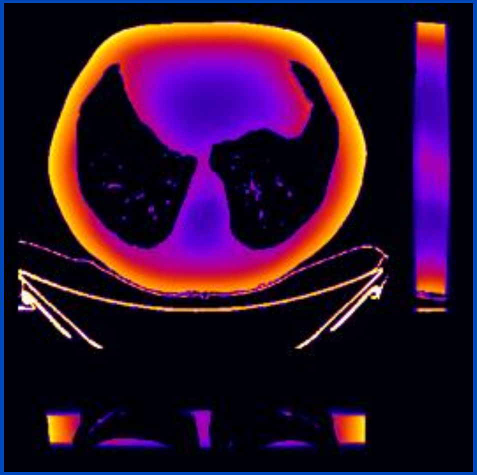
Results

Thorax, tube A, 120 kV, no bowtie

CT image



First order dose

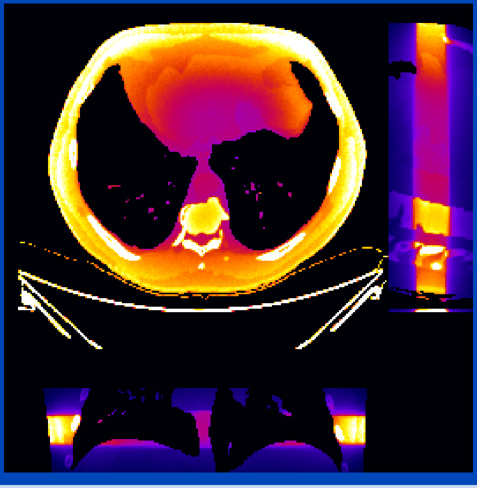


	MC	DDE
48 slices	1 h	0.25 s
whole body	20 h	5 s

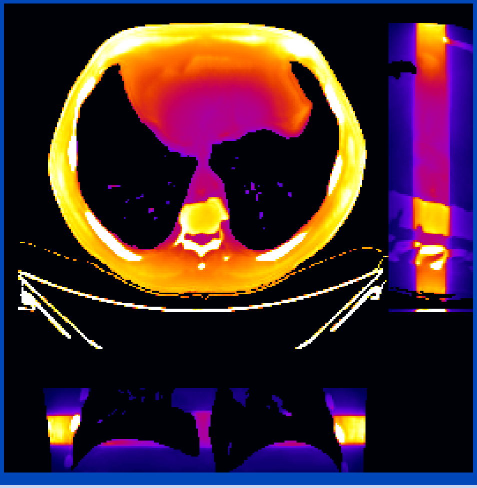
MC uses 16 CPU kernels
DDE uses one Nvidia Quadro P600 GPU

DDE training took 74 h for 300 epochs, 1440 samples, 48 slices per sample

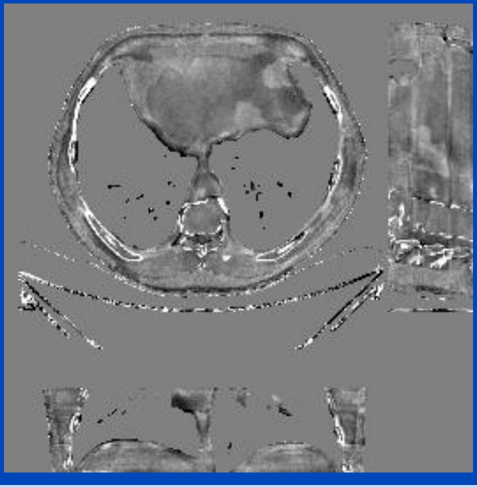
MC ground truth



DDE



Relative error



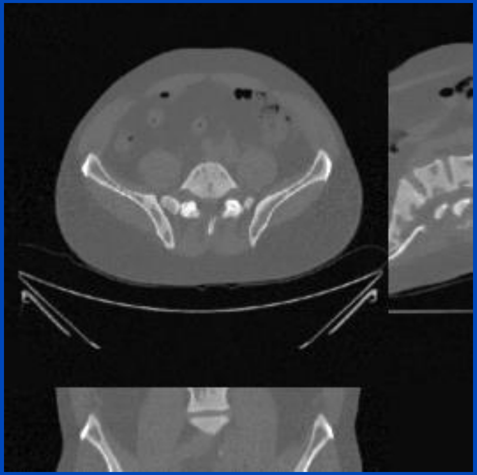
C = 0%
W = 40%

J. Maier, L. Klein, E. Eulig, S. Sawall, and M. Kachelrieß. Real-time estimation of patient-specific dose distributions for medical CT using the deep dose estimation. Med. Phys. 49(4):2259-2269, April 2022. *Best Paper within Machine Learning at ECR 2019!*

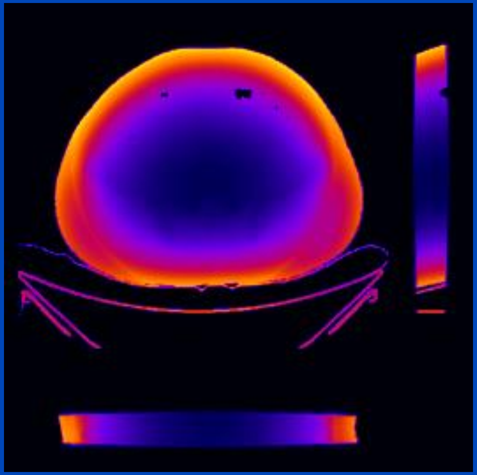
Results

Pelvis, tube B, 120 kV, no bowtie

CT image



First order dose

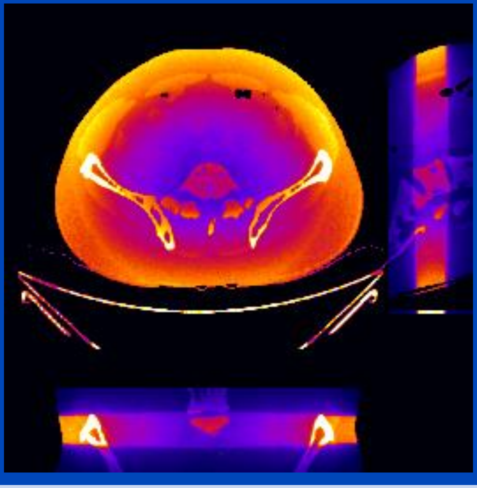


	MC	DDE
48 slices	1 h	0.25 s
whole body	20 h	5 s

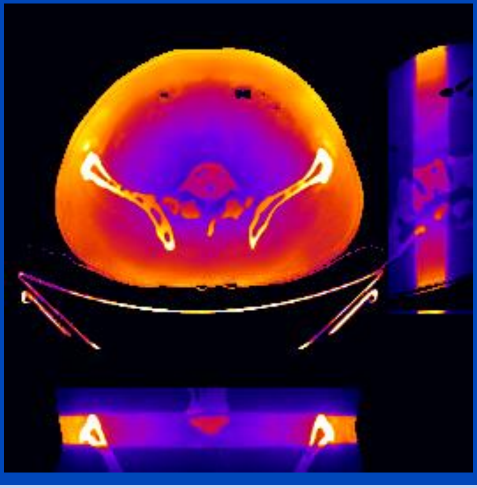
MC uses 16 CPU kernels
DDE uses one Nvidia Quadro P600 GPU

DDE training took 74 h for 300 epochs, 1440 samples, 48 slices per sample

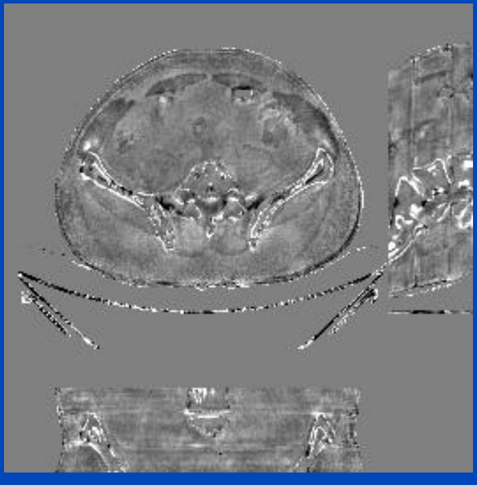
MC ground truth



DDE



Relative error



C = 0%
W = 40%

J. Maier, L. Klein, E. Eulig, S. Sawall, and M. Kachelrieß. Real-time estimation of patient-specific dose distributions for medical CT using the deep dose estimation. Med. Phys. 49(4):2259-2269, April 2022. *Best Paper within Machine Learning at ECR 2019!*

DDE's Organ Dose and D_{eff} MAPEs

Organ and ICRP weight		80 kV	100 kV	120 kV
Bone marrow	0.12	5.2%	6.7%	7.1%
Bone surface	0.01	5.7%	7.0%	7.2%
Brain	0.01	5.1%	4.9%	5.3%
Breast	0.12	1.0%	1.4%	2.1%
Colon	0.12	0.9%	1.7%	1.9%
Esophagus	0.04	1.3%	2.4%	2.3%
Gonads	0.08	3.2%	2.7%	2.2%
Liver	0.04	2.9%	1.1%	0.8%
Lung	0.12	1.7%	3.5%	4.0%
Remainder	0.12	0.9%	1.9%	2.3%
Salivary glands	0.01	4.9%	5.1%	5.3%
Skin	0.01	2.8%	3.3%	4.2%
Stomach	0.12	2.3%	1.1%	0.8%
Thyroid gland	0.04	3.1%	3.0%	2.3%
Urinary bladder	0.04	1.7%	1.7%	1.3%
Effective dose		**1.2%**	**2.5%**	**2.7%**

Weighting factors and mean absolute percentage error of the DDE organ dose values with respect to the ground truth Monte Carlo organ dose values.

Conclusions on DDE

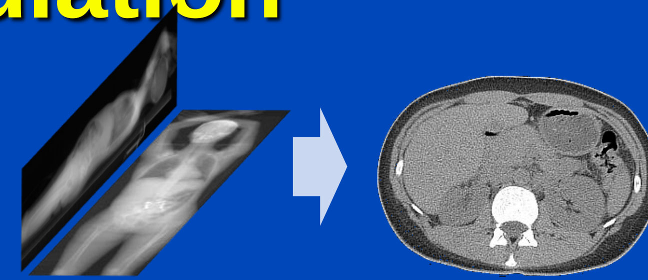
- **DDE provides accurate dose predictions**
  - **for circle scans**
  - **for sequence scans**
  - **for partial scans (less than 360°)**
  - **for limited angle scans (less than 180°)**
  - **for spiral scans**
  - **for different tube voltages**
  - **for scans with and without bowtie filtration**
  - **for scans with tube current modulation**
- **In practice it may therefore be not necessary to perform separate training runs for these cases.**
- **Thus, accurate real-time patient dose estimation may become feasible with DDE.**

J. Maier, L. Klein, E. Eulig, S. Sawall, and M. Kachelrieß. Real-time estimation of patient-specific dose distributions for medical CT using the deep dose estimation. Med. Phys. 49(4):2259-2269, April 2022. ***Best Paper within Machine Learning at ECR 2019!***

Patient Risk-Minimizing Tube Current Modulation

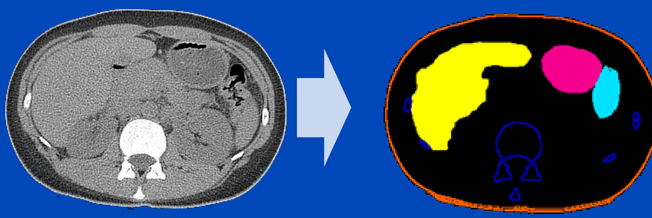
1. Coarse reconstruction from two scout views

– E.g. X. Ying, et al. X2CT-GAN: Reconstructing CT from biplanar x-rays with generative adversarial networks. CVPR 2019.



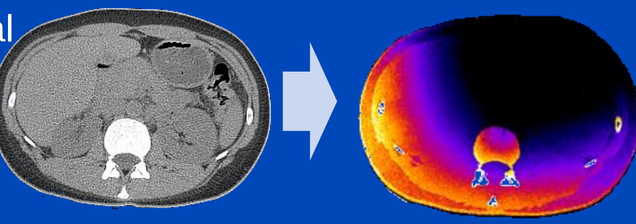
2. Segmentation of radiation-sensitive organs

– E.g. S. Chen, M. Kachelrieß et al., Automatic multi-organ segmentation in dual-energy CT (DECT) with dedicated 3D fully convolutional DECT networks. Med. Phys. 2019.



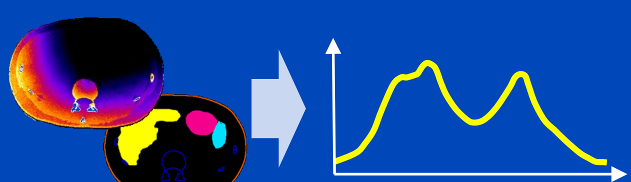
3. Calculation of the effective dose per view using the deep dose estimation (DDE)

– J. Maier, E. Eulig, S. Dorn, S. Sawall and M. Kachelrieß. Real-time patient-specific CT dose estimation using a deep convolutional neural network. IEEE Medical Imaging Conference Record, M-03-178: 3 pages, Nov. 2018.



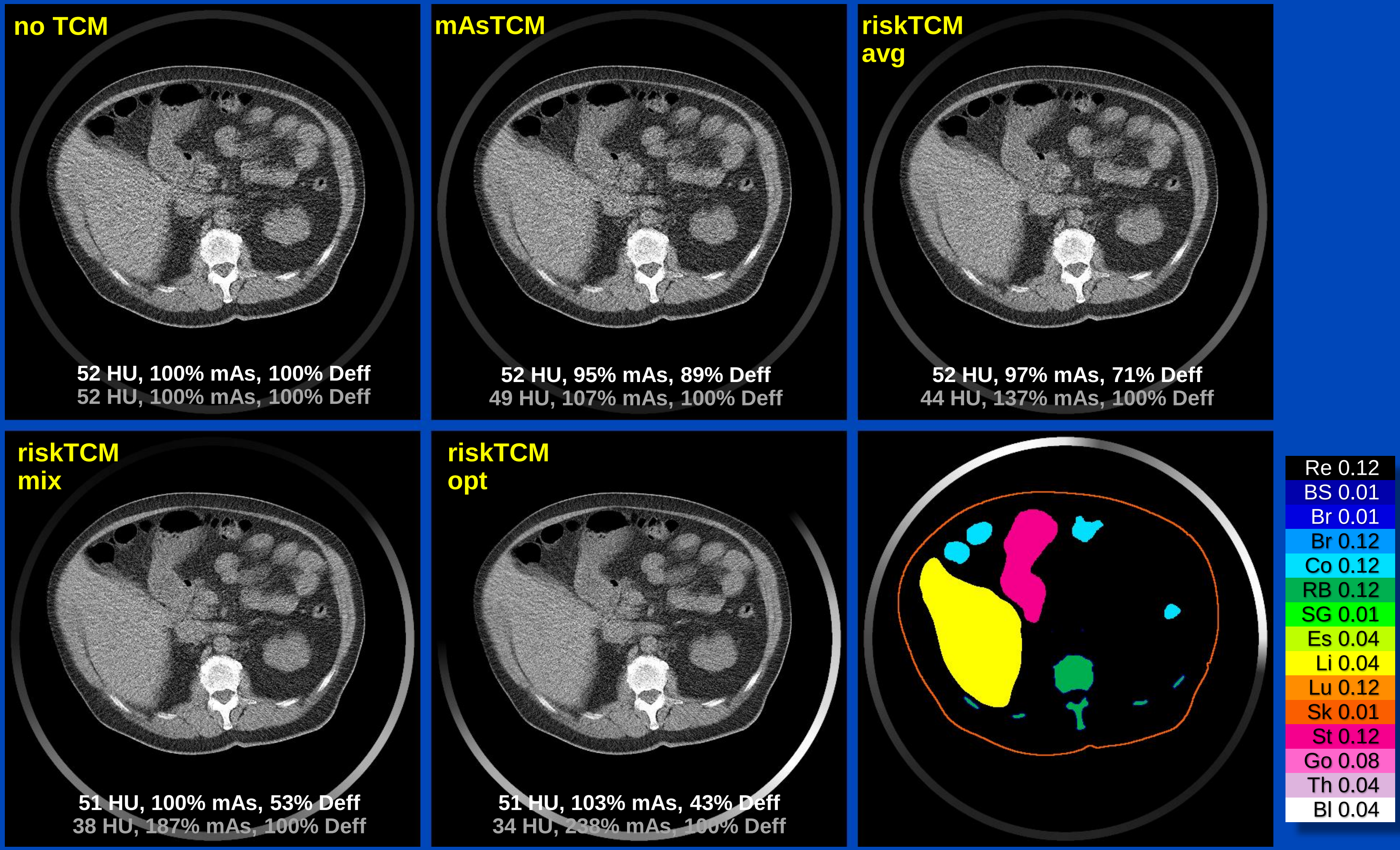
4. Determination of the tube current modulation curve that minimizes the radiation risk

– L. Klein, C. Liu, J. Steidel, L. Enzmann, M. Knaup, S. Sawall, A. Maier, M. Lell, J. Maier, and M. Kachelrieß. Patient-specific radiation risk-based tube current modulation for diagnostic CT. Med. Phys. 49(7):4391-4403, July 2022.



[1]L. Klein, C. Liu, J. Steidel, L. Enzmann, M. Knaup, S. Sawall, A. Maier, M. Lell, J. Maier, and M. Kachelrieß. Patient-specific radiation risk-based tube current modulation for diagnostic CT. Med. Phys. 49(7):4391-4403, July 2022.

Patient 04 - Abdomen



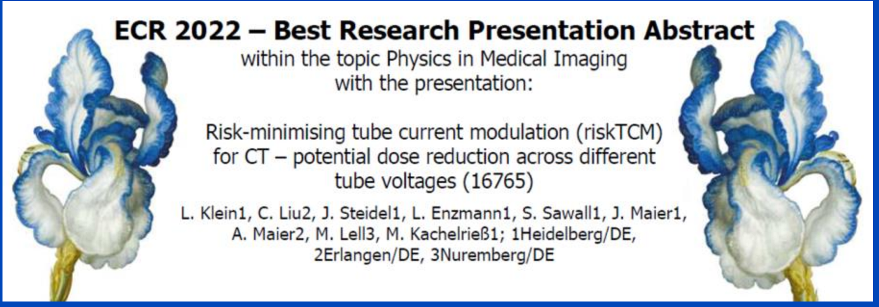
C = 25 HU, *W* = 400 HU

[1]L. Klein, C. Liu, J. Steidel, L. Enzmann, M. Knaup, S. Sawall, A. Maier, M. Lell, J. Maier, and M. Kachelrieß. Patient-specific radiation risk-based tube current modulation for diagnostic CT. Med. Phys. 49(7):4391-4403, July 2022.

Conclusions on RiskTCM

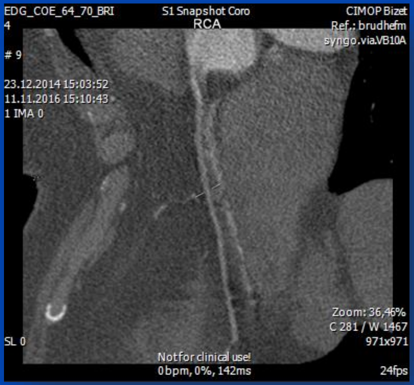
- Risk-specific TCM minimizes the patient risk.
- With D_{eff} as a risk model riskTCM can reduce risk by up to 30%, compared with the gold standard mAsTCM.
- Other risk models, in particular age-, weight- and sex-specific models, can be used with riskTCM [redacted]ell.
- Note:
  - [redacted]
  - [redacted] good for the patient
  - detector flux equalizing TCM = good for the detector

It is up to the vendors to take action!

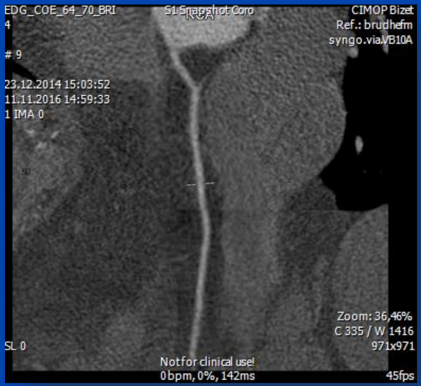


[1]L. Klein, C. Liu, J. Steidel, L. Enzmann, M. Knaup, S. Sawall, A. Maier, M. Lell, J. Maier, and M. Kachelrieß.
Patient-specific radiation risk-based tube current modulation for diagnostic CT. Med. Phys. 49(7):4391-4403, July 2022.

Deep Cardiac Motion Compensation



???



Motivation

C = 0 HU, *W* = 1200 HU

Motion artifacts

High noise levels

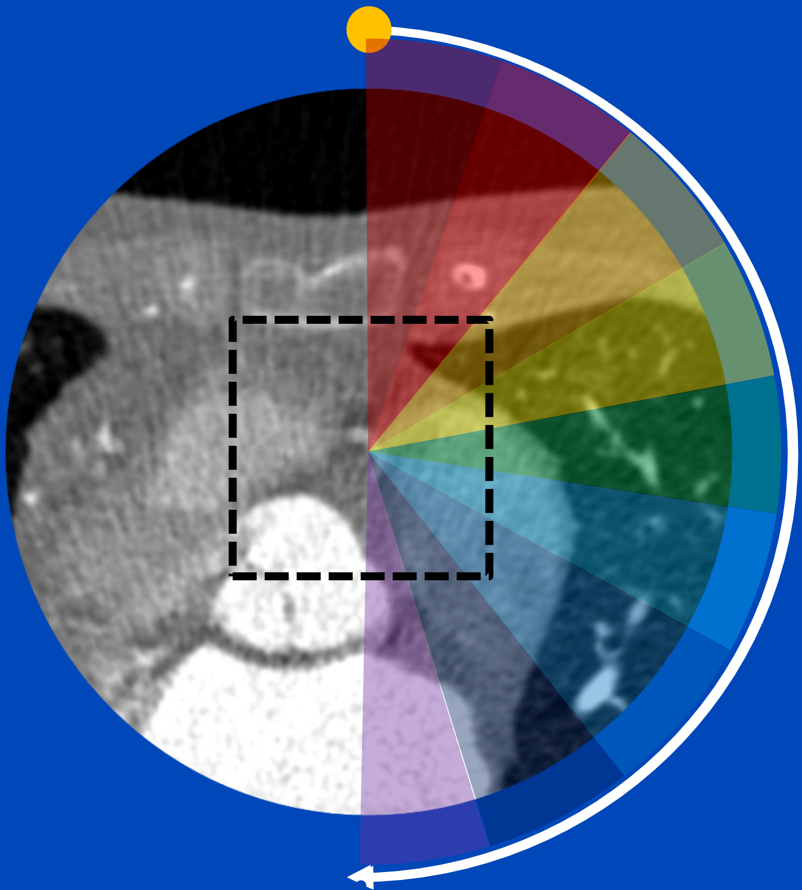
Table 3: Reason for FFR_{CT} Rejection in the ADVANCE Registry and Clinical Cohort *

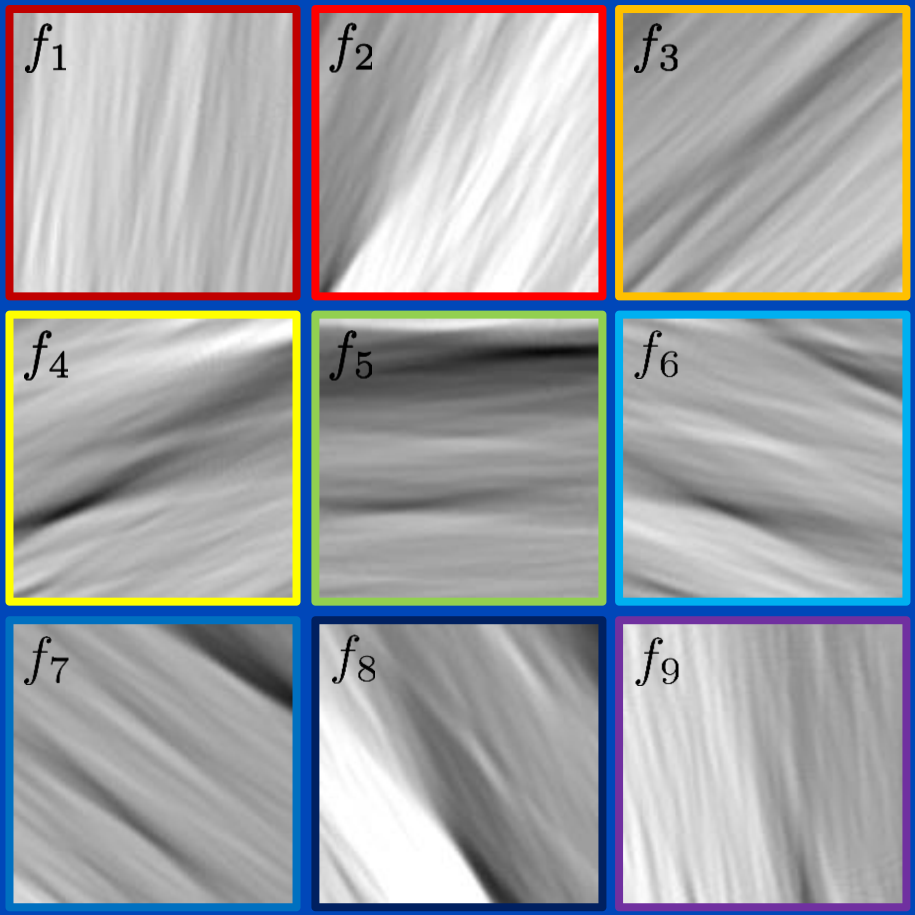
	FFR_{CT} Rejected*	
Reason for Rejection	ADVANCE Registry (n = 80)	Clinical Cohort (n = 892)
Inadequate image quality†		
Blooming	4 (5.0)	29 (3.0)
Clipped structure	4 (5.0)	39 (4.3)
Motion artifacts	63 (78.0)	729 (81.4)
Image noise	2 (2.5)	198 (22.1)
Inappropriate submission		
Stent or previous coronary artery bypass graft present	5 (6.2)	116 (13.0)
Cardiac hardware present	2 (2.5)	29 (3.2)

The rejection rate was 892 of 10416 cases submitted

* G. Pontone et al., "Determinants of Rejection Rate for Coronary CT Angiography Fractional Flow Reserve Analysis", *Radiology*, *292*(3), 597–605 (2019)

Partial Angle-Based Motion Compensation (PAMoCo)

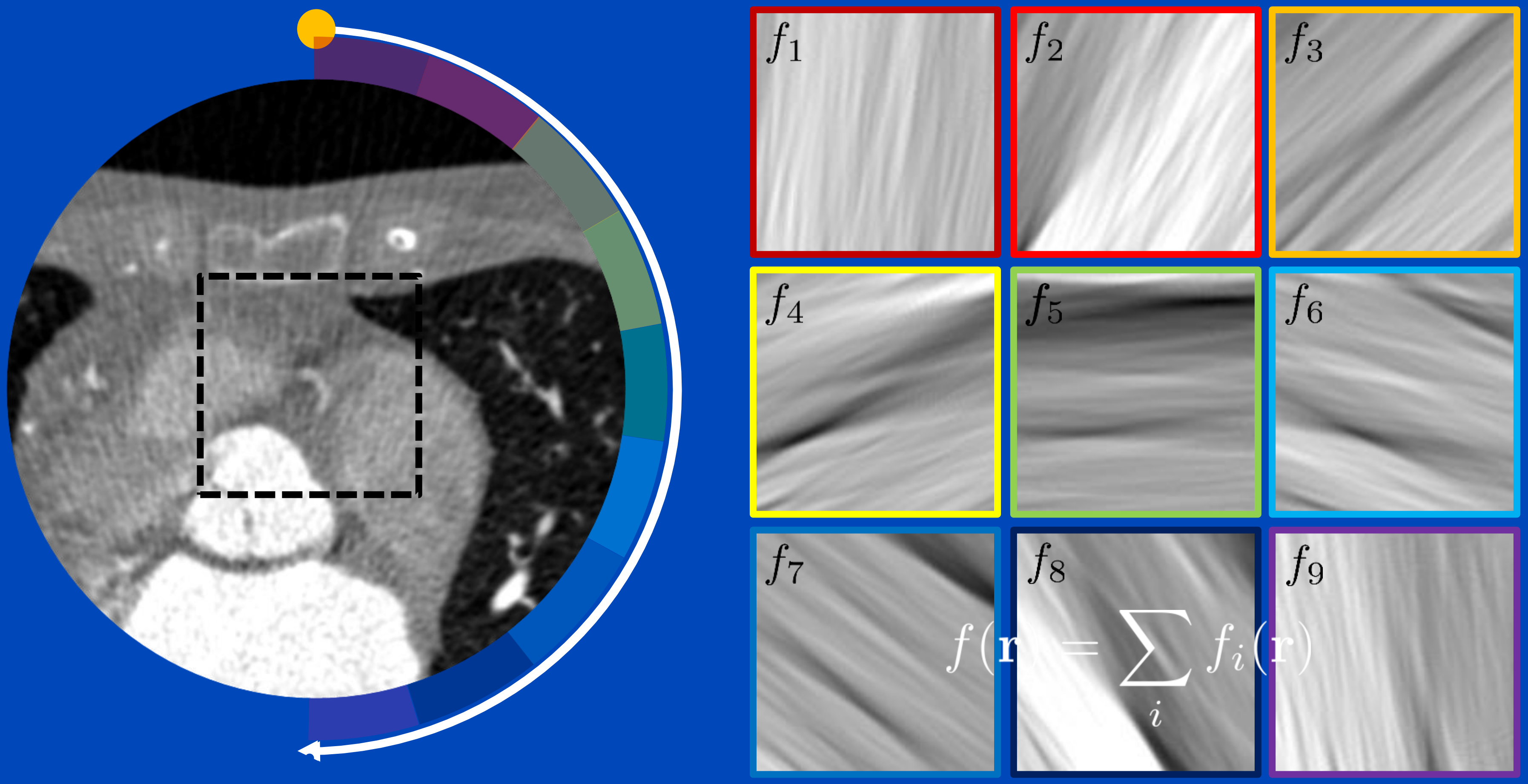




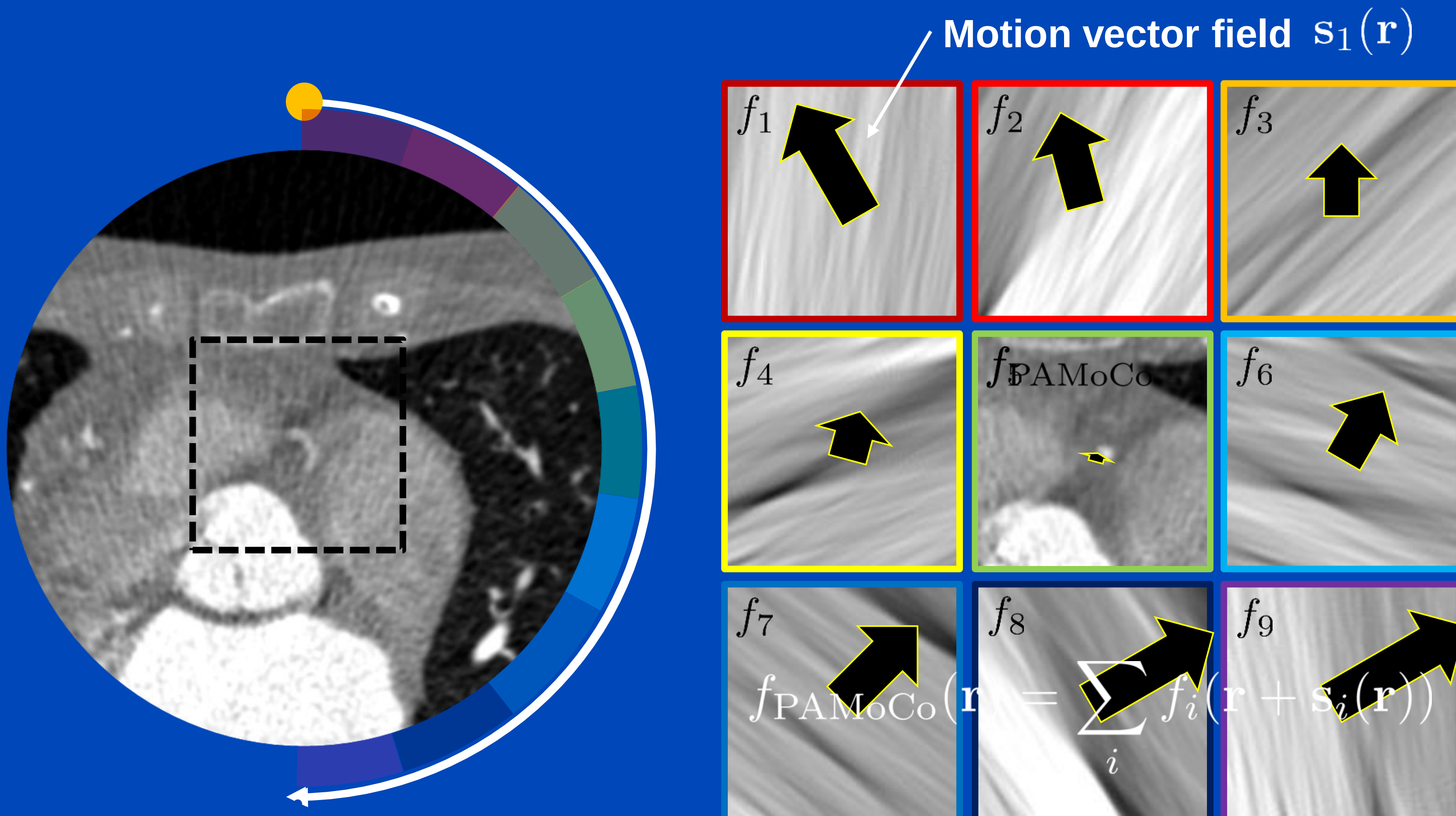
Animated rotation time = 100 × real rotation time

dkfz.

Partial Angle-Based Motion Compensation (PAMoCo)



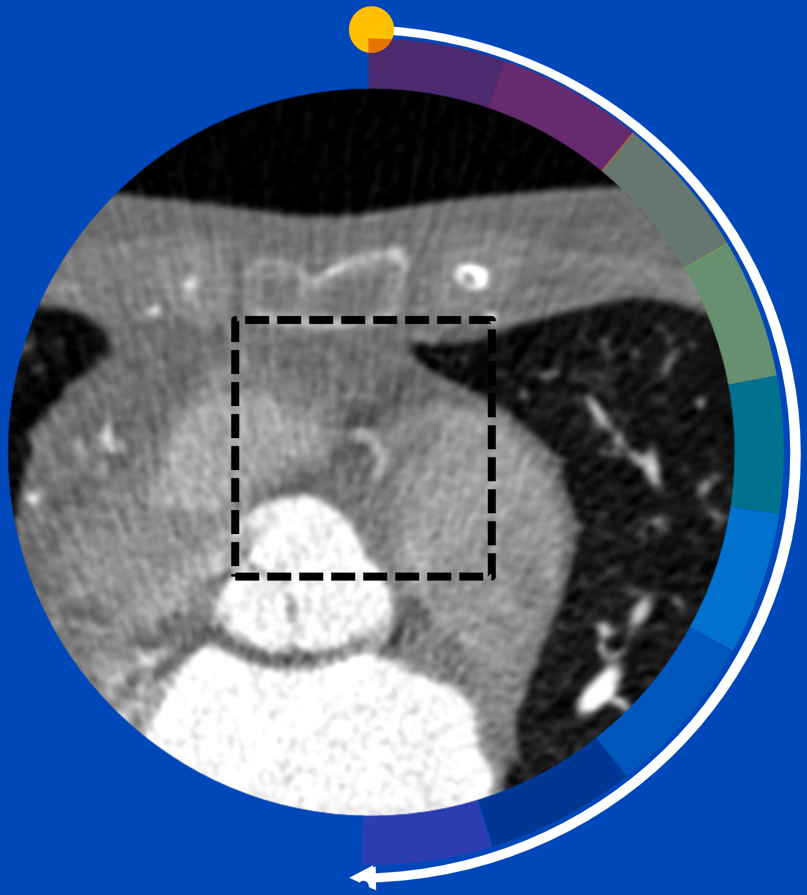
Partial Angle-Based Motion Compensation (PAMoCo)



$$f_{\mathrm{PAMoCo}}(\mathbf{r}) = \sum_i f_i(\mathbf{r} + \mathbf{s}_i(\mathbf{r}))$$

Apply motion vector fields (MVFs) to partial angle reconstructions

Partial Angle-Based Motion Compensation (PAMoCo)



Prior work:

[1] S. Kim et al., "Cardiac motion correction based on partial angle reconstructed images in x-ray CT", Med. Phys. 42 (5): 2560–2571 (2015).

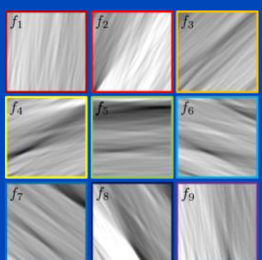
[2] J. Hahn et al., "Motion compensation in the region of the coronary arteries based on partial angle reconstructions from short-scan CT data", Med. Phys. 44 (11): 5795–5813 (2017).

[3] S. Kim et al., "Cardiac motion correction for helical CT scan with an ordinary pitch", IEEE TMI 37 (7): 1587–1596 (2018).

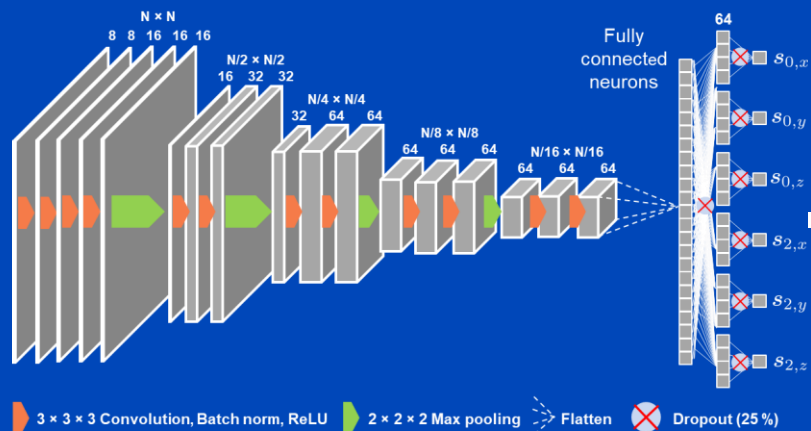
→ **Limitation: Challenging / time-consuming optimization**

Deep Partial Angle-Based Motion Compensation (Deep PAMoCo)

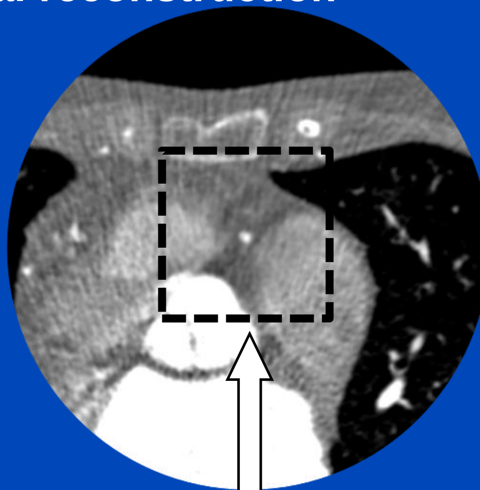
PARs centered around coronary artery

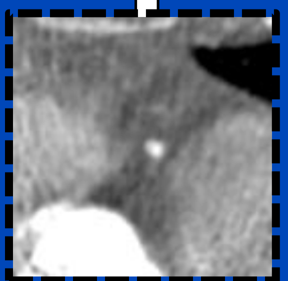


Neural network to predict parameters of a motion model



Reinsertion of patch into initial reconstruction





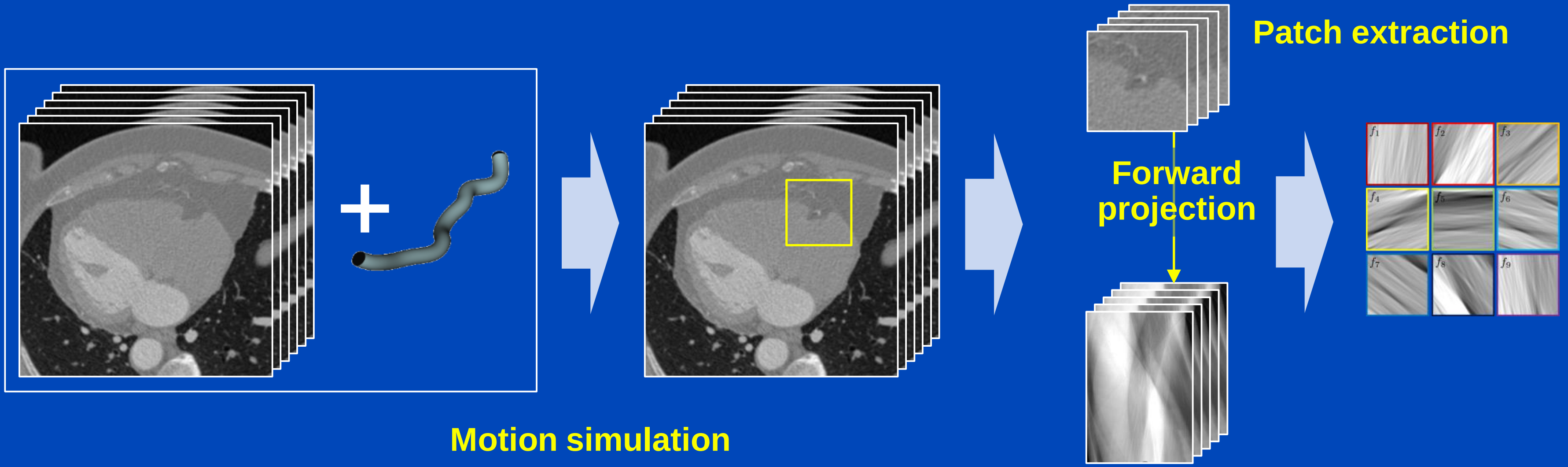
Spatial transformer

Application of the motion model to the PARs via a spatial transformer[1]

[1] M. Jaderberg et al., “Spatial transformer networks”, NIPS 2015: 2017–2025 (2015).

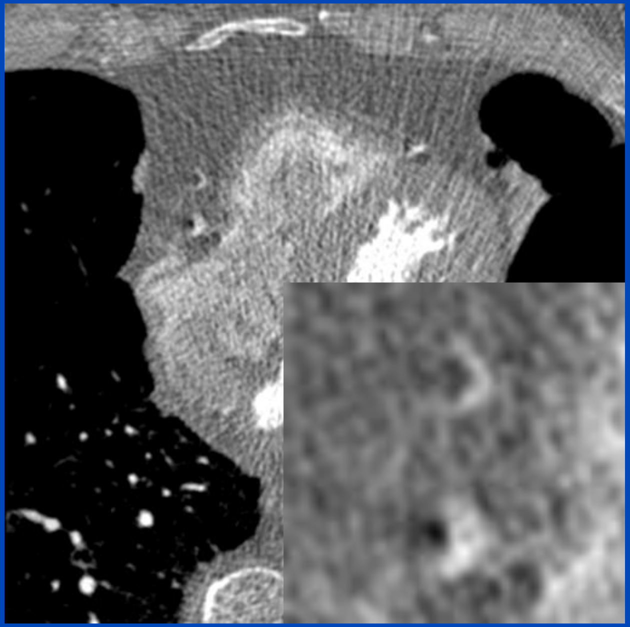
Training Data Generation

- Removal of coronary arteries from real CT reconstructions.
- Insertion of artificial coronary arteries with different shape, size, and contrast.
- Simulation of CT scans with coronary artery motion.

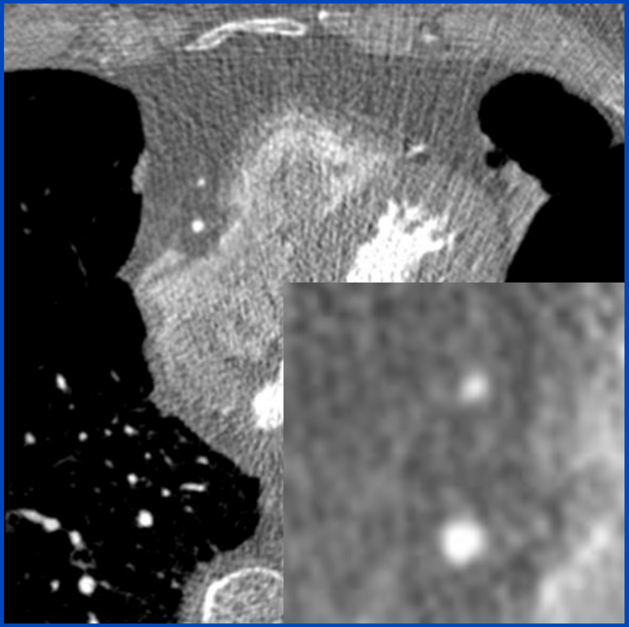


Patient 1

Original



Deep PAMoCo

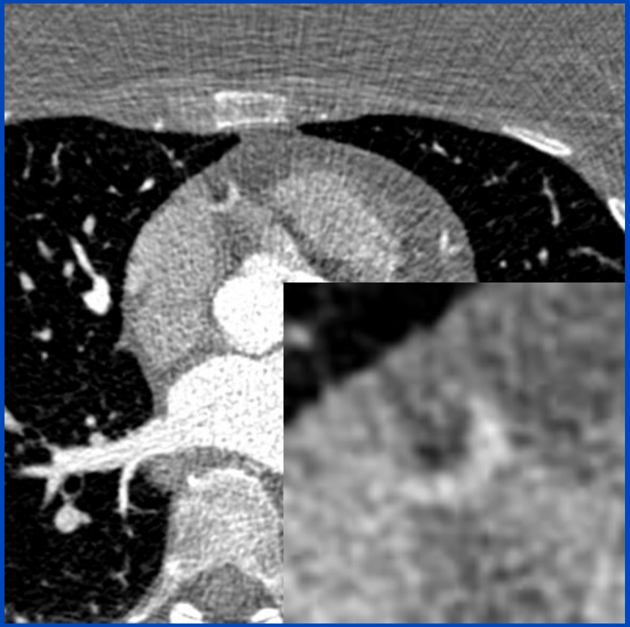


C = 0 HU, *W* = 1400 HU

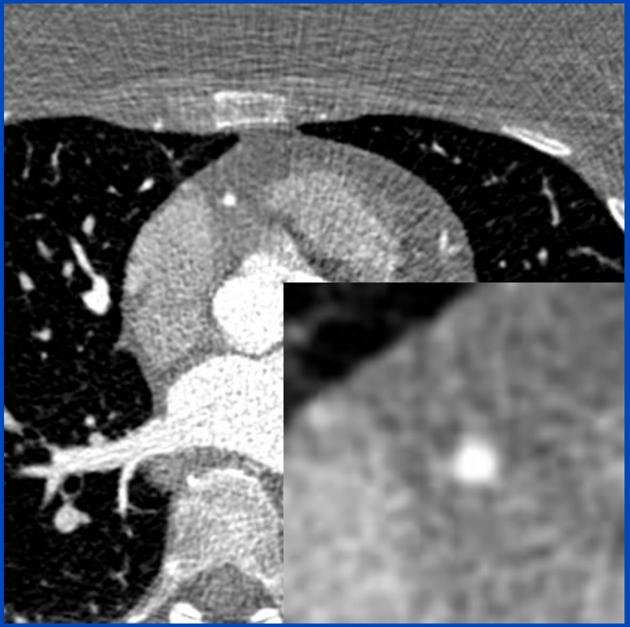
J. Maier, S. Lebedev, J. Erath, E. Eulig, S. Sawall, E. Fournié, K. Stierstorfer, M. Lell, and M. Kachelrieß. Deep learning-based coronary artery motion estimation and compensation for short-scan cardiac CT. Med. Phys. 48(7):3559-3571, July 2021.

Patient 2

Original



Deep PAMoCo

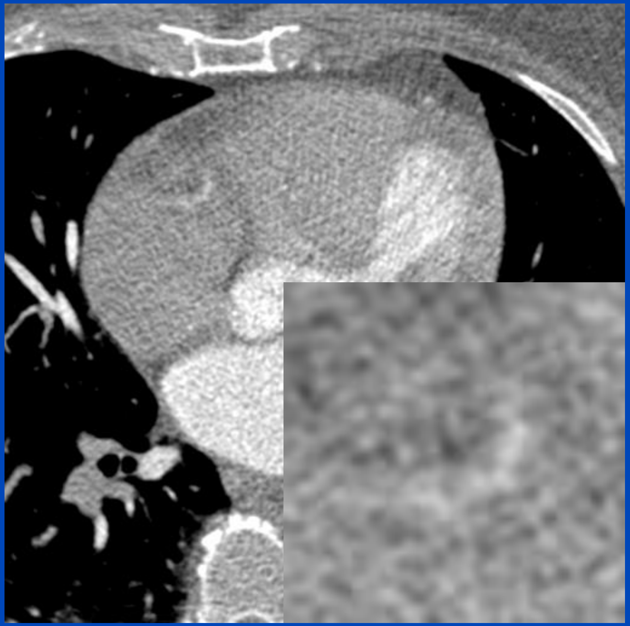


C = 0 HU, W = 1600 HU

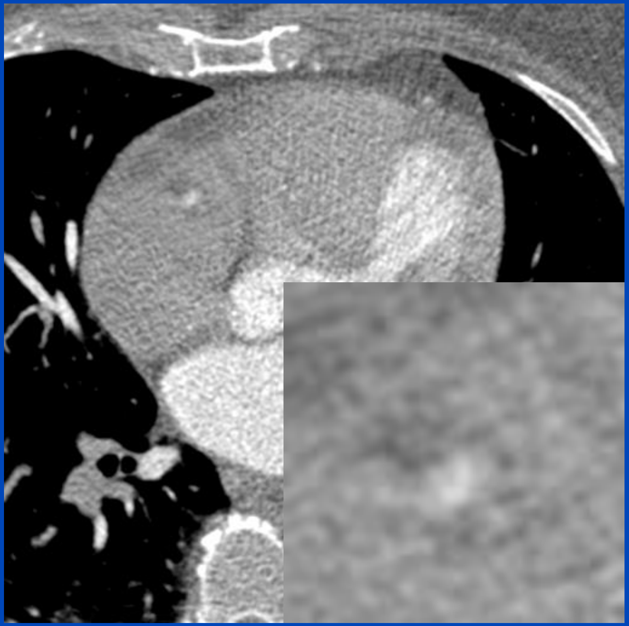
J. Maier, S. Lebedev, J. Erath, E. Eulig, S. Sawall, E. Fournié, K. Stierstorfer, M. Lell, and M. Kachelrieß. Deep learning-based coronary artery motion estimation and compensation for short-scan cardiac CT. Med. Phys. 48(7):3559-3571, July 2021.

Patient 3

Original



Deep PAMoCo

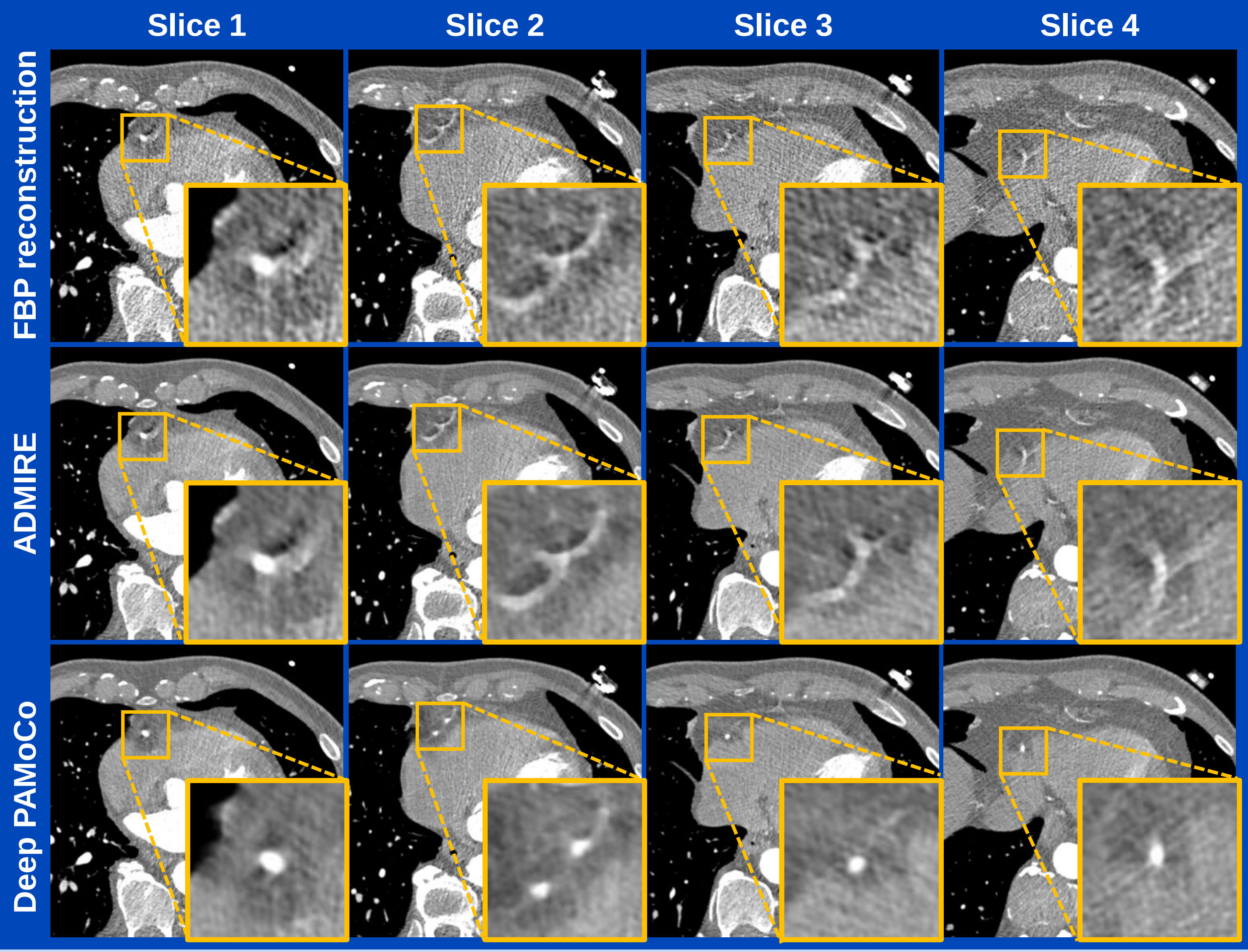


C **= 0 HU,** ***W*** **= 1000 HU**

J. Maier, S. Lebedev, J. Erath, E. Eulig, S. Sawall, E. Fournié, K. Stierstorfer, M. Lell, and M. Kachelrieß. Deep learning-based coronary artery motion estimation and compensation for short-scan cardiac CT. Med. Phys. 48(7):3559-3571, July 2021.

Patient 4 (Iterative Recon)

Measurements at a Siemens Somatom AS, patient 1



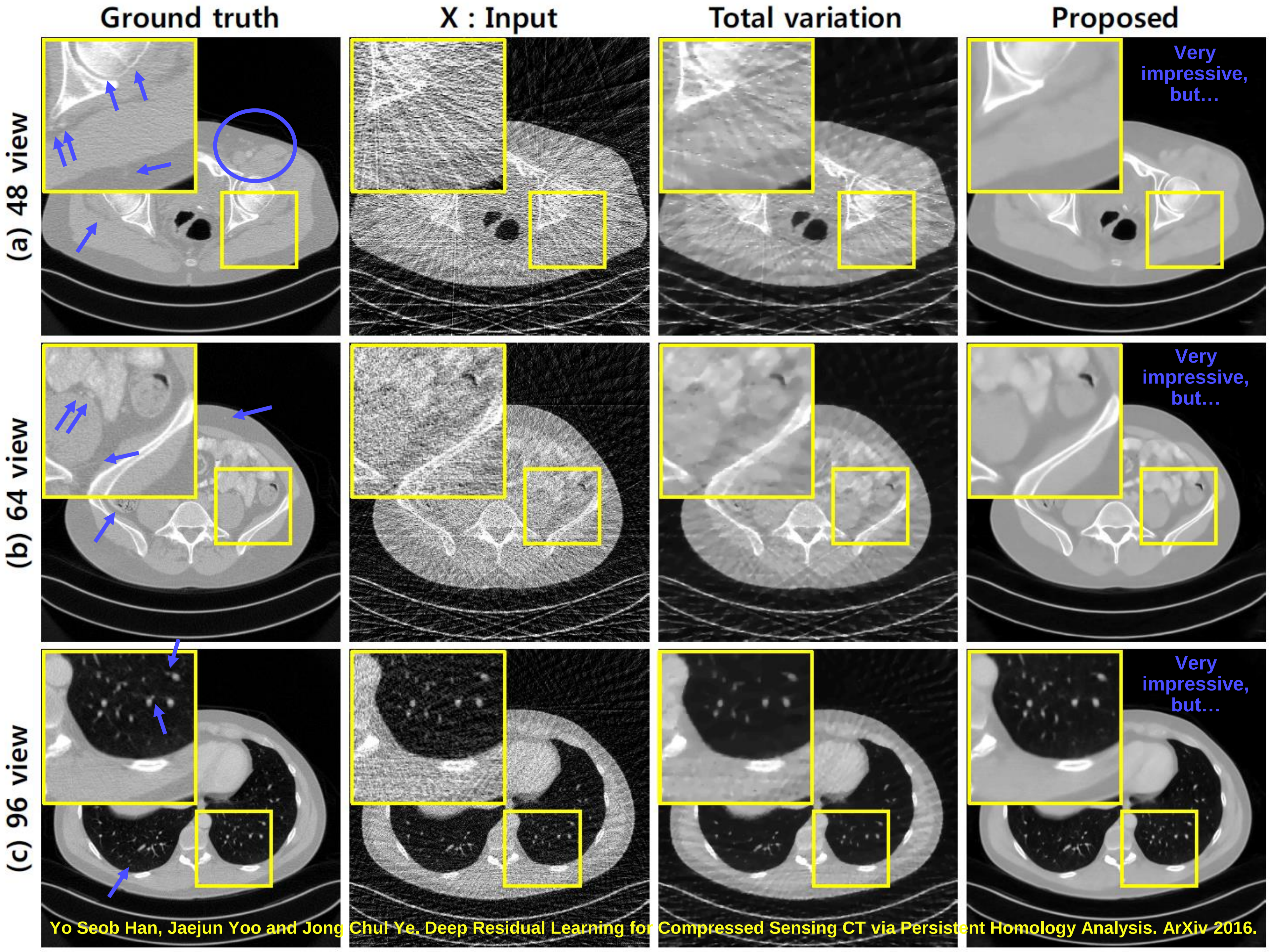
C = 0 HU, *W* = 1200 HU

Are the Methods Reliable?

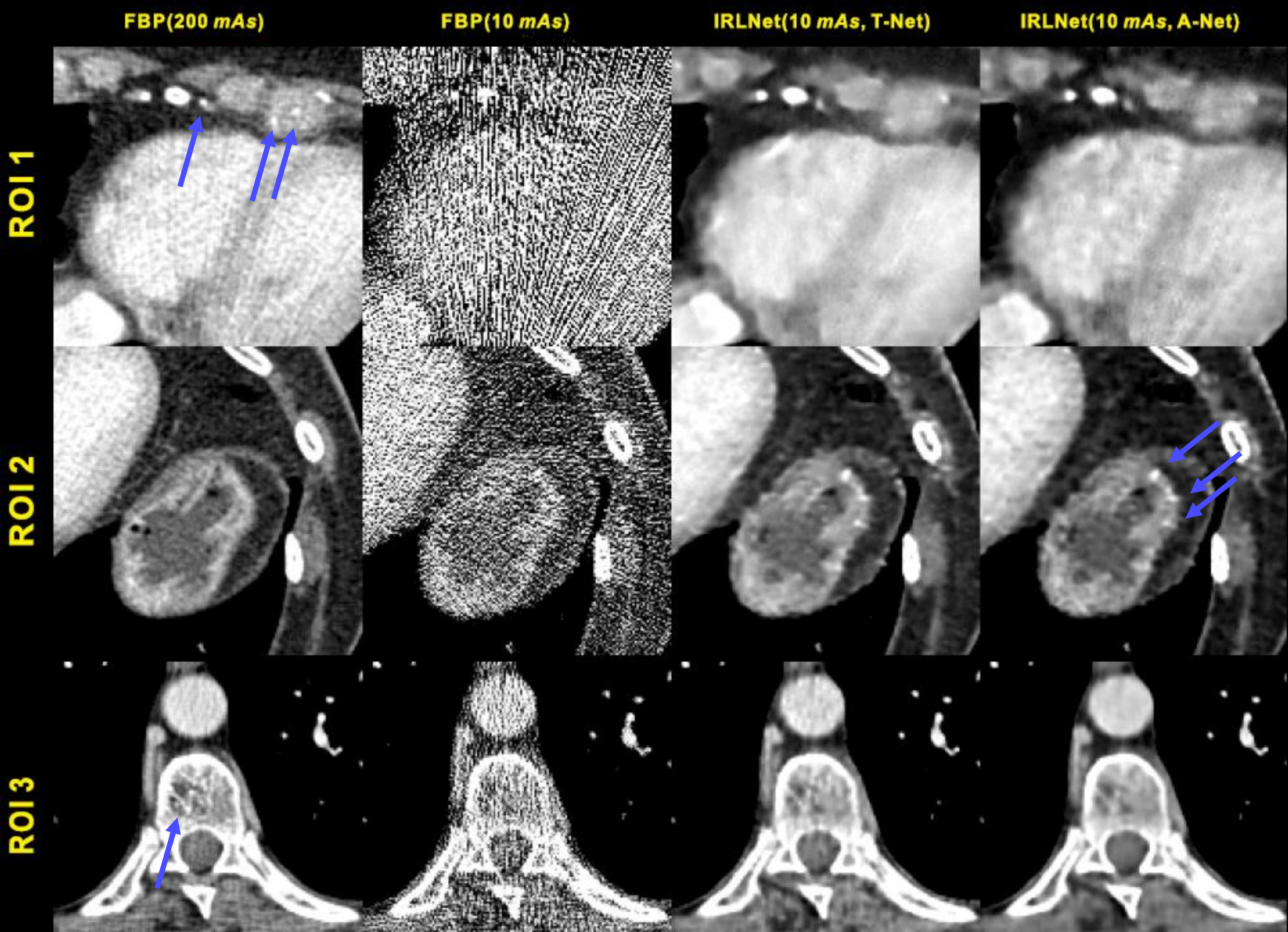
- Studies about explainability of AI in CT image formation are more than sparse.
- My thoughts:
  - Cosmetic corrections: Unclear if noise reduction, metal artifact reduction etc. is removing/adding lesions. The whole process is a black box.
  - Physical corrections: A clear physical meaning and rawdata fidelity appear more reliable. Examples:
    - » MAR or detruncation networks where the NN output is used only to forward project and inpaint/extrapolate the rawdata
    - » Scatter correction that estimates a smooth physically realistic (trained with MC) scatter signal in intensity domain
    - » Motion correction networks that estimate motion vectors rather than manipulating the voxel values

Explainable AI for CT: Analyzing CT Image Denoising Networks by Reconstructing their Invariances

- Elias Eulig, Björn Ommer, and Marc Kachelrieß
- RSNA 2022



Ground truth
X : Input
Total variation
Proposed
(a) 48 view
(b) 64 view
(c) 96 view
Very impressive, but…
Very impressive, but…
Very impressive, but…
Yo Seob Han, Jaejun Yoo and Jong Chul Ye. Deep Residual Learning for Compressed Sensing CT via Persistent Homology Analysis. ArXiv 2016.



Y. Wang et al. Iterative quality enhancement via residual-artifact learning networks for low-dose CT. Phys. Med. Biol. 63:215004, 2018.

Motivation

In general:

- Deep learning methods are employed for many problems in medical image formation, including image-based noise reduction.
- However, they lack interpretability due to black-box nature of DNNs. Recent advancement in generative modelling signal false confidence.

Aim of this work:

- Lay fundamentals for post-hoc interpretability and robustness analysis of denoising DNNs.
- Use two simple denoising networks *f* as initial examples:
  - Chen‘s simple 3-layer CNN trained with $\mathcal{L}_2$ loss[1]
  - Yang‘s Wasserstein GAN with additional perceptual loss[2]
- See what they have learned to represent and what to ignore: For a given output *x´* there are many inputs *x* that produce the same output *x´* = *f*(*x*).
- Employ low dose CT image and projection dataset for all studies.[3]

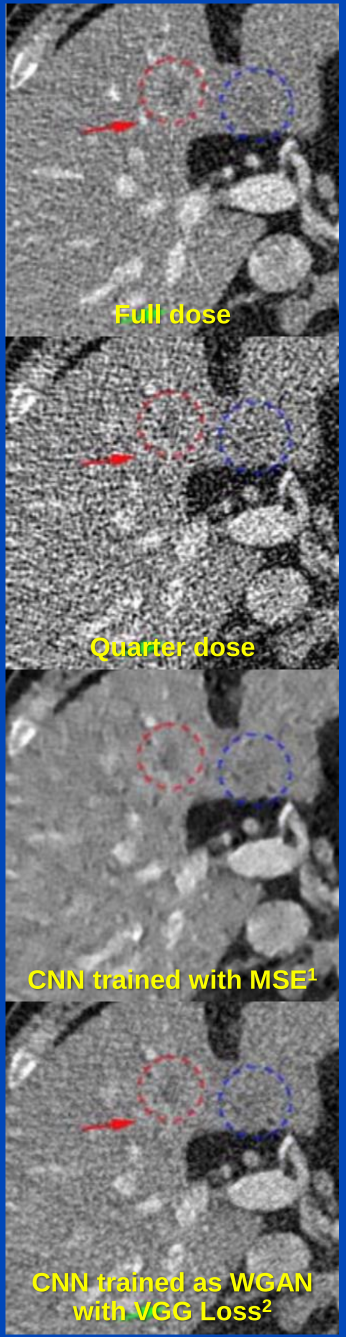


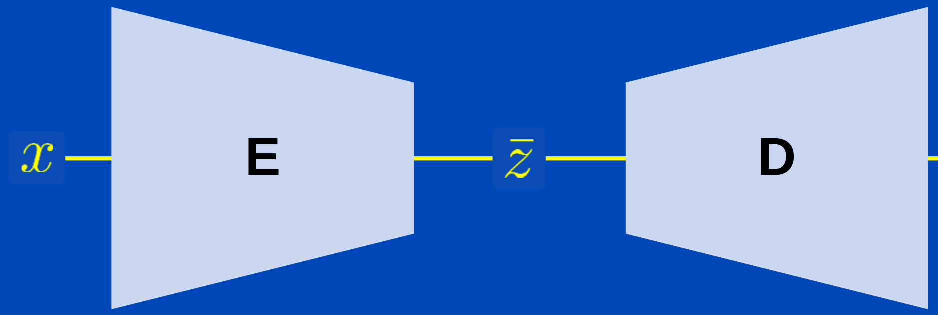
Figure from reference [2]

[1] H. Chen et al., "Low-dose CT denoising with convolutional neural network”, ISBI 2017, 2017.
[2] Q. Yang *et al*., "Low-Dose CT Image Denoising Using a Generative Adversarial Network […]”, in *IEEE TMI*, vol. 37, no. 6, 2018.
[3] C. McCollough et al., “Data from low dose CT image and projection data [data set],” The Cancer Imaging Archive, 2020.

dkfz.

Recap 1: What is an Autoencoder (AE)?

- In and output domain are the same, here *x*.
- Bottleneck *z* enforces the encoder and decoder to do a good job.

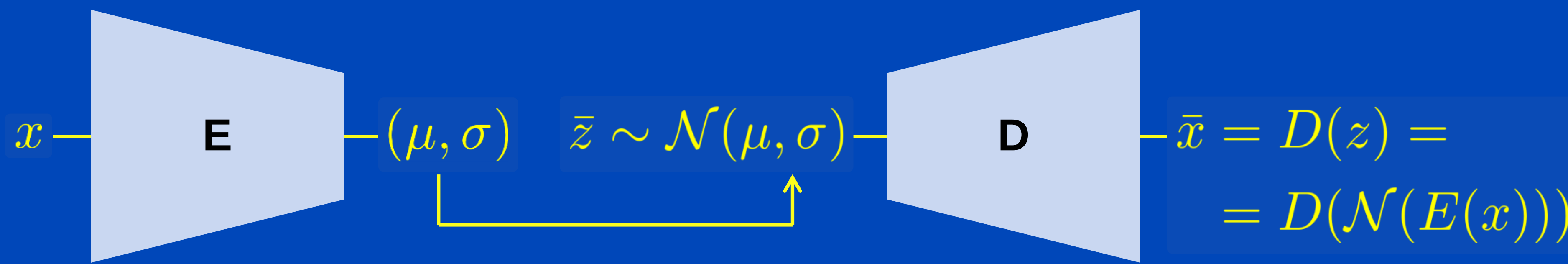


$$\bar{x} = D(z) = \\ = D(E(x))$$

- Examples:
  - Principal component analysis (linear autoencoder), lossless
  - PCA with dimensionality reduction (nonlinear due to clipping), lossy
  - Image compression and decoding, e.g. jpeg, lossy
- Latent space typically not interpretable.

Recap 2: What is a Variational AE (VAE)?

- Make latent space regular.
- Allow to sample in latent space from a given distribution, here: normal distribution.



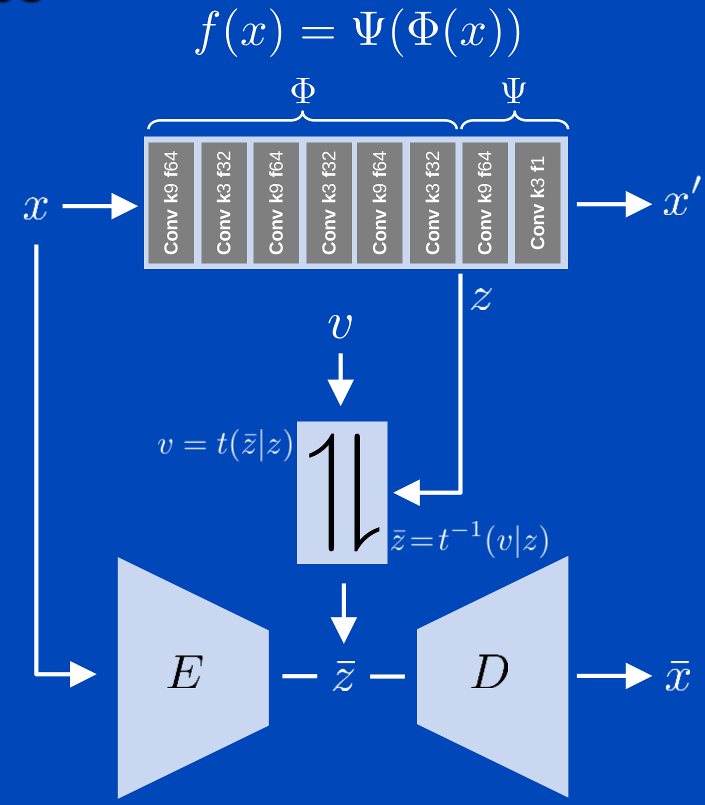
- The VAE is a generative model.
- It allows to generate new data by sampling new values from the normal distribution.

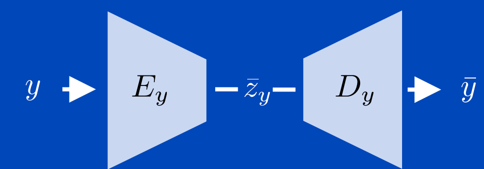
Method

Recovering Invariances

- **Our work is based on Rombach et al.[1]**
- **Given a function or network $f(x) = \Psi(\Phi(x))$ we analyze its internal latent representations $z = \Phi(x)$.**
- **Train a VAE to learn a complete data representation $\bar{z} = E(x)$ of low dose images.**
- **Disentangle information captured in z and invariances v by learning a mapping**
$$v = t(\bar{z}|z), \;\; \mathcal{L}(v) = \mathcal{N}(0,1)$$
- **$t(\cdot|z)$ is realized by a conditional invertible neural network (cINN).**
- **Generate new images varying only by their invariances**

$$\bar{x} = D(t^{-1}(v|z)) \qquad v \sim \mathcal{N}(0,1)$$





Alternative: Use VAE in high dose domain, i.e. VAE_y, to visualize the invariances.

[1]Robin Rombach, Patrick Esser, and Björn Ommer. "Making sense of CNNs: Interpreting deep representations and their invariances with INNs", ECCV 2020.

Method

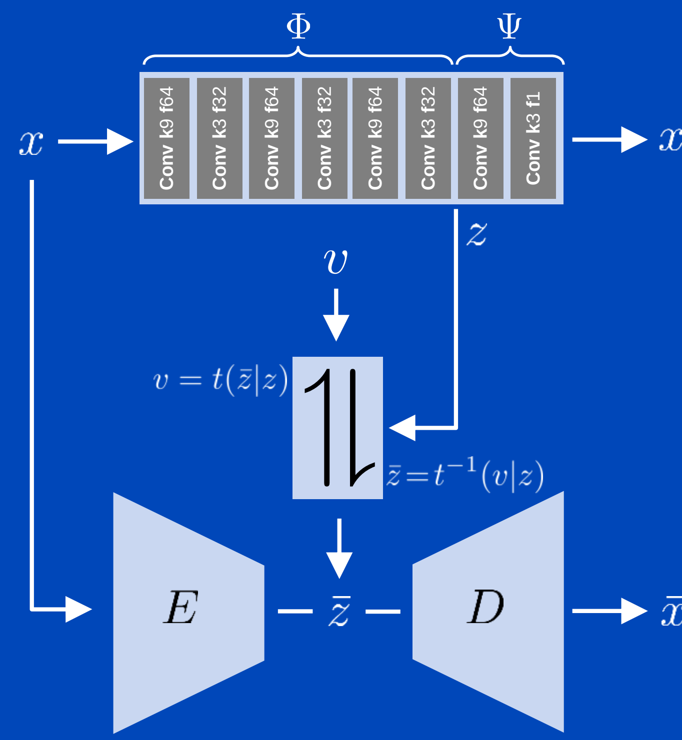
Recovering Invariances

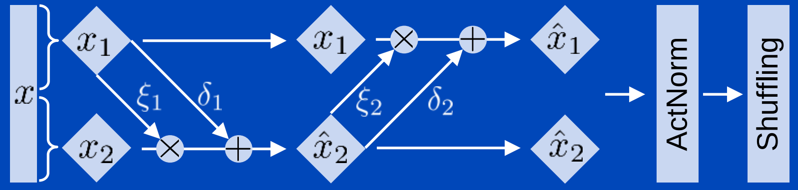
1. **Our work is based on Rombach et al.[1]**
2. **Train denoising methods Chen et al. & Yang et al.**
3. **Train VAE to learn a complete data representation of the low dose images *x*.**
4. **For each denoising method and layer in the network we wish to evaluate, train a cINN to recover the invariances.**
5. **For a given test image, sample 250 invariances *v*, apply the inverse mapping t^{-1} and apply the pretrained decoder *D*.**

$$t^{-1} \text{ maps } \mathcal{N}(0,1) \text{ onto } p(\bar{z}|z).$$

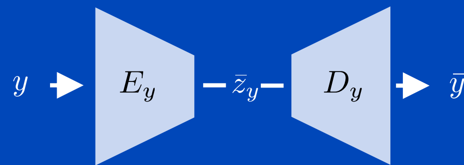
Thus it produces only images that are likely under the training distribution of the AE.

$$f(x) = \Psi(\Phi(x))$$





Building block of INN: Invertible block, ξ_{12} and δ_{12} are CNNs or NNs



Alternative: Use VAE in high dose domain, i.e. VAE_y, to visualize the invariances.

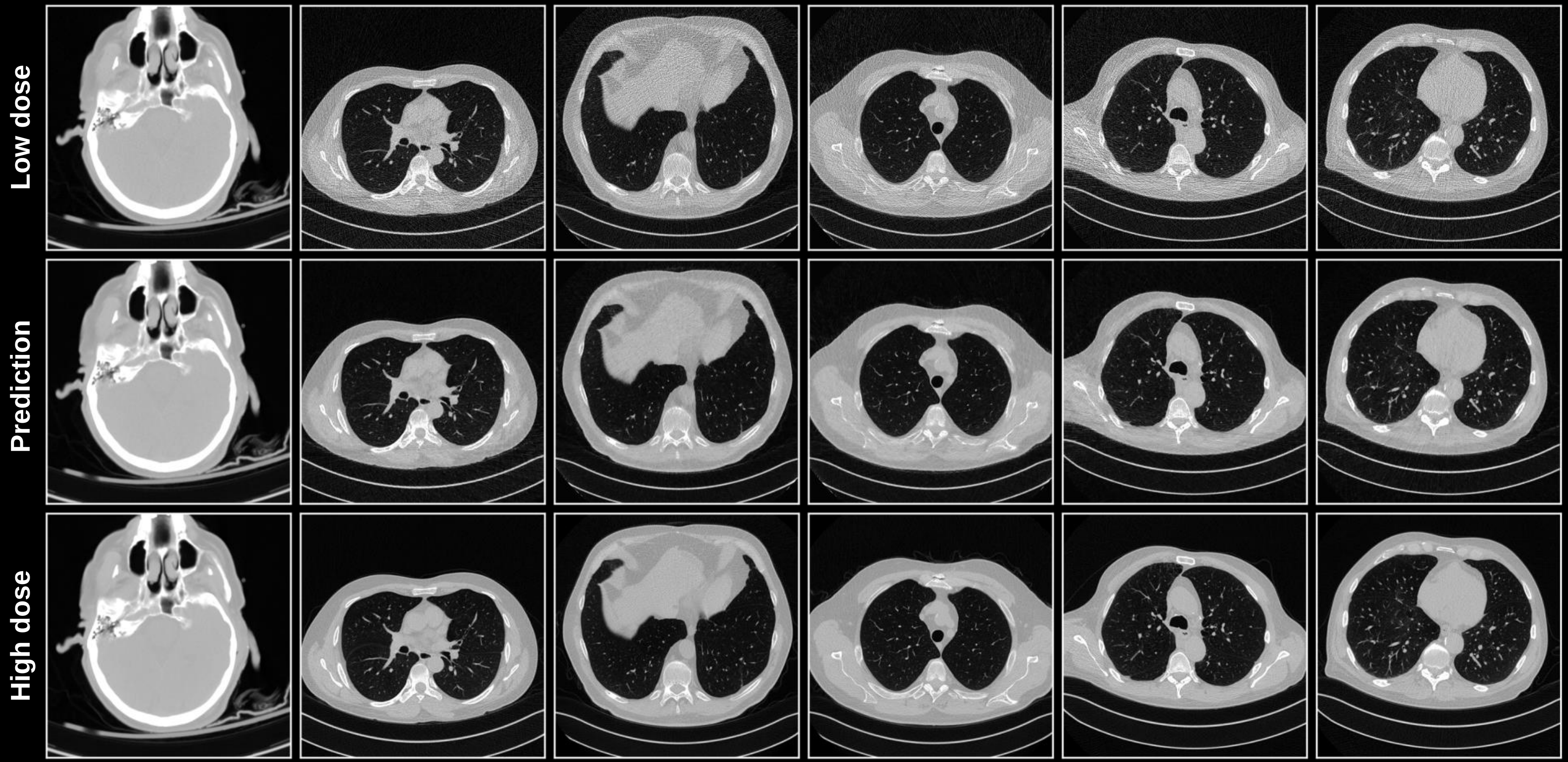
$$x_1 \exp(\xi_2(\hat{x}_2)) + \delta_2(\hat{x}_2) = \hat{x}_1$$
$$x_2 \exp(\xi_1(x_2)) + \delta_1(x_1) = \hat{x}_2$$

[1]Rombach et al. "Making sense of CNNs: Interpreting deep representations and their invariances with INNs", ECCV 2020.

dkfz.

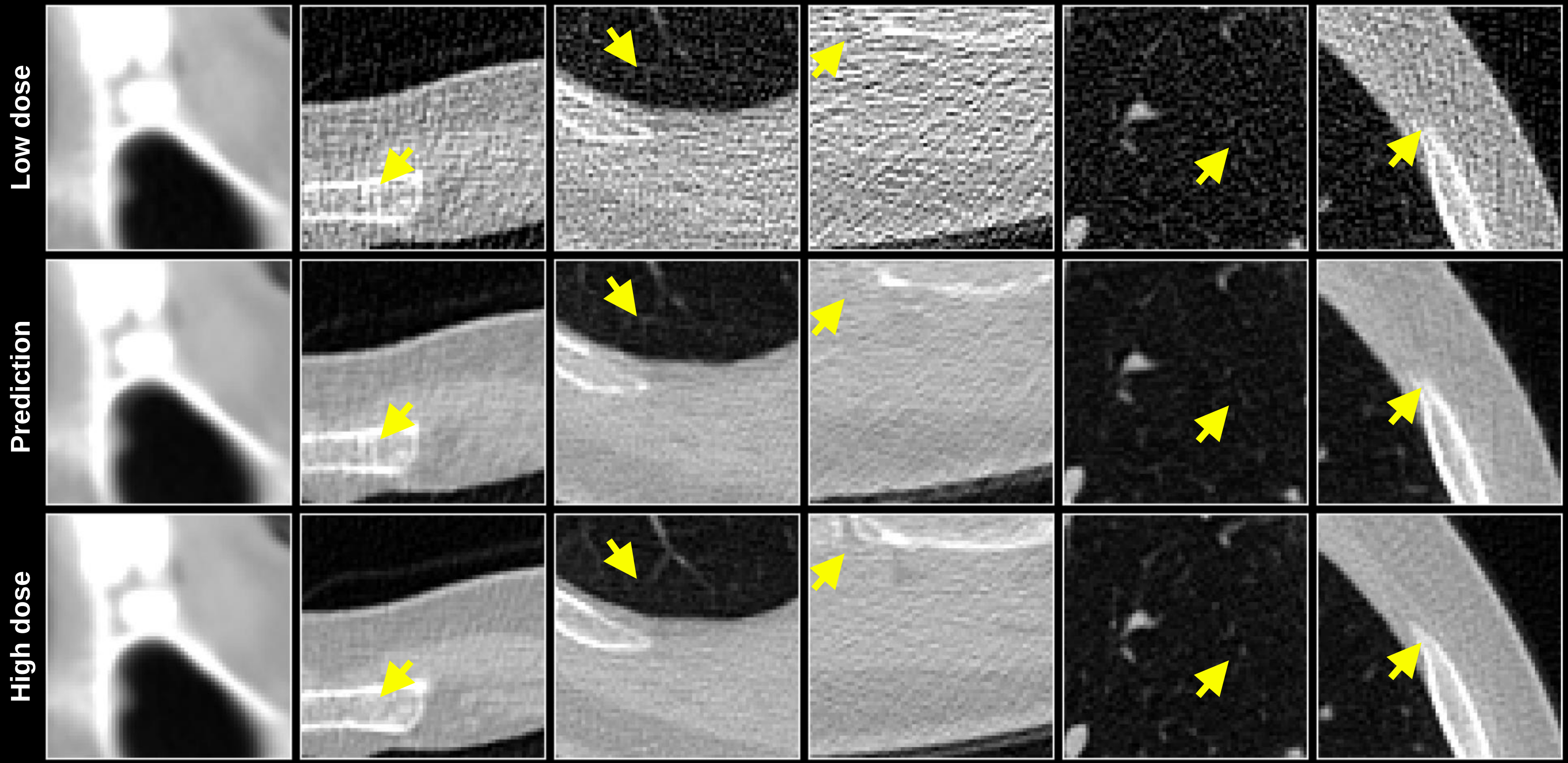
Results

Denoising (Yang et al.) $f = \Psi \circ \Phi$

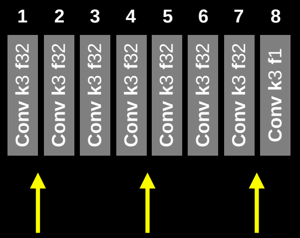


Results

Denoising (Yang et al.) $f = \Psi \circ \Phi$

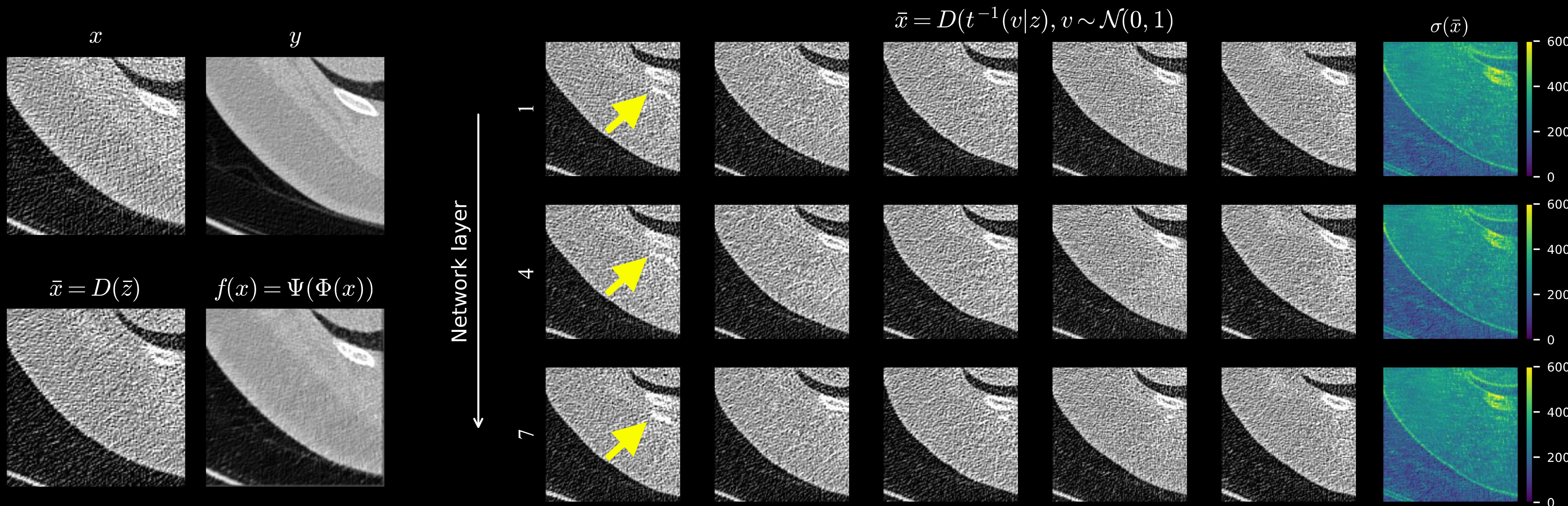


Arrows point at selected differences between prediction and ground truth.



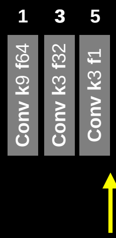
Results

Sampling Invariances in Yang et al.'s Net

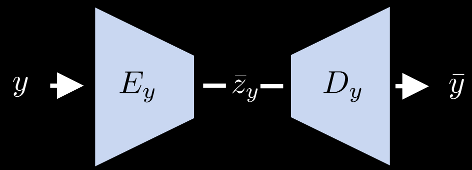


$$\Phi(x) = \Phi(\bar{x}) \quad \forall\ \bar{x}$$

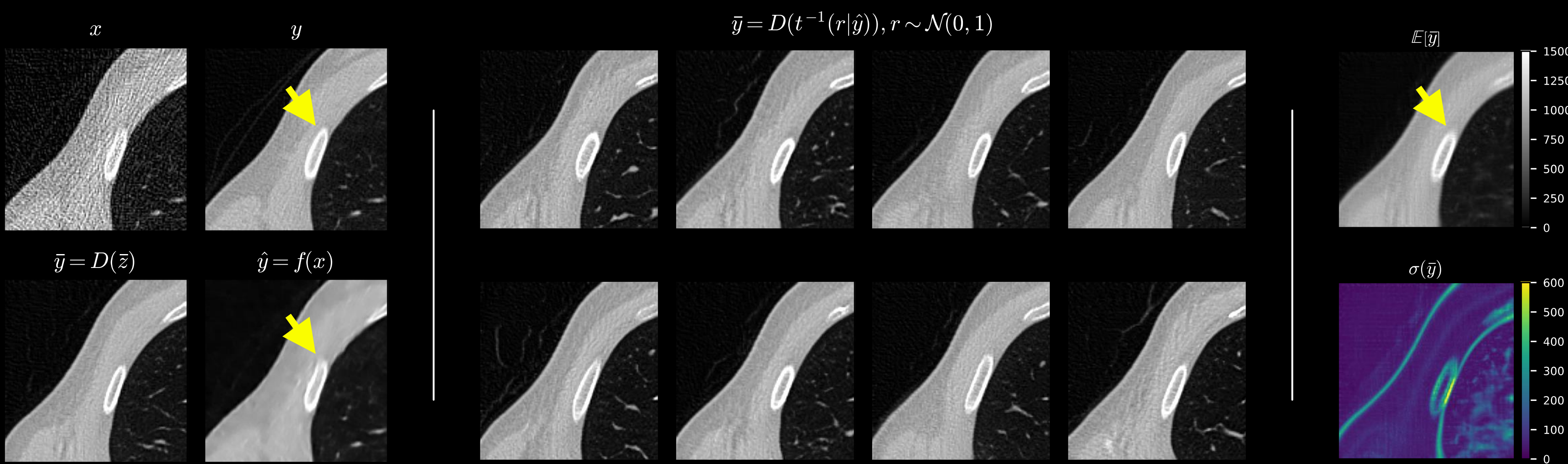
$$x' = f(x) = f(\bar{x}) \quad \forall\ \bar{x}$$



Results



Sampling Invariances in Target Domain in Chen et al.'s Net



$$x' = f(x) = f(\mathsf{R}^{-1}\mathcal{P}\,\mathsf{R}\,\bar{y}) \quad \forall\ \bar{y}$$

forward project
add Poisson noise
reconstruct

Wiretapping after last layer.

dkfz.

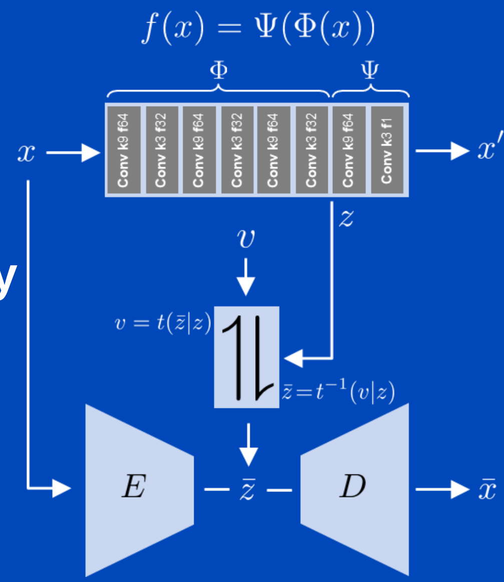
Conclusions & Outlook

Conclusions

- Designed a method to highlight invariances of a given network.
- Algorithm agnostic, not restricted to denosing .
- Architecture agnostic, not restricted to CT.
- Both denoising methods are invariant to some anatomical features to some extent.

Outlook

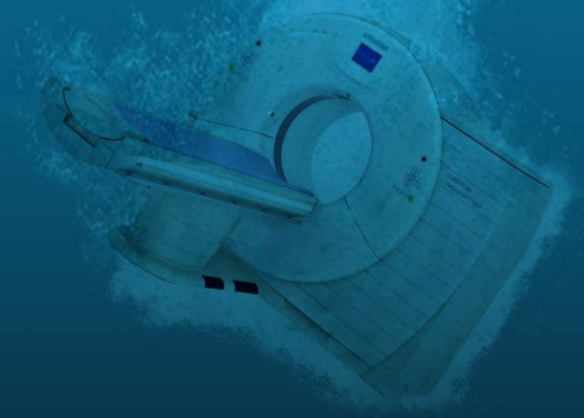
- Improve interpretability by
  - improving the embedding of the VAEs,
  - mapping sampled invariance images to semantically meaningful space (disentangled representations of e.g. tumors).
- One could use the undesired invariances to finetune the denoising methods.



Conclusions on Deep CT

- Machine learning plays and will play a significant role in CT image formation.
- High potential for
  - Artifact correction
  - Noise and dose reduction
  - Real-time dose assessment (also for RT)
  - …
- Care has to be taken
  - Underdetermined acquisition, e.g. sparse view or limited angle CT, require the net to make up information!
  - Nice looking images do not necessarily represent the ground truth.
  - Data consistency layers and variational networks with rawdata access may ensure that the information that is made up is consistent with the measured data.
  - …

Thank You!



Job opportunities through DKFZ’s international PhD or Postdoctoral Fellowship programs (marc.kachelriess@dkfz.de).

Parts of the reconstruction software were provided by RayConStruct® GmbH, Nürnberg, Germany.