

Advances in AI for CT and CBCT Image Formation

Marc Kachelrieß

German Cancer Research Center (DKFZ)

Heidelberg, Germany

www.dkfz.de/ct

RESEARCH ARTICLE

MEDICAL PHYSICS

Benchmarking deep learning-based low-dose CT image denoising algorithms

Elias Eulig^{1,2} | Björn Ommer³ | Marc Kachelrieß^{1,4}

¹Division of X-Ray Imaging and Computed Tomography, German Cancer Research Center (DKFZ), Heidelberg, Germany

²Faculty of Physics and Astronomy, Heidelberg University, Heidelberg, Germany

³CompVis @ LMU Munich, MCML, Munich, Germany

⁴Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany

Abstract

Background: Long-lasting efforts have been made to reduce radiation dose and thus the potential radiation risk to the patient for computed tomography (CT) acquisitions without severe deterioration of image quality. To this end, various techniques have been employed over the years including iterative reconstruction methods and noise reduction algorithms.

Purpose: Recently, deep learning-based methods for noise reduction became increasingly popular and a multitude of papers claim ever improving performance both quantitatively and qualitatively. However the lack of a standardized

LDCT Benchmark

- Algorithms used for our benchmark:

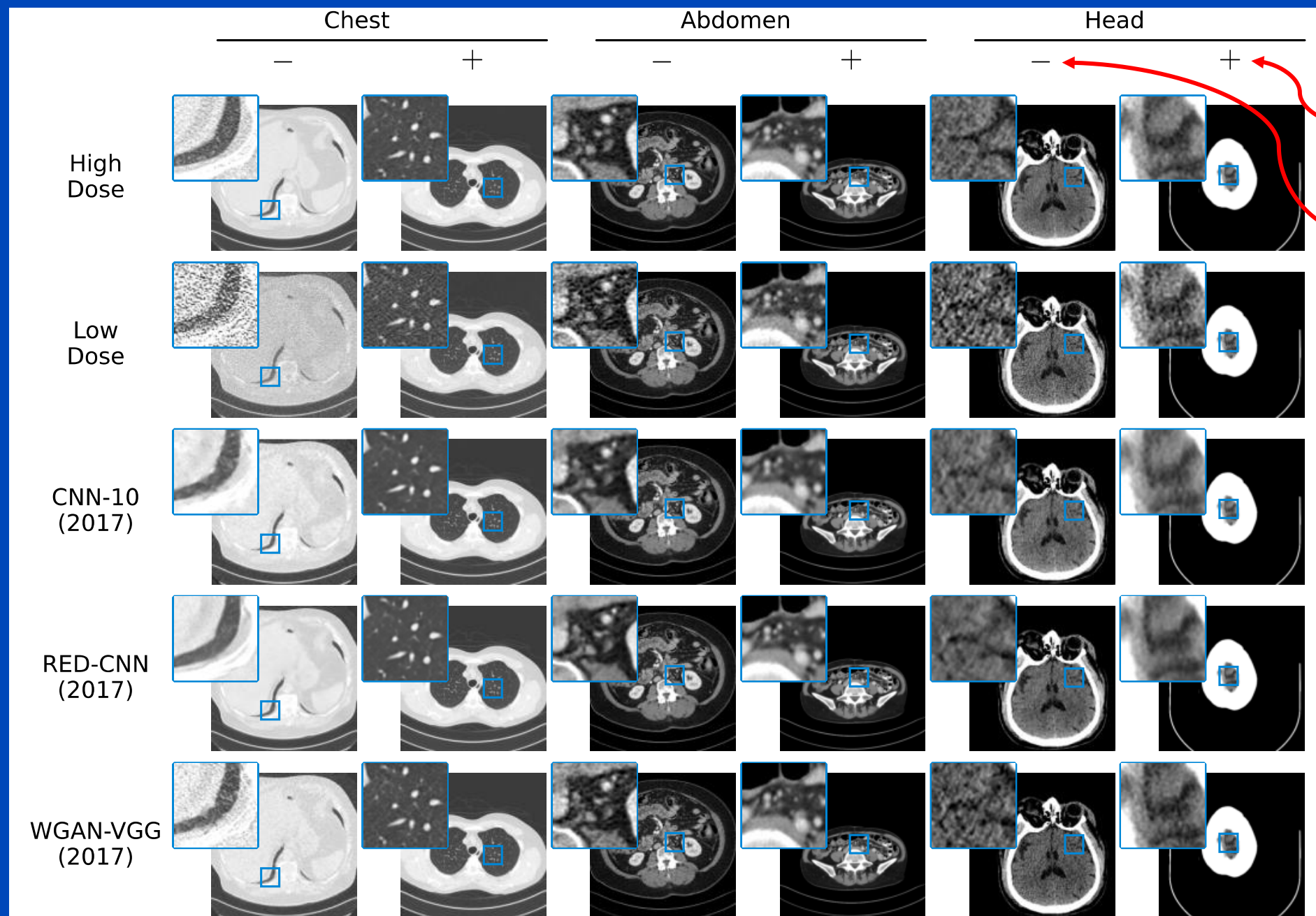
- CNN-10 (2017)
 - RED-CNN (2017)
 - ResNet (2018)
 - WGAN-VGG (2017)
 - QAE (2019)
 - DU-GAN (2021)
 - TransCT (2021)
 - Bilateral (2022)
- Standard CNNs trained with pixelwise losses
- CNNs trained with adversarial losses
- Specialized architectures trained with pixelwise losses

- All tested methods

- do the same hyperparameter optimization
- use the same train/validation set
- were evaluated on the same test set

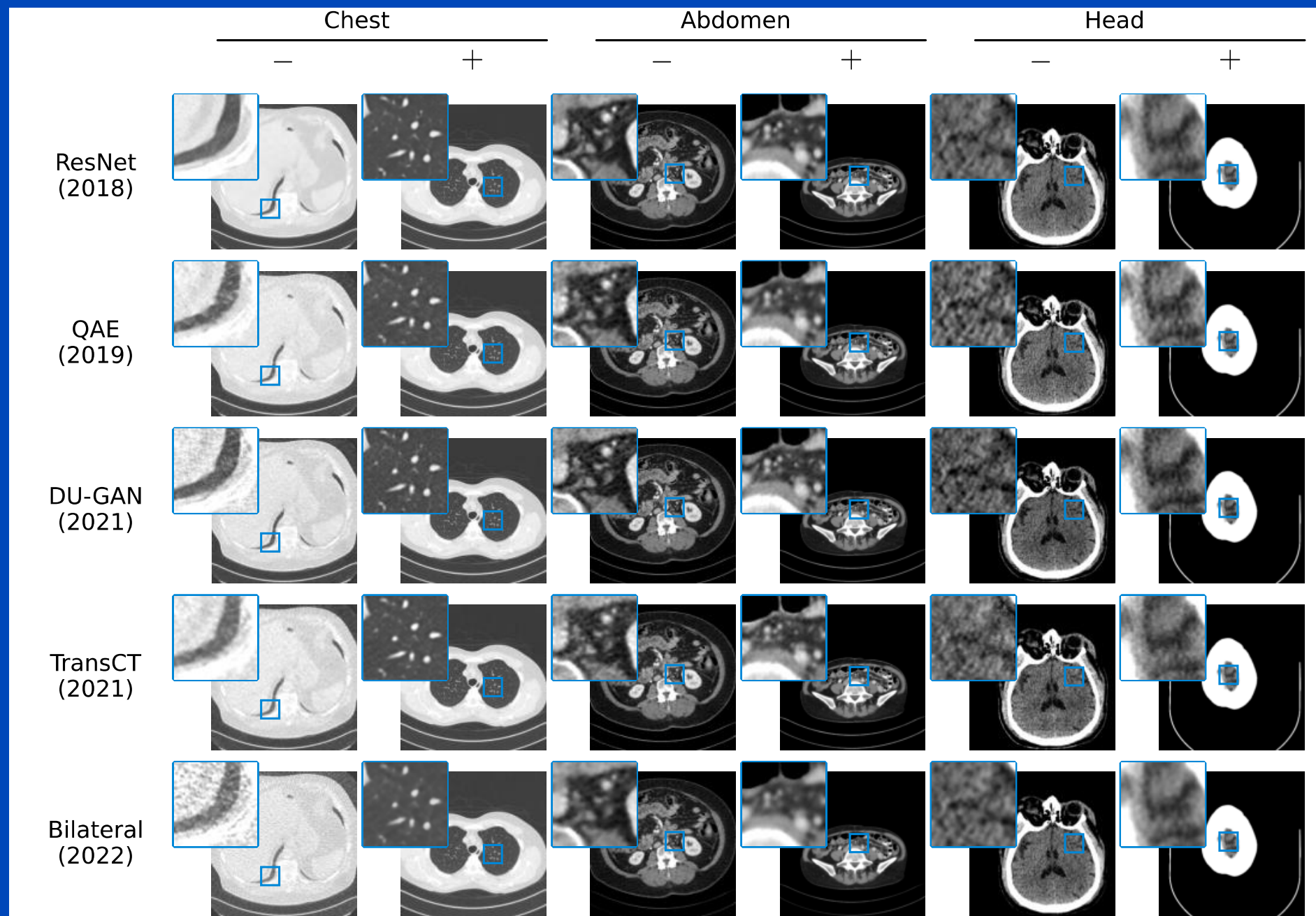


github.com/eeulig/ldct-benchmark



Slice where the average SSIM across all head slices and methods is **highest**.

Slice where the average SSIM across all head slices and methods is **lowest**.



Quantitative Results

PSNR units are decibel (dB)	Head (25% dose)				Chest (10% dose)				Abdomen (25% dose)			
	SSIM	PSNR	VIF	RFS	SSIM	PSNR	VIF	RFS	SSIM	PSNR	VIF	RFS
Low dose scan	26.40	0.55	0.71	0.34	18.77	0.09	0.70	0.84	28.67	0.34	0.75	0.88
CNN-10 (2017)	28.86	0.62	0.94	0.59	27.71	0.19	0.80	0.90	32.39	0.45	0.88	0.90
RED-CNN (2017)	30.41	0.69	0.95	0.61	28.36	0.22	0.76	0.90	33.22	0.49	0.80	0.90
WGAN-VGG (2017)	25.36	0.53	0.86	0.51	25.54	0.15	0.98	0.88	30.51	0.38	0.92	0.88
ResNet (2018)	29.64	0.67	0.91	0.61	28.42	0.22	0.75	0.90	33.15	0.49	0.79	0.90
QAE (2019)	28.51	0.59	0.95	0.58	27.62	0.19	0.83	0.89	32.02	0.42	0.96	0.90
DU-GAN (2021)	28.76	0.62	0.94	0.57	26.68	0.17	0.96	0.89	32.13	0.43	0.97	0.90
TransCT (2021)	24.65	0.44	0.88	0.56	26.99	0.17	0.83	0.88	30.53	0.37	0.92	0.85
Bilateral (2022)	26.60	0.50	0.87	0.55	25.59	0.16	0.64	0.86	27.13	0.36	0.87	0.87

Green numbers indicate that a method is significantly better than the previously published best method.
 Orange numbers indicate that it is significantly worse.

If you are interested to benchmark your noise reduction algorithm:
<https://github.com/eeulig/ldct-benchmark>

Noise-Augmented Deep Denoising of CT (NADD)

Gernot Kristof, Achim Byl, Elias Eulig, and Marc Kachelrieß

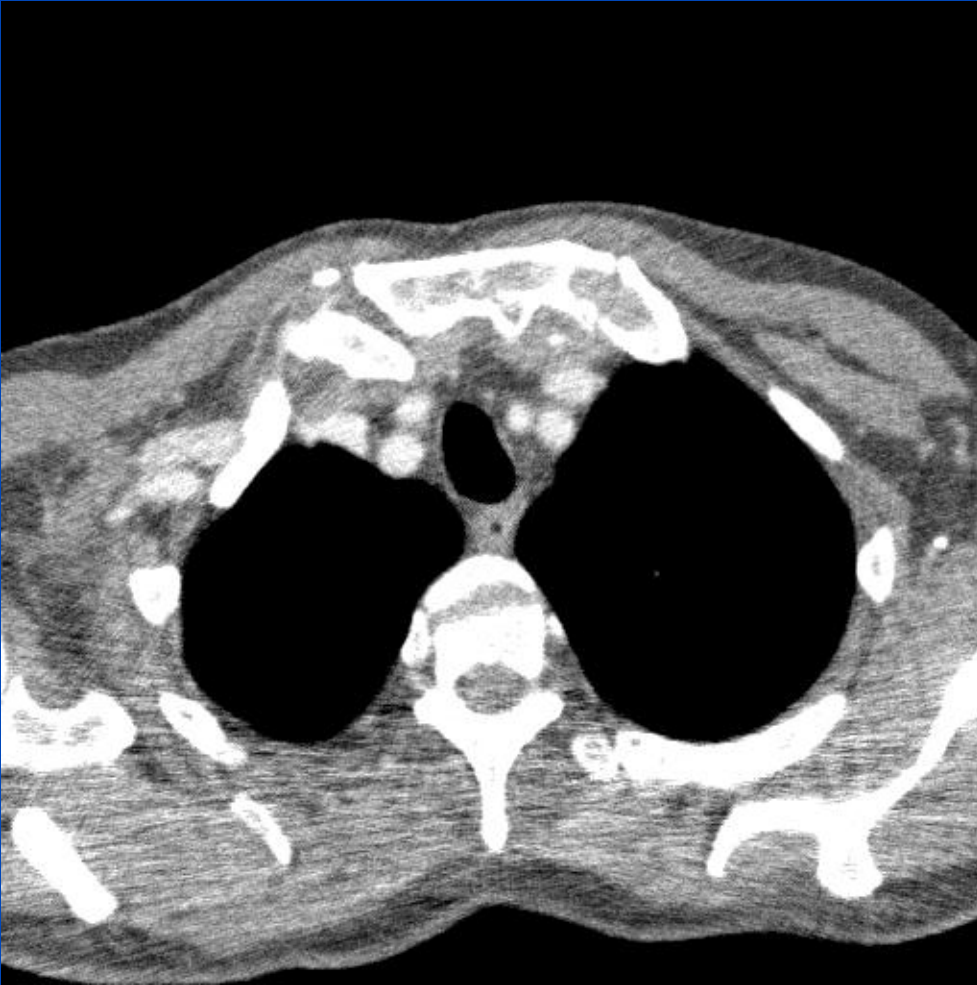
Abstract—Denoising low dose CT images can have great advantages for the aim of minimizing patient risk, as it can help lower the effective dose to the patient, while providing constant image quality. Conventional deep denoising algorithms cannot handle the correlation between neighboring pixels or voxels, because the noise structure in CT is a resultant of the global attenuation properties of the patient and because the receptive field of denoising approaches is rather small. In this work additional noise realizations were generated, reconstructed, and used as additional input into a denoising network to guide the denoising process. The network was compared to a similar network, without additional noise augmentation. It was shown, that the noise-augmented deep denoising network outperformed the conventional denoising network. Due to the additional in-

to 2D images but will be extended to 3D volumes in the near future.

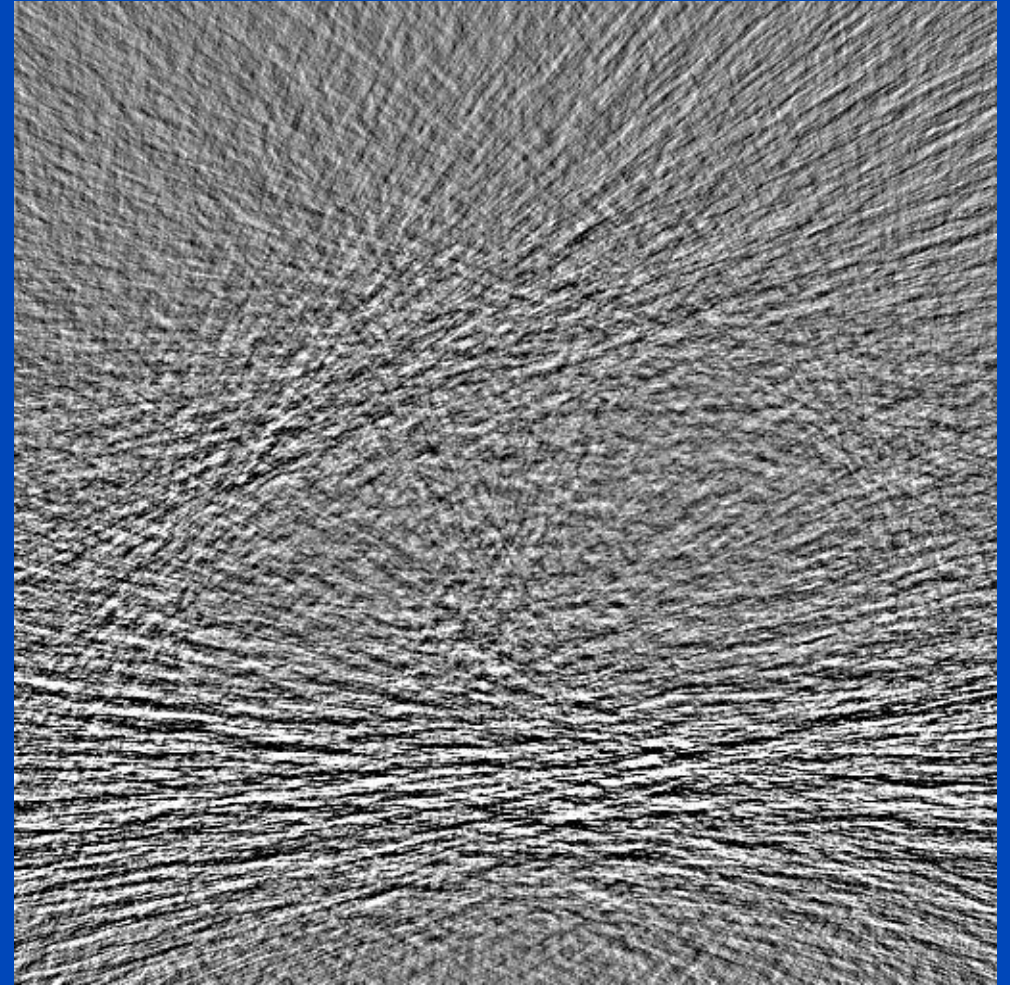
II. METHOD

The idea of this project is to generate, by noise injection into the rawdata followed by image reconstruction, a multitude of new noise realizations $h_c(\mathbf{r})$ of the CT image $g(\mathbf{r})$ that is to be denoised. Here we use $C = 10$ noise realizations. These are input to the NADD network, together with the original CT image that is to be denoised. With these ten additional input channels NADD can learn about the noise correlation

Low Dose and Noise Only Images

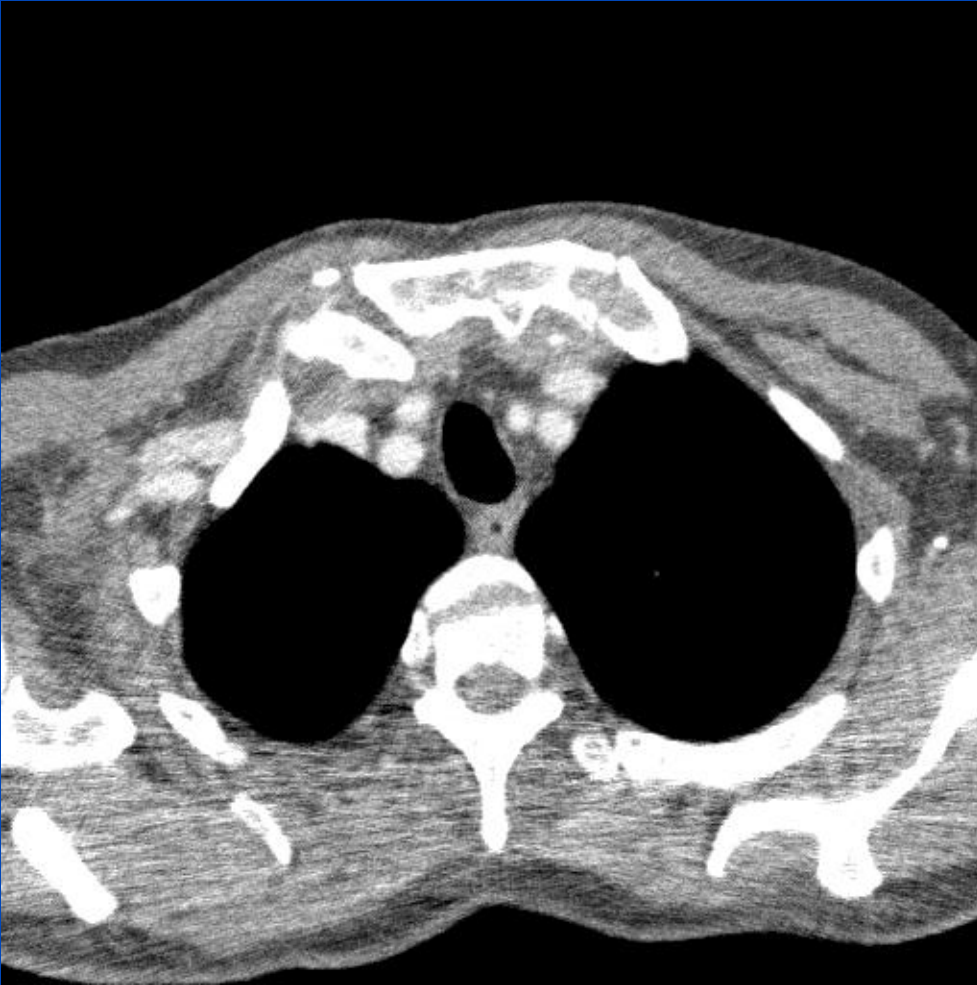


$C = 0 \text{ HU}$, $W = 600 \text{ HU}$

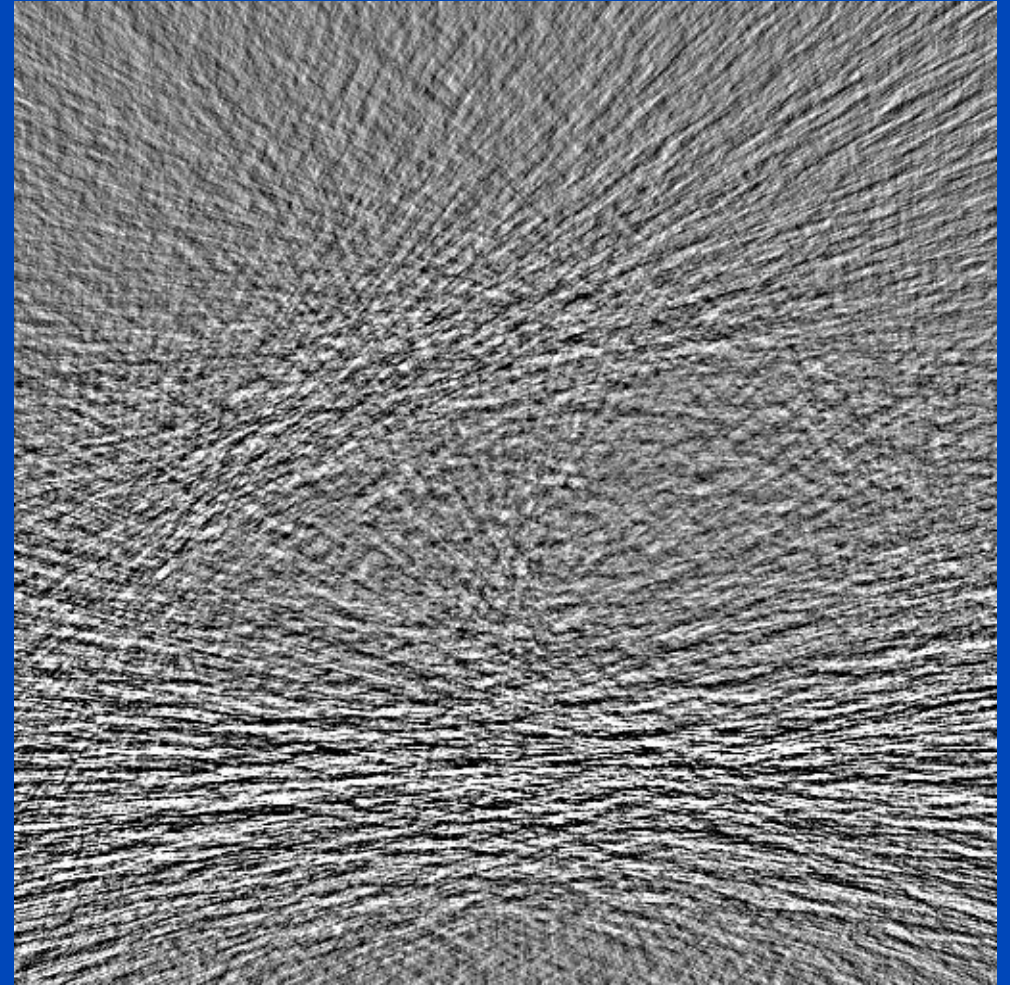


$C = 0 \text{ HU}$, $W = 150 \text{ HU}$

Low Dose and Noise Only Images

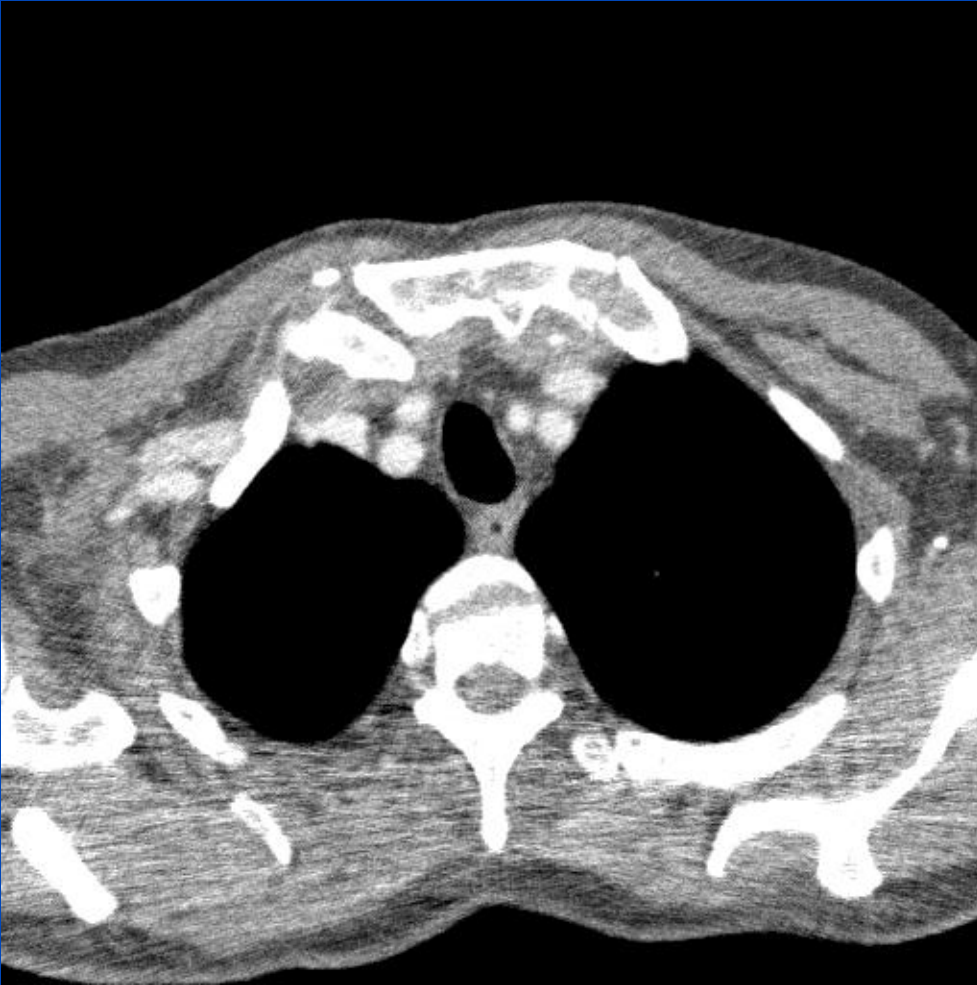


$C = 0 \text{ HU}$, $W = 600 \text{ HU}$

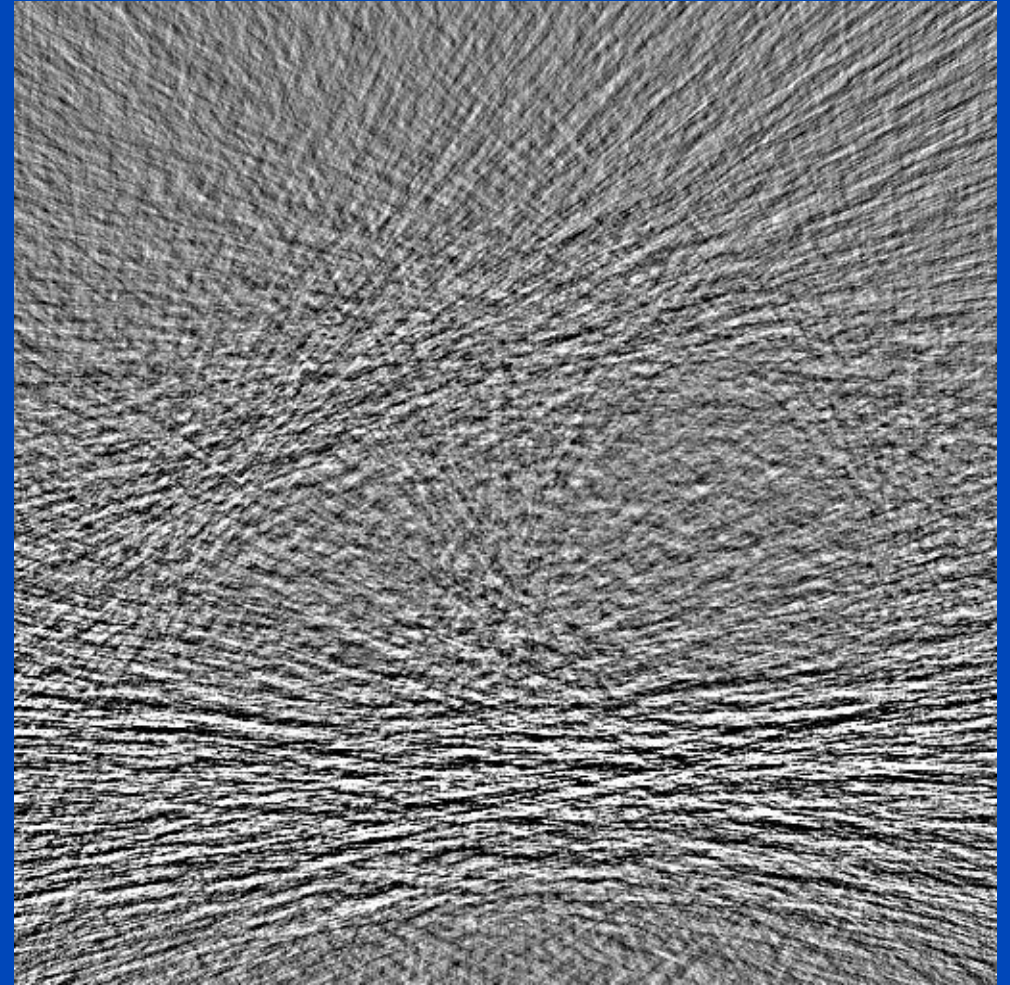


$C = 0 \text{ HU}$, $W = 150 \text{ HU}$

Low Dose and Noise Only Images



$C = 0 \text{ HU}, W = 600 \text{ HU}$



$C = 0 \text{ HU}, W = 150 \text{ HU}$

NADD: Background, Aim and Idea

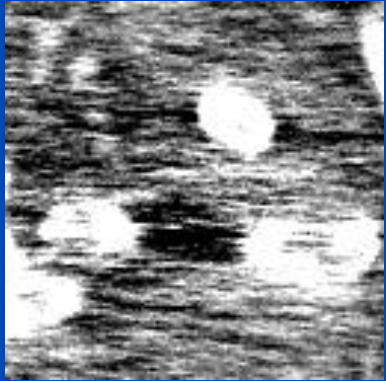
- **Background:** CT noise is strongly correlated between pixels. The correlation depends on the patient's attenuation properties and on the tube current curve that was used to generate the images.
- **Aim:** To find out whether noise reduction networks in CT benefit from seeing more than one noise realization.
- **Idea:** Generate, by rawdata noise injection and reconstruction, several noise realizations and provide them to existing denoising networks in addition to the noisy image.

NADD = 1 measurement + 10 simulated noise realizations

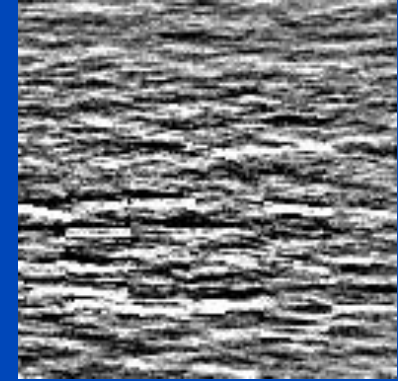
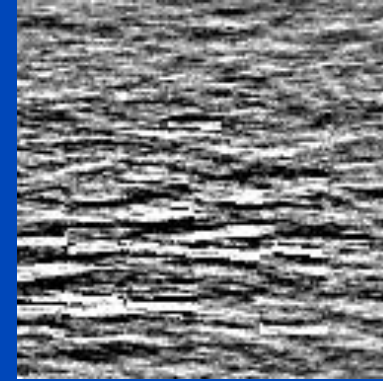
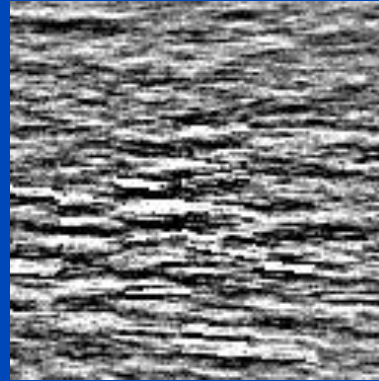
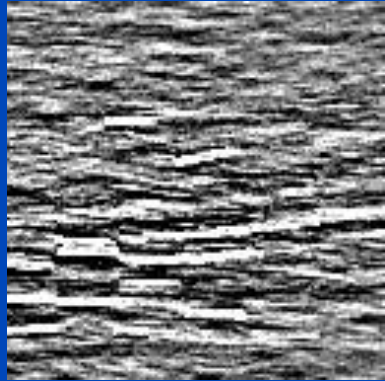
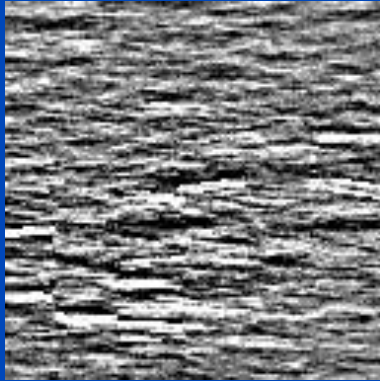
Low Dose, Std Dose and all 10 Noise Only Realizations

Network input

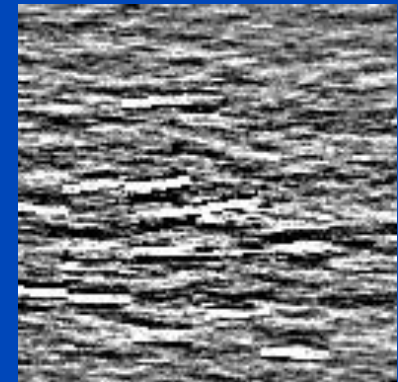
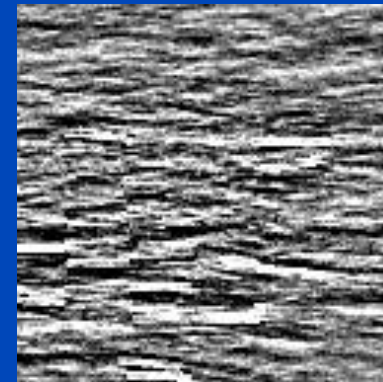
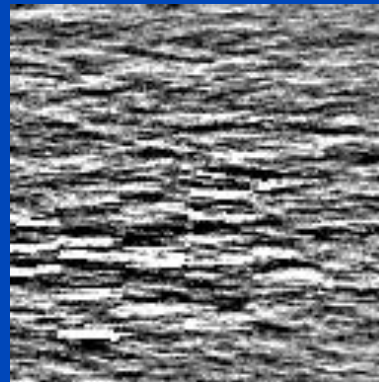
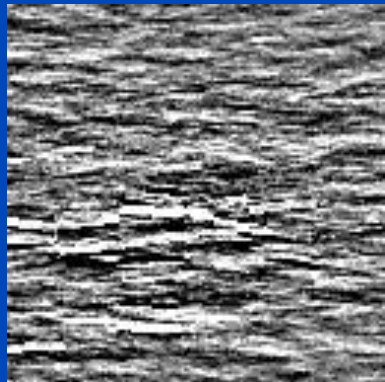
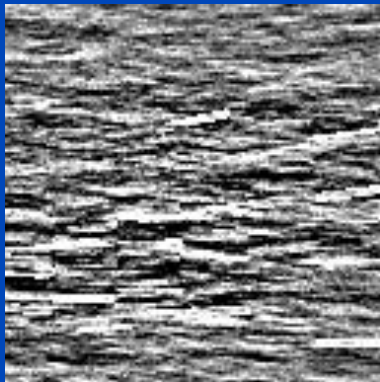
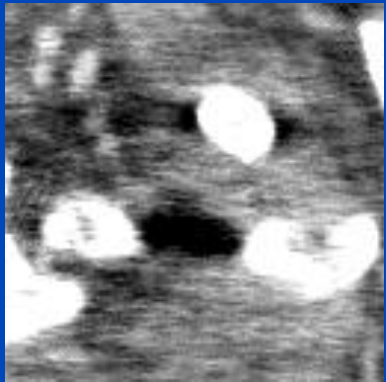
Low Dose



+10 simulated noise realizations (simulation must be done in rawdata domain)



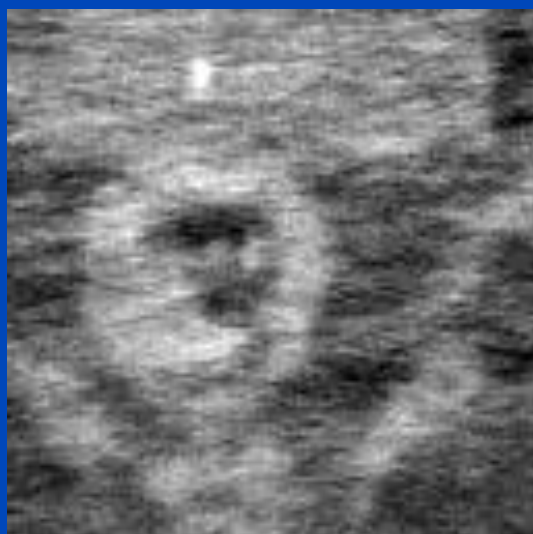
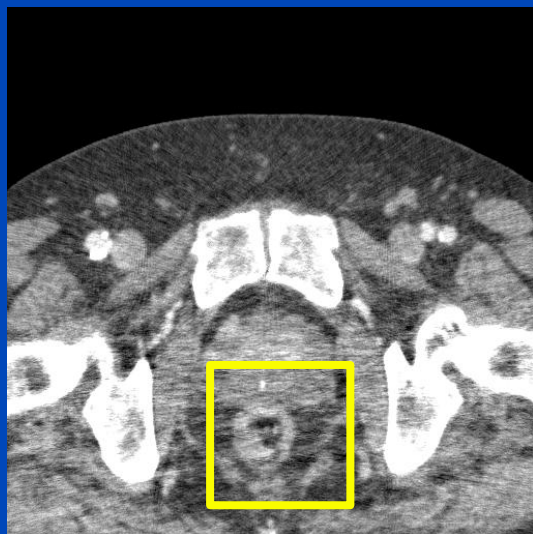
Std Dose



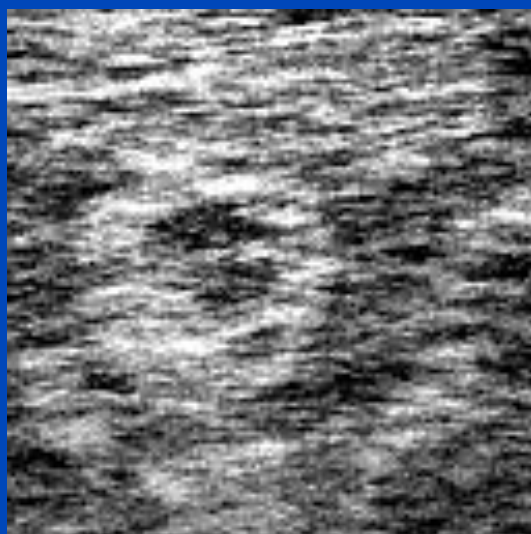
CT Images: $C = 0$ HU, $W = 600$ HU. Noise only images: $C = 0$ HU, $W = 550$ HU.

Results

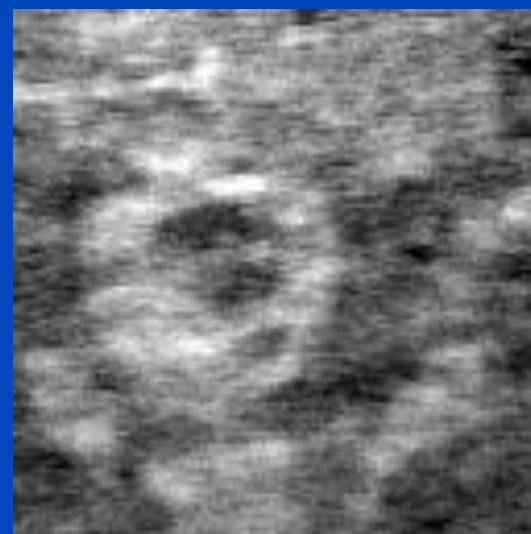
Std Dose



Low Dose



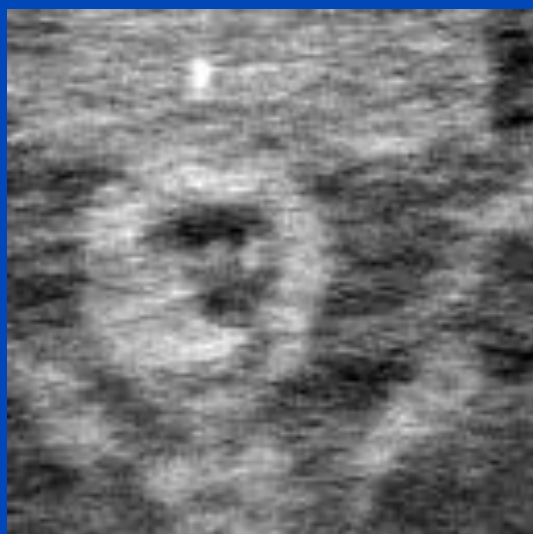
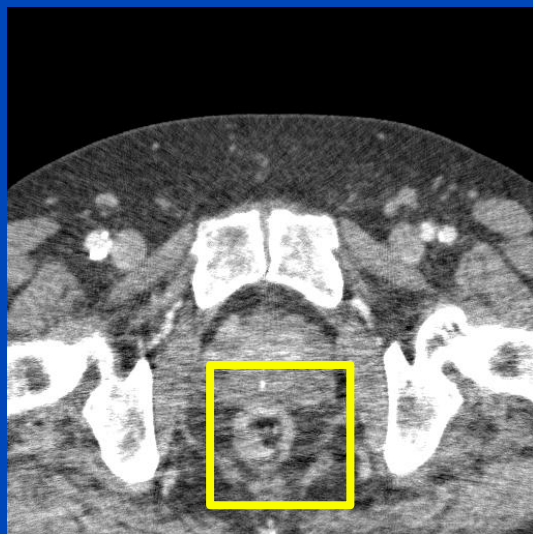
WGAN



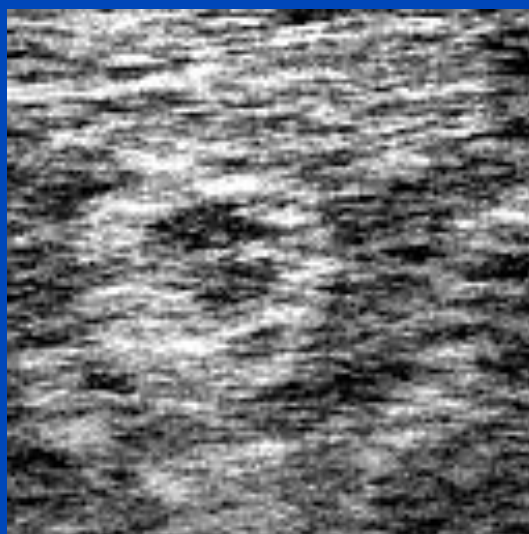
$C = 0 \text{ HU}$, $W = 500 \text{ HU}$

Results

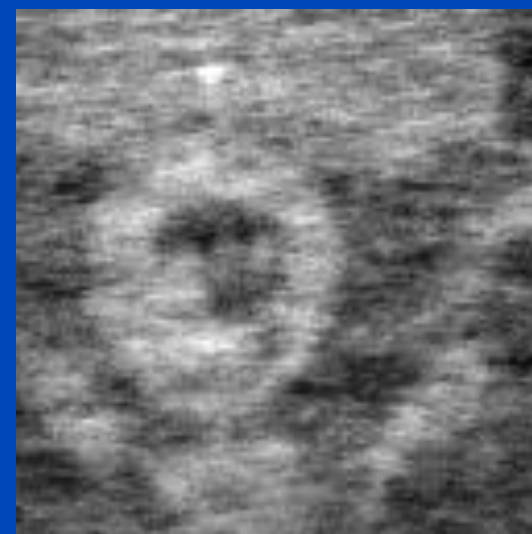
Std Dose



Low Dose



WGAN₊₁₀



C = 0 HU, W = 500 HU

Low Dose



CNN10



ResNet



WGAN



Std Dose



CNN10₊₁₀



ResNet₊₁₀

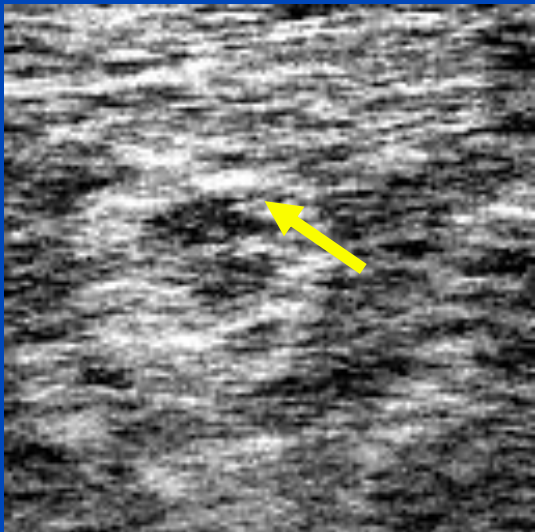


WGAN₊₁₀

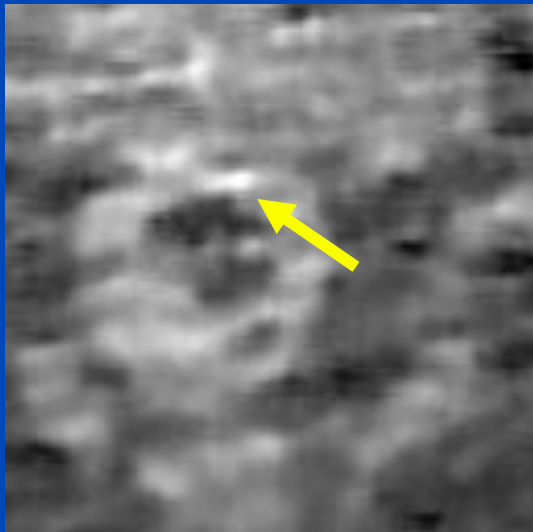


C = 0 HU, W = 500 HU

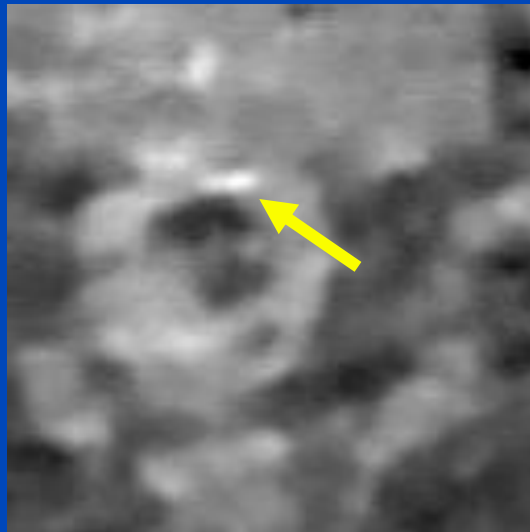
Low Dose



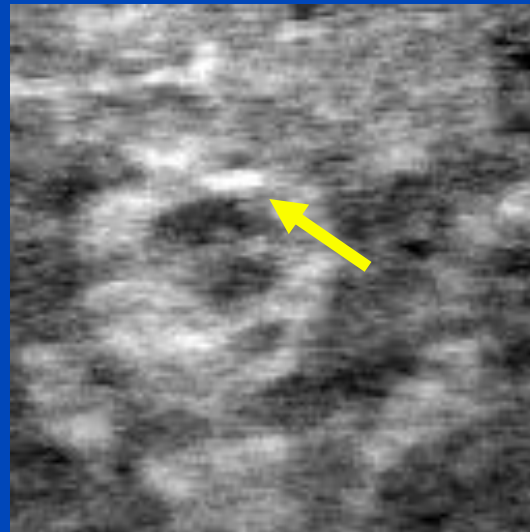
CNN10



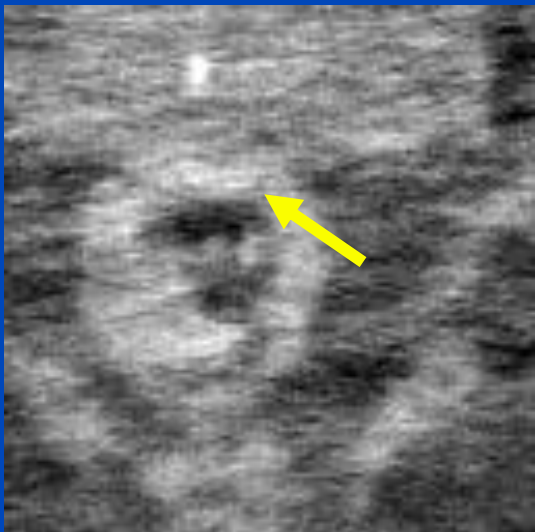
ResNet



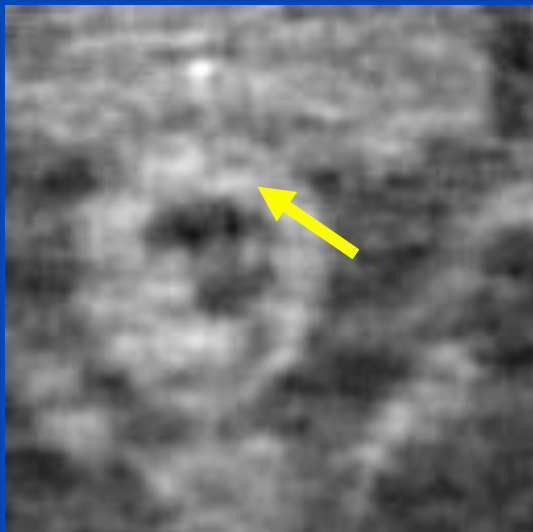
WGAN



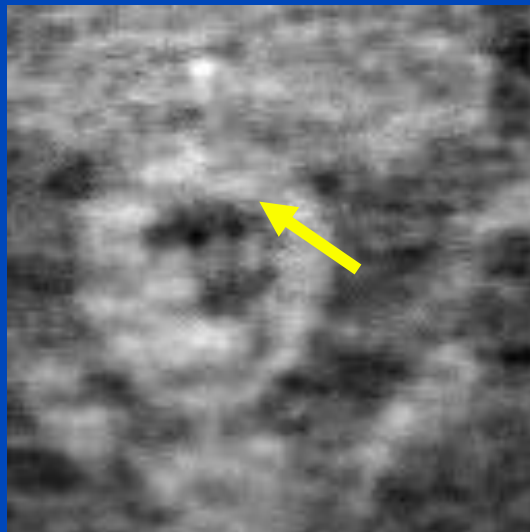
Std Dose



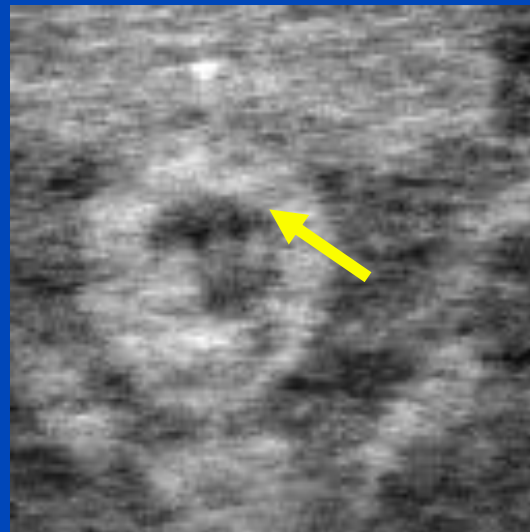
CNN10₊₁₀



ResNet₊₁₀

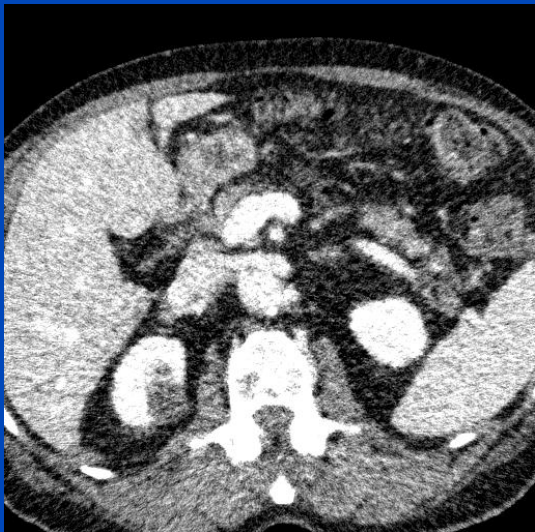


WGAN₊₁₀



C = 0 HU, W = 500 HU

Low Dose



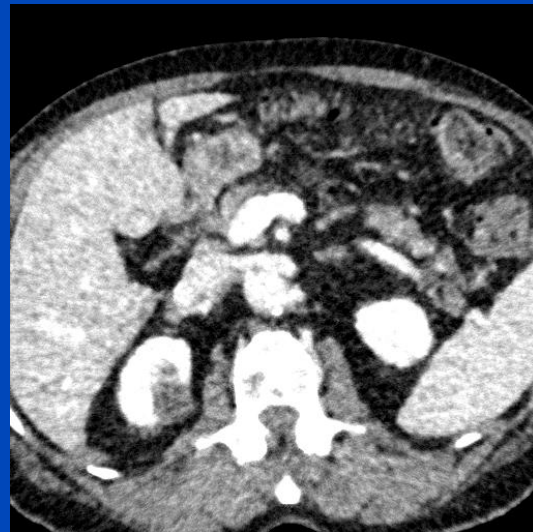
CNN10



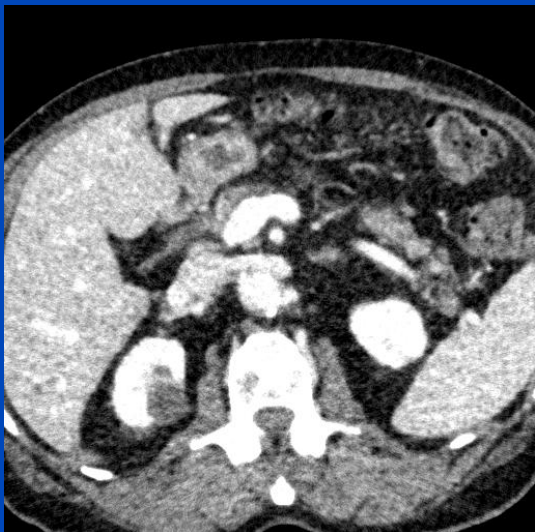
ResNet



WGAN



Std Dose



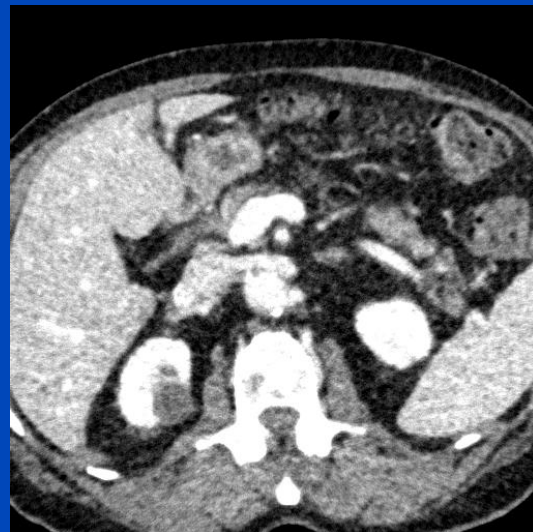
CNN10₊₁₀



ResNet₊₁₀

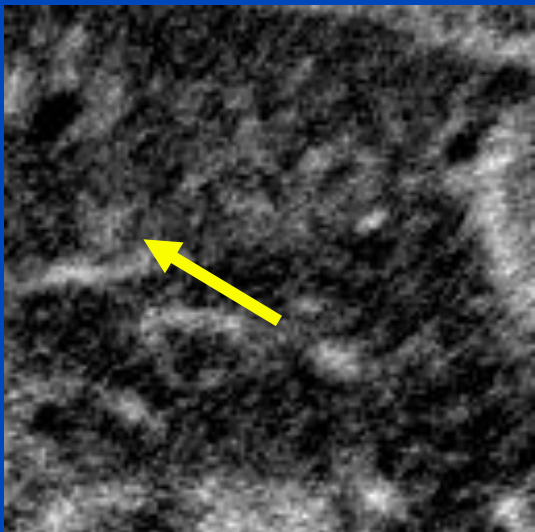


WGAN₊₁₀

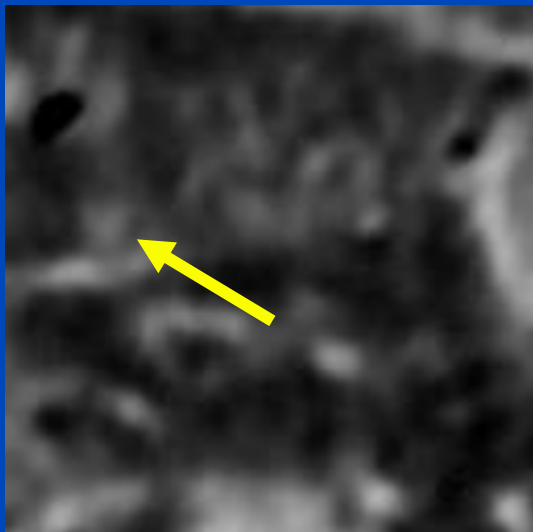


$C = 40 \text{ HU}, W = 400 \text{ HU}$

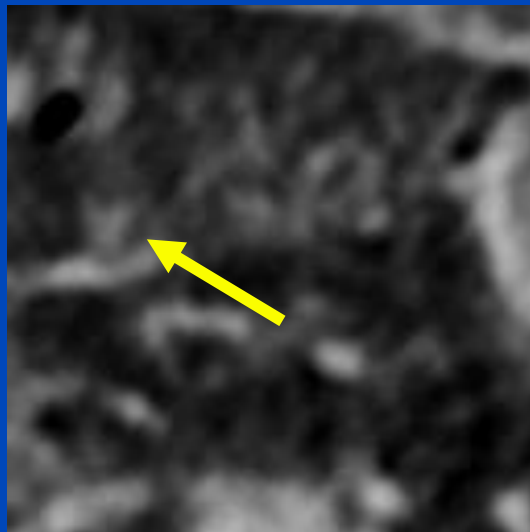
Low Dose



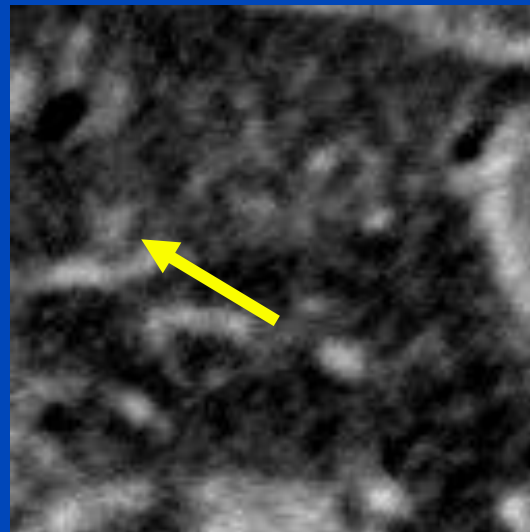
CNN10



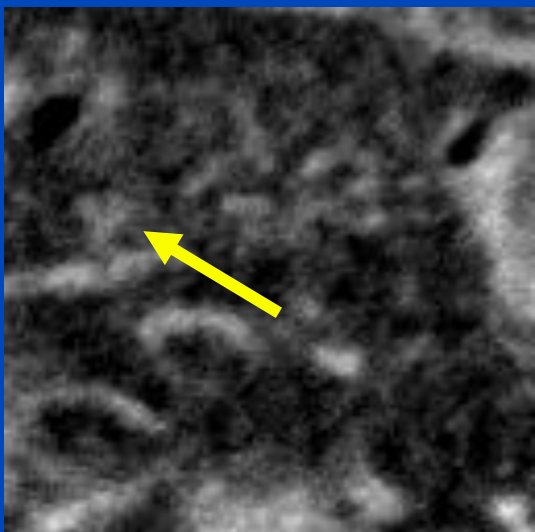
ResNet



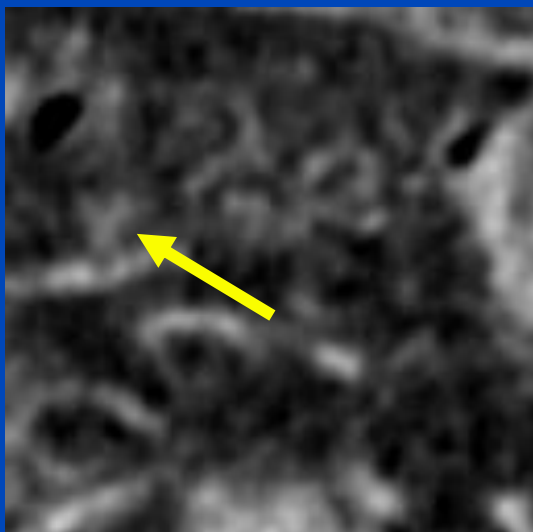
WGAN



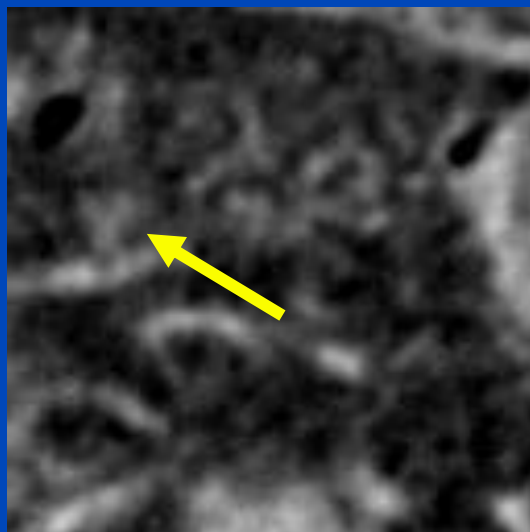
Std Dose



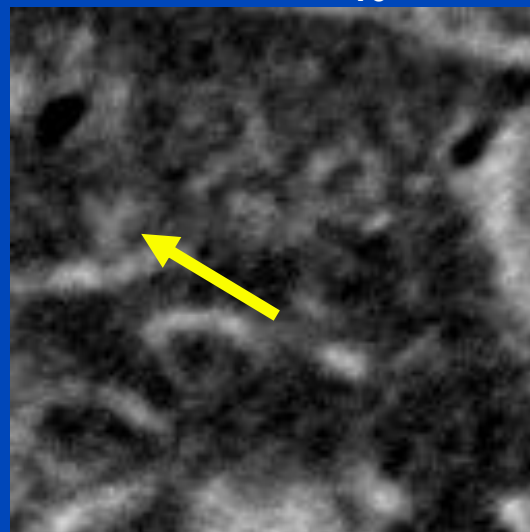
CNN10₊₁₀



ResNet₊₁₀

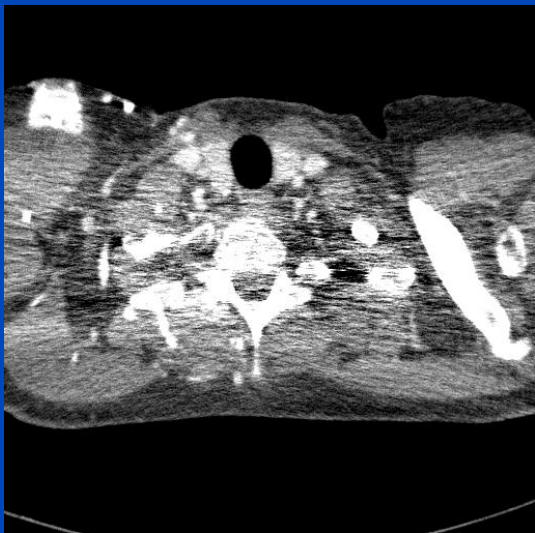


WGAN₊₁₀

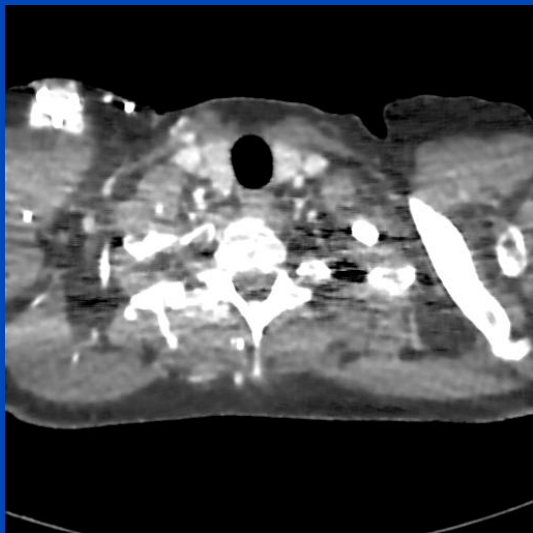


$C = 40 \text{ HU}$, $W = 400 \text{ HU}$

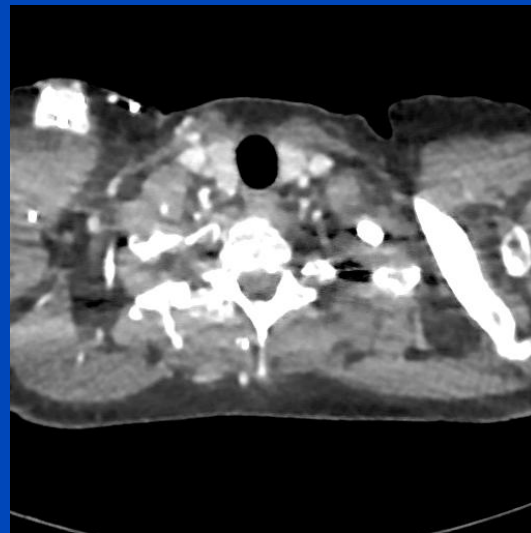
Low Dose



CNN10



ResNet



WGAN



Std Dose



CNN10₊₁₀



ResNet₊₁₀

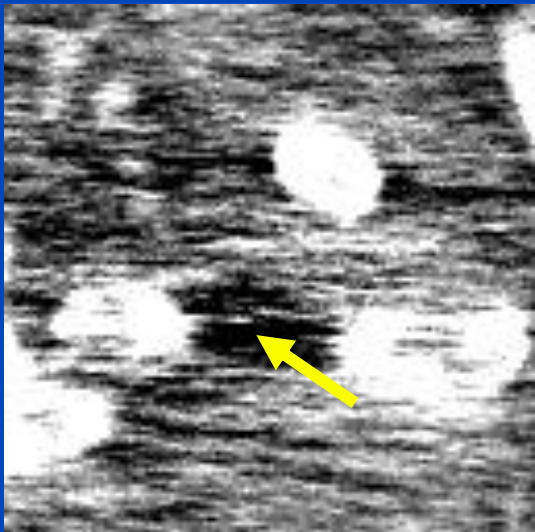


WGAN₊₁₀

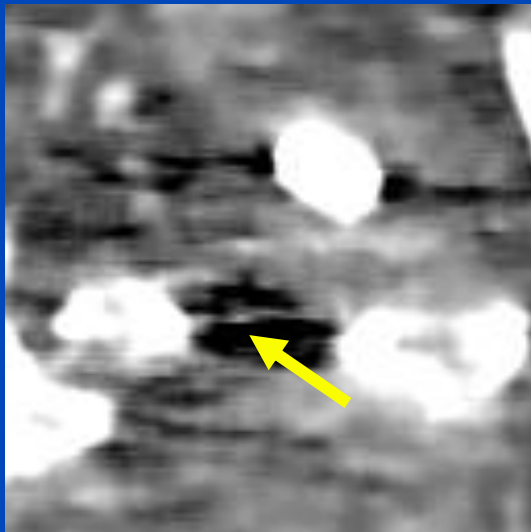


C = 0 HU, W = 600 HU

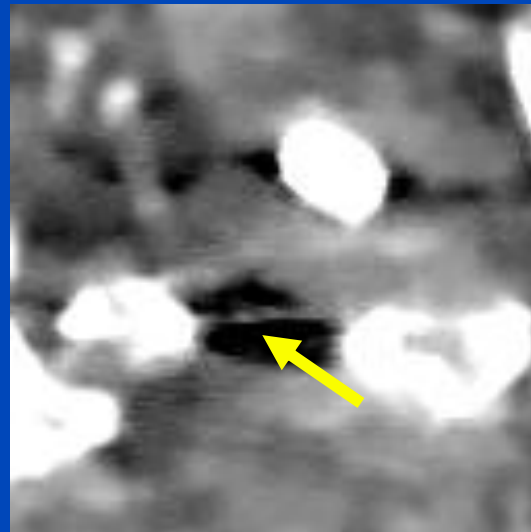
Low Dose



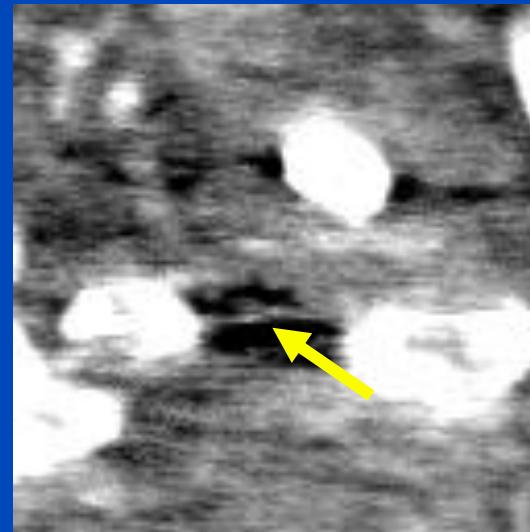
CNN10



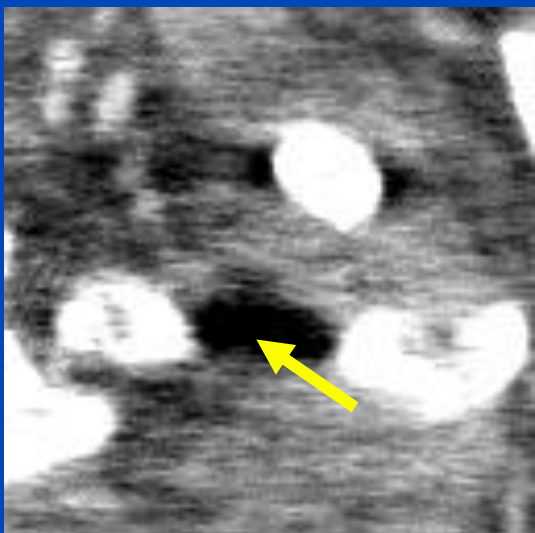
ResNet



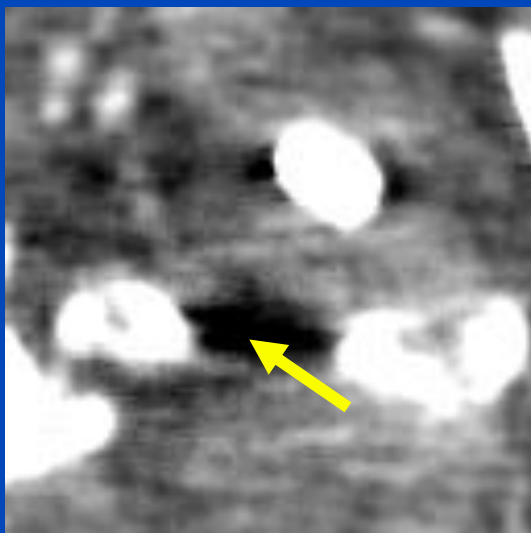
WGAN



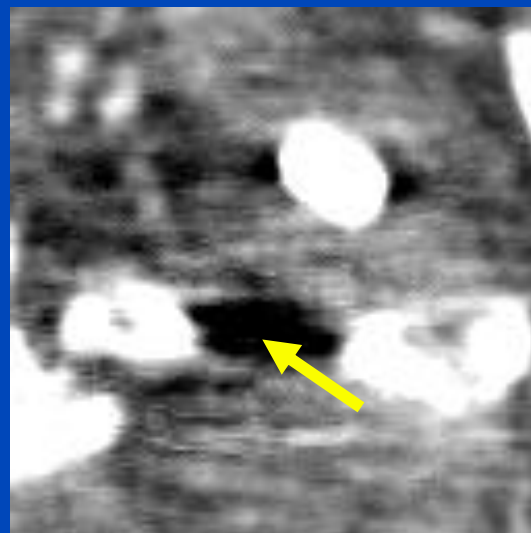
Std Dose



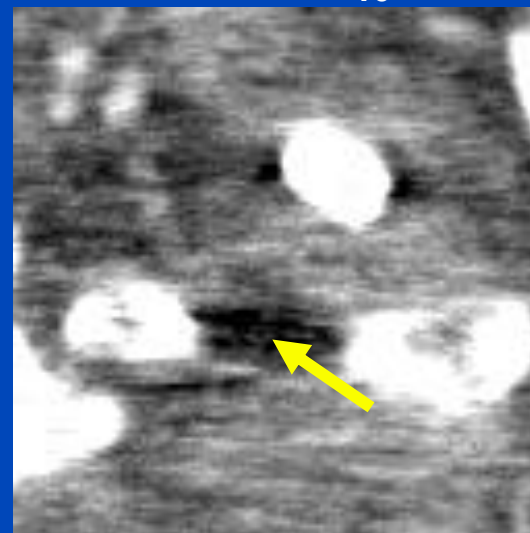
CNN10₊₁₀



ResNet₊₁₀



WGAN₊₁₀



C = 0 HU, W = 600 HU

RESEARCH ARTICLE

MEDICAL PHYSICS

Deep learning-based cone-beam CT motion compensation with single-view temporal resolution

Joscha Maier¹ | Stefan Sawall^{1,2} | Marcel Arheit³ | Pascal Paysan³ |
Marc Kachelrieß^{1,2}

¹Division of X-Ray Imaging and Computed Tomography, German Cancer Research Center (DKFZ), Heidelberg, Germany

²Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany

³Varian Medical Systems Imaging Laboratory, GmbH, Baden-Daettwil, Switzerland

Correspondence

Marc Kachelrieß, Division of X-Ray Imaging and Computed Tomography, German Cancer Research Center (DKFZ), Heidelberg, Germany.
Email: marc.kachelriess@dkfz.de

Abstract

Background: Cone-beam CT (CBCT) scans that are affected by motion often require motion compensation to reduce artifacts or to reconstruct 4D (3D+time) representations of the patient. To do so, most existing strategies rely on some sort of gating strategy that sorts the acquired projections into motion bins. Subsequently, these bins can be reconstructed individually before further post-processing may be applied to improve image quality. While this concept is useful for periodic motion patterns, it fails in case of non-periodic motion as observed, for example, in irregularly breathing patients.

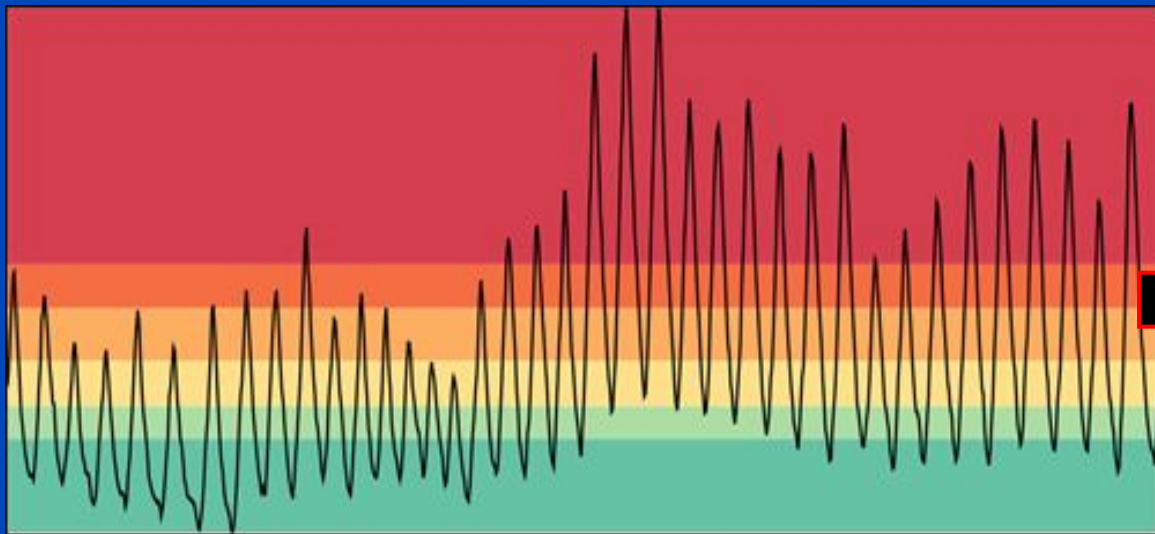
Purpose: To address this issue and to increase temporal resolution, we propose the deep single angle-based motion compensation (SAMoCo).

Methods: To avoid gating, and therefore its downsides, the deep SAMoCo trains

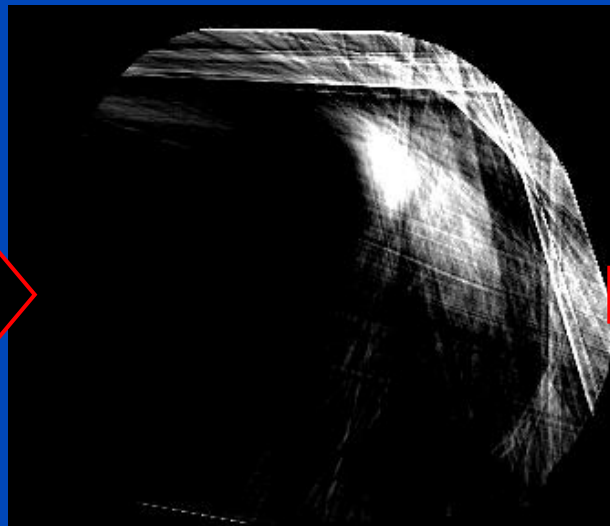
Yet Unsolved Problems

Gating and gating-based motion compensation (MoCo)

- require gating signal,
- assume periodic motion,
- have low temporal resolution,
- fail on irregular breathing:



Patient with irregular breathing pattern:

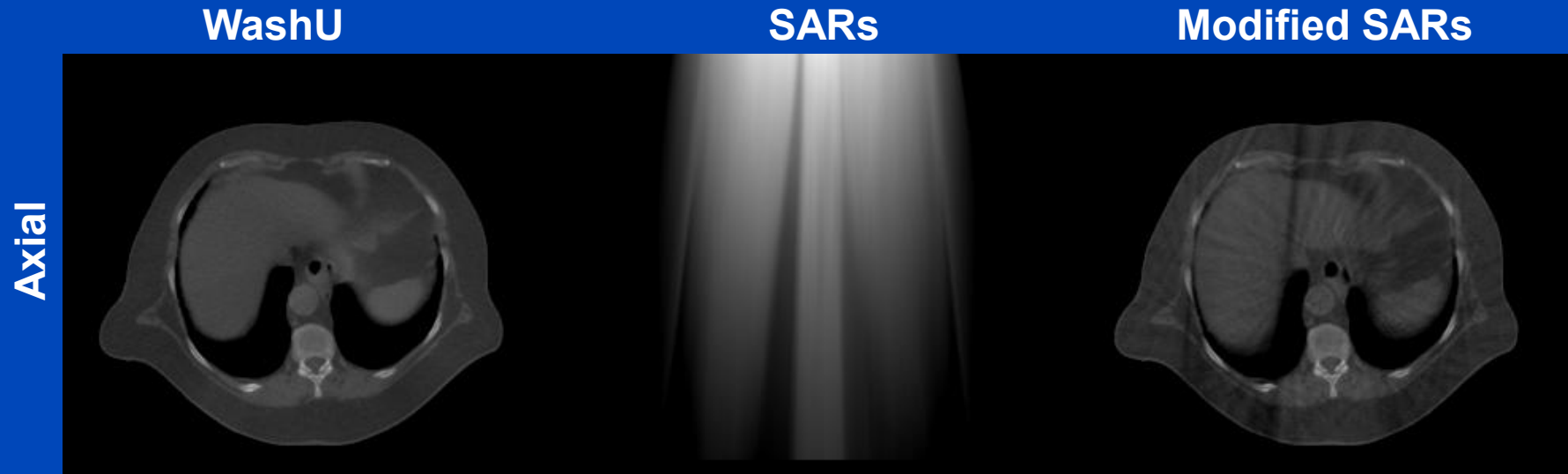


Gating = bad



Gating + MoCo = still bad

Single Angle Reconstructions (SARs)

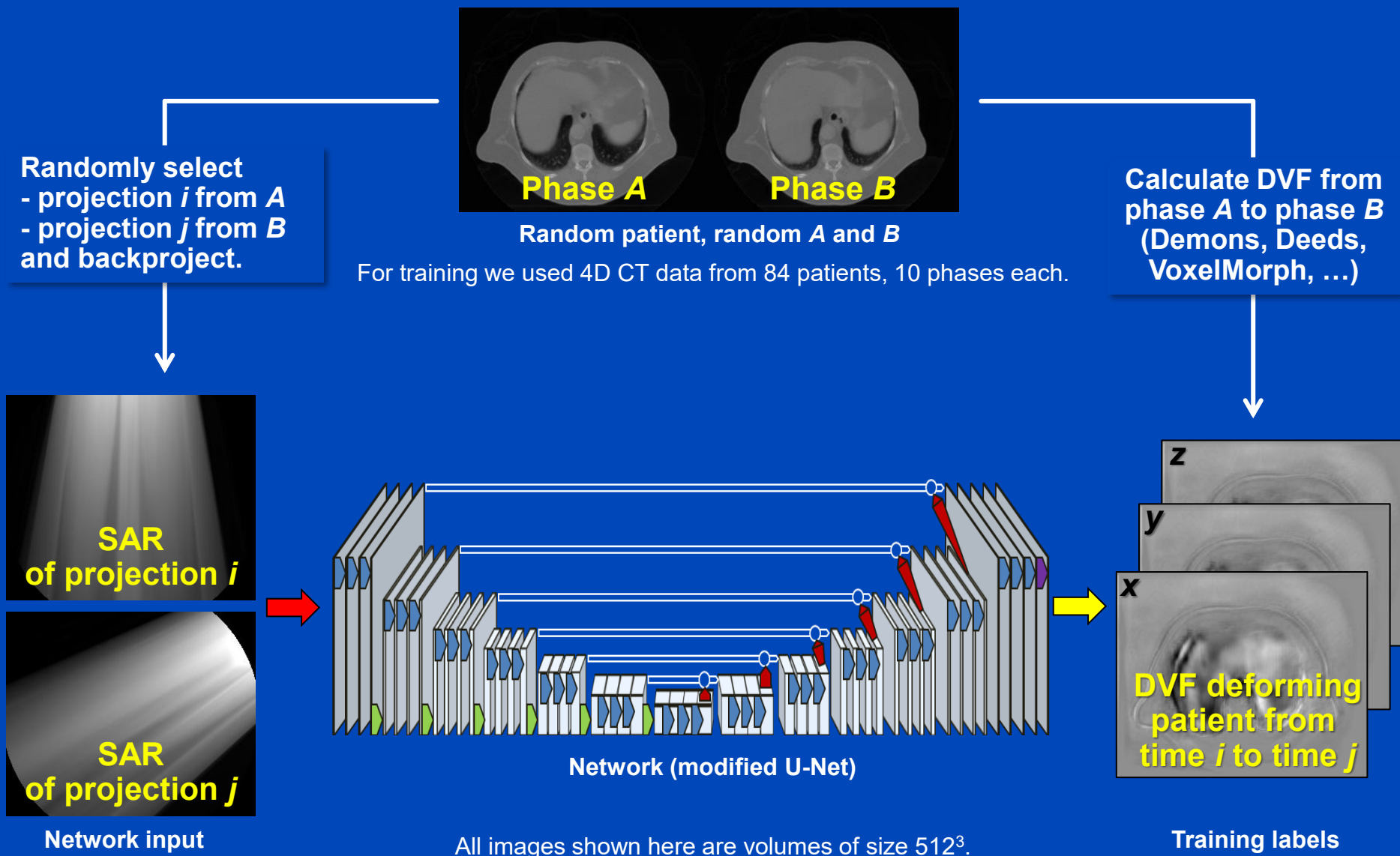


84 4D CT scans (no artifacts, high temporal resolution)

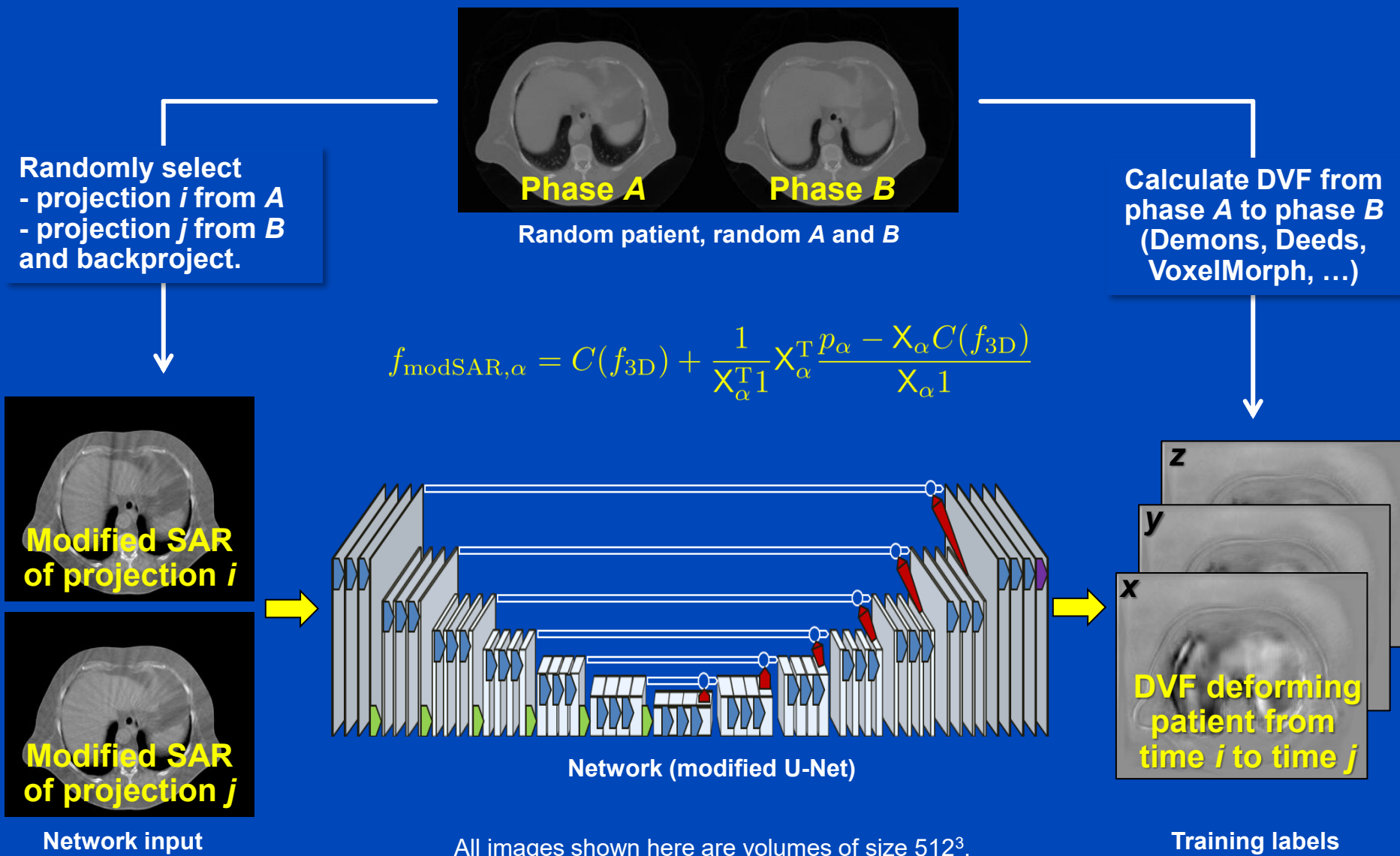
- 10 respiratory phases each (WashU/Colorado dataset)
- 10·10 combinations of phase A and B possible (including $A=B$)
- 84·10·10 displacement vector fields (DVF) known
- 720 CBCT projections¹ simulated for each CT scan (each phase)
- 84·10·720·10·720 projection pairs with known DVF

¹The actual projection numbers are between 420 and 900 and depend on the scan mode.

Training Workflow of Deep SAMoCo



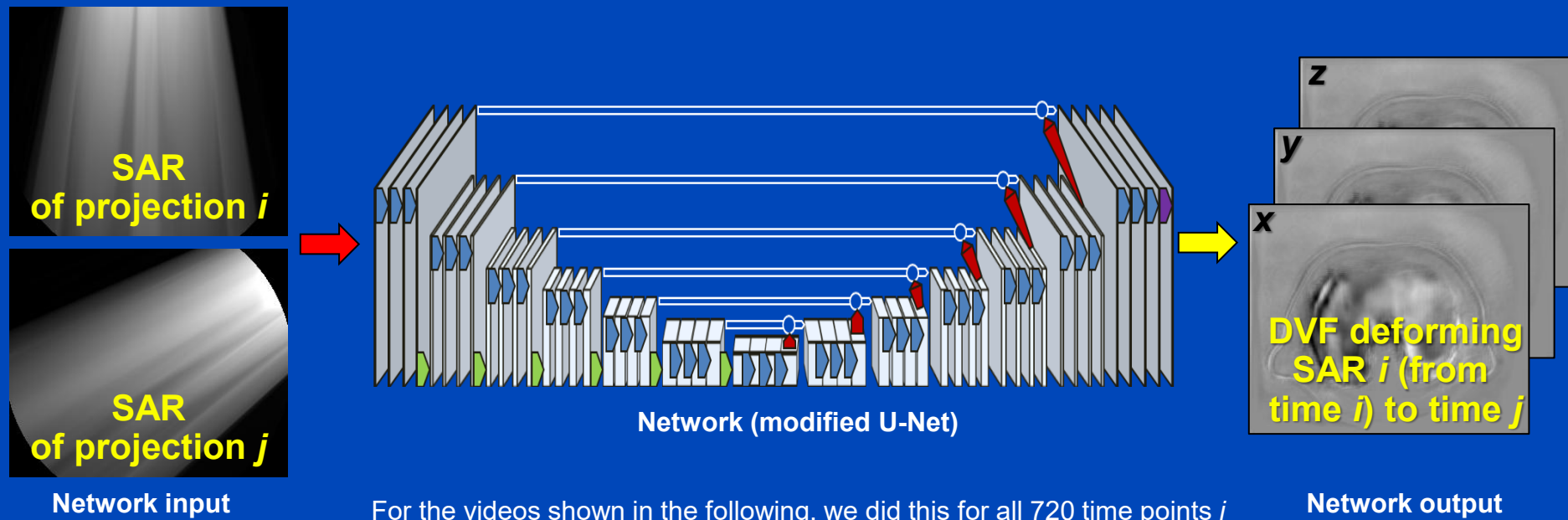
Training Workflow of Deep SAMoCo



Inference Workflow of Deep SAMoCo

- For a new patient
 - decide for the desired time point j , e.g. the one from 1 millisecond ago
 - for all $i \neq j$ get the DVFs pointing from i to j from the neural network
 - deform SARs for all $i \neq j$ into time point j
 - add all the volumes

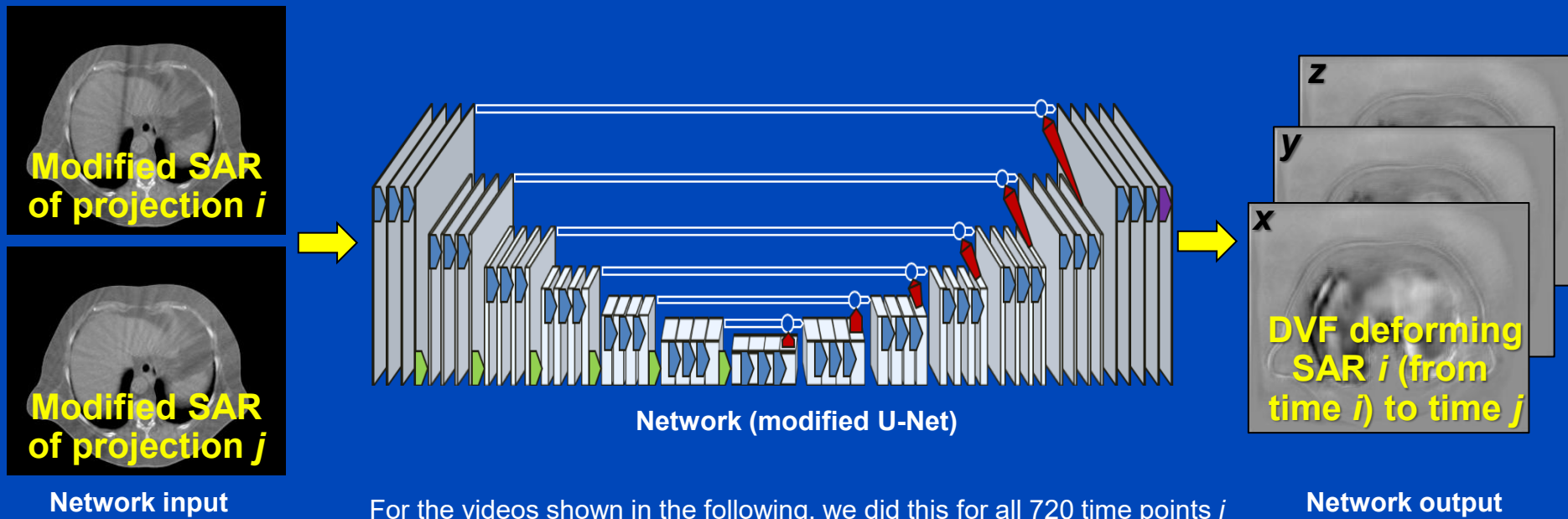
For all $i \neq j$ do



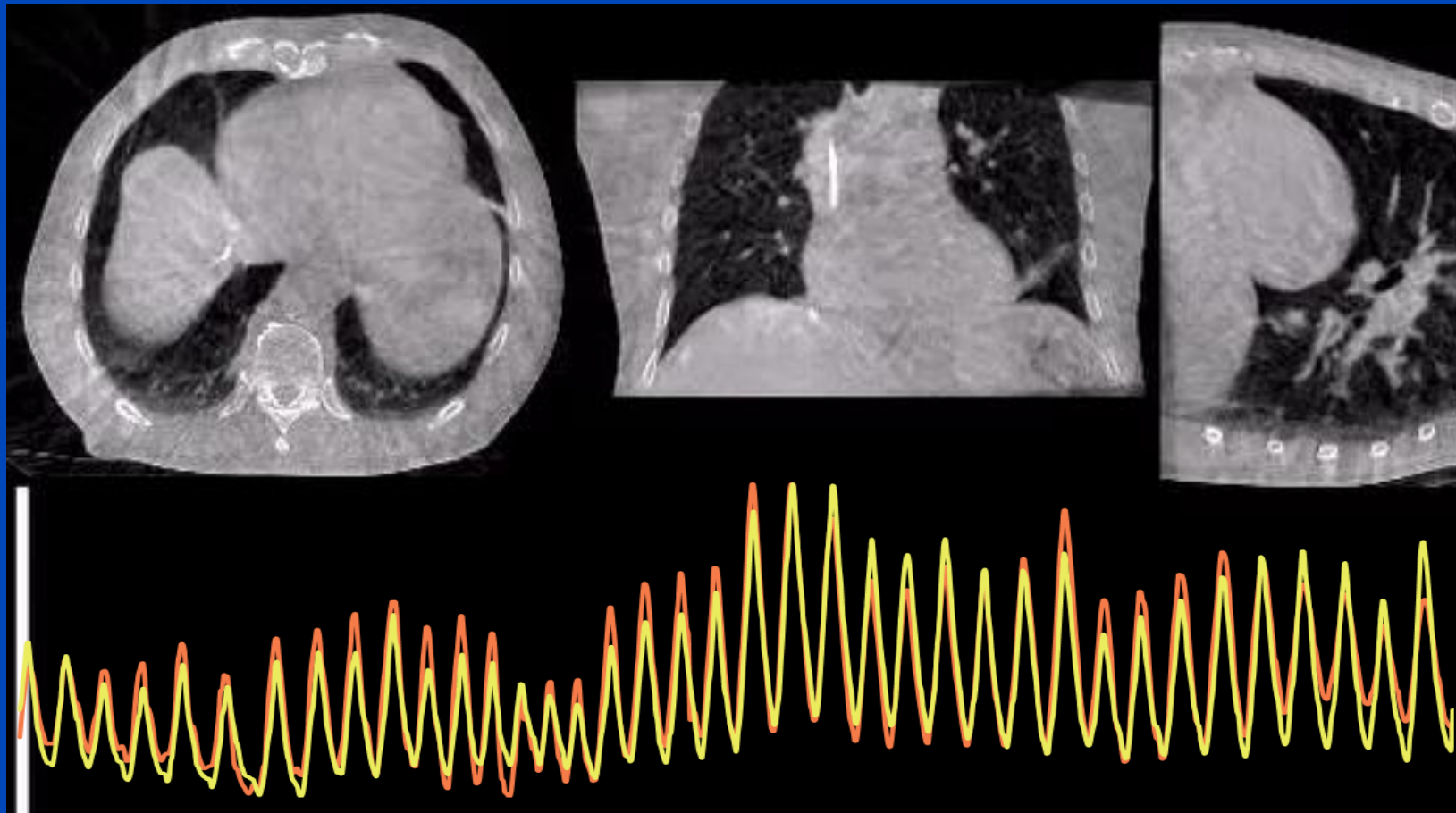
Inference Workflow of Deep SAMoCo

- For a new patient
 - decide for the desired time point j , e.g. the one from 1 millisecond ago
 - for all $i \neq j$ get the DVFs pointing from i to j from the neural network
 - deform SARs for all $i \neq j$ into time point j
 - add all the volumes

For all $i \neq j$ do

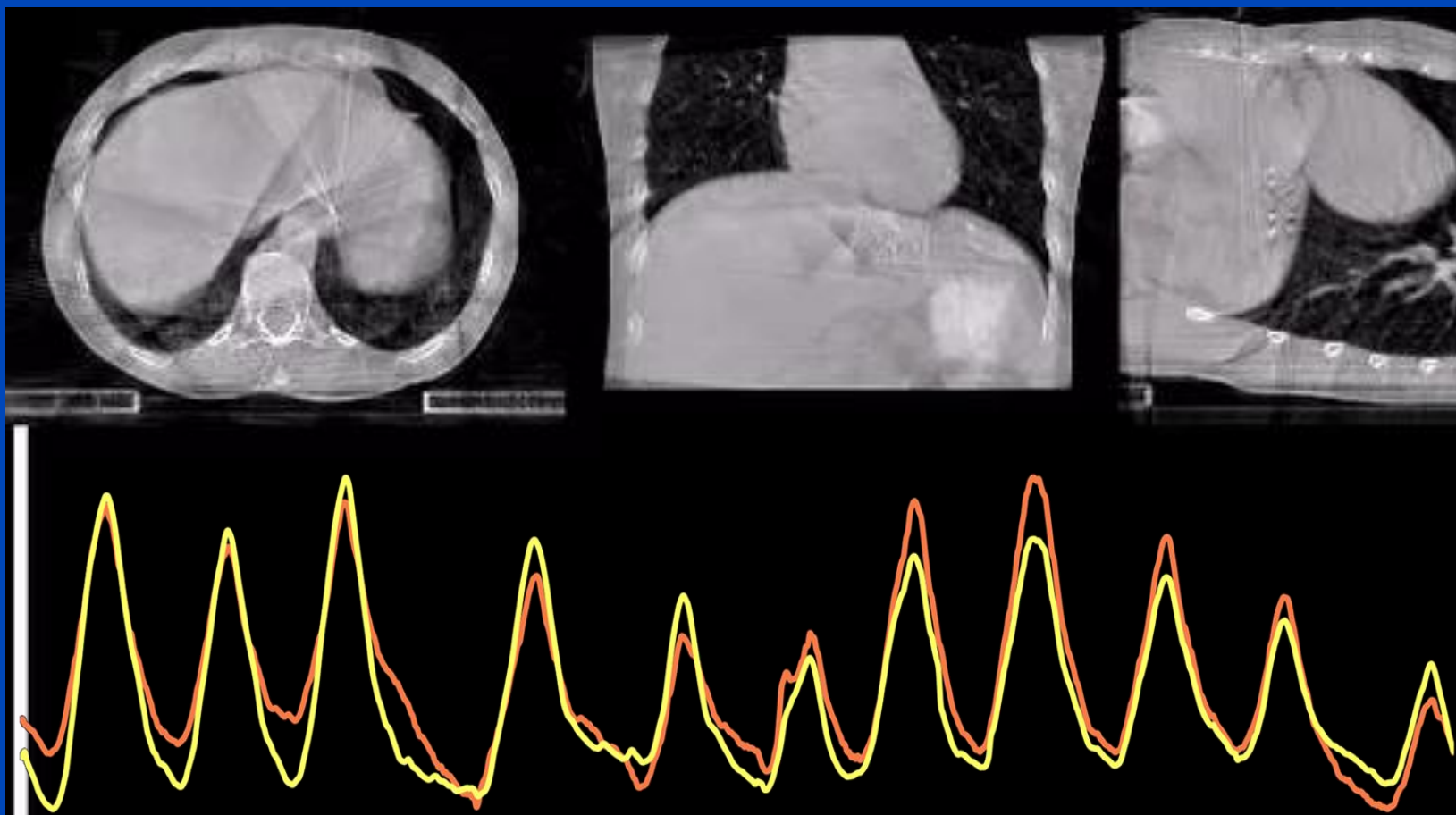


VUMC_4DThorax



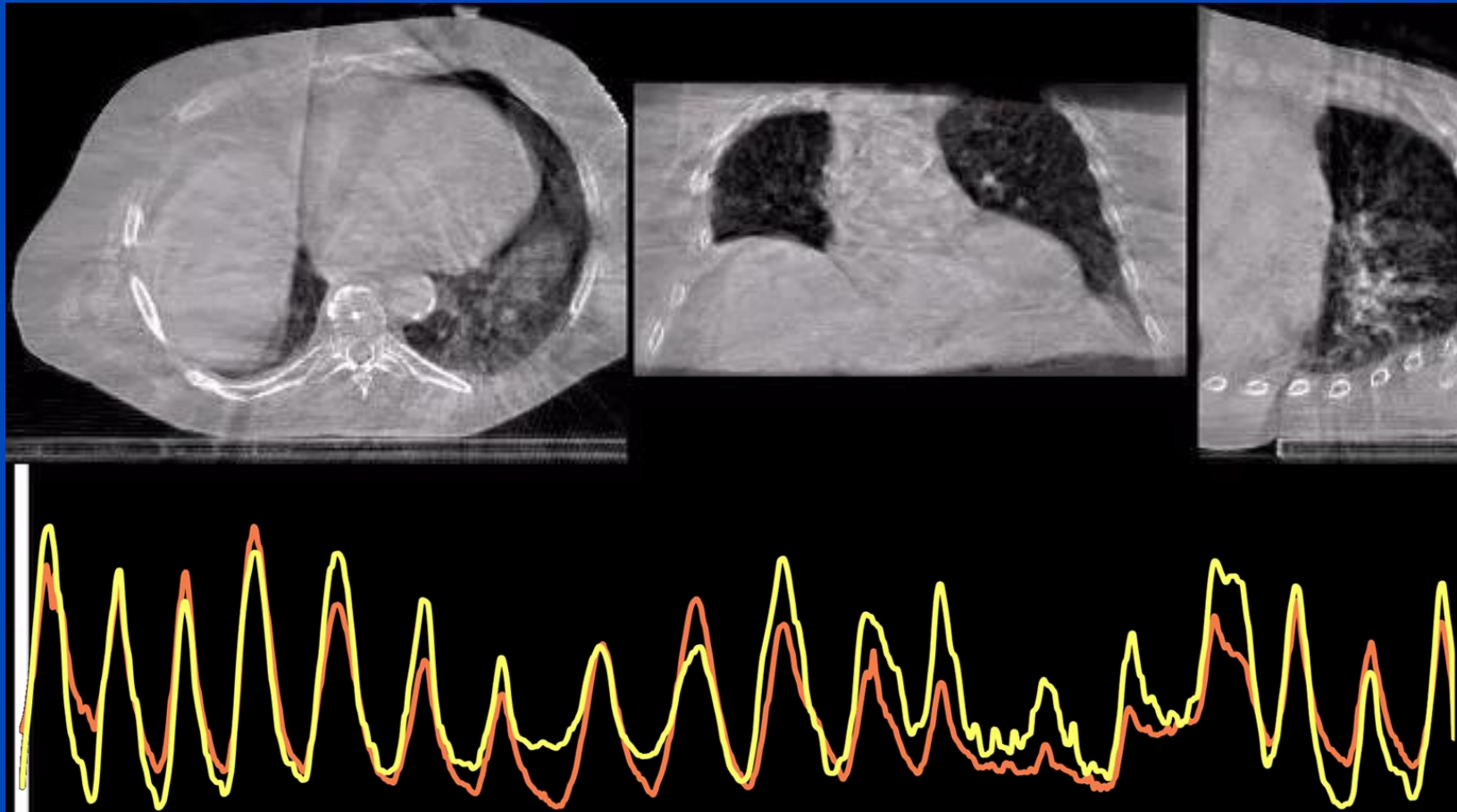
Red: RPM signal (external signal – not used for recon)
Yellow: Diaphragm motion (intrinsic signal – from PAMoCo recon)

MSK 7



Red: RPM signal (external signal – not used for recon)
Yellow: Diaphragm motion (intrinsic signal – from PAMoCo recon)

MSK 1



Red: RPM signal (external signal – not used for recon)
Yellow: Diaphragm motion (intrinsic signal – from PAMoCo recon)

RESEARCH ARTICLE

MEDICAL PHYSICS

Patient-specific radiation risk-based tube current modulation for diagnostic CT

Laura Klein^{1,2} | Chang Liu³ | Jörg Steidel^{1,2} | Lucia Enzmann^{1,2} |
Michael Knaup¹ | Stefan Sawall^{1,4} | Andreas Maier³ | Michael Lell⁵ |
Joscha Maier¹ | Marc Kachelrieß^{1,4}

¹Division of X-Ray Imaging and CT, German Cancer Research Center (DKFZ), Heidelberg, Germany

²Department of Physics and Astronomy, Ruprecht-Karls-University Heidelberg, Heidelberg, Germany

³Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany

Abstract

Purpose: Modern CT scanners use automatic exposure control (AEC) techniques, such as tube current modulation (TCM), to reduce dose delivered to patients while maintaining image quality. In contrast to conventional approaches that minimize the tube current time product of the CT scan, referred to as mAsTCM in the following, we herein propose a new method referred to as risk-TCM, which aims at reducing the radiation risk to the patient by taking into account the specific radiation risk of every dose-sensitive organ.

Patient Risk-Minimizing Tube Current Modulation (riskTCM)

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

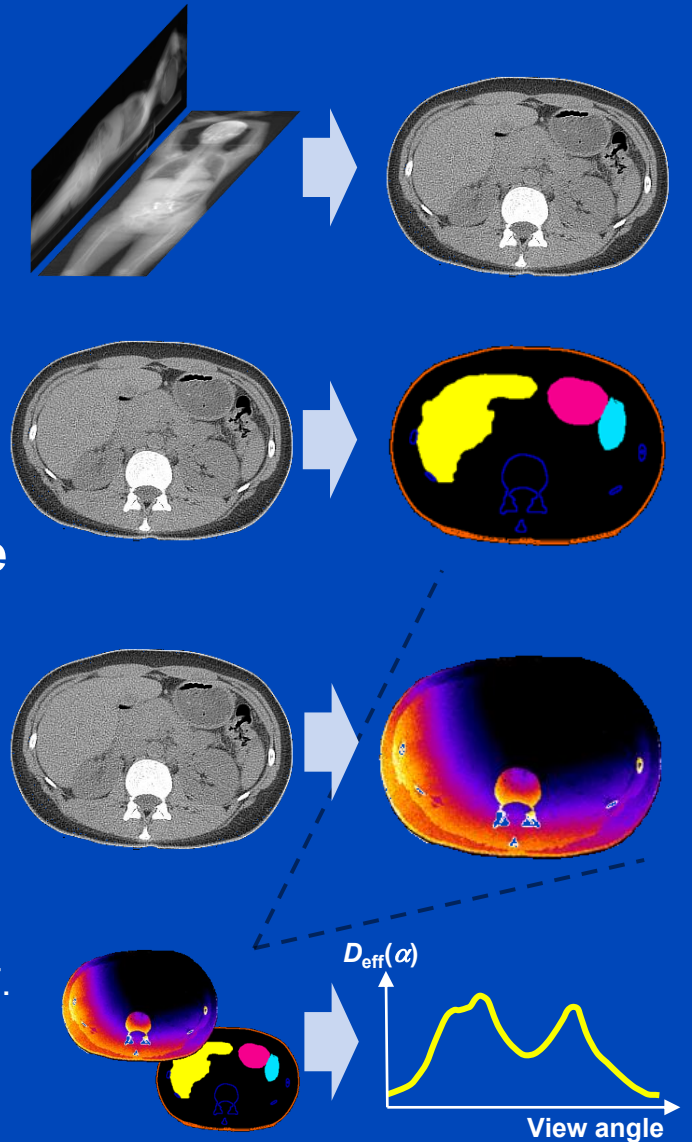
- E.g. S. Chen, M. Kachelrieß et al., Automatic multi-organ segmentation in dual-energy CT (DECT) with dedicated 3D fully convolutional DECT networks. Med. Phys. 2019.

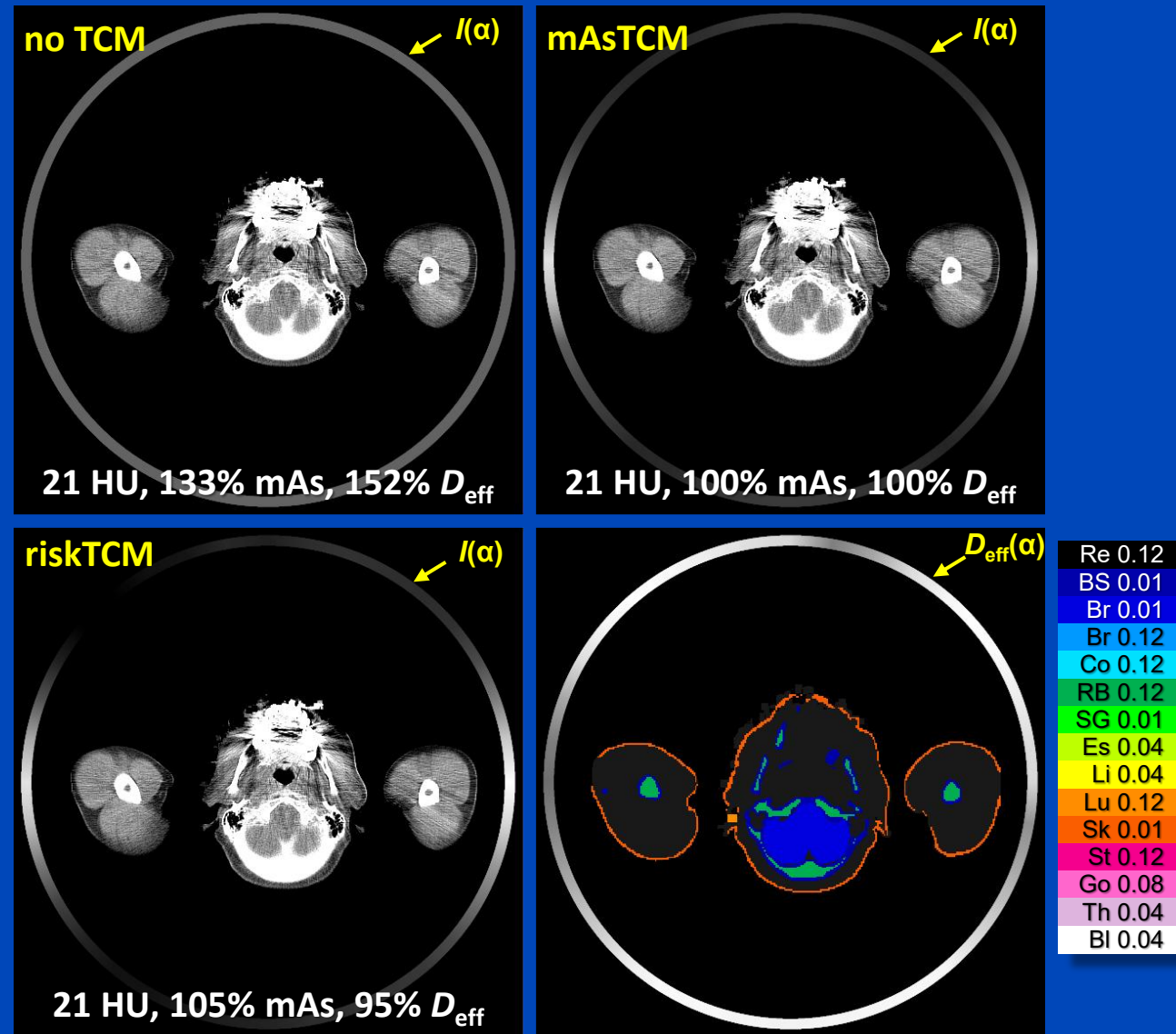
3. Calculation of the effective dose per view using the deep dose estimation (DDE)

- 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.

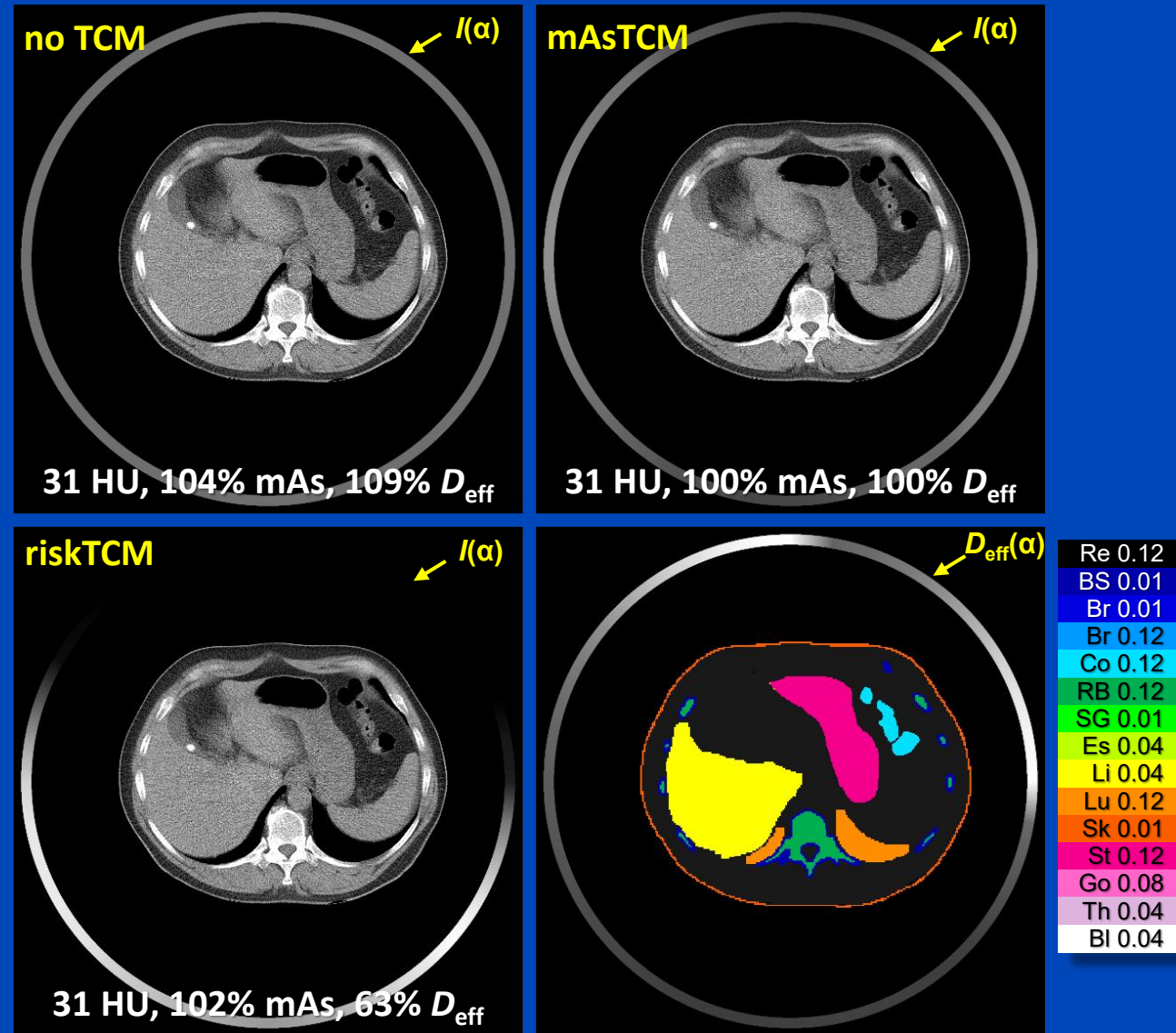
4. Determination of the tube current modulation curve that minimizes the radiation risk

- 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.





$C = 25 \text{ HU}, W = 400 \text{ HU}$

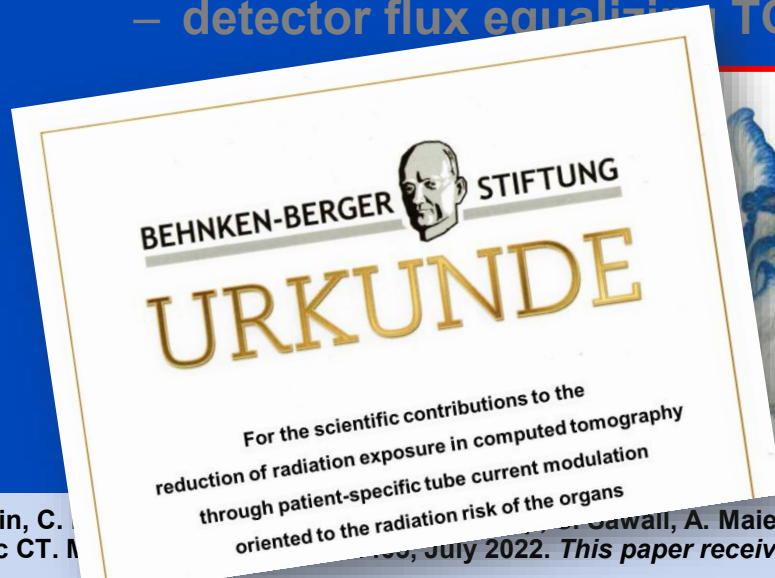


C = 25 HU, W = 400 HU

Conclusions on riskTCM

- Risk-specific TCM minimizes the patient risk.
- With D_{eff} as a risk model riskTCM can reduce risk to 30%, compared with the gold standard.
- Other risk models, e.g., weight- and sex-specific models, can be used with riskTCM as well.
- Note:
 - mAsTCM = good for the x-ray tube
 - **riskTCM = good for the patient**
 - detector flux equalizing TCM = good for the detector

It is up to the vendors to take action!



ECR 2022 – Best Research Presentation Abstract

within the topic Physics in Medical Imaging
with the presentation:

Risk-minimising tube current modulation (riskTCM)
for CT – potential dose reduction across different
tube voltages (16765)

L. Klein¹, C. Liu², J. Steidel¹, L. Enzmann¹, S. Sawall¹, J. Maier¹,
A. Maier², M. Lell³, M. Kachelrieß¹; ¹Heidelberg/DE,
²Erlangen/DE, ³Nuremberg/DE



RESEARCH ARTICLE

MEDICAL PHYSICS

Latent space reconstruction for missing data problems in CT e.g. for metal inpainting, detruncation, limited angle extrapolation, ...

Anton Kabelac^{1,2} | Elias Eulig^{1,2} | Joscha Maier¹ | Maximilian Hammermann^{1,2} |
Michael Knaup¹ | Marc Kachelrieß^{1,3}

¹Division of X-Ray Imaging and Computed Tomography, German Cancer Research Center (DKFZ), Heidelberg, Germany

²Faculty of Physics and Astronomy, Heidelberg University, Heidelberg, Germany

³Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany

Correspondence

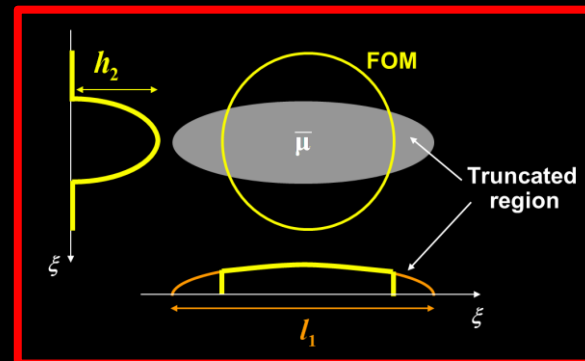
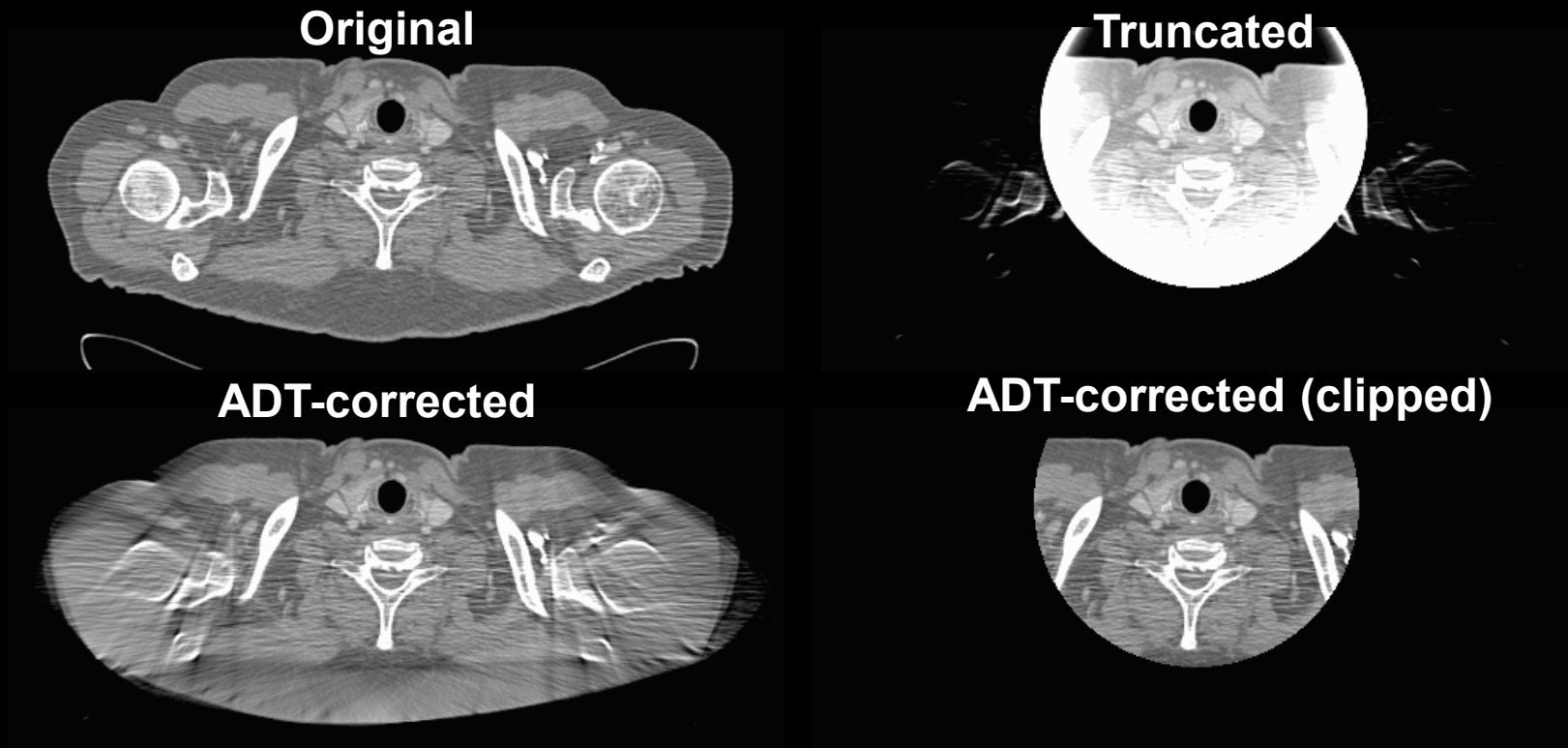
Anton Kabelac, Division of X-Ray Imaging and Computed Tomography, German Cancer Research Center (DKFZ), Heidelberg, Germany.
Email: anton.kabelac@dkfz.de

Abstract

Background: The reconstruction of a computed tomography (CT) image can be compromised by artifacts, which, in many cases, reduce the diagnostic value of the image. These artifacts often result from missing or corrupt regions in the projection data, for example, by truncation, metal, or limited angle acquisitions.

Purpose: In this work, we introduce a novel deep learning-based framework, latent space reconstruction (LSR), which enables correction of various types of artifacts arising from missing or corrupted data.

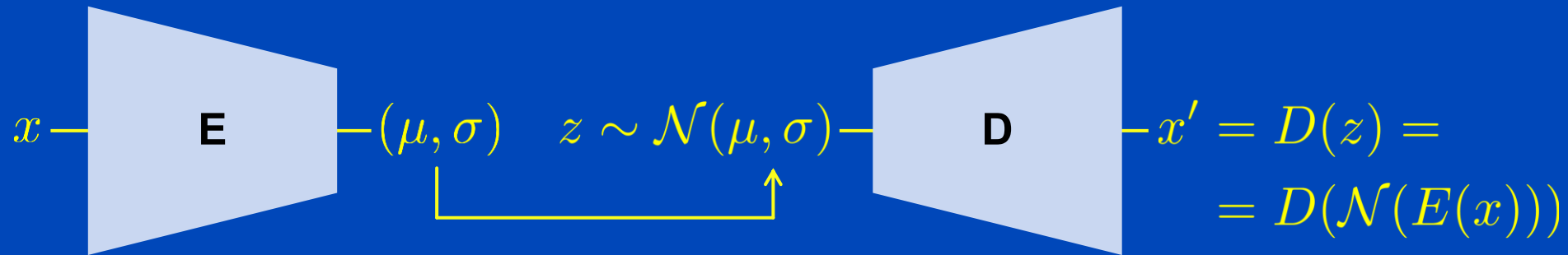
Methods: First, we train a generative neural network on uncorrupted CT images. After training, we iteratively search for the point in the latent space of this network that best matches the compromised projection data we measured. Once an optimal point is found, forward-projection of the generated CT image can be used to inpaint the corrupted or incomplete regions of the measured raw data.



$C = 0 \text{ HU}, W = 1000 \text{ HU}$

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.

Latent Space Reconstruction (LSR) for Detruncation

- Train VAE on very many untruncated CT images f_n

$$\theta = \arg \min_{\theta} \sum_n \|D(\mathcal{N}(E(f_n(\mathbf{r})))) - f_n(\mathbf{r})\|$$

- Find latent space point z to best match the truncated rawdata p

$$z = \arg \min_z \|XD(z) - p\|$$

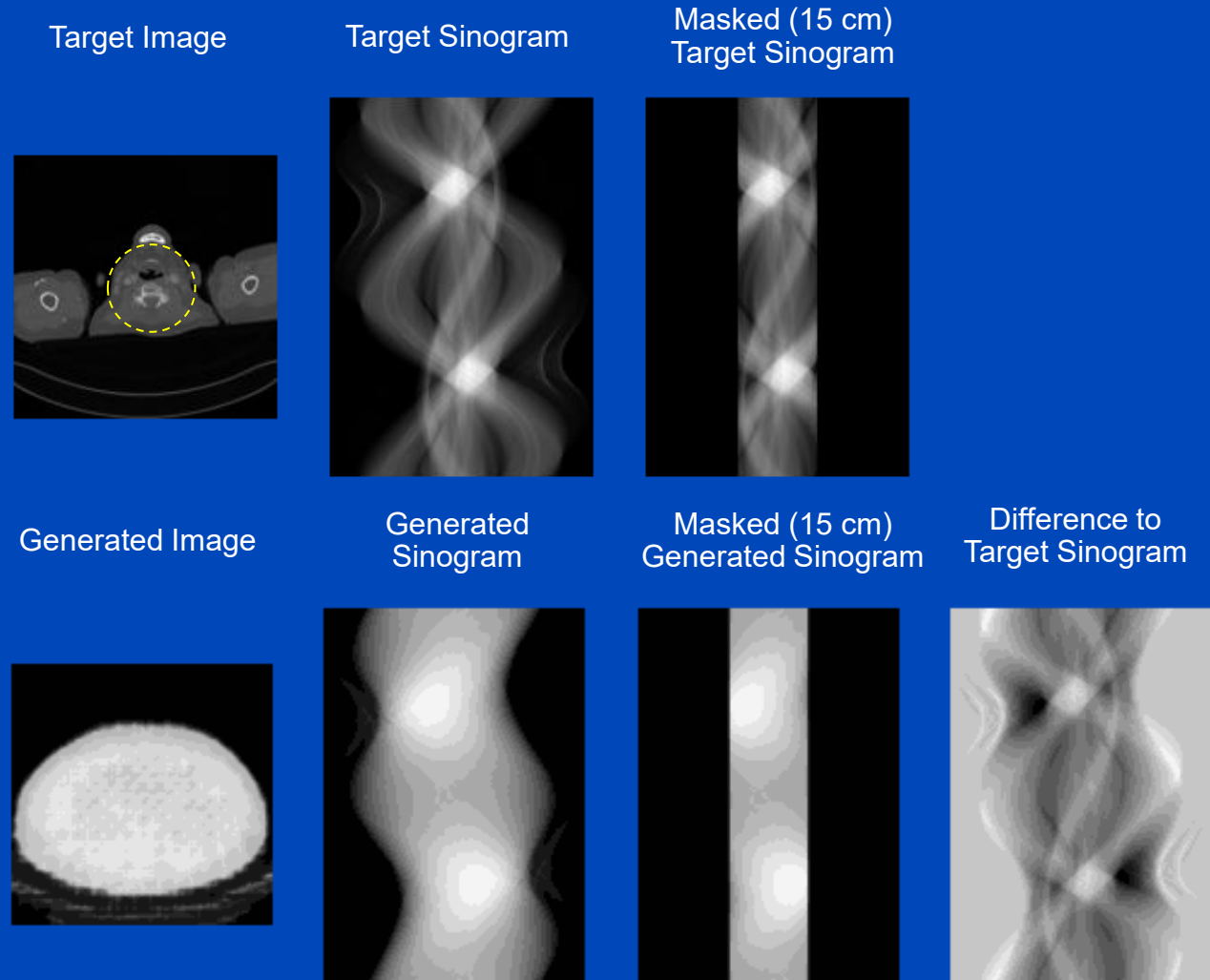
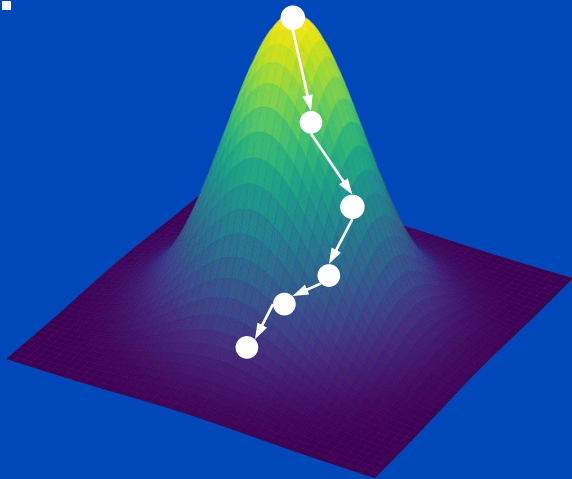
- Forward project $D(z)$ and use the resulting rawdata to extrapolate the measured rawdata.
- Do a final image reconstruction of the detruncated sinogram.

Search in Latent Space

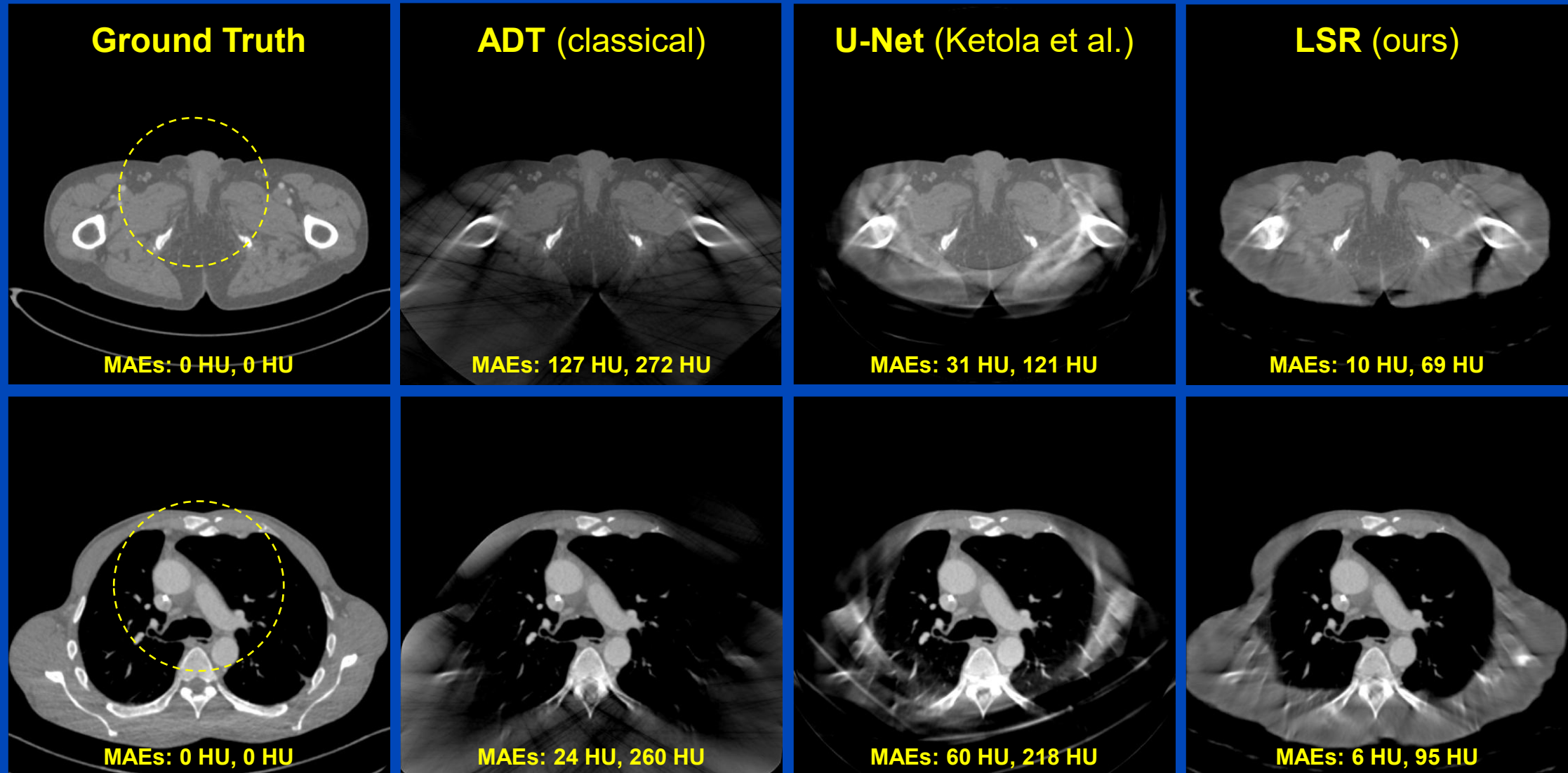
- Optimization of latent space vector in projection domain

$$z = \arg \min_z \|XD(z) - p\|_{15 \text{ cm}}$$

- Video showing intermediate images of selected iteration steps.



Results



C = 50 HU, W = 1200 HU

Summary on Deep Detruncation

- No need for machine learning to restore the gray values within the FOM.
- Image domain cosmetic detruncation can serve as an intermediate step to detruncate CT data.
- Latent space reconstruction (LSR) is an interesting way that simultaneously guarantees rawdata fidelity and nice CT images.

Thank You!



This presentation will soon be available at www.dkfz.de/ct.
Job opportunities through marc.kachelriess@dkfz.de or through DKFZ's PhD program.
Parts of the reconstruction software were provided by RayConStruct® GmbH, Nürnberg, Germany.