

Dissertation

submitted to the
Combined Faculty
of Mathematics, Engineering and Natural Sciences
of Heidelberg University, Germany
for the degree of
Doctor of Natural Sciences

Put forward by

M. Sc. Achim Byl

Born in: Gifhorn, Germany

Oral examination: 06.06.2024

Data-driven Measures for Dose Reduction and Image Quality Enhancement in Computed Tomography

Referees: Prof. Dr. Joao Seco
Prof. Dr. Marc Kachelrieß

Data-driven Measures for Dose Reduction and Image Quality Enhancement in Computed Tomography

Dose reduction without sacrificing image quality is one of the primary aims of computed tomography (CT) research. Especially in interventional applications, which require multiple scans, dose reduction is paramount. Two pathways of arbitrarily reducing dose are the reduction of tube current (low-mAs CT) and reduction of number of projections (sparse-view CT). The former will increase noise, the latter sparseness artifacts, such that there is a trade-off between dose and image quality in either case. Additionally, if the patient does not fully fit in the field of measurement, as is typical for interventional cone-beam CT, the image will suffer from truncation artifacts and a small field of view. Novel deep neural networks have shown promising results in a variety of image processing tasks. The aim of this thesis is therefore to analyze the different low-dose CT realizations in conjunction with deep learning-based image processing and facilitate reconstruction from truncated projections with both a data-driven and iterative method. Quantitative image quality analysis of low-dose CT is performed with several conventional and task-based methods. The latter are able to distinguish between sufficiently and insufficiently trained networks, ensuring a safe utilization of deep learning-based methods. The well-trained neural networks are able to support the tested 80% dose reduction by restoring image quality. Between the different realizations of low-dose CT, low-mAs CT is determined as preferable. Both detruncation methods achieve satisfactory results. However, the computational cost of DART remains prohibitive while the deep learning-based detruncation promises to increase the field of view in real-time. This in turn may improve image guidance and secondary algorithms during operations.

Datengetriebene Maßnahmen zur Dosisreduktion und Bildqualitätsverbesserung in der Computertomographie

Dosisreduktion ohne Verringerung der Bildqualität ist einer der Forschungsschwerpunkte im Bereich Computertomographie (CT). Besonders in interventionellen Anwendungen, die oft wiederholte Aufnahmen benötigen, ist Dosisreduktion entscheidend. Zwei mögliche Wege die Dosis beliebig zu verringern, sind die Senkung des Röhrenstroms und die Verringerung der Projektionsanzahl. Ersteres bewirkt ein Ansteigen des Bildrauschens, letzteres verursacht Streifenartefakte. Somit gibt es in beiden Fällen einen Kompromiss zwischen Dosis und Bildqualität. Falls der Patient nicht in das Messfeld passt, was üblich für interventielle C-Arm Systeme ist, entstehen Trunkierungsartefakte und das Field-of-View ist verkleinert. Neuartige neuronale Netze haben bereits für diverse Anwendungen aus der Bildverarbeitung hervorragende Ergebnisse geliefert. Das Ziel dieser Arbeit ist daher, die verschiedenen Niedrigdosis-CT-Implementierungen im Verbund mit Deep Learning-basierter Bildverbesserung zu analysieren, sowie die Rekonstruktion aus trunkierten Rohdaten mithilfe einer Deep Learning-basierten und einer iterative Methode zu ermöglichen. Zur quantitativen Bildqualitätsanalyse der Niedrigdosis-Scans wurden mehrere konventionelle und Aufgaben-basierte Metriken implementiert. Letztere sind in der Lage zwischen ausreichend und unzureichend trainierten Netzen zu unterscheiden und gewährleisten so die sichere Verwendung von Deep Learning Algorithmen. Die Netze ermöglichen eine Dosisreduktion von 80% durch die Wiederherstellung der Bildqualität. Von den Methoden zur Dosisreduktion wird die Senkung des Röhrenstroms als zu bevorzugen bewertet. Beide Methoden zur Detrunkierung liefern zufriedenstellende Ergebnisse. Allerdings ist der mit DART verbundene Rechenaufwand derzeit noch weitaus größer als die Inferenzzeit des Netzes, welches Korrekturen in Echtzeit verspricht. Diese können möglicherweise zur Verbesserung der Bildsteuerung während Operationen oder nachfolgenden Algorithmen beitragen.

Contents

List of Acronyms	IX
List of Tables	XI
List of Figures	XIII
1 Introduction	1
2 Fundamentals	5
2.1 Interactions of X-Rays with Matter	5
2.1.1 Lambert-Beer Law	5
2.1.2 Rayleigh Scattering	7
2.1.3 Compton Scattering	7
2.1.4 Photoelectric Effect	8
2.2 Data Acquisition	8
2.3 Image Reconstruction	9
2.4 Image Artifacts	11
2.4.1 Noise	11
2.4.2 Sparseness	13
2.4.3 Truncation	14
2.5 Neural Networks	15
2.5.1 Neurons	15
2.5.2 Layers	16
2.5.3 Network Training	18
3 Materials & Methods	21
3.1 Low-Dose CT Correction	21
3.1.1 Simulating Low-Dose CT Data	21
3.1.2 CNN Denoising	26
3.2 Image Quality Assessment	28
3.2.1 Conventional Analysis	28
3.2.2 Lesion Segmentation	29
3.2.3 Detectability Metrics	33
3.3 Detruncation	37

3.3.1	Simulating Truncated CT Data	37
3.3.2	Cosine Detruncation	37
3.3.3	DART Detruncation	38
3.3.4	CNN Detruncation	39
3.3.5	Analysis	39
4	Results	41
4.1	Comparison of Dose Reduction Approaches	41
4.1.1	Conventional Image Quality Metrics	45
4.1.2	Task-based Image Quality Metrics	46
4.2	Validation of Task-Based Metrics	51
4.2.1	Segmentation Metrics	51
4.2.2	Model Observer	53
4.3	Ablation Study	58
4.3.1	Network Architecture	58
4.3.2	Loss Function	65
4.4	Quality Assurance of Insufficiently Trained Networks	73
4.4.1	Reducing Training Set Size	73
4.4.2	Reducing Number of Epochs	82
4.5	Detruncation	89
4.5.1	50% Truncation	89
4.5.2	25% Truncation	94
5	Summary & Discussion	99
5.1	Low-Dose CT Denoising	99
5.1.1	Low-Dose CT Realizations	100
5.1.2	Image Metrics	102
5.2	Detruncation	103
6	Conclusions	105
	Bibliography	107

List of Acronyms

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
ASD	Average symmetric surface distance
AUC	Area under receiver operating characteristic curve
BKE	Background known exactly
BKS	Background known statistically
BP	Back-projection
CBCT	Cone-beam computed tomography
CHO	Channelized hotelling observer
CNN	Convolutional neural network
CT	Computed tomography
DART	Discrete algebraic reconstruction technique
eFOV	Extended field of view
FBP	Filtered back-projection
FC	Fully-connected
FOM	Field of measurement
FOV	Field of view
FP	Forwards-projection
H-DenseUNet	Hybrid dense U-Net

HU	Hounsfield units
ISBI	International Symposium on Biomedical Imaging
LiTS	Liver Tumor Segmentation
MAR	Metal artifact reduction
MICCAI	Medical Image Computing and Computer Assisted Intervention Society
MO	Model observer
MSD	Maximum symmetric surface distance
MSE	Mean squared error
NN	Neural network
PET	Positron emission tomography
QA	Quality assurance
ReLU	Rectified linear unit
RMSD	Root mean squared symmetric surface distance
RMSE	Root mean squared error
ROC	Receiver operating characteristic
ROI	Region of interest
RVD	Relative volume difference
SART	Simultaneous algebraic reconstruction technique
SKE	Signal known exactly
SKS	Signal known statistically
SNR	Signal-to-noise ratio
SSIM	Structural similarity index metric
TCM	Tube current modulation

List of Tables

4.1	Conventional quantitative image quality analysis of the W-Net denoising.	45
4.2	95% confidence intervals of AUC estimates for the SKE-BKS model observer for 2 mm and 4 mm lesions with and without W-Net denoising.	46
4.3	Quantitative results of the segmentation of liver lesions after W-Net denoising.	49
4.4	Validation results of the segmentation metrics.	52
4.5	Validation results of the segmentation metrics calculated in reference to the CNN segmentations of the ground truth CT images.	52
4.6	Conventional quantitative results using RMSE and SSIM for the ablation study with $N = 229$	58
4.7	95% confidence intervals of AUC estimates of the SKE-BKS model observer for the network architecture ablation study with $N = 229$	65
4.8	Network architecture ablation study results of the segmentation of liver lesions after CNN denoising with $N = 229$	65
4.9	Conventional quantitative results for the loss function study.	66
4.10	95% confidence intervals of AUC of the model observer study for MSE and MSE with perceptual loss functions.	66
4.11	Quantitative results of the segmentation of liver lesions after CNN denoising with MSE loss function.	73
4.12	Conventional quantitative results for the train set size study.	81
4.13	95% confidence intervals of AUC of the model observer study for different train set sizes relative to full set and two lesion sizes.	81
4.14	Quantitative results of the segmentation of liver lesions after CNN denoising with different relative train set sizes.	81
4.15	Conventional quantitative results for the epoch number study.	82
4.16	95% confidence intervals of AUC of the model observer study for different number of epochs with $N = 229$ and two lesion sizes.	89
4.17	Quantitative results of the segmentation of liver lesions after CNN denoising with different number of epochs.	89
4.18	Quantitative results for 50% truncation.	90
4.19	Quantitative results for 25% truncation.	97

List of Figures

2.1	Contributions of Rayleigh scattering, the photoelectric effect and Compton scattering to the total attenuation coefficient.	6
2.2	Relevant interactions of X-rays with matter.	7
2.3	Schematic of an X-ray tube.	9
2.4	2D parallel beam geometry for the filtered backprojection	10
2.5	Examples of increasing image noise with decreasing number of photons relative to full dose.	13
2.6	Examples of increasing sparseness artifacts with decreasing number of projections.	14
2.7	Examples of increasing truncation artifacts with decreasing number of detector pixels.	15
2.8	Fully-connected layer with two inputs, three neurons, and one output. .	16
2.9	Schematic of a 1D convolutional layer.	18
3.1	Example data from the LiTS dataset.	22
3.2	Relationship between number of projections and tube current at constant dose for different dose implementations. Red dashed lines mark the tested configurations.	24
3.3	Tube current vs. number of projections for the five tested dose implementations on a log-log scale. Red dashed lines mark the tested configurations.	25
3.4	Network architecture of the W-Net	27
3.5	U-Net architecture	27
3.6	Architecture of the H-DenseUNet used for liver and lesion segmentation. .	30
3.7	H-DenseUNet segmentations of liver and lesions on the ground truth LiTS dataset.	30
3.8	Possible issues in the mapping of predicted lesions to ground truth lesions. .	31
3.9	Two examples of the correspondence algorithm to find matched lesions in ground truth and prediction.	31
3.10	Pipeline for the model observer study.	34
3.11	Gabor channels used for the channelized hotelling observer.	36
3.12	Scheme of the DART reconstruction.	38

4.1	Denoising results for a pelvis slice.	42
4.2	Denoising results for an abdomen slice.	43
4.3	Denoising results for a thorax slice.	44
4.4	Denoising results for the model observer study with a 2 mm lesion.	47
4.5	Denoising results for the model observer study with a 4 mm lesion.	48
4.6	CNN-based segmentation contours after W-Net denoising of low-dose CT.	50
4.7	95% confidence interval estimates of the AUC for the repeated SKE-BKS model observer with low-dose images.	54
4.8	95% confidence interval estimates of the AUC for the repeated SKE-BKS model observer with denoised images.	55
4.9	AUC spread for the SKE-BKE model observer with low-dose images.	56
4.10	AUC spread for the SKE-BKE model observer with denoising images.	57
4.11	Network architecture ablation study denoising results for $N = 229$	59
4.12	Network architecture ablation study denoising results for $N = 229$	60
4.13	Network architecture ablation study denoising results for $N = 229$	61
4.14	Network architecture ablation study denoising results for the model observer study with a 2 mm lesion and $N = 229$	62
4.15	Network architecture ablation study denoising results for the model observer study with a 4 mm lesion and $N = 229$	63
4.16	Network architecture ablation study segmentation contours of liver and liver lesions.	64
4.17	Loss function study denoising results for pelvis scan.	67
4.18	Loss function study denoising results for abdomen scan.	68
4.19	Loss function study denoising results for thorax scan.	69
4.20	Loss function study denoising results for the model observer study with a 2 mm lesion.	70
4.21	Loss function study denoising results for the model observer study with a 4 mm lesion.	71
4.22	Loss function study segmentation contours of liver and liver lesions after denoising of low-dose CT.	72
4.23	Loss curves for small train sets	74
4.24	Train set size study denoising results for pelvis scan.	75
4.25	Train set size study denoising results for abdomen scan.	76
4.26	Train set size study denoising results for thorax scan.	77
4.27	Train set size study denoising results for the model observer study with a 2 mm lesion.	78
4.28	Train set size study denoising results for the model observer study with a 4 mm lesion.	79
4.29	Train set size study segmentation contours of liver and liver lesions after denoising of low-dose CT.	80
4.30	Epoch number study denoising results for pelvis scan.	83
4.31	Epoch number study denoising results for abdomen scan.	84
4.32	Epoch number study denoising results for thorax scan.	85

4.33	Epoch number study denoising results for the model observer study with a 2 mm lesion.	86
4.34	Train set size study denoising results for the model observer study with a 4 mm lesion.	87
4.35	Epoch number study segmentation contours of liver and liver lesions after denoising of low-dose CT.	88
4.36	DART iterations with 50% truncation.	91
4.37	Detruncation result images with 50% truncation.	92
4.38	Detruncation result projections with 50% truncation.	93
4.39	Detruncation result images with 25% truncation.	95
4.40	Detruncation result projections with 25% truncation.	96

1 | Introduction

Computed tomography (CT) is one of the main imaging modalities in the clinical workflow. It generates cross-sectional slices of the human body from several X-ray projections in scan times of a few seconds [1]. Historically, it was the first modality to provide images that were not inhibited by the superposition of different structures [2]. Four years after its invention by Hounsfield in 1973 [3], about 1000 systems were installed [4]. In the modern clinical landscape, virtually every hospital has at least one CT [5], which highlights its unequivocal importance to current diagnostics.

While CT was initially used primarily for diagnostics, developments in flat panel detector technology have made three-dimensional (3D) imaging available to interventional applications via C-arm cone-beam computed tomography (CBCT) systems [6]–[10]. This includes image guidance for minimally invasive procedures, e.g. for vascular stenting [7], [11], [12], embolization [13], spine surgery [14], and follow-up treatments after hip replacement [15]. Other scenarios in which CBCT is performed are chemo- and radioembolization for cancer treatment [13], [16], [17], radiotherapy [6], [9], [18], and dental scans [19], [20]. In general, the scans may be used both for image guidance, as well as verification of treatment success.

The main concern when employing CT stems from the X-rays that are used in the image formation: each scan is associated with a significant radiation dose to the patient, and in the interventional context, to the clinical staff. The exact dose varies between patients, body region and type of exam. Wider cross-sections, e.g. shoulders or abdomen, or obese patients, require more radiation to achieve satisfactory image qualities [1]. While a single abdominal may only cause an effective dose of 4.5 mSv, oftentimes it is necessary to perform multiple acquisitions, increasing the total dose to the patient. For example, treatment planning in radiotherapy requires follow-up scans to tailor the plan to the actual patient position and identify treatment progress [9], [16], [18]. Likewise, image-guided surgeries necessitate continued updates on the position of surgical instruments [7], [21]. As a 3D CBCT needs a large number of projections, two-dimensional (2D) X-ray imaging, e.g. 2D fluoroscopy, is still prevalent for interventions [22]. In order to enable more frequent utilization of 3D scans, thus providing the surgeon with additional information leading to improved surgeries, dose reductions are mandatory.

Dose reduction has therefore long been one of the primary research aims related to CT. Several hardware and software-based approaches have been investigated and implemented in commercial systems. Methods that act before the reconstruction

process include prefilters that shape the spectrum of the X-ray beam in a manner that reduces low energy photons [1], [23]–[25]. This decreases the fraction of photons that contribute to patient dose but not to the detected signal and increases image quality. The development of novel photon counting CT detectors promises further potential for low-dose CT, partially due to a lack of electronic noise [26]–[30]. As stated before, the dose required for a diagnostically valuable image changes with the cross-section of the patient. Because the thickness of the patient is not the same for all directions, tube current modulation (TCM) is able to reduce the dose for projections where the beam has to penetrate less tissue, and vice versa [23], [31]–[35]. This leads to a more efficient utilization of radiation in terms of image noise. Likewise, the TCM can be used to protect specific organs or minimize the effective dose of the acquisition [29], [36].

In general, the dose can be arbitrarily reduced by decreasing the tube current, or the number of projections. Both will lead to a net decrease in X-rays and thus dose, but will deteriorate the image quality if not sufficiently corrected for. Reducing the tube current in so-called low-mAs CT increases the noise of the measured projections according to the Poisson distribution of X-ray detection [1], [37]. This noise is then transferred to the image in the reconstruction. A plethora of methods have been developed to reduce image noise. These include adaptive filters [24], [38]–[41] and advanced iterative reconstructions that may include a sophisticated model of the CT data generation process [42]–[46]. Likewise, sparse-view CT, i.e. acquisitions with fewer projections, causes streak-like artifacts in the image due to the reduced angular sampling. A variety of iterative [47]–[50] and compressed sensing methods [47], [51]–[55] have been established to reconstruct from sparse data.

While conventional methods have enabled moderate dose reductions already, novel deep learning algorithms have shown highly promising results in a multitude of image processing tasks [56]–[60]. This includes classification [61], [62] and segmentation of medical images [63]–[66]. Neural networks (NNs) have also been applied to low-mAs scans, yielding favorable results [41], [67]–[79]. Recently, the first deep learning methods have been implemented in commercial CT systems [80]–[82]. NNs incorporate prior information via the training process into the image correction, thus achieving results superior to conventional methods. Similarly, deep learning-based algorithms have successfully reduced artifacts in sparse-view acquisitions [55], [58], [76], [83]–[104]. Despite the large amount of literature on either method of dose reduction, there is scarcely any research on which approach is preferable [76].

Another case of reconstructing scans with insufficient raw data is the problem of truncated projections [1]. This occurs frequently in interventional CBCT, where the detector is often only 30 cm wide in axial direction, such that the patient does not fully fit in the field of measurement (FOM). The collimation of the X-ray beam may be further reduced, to decrease the patient dose [9]. Conventional reconstruction algorithms applied to these data will create cupping artifacts inside the field of view (FOV) [1], [105]. In addition, information outside the FOV is lost, which may be beneficial for secondary algorithms such as positron emission tomography (PET) attenuation correction, scatter correction or beam hardening correction [106]. Established methods to reduce truncation artifacts typically extrapolate the sinograms based on some simplified patient model,

e.g. a water cylinder [106]–[112]. Another algorithm that has been employed for the reconstruction from insufficient raw data, primarily limited angle and sparse acquisitions but also truncation, is the discrete algebraic reconstruction technique (DART) method [113]–[123]. Recently, deep learning-based methods have shown promising results for the detruncation task [95], [124]–[131].

While state-of-the-art NN appear to outperform analytic or iterative methods in image reconstruction and processing, they are significantly less predictable [132]. Indeed, the networks themselves can be considered a black box due to the large number of parameters. One approach to remedy this lack of interpretability, is to use the network to find parameters of an analytic or iterative method [42]. To ensure the safe utilization of networks in general, sophisticated quality assurance is necessary. This is especially important in medical imaging. For instance, the removal of a lesion by a denoising algorithm may lead to a misdiagnosis.

The aim of this thesis is to investigate the potential of deep learning-based methods to reduce CT dose in conjunction with low-mAs and sparse-view CT, and to reconstruct images from truncated data. Specifically, this work analyses different approaches to low-dose CT and seeks to determine the optimal combination of reducing tube current and number of projections. The low-dose scans are corrected with dedicated NNs, and evaluated on classical image quality measures as well as task-based metrics. For the latter, low-dose scans of patients with liver lesions are simulated and corrected. Furthermore, several metrics of lesion segmentation accuracy are computed with a state-of-the-art deep segmentation algorithm. In addition, analytical model observers are implemented to determine lesion detectability in proximity to human perception [133]–[141]. The work then investigates whether the metrics are able to distinguish between sufficiently and insufficiently trained networks. For the task of detruncation, truncated scans are simulated and subsequently reconstructed with a dual-domain NN, as well as DART-based detruncation. The results are then quantitatively analyzed inside and outside the original FOV.

The structure of this work is as follows: Chapter 2 contains the fundamentals of X-ray computed tomography, including the underlying physical processes, image reconstruction methods and artifacts. In addition, it discusses the basics of machine learning and NNs. Chapter 3 describes the methodology of this work. It outlines the simulation of low-dose CT, network architectures and training schemes. Furthermore, it introduces the task-based image quality metrics. Moreover, the detruncation algorithms are presented. Chapter 4 provides the experimental results. First, dose reduction approaches are evaluated and compared. This is followed by an ablation study of the network architecture and the loss function. Subsequently, the developed metrics are applied to insufficiently trained networks. Finally, the DART and NN detruncation results are shown. Chapter 5, summarizes and discusses the results of this work. Lastly, chapter 6 draws the conclusions of this thesis.

2 | Fundamentals

2.1 Interactions of X-Rays with Matter

CT is made possible by the interaction of X-rays with the patient and detector. X-rays are photons generated as bremsstrahlung or characteristic X-rays within a wavelength range of 10^{-8} m to 10^{-13} m [2]. Notably, photons originating from radioactive decay are called γ -rays, even if they fall into the wavelength range above. The diagnostic value of X-rays was acknowledged immediately, such that they have been used for medical imaging since their discovery by W. C. Röntgen in 1898 [1], [142].

2.1.1 Lambert-Beer Law

As an X-ray beam passes through matter, a fraction of the photons will interact with the object and are attenuated. Two parameters govern how strongly the beam is attenuated: the linear attenuation coefficient μ and the thickness of the material d . The attenuation coefficient depends on the energy of the photon E and the object material, i.e. its density. For simplicity, the following examines a mono-energetic X-ray pencil beam with an initial intensity I_0 and a homogeneous object. After passing through the object, the intensity of the beam is

$$I = I_0 \cdot e^{-\mu \cdot d}, \quad (2.1)$$

according to the Lambert-Beer law. In diagnostic imaging, most X-rays have energies between 30 keV and 150 keV. In this case, three physical effects contribute to the total attenuation coefficient: Rayleigh scattering (μ_R), the photoelectric effect (μ_P), and Compton scattering (μ_C) [1], [105]. Figure 2.1 illustrates the partial and total attenuation coefficients of water and lead for different photon energies. A fourth effect, pair production, becomes relevant for high energy X-rays, i.e. above 1.022 MeV. Figure 2.2 details the physical processes involved in the relevant interactions, which will be discussed in detail in the following subsections. Note that the Lambert-Beer law only holds for a pencil beam, where all scattered photons are removed from the beam. For a brief overview of scatter in CT, see section 2.4. The total attenuation μ is then the sum of the three partial attenuation coefficients:

$$\mu = \mu_R + \mu_P + \mu_C. \quad (2.2)$$

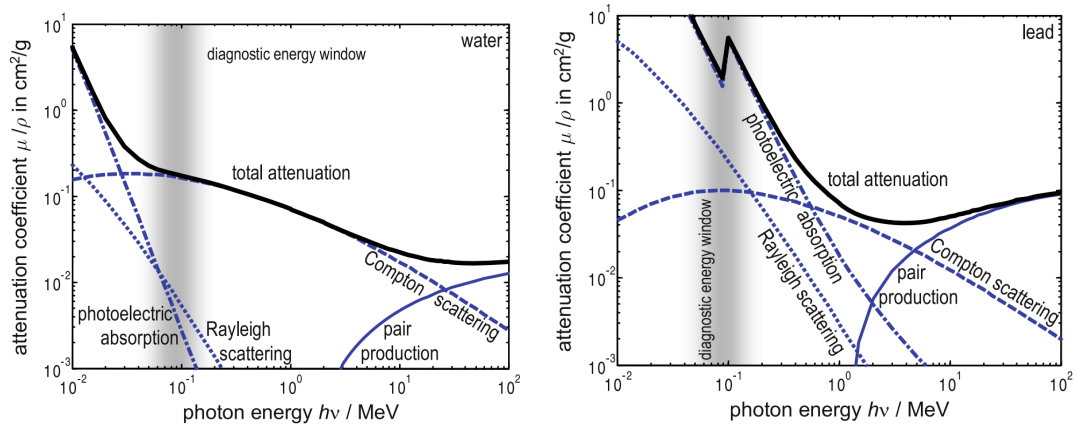


Figure 2.1: Contributions of Rayleigh scattering, the photoelectric effect and Compton scattering to the total attenuation coefficient of water (left) and lead (right). The diagnostic energy window that is relevant for CT is indicated. From [2].

Typically, interaction probabilities are indicated by the cross-section σ . Expressing μ in terms of σ yields

$$\mu = \frac{\rho N_A}{A} \sigma, \quad (2.3)$$

where ρ is the density of the material, N_A is Avogadro's number and A is the nucleon number. The first term in the equation above is the number of atoms per unit volume. Note that since the density of a material can fluctuate, databases primarily list the mass attenuation coefficient $\frac{\mu}{\rho}$. The next sections elaborate on the different photon-matter interactions and discuss their relevance for varying energies and materials based on the cross-section.

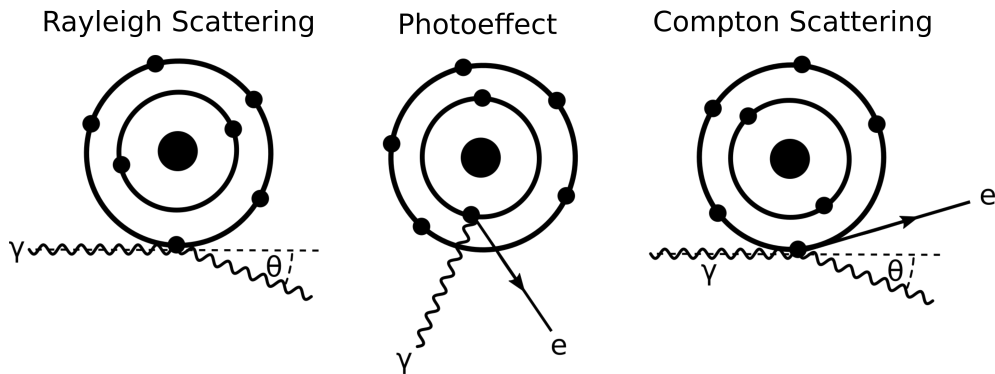


Figure 2.2: Relevant interactions of X-rays with matter, including Rayleigh scattering (left), the photoelectric effect (middle) and Compton scattering (right).

2.1.2 Rayleigh Scattering

As illustrated in Figure 2.2, Rayleigh scattering describes an event where an X-ray photon is scattered off a bound electron such that no photon energy is lost [1], [2], [105], [143]. It is therefore also referred to as elastic or coherent scattering. Although no energy is transferred, the direction of the photon path changes slightly. As evident in Figure 2.1, Rayleigh scattering is more important for low energies and heavy atoms. The cross-section is proportional to

$$\sigma_R \propto Z^2/E^2, \quad (2.4)$$

where E is the photon energy and Z is the atomic number. While Rayleigh scattering is relevant for very dense materials, as seen in Figure 2.1, in the diagnostic range it is the least important effect for water.

2.1.3 Compton Scattering

If the scattering process transfers energy to the electron, the process is called Compton scattering, or inelastic scattering. In contrast to Rayleigh scattering, where the photon interacts with a bound electron, Compton scattering involves a quasi-free electron, typically a valence electron in an outer shell [1], [2], [105], [143]. As shown in Figure 2.2, the photon path is changed by a scattering angle, and the electron leaves the atomic shell. Due to the energy transfer, the photon energy after scattering is

$$E' = E \frac{1}{1 + \frac{E}{m_e c} (1 - \cos \vartheta)}. \quad (2.5)$$

Here, ϑ is the scattering angle and m_e is the electron mass. The total cross-section is

$$\sigma_C \propto \frac{Z}{\sqrt{E}}. \quad (2.6)$$

Figure 2.1 shows that in the diagnostic range, Compton scattering is the dominant interaction for water. In addition, the attenuation coefficient is significantly more constant over the energy range than for Rayleigh scattering or the photoelectric effect. Thus, Compton scattering primarily provides low-contrast information [2].

2.1.4 Photoelectric Effect

The third relevant photon-matter interaction is the photoeffect. Here, the X-ray photon is fully absorbed by the atom, removing an electron as visualized in Figure 2.2 [1], [2], [105], [143]. A lower shell now has a vacancy, which is energetically unfavorable. Thus, an electron from an outer shell will drop into the lower shell under emission of a photon. Since the energy of the emitted photon is equal to the energy difference between the two levels, it is called characteristic photon. Likewise, the photoelectric absorption requires photon energies above the binding energy of the respective shell. Thus, the graph of the attenuation coefficient has characteristic edges, as seen in Figure 2.1 for lead. The cross-section of the photoelectric absorption follows

$$\sigma_P \propto Z^4/E^3. \tag{2.7}$$

Due to the Z dependence, the photoelectric effect dominates the total attenuation for high- Z materials.

2.2 Data Acquisition

The first step of data acquisition in CT is the generation of X-rays. Figure 2.3 illustrates an X-ray tube. A cathode produces electrons, which are accelerated towards an anode target. A magnetic field deflects the electrons to achieve the correct path. When electrons hit the anode, they create bremsstrahlung and characteristic radiation. Most of the energy, about 99%, is wasted as heat [2]. Therefore, the target must have a high melting point and is cooled during the scan. For the same reason the most common anode material is tungsten. Another benefit of tungsten is its high atomic number, which increases the cross-section of the photoelectric effect, as presented in the previous section. This leads to better tube performance due to an improved X-ray efficiency [1]. After tube generates an X-ray beam, it is collimated and pre-filtered to adjust the shape and spectrum. This reduces the patient dose, which allows for higher image quality by having a larger fraction of X-rays contribute to the image formation [1], [23]. In addition, changing the spectrum also alters contrast between tissues in the reconstructed images. After pre-filtration, the beam passes through the patient. Parts of the beam are attenuated according to section 2.1. The remaining X-rays reach the detector, which converts the X-rays into an electronic signal. In a conventional energy-integrating detector, the X-rays first interact with a scintillator, commonly $\text{Gd}_2\text{O}_2\text{S}$ [144]. This produces photons in the visible spectrum that are subsequently measured in a photo diode. After digitization this yields an intensity value I . The projection value p is finally calculated by solving Equation (2.1) for $\mu \cdot d$, i.e. dividing by a pre-calibrated

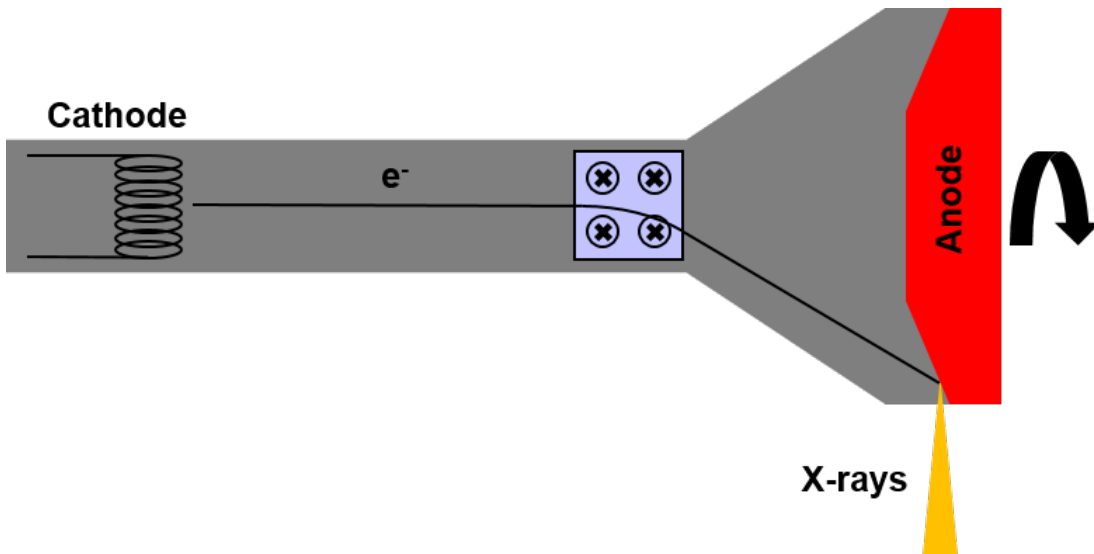


Figure 2.3: Schematic of an X-ray tube. The anode is rotating for better heat dissipation.

intensity without patient I_0 and taking the negative logarithm. This value is finally saved for each rotation angle and detector pixel in the so-called sinogram.

2.3 Image Reconstruction

After the CT data are measured, a reconstruction algorithm is required to produce slice images from the projection data. In order to adequately sample the patient, a 2D scan needs to cover 180 degrees. Besides the standard filtered back-projection (FBP), several more sophisticated reconstruction techniques are in use, as illustrated in chapter 1. This primarily includes iterative reconstruction and state-of-the-art deep learning algorithms [2], [29]. In view of the multitude of algorithms and different scan geometries, only the analytic FBP will be derived here with a parallel beam geometry. In current clinical practice, fan- and cone-beam geometries with spiral trajectories are prevalent [1]. The algorithms used for these more complex geometries are often related to the FBP.

Figure 2.4 illustrates the geometry used in this study. Each pixel in the sinogram, i.e. each projection value p , is associated with its position along the detector ξ and the rotation angle of the X-ray source ϑ . The projection p is thus an integral along the X-ray beam, which can be described by a line with $x \cos \vartheta + y \sin \vartheta = \xi$. Ignoring effects such as scatter, the projection value is consequently related to the attenuation of the patient μ as follows:

$$p(\vartheta, \xi) = \int \mu(x, y) \delta(x \cos \vartheta + y \sin \vartheta - \xi) dx dy, \quad (2.8)$$

where $\delta(\cdot)$ is the Dirac delta function. The X-ray transform follows as the sum of all possible line integrals [105]. For 2D acquisitions, the X-ray transform is also referred to

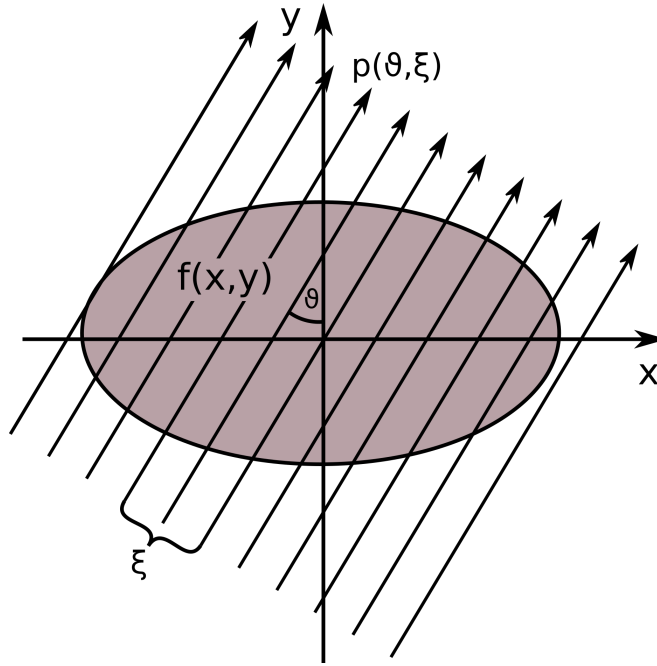


Figure 2.4: 2D parallel beam geometry for the filtered backprojection

as Radon transform [2]. Equation (2.8) describes the forward process in CT. However, the desired variable is the attenuation distribution μ , as this constitutes the cross-sectional image. Therefore, the equation must be inverted. This is possible with the Fourier-Slice theorem, defined as

$$P(\vartheta, u) = M(u \cos \vartheta, u \sin \vartheta). \quad (2.9)$$

Here, $P(\vartheta, u) = \mathcal{F}p(\vartheta, \xi)$ is the Fourier transform of the sinogram and similarly $M(u_x, u_y) = \mathcal{F}\mu(x, y)$ is the Fourier transform of the image. In the above equation, it holds that $u_x = u \cos \vartheta$, $u_y = u \sin \vartheta$ and $du_x du_y = |u| du d\vartheta$. Although it seems that an inverse Fourier transform of M would directly lead to the desired image, the discrete angular sampling of CT would create artifacts. Instead, using Equations (2.8) and (2.9) together gives the FBP:

$$\begin{aligned} \mu(x, y) &= \int_0^\pi d\vartheta \int_{-\infty}^{\infty} |u| P(\vartheta, u) e^{2\pi i u (x \cos \vartheta + y \sin \vartheta)} du \\ &= \int_0^\pi d\vartheta \int_{-\infty}^{\infty} K(u) P(\vartheta, u) e^{2\pi i u \xi} du. \end{aligned} \quad (2.10)$$

The ramp filter $K(u) = |u|$ is required to prevent the aforementioned artifacts. The expression can be further reduced with the convolution theorem:

$$\mu(x, y) = \int_0^\pi d\vartheta p(\vartheta, \xi) * k(\xi)|_{\xi=x \cos \vartheta + y \sin \vartheta}. \quad (2.11)$$

The kernel $k(\xi)$ that is convolved with the sinogram follows as

$$k(\xi) = \mathcal{F}^{-1}K(u) = \frac{-1}{2\pi^2\xi^2}. \quad (2.12)$$

The discretized version of this kernel, under consideration of the Nyquist frequency of the data, is known as the Ram-Lak kernel [1]. On commercial systems, a variety of vendor-specific kernels are in use. This way, characteristics of the image can be adjusted in the reconstruction process. For instance, a sharper kernel will lead to higher resolution at the cost of higher noise. Generally, the attenuation values μ are not used for diagnosis. Instead, they are re-scaled in reference to the attenuation value of water and presented in Hounsfield units, defined as

$$\text{CT}(x, y) = \frac{\mu(x, y) - \mu_{\text{Water}}}{\mu_{\text{Water}}} \cdot 1000 \text{ HU}. \quad (2.13)$$

By definition, water has a value of 0 HU. Other values include air at -1000 HU, fat at -80 HU, liver at 50 HU, spongy bone around 200 HU and cortical bone between 350 and 2000 HU [1]. CT values are typically represented as 12 Bit integers with an offset of -1024 HU. Consequently, highly attenuating materials such as metals may have values above the maximum value of 3071 HU [105]. In non-medical applications, where water is not relevant, Hounsfield units (HU) values are not in use. Instead, the attenuation values themselves are used, or re-scaled to the dominant material [145].

2.4 Image Artifacts

CT image reconstruction heavily revolves around avoiding and correcting image artifacts. In the context of low-dose CT, image noise and sparseness artifacts are most relevant. Other artifacts include beam hardening, scatter, motion artifacts, metal artifacts, and truncation [105]. The latter is especially relevant in applications using CBCT, as detectors typically are not able to cover the full patient and is the secondary focus of this work.

2.4.1 Noise

In order for an X-ray to contribute to the sinogram, it must pass through the patient and interact with the detector. However, as described in section 2.1, the photon-object interactions are associated with a probability. As a consequence, the number of X-ray photons that reach the detector follows a Poisson distribution [1], where the probability of an individual photon to be measured is

$$\left(1 - e^{-\mu_{\text{D}}(E)d_{\text{D}}}\right) e^{-p(E)}. \quad (2.14)$$

Here, $\mu_{\text{D}}(E)$ is the attenuation coefficient of the detector material at X-ray energy E , d_{D} is the intersection length of photon and detector, and $p(E)$ is the line integral of

the attenuation of the patient along the X-ray beam. The detected signal S is now equivalent to

$$S = \int dE a(E)N(E), \quad (2.15)$$

where $a(E)$ and $N(E)$ are the detector response and number of photons at energy E , respectively. Typically, $a(E) \propto E$. The attenuation value is defined as $Q = -\ln S/ES_0$, where ES_0 is the expected value of the detected signal without patient. Due to the Poisson characteristic of detected X-rays, Q can be estimated as

$$EQ \approx -\ln \frac{ES}{ES_0} \quad (2.16)$$

and

$$\text{Var}Q \approx \frac{\text{Var}S}{(ES)^2}, \quad (2.17)$$

where Var signifies the variance. Due to the linear nature of FBP, the projection noise will linearly propagate to the image domain [1], [2], [105]. Under the assumption of a monochromatic X-ray beam, the equations simplify to

$$EQ \approx -\ln \frac{EN}{EN_0} \quad (2.18)$$

and

$$\text{Var}Q \approx \frac{1}{EN_0} e^{EQ}. \quad (2.19)$$

As evident, the variance is indirectly proportional to the number of X-ray photons before the patient. Decreasing the tube current, and therefore the dose, by a factor of two will consequently lead to an increase of image noise to $\sqrt{2} = 141\%$ of the full dose image. Figure 2.5 illustrates how the image noise increases with decreasing N_0 , i.e. tube current. In addition, due to the exponential term, the noise will further increase for obese patients. Another source of noise is the electronic noise of the detector element itself [37].

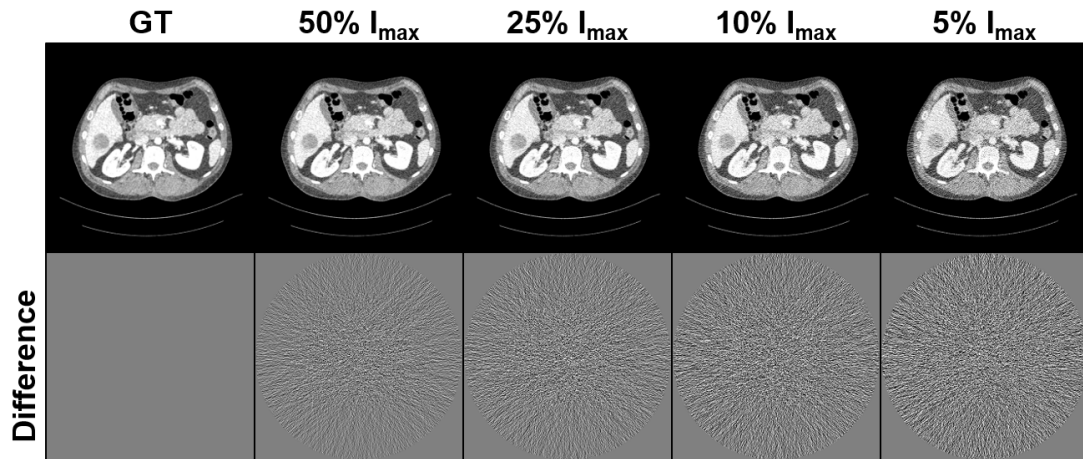


Figure 2.5: Examples of increasing image noise with decreasing number of photons relative to full dose. Bottom row shows difference to ground truth. $C = 0$ HU, $W = 500$ HU for CT images, $C = 0$ HU, $W = 100$ HU for difference images.

2.4.2 Sparseness

According to Equation (2.11), a CT image can be reconstructed from the integral of all rays passing through the patient over an angle of 180 deg. However, due to the discretization of CT system and reconstruction software, only a finite number of line integrals are available. Thus, the integral in Equation (2.11) becomes a sum [2]. Provided that the number of projections is high, this will not impact the image. However, if the number of projections is reduced below a certain limit, the patient is insufficiently sampled in the angular direction. Then, the image reconstruction presents an ill-posed inverse problem and conventional algorithms such as FBP cause streak artifacts in the image [58], [105]. Figure 2.6 provides an example of sparseness artifacts increasing with decreasing number of projections. Interestingly, reducing the projections also increases image noise, as fewer rays are averaged for each pixel during the backprojection process.

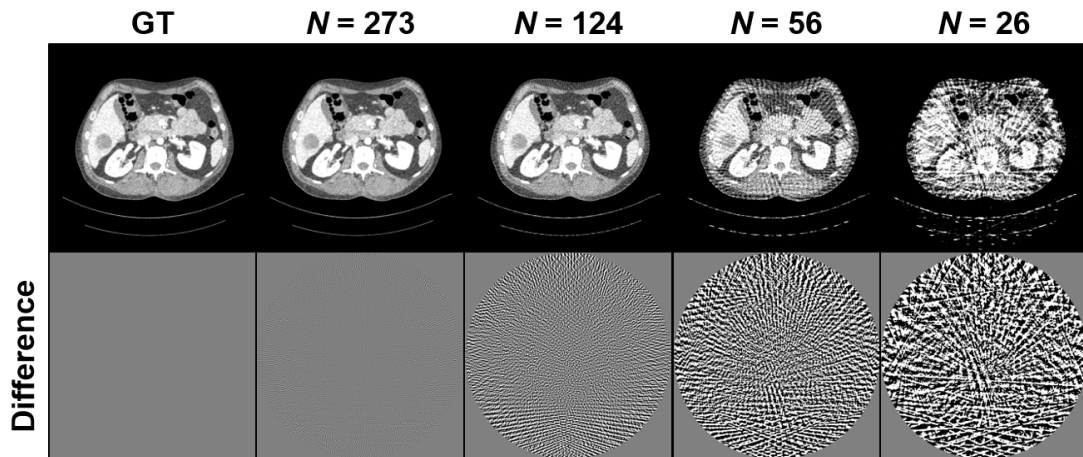


Figure 2.6: Examples of increasing sparseness artifacts with decreasing number of projections. Bottom row shows difference to ground truth. $C = 0$ HU, $W = 500$ HU for CT images, $C = 0$ HU, $W = 100$ HU for difference images.

2.4.3 Truncation

If parts of the patient are not fully inside the FOM, the projections will be insufficient for conventional reconstruction. Firstly, the FOV will be reduced in size. Secondly, truncation artifacts appear as hyperdense areas close to the edge of the FOV [1]. Depending on the amount of truncation, CT values accuracy deteriorates inside the whole image. Figure 2.7 presents reconstructions with varying number of equally-sized detector pixels. Note that values outside of the FOM are set to zero. To prevent truncation artifacts, the sinogram is extrapolated to simulate a laterally larger detector. With conventional algorithms, only the data inside the field of measurement can be returned to high image quality. For data outside of the field of measurement, more sophisticated methods are required [1].

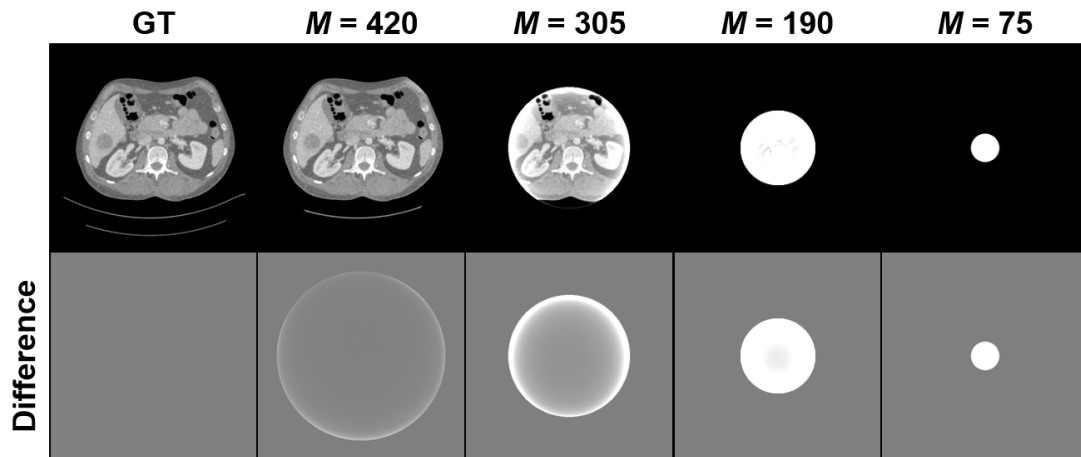


Figure 2.7: Examples of increasing truncation artifacts with decreasing number of detector pixels. Bottom row shows difference to ground truth. $C = 0$ HU, $W = 1000$ HU for CT images, $C = 0$ HU, $W = 1000$ HU for difference images.

2.5 Neural Networks

neural network are a major component of the current machine learning landscape. In contrast to traditional methods, where the relationship between input and output is hardcoded into the algorithm, neural networks learn this transformation from the data [56]. Currently, NNs produce state-of-the-art results in many image processing tasks [57]–[60].

2.5.1 Neurons

Neurons, as in the human brain, are the foundational unit of NN processing. A single neuron is defined mathematically as

$$a' = \sigma \left(\sum_k a_k w_k + b \right), \quad (2.20)$$

where a_k are some real input values, and a' is the output. Together with the weights w_k and bias b , which are later determined during the training process, the inputs are linearly transformed. In order to be able to model non-linear transformations, the so-called activation function σ must be a non-linear function itself [56]. One popular activation function is the rectified linear unit (ReLU) [146], [147], which is defined as

$$\sigma(x) = \max(0, x). \quad (2.21)$$

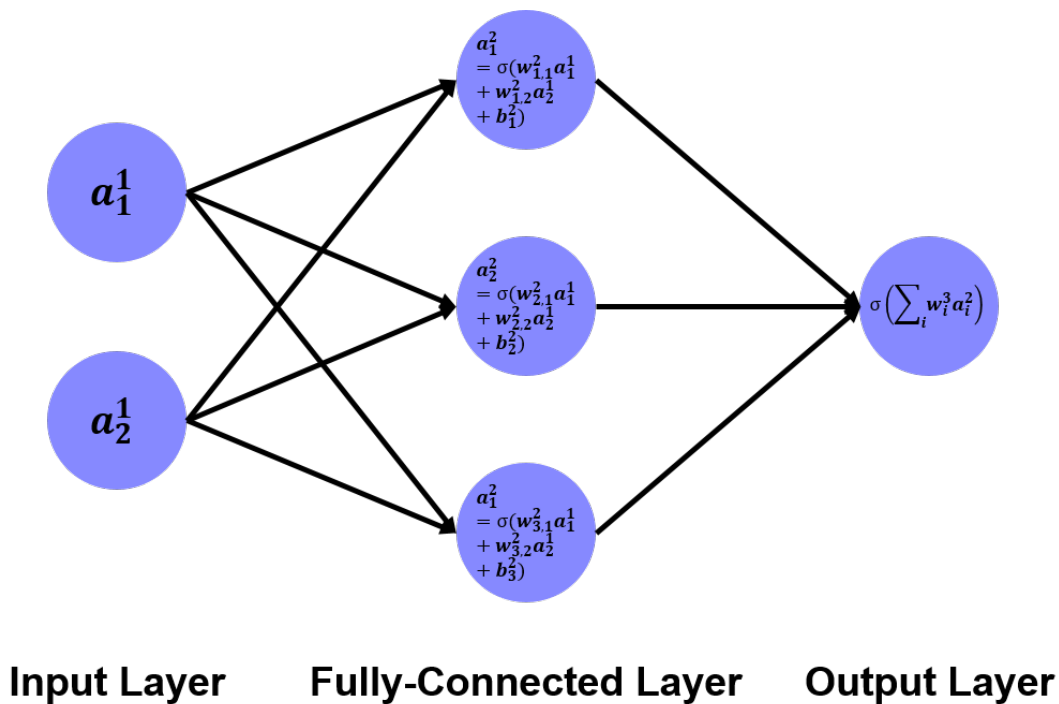


Figure 2.8: Fully-connected layer with two inputs, three neurons, and one output.

2.5.2 Layers

A single neuron has not enough computational power to perform the desired processing tasks. Therefore, neurons are arranged in layers, such that the outputs of neurons of the first layer will be used as input for neurons of the next layer and so forth. How exactly the layers are connected is an important aspect of network design. Two of the most common layers are fully-connected (FC) layers and convolutional layers.

i.) Fully-Connected Layers

As the name suggests, in a FC layer, each neuron is connected to all neurons of the previous layer [56]. Figure 2.8 illustrates a simple neural network with two inputs, an FC layer with three neurons, and one output. Note that the middle layer is also referred to as hidden layer, as the user only interacts with the input and output layers. For large numbers of neurons and layers, this type of layer will create a significant computational cost. Increasing the number of neurons by a factor of two for two FC layers will quadruple the number of weights between them. Therefore, this type of layer is most commonly used for low-dimensional inputs. These may occur in low-resolution imaging or after other network layers have reduced the dimensionality [56], [57], [59], [132]. For example, classification tasks usually include FC layers at the end of the network to produce the final estimates.

ii.) Convolutional Layers

In convolutional layers, each neuron only sees a small subset of neurons from the previous layer, significantly reducing computational costs. Here, the weights are employed as elements of a convolution kernel, such that the i th neuron of layer j is defined as

$$a_i^j = \sigma \left(\sum_{k=-N}^N a_{i+k}^{j-1} w_k^j + b^j \right), \quad (2.22)$$

where the input neurons k only include those within the convolution kernel of size $2N + 1$, given by w_k . Figure 2.9 visualizes a one-dimensional (1D) convolution layer with a 3×1 kernel. In order to retain the tensor size, the input layer is padded with zeros. Otherwise, the output tensor will have a reduced size depending on the stride and kernel size of the convolution layer. As convolutions are already omnipresent in traditional image processing, it is not surprising that convolutional layers are used extensively for deep learning-based approaches as well. Networks based on convolutional layers are called convolutional neural network. Equation (2.22) can be adapted to any input dimension, e.g. 2D images and 3D volumes, by converting the 1D convolution to higher dimensional equivalents. For inputs with multiple channels, for instance color images or outputs from deeper network layers, one would typically use one kernel for each combination of input and output channel [56], [57].

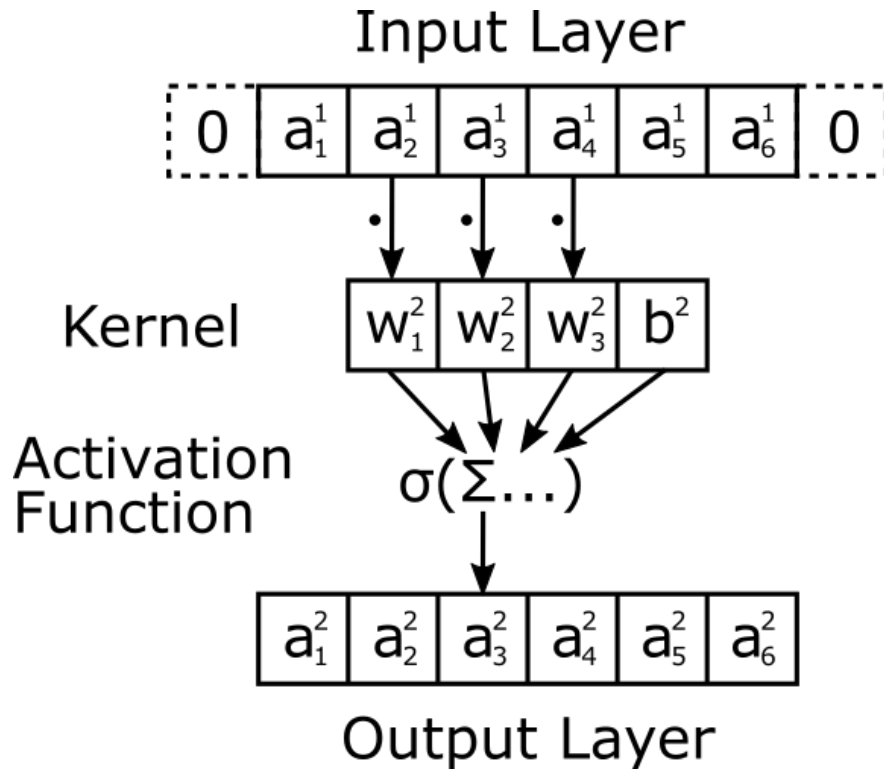


Figure 2.9: Schematic of a 1D convolutional layer. The input layer is zero-padded to ensure that the output is the same size.

2.5.3 Network Training

As stated above, neural networks determine their internal parameters from the data. Since convolutional neural networks (CNNs) can have billions of weights and biases to be calculated, the so-called training process cannot be performed manually. The most common approach is supervised training [57]. Here, the network is provided with inputs, e.g. noisy images, and corresponding desired target, e.g. clean images. Since each image is perfectly sorted into either category, they are called labeled. In principle, the training consists of three steps. First, the input is fed into the network and processed. Secondly, the difference to the desired target is computed according to some loss function. The loss function has significant impact on the training result, as it defines what constitutes an accurate network output. A common choice for image correction tasks is mean squared error (MSE) [57], [58]. Thirdly, the network parameters are updated depending on the value of the loss function. This is typically performed via backpropagation or gradient-descent optimization, for instance the Adam optimizer [148], [149]. Usually, the network weights are updated after a small number of samples, called batches, have passed through the network. After all training data have been used, one epoch concludes. This is ideally followed by evaluating the network on validation data, which have been excluded from training. The validation step ensures that the network is

capable of performing well on unseen data. It also aids in determining hyperparameters of the network training [56], [57]. The training process is finally stopped after some predetermined amount of epochs has passed, the improvement becomes too small, or a target metric is reached. Although supervised training is conceptually the simplest form of network training, it requires accurate input-target-pairs. If these cannot be generated, e.g. due to unrealistic simulations, more sophisticated training schemes are required [57], [150], [151].

3 | Materials & Methods

3.1 Low-Dose CT Correction

3.1.1 Simulating Low-Dose CT Data

In order to train a CNN for low-dose CT denoising and destreaking, a large training dataset is required. Ideally, the data consists of real low-dose and corresponding high-dose scans. While this is possible with phantom measurements, performing multiple acquisitions is infeasible for patients due to dose concerns and patient movement. Nevertheless, supervised learning is still possible as low-dose scans can be simulated from diagnostic CT volumes. As an alternative to using labeled data, several unsupervised methods have been proposed in the literature, as discussed in section 2.4.

This work uses the publicly available Liver Tumor Segmentation (LiTS) Benchmark dataset, which contains 130 CT volumes with corresponding liver and lesion segmentation [66]. Figure 3.1 provides examples of the LiTS dataset. So far, there have been three LiTS challenges, namely at International Symposium on Biomedical Imaging (ISBI) 2017, Medical Image Computing and Computer Assisted Intervention Society (MICCAI) 2017, and MICCAI 2018. The first two were dedicated to lesion segmentation, and joint liver and lesion segmentation, respectively. The latter was part of a larger image segmentation challenge, including nine other tasks. Various state-of-the-art segmentation methods have been developed for and tested on the dataset [64]–[66]. Other large public and private datasets exist and were considered. Ultimately, the LiTS data were chosen due to the availability of high-quality pretrained segmentation networks for task-based image quality assessment, as described in section 3.2.2. Besides the 130 volumes with ground truth segmentation, the dataset includes an additional 70 volumes without segmentation intended for the final challenge benchmark. These volumes were excluded from this study, as the ground truth segmentation is necessary for testing. Note that several institutions and radiologists were involved in the creation of the dataset, such that the detailed look of the segmentations differs from patient to patient. In Figure 3.1 this is evident in the jagged liver segmentation for the leftmost case.

The LiTS scans typically feature a pixel size of around 0.8 mm. Although the dataset is for lesion segmentation, the volumes are mostly full body CT acquisitions ranging from hip to shoulders. Since only image data is available, the corresponding raw data must be generated artificially. A monochromatic forward projection of the CT volumes yields the ground truth sinograms. The forward projection is performed in parallel beam

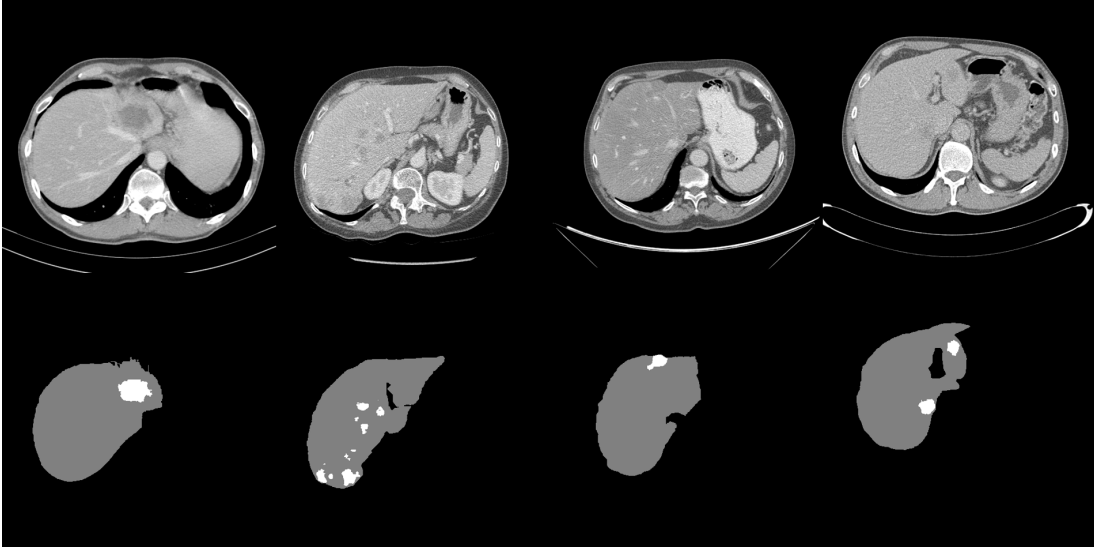


Figure 3.1: Example data from the LiTS dataset. Top row shows the CT images, bottom row the corresponding segmentation of liver (grey), and lesions (white). $C = 0$ HU, $W = 500$ HU.

geometry with $N_{\max} = 512$ equally-spaced projections covering 180° . Each projection contains $M = 512$ detector pixels 0.8 mm apart.

Under the assumption that the ground truth images do not contain a significant amount of noise, low-mAs scans were simulated by directly adding Poisson noise to the sinograms. This models the Poisson behavior of photons being detected by an X-ray detector, as described in section 2.4.1. Electronic noise is not simulated. With a photon number of I_0 , the noisy projections $p_{\text{low-mAs}}$ are calculated as

$$p_{\text{low-mAs}}(\theta, \xi) = -\ln(\text{Poisson}(I_0 \exp p(\theta, \xi))/I_0), \quad (3.1)$$

where $p(\theta, \xi)$ is the sinogram value at rotation angle θ and detector position ξ .

For sparse-view CT, the number of projections N were reduced by the appropriate factor. Each instance of sparse-view CT was generated with a new forward projection, as the angular spacing is not always an integer multiple of the non-sparse angular increments.

To combine sparse-view CT with low-mAs CT, apply Equation (3.1) to the sparse sinogram. Since the dose should remain the same for all instances of low-dose CT, the total number of photons must be identical. With the simulation parameters N and I_0 , this equates to

$$D \propto I_0 N, \quad (3.2)$$

where D is the patient dose. Given a desired dose reduction by a factor α , the low-dose

CT with photon number $I_{0,\text{LD}}$ and number of projections N_{LD} fulfills

$$\begin{aligned} D_{\text{LD}} &= \frac{1}{\alpha} D \propto \frac{1}{\alpha} I_0 N = \frac{I_0}{\alpha_I} \frac{N}{\alpha_N} = I_{0,\text{LD}} N_{\text{LD}}, \\ I_{0,\text{LD}} &= \frac{I_0}{\alpha_I}, \\ N_{0,\text{LD}} &= \frac{N}{\alpha_N}, \end{aligned} \tag{3.3}$$

where α_I and α_N are the factors reducing the tube current and number of projections, respectively. Since the total dose reduction factor is always equal to α , regardless of dose implementation, it holds that

$$\alpha = \alpha_I \alpha_N. \tag{3.4}$$

Evidently, α_I and α_N are inversely proportional. Figure 3.2 visualizes the relationship between tube current and number of projections at equal dose. Therefore, to have fairly spaced pairs of $\alpha_{I,N}$, they are sampled as

$$\alpha_I = \frac{\alpha}{\alpha_N} = e^{\frac{k}{K-1} \ln(\alpha)}, \text{ and } \alpha_N = e^{(1-\frac{k}{K-1}) \ln(\alpha)}, \tag{3.5}$$

where $0 \leq k \leq K - 1$, $k, K \in \mathbb{Z}$ determines what fraction of the dose reduction is achieved by reducing the tube current, and K is the number of dose implementations. Figure 3.3 plots the trade-off between number of projections and tube current on a log-log scale.

This study investigated a dose reduction 80%, i.e. $\alpha = 5$, with five different realizations of low-dose CT. The high-dose photon number of the ground truth scans I_{max} was set to 1.5×10^6 photons and the number of projections to $N = 512$. These parameters serve to generate high-dose images, which can also be compared to the ground truth. To simulate a dose reduction of 80%, low-dose images were generated according to Equation (3.5). This yields $N \in \{512, 342, 229, 153, 102\}$ and $I_0 \in I_{\text{max}}\{0.2, 0.30, 0.45, 0.67, 1.0\}$ as highlighted in Figure 3.3.

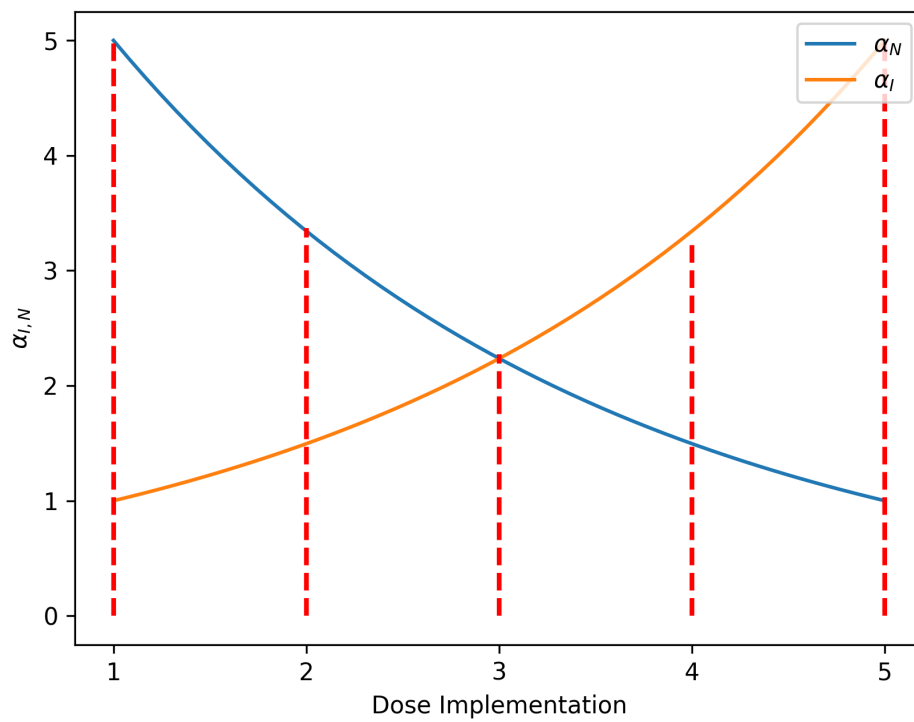


Figure 3.2: Relationship between number of projections and tube current at constant dose for different dose implementations. Red dashed lines mark the tested configurations.

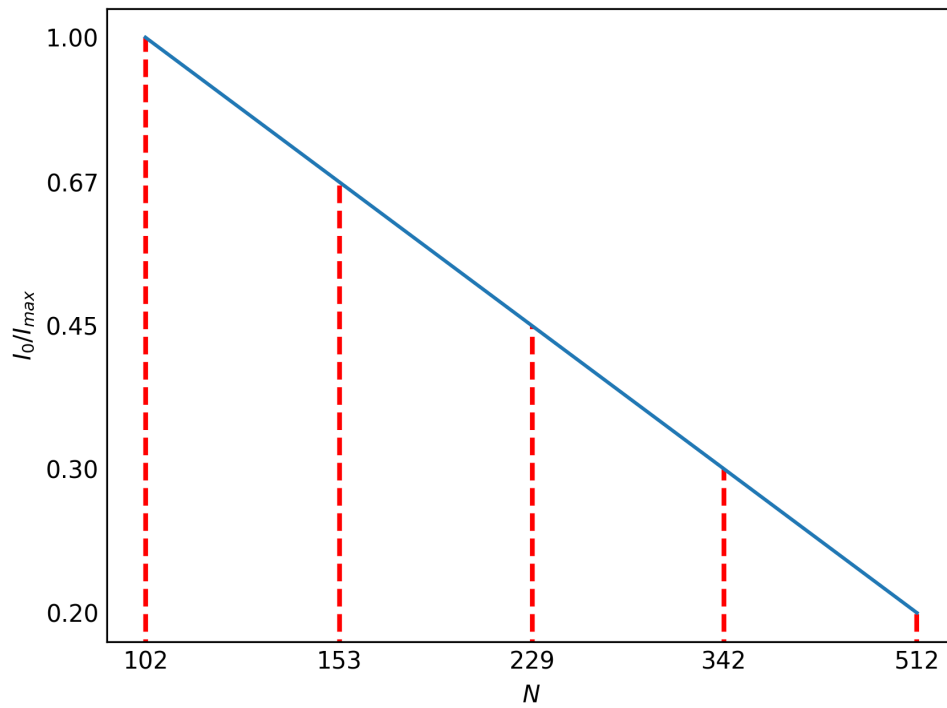


Figure 3.3: Tube current vs. number of projections for the five tested dose implementations on a log-log scale. Red dashed lines mark the tested configurations.

3.1.2 CNN Denoising

i.) Network Architecture

This study uses a W-Net to correct the low-dose images. A similar architecture was previously employed for sparse-view CT reconstruction [87]. Figure 3.4 describes the W-Net architecture. It consists of three components: First, a sinogram domain residual U-Net (U-Net_{Sino}) [152] corrects sparseness artifacts and noise in the sinogram. Secondly, a differentiable FBP layer reconstructs the image. This layer uses a Ram-Lak kernel for the filtering step. Note that in [87], the FBP layer is also trainable. Finally, an image domain residual U-Net (U-Net_{Img}) removes leftover artifacts and image noise. Figure 3.5 illustrates the U-Net architecture in more detail. The architecture is 2D and identical for both the sinogram and image domain network. Note that the different low-dose CT realizations feature differently sized sinograms. In addition, image correction is easier for networks to learn than image generation. Thus, the low-dose images are forward projected in the full-view geometry to produce the input sinograms. Consequently, the network corrects the full, but corrupted sinograms.

ii.) Network Training

The training of the W-Net combines three loss functions. Let S_N , S_{GT} , and f_S be the noisy sinogram, ground truth sinogram, and sinogram domain U-Net. Then, the sinogram loss is given by

$$\mathcal{L}_S = \|f_S(S_N) - S_{GT}\|^2, \quad (3.6)$$

i.e. the MSE of network output and ground truth. Similarly, the FBP loss is

$$\mathcal{L}_{\text{FBP}} = 0.9995\|f_{\text{FBP}}(S_S) - I_{\text{GT}}\|^2 + 0.0005\mathcal{L}_{\text{Perc}}(f_{\text{FBP}}(S_S), I_{\text{GT}}), \quad (3.7)$$

where f_{FBP} is the reconstruction layer, I_{GT} the ground truth image and $S_S = f_S(S_N)$ the output of the sinogram network. In the above equation, the second part of the loss function is perceptual loss. Perceptual loss is calculated by feeding the ground truth and test image through a pre-trained VGG-19 network [153] and extracting the features of the second convolution layer in second and third block [75], [78], [126], [154]. Then, the MSE is computed between ground truth and test image over the concatenated output features. For the perceptual loss, the images are windowed to $C = 100$ HU and $W = 1000$ HU and scaled to the range [0 1]. The last loss function is

$$\mathcal{L}_I = 0.9995\|f_I(I_{\text{FBP}}) - I_{\text{GT}}\|^2 + 0.0005\mathcal{L}_{\text{Perc}}(f_I(I_{\text{FBP}}), I_{\text{GT}}), \quad (3.8)$$

where f_I is the image domain U-Net, and I_{FBP} is the output of the FBP layer. The network is trained end-to-end, therefore the loss functions are combined as

$$\mathcal{L}_{\text{tot}} = \alpha_S\mathcal{L}_S + \alpha_{\text{FBP}}\mathcal{L}_{\text{FBP}} + \alpha_I\mathcal{L}_I, \quad (3.9)$$

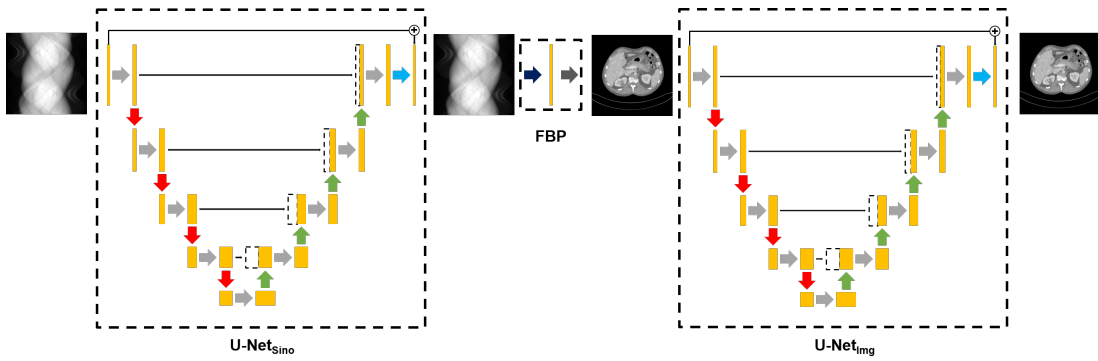


Figure 3.4: Network architecture of the W-Net used for denoising of the low-dose images. The FBP module consists of a convolution with the Ramlak kernel and a differentiable backprojection layer. For a detailed description of the U-Nets, see Figure 3.5

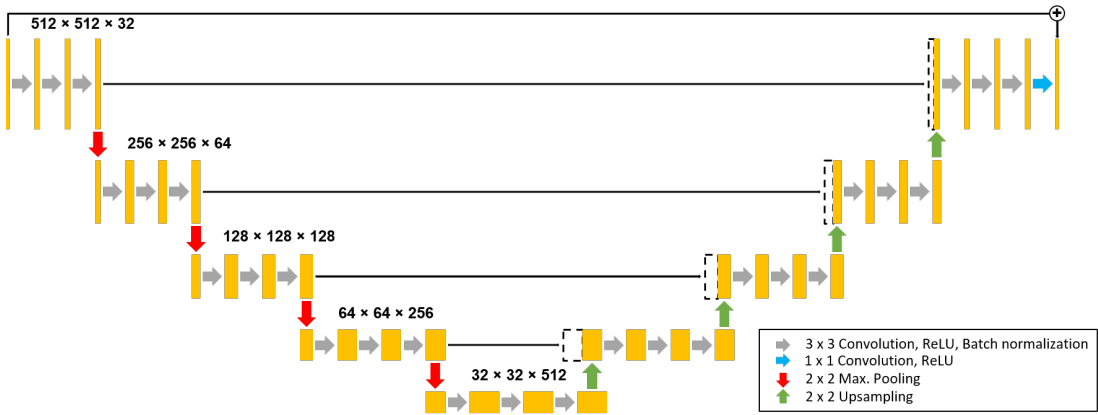


Figure 3.5: Detailed architecture of the U-Nets for denoising the sinogram and images. Input and output are connected with a residual connection.

where $\alpha_{\text{FBP}} = \alpha_{\text{I}} = 1$ and $\alpha_{\text{S}} = 0.001$ are set empirically.

To train the networks, the paired low- and high-dose images based on the LiTS dataset are split into training, validation and test set. This study uses 90 patients for training, 20 for validation and 20 for testing. The networks are trained for 20 epochs using an Adam optimizer [148], [149] with $\beta_1 = 0.9$, $\beta_2 = 0.999$ and a learning rate of 0.005. All models are implemented in Tensorflow and trained on an NVIDIA Geforce GTX 2080Ti.

iii.) Insufficiently Trained Networks

One critical aspect of NNs is that the network mostly functions as a black box [57], [132]. The task of quality assurance (QA) therefore becomes even more important than for conventional, analytical algorithms. When CNNs are designed for clinical applications,

the necessity of QA is further exacerbated. This study investigates if the methods of image quality assessment described in section 3.2 can be utilized to identify networks that were insufficiently trained. Networks can be trained insufficiently in a plethora of ways. One common issue is when the training data belong to a different domain or distribution than the test data. For example, training on phantom data while testing on patient data. The problem also occurs when the simulation process cannot produce images that are accurately mimicking the real data [132], [151]. In CT, this is often the case with deep learning-based metal artifact reduction (MAR), as metal artifacts are highly complex. Some methods therefore include a domain adaption step, to adjust the CNN to the real data [155]. If the training data is drawn from the correct distribution, but the set of training data is too small, the network will overfit on the given data. Then, it will not generalize to the overall distribution, i.e. produce suboptimal results on the test data. Finally, when the number of epochs is too small, the training process will not find the optimal network weights, leading to a decrease in overall performance [132]. This study artificially reduces the quality of the network training by either reducing the size of the training set or reducing the number of epochs.

3.2 Image Quality Assessment

3.2.1 Conventional Analysis

For quantitative analysis, calculate root mean squared error (RMSE) and structural similarity index metric (SSIM) of the test set compared to the ground truth images. RMSE is defined as

$$\text{RMSE}(x, y) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2}, \quad (3.10)$$

where n is the number of pixels, y is the tested image, and x is the ground truth. For SSIM, this work uses the mean local SSIM, defined as

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (3.11)$$

where μ , σ , and σ_{xy} are the local mean, standard deviation, and cross-covariance, respectively. The constants are set to $C_1 = (0.01 * L)^2$ and $C_2 = (0.03 * L)^2$, with the dynamic range L . SSIM was not developed for images with negative numbers, such as CT images in the Hounsfield scale [156]. Therefore, CT images are mapped from [-400, 600] HU to the range [0, 1], i.e.

$$\hat{I} = \begin{cases} 0, & I < -400 \\ \frac{I+400}{1000}, & -400 \leq I \leq 600 \\ 1, & I > 600, \end{cases} \quad (3.12)$$

with the original image I in HU. This corresponds to a window with center of $C = 100$ HU and width of $W = 1000$ HU. This prioritizes relevant soft tissue information in the images, as well as bone-tissue boundaries. The dynamic range of the rescaled images is consequently $L = 1$.

3.2.2 Lesion Segmentation

i.) Data Preparation

To fully evaluate the image quality before and after CNN processing, conventional image quality metrics are insufficient. Instead, task-based metrics are required. One potential low-dose imaging task is liver and lesion segmentation. For this purpose, a segmentation of liver and liver lesions is performed in the CT volumes. Since the LiTS dataset was originally compiled for a segmentation challenge, there are a multitude of state-of-the-art deep learning-based approaches in the literature that are trained on these data [64]–[66]. The hybrid dense U-Net (H-DenseUNet) by Li *et al.* [64] placed first in the original challenge and is publicly available including the trained weights. Figure 3.6 illustrates the network architecture. The H-DenseUNet consists of a 2D U-Net for intra-slice features, a 3D U-Net for inter-slice features, and a hybrid feature fusion layer to obtain the final segmentation. This way, the network can leverage 3D information at relatively low computational cost. Note that the input volumes are truncated to a range of -200 to 250 HU to further emphasize the relevant CT values. The network output is further processed by hole-filling, selecting the largest connected component for the liver, and removing lesions outside the liver.

In order to calculate the segmentation metrics below, the predicted lesions must be matched to the ground truth lesions. This is done in several steps, following the algorithm by [66]. First, 3D-connected components are labeled in both prediction and ground truth to determine individual lesions. At this point, several issues might prevent a simple one-to-one mapping of prediction and ground truth. Figure 3.8 illustrates potential problems: A ground truth lesion being represented by multiple lesions in the prediction constitutes a split error. Conversely, a merge error is when multiple ground truth lesions are combined in the prediction. Finally, both problems can occur together as a split and merge error.

In order to facilitate the analysis, the metrics are not calculated per individual lesion, but for small sets of lesions. A correspondence algorithm matches lesion from ground truth and prediction. In essence, the algorithm transforms the many-to-many mapping problem to two many-to-one mappings. Figure 3.9 shows the algorithm on two examples. First, lesions in the ground truth are merged if they are overlapped by a lesion in the prediction. Here, only the largest overlap is counted for each ground truth lesion, i.e. each ground truth lesion corresponds either to no predicted lesion or to exactly one. In the second step, lesions in the predicted segmentation are merged if they overlap the same ground truth lesion. Again, only the largest overlap is taken into consideration. This ensures a one-to-one mapping of sets of lesions.

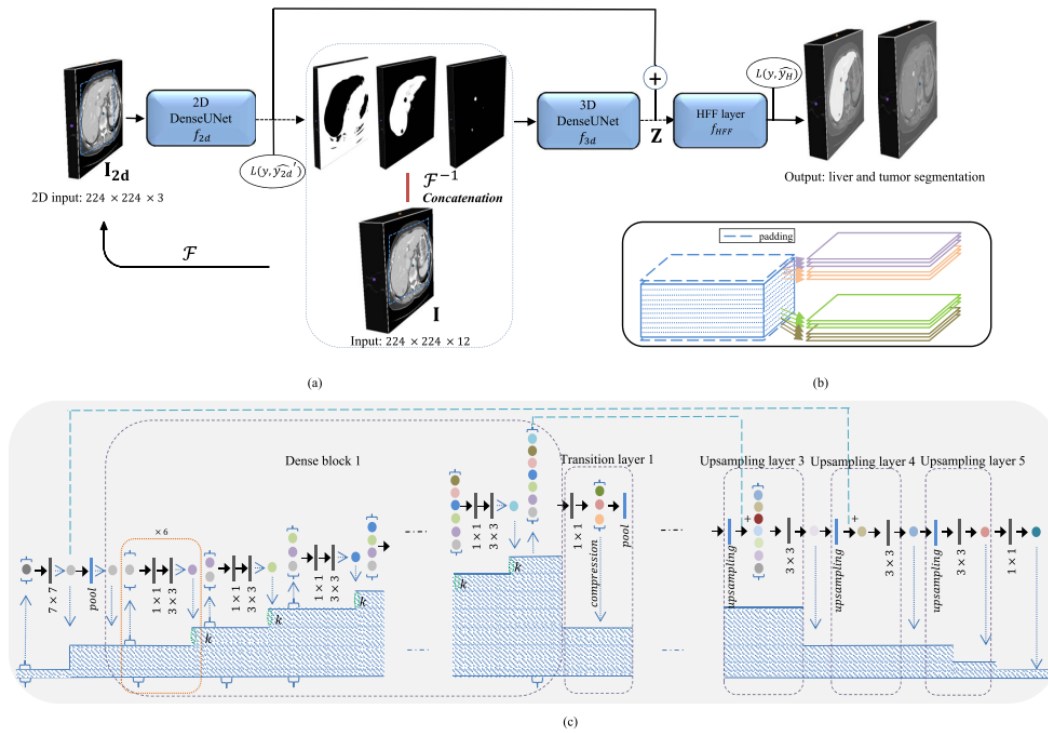


Figure 3.6: Architecture of the H-DenseUNet used for liver and lesion segmentation. From [64].

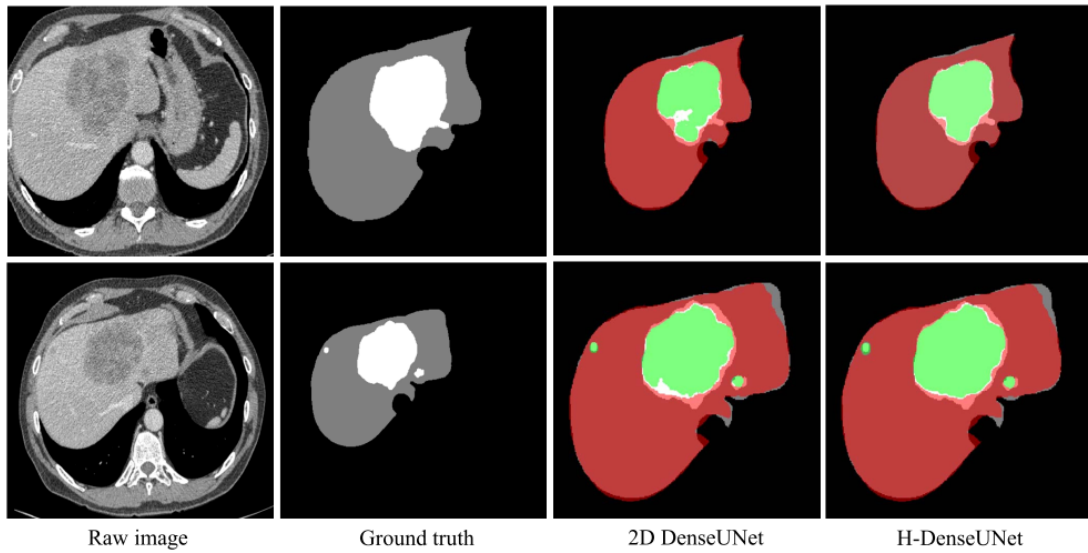


Figure 3.7: H-DenseUNet segmentations of liver and lesions on the ground truth LiTS dataset. Ground truth liver in grey, ground truth lesions in white, predicted liver in red, and predicted lesions in green. From [64].

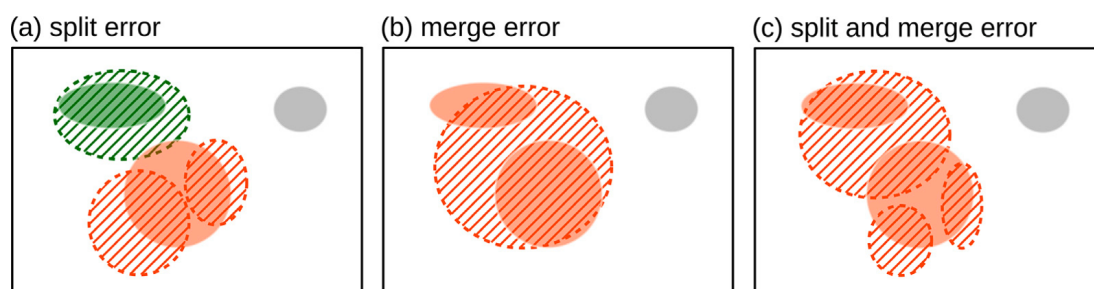


Figure 3.8: Possible issues in the mapping of predicted lesions to ground truth lesions. Solid fill indicates ground truth lesions, hatched interior a predicted lesion. Corresponding lesions have the same color. From [66].

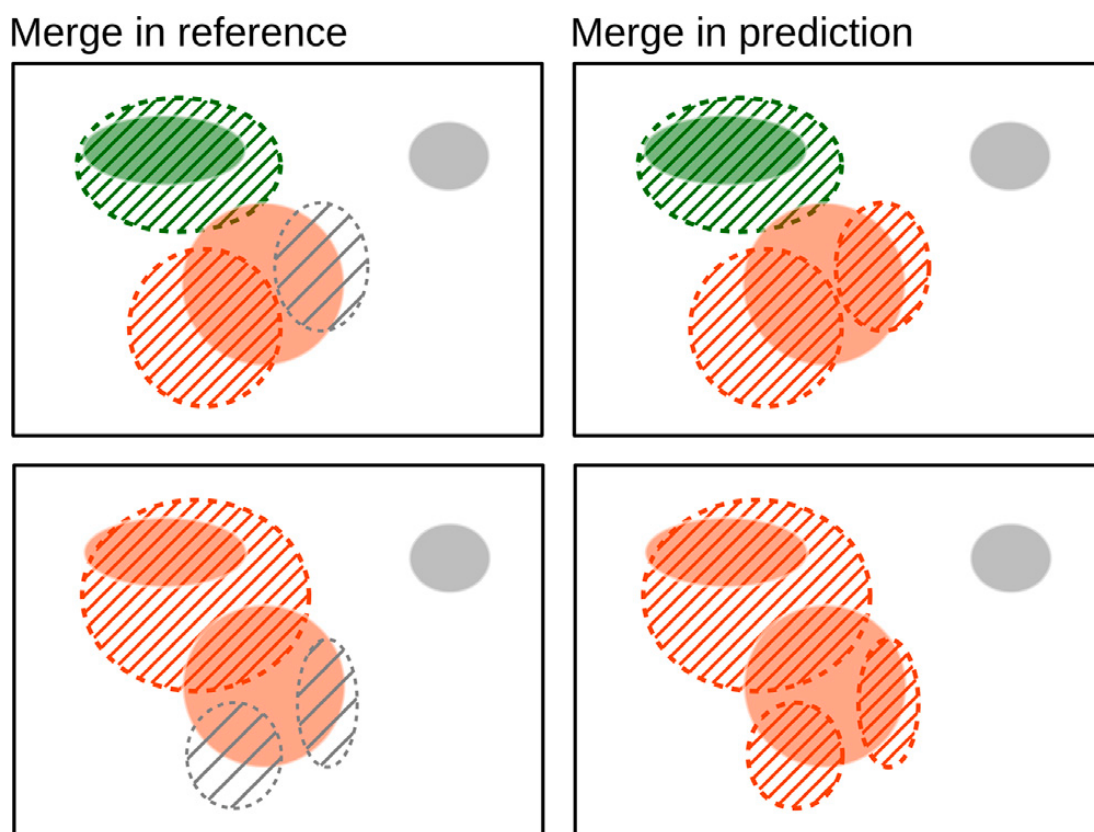


Figure 3.9: Two examples of the correspondence algorithm to find matched lesions in ground truth (solid fill) and prediction (hatched). First, the ground truth lesions are merged if a predicted lesion overlaps them. Second, predicted lesions are merged if they overlap the same (merged) ground truth lesion. Corresponding lesions have the same color. From [66].

ii.) Metrics

There is a plethora of metrics to analyze segmentation quality [157]. This work compares dice score, relative volume difference (RVD), average symmetric surface distance (ASD), root mean squared symmetric surface distance (RMSD), maximum symmetric surface distance (MSD), lesion recall, and lesion precision. The latter two are technically detectability metrics, but can be calculated due to the correspondence algorithm and are therefore included in this part of the analysis. To ensure a standardized implementation of the metrics, the publicly available LiTS evaluation code is used [66]. Dice score measures the overlap of two binary masks A and B as

$$\text{dice}(A, B) = \frac{2|A \cap B|}{|A| + |B|}, \quad (3.13)$$

where $|A|$ indicates the number of voxels in set A . For a perfect segmentation, the dice score is equal to 1. As the binary masks can be analyzed per individual lesion, per patient and per data set, there are several possible dice scores. This work investigates both per case dice, and global dice. The former gives a higher weight to patients with low tumor burden.

RVD is similar to dice, however does not take into account the overlap of the two masks, only their size difference. This metric is meaningful in cases where the exact position of lesions is less important than the absolute size of the tumor. It is defined as

$$\text{RVD}(A, B) = \frac{|B| - |A|}{|A|}, \quad (3.14)$$

where A is the reference volume. RVD goes to 0 for a perfect segmentation.

ASD provides a metric to determine how far the surfaces of ground truth and predicted lesion are apart. Let S_A be the surface pixels of mask A and $d(x, S_A)$ the shortest distance from voxel x to the surface of A . Here, shortest is defined as smallest Euclidean distance. Then, ASD is given by

$$\text{ASD}(A, B) = \frac{1}{|S_A| + |S_B|} \left(\sum_{x \in S_A} d(x, S_B) + \sum_{x \in S_B} d(x, S_A) \right). \quad (3.15)$$

ASD goes to 0 as the segmentation improves.

RMSD is similar to ASD, but sums the square distances as follows:

$$\text{RMSD}(A, B) = \sqrt{\frac{1}{|S_A| + |S_B|} \left(\sum_{x \in S_A} d^2(x, S_B) + \sum_{x \in S_B} d^2(x, S_A) \right)}. \quad (3.16)$$

Again, a perfect segmentation yields a value of 0. Naturally, RMSD is strongly correlated with ASD. However, using squared distances causes larger surface errors to more heavily impact the final metric.

MSD is another surface metric and also referred to as Hausdorff distance. Instead of

calculating average values over both surfaces, only the maximum distances are taken into account:

$$\text{MSD}(A, B) = \max(\max_{x \in S_A} d(x, S_B), \max_{x \in S_B} d(x, S_A)). \quad (3.17)$$

As with the other surface distances, a value of 0 indicates a perfect prediction. Since effectively only a single voxel pair is considered in the final metric, MSD is highly sensitive to outliers. Simultaneously, evaluating the maximum error can be clinically interesting, especially for surgical applications.

Lesion recall, also known as sensitivity, is a measure of lesion detection. It is defined as

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (3.18)$$

where TP and FN are the true positives and false negatives, respectively. For a perfect segmentation, there would be no false negatives and recall would equal 1. This measure emphasizes the identification of true lesions, regardless of the number of false positives. Finally, lesion precision indicates how reliable a positive prediction is and is given by

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (3.19)$$

where FP are the false positives. As with recall, a perfect prediction leads to a value of 1. Generally, a high recall may be considered preferable to a high precision. This is because false positives can be ruled out by the radiologist, while missing lesions can lead to a potentially dangerous misdiagnosis.

3.2.3 Detectability Metrics

i.) Data Preparation

To complement the deep lesion segmentation analysis, a signal known exactly (SKE) model observer study is implemented. For this purpose, a dataset of images with and without lesion present is generated. Artificial 2D lesions L are inserted into the volumes as

$$L(x, y) = \frac{C}{2} \left(1 - \text{erf} \left(\frac{r(x, y) - R}{n} \right) \right), \quad (3.20)$$

where C is the contrast to the liver, r is the distance to the lesion center, R is the radius of the lesion, and n determines how smooth the lesion goes down to zero. The function is chosen in accordance with [158] in order to simulate realistic liver-tumor transitions. In general, R can be a 3D function. This allows the parameterization of complicated lesion shapes, e.g. by fitting a spherical function to an existing lesion. For simplicity and reproducibility, this work only considers circular lesions, i.e. $R(x, y, z) = R$.

Figure 3.10 shows the data preparation process. First, a random lesion location is chosen. To ensure that the simulated lesion is within the liver and does not overlap

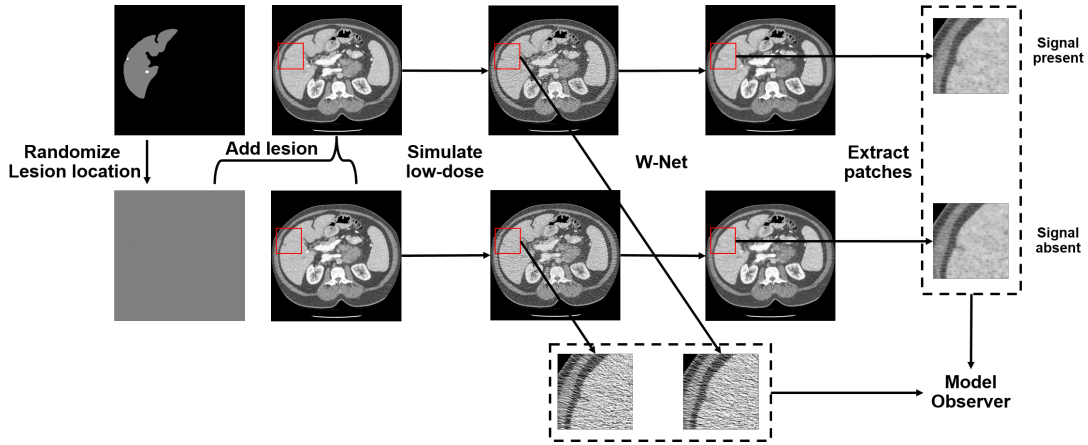


Figure 3.10: Pipeline for the model observer study.

any existing lesions, the liver mask excluding the ground truth lesions is eroded with a square structuring element that corresponds to the simulated lesion size. This leaves only areas far from a real lesion or air. Then, for each slice, the artificial lesion is inserted at a random point on this binary map. This yields the signal-present ground truth images. The process in section 3.1.1 subsequently produces low-dose CT images as previously. Note that the dataset without lesions, i.e. the signal-absent images, does not have to be generated again, as it is available from the previous experiment. Finally, the CNNs correct the images, and patches of size 100×100 centered around the lesion position are extracted from the signal-present and signal-absent images. This study analyzes two kinds of lesions: small lesions with $R = 2$ mm and $C = -40$ HU, and medium lesions with $R = 4$ mm and $C = -20$ HU.

Model observer (MO) studies are typically categorized by the type of signal and background in the test images, divided into "known-exactly" and "known-statistically" [139], [141], [159]. This work implements a SKE study, i.e. all the parameters of the lesion signal are known beforehand. One possibility of a signal known statistically (SKS) task would be to randomize the lesion location within the image patch. This work generates a dataset such that each image patch is from a different patient slice. Although all lesions are placed within the liver and not on an existing lesion, the background may still contain distinct features other than image noise, e.g. vessels or existing lesions far from the center. Therefore, the background is only known statistically and the MO study is SKE-background known statistically (BKS) [160]. In addition, section 4.2.2 investigates a background known exactly (BKE) study, with 100 different noise realizations for each image patch.

ii.) Model Observer Analysis

To analyze the signal absent and signal present patches, a channelized hotelling observer (CHO) with 40 Gabor channels is employed. The task for the the model observer is to determine whether a lesion is present in the test image or not. Internally, the image is

channelized by multiplying with the channels. This reduces the problem dimensionality and emulates the human visual system [134], [139]. Subsequently, the observer calculates a decision variable λ . If λ is above a certain threshold T , a lesion is detected. The model observer is implemented utilizing the publicly available IQModelo software [137]. Figure 3.11 shows the Gabor channels used in this study. The channels are not rotationally symmetric, unlike many other common CHO channel types, which aids in signal detection. It follows that they typically feature a high detection performance [137]. Although the simulated lesions are circular, the image noise is object dependent and therefore not isotropic. Thus, Gabor channels still appear an adequate choice for the channels. Each 2D Gabor channel $G(i, j)$ is defined as

$$G(i, j) = \exp\left(-4(\ln 2)((i - i_0)^2 + (j - j_0)^2)/w_s^2\right) \times \cos 2\pi f_c((i - i_0) \cos \theta + (j - j_0) \sin \theta) + \xi, \quad (3.21)$$

where i_0, j_0 is the center of the channel, w_s is the spatial width, θ the orientation, ξ a phase offset, and f_c the center frequency [159]. The spatial width is related to the bandwidth of the filter w_f by $w_s = 4(\ln 2)/(\pi w_f)$. To generate the channels, four bandwidths were used, namely $w_f = [1/32, 1/16, 1/8, 1/4]$ cycles/pixel with center $f_c = [3/64, 3/32, 3/16, 3/8]$. The orientation was chosen as $\theta = [0, 1/5, 2/5, 3/5, 4/5]\pi$ and the phase offset as $\xi = [0, \pi/2]$.

The model observer can be evaluated based on its receiver operating characteristic (ROC) curve [139]. The ROC curve plots the true positive rate vs. the false positive rate for different decision thresholds T . Thus it provides information about how well λ is able to differentiate between signal present and absent images. A more convenient metric is the area under receiver operating characteristic curve (AUC), which is directly related to the signal-to-noise ratio (SNR) of the observer. Here, $AUC = 1$ indicates a perfect observer, while $AUC = 0.5$ represents random guessing. If the AUC were below 0.5, the decision of the model observer could simply be inverted. Therefore, the value is always $0 \leq AUC \leq 1$ [139], [161]. Notably, IQModelo provides an unbiased estimate of AUC confidence intervals directly from the test set. It does not require separate training and test sets to give an estimate. In this study, the two-sided 95% confidence intervals are compared.

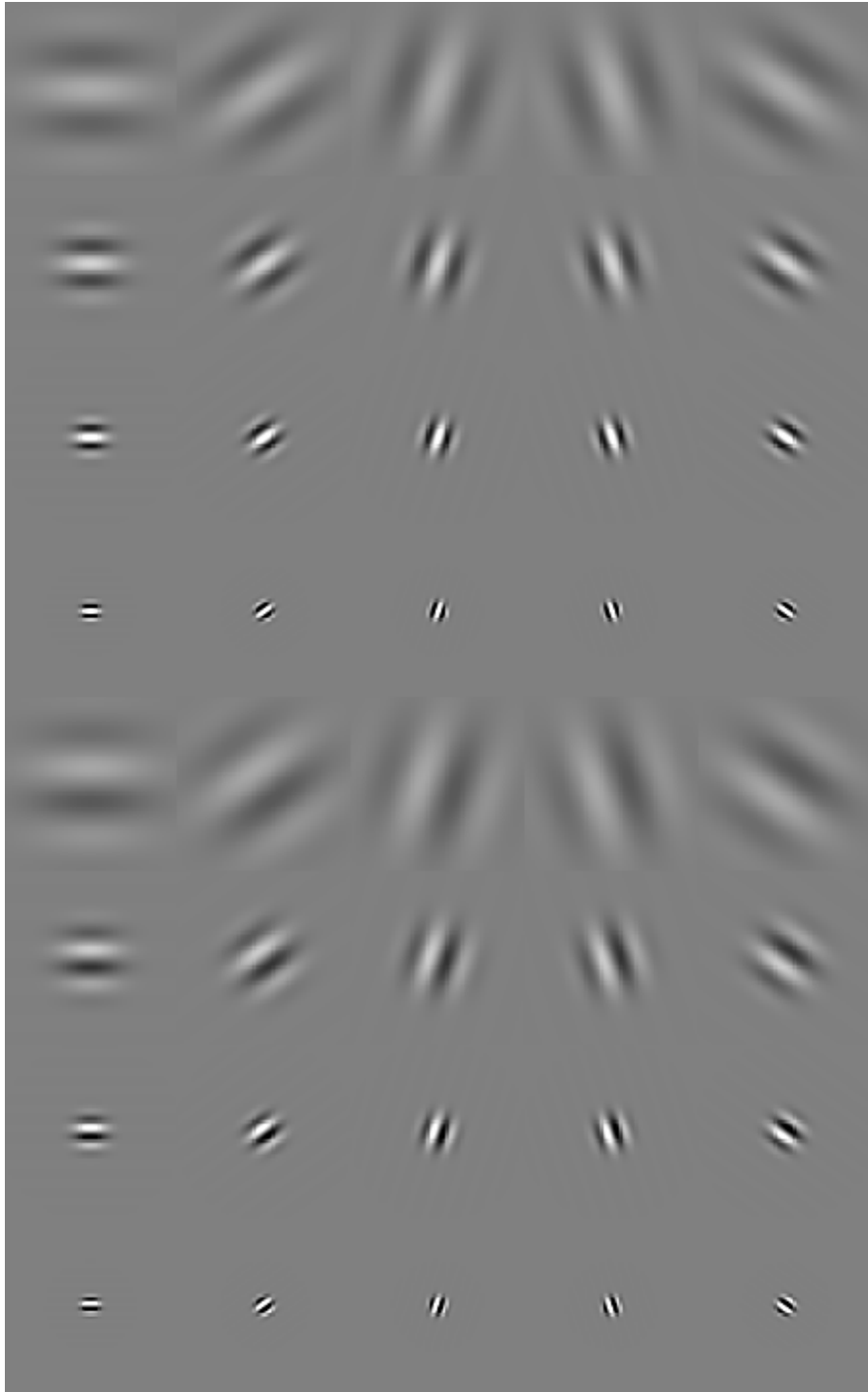


Figure 3.11: Gabor channels used for the channelized hotelling observer.

3.3 Detruncation

Ideally, the CT system acquires projections that cover the full patient cross-section for all angles. This is the case for the simulations for the low-dose CT experiments. Assuming a parallel beam geometry as before, the number of detector pixels M_f of size $\Delta\xi$ is sufficient to reconstruct a FOV with radius

$$R_{M,f} = \xi_{f,\max} = \frac{1}{2}(M_f - 1)\Delta\xi. \quad (3.22)$$

Note that the detector is centered, such that the detector width directly corresponds to the maximum FOV.

If the sinogram is truncated, only $M_t < M_f$ detector pixels are available. Following Equation (3.22), this will decrease the maximum FOV accordingly. In addition, even the reduced FOV will suffer from cupping artifacts as illustrated in Figure 2.7 [1], [105]. In order to restore image quality inside the FOV and possibly reconstruct structures in the extended field of view (eFOV), the sinogram must be extrapolated by $\frac{1}{2}(M_f - M_t)$ pixels on either side of the truncated sinogram. Note that the maximum reasonable eFOV is defined by the bore size of the system.

3.3.1 Simulating Truncated CT Data

To generate full and truncated sinograms, the ground truth images from the LiTS dataset are monochromatically forward-projected in parallel beam geometry. The geometry is identical to the one described in section 3.1.1, i.e. $N = 512$ projections with an angular range of 0 to 180° and $M_f = 512$ detector pixels with a pixel size of 0.8 mm. Two levels of truncation are simulated: 25% and 50%. Thus, the truncated detectors have $M_t = 384$, 256 pixels, respectively. No additional noise is added to the projections.

3.3.2 Cosine Detruncation

One method of extrapolating the missing data in the sinogram is a simple cosine detruncation [106]. Here, the sinogram smoothly goes to zero between the outermost acquired pixels at position $\pm\xi_{t,\max}$ and the new outermost pixels at $\pm\xi_{f,\max}$. The detruncated sinogram p_d is defined as

$$p_d(\vartheta, \xi) = \begin{cases} p_d(\vartheta, \xi), & |\xi| < \xi_{t,\max} \\ \cos \frac{|\xi| - \xi_{t,\max}}{2\pi(\xi_{f,\max} - \xi_{t,\max})} p_d(\vartheta, \pm\xi_{t,\max}), & \xi_{t,\max} < |\xi| < \xi_{f,\max} \\ 0, & |\xi| > \xi_{f,\max}, \end{cases} \quad (3.23)$$

where the value for extrapolation depends on whether ξ is positive or negative.

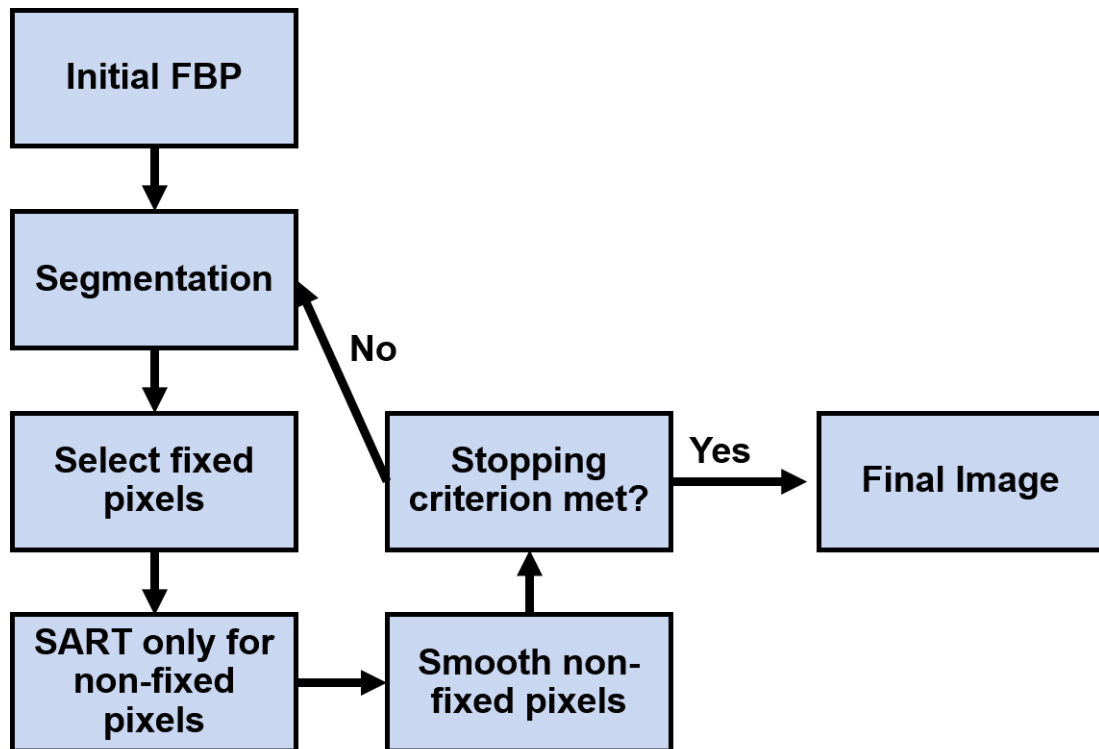


Figure 3.12: Scheme of the DART reconstruction.

3.3.3 DART Detruncation

DART combines prior information in the form of discretization with real-valued algebraic reconstruction [118]–[120]. Figure 3.12 shows the basic DART scheme. The main principle of DART is to leave some pixels untouched during each iteration of some iterative reconstruction. These fixed pixels are set to constant values, e.g. 0 HU for soft tissue pixels and -1000 HU for air. In this manner, prior knowledge is introduced into the system, and the number of equations in the reconstruction are reduced, improving the reconstruction.

There are three main design concerns with DART: determining which pixels to fix, the value of said fixed pixels, and the iterative reconstruction of the free pixels. After an initial reconstruction, DART determines the free and the fixed pixels. For this purpose, the image is segmented into several tissues via thresholding. In this work, pixels are segmented into soft tissue and air. Subsequently, pixels that are fully surrounded by pixels of the same class are fixed and set to the average value of their class. This process is essentially a grayscale erosion of the tissue map with a 3×3 structuring element. In addition, each fixed pixel is given an additional probability of 65% to be classified as free. This allows the DART algorithm to change values inside homogeneous patches, i.e. create and remove holes. The fixed pixels are forward projected, and their contribution is then subtracted from the raw-data. Five simultaneous algebraic

reconstruction technique (SART) iterations finally update the free image pixels only [162]. One SART iteration is defined as

$$f_{\text{new}} = f + \lambda \frac{1}{X^T \mathbf{1}} X^T \left(\frac{p - Xf}{X \mathbf{1}} \right), \quad (3.24)$$

where f is the current image estimate, f_{new} is the new estimate, p are the raw data, λ is a relaxation factor, X is the forward-projection and X^T is the back-projection. Note that other iterative reconstruction algorithms are also feasible in the context of DART [118].

One issue in DART is the accumulation of noise on the free pixels, as all noise in the raw data is distributed only on the free pixels. To remedy this, a bilateral filter smooths the free pixels after SART, which concludes the DART iteration. DART has no well-defined stopping criterion. Instead, a fixed number of iterations $N_{\text{iter}} = 1500$ is performed. To increase convergence speed, the bilateral filter is set stronger in early iterations. A conventional cosine detruncation as described in the previous section provides the initial reconstruction. Note that while DART was designed to produce segmented images, the algorithm internally uses real valued images, which are used in this study [118]. The DART image is finally utilized as a prior image and forward projected with M_f . This DART sinogram is used to complete the original raw data in order to expand the FOV.

3.3.4 CNN Detruncation

As a data-driven comparison method to the DART detruncation, a W-Net as introduced in section 3.1.2, is trained to predict ground truth images from truncated CT sinograms. As with the network for low-dose CT, 90 patients are used for training, 20 for validation and 20 for testing. The network is trained separately for the 25% and 50% truncated sinograms. Besides the number of training epochs, which is set to 30, the other training parameters are as detailed in section 3.1.2. The W-Net results are again used as prior images to complete the original raw data.

3.3.5 Analysis

For quantitative analysis, several metrics are employed. First, the image quality inside the original FOV is measured by RMSE_{FOV} in comparison to the ground truth images. Similarly, the image quality inside the larger eFOV is computed as $\text{RMSE}_{\text{eFOV}}$. Furthermore, SSIM is calculated over the whole image. Finally, the images are segmented by thresholding into air and patient with a threshold of -600 HU. On these segmentations, the dice score is computed.

4 | Results

This chapter presents the experimental results of the CNN denoising and detruncation. First, the different combinations of sparse-view and low-mAs CT are compared based on the image metrics introduced in the previous section. Earlier experiments on this subject were presented at IEEE Medical Imaging Conference 2023 [163]. Second, an ablation study analyzes the importance of the different network components and loss function. Third, the image metrics are applied to denoised images inferred from insufficiently trained networks to determine whether the metrics can be used for quality assurance. Finally, the detruncation results are shown. Initial results with the DART detruncation were presented at the CT Meeting 2022 [164].

4.1 Comparison of Dose Reduction Approaches

This subsection analyzes how the W-Net is able to improve the image quality of images from different dose reduction approaches. Figure 4.1 illustrates the denoising performance of the W-Net on a pelvis scan. Clearly, all denoised images are visually superior to the low-dose images. Both image noise and sparseness artifacts are strongly reduced or removed. The noise texture of the denoised images closely resembles the ground truth and no oversmoothing is apparent. Nevertheless, some texture differences are visible that originate from the perceptual loss. In some instances, fine structures in the ground truth are not visible or altered in the CNN output, in particular for the case with $N = 102$. However, the structures are heavily obscured by noise and artifacts in the low-dose images. The difference images further indicate that the network enhanced the edges of some of the bone structures. To further showcase the denoising ability of the CNN on different anatomies, Figures 4.2 and 4.3 give the network outputs for an abdomen and thorax case, respectively. Similarly to the previous case, the CNNs are able to restore image quality for all low-dose realizations in both body regions. Still, some inconsistencies compared to the ground truth remain. While the low-dose images clearly decrease in image quality as N becomes smaller, the differences are less significant between the dose realizations after CNN processing.

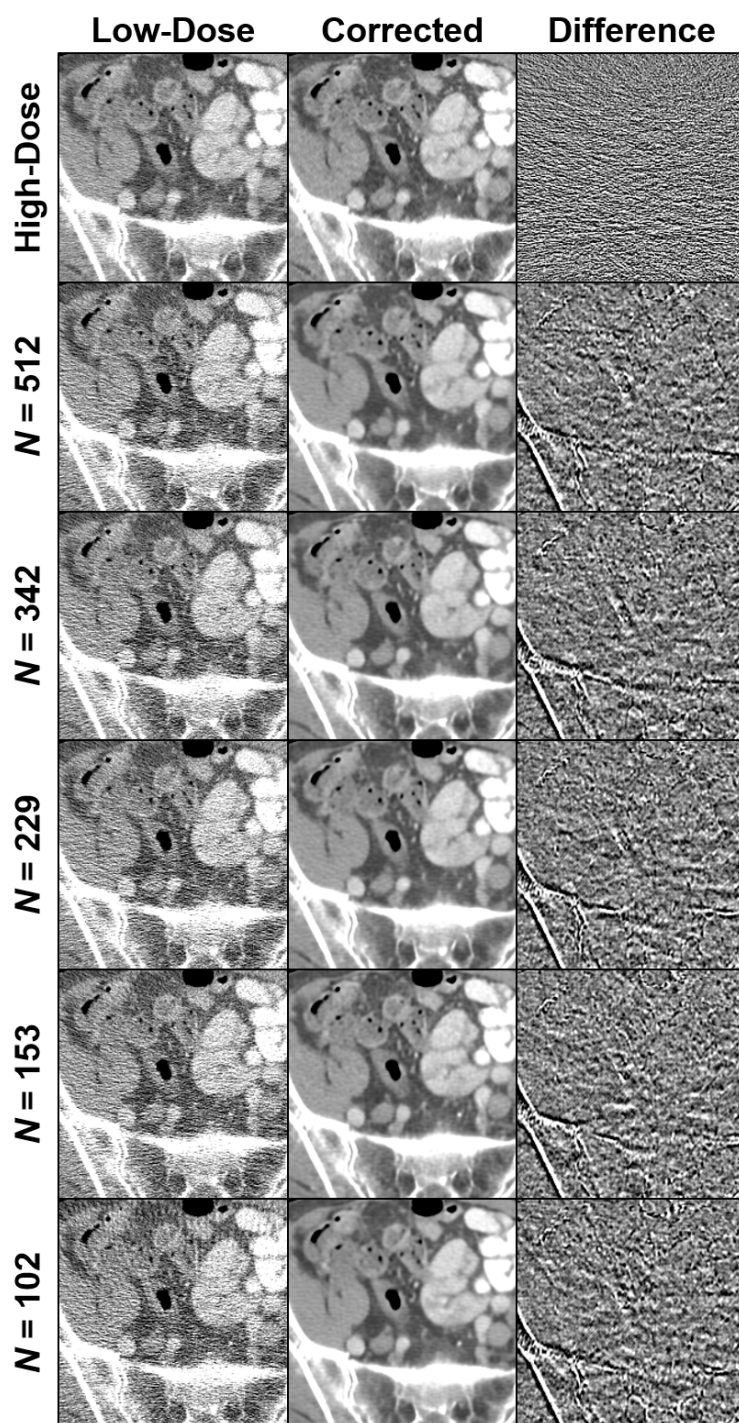


Figure 4.1: Denoising results for a pelvis slice. First row shows high-dose image, ground truth, and difference of high-dose to ground truth. Difference images are with respect to the ground truth. $C = 0$ HU, $W = 500$ HU for CT images, $C = 0$ HU, $W = 100$ HU for difference images.

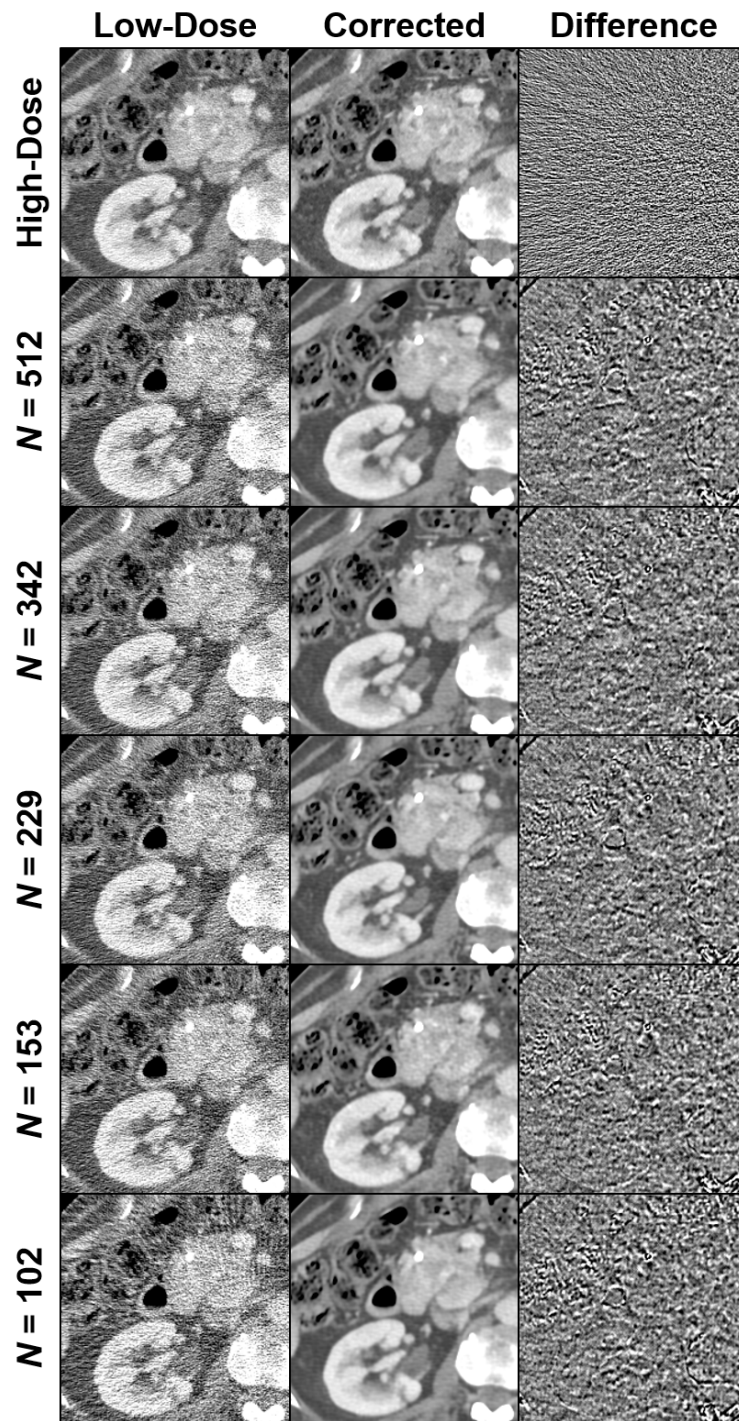


Figure 4.2: Denoising results for an abdomen slice. First row shows high-dose image, ground truth, and difference of high-dose to ground truth. Difference images are with respect to the ground truth. $C = 0$ HU, $W = 500$ HU for CT images, $C = 0$ HU, $W = 100$ HU for difference images.

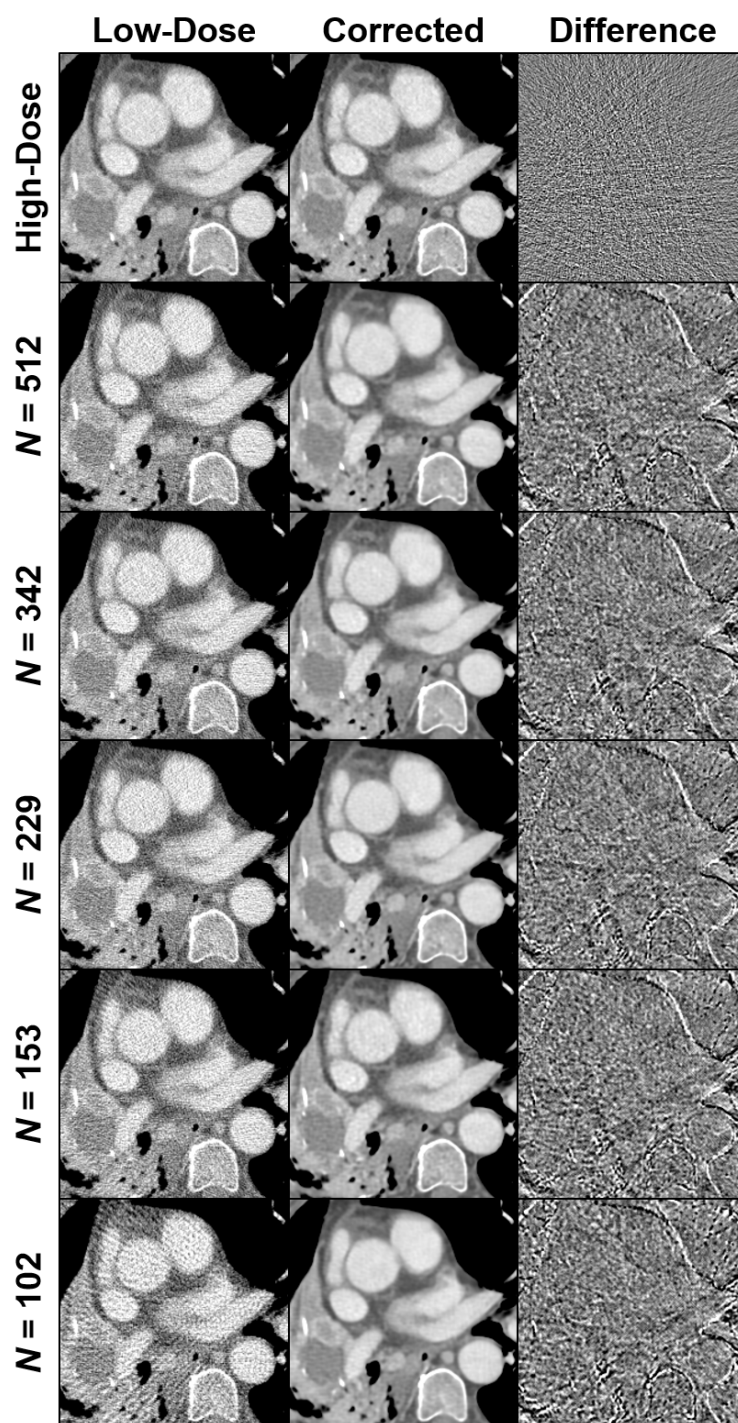


Figure 4.3: Denoising results for a thorax slice. First row shows high-dose image, ground truth, and difference of high-dose to ground truth. Difference images are with respect to the ground truth. $C = 0$ HU, $W = 500$ HU for CT images, $C = 0$ HU, $W = 100$ HU for difference images.

4.1.1 Conventional Image Quality Metrics

In addition to the qualitative comparison, a quantitative analysis was performed. For this purpose, RMSE and SSIM were calculated for each slice of the test set and averaged. Table 4.1 summarizes the results. Notably, the image quality of the uncorrected low-dose images is not equal when measured with either metric. Instead, there is a trend of improved image quality, i.e. lower RMSE and higher SSIM, for higher numbers of projections. This indicates that despite equal dose, sparseness artifacts are more detrimental to these conventional measures. Going from $N = 512$ to $N = 102$ on increases RMSE by 50% and decreases SSIM by 6%. After denoising, all dose realizations feature significantly reduced RMSE and increased SSIM. All RMSE values are superior to the simulated high-dose images. This highlights the ability of the W-Net to correct all investigated scan types. While all networks perform well, the trend of decreasing image quality as N decreases remains, with $N = 102$ performing worst in both metrics. However, as the image quality of all images is high, the differences are not as relevant. The mean relative differences between full sparse-view and full low-mAs CT are now 6% and 0.7% for MSE and SSIM, respectively. Based on a Wilcoxon signed-rank test, the differences between uncorrected and corresponding corrected images, as well as between the corrected images are all statistically significant.

Metric	Dataset	$N = 512$	$N = 342$	$N = 229$	$N = 153$	$N = 102$	High-Dose
RMSE	Low-Dose	47.056	47.765	49.460	46.229	72.290	21.417
	W-Net	17.930	17.783	17.661	17.992	19.064	21.417
SSIM	Low-Dose	0.7846	0.7823	0.7790	0.7661	0.7385	0.9020
	W-Net	0.9479	0.9480	0.9482	0.9461	0.9415	0.9020

Table 4.1: Conventional quantitative image quality analysis of the W-Net denoising using RMSE (in HU) and SSIM.

4.1.2 Task-based Image Quality Metrics

i.) Model Observer Study

Although widely used, conventional image quality metrics do not accurately represent the diagnostic value of a CT image. Instead, task-based metrics can better summarize image quality with respect to specific clinical scenarios. Here, results for the lesion segmentation and MO study will be shown. Figure 4.4 presents signal present patches of the 2 mm lesion study before and after denoising. In addition, the signal-only image is computed by subtracting . Again, all networks were able to restore image quality in the low-dose image patches. The signal is barely visible in both the high-dose and corrected images, making a qualitative assessment difficult. Figure 4.5 shows the denoising results of the 4 mm lesion study. As before, the W-Net is able to reduce the noise in all cases. The large lesion is clearly visible in all network outputs. Similar to Figure 4.1, some edges are enhanced.

Table 4.2 provides the quantitative SKE-BKS results in form of 95% confidence intervals of the AUC estimates. As with the conventional metrics, the AUC estimates indicate a trend of decreasing image quality for low-dose CT with fewer projections. The trend is valid for both investigated lesion sizes. For instance, the lower bound of the AUC estimate decreases by 4% from $N = 512$ to $N = 102$. After CNN denoising, all but one AUC estimate (for $N = 342$) increases. This indicates that the denoising process did not remove any details from the image as would for instance be expected from a simple smoothing. The increase in AUC is more pronounced for the larger lesion size, possibly because the network more easily distinguishes large, low contrast lesions from the background. Although $N = 512$ yields the best results for both lesion sizes, there is no clear trend of increasing AUC for increasing number of projections. In contrast to the conventional metrics, the AUC of W-Net images does not reach the level of the high-dose image, implying that not all information lost in the low-dose simulation could be recovered.

Dataset		$N = 512$	$N = 342$	$N = 229$	$N = 153$	$N = 102$	High-Dose
2 mm	Low-Dose	0.8720	0.8705	0.8691	0.8651	0.8376	0.9348
		0.8773	0.8758	0.8745	0.8705	0.8430	0.9391
	W-Net	0.8800	0.8641	0.8704	0.8736	0.8661	0.9348
		0.8853	0.8696	0.8758	0.8790	0.8715	0.9391
4 mm	Low-Dose	0.8290	0.8273	0.8258	0.8211	0.8095	0.8810
		0.8327	0.8327	0.8312	0.8265	0.8149	0.8863
	W-Net	0.8491	0.8405	0.8401	0.8384	0.8422	0.8810
		0.8546	0.8460	0.8455	0.8438	0.8476	0.8863

Table 4.2: 95% confidence intervals of AUC estimates for the SKE-BKS model observer for 2 mm and 4 mm lesions with and without W-Net denoising.

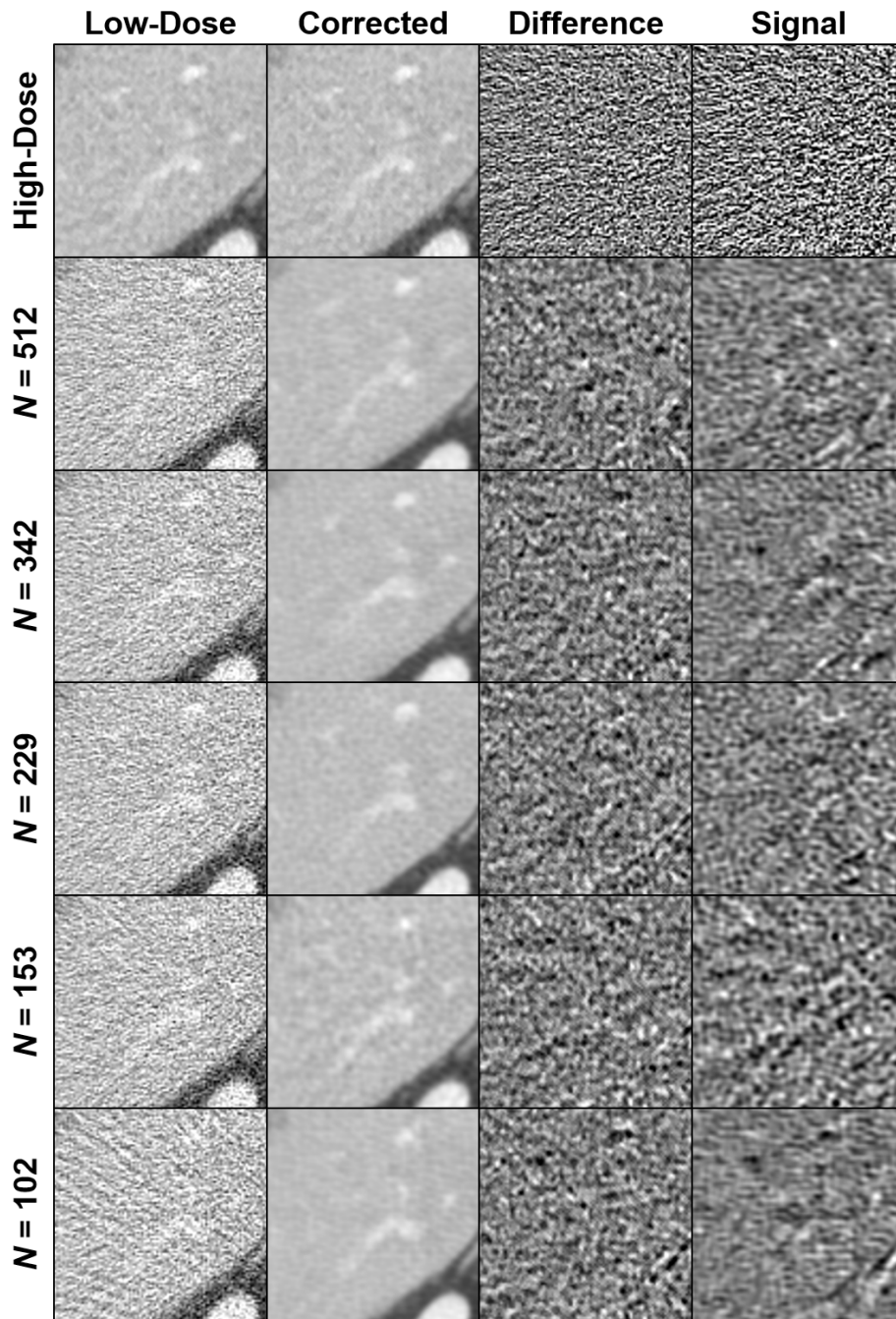


Figure 4.4: Denoising results for the model observer study with a 2 mm lesion. First row shows high-dose image, ground truth, difference of high-dose to ground truth, and high-dose signal image. Difference images are with respect to the ground truth. Signal images are calculated as the difference between signal-present and signal-absent images. $C = 0$ HU, $W = 500$ HU for CT images, $C = 0$ HU, $W = 100$ HU for difference images.

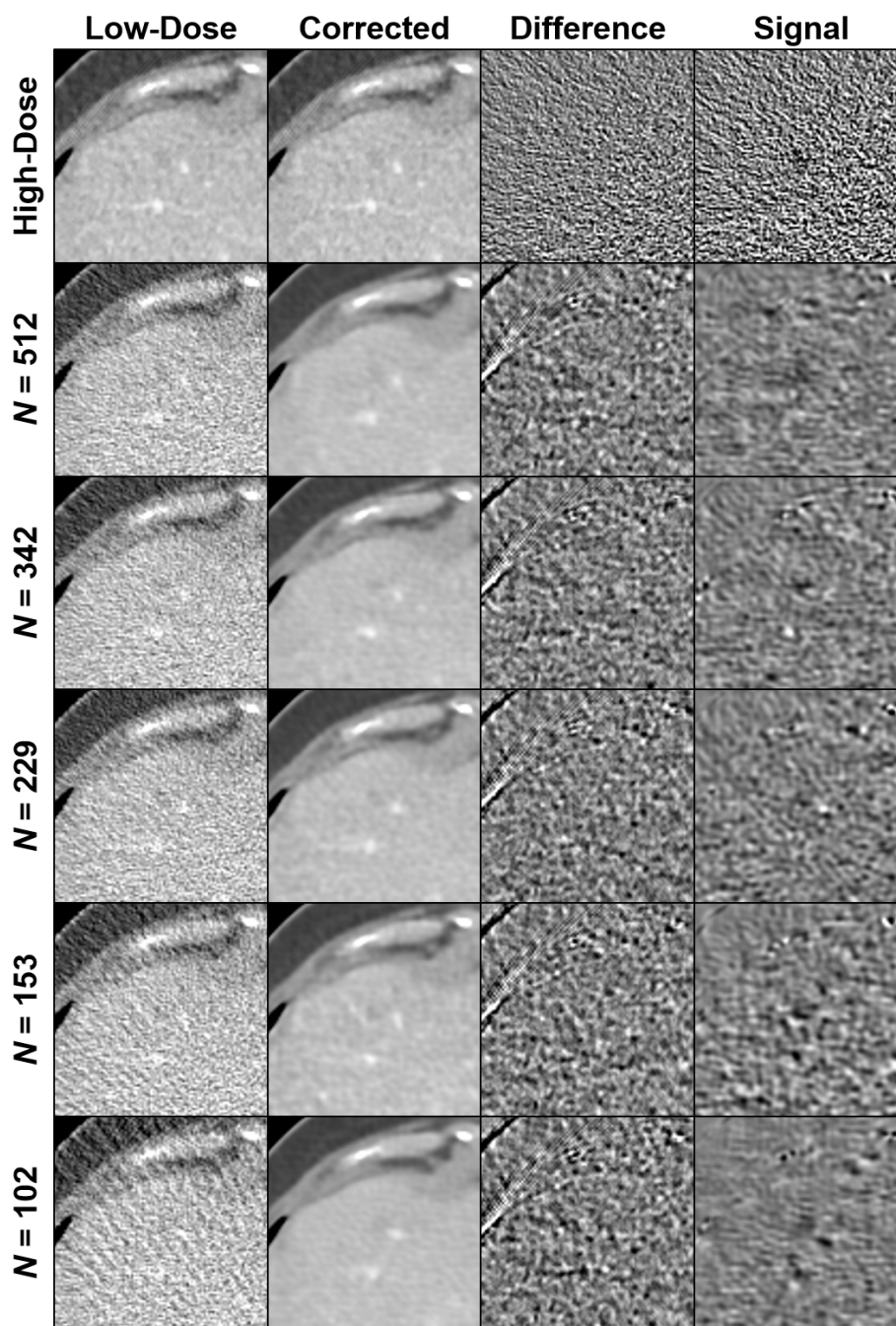


Figure 4.5: Denoising results for the model observer study with a 4 mm lesion. First row shows high-dose image, ground truth, difference of high-dose to ground truth, and high-dose signal image. Difference images are with respect to the ground truth. Signal images are calculated as the difference between signal-present and signal-absent images. $C = 0$ HU, $W = 500$ HU for CT images, $C = 0$ HU, $W = 100$ HU for difference images.

Metric	$N = 512$	$N = 342$	$N = 229$	$N = 153$	$N = 102$
Per Case Dice	0.832	0.830	0.819	0.848	0.806
Global Dice	0.923	0.926	0.926	0.928	0.923
RVD	-0.052	-0.003	-0.004	-0.014	-0.022
ASD	0.671	0.644	0.637	0.633	0.650
RMSD	0.974	0.943	0.928	0.929	0.945
MSD	4.226	4.270	4.112	4.109	4.229
Recall	0.653	0.651	0.644	0.662	0.642
Precision	0.614	0.584	0.673	0.721	0.627

Table 4.3: Quantitative results of the segmentation of liver lesions after W-Net denoising.

ii.) Segmentation

Besides the MO study, a segmentation-based approach was used to analyze the denoised CT images. First, the liver and liver lesions are segmented in the images after W-Net denoising. Subsequently, the quality of the segmentation is quantified with several metrics as described in section 3.2.2. Figure 4.6 shows two example segmentation results from the test set. In all cases, the algorithms identified all regions containing. Despite the overall good segmentation, in some instances the segmented lesions merge multiple ground truth lesions. It should be noted, however, that these merges occur for lesions that belong to the same 3D object. In the first case, the predicted segmentation is significantly smoother than the ground truth. No significant difference in segmentation accuracy is evident between the dose instances. Table 4.3 provides the quantitative results of the segmentation. $N = 153$ produced the best metrics, besides RVD (third best) and RMSD (second best). $N = 512$ and $N = 102$ generally perform worse than the other dose instances in RVD and surface metrics. Otherwise, there is no clear trend visible.

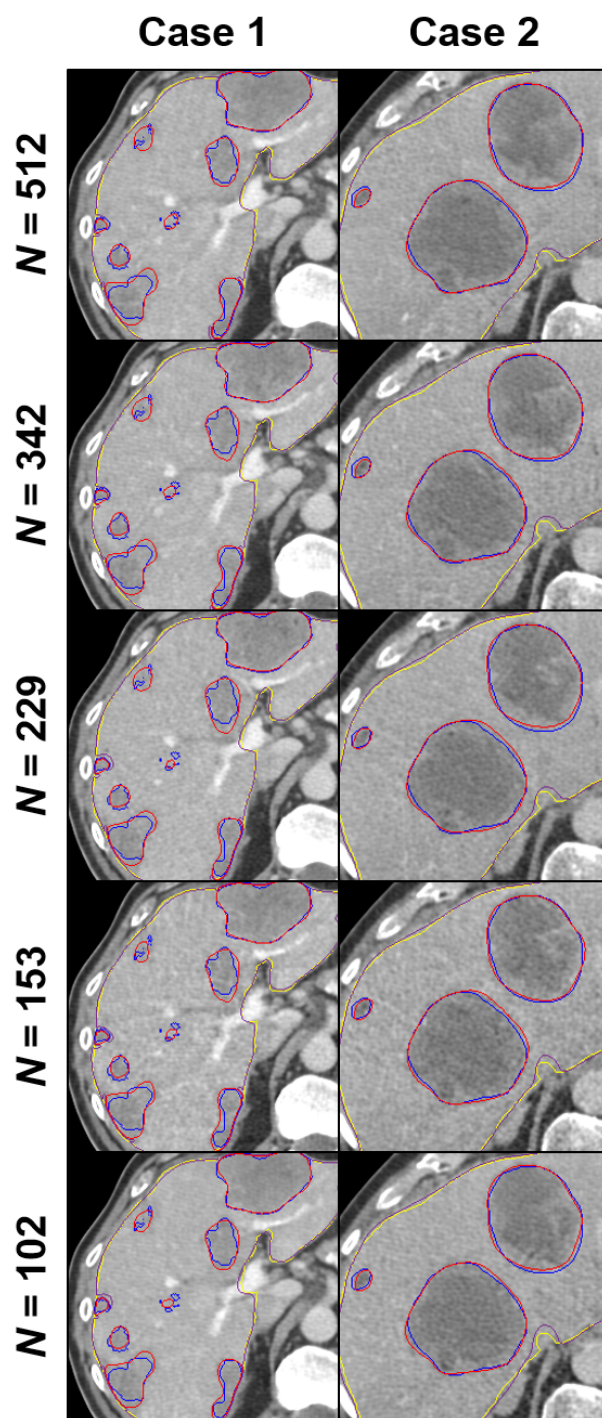


Figure 4.6: CNN-based segmentation contours of liver and liver lesions after W-Net denoising of low-dose CT. Liver ground truth in yellow, segmented liver in purple, ground truth lesions in blue, segmented lesions in red. $C = 25$ HU, $W = 450$ HU.

4.2 Validation of Task-Based Metrics

4.2.1 Segmentation Metrics

From the data above it is unclear how accurately the segmentation-based metrics are able to identify high quality CT images. Therefore, to validate the findings, a new test set was created consisting of ten different noise realizations of five volumes. As before, all volumes were denoised by the W-Net and subsequently segmented by the H-DenseUNet as described in section 3.2.2 [64]. This set can now be used to validate the segmentation metrics. Table 4.4 provides the results of the validation study, i.e. the mean and standard deviation of each metric. Apart from global dice, the standard deviations are relatively high, especially for RVD. In addition, a Wilcoxon signed-rank test was performed. The test indicated that only very few of the measures yielded significant differences between dose implementations. For example, dice per case was not significant. On the contrary, global dice differences are significant, apart from the comparisons between $N = 512$, 342, and 102. This indicates that the low-dose realizations are equivalent in terms of segmentation performance within the accuracy of the test.

Another point of interest is what the test segmentations should be compared to. The segmentation metrics might be considered a surrogate to estimate the proximity of the test image to the distribution of the ground truth. However, if the test image were identical to the ground truth image, the H-DenseUNet segmentation would not yield the ground truth segmentation and consequently not yield perfect metric scores. Thus, one should examine the efficacy of comparing the test segmentation to the H-DenseUNet output for the ground truth CT image. Table 4.5 gives the results for the validation set. When compared to Table 4.4, it is noticeable that most metrics shift in a common direction. In some cases, the ranking of the noise implementations changes slightly, but most comparisons remain as before.

Metric	$N = 512$	$N = 342$	$N = 229$	$N = 153$	$N = 102$
Per Case Dice	0.802	0.796	0.800	0.811	0.784
	± 0.037	± 0.028	± 0.032	± 0.039	± 0.022
Global Dice	0.952	0.953	0.954	0.959	0.953
	± 0.0003	± 0.0005	± 0.0005	± 0.0003	± 0.0005
RVD	0.077	0.071	0.085	0.015	0.039
	± 0.024	± 0.014	± 0.023	± 0.023	± 0.021
ASD	0.717	0.670	0.629	0.562	0.670
	± 0.022	± 0.018	± 0.014	± 0.015	± 0.015
RMSD	1.025	0.953	0.911	0.834	0.963
	± 0.047	± 0.025	± 0.023	± 0.017	± 0.023
MSD	5.133	4.841	4.966	4.414	5.077
	± 0.325	± 0.247	± 0.306	± 0.206	± 0.302
Recall	0.829	0.829	0.856	0.842	0.841
	± 0.020	± 0.021	± 0.020	± 0.027	± 0.017
Precision	0.529	0.543	0.568	0.638	0.539
	± 0.034	± 0.030	± 0.043	± 0.035	± 0.031

Table 4.4: Validation results of the segmentation metrics based on ten repeated measurements of five volumes. Metric values are given as mean and standard deviation over the ten noise instances.

Metric	$N = 512$	$N = 342$	$N = 229$	$N = 153$	$N = 102$
Per Case Dice	0.817	0.812	0.811	0.822	0.793
	± 0.037	± 0.033	± 0.037	± 0.037	± 0.024
Global Dice	0.962	0.965	0.967	0.971	0.964
	± 0.000	± 0.001	± 0.000	± 0.000	± 0.001
RVD	0.033	0.042	0.090	0.011	0.005
	± 0.038	± 0.024	± 0.046	± 0.031	± 0.018
ASD	0.791	0.681	0.744	0.622	0.805
	± 0.043	± 0.054	± 0.038	± 0.042	± 0.025
RMSD	1.167	1.014	1.087	0.961	1.160
	± 0.070	± 0.063	± 0.043	± 0.044	± 0.030
MSD	5.485	4.710	5.014	4.591	5.576
	± 0.338	± 0.296	± 0.229	± 0.180	± 0.318
Recall	0.828	0.839	0.870	0.870	0.839
	± 0.029	± 0.026	± 0.028	± 0.036	± 0.016
Precision	0.617	0.642	0.673	0.769	0.627
	± 0.042	± 0.042	± 0.052	± 0.038	± 0.032

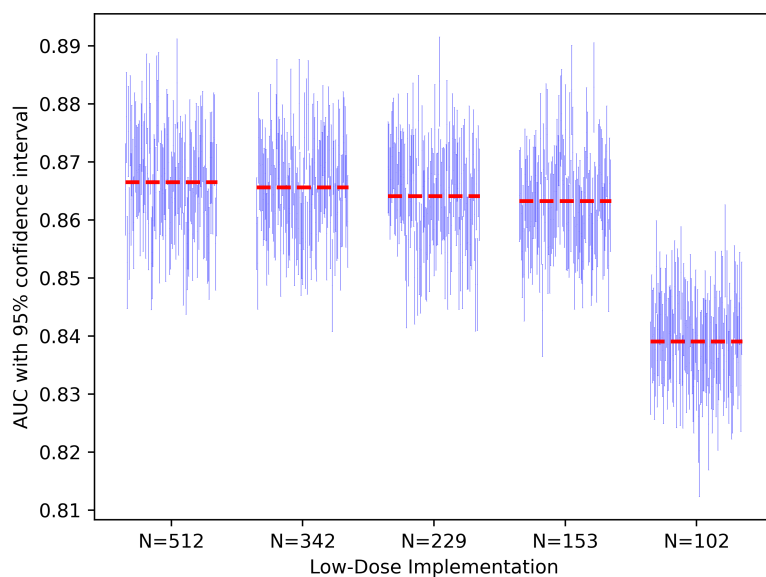
Table 4.5: Validation results of the segmentation metrics calculated in reference to the CNN segmentations of the ground truth CT images. Values given as mean and standard deviation over the ten noise instances.

4.2.2 Model Observer

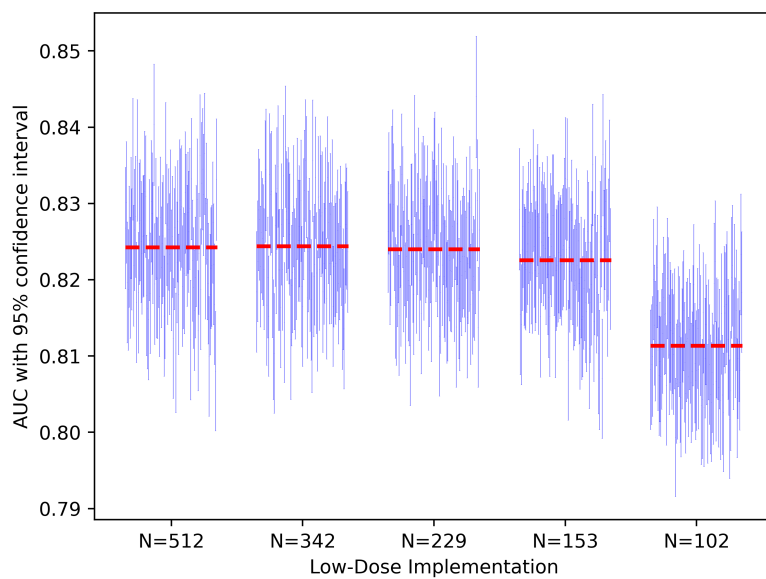
In addition to validating the segmentation-based metrics, it is also of interest to investigate the MO study in-depth. While SKE-BKS studies are highly clinically relevant, they might not yield the best assessment of image quality after CNN processing. For this purpose, for each lesion position from five of the test volumes, 100 noise instances were generated with and without lesions and corrected with the CNN. This leads to two datasets for both lesion sizes: 904 SKE-BKE datasets with 100 signal-absent and present images each, and 100 SKE-BKS datasets with 904 different backgrounds.

Figure 4.7a and 4.7b plot the SKE-BKS AUC confidence intervals results for the 2 mm and 4 mm lesion study without W-Net denoising, respectively. As evident, the image quality is reduced as the number of projections decreases. This agrees with the previous findings on conventional image quality metrics in section 4.1.1. Figure 4.8a and 4.8b show the same graph but for CNN corrected images. According to the data, the previously seen trend is gone. Compared to Table 4.2, the relative ranking of the dose implementations has not changed, suggesting that the SKE-BKS study yields consistent results.

Figure 4.9a and 4.9b illustrate the SKE-BKE results for the uncorrected low-dose CT. Since the MO is able to virtually perfectly distinguish the signal-present and -absent images, internal Gaussian noise with standard deviation of 100 HU was added to the MO. All low-dose instances feature almost identical AUC values at the same lesion size. Figure 4.10a and 4.10b show the corresponding results for corrected images. Firstly, $N = 342$ and 229 are being outperformed by the other low-dose implementations. Secondly, compared to Figure 4.9, the AUC for corrected images decreases, indicating a loss of image quality due to the W-Net processing. Thus, BKE and BKS studies yield conflicting results.

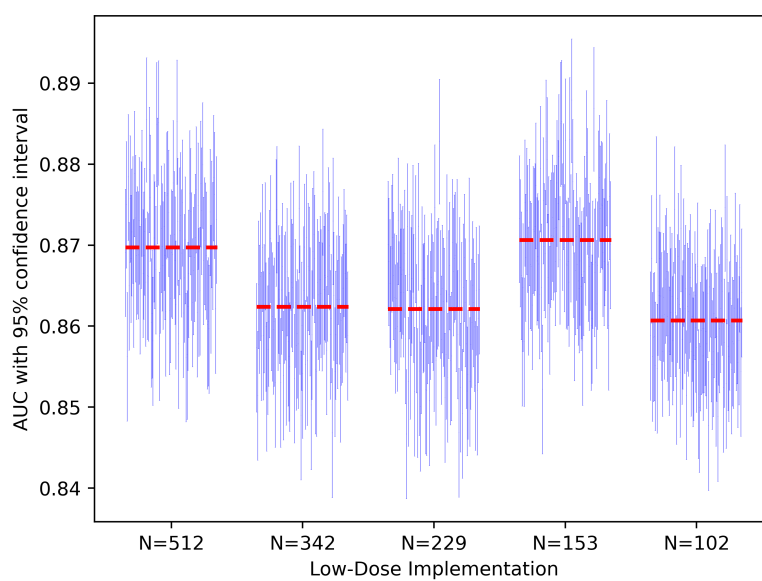


(a) 2 mm lesions.

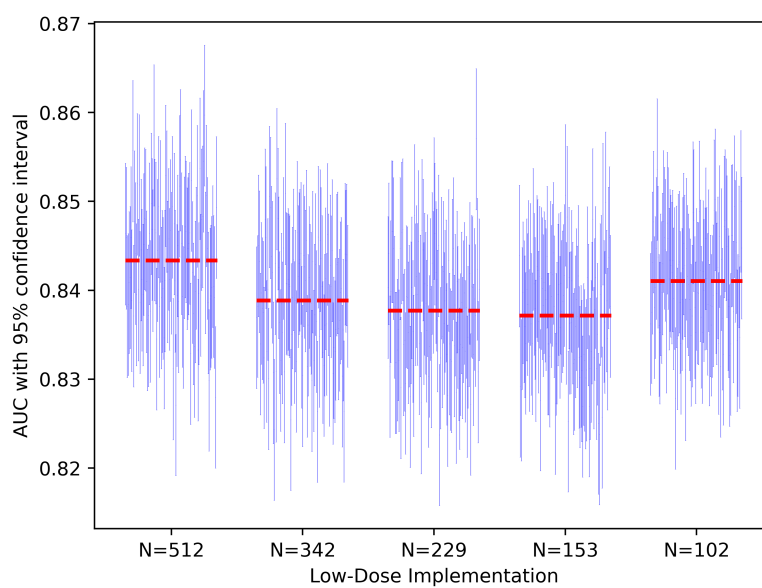


(b) 4 mm lesions.

Figure 4.7: 95% confidence interval estimates of the AUC for the repeated SKE-BKS model observer with low-dose images. Red dashed lines indicate mean value.

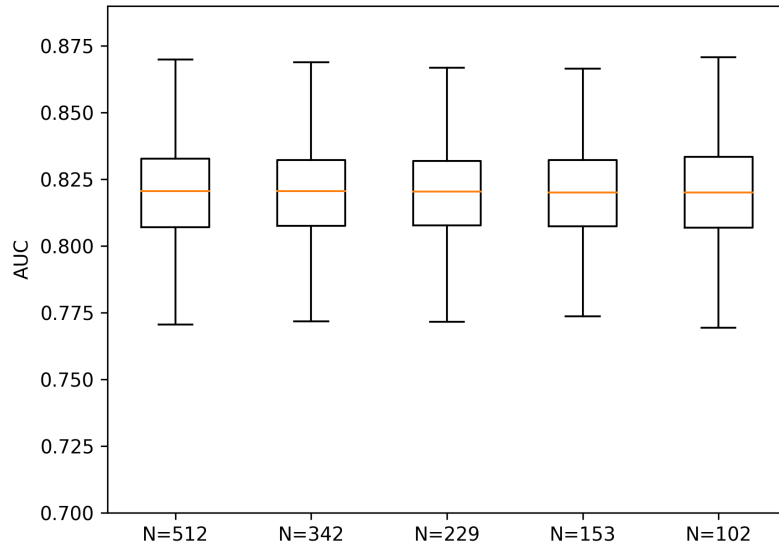


(a) 2 mm lesions.

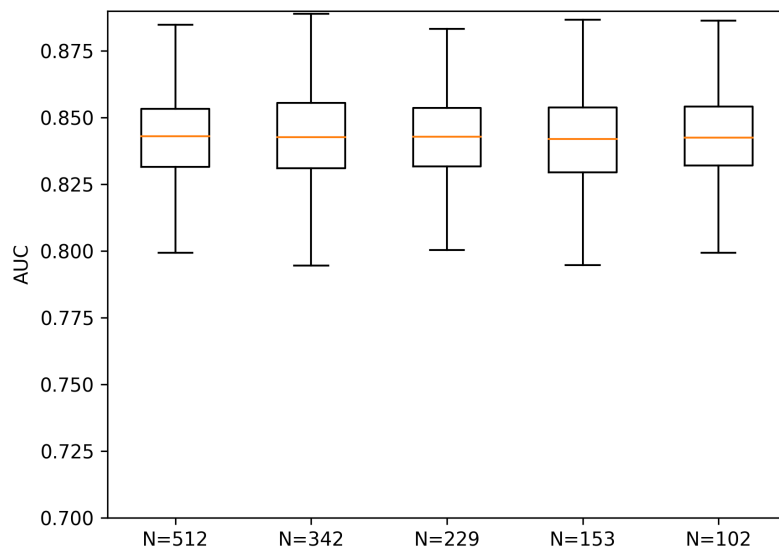


(b) 4 mm lesions.

Figure 4.8: 95% confidence interval estimates of the AUC for the repeated SKE-BKS model observer with denoised images. Red dashed lines indicate mean value.

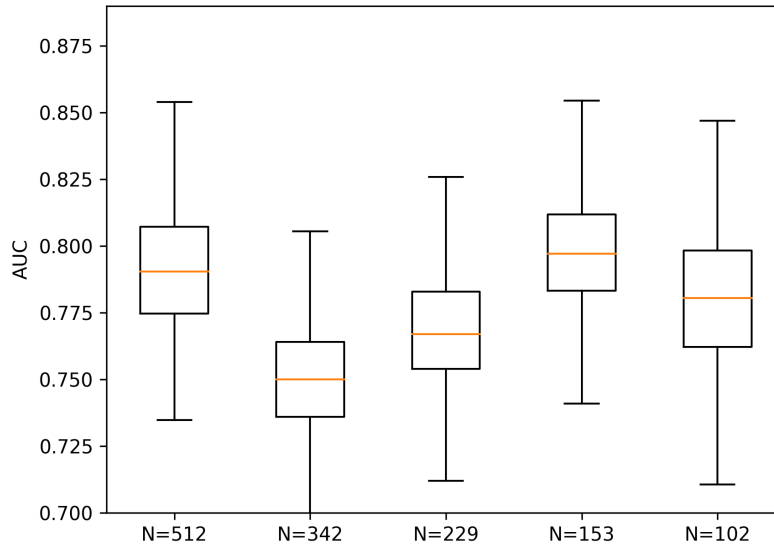


(a) 2 mm lesions.

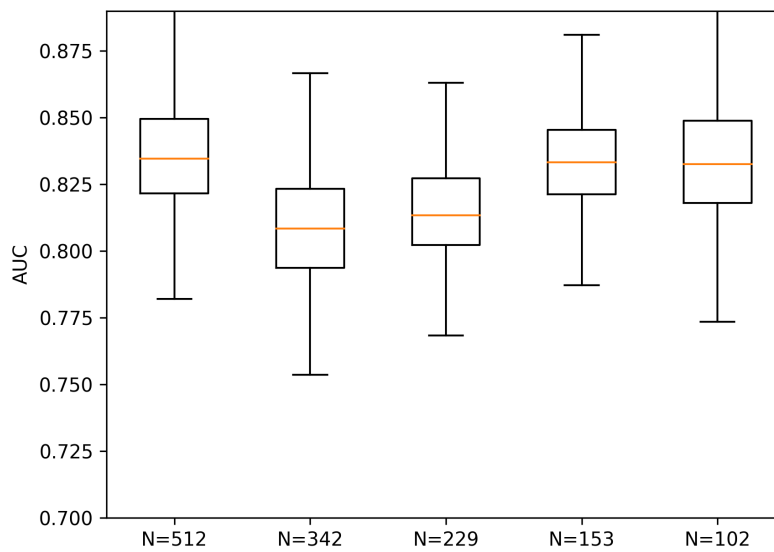


(b) 4 mm lesions.

Figure 4.9: AUC spread for the SKE-BKE model observer with low-dose images.



(a) 2 mm lesions.



(b) 4 mm lesions.

Figure 4.10: AUC spread for the SKE-BKE model observer with denoised images.

4.3 Ablation Study

This section explores the network components and loss function of the W-Net and how they influence the denoising performance.

4.3.1 Network Architecture

As detailed in section 3.1.2, the W-Net consists of three parts: a projection domain U-Net, a reconstruction layer, and an image domain U-Net. For the ablation study, three network architectures are investigated: projection domain U-Net with FBP layer $U\text{-Net}_{\text{Proj}}$, only the image domain U-Net $U\text{-Net}_{\text{Img}}$, and the full W-Net. For simplicity, not all dose configurations are tested. Instead, to show both denoising and destreaking performance, the number of projections is set to $N = 229$.

Figure 4.11 shows the CNN results for a pelvis case. The $U\text{-Net}_{\text{Sino}}$ output features remaining streaks. Furthermore, the image is smoothed, as highlighted by the difference image. The image domain and dual domain networks both provide good results, the latter more closely resembling the ground truth. The W-Net again enhances some bone edges. Figure 4.12 and 4.13 present an abdomen and thorax case, respectively. Again, the sinogram method performs worst. Table 4.6 provides the quantitative results for the conventional metrics. $U\text{-Net}_{\text{Img}}$ features the lowest RMSE, and best SSIM. The full W-Net has the second lowest RMSE and second highest SSIM, yielding better metrics than the high-dose image. The sinogram domain network yields the worst results and is inferior to the high-dose images in terms of RMSE.

Figure 4.14 and 4.15 showcase the performance on the 2 mm and 4 mm lesion model observer dataset, respectively. Here, the smoothing performed by $U\text{-Net}_{\text{Sino}}$ and $U\text{-Net}_{\text{Img}}$ becomes more apparent. The signal is most visible for the image domain and dual domain results. Table 4.7 lists the quantitative results of the model observer study. For both lesion sizes, the full W-Net yields the best results, with an increase of up to 9% in the lower AUC bound compared to the partial networks. Surprisingly, $U\text{-Net}_{\text{Img}}$ performs worse than $U\text{-Net}_{\text{Sino}}$.

Finally, Figure 4.16 and Table 4.8 highlight the segmentation results. Overall all tested network architectures feature similar qualitative segmentation performance. The full W-Net yields superior results in global dice, RVD, ASD, RMSD, MSD and recall. The image domain network gives the worst results in all tested metrics besides recall, where it performs second best.

Metric	$U\text{-Net}_{\text{Sino}}$	$U\text{-Net}_{\text{Img}}$	W-Net	Low-Dose	High-Dose
RMSE	22.436	13.441	17.661	49.460	21.417
SSIM	0.930	0.953	0.948	0.7790	0.9020

Table 4.6: Conventional quantitative results using RMSE (in HU) and SSIM for the ablation study with $N = 229$.

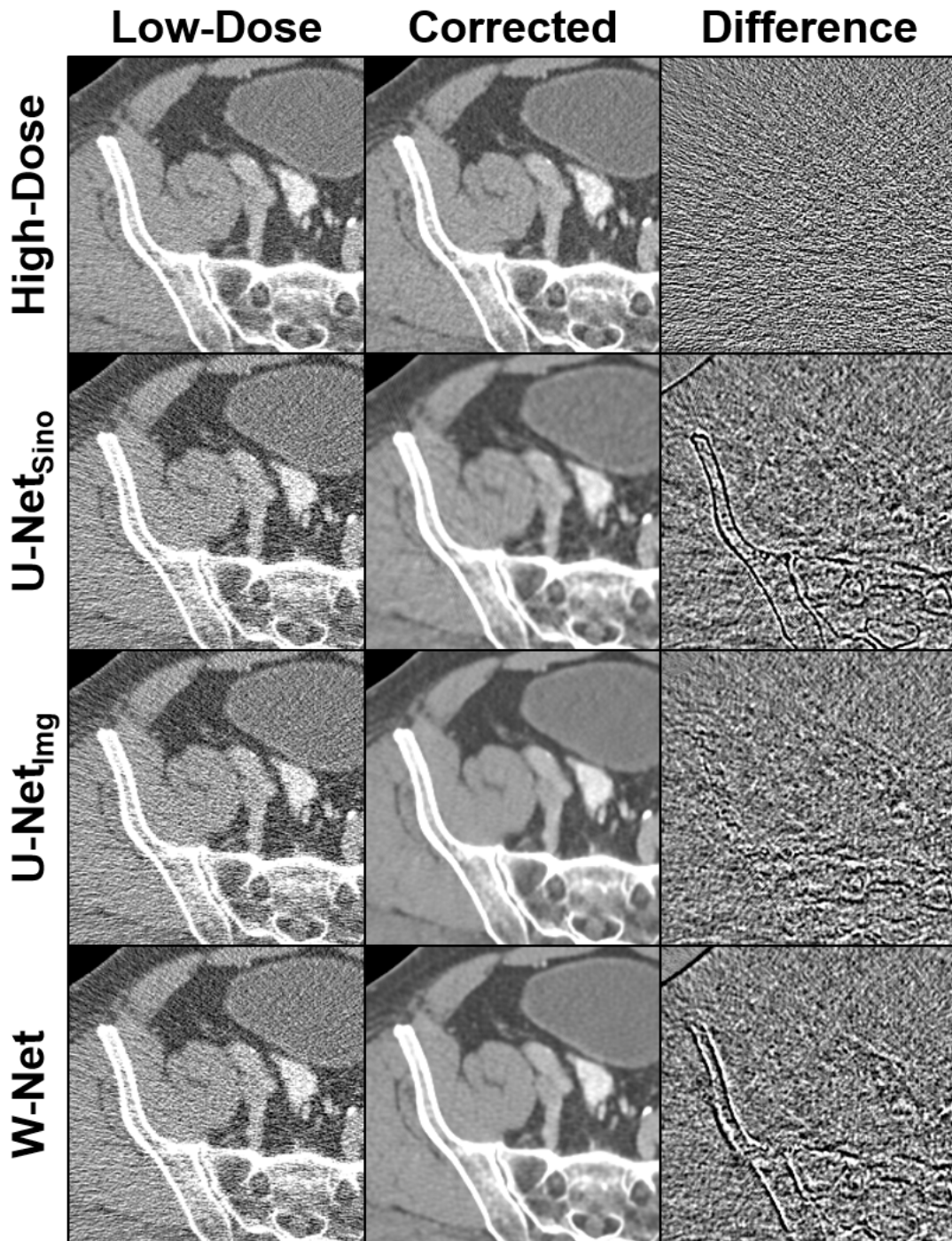


Figure 4.11: Network architecture ablation study denoising results for $N = 229$. Difference images are with respect to the ground truth. $C = 0$ HU, $W = 500$ HU for CT images, $C = 0$ HU, $W = 100$ HU for difference images.

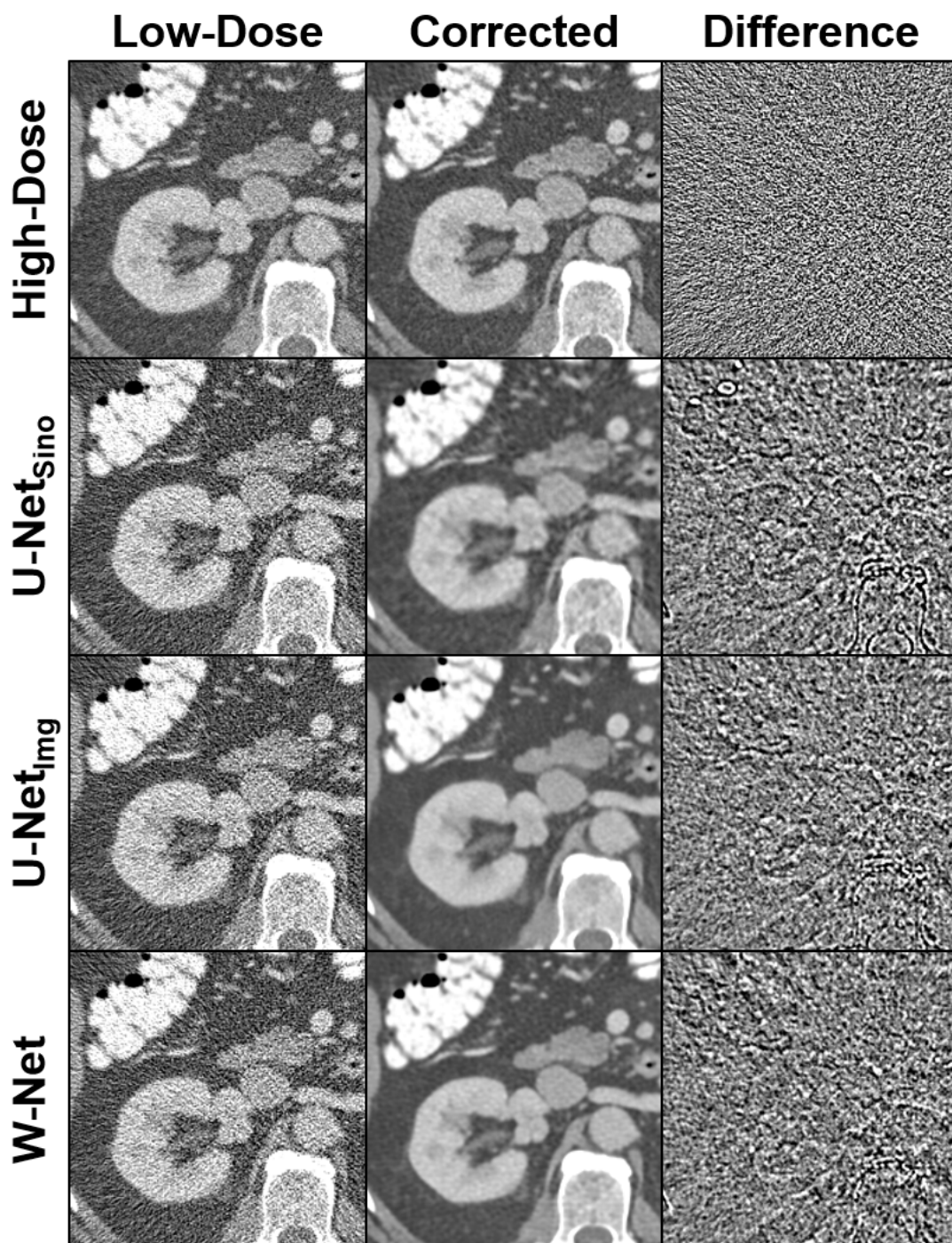


Figure 4.12: Network architecture ablation study denoising results for $N = 229$. Difference images are with respect to the ground truth. $C = 0$ HU, $W = 500$ HU for CT images, $C = 0$ HU, $W = 100$ HU for difference images.

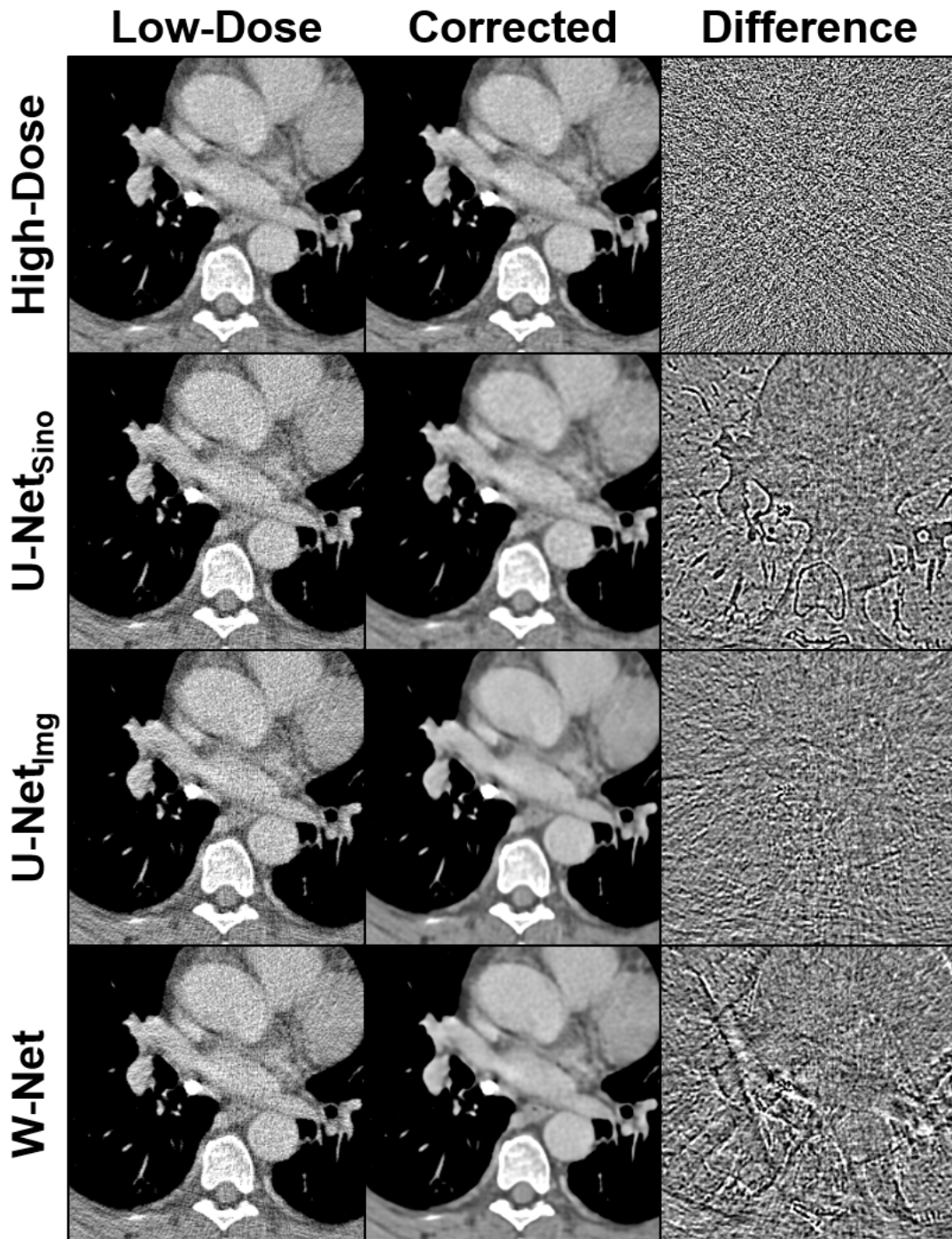


Figure 4.13: Network architecture ablation study denoising results for $N = 229$. Difference images are with respect to the ground truth. $C = 0$ HU, $W = 500$ HU for CT images, $C = 0$ HU, $W = 100$ HU for difference images.

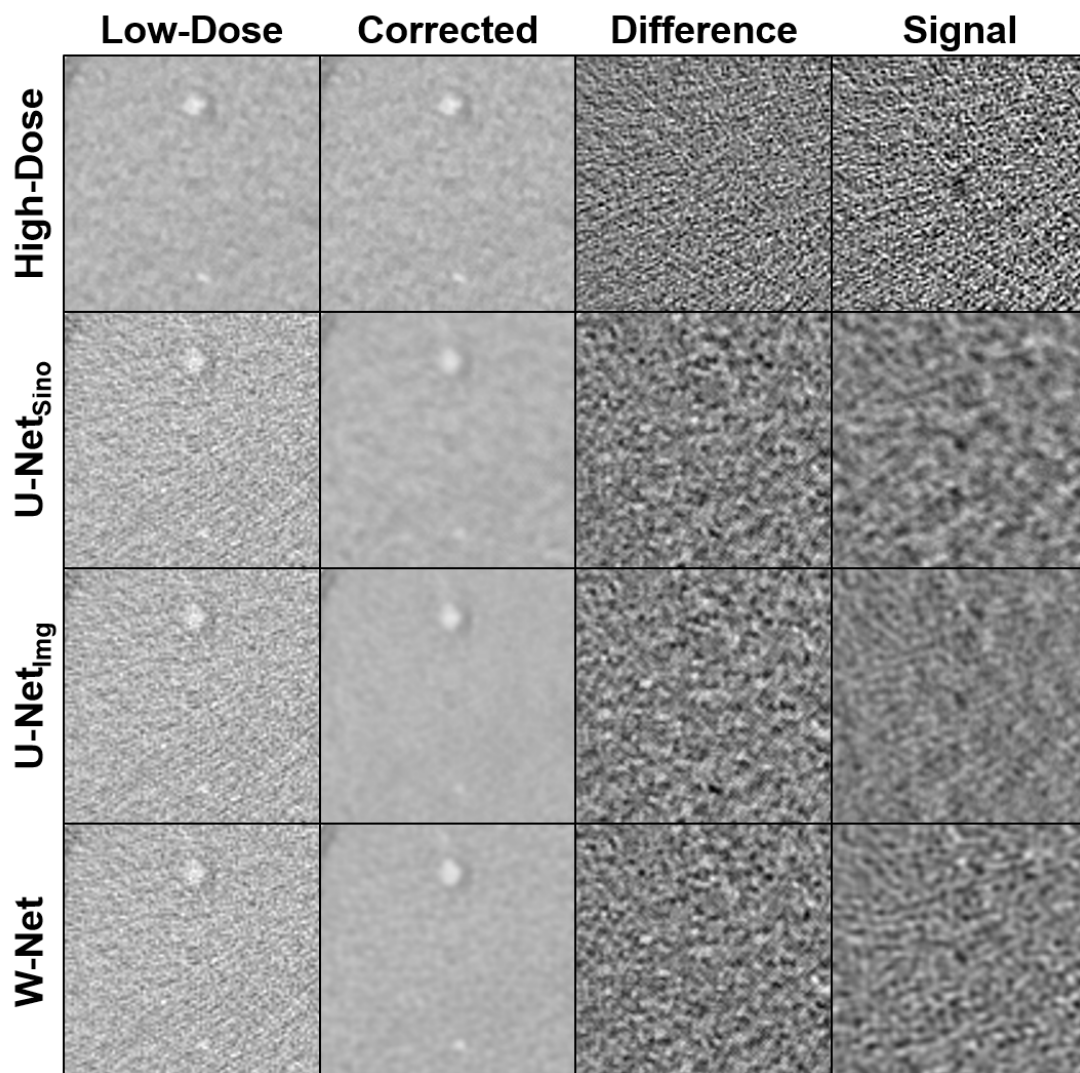


Figure 4.14: Network ablation study denoising results for the model observer study with a 2 mm lesion and $N = 229$. First row shows high-dose image, ground truth, difference of high-dose to ground truth, and high-dose signal image. Difference images are with respect to the ground truth. $C = 0$ HU, $W = 500$ HU for CT images, $C = 0$ HU, $W = 100$ HU for difference images.

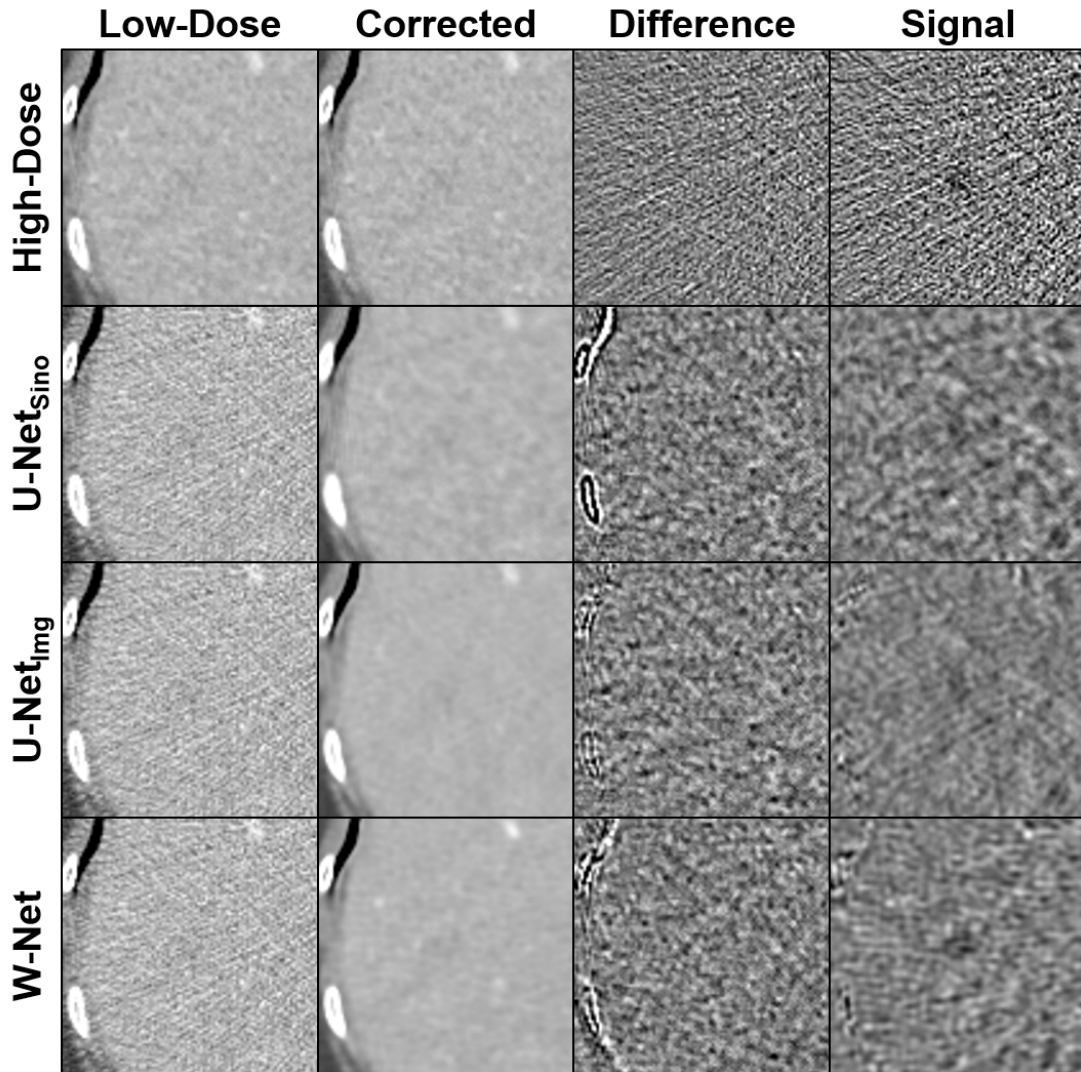


Figure 4.15: Network architecture ablation study denoising results for the model observer study with a 4 mm lesion and $N = 229$. First row shows high-dose image, ground truth, difference of high-dose to ground truth, and high-dose signal image. Difference images are with respect to the ground truth. $C = 0$ HU, $W = 500$ HU for CT images, $C = 0$ HU, $W = 100$ HU for difference images.

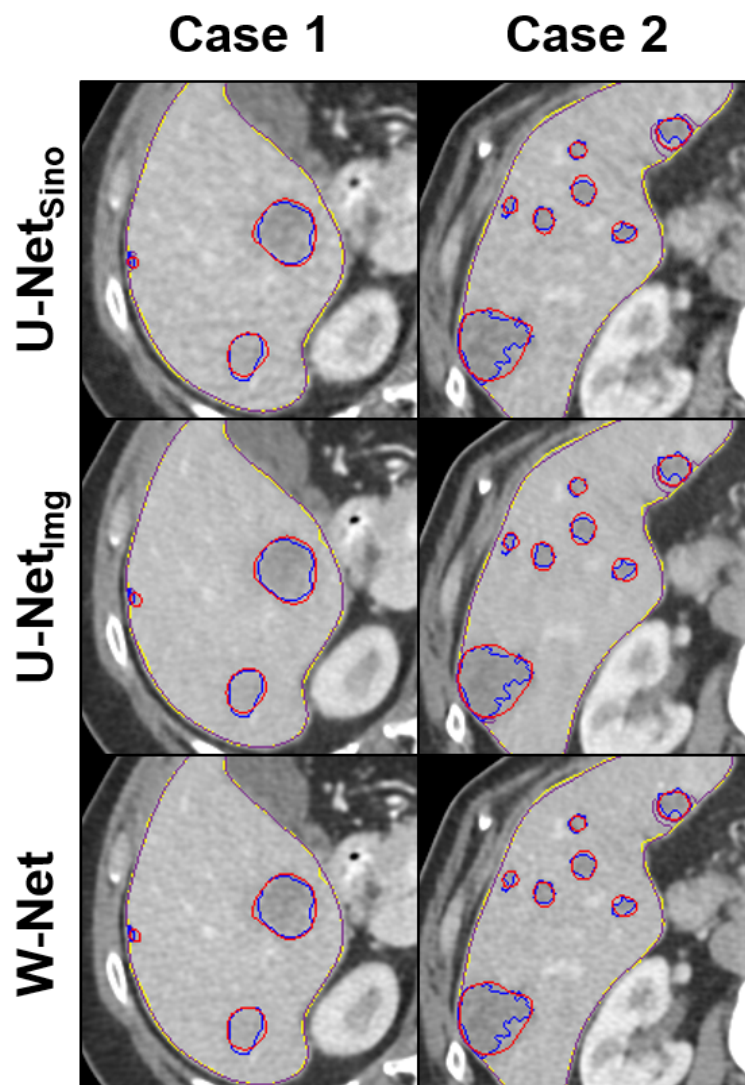


Figure 4.16: Network architecture ablation study segmentation contours of liver and liver lesions after denoising of low-dose CT with $N = 229$. First row shows high-dose image, ground truth, difference of high-dose to ground truth, and high-dose signal image. Liver ground truth in yellow, segmented liver in purple, ground truth lesions in blue, segmented lesions in red. $C = 25$ HU, $W = 450$ HU.

Lesion	U-Net _{Sino}	U-Net _{Img}	W-Net	Low-Dose	High-Dose
2 mm	0.8702	0.8625	0.8704	0.8691	0.9348
	0.8755	0.8680	0.8758	0.8745	0.9391
4 mm	0.8347	0.8327	0.8401	0.8258	0.8810
	0.8402	0.8381	0.8455	0.8312	0.8863

Table 4.7: 95% confidence intervals of AUC estimates of the SKE-BKS model observer for the network architecture ablation study with $N = 229$.

Metric	U-Net _{Sino}	U-Net _{Img}	W-Net
Per Case Dice	0.825	0.811	0.819
Global Dice	0.925	0.923	0.926
RVD	-0.020	-0.058	-0.004
ASD	0.661	0.678	0.637
RMSD	0.970	0.986	0.928
MSD	4.351	4.279	4.112
Recall	0.630	0.639	0.644
Precision	0.739	0.595	0.673

Table 4.8: Network architecture ablation study results of the segmentation of liver lesions after CNN denoising with $N = 229$.

4.3.2 Loss Function

Apart from the network architecture itself, the loss function plays a vital part in how the denoising will be performed. Therefore, this subsection investigates how removing the perceptual loss component and only using MSE for training influences the W-Net output. Figure 4.17, 4.18, and 4.19 show three cases that were corrected with either MSE or MSE with perceptual, as presented in section 4.1. As apparent, the MSE loss causes reduced image noise, but only by smoothing the images. Using perceptual loss ensures a similar noise texture compared to the ground truth. The same is visible in Figures 4.20 and 4.21, which highlight the MO patches. The signal is generally more visible when using perceptual loss. Table 4.9 gives the conventional image metrics. The network trained only on MSE yields significantly lower RMSE values and better SSIM for $N = 342$, 229, and 153. Both networks are able to reach the image quality of the high-dose reference.

Table 4.10 presents the quantitative results of the MO study. In contrast to the conventional metrics, the model observer study indicates that training with perceptual loss yields far superior metric scores for both lesion sizes. Notably, MSE does not improve the AUC score relative to the low-dose CT for several cases. Finally, Figure 4.22 and Table 4.11 display the segmentation results. For sparse-view CT, a small non-existing lesion is predicted by both networks. Comparing the values to Table 4.3,

Dataset		$N = 512$	$N = 342$	$N = 229$	$N = 153$	$N = 102$	High-Dose
RMSE	Low-Dose	47.056	47.765	49.460	46.229	72.290	21.417
	MSE	20.891	13.466	13.733	13.945	21.816	21.417
	MSE+Perc	17.783	17.661	17.992	17.992	19.064	21.417
SSIM	Low-Dose	0.7846	0.7823	0.7790	0.7661	0.7385	0.9020
	MSE	0.9431	0.9536	0.9528	0.9519	0.9397	0.9020
	MSE+Perc	0.9479	0.9480	0.9482	0.9461	0.9415	0.9020

Table 4.9: Conventional quantitative results for MSE and MSE with perceptual loss functions using RMSE (in HU) and SSIM.

Dataset		$N = 512$	$N = 342$	$N = 229$	$N = 153$	$N = 102$	High-Dose
2 mm	Low-Dose	0.8720	0.8705	0.8691	0.8651	0.8376	0.9348
		0.8773	0.8758	0.8745	0.8705	0.8430	0.9391
	MSE	0.8687	0.8499	0.8475	0.8593	0.8385	0.9348
		0.8741	0.8554	0.8530	0.8647	0.8457	0.9391
	MSE+Perc	0.8800	0.8641	0.8704	0.8736	0.8661	0.9348
		0.8853	0.8696	0.8758	0.8790	0.8715	0.9391
4 mm	Low-Dose	0.8290	0.8273	0.8258	0.8211	0.8095	0.8810
		0.8327	0.8327	0.8312	0.8265	0.8149	0.8863
	MSE	0.8460	0.8158	0.8119	0.8295	0.8119	0.8810
		0.8515	0.8212	0.8173	0.8349	0.8173	0.8863
	MSE+Perc	0.8491	0.8405	0.8401	0.8384	0.8422	0.8810
		0.8546	0.8460	0.8455	0.8438	0.8476	0.8863

Table 4.10: 95% confidence intervals of AUC of the model observer study for MSE and MSE with perceptual loss functions.

the surface metrics indicate that perceptual loss improves image quality. Otherwise, no clear trends are noticeable.

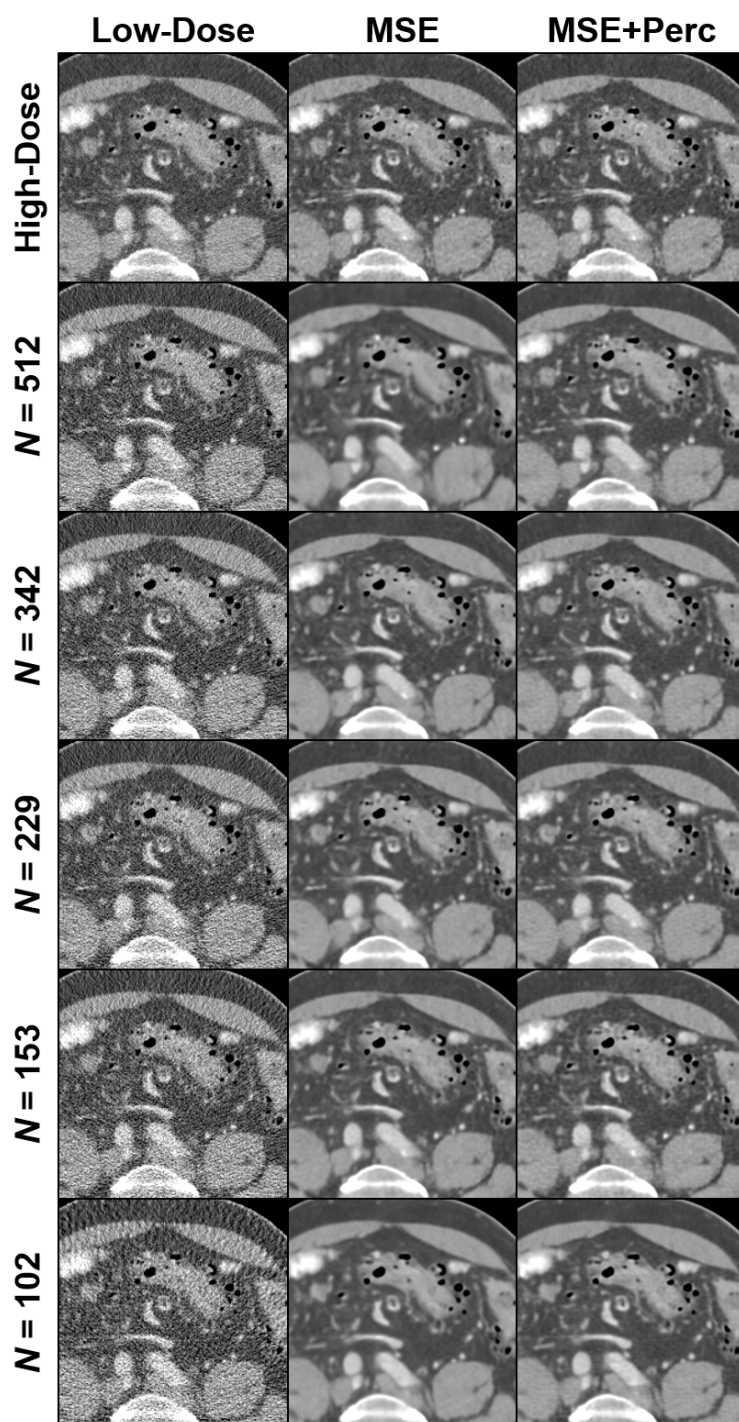


Figure 4.17: Loss function study denoising results for MSE and MSE with perceptual loss for a pelvis scan. $C = 0$ HU, $W = 500$ HU.

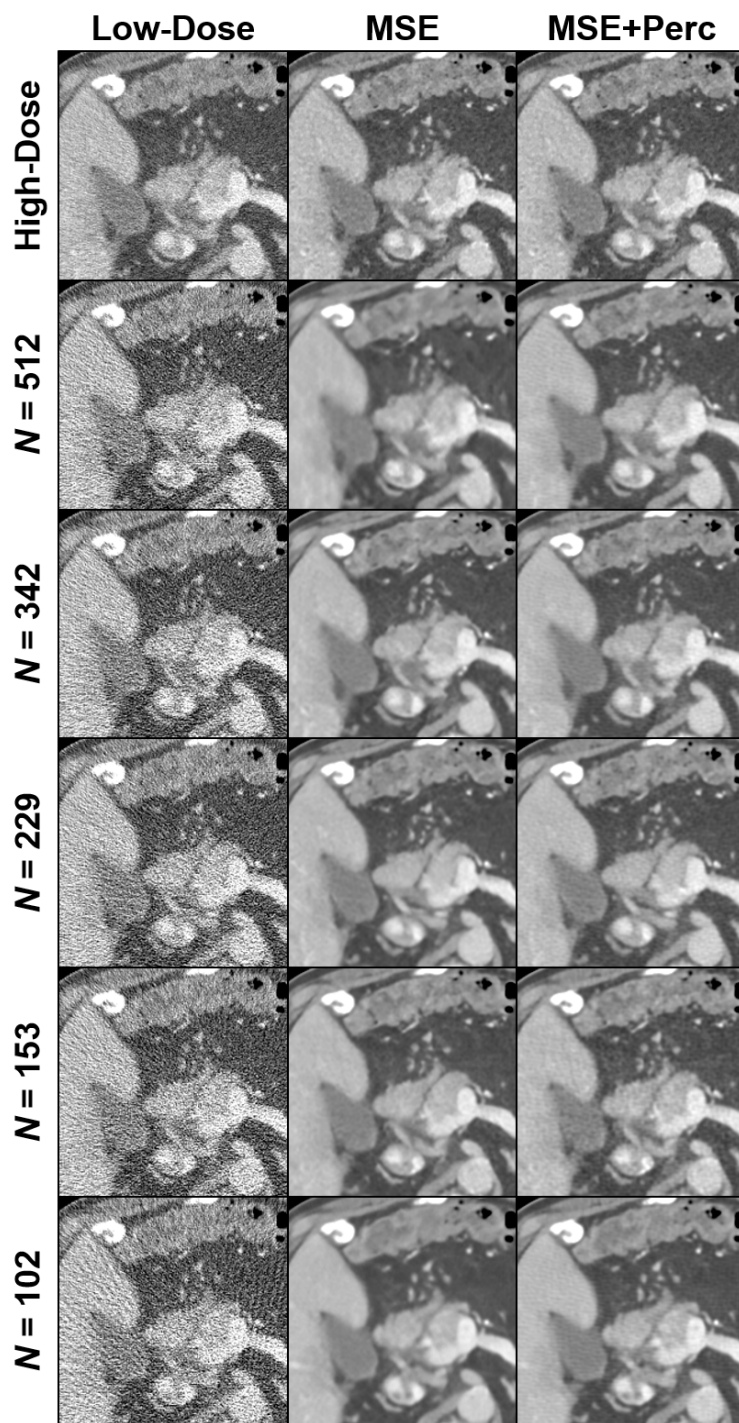


Figure 4.18: Loss function study denoising results for MSE and MSE with perceptual loss for an abdomen scan. $C = 0$ HU, $W = 500$ HU.

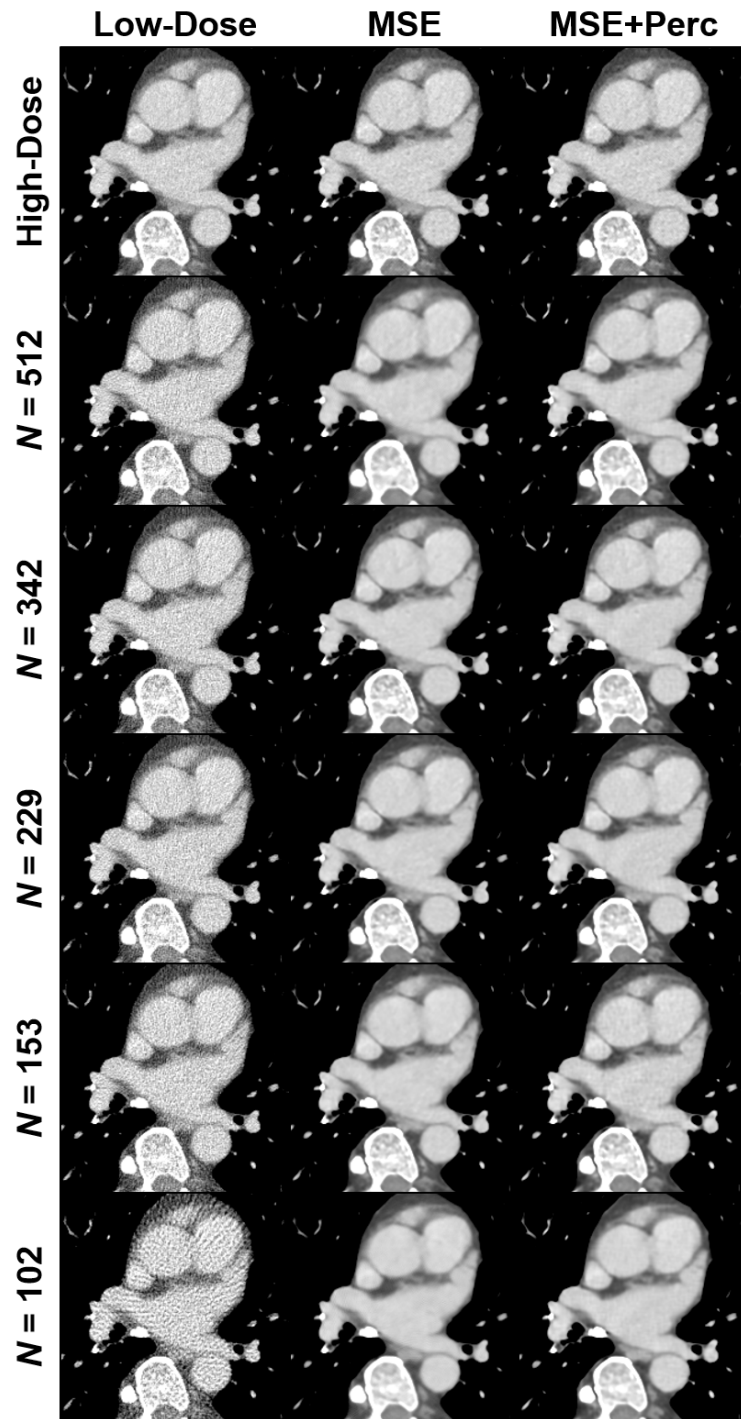


Figure 4.19: Loss function study denoising results for MSE and MSE with perceptual loss for a thorax scan. $C = 0$ HU, $W = 500$ HU.

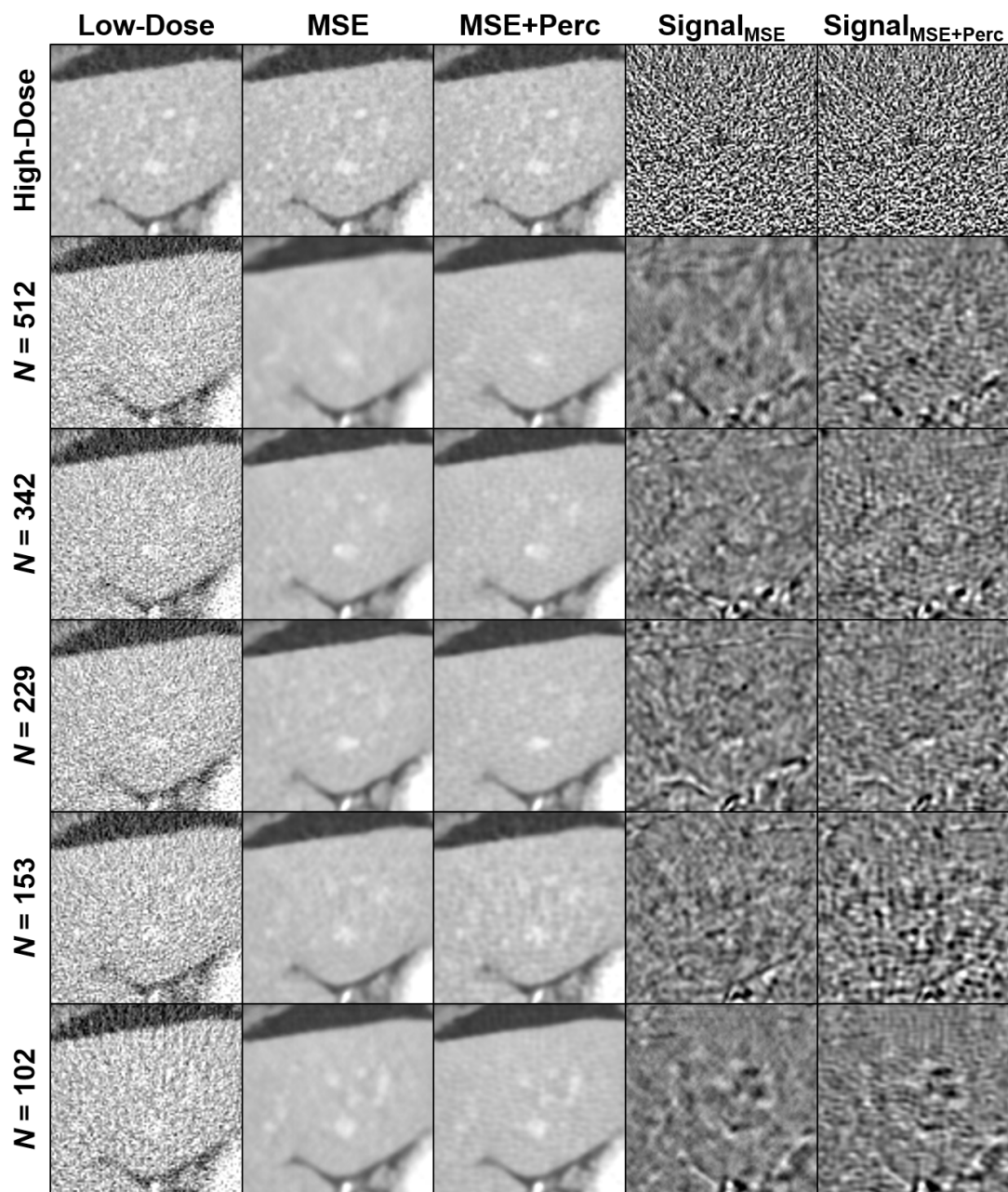


Figure 4.20: Loss function study denoising results for the model observer study with a 2 mm lesion. First row shows high-dose image, ground truth, difference of high-dose to ground truth, and high-dose signal image. $C = 0$ HU, $W = 500$ HU.

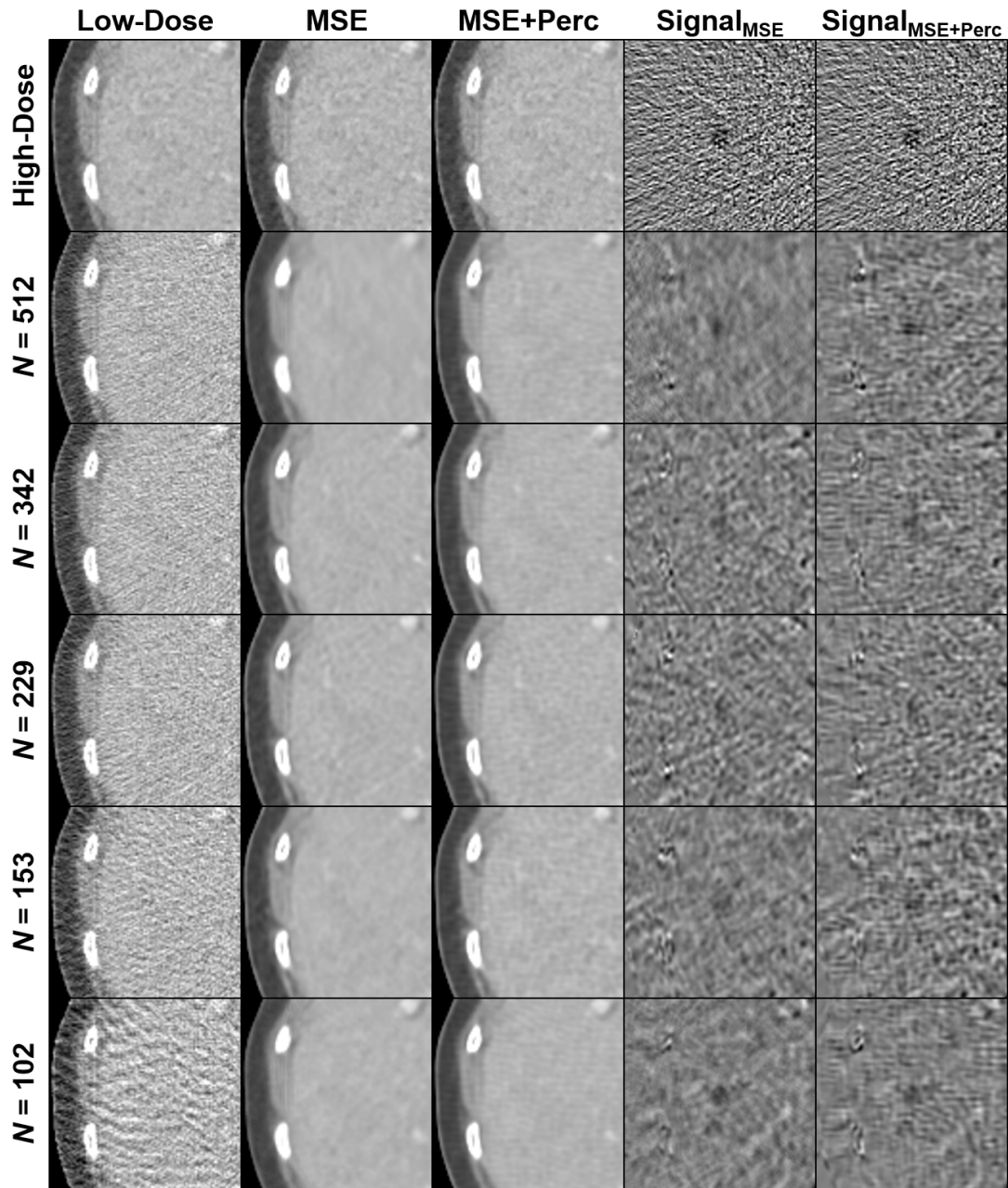


Figure 4.21: Loss function study denoising results for the model observer study with a 4 mm lesion. First row shows high-dose image, ground truth, difference of high-dose to ground truth, and high-dose signal image. $C = 0$ HU, $W = 500$ HU.

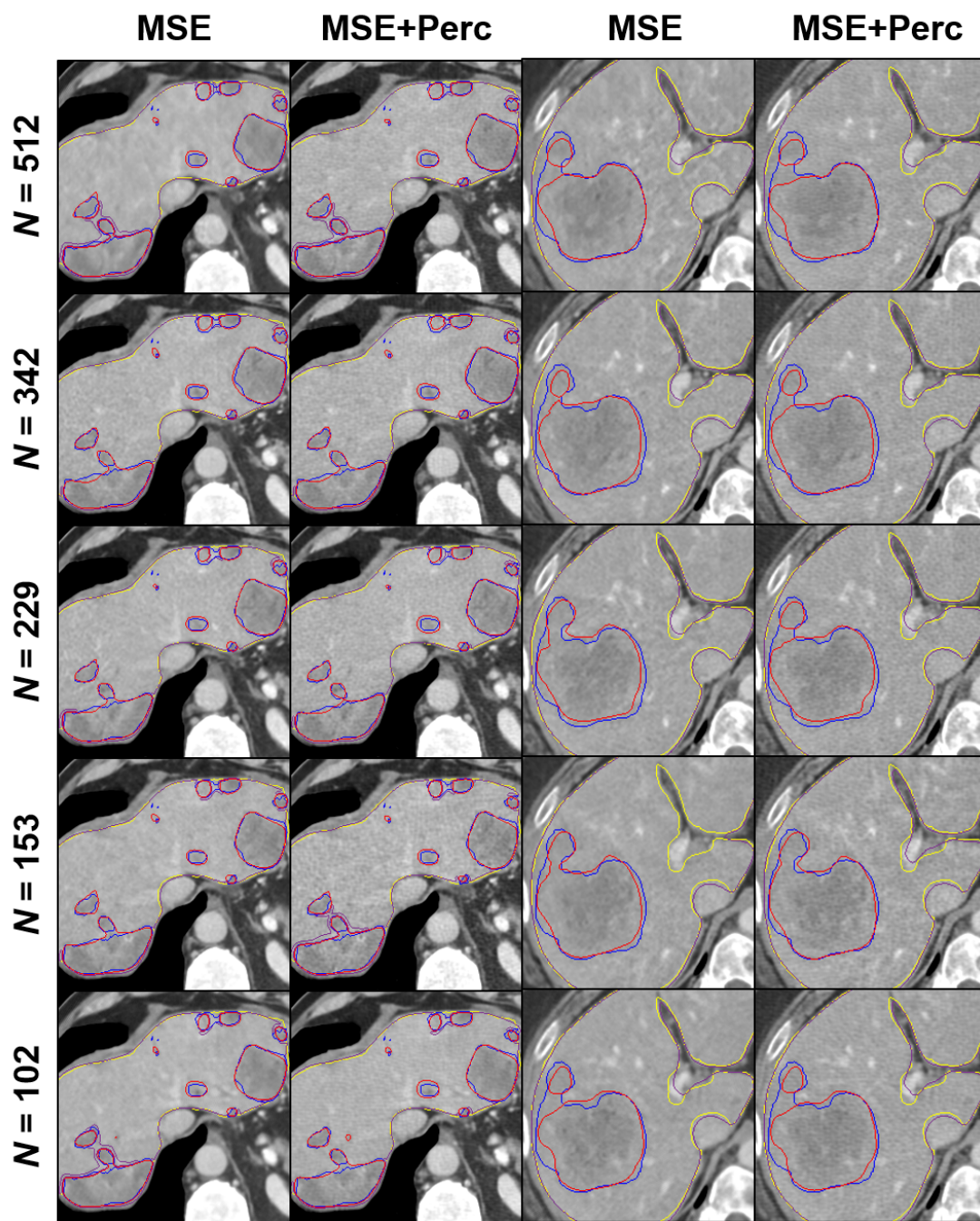


Figure 4.22: Loss function study segmentation contours of liver and liver lesions after denoising of low-dose CT. Liver ground truth in yellow, segmented liver in purple, ground truth lesions in blue, segmented lesions in red. $C = 25$ HU, $W = 450$ HU.

Metric	$N = 512$	$N = 342$	$N = 229$	$N = 153$	$N = 102$
Per Case Dice	0.809	0.834	0.818	0.835	0.805
Global Dice	0.917	0.925	0.925	0.924	0.921
RVD	-0.124	0.009	-0.001	-0.045	0.008
ASD	0.728	0.668	0.650	0.657	0.678
RMSD	1.037	0.984	0.942	0.949	0.990
MSD	4.331	4.341	4.184	4.070	4.402
Recall	0.635	0.651	0.653	0.657	0.655
Precision	0.553	0.652	0.690	0.636	0.572

Table 4.11: Quantitative results of the segmentation of liver lesions after CNN denoising with MSE loss function.

4.4 Quality Assurance of Insufficiently Trained Networks

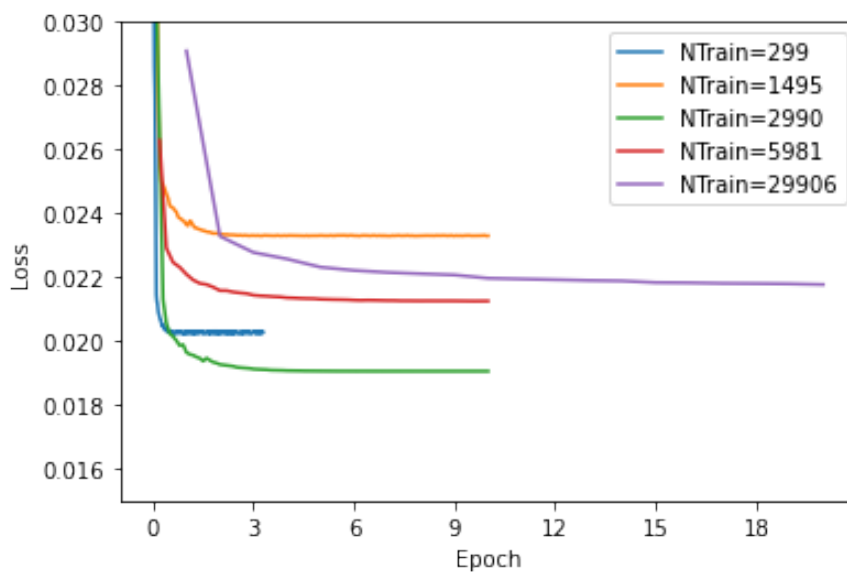
Since CNNs can be considered a black box, it is critically important that image metrics are able to identify networks that were insufficiently trained. This includes training with too few samples and stopping training too early. This section analyzes both issues, using $N = 229$ projections for all tested networks.

4.4.1 Reducing Training Set Size

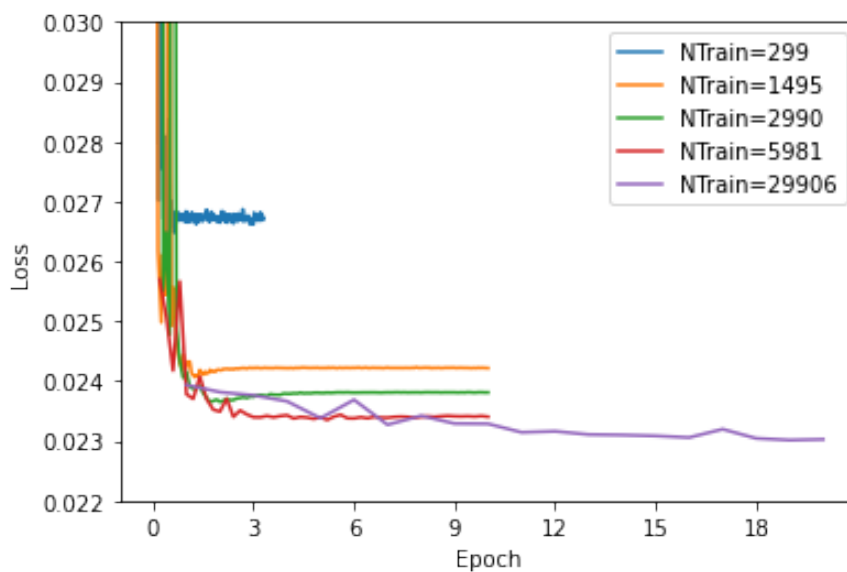
Here, the W-Net was trained with fewer than the full $N_{\text{Train}} = 29906$. Figure 4.23a and 4.23b plot the training and validation loss curves for different sizes of the train set, respectively. Note that network training was stopped early when there was no further improvement in the validation loss. It is evident that although the network with only 1% of the training images yields similar training loss than the other networks, it is clearly inferior in terms of validation loss. In general, fewer training samples correlate with lower validation loss, but not with lower training loss. This highlights the necessity of large datasets to ensure that the trained networks are able to generalize.

Figures 4.24, 4.25, 4.26 provide network results for a pelvis, abdomen and thorax scan, respectively. Image quality noticeably deteriorates as the training set is reduced. The by far worst results occur when only 1% of the full dataset is used, as prominently illustrated in the difference images. Figure 4.27 and 4.28 show the MO study patches. Again, image quality decreases when training on fewer samples. When training on 1% of the data, streak artifacts are prominent. Figure 4.29 presents the segmentation images. The segmentations, in contrast to the CT images, are similar for all training sets.

The quantitative results are listed in Table 4.12 for conventional metrics, 4.13 for AUC, and 4.14 for segmentation-based metrics. The conventional metrics suffer as fewer training samples are used, although using 20% of the dataset yielded the second worst results. However, all SSIM values are above the high-dose reference. MO and segmentation performance also decrease almost monotonically as the dataset is shrunk.



(a) Training loss



(b) Validation loss.

Figure 4.23: Loss curves for trainings with different number of samples N_{Train} . The x-axis indicates number of updates equivalent to an epoch with full train set.

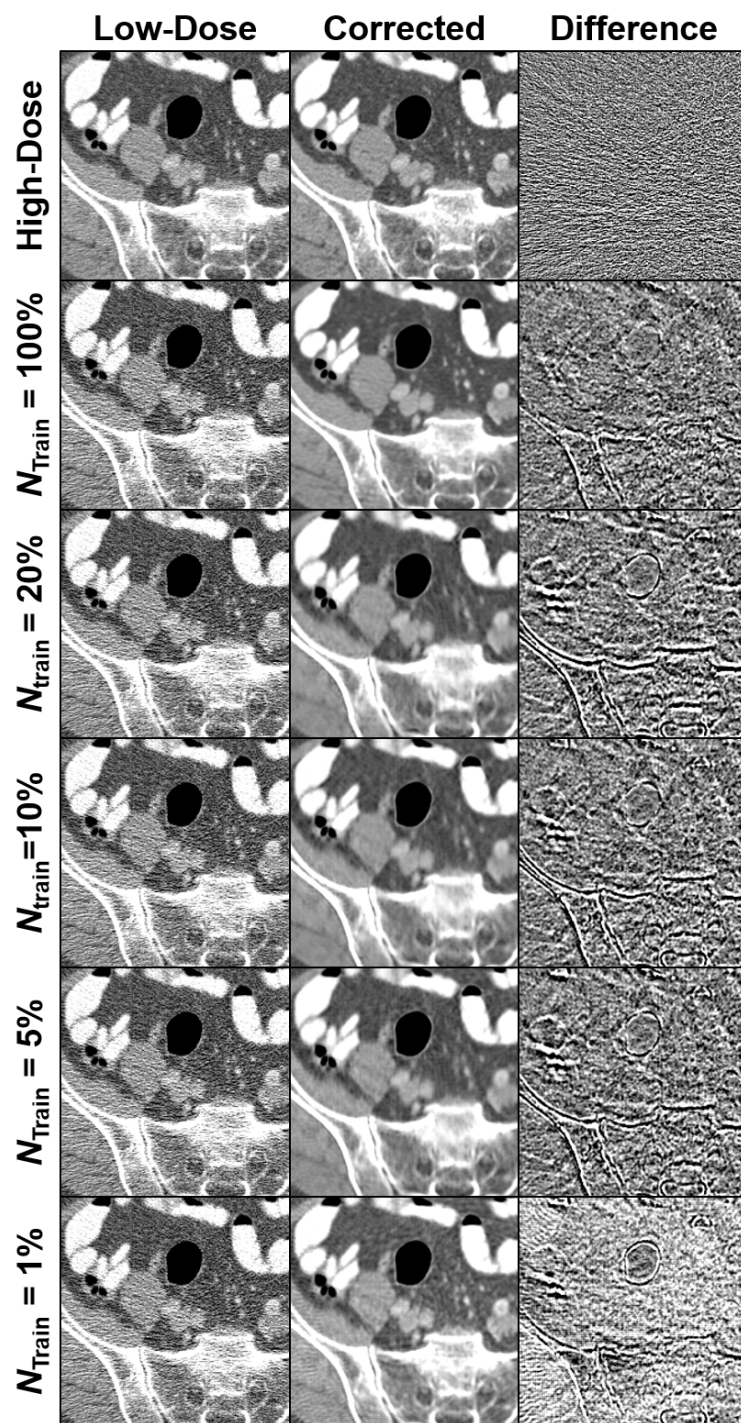


Figure 4.24: Denoising results for different sizes of the train set for a pelvis scan. $N = 229$ and $N_{\text{Train}, 100\%} = 29904$. $C = 0$ HU, $W = 500$ HU.

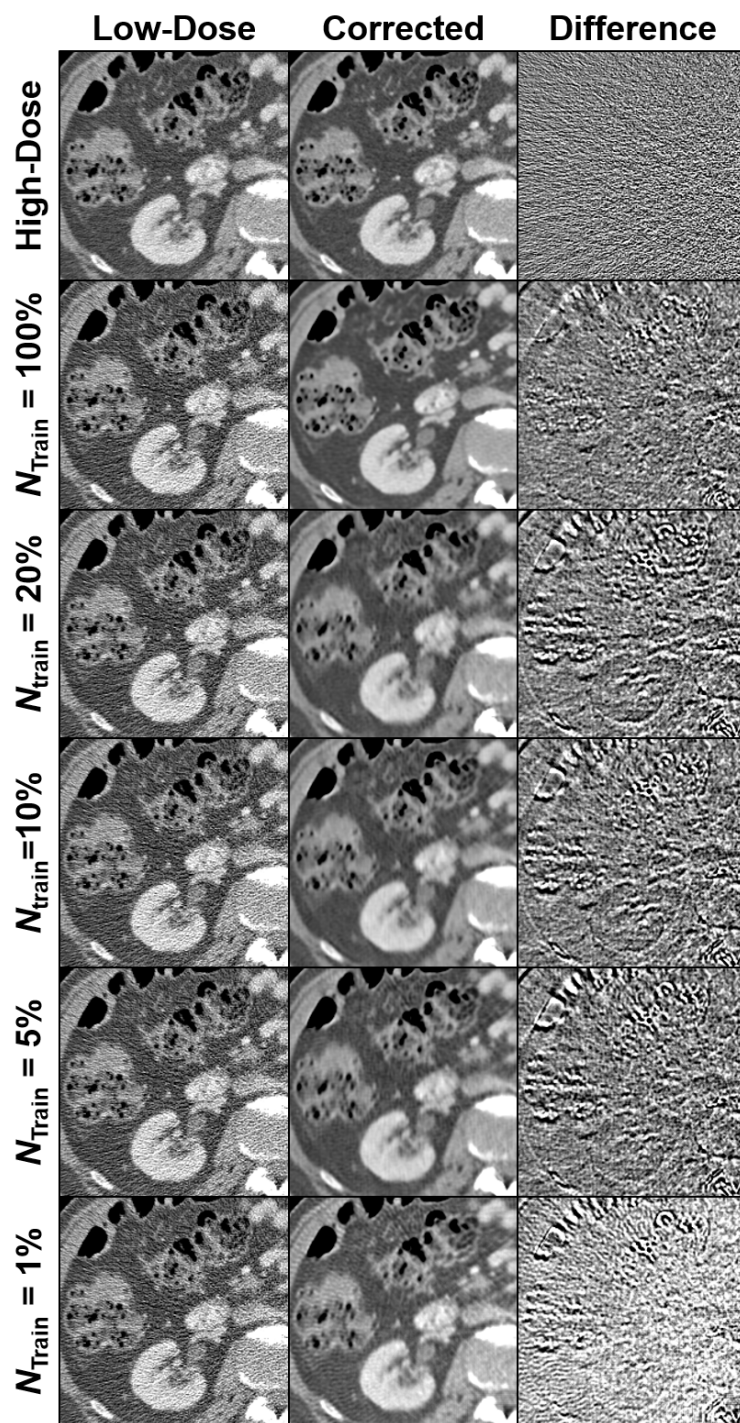


Figure 4.25: Denoising results for different sizes of the train set for an abdomen scan. $N = 229$ and $N_{\text{Train}, 100\%} = 29904$. $C = 0$ HU, $W = 500$ HU.

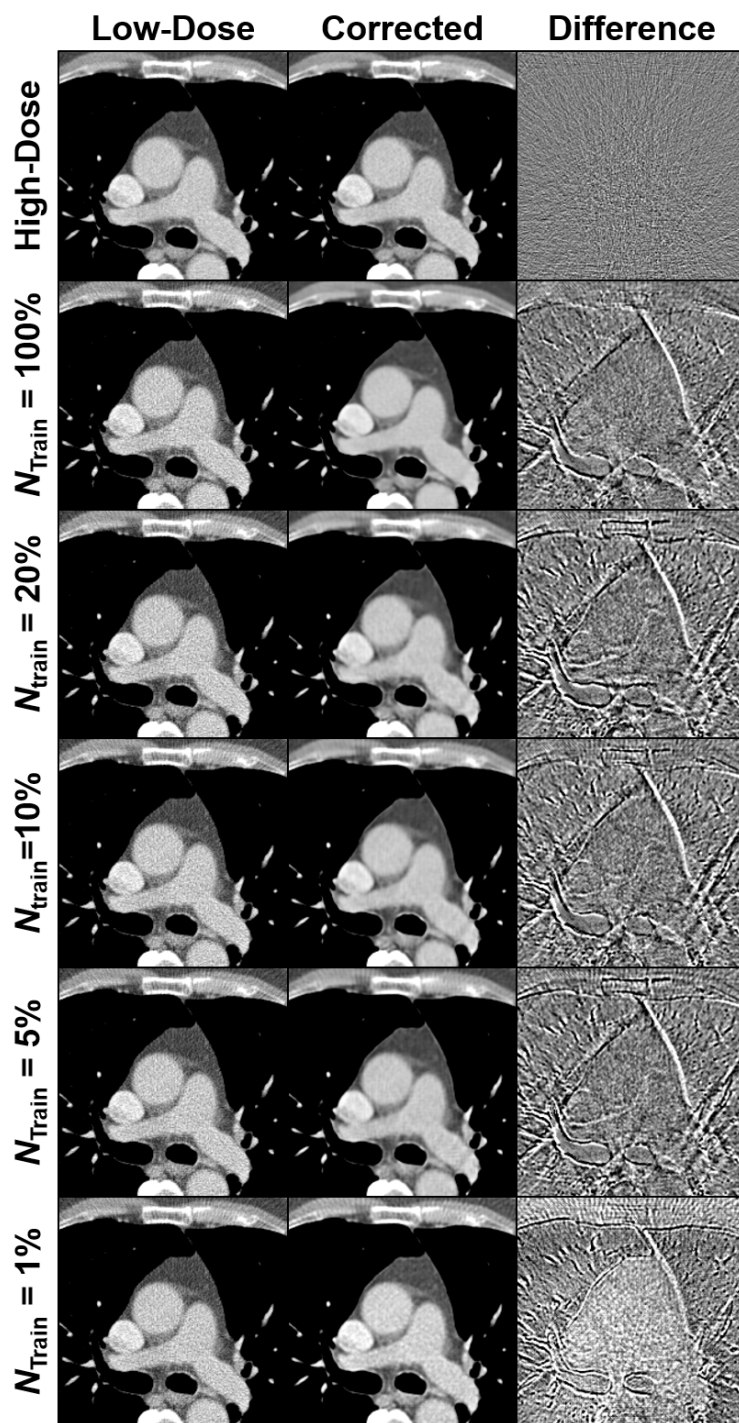


Figure 4.26: Denoising results for different sizes of the train set for a thorax scan. $N = 229$ and $N_{\text{Train}, 100\%} = 29904$. $C = 0$ HU, $W = 500$ HU.

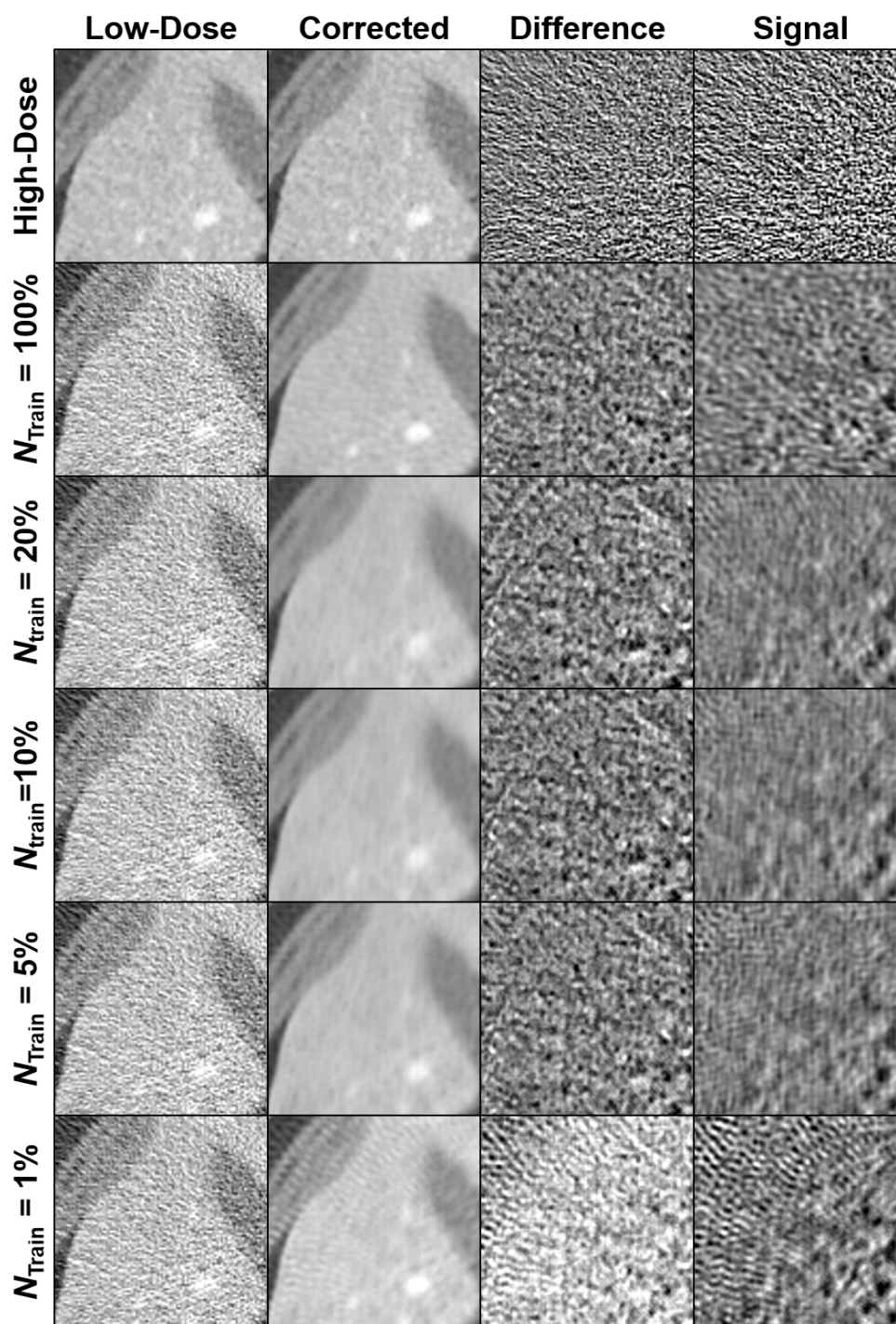


Figure 4.27: Train set size study denoising results for the model observer study with a 2 mm lesion. First row shows high-dose image, ground truth, difference of high-dose to ground truth, and high-dose signal image. $C = 0$ HU, $W = 500$ HU.

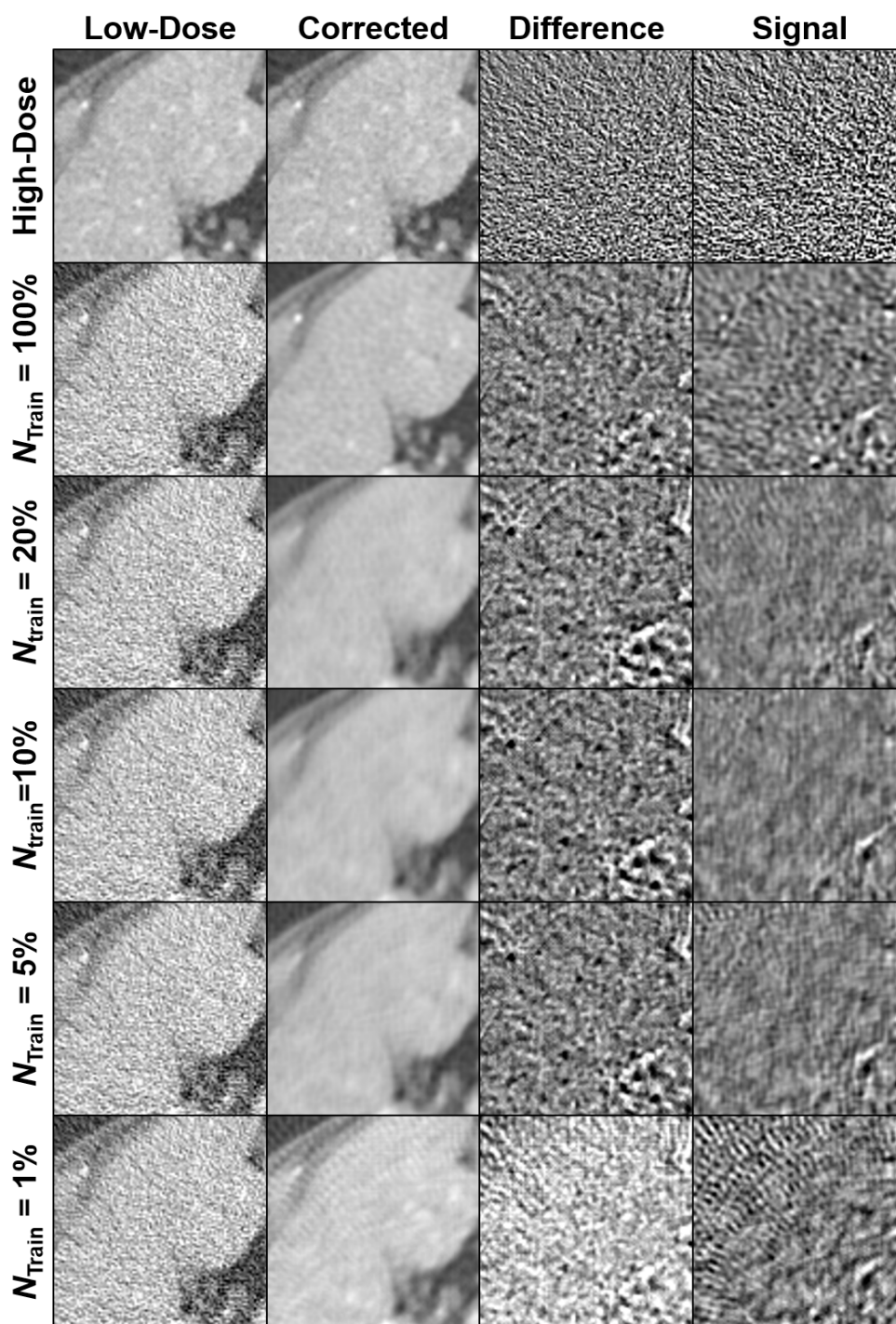


Figure 4.28: Train set size study denoising results for the model observer study with a 4 mm lesion. First row shows high-dose image, ground truth, difference of high-dose to ground truth, and high-dose signal image. $C = 0$ HU, $W = 500$ HU.

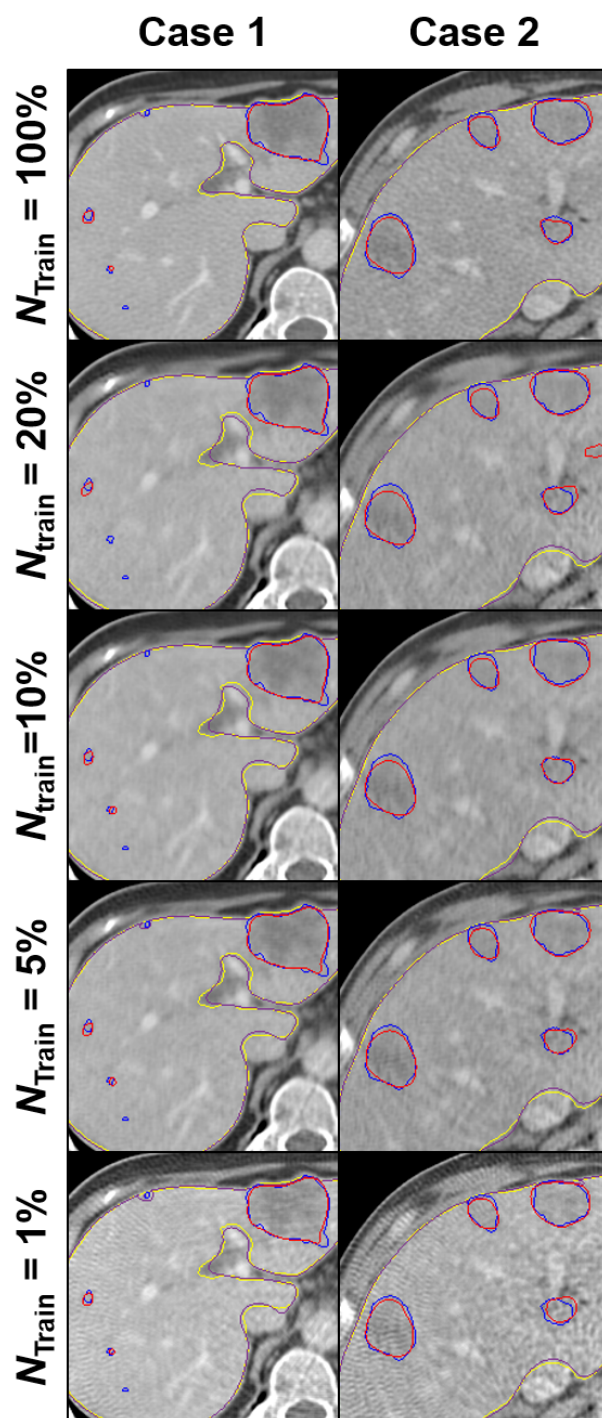


Figure 4.29: Train set size study segmentation contours of liver and liver lesions after denoising of low-dose CT. Liver ground truth in yellow, segmented liver in purple, ground truth lesions in blue, segmented lesions in red. $C = 25$ HU, $W = 450$ HU.

4.4. QUALITY ASSURANCE OF INSUFFICIENTLY TRAINED NETWORKS

Metric	100%	20%	10%	5%	1%	High-Dose
RMSE	17.992	23.634	21.605	23.002	34.094	21.417
SSIM	0.9482	0.9267	0.9340	0.9308	0.9086	0.9020

Table 4.12: Conventional quantitative results for different train set sizes relative to full set using RMSE (in HU) and SSIM.

	100%	20%	10%	5%	1%	High-Dose
2 mm	0.8704	0.8693	0.8686	0.8666	0.8596	0.9348
	0.8758	0.8748	0.8740	0.8720	0.8651	0.9391
4 mm	0.8401	0.8397	0.8361	0.8372	0.8353	0.8810
	0.8455	0.8452	0.8416	0.8426	0.8407	0.8863

Table 4.13: 95% confidence intervals of AUC of the model observer study for different train set sizes relative to full set and two lesion sizes.

Metric	100%	20%	10%	5%	1%
Per Case Dice	0.819	0.822	0.800	0.792	0.788
Global Dice	0.926	0.923	0.918	0.915	0.915
RVD	-0.004	0.018	-0.031	-0.047	-0.029
ASD	0.637	0.678	0.699	0.719	0.755
RMSD	0.928	1.000	1.003	1.031	1.099
MSD	4.112	4.521	4.560	4.763	4.789
Recall	0.644	0.646	0.638	0.626	0.621
Precision	0.673	0.706	0.615	0.589	0.608

Table 4.14: Quantitative results of the segmentation of liver lesions after CNN denoising with different relative train set sizes.

4.4.2 Reducing Number of Epochs

For this experiment, the number of training epochs was reduced. Figure 4.30, 4.31 and 4.32 show a hip, abdomen and thorax scan, respectively. For all patients, the image quality decreases as the number of epochs is reduced. This is especially evident in the hip scan, where the network leaves streaks between the bones when trained for only one or two epochs. The model observer results are depicted in Figures 4.33 and 4.34, where similar trends are visible. The difference images indicate that bone-tissue transitions are smoothed for low number of epochs. Figure 4.35 provides the segmentation images. As for previous experiments, all networks perform similarly well.

Table 4.15, 4.16, and 4.17 list the conventional, MO, and segmentation-based quantitative results. In the conventional metrics, higher number of epochs seem to correlate with worse RMSE, while SSIM is similar for all networks. Notably, $N_{\text{Epoch}} = 5$ yields the best results. Likewise, $N_{\text{Epoch}} = 5$ produces the best AUC values. Otherwise, lower number of epochs correlate with a lower image quality. Finally, no clear trend is noticeable in the segmentation metrics, which agrees with the qualitative assessment of Figure 4.35.

Metric	20	10	5	2	1	High-Dose
RMSE	17.992	17.363	14.051	15.002	15.529	21.417
SSIM	0.9482	0.9484	0.9507	0.9471	0.9465	0.9020

Table 4.15: Conventional quantitative results for different number of epochs with $N = 229$ using RMSE (in HU) and SSIM.

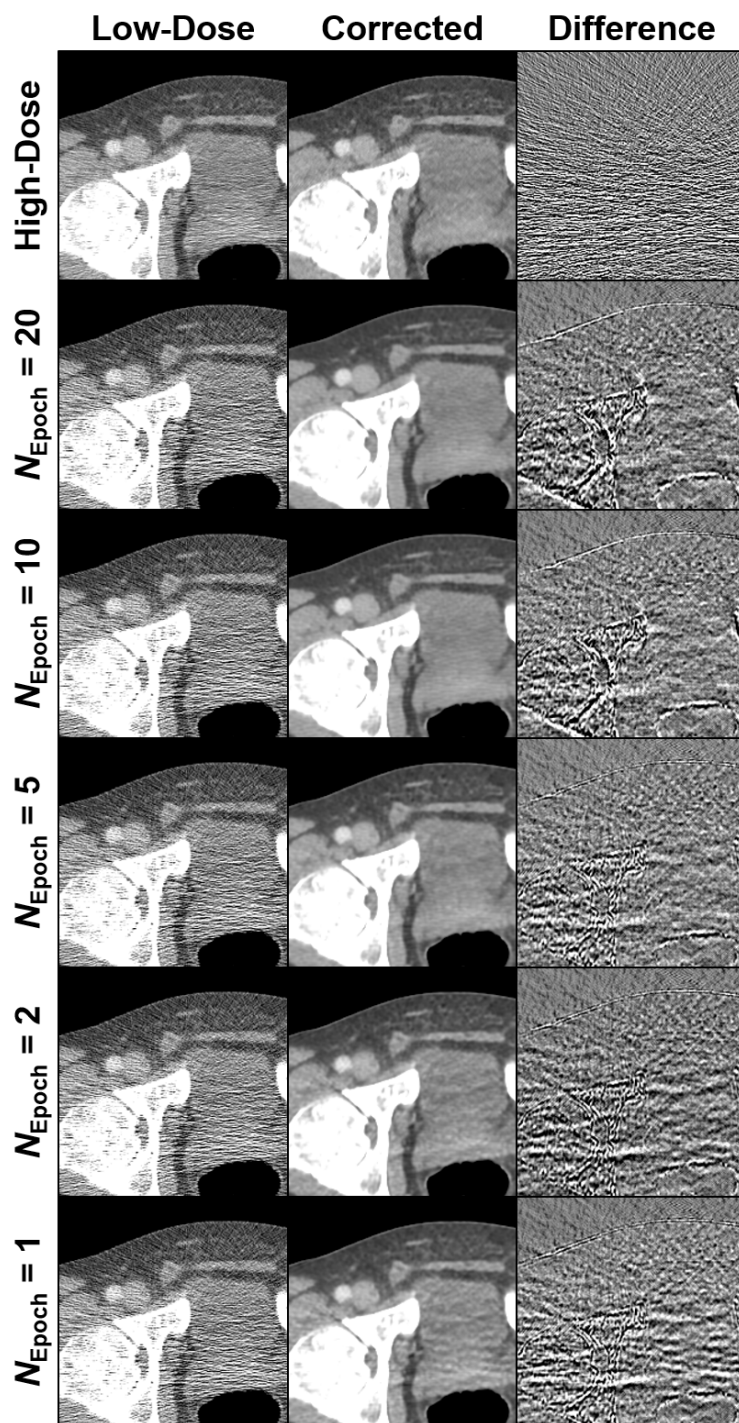


Figure 4.30: Denoising results for different numbers of epochs for a hip scan with $N = 229$. $C = 0$ HU, $W = 500$ HU.

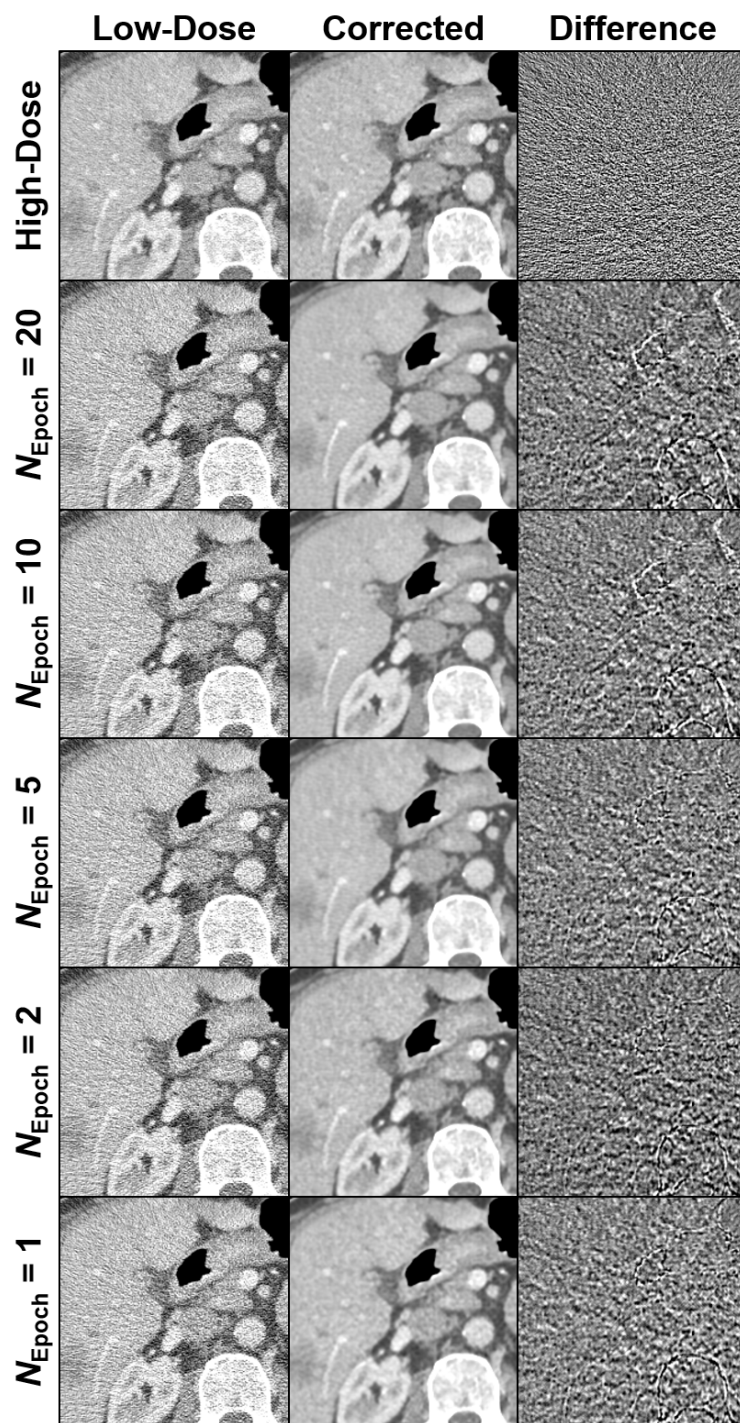


Figure 4.31: Denoising results for different numbers of epochs for an abdomen scan with $N = 229$. $C = 0$ HU, $W = 500$ HU.

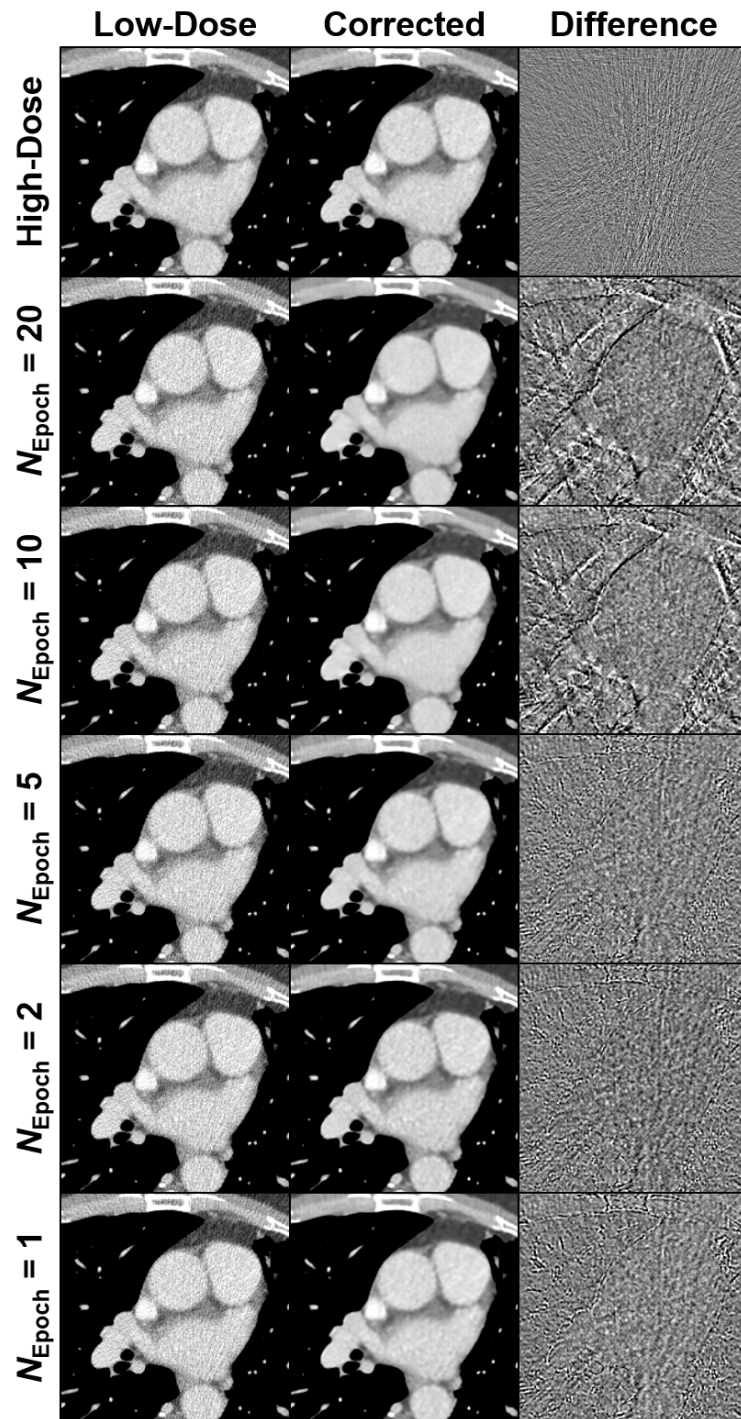


Figure 4.32: Denoising results for different numbers of epochs for a thorax scan with $N = 229$. $C = 0$ HU, $W = 500$ HU.

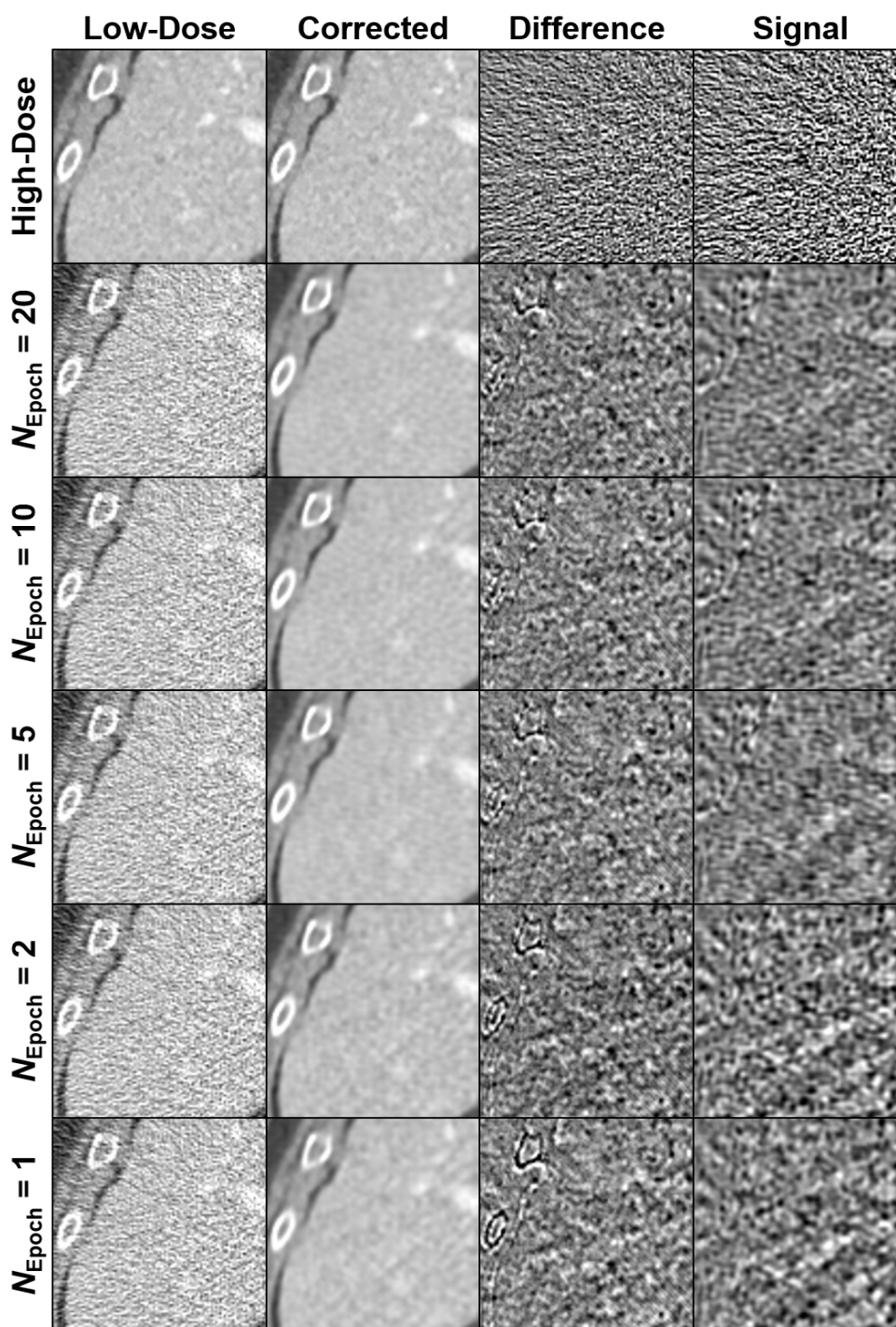


Figure 4.33: Epoch number study denoising results for the model observer study with a 2 mm lesion. First row shows high-dose image, ground truth, difference of high-dose to ground truth, and high-dose signal image. $C = 0$ HU, $W = 500$ HU.

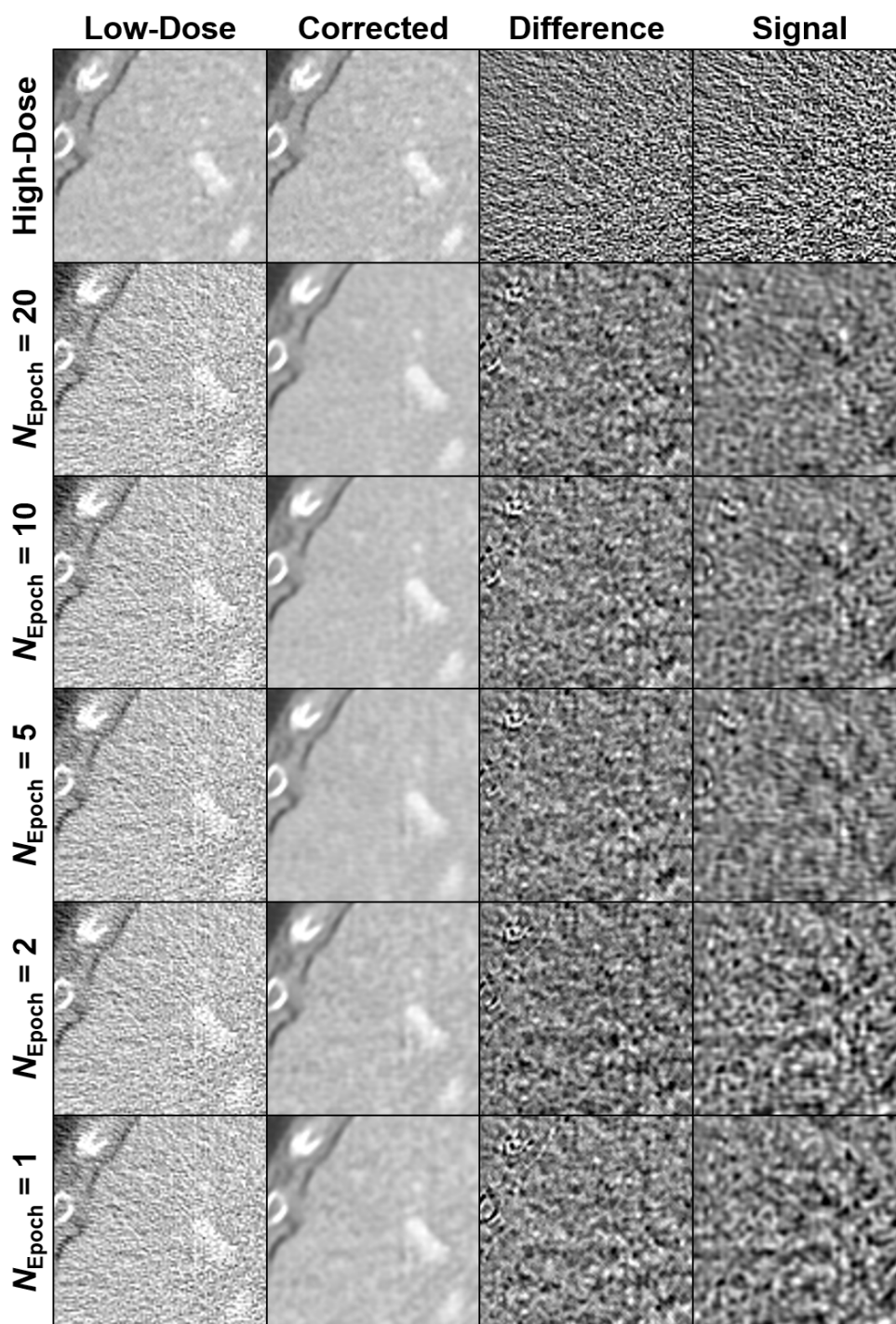


Figure 4.34: Epoch number study denoising results for the model observer study with a 4 mm lesion. First row shows high-dose image, ground truth, difference of high-dose to ground truth, and high-dose signal image. $C = 0$ HU, $W = 500$ HU.

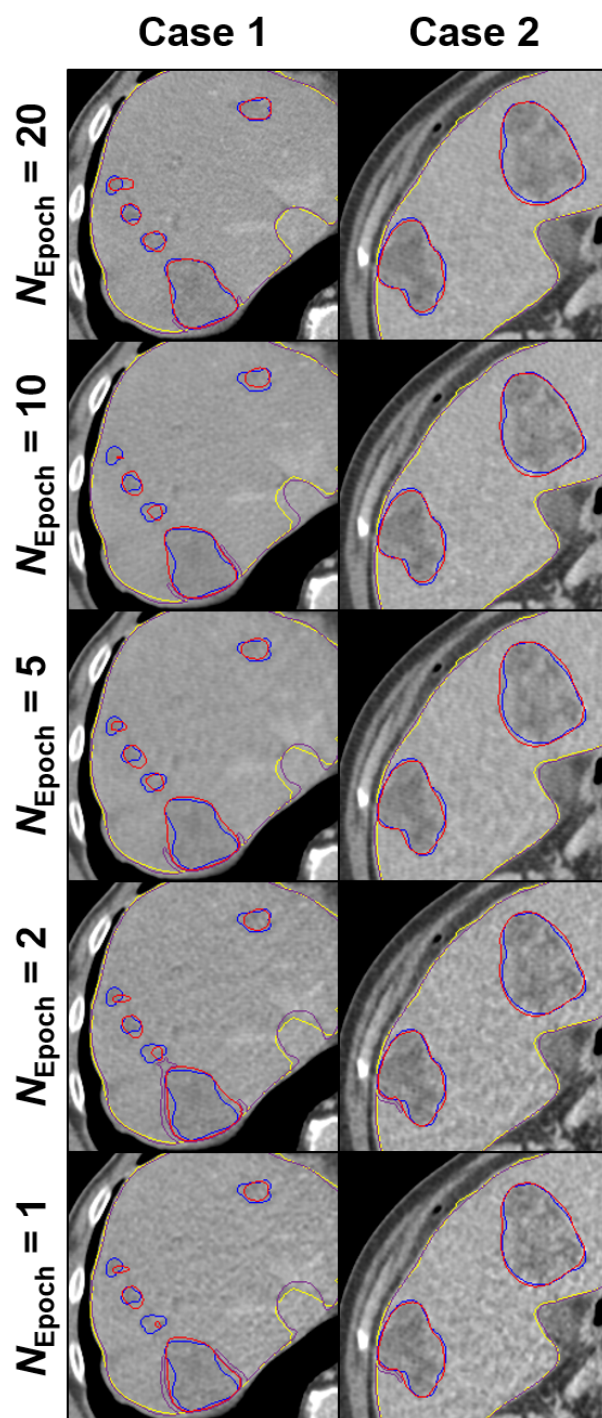


Figure 4.35: Epoch number segmentation contours of liver and liver lesions after denoising of low-dose CT. Liver ground truth in yellow, segmented liver in purple, ground truth lesions in blue, segmented lesions in red. $C = 25$ HU, $W = 450$ HU.

	20	10	5	2	1	High-Dose
2 mm	0.8704	0.8701	0.8709	0.8688	0.8521	0.9348
	0.8758	0.8754	0.8763	0.8741	0.8576	0.9391
4 mm	0.8401	0.8390	0.8401	0.8398	0.8147	0.8810
	0.8455	0.8446	0.8456	0.8453	0.8201	0.8863

Table 4.16: 95% confidence intervals of AUC of the model observer study for different number of epochs with $N = 229$ and two lesion sizes.

Metric	20	10	5	2	1
Per Case Dice	0.819	0.823	0.819	0.821	0.822
Global Dice	0.926	0.927	0.926	0.925	0.928
RVD	-0.004	0.016	0.013	-0.040	0.012
ASD	0.637	0.668	0.673	0.665	0.626
RMSD	0.928	0.990	1.001	0.976	0.916
MSD	4.112	4.494	4.550	4.350	4.092
Recall	0.644	0.621	0.623	0.646	0.646
Precision	0.673	0.793	0.768	0.659	0.685

Table 4.17: Quantitative results of the segmentation of liver lesions after CNN denoising with different number of epochs.

4.5 Detruncation

This section showcases the reconstruction of CT images from truncated raw data. Two truncation amounts were investigated, 50% and 25%.

4.5.1 50% Truncation

Figure 4.36 shows intermediate reconstruction of the DART algorithm. The algorithm is initialized with an FBP of a cosine detruncation. DART seems to converge slower for the thorax scan, which also features a hypodense artifact below the spine. The earlier iterations are blurrier, due to the stronger filtering step. The final images are close to the ground truth images in outline, but bone structures and fat tissue are not correctly resolved due to the binary nature of the DART implementation. In the third case, the DART result is significantly darker in the center than the ground truth image.

Figure 4.37 and 4.38 provide the detruncated images and sinograms, respectively. Note that for the DART and W-Net, the algorithm output is used to complete the original sinograms. DART, the W-Net, and the conventional cosine detruncation are all able to improve image quality in the original FOV. However, DART and W-Net also restore information in the eFOV. In both cases, there are inconsistencies in the transition from FOV to eFOV.

Metric	Truncated	Cosine	W-Net _{Prior}	W-Net	DART _{Prior}	DART
RMSE _{FOV}	1262.166	53.013	26.691	28.631	65.475	20.557
	702.760	58.138	23.865	25.281	87.737	92.844
	1355.924	103.173	18.134	16.454	163.097	42.402
	1270.852	127.749	27.364	29.360	93.322	16.315
RMSE _{eFOV}	873.312	258.439	121.997	120.201	140.840	127.996
	538.059	290.387	122.305	119.128	258.969	258.569
	900.285	282.489	84.569	82.512	212.77	190.303
	852.035	279.175	97.245	95.635	184.214	175.793
SSIM	0.6891	0.8252	0.8901	0.8841	0.8532	0.8807
	0.7730	0.8415	0.9252	0.9221	0.8285	0.8424
	0.6115	0.7594	0.9279	0.9206	0.7728	0.8321
	0.5746	0.8057	0.9172	0.9011	0.8460	0.8728
Dice	0.4652	0.8810	0.9780	0.9782	0.9818	0.9820
	0.5761	0.7350	0.9532	0.9563	0.8781	0.8776
	0.6069	0.8725	0.9862	0.9863	0.9604	0.9603
	0.6117	0.9009	0.9843	0.9848	0.9634	0.9634

Table 4.18: Quantitative results for 50% truncation with RMSE in HU for four test cases.

Table 4.18 gives the quantitative results for the detruncation. Here, the prior images indicate the immediate algorithm output without projection completion of the original raw data. All algorithms successfully restore image quality in the FOV, confirming the qualitative results. Both W-Net and DART are superior to the cosine detruncation. In addition, the W-Net generally outperforms the DART detruncation, both with and without projection completion.

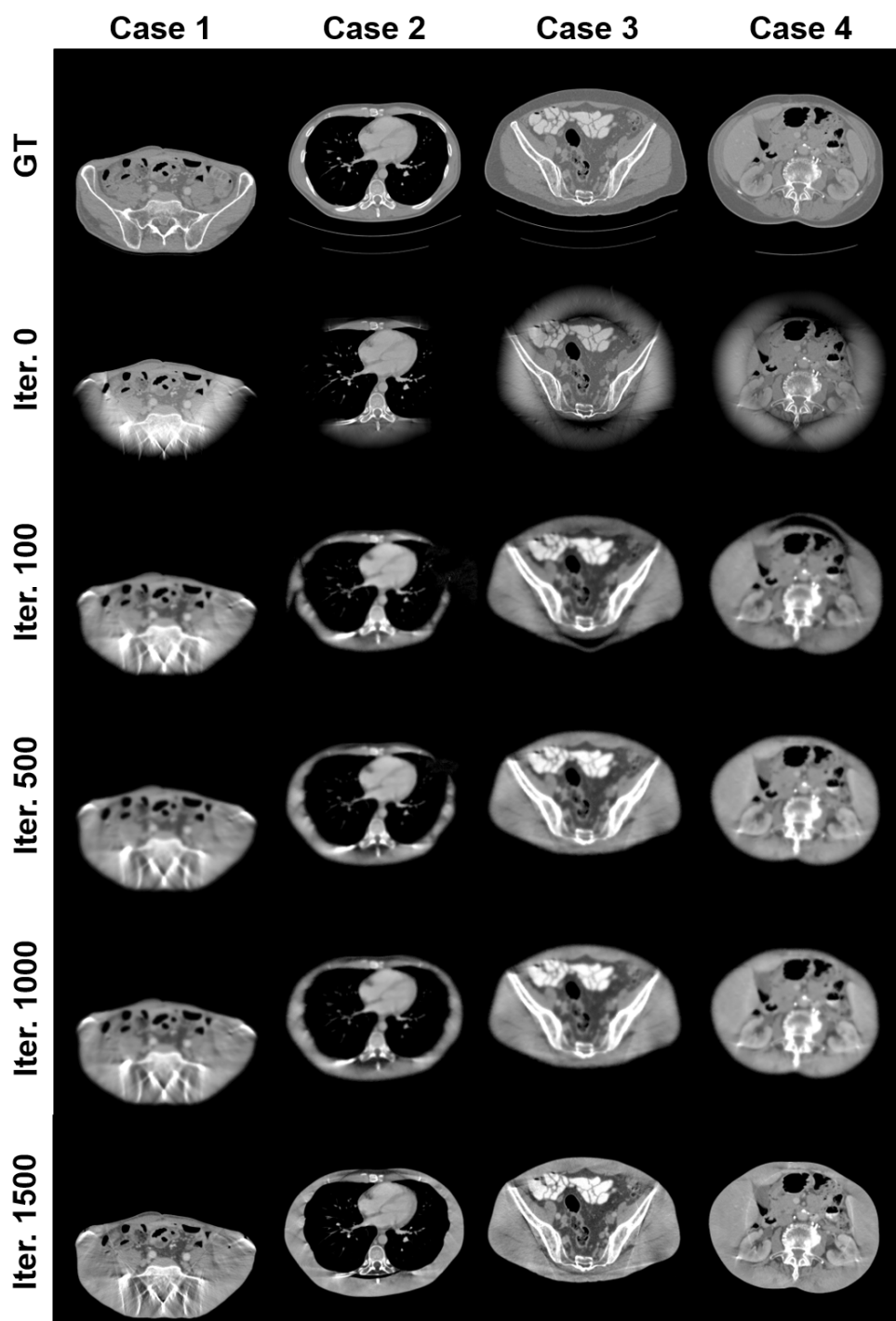


Figure 4.36: Intermediate DART iterations with 50% truncation. $C = 0$ HU, $W = 1000$ HU.

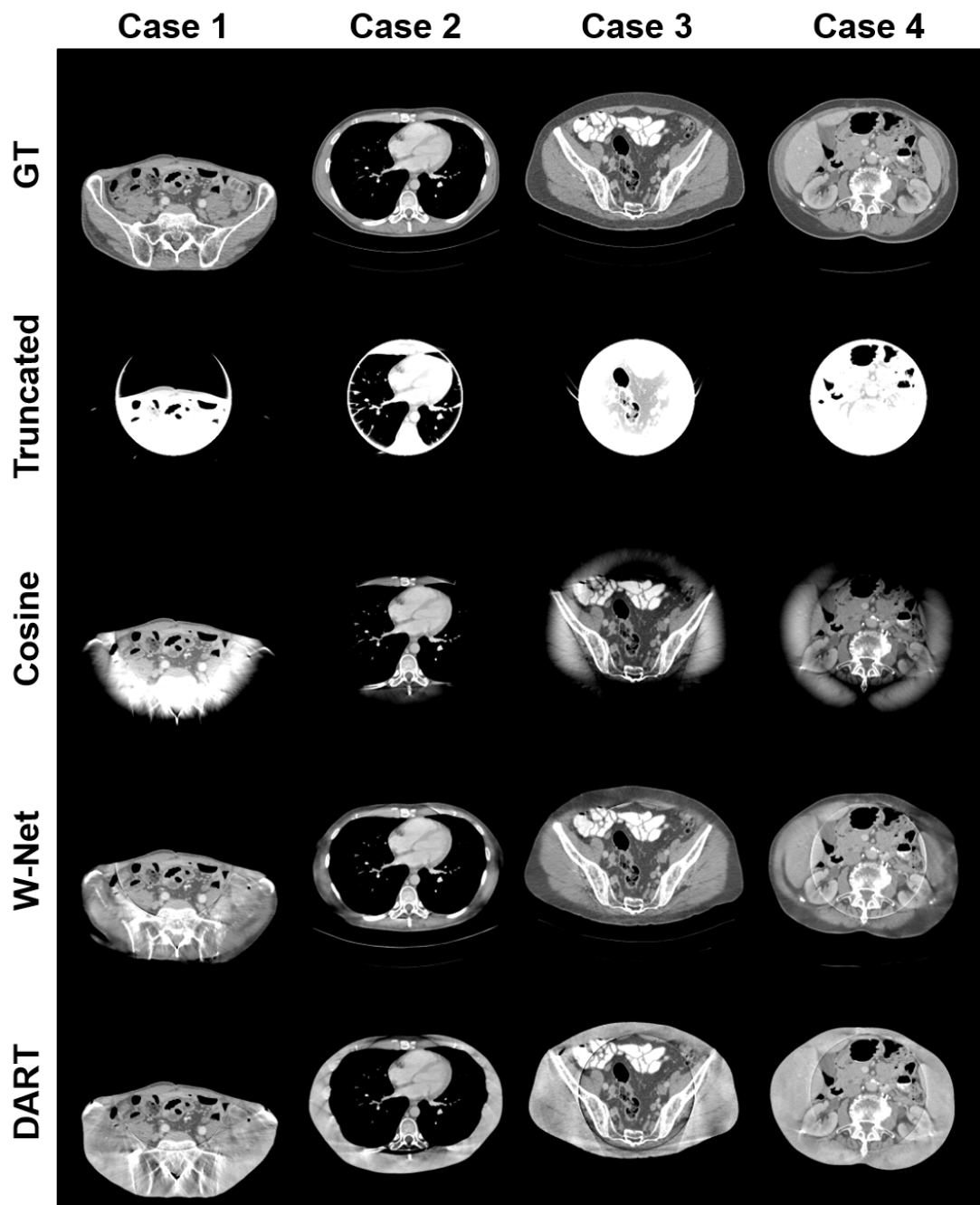


Figure 4.37: Detruncation result images with 50% truncation. $C = 0$ HU, $W = 500$ HU.

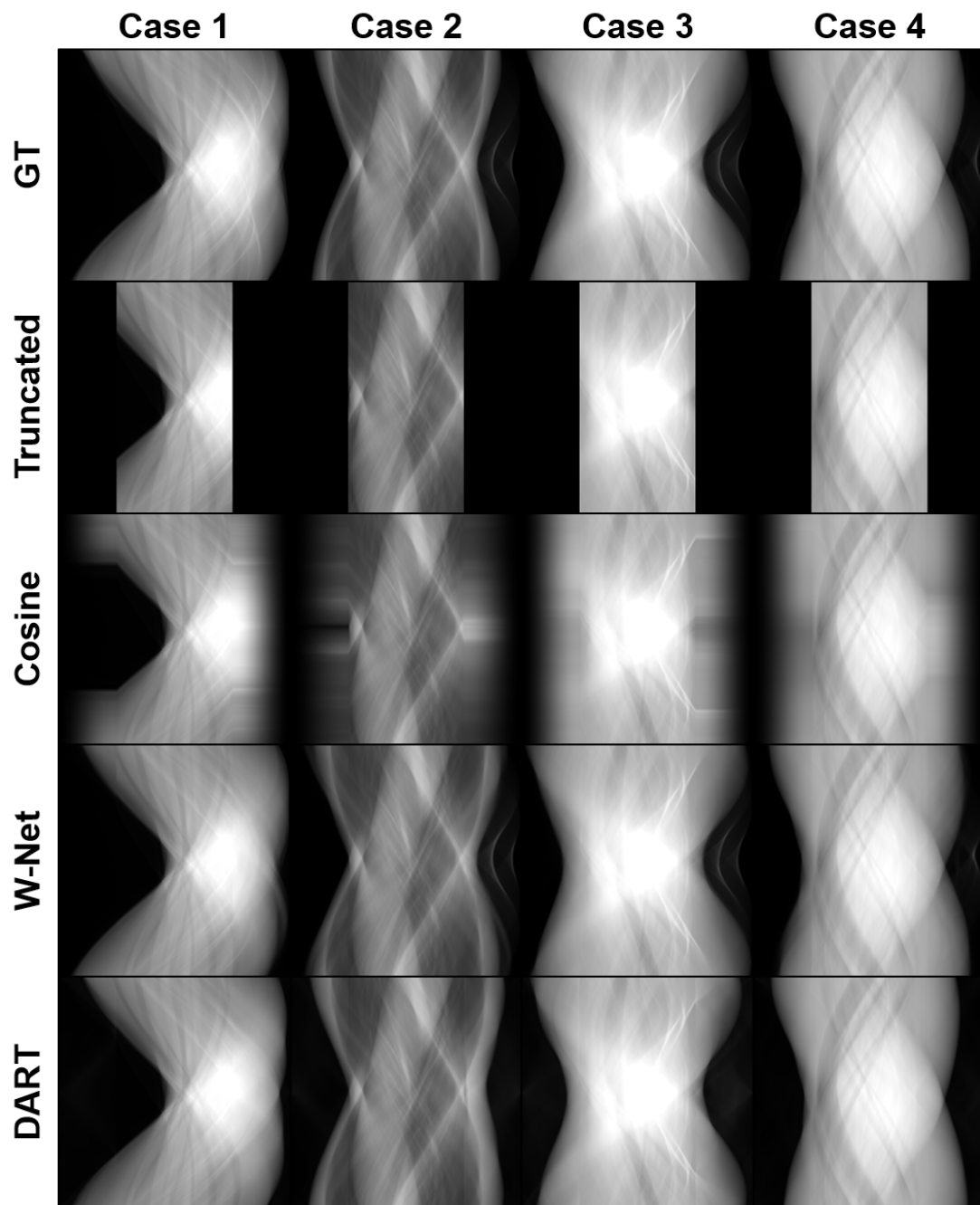


Figure 4.38: Detruncation result projections with 50% truncation. $C = 0$ HU, $W = 500$ HU.

4.5.2 25% Truncation

Figure 4.39 and 4.40 illustrate the detruncation results for a truncation of 25%. Compared to the previous experiment, the truncation artifacts are less significant. Again, all three algorithms are able to correct the CT values in the FOV. Only the W-Net and DART restore information in the eFOV. However, DART again suffers from artifacts at the edge of the original FOV, and cannot adequately reconstruct fat tissue in the extrapolated regions.

Table 4.19 gives the corresponding quantitative results. Inside the FOV, the DART prior image produces the worst RMSE results among the detruncation algorithms. This is partially due to the high noise in the DART reconstruction. Otherwise, the DART and W-Net results are superior to the conventional cosine detruncation. Overall, the W-Net yields the best metrics.

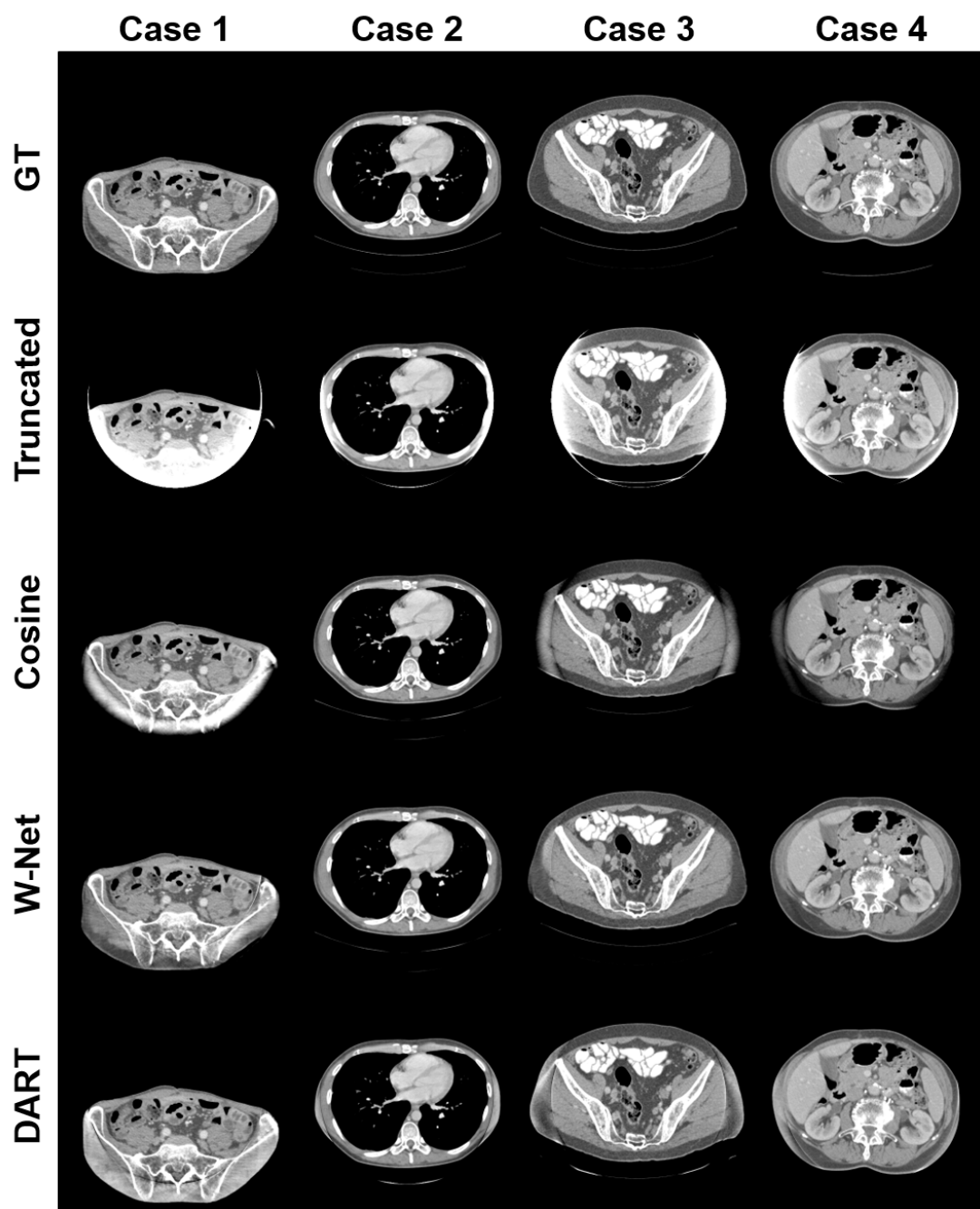


Figure 4.39: Detruncation result images with 25% truncation. $C = 0$ HU, $W = 500$ HU.

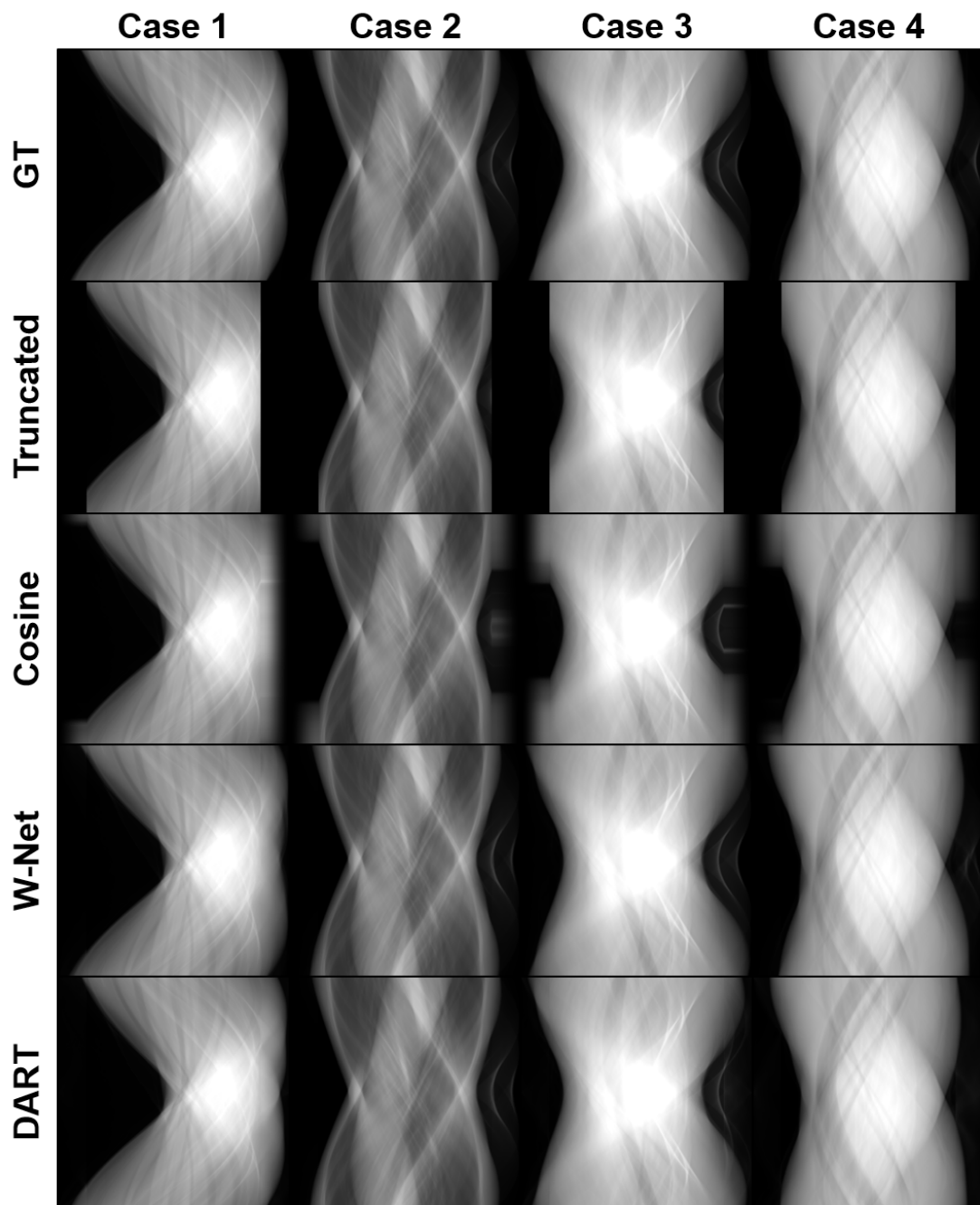


Figure 4.40: Detruncation result projections with 25% truncation. $C = 3.5$, $W = 7$.

Metric	Truncated	Cosine	W-Net _{Prior}	W-Net	DART _{Prior}	DART
RMSE _{FOV}	566.596	19.429	9.302	15.264	88.907	20.481
	176.168	42.917	6.054	15.368	111.269	35.498
	351.551	30.964	3.241	11.000	147.653	21.641
	252.080	43.372	4.172	8.824	88.648	17.541
RMSE _{eFOV}	591.62	136.113	73.581	71.280	116.073	88.590
	207.101	136.201	58.209	57.412	113.985	80.717
	378.072	118.302	28.379	28.572	153.573	101.884
	281.715	138.703	52.563	52.317	121.000	100.667
SSIM	0.817	0.9145	0.9477	0.9432	0.8972	0.9420
	0.9289	0.9406	0.9705	0.9705	0.9322	0.9574
	0.8475	0.9270	0.9830	0.9770	0.8188	0.9340
	0.9089	0.9416	0.9862	0.9829	0.8867	0.9542
Dice	0.7892	0.9635	0.9855	0.9861	0.9874	0.9878
	0.9358	0.9564	0.9819	0.9808	0.9768	0.9789
	0.9067	0.9831	0.9967	0.9966	0.9856	0.9863
	0.9480	0.9807	0.9958	0.9957	0.9861	0.9867

Table 4.19: Quantitative results for 25% truncation with RMSE in HU for four test cases.

5 | Summary & Discussion

This thesis investigated two applications of deep learning-based image processing: the removal of noise and streak artifacts in low-dose CT, and the sinogram completion of truncated CT scans. The latter was also accomplished with an iterative approach. This chapter will first discuss the low-dose CT results, then the detruncation.

5.1 Low-Dose CT Denoising

Dose in CT can almost arbitrarily be saved by either reducing the number of projections or the tube current. However, this kind of dose reduction entails a deterioration of image quality. In order to remove the inevitable noise and sparseness artifacts in low-dose CT images, a W-Net was implemented. The network employs two U-Nets, which correct the sinogram and image, respectively. To enable an end-to-end training of the network, the two U-Nets were connected with a differentiable FBP layer.

The W-Net was trained for five different dose implementations, ranging from low-mAs CT with conventional number of projections, to sparse-view CT with full tube current per projection. All scans were simulated with a 80% dose reduction compared to the high dose image. As shown in section 4.1, the W-Net was able to significantly improve image quality for all tested scan configurations. This suggests that deep learning-based methods may enable significant dose reductions in CT. Qualitatively, some discrepancies remained when comparing the images to the ground truth.

There are two main interpretations for these inaccuracies: Firstly, the network is unable to sufficiently discriminate between content and noise. Secondly, some image details are completely covered by noise or streaks in the low-dose simulation process.

In the first interpretation, the network cannot fully separate image content from noise and artifacts, although enough information is available. The bottleneck is consequently either the network architecture or the training process. U-Nets, and correspondingly also W-Nets, have been shown to perform well on a variety of image processing tasks [60], [132], [152]. While the network architecture may thus be not fully optimal, it should be sufficient, as indicated by the satisfactory denoising results. Section 4.3.1 further investigated the network components of the W-Net, illustrating the effectiveness of combined sinogram and image denoising. Sinogram-only denoising produced worse results and did not accurately restore image details. Ideally, a network architecture search should be performed to further optimize the parameters of the two U-Nets.

In the training process, the primary design choices are the training data and the loss function. In this study, the test cases are generated with the same process as the training data, such that there is no distribution mismatch between training and testing stage. As the Adam optimizer is state-of-the-art [149], the loss function becomes the most influential remaining factor in the network performance. This work employs a combination of MSE and perceptual loss, implemented by comparing intermediate feature layers of a pre-trained VGG-19 network [153]. The perceptual loss ensures that the final image has a similar noise texture and resolution as the ground truth data [77]. Notably, the VGG-19 itself is trained on photographs, not on CT images. Therefore, the rigor of this type of perceptual loss function is debatable in the context of CT imaging. In addition, a suboptimal choice of hyperparameters in the loss function may lead to texture artifacts in the CNN output. This includes choosing VGG-19 layers that correspond to content instead of texture, and setting the MSE contribution too low in the total loss. In contrast, using MSE loss without perceptual loss lead to a further reduction in image noise, but at the cost of overly smoothing the images. This resulted in lower task-based metrics, as presented in section 4.3.2. It is therefore evident, that standalone MSE is not optimal for the denoising problem as structures are blurred. Another aspect is the fact that the ground truth images in the LiTS dataset still contained a small amount of noise themselves [66]. This means that a comparison to the ground truth images does not perfectly represent a comparison to the ideal patient image. Thus, a noiseless image would neither produce minimal loss nor optimal conventional image quality metrics.

If on the contrary, the image details are fully disguised by noise or artifacts, the low-dose CTs may not contain enough information for the CNN to restore the structures. If this is the case, a denoising or destreaking algorithm performs adequately when removing artifacts but not resolving the details. Otherwise, the algorithm might insert new features into the images, so-called hallucinations. Although the W-Net does use prior information by merit of the network training process, using additional patient-specific information, e.g. a previous scan, may further enable the network to infer fine structures and reduce artifacts [67]. Conversely, this would enable more drastic dose reductions. This is especially interesting for interventional applications that require multiple scans in short succession.

5.1.1 Low-Dose CT Realizations

As stated above, the W-Net successfully reduced noise and sparseness artifacts in all investigated dose configurations. At the same time, the image quality differed between the setups. Before correction, there is a clear trend of decreasing image quality as the number of projections decreases. Thus, in a conventional imaging scenario, it is decisively preferable to reduce the tube current instead. In conjunction with W-Net processing, the differences between the acquisitions become less pronounced. Generally, scans with more projections still perform better in both conventional and task-based metrics, but the trend is not as clear as before. As low-mAs CT performs best both with and without denoising, it appears beneficial to always use the maximum number of

projections. Conversely, the relatively small difference after CNN correction imply that the comparative advantage of deep learning-based image processing is in sparse-view CT, i.e. CNNs are more relevant for sparse acquisitions than for low-mAs scans. However, a more sophisticated analysis of information content is needed to properly support this claim.

While the results of this work support a preference of low-mAs over sparse-view CT from an image processing perspective, another crucial aspect is the technical feasibility. In a modern diagnostic CT system, the X-ray source and detector do not in step-and-shoot mode. Instead, the gantry is continuously rotating during the scan [1]. As a consequence, each projection is measured over a small angular range instead of a single position. Accordingly, if the number of projections is large, the angle covered by each projection is negligibly small. When the number of projections decreases, a single projection will cover a larger angle, essentially smearing information in the angular direction. This will further deteriorate image quality. In the simulations shown in this thesis, this smearing is not taken into account, exacerbating the advantage of low-mAs CT. In addition, this procedure involves raising the rotation speed of the system, which is typically not supported by commercial CT systems. One could avoid this by simply intermittently turning off the X-ray tube during the rotation. This rapid on-off switching, however, is not available for the high power X-ray tubes in current diagnostic CT. Pulsed X-ray tubes exist do exist for several interventional CBCT machines, making sparse-view acquisitions more interesting for these system [165].

C-arm CBCT systems are predominantly used for 2D imaging, such as X-rays imaging or 2D fluoroscopy [9]. Yet they are capable of 3D protocols, albeit with significantly higher acquisition times and lower image quality compared to diagnostic CT. The benefit of sparse-view CBCT would be faster scan times, as well as reduced dose. In addition, interventional applications often require multiple successive scans, which add a significant dose penalty to the use of 3D protocols. Algorithms have been proposed utilizing the inherent sparseness of the difference between follow-up scans, allowing for a drastic reduction of number of projections [99].

One particular application of sparse-view CBCT are retroactively gated cardiac acquisitions [12]. Here, the sinogram is divided according to the phase of the heart beat while the C-arm slowly rotates around the patient. As a result, each phase-gated sinogram only contains a subset of the full sinogram, with a correspondingly lower number of projections. This imaging scenario is distinctly removed from the low-dose CT discussed in this work, as there is no trade-off between projections and tube current in cardiac CT. Furthermore, the projections are not equally distributed over all angles, such that the reconstruction must also consider limited angle artifacts.

While low-mAs CT should be preferred for most low-dose acquisitions, there is a limit to how low the tube current can be set. If the X-ray flux drops below a certain threshold, the signal will be indistinguishable from the electronic noise of the detector [105]. Then, image reconstruction will be impossible. Photon starvation typically occurs in the context of metal artifacts, where the affected projections will be replaced via interpolation [1]. Of course, this is not feasible for low-dose CT, where the whole sinogram would have to be replaced. Under these circumstances, and at equal dose, the

only viable solution is decreasing the number of projections and increasing tube current. This study has not considered electronic noise in the noise simulation, which is one of its limitations. At the dose level simulated, the photon noise should still be dominant, such that the exclusion of electronic noise is not significant. Electronic noise can be avoided in patient scans by using novel photon counting detectors, which do not suffer from this type of noise [28].

5.1.2 Image Metrics

Quantitative analysis is an integral component of medical image processing. Many established and widely-used metrics, however, do not properly relay the diagnostic value of the images. This necessitates more refined pathways of quantifying image quality, for instance via task-based metrics. In this study, three kinds of metrics have been used: conventional (RMSE, SSIM), segmentation-based (dice, RVD, ASD, RMSD, MSD, recall, precision), and model observer-based (AUC). In addition, perceptual loss and MSE were used in network training.

In terms of ease of use, the conventional metrics are far superior to the task-based ones. They are fast to calculate and only require the ground truth images as reference. If no ground truth was available, one could place a region of interest (ROI) and measure the image noise as the standard deviation inside the ROI. The low computation times and simplicity make them a natural choice as loss function for network training. Despite their advantages, in this study, the conventional metrics did not perfectly align with the human perception of image quality.

The shortcomings of RMSE, MSE, and SSIM are evident in several parts of this thesis. MSE as a loss function was deemed inferior to MSE with perceptual loss in section 4.3.2. Although the latter could be considered a conventional metric, as it is computed as the MSE of two feature layers, it requires a sophisticated, pre-trained CNN. Moreover, the network trained only with MSE performed better on traditional metrics, while yielding worse images qualitatively and measured on task-based metrics. This highlights the discrepancy between image quality and performance on conventional image quality metrics. The same could be observed for the insufficiently trained networks in section 4.4.

In contrast to conventional metrics, task-based image quality analysis requires a significant effort. For the segmentation-based metrics, each volume had to be segmented into liver and liver lesions, followed by an elaborate algorithm to determine the final metrics. To measure detectability, a separate lesion dataset was simulated, and evaluated with a model observer. Despite the computation cost associated with these relatively complicated metrics, the results are decidedly more meaningful and interpretable. Compared to the conventional and segmentation-based metrics, AUC was an effective tool for identifying insufficiently trained networks and comparing loss functions, as shown in section 4.4 and 4.3.2, respectively. Two lesion sizes were investigated in the context of the MO study. The CNN denoising was generally more effective for the larger lesion with 4 mm radius and 20 HU contrast. In some cases for the small lesion with 2 mm radius and 40 HU contrast, AUC was measured higher without denoising. This indicates

that the W-Net is better at preserving larger structures, even if they have lower contrast. Besides the SKE-BKS study setup used for the majority of this thesis, 4.2.2 also shows SKE-BKE results. Since the ability of the CHO to distinguish between signal and no-signal images is almost perfect on BKE datasets, internal noise had to be added to give meaningful values. Here, the W-Net did not improve image quality. This suggests that the network output is inconsistent when the same image with different noise is used as input.

In this work, segmentation-based metrics were used to evaluate denoising performance. In principle, lowering noise without affecting image content should yield better a better segmentation. Likewise, if the noise is removed together with a loss of details or resolution, the segmentation should suffer. Overall, the segmentation metrics did not show the desired level of accuracy, as presented in section 4.2.1. Consequently, the differences of segmentation metrics visible in the other parts of this thesis are within the margin of error of the metrics themselves. Since the H-DenseUNet [64] was trained on the ground truth data, a distribution mismatch might be the cause for the large deviations. This could then require a complete retraining of the H-DenseUNet for tested dose configuration and denoising network. Another approach that was tested was to compare the H-DenseUNet predictions not to the ground truth segmentations, but to the H-DenseUNet predictions on the ground truth CT volumes. Then, if the distribution of the denoised volumes is close to the distribution of ground truth volumes, the H-DenseUNet predictions should be similar. This, in effect, is somewhat comparable to the principles of perceptual loss. The resulting segmentation metrics were generally higher than when comparing with the ground truth, but otherwise no significant difference was apparent. In essence, the tested CT datasets feature comparable image quality within the accuracy of the segmentation-based analysis. One important factor is that the H-DenseUNet performs several pre- and post-processing steps that may cause the network output to be highly similar for the tested images.

5.2 Detruncation

The second part of this thesis considered the CT reconstruction from truncated projections, which is a frequent problem especially in interventional CBCT. Although basic extrapolation methods can alleviate the truncation artifacts in the FOV, the remaining information outside the FOV is lost. While this information should not be used for clinical diagnosis, it can be utilized for secondary algorithms such as beam-hardening correction. In this work, two algorithms were implemented to restore image details in both the FOV and the eFOV: the iterative DART detruncation [164], and the deep learning-based W-Net.

In the case of 25% truncation, where only a small fraction of the patient is outside the FOV, both algorithms provided overall satisfactory results. For the case of 50% truncation, the accuracy in the eFOV decreases for both DART and the W-Net, as expected. Here, each algorithm suffers from different issues, while the W-Net generally outperforms the DART detruncation. The DART image does not correctly estimate different tissue types, as it uses a binarization step to enable the reconstruction. This

leads to an overestimation of patient size in regions with bone, and an underestimation in regions with fat tissue, and produces the wrong CT values. A possible approach to remedy this, is to increase the number of thresholds in the discretization step [114]. One would, however, have to consider a potential increase in the time until convergence as the system is now more complex. In contrast, the W-Net is able to differentiate between different tissues but does not consistently produce realistic patient outlines. Although the perceptual loss performed well for the denoising task, it may be preferable to add a loss that penalizes deviations in patient outline, such as dice score.

Another benefit of the W-Net is its low inference time compared to DART. The latter requires a large number of iterations, each including multiple forwards-projection (FP) and back-projection (BP) steps, until convergence. Meanwhile, W-Net only uses a single FBP. There are several ways of decreasing the computational cost of DART. For instance, the problem can be downsampled for the early iterations of the algorithm [122]. In this work, the smoothing filter was stronger in early iterations to accelerate convergence. A reduction of matrix size may yield a similar effect, besides reducing the computational burden of the individual FPs and BPs. Later iterations should be performed on the original pixel size, to ensure high resolution results. Another way is to combine DART with other methods. A deep learning-guided update step could speed up convergence immensely. If an analytic method is desired, region growing methods could steer the early iterations to more quickly expand the patient outline. A third option could be to include other prior information. For example, a camera image from above the patient could provide the patient outline from above. Then, the algorithm only needs to find the part of the outline not visible on the camera and fill in the tissues. Several other approaches have previously been discussed in the literature [113], [115], [116], [121].

The detruncation analysis in this work has four major limitations: First, due to the computational cost of DART, only individual slices were compared instead of the full test set. Thus, section 4.5 functions more as a proof-of-principle case study, than an extensive comparison. Second, only basic DART [118] was implemented while there are more sophisticated variations that may produce better detruncation results. Third, modern scanners do not use parallel beam geometry but fan-beam or cone-beam geometries. The general results of this work should be transferable to these geometries, e.g. by rebinning, but other geometries might influence the performance of either DART or the W-Net. Fourth, the detruncation results were only analyzed with conventional image metrics. If the results are to be used in a secondary algorithm, e.g. scatter estimation, the detruncation algorithms should be evaluated on metrics related to said application [106].

6 | Conclusions

Dose reduction is one of the primary concerns of CT research. Several pathways of arbitrarily reducing dose exist, including decreasing the tube current and lowering the number of projections. However, these methods inadvertently cause a deterioration of image quality. This thesis has shown that data-driven CNNs can drastically improve image quality of low-dose CT. They can thus enable dose reduction without sacrificing diagnostic value of the images. This in turn will permit more frequent usage of 3D image guidance during surgeries, providing the surgeon with spatial information. In order to ensure consistent network outputs, task-based metrics were used to differentiate between well- and insufficiently trained networks. Here, detectability as measured by a model observer study allowed for a quantitative analysis close to human perception and identified inferior networks. This paves the way for a safe utilization of CNN-based methods in the clinical workflow, where accurate and reliable algorithms are paramount. In addition, low-mAs protocols were deemed preferable to sparse-view scans, although the CNN was able to decrease the differences. Finally, truncated CT scans were successfully reconstructed both iteratively with a DART-based method, as well as in data-driven manner with a dual domain CNN. Especially the latter has shown the potential of improving the image quality of acquisitions with small detectors. The additional information outside of the FOV may help guide the surgeon during an operation by placing the FOV in the context of the remaining anatomy. In addition, the larger FOV increases the performance of secondary algorithms such as beam-hardening or attenuation correction, further improving treatment accuracy.

Bibliography

- [1] W. Schlegel, C. P. Karger, and O. Jäkel, Eds., *Medizinische Physik: Grundlagen – Bildgebung – Therapie – Technik*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2018.
- [2] T. M. Buzug, *Computed Tomography*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008.
- [3] G. N. Hounsfield, “Computerized transverse axial scanning (tomography): Part 1. Description of system,” *The British Journal of Radiology*, volume 46, number 552, pages 1016–1022, 1973.
- [4] M. M. Ter-Pogossian, “Basic principles of computed axial tomography,” *Seminars in Nuclear Medicine*, volume 7, number 2, pages 109–127, 1977.
- [5] A. A. Ginde, A. Foianini, D. M. Renner, M. Valley, and C. A. Camargo Jr, “Availability and Quality of Computed Tomography and Magnetic Resonance Imaging Equipment in U.S. Emergency Departments,” *Academic Emergency Medicine*, volume 15, number 8, pages 780–783, 2008.
- [6] V. Tacher and H. Kobeiter, “State of the Art of Image Guidance in Interventional Radiology,” *Journal of the Belgian Society of Radiology*, volume 102, number S1, page 7, 2018.
- [7] J. H. Siewerdsen, M. J. Daly, G. Bachar, D. J. Moseley, G. Bootsma, K. K. Brock, S. Ansell, G. A. Wilson, S. Chhabra, D. A. Jaffray, and J. C. Irish, “Multimode C-arm fluoroscopy, tomosynthesis, and cone-beam CT for image-guided interventions: From proof of principle to patient protocols,” presented at the Medical Imaging, J. Hsieh and M. J. Flynn, Eds., San Diego, CA, 2007, 65101A.
- [8] S. Raj, F. G. Irani, K. H. Tay, and B. S. Tan, “C-arm Cone Beam Computed Tomography: A New Tool in the Interventional Suite,” *Annals of the Academy of Medicine, Singapore*, volume 42, number 11, pages 585–592, 2013.
- [9] R. C. Orth, M. J. Wallace, and M. D. Kuo, “C-arm Cone-beam CT: General Principles and Technical Considerations for Use in Interventional Radiology,” *Journal of Vascular and Interventional Radiology*, volume 20, number 7, S538–S544, 2009.

- [10] L. De Beuckeleer, “Cone Beam CT-arthrography of the wrist: High resolution images at low radiation dose,” in collaboration with K. Carpentier, B. De Foer, and M. Pouillon, 857 words, 2014.
- [11] G. Böning, W. M. Lüdemann, J. Chapiro, M. Jonczyk, B. Hamm, R. W. Günther, B. Gebauer, and F. Streitparth, “Clinical Experience with Real-Time 3-D Guidance Based on C-Arm-Acquired Cone-Beam CT (CBCT) in Transjugular Intrahepatic Portosystemic Stent Shunt (TIPSS) Placement,” *CardioVascular and Interventional Radiology*, volume 41, number 7, pages 1035–1042, 2018.
- [12] J. Rieber, C. Rohkohl, G. Lauritsch, H. Rittger, and O. Meissner, “Kardiale Anwendung der C-Arm-Computertomographie,” *Der Radiologe*, volume 49, number 9, pages 862–867, 2009.
- [13] S. Kakeda, Y. Korogi, N. Ohnari, J. Moriya, N. Oda, K. Nishino, and W. Miyamoto, “Usefulness of Cone-Beam Volume CT with Flat Panel Detectors in Conjunction with Catheter Angiography for Transcatheter Arterial Embolization,” *Journal of Vascular and Interventional Radiology*, volume 18, number 12, pages 1508–1516, 2007.
- [14] S. Schafer, S. Nithiananthan, D. J. Mirotta, A. Uneri, J. W. Stayman, W. Zbijewski, C. Schmidgunst, G. Kleinszig, A. J. Khanna, and J. H. Siewerdsen, “Mobile C-arm cone-beam CT for guidance of spine surgery: Image quality, radiation dose, and integration with interventional guidance,” *Medical Physics*, volume 38, number 8, pages 4563–4574, 2011.
- [15] Y. Otake, J. W. Stayman, W. Zbijewski, R. J. Murphy, M. D. Kutzer, R. H. Taylor, J. H. Siewerdsen, and M. Armand, “Model-based cone-beam CT reconstruction for image-guided minimally invasive treatment of hip osteolysis,” presented at the SPIE Medical Imaging, D. R. Holmes and Z. R. Yaniv, Eds., Lake Buena Vista (Orlando Area), Florida, USA, 2013, 86710Y.
- [16] V. Tacher, A. Radaelli, M. Lin, and J.-F. Geschwind, “How I Do It: Cone-Beam CT during Transarterial Chemoembolization for Liver Cancer,” *Radiology*, volume 274, number 2, pages 320–334, 2015.
- [17] S. B. Solomon and S. G. Silverman, “Imaging in Interventional Oncology,” *Radiology*, volume 257, number 3, pages 624–640, 2010.
- [18] G. Landry and C. Hua, “Current state and future applications of radiological image guidance for particle therapy,” *Medical Physics*, volume 45, number 11, 2018.
- [19] J.-J. Yu, G.-T. Kim, Y.-S. Choi, E.-H. Hwang, J. Paek, S.-H. Kim, and J. C. Huang, “Accuracy of a cone beam computed tomography-guided surgical stent for orthodontic mini-implant placement,” *The Angle Orthodontist*, volume 82, number 2, pages 275–283, 2012.

-
- [20] C. Holberg, S. Steinhäuser, P. Geis, and I. Rudzki-Janson, “Cone-Beam Computed Tomography in Orthodontics: Benefits and Limitations,” *Journal of Orofacial Orthopedics / Fortschritte der Kieferorthopädie*, volume 66, number 6, pages 434–444, 2005.
- [21] T. M. Peters, “Image-guidance for surgical procedures,” *Physics in Medicine and Biology*, volume 51, number 14, R505–R540, 2006.
- [22] R. Gupta, C. Walsh, I. S. Wang, M. Kachelrieß, J. Kuntz, and S. Bartling, “CT-Guided Interventions: Current Practice and Future Directions,” in *Intraoperative Imaging and Image-Guided Therapy*, F. A. Jolesz, Ed., New York, NY: Springer New York, 2014, pages 173–191.
- [23] K. Rajendran, B. A. Voss, W. Zhou, S. Tao, D. R. DeLone, J. I. Lane, J. M. Weaver, M. L. Carlson, J. G. Fletcher, C. H. McCollough, and S. Leng, “Dose Reduction for Sinus and Temporal Bone Imaging Using Photon-Counting Detector CT With an Additional Tin Filter,” *Investigative Radiology*, volume 55, number 2, pages 91–100, 2020.
- [24] C. H. McCollough, M. R. Bruesewitz, and J. M. Kofler, “CT Dose Reduction and Dose Management Tools: Overview of Available Options,” *RadioGraphics*, volume 26, number 2, pages 503–512, 2006.
- [25] J. Steidel, J. Maier, S. Sawall, and M. Kachelrieß, “Dose reduction potential in diagnostic single energy CT through patient-specific prefilters and a wider range of tube voltages,” *Medical Physics*, volume 49, number 1, pages 93–106, 2022.
- [26] X. Wang, A. Zamyatin, and D. Shi, “Dose reduction potential with photon counting computed tomography,” presented at the SPIE Medical Imaging, N. J. Pelc, R. M. Nishikawa, and B. R. Whiting, Eds., San Diego, California, USA, 2012, page 831 349.
- [27] K. Rajendran, M. Petersilka, A. Henning, E. R. Shanblatt, B. Schmidt, T. G. Flohr, A. Ferrero, F. Baffour, F. E. Diehn, L. Yu, P. Rajiah, J. G. Fletcher, S. Leng, and C. H. McCollough, “First Clinical Photon-counting Detector CT System: Technical Evaluation,” *Radiology*, volume 303, number 1, pages 130–138, 2022.
- [28] M. J. Willemink, M. Persson, A. Pourmorteza, N. J. Pelc, and D. Fleischmann, “Photon-counting CT: Technical Principles and Clinical Prospects,” *Radiology*, volume 289, number 2, pages 293–312, 2018.
- [29] M. M. Lell and M. Kachelrieß, “Recent and Upcoming Technological Developments in Computed Tomography: High Speed, Low Dose, Deep Learning, Multienergy,” *Investigative Radiology*, volume 55, number 1, pages 8–19, 2020.
- [30] H. Huflage, R. Hendel, A. S. Kunz, S. Ergün, S. Afat, N. Petri, V. Hartung, P. Gruschwitz, T. A. Bley, and J.-P. Grunz, “Investigating the Small Pixel Effect in Ultra-High Resolution Photon-Counting CT of the Lung,” *Investigative Radiology*, volume 59, number 4, pages 293–297, 2024.

- [31] J. N. Althén, “Automatic tube-current modulation in CT—a comparison between different solutions,” *Radiation Protection Dosimetry*, volume 114, number 1-3, pages 308–312, 2005.
- [32] S. Rizzo, M. Kalra, B. Schmidt, T. Dalal, C. Suess, T. Flohr, M. Blake, and S. Saini, “Comparison of Angular and Combined Automatic Tube Current Modulation Techniques with Constant Tube Current CT of the Abdomen and Pelvis,” *American Journal of Roentgenology*, volume 186, number 3, pages 673–679, 2006.
- [33] M. Gies, W. A. Kalender, H. Wolf, C. Suess, and M. T. Madsen, “Dose reduction in CT by anatomically adapted tube current modulation. I. Simulation studies,” *Medical Physics*, volume 26, number 11, pages 2235–2247, 1999.
- [34] M. K. Kalra, M. M. Maher, T. L. Toth, B. Schmidt, B. L. Westerman, H. T. Morgan, and S. Saini, “Techniques and Applications of Automatic Tube Current Modulation for CT,” *Radiology*, volume 233, number 3, pages 649–657, 2004.
- [35] X. Li, W. P. Segars, and E. Samei, “The impact on CT dose of the variability in tube current modulation technology: A theoretical investigation,” *Physics in Medicine and Biology*, volume 59, number 16, pages 4525–4548, 2014.
- [36] 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,” *Medical Physics*, volume 49, number 7, pages 4391–4403, 2022.
- [37] M. Cesarelli, P. Bifulco, T. Cerciello, M. Romano, and L. Paura, “X-ray fluoroscopy noise modeling for filter design,” *International Journal of Computer Assisted Radiology and Surgery*, volume 8, number 2, pages 269–278, 2013.
- [38] A. Maier, L. Wigström, H. G. Hofmann, J. Hornegger, L. Zhu, N. Strobel, and R. Fahrig, “Three-dimensional anisotropic adaptive filtering of projection data for noise reduction in cone beam CT,” *Medical Physics*, volume 38, number 11, pages 5896–5909, 2011.
- [39] M. Kachelrieß, O. Watzke, and W. A. Kalender, “Generalized multi-dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam CT,” *Medical Physics*, volume 28, number 4, pages 475–490, 2001.
- [40] Y. Zhang, J. Zhang, and H. Lu, “Statistical Sinogram Smoothing for Low-Dose CT With Segmentation-Based Adaptive Filtering,” *IEEE Transactions on Nuclear Science*, volume 57, number 5, pages 2587–2598, 2010.
- [41] S. V. Mohd Sagheer and S. N. George, “A review on medical image denoising algorithms,” *Biomedical Signal Processing and Control*, volume 61, page 102 036, 2020.
- [42] J. He, Y. Yang, Y. Wang, D. Zeng, Z. Bian, H. Zhang, J. Sun, Z. Xu, and J. Ma, “Optimizing a Parameterized Plug-and-Play ADMM for Iterative Low-Dose CT Reconstruction,” *IEEE Transactions on Medical Imaging*, volume 38, number 2, pages 371–382, 2019.

-
- [43] L. L. Geyer, U. J. Schoepf, F. G. Meinel, J. W. Nance, G. Bastarrika, J. A. Leipsic, N. S. Paul, M. Rengo, A. Laghi, and C. N. De Cecco, “State of the Art: Iterative CT Reconstruction Techniques,” *Radiology*, volume 276, number 2, pages 339–357, 2015.
- [44] P. B. Noël, S. Engels, T. Köhler, D. Muenzel, D. Franz, M. Rasper, E. J. Rummeny, M. Dobritz, and A. A. Fingerle, “Evaluation of an iterative model-based CT reconstruction algorithm by intra-patient comparison of standard and ultra-low-dose examinations,” *Acta Radiologica*, volume 59, number 10, pages 1225–1231, 2018.
- [45] M. Beister, D. Kolditz, and W. A. Kalender, “Iterative reconstruction methods in X-ray CT,” *Physica Medica*, volume 28, number 2, pages 94–108, 2012.
- [46] H. Zhang, J. Wang, D. Zeng, X. Tao, and J. Ma, “Regularization strategies in statistical image reconstruction of low-dose x-ray CT : A review,” *Medical Physics*, volume 45, number 10, 2018.
- [47] E. Y. Sidky, C.-M. Kao, and X. Pan. “Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT.” (2009), [Online]. Available: <http://arxiv.org/abs/0904.4495> (visited on 03/26/2024), preprint.
- [48] S. Niu, Y. Gao, Z. Bian, J. Huang, W. Chen, G. Yu, Z. Liang, and J. Ma, “Sparse-view x-ray CT reconstruction via total generalized variation regularization,” *Physics in Medicine and Biology*, volume 59, number 12, pages 2997–3017, 2014.
- [49] J. Bian, J. H. Siewerdsen, X. Han, E. Y. Sidky, J. L. Prince, C. A. Pelizzari, and X. Pan, “Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT,” *Physics in Medicine and Biology*, volume 55, number 22, pages 6575–6599, 2010.
- [50] J. Huang, Y. Zhang, J. Ma, D. Zeng, Z. Bian, S. Niu, Q. Feng, Z. Liang, and W. Chen, “Iterative Image Reconstruction for Sparse-View CT Using Normal-Dose Image Induced Total Variation Prior,” *PLoS ONE*, volume 8, number 11, G. Wang, Ed., e79709, 2013.
- [51] S. Hashemi, S. Beheshti, P. R. Gill, N. S. Paul, and R. S. C. Cobbold, “Accelerated Compressed Sensing Based CT Image Reconstruction,” *Computational and Mathematical Methods in Medicine*, volume 2015, pages 1–16, 2015.
- [52] G.-H. Chen, J. Tang, and S. Leng, “Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets: Prior image constrained compressed sensing (PICCS),” *Medical Physics*, volume 35, number 2, pages 660–663, 2008.
- [53] H. Kudo, T. Suzuki, and E. A. Rashed, “Image reconstruction for sparse-view CT and interior CT— introduction to compressed sensing and differentiated back-projection,” *Quantitative Imaging in Medicine and Surgery*, volume 3, number 3, 2013.

- [54] Z. Zhu, K. Wahid, P. Babyn, D. Cooper, I. Pratt, and Y. Carter, “Improved Compressed Sensing-Based Algorithm for Sparse-View CT Image Reconstruction,” *Computational and Mathematical Methods in Medicine*, volume 2013, pages 1–15, 2013.
- [55] S. Kazuo, K. Kawamata, and H. Kudo, “Combining compressed sensing and deep learning using multi-channel CNN for image reconstruction in low-dose and sparse-view CT,” in *International Forum on Medical Imaging in Asia 2021*, R.-F. Chang, Ed., Taipei, Taiwan: SPIE, 2021, page 23.
- [56] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, volume 521, number 7553, pages 436–444, 2015.
- [57] A. Maier, C. Syben, T. Lasser, and C. Riess, “A gentle introduction to deep learning in medical image processing,” *Zeitschrift für Medizinische Physik*, volume 29, number 2, pages 86–101, 2019.
- [58] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, “Deep Convolutional Neural Network for Inverse Problems in Imaging,” *IEEE Transactions on Image Processing*, volume 26, number 9, pages 4509–4522, 2017.
- [59] M. I. Razzak, S. Naz, and A. Zaib, “Deep Learning for Medical Image Processing: Overview, Challenges and Future,” 2017.
- [60] P. P. Shinde and S. Shah, “A Review of Machine Learning and Deep Learning Applications,” in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCCUBEA)*, Pune, India: IEEE, 2018, pages 1–6.
- [61] J. Zhang, Y. Xie, Q. Wu, and Y. Xia, “Medical image classification using synergic deep learning,” *Medical Image Analysis*, volume 54, pages 10–19, 2019.
- [62] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng, and M. Chen, “Medical image classification with convolutional neural network,” in *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*, Singapore: IEEE, 2014, pages 844–848.
- [63] N. Mahmoodian, H. Thadesar, M. Sadeghi, M. Georgiades, M. Pech, and C. Hoeschen, “Segmentation of Living and ablated Tumor parts in CT images Using ResLU-Net,” *Current Directions in Biomedical Engineering*, volume 8, number 2, pages 49–52, 2022.
- [64] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P. A. Heng. “H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation from CT Volumes.” (2018), [Online]. Available: <http://arxiv.org/abs/1709.07330> (visited on 03/24/2024), preprint.
- [65] K. Roth, T. Konopczyński, and J. Hesser. “Liver Lesion Segmentation with slice-wise 2D Tiramisu and Tversky loss function.” (2019), [Online]. Available: <http://arxiv.org/abs/1905.03639> (visited on 03/24/2024), preprint.

-
- [66] P. Bilic, P. Christ, H. B. Li, *et al.*, “The Liver Tumor Segmentation Benchmark (LiTS),” *Medical Image Analysis*, volume 84, page 102680, 2023.
- [67] C. Wu, P. Zhang, Y. Xu, and J. Yao, “Combined spatial and temporal deep learning for image noise reduction of fluoroscopic x-ray sequences,” in *Medical Imaging 2020: Physics of Medical Imaging*, H. Bosmans and G.-H. Chen, Eds., Houston, United States: SPIE, 2020, page 145.
- [68] K. A. S. H. Kulathilake, N. A. Abdullah, A. Q. M. Sabri, and K. W. Lai, “A review on Deep Learning approaches for low-dose Computed Tomography restoration,” *Complex & Intelligent Systems*, volume 9, number 3, pages 2713–2745, 2023.
- [69] K. Choi, S. H. Kim, and S. Kim, “Self-supervised denoising of projection data for low-dose cone-beam CT,” *Medical Physics*, volume 50, number 10, pages 6319–6333, 2023.
- [70] A. A. Zamyatin, L. Yu, and D. Rozas, “3D residual convolutional neural network for low dose CT denoising,” in *Medical Imaging 2022: Physics of Medical Imaging*, W. Zhao and L. Yu, Eds., San Diego, United States: SPIE, 2022, page 165.
- [71] F. Wagner, M. Thies, L. Pfaff, O. Aust, S. Pechmann, D. Weidner, N. Maul, M. Rohleder, M. Gu, J. Utz, F. Denzinger, and A. Maier. “On the Benefit of Dual-domain Denoising in a Self-supervised Low-dose CT Setting.” (2022), [Online]. Available: <http://arxiv.org/abs/2211.01111> (visited on 03/24/2024), preprint.
- [72] C. Niu, M. Li, X. Guo, and G. Wang, “Self-supervised dual-domain network for low-dose CT denoising,” in *Developments in X-Ray Tomography XIV*, B. Müller and G. Wang, Eds., San Diego, United States: SPIE, 2022, page 15.
- [73] L. Chao, P. Zhang, Y. Wang, Z. Wang, W. Xu, and Q. Li, “Dual-domain attention-guided convolutional neural network for low-dose cone-beam computed tomography reconstruction,” *Knowledge-Based Systems*, volume 251, page 109295, 2022.
- [74] Z. Zhang, X. Liang, W. Zhao, and L. Xing, “Noise2Context: Context-assisted learning 3D thin-layer for low-dose CT,” *Medical Physics*, volume 48, number 10, pages 5794–5803, 2021.
- [75] M. Gholizadeh-Ansari, J. Alirezaie, and P. Babyn, “Deep Learning for Low-Dose CT Denoising Using Perceptual Loss and Edge Detection Layer,” *Journal of Digital Imaging*, volume 33, number 2, pages 504–515, 2020.
- [76] T. Humphries, S. Coulter, D. Si, M. Simms, and R. Xing, “Comparison of deep learning approaches to low dose CT using low intensity and sparse view data,” in *Medical Imaging 2019: Physics of Medical Imaging*, H. Bosmans, G.-H. Chen, and T. Gilat Schmidt, Eds., San Diego, United States: SPIE, 2019, page 156.
- [77] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang, “Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss,” *IEEE Transactions on Medical Imaging*, volume 37, number 6, pages 1348–1357, 2018.

- [78] M. P. Heinrich, M. Stille, and T. M. Buzug, “Residual U-Net Convolutional Neural Network Architecture for Low-Dose CT Denoising,” *Current Directions in Biomedical Engineering*, volume 4, number 1, pages 297–300, 2018.
- [79] H. Chen, Y. Zhang, W. Zhang, P. Liao, K. Li, J. Zhou, and G. Wang, “Low-dose CT denoising with convolutional neural network,” in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, Melbourne, Australia: IEEE, 2017, pages 143–146.
- [80] P. Barca, S. Domenichelli, R. Golfieri, L. Pierotti, L. Spagnoli, S. Tomasi, and L. Strigari, “Image quality evaluation of the Precise image CT deep learning reconstruction algorithm compared to Filtered Back-projection and iDose4: A phantom study at different dose levels,” *Physica Medica*, volume 106, page 102 517, 2023.
- [81] R. Singh, S. R. Digumarthy, V. V. Muse, A. R. Kambadakone, M. A. Blake, A. Tabari, Y. Hoi, N. Akino, E. Angel, R. Madan, and M. K. Kalra, “Image Quality and Lesion Detection on Deep Learning Reconstruction and Iterative Reconstruction of Submillisievert Chest and Abdominal CT,” *American Journal of Roentgenology*, volume 214, number 3, pages 566–573, 2020.
- [82] Y. Noda, T. Kaga, N. Kawai, T. Miyoshi, H. Kawada, F. Hyodo, A. Kambadakone, and M. Matsuo, “Low-dose whole-body CT using deep learning image reconstruction: Image quality and lesion detection,” *The British Journal of Radiology*, volume 94, number 1121, page 20 201 329, 2021.
- [83] B. Guan, C. Yang, L. Zhang, S. Niu, M. Zhang, Y. Wang, W. Wu, and Q. Liu, “Generative Modeling in Sinogram Domain for Sparse-View CT Reconstruction,” *IEEE Transactions on Radiation and Plasma Medical Sciences*, volume 8, number 2, pages 195–207, 2024.
- [84] Y. Zhang, H. Chen, W. Xia, Y. Chen, B. Liu, Y. Liu, H. Sun, and J. Zhou, “LEARN++: Recurrent Dual-Domain Reconstruction Network for Compressed Sensing CT,” *IEEE Transactions on Radiation and Plasma Medical Sciences*, volume 7, number 2, pages 132–142, 2023.
- [85] Z. Li, C. Ma, J. Chen, J. Zhang, and H. Shan, “Learning to Distill Global Representation for Sparse-View CT,” in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, France: IEEE, 2023, pages 21 139–21 150.
- [86] S. Kim, B. Kim, J. Lee, and J. Baek, “Sparsier2Sparse: Self-supervised convolutional neural network-based streak artifacts reduction in sparse-view CT images,” *Medical Physics*, volume 50, number 12, pages 7731–7747, 2023.
- [87] T. Cheslerean-Boghiu, F. C. Hofmann, M. Schulthei, F. Pfeiffer, D. Pfeiffer, and T. Lasser, “WNet: A Data-Driven Dual-Domain Denoising Model for Sparse-View Computed Tomography With a Trainable Reconstruction Layer,” *IEEE Transactions on Computational Imaging*, volume 9, pages 120–132, 2023.
- [88] S. Wang, X. Li, and P. Chen, “ADMM-SVNet: An ADMM-Based Sparse-View CT Reconstruction Network,” *Photonics*, volume 9, number 3, page 186, 2022.

-
- [89] E. Y. Sidky and X. Pan, “Report on the AAPM deep-learning sparse-view CT grand challenge,” *Medical Physics*, volume 49, number 8, pages 4935–4943, 2022.
- [90] H. Shibata, S. Hanaoka, Y. Nomura, T. Nakao, T. Takenaga, N. Hayashi, and O. Abe, “On the Simulation of Ultra-Sparse-View and Ultra-Low-Dose Computed Tomography with Maximum a Posteriori Reconstruction Using a Progressive Flow-Based Deep Generative Model,” *Tomography*, volume 8, number 5, pages 2129–2152, 2022.
- [91] Y. Guo, Y. Wang, M. Zhu, D. Zeng, Z. Bian, X. Tao, and J. Ma, “Dual domain closed-loop learning for sparse-view CT reconstruction,” in *7th International Conference on Image Formation in X-Ray Computed Tomography*, J. W. Stayman, Ed., Baltimore, United States: SPIE, 2022, page 60.
- [92] Y. Zhang, H. Chen, W. Xia, Y. Chen, B. Liu, Y. Liu, H. Sun, and J. Zhou, “Dual-domain reconstruction network for sparse-view CT,” in *Developments in X-Ray Tomography XIII*, B. Müller and G. Wang, Eds., San Diego, United States: SPIE, 2021, page 36.
- [93] C. Zhang, Y. Li, and G. Chen, “Accurate and robust sparse-view angle CT image reconstruction using deep learning and prior image constrained compressed sensing (DL-PICCS),” *Medical Physics*, volume 48, number 10, pages 5765–5781, 2021.
- [94] W. Wu, D. Hu, C. Niu, H. Yu, V. Vardhanabhuti, and G. Wang, “DRONE: Dual-Domain Residual-based Optimization Network for Sparse-View CT Reconstruction,” *IEEE Transactions on Medical Imaging*, volume 40, number 11, pages 3002–3014, 2021.
- [95] A. R. Podgorsak, M. M. Shiraz Bhurwani, and C. N. Ionita, “CT artifact correction for sparse and truncated projection data using generative adversarial networks,” *Medical Physics*, volume 48, number 2, pages 615–626, 2021.
- [96] S. Majee, T. Balke, C. Kemp, G. Buzzard, and C. Bouman, “Multi-Slice Fusion for Sparse-View and Limited-Angle 4D CT Reconstruction,” *IEEE Transactions on Computational Imaging*, volume 7, pages 448–462, 2021.
- [97] D. Hu, J. Liu, T. Lv, Q. Zhao, Y. Zhang, G. Quan, J. Feng, Y. Chen, and L. Luo, “Hybrid-Domain Neural Network Processing for Sparse-View CT Reconstruction,” *IEEE Transactions on Radiation and Plasma Medical Sciences*, volume 5, number 1, pages 88–98, 2021.
- [98] H. Zhang, B. Liu, H. Yu, and B. Dong. “MetaInv-Net: Meta Inversion Network for Sparse View CT Image Reconstruction.” (2020), [Online]. Available: <http://arxiv.org/abs/2006.00171> (visited on 03/26/2024), preprint.
- [99] C. Shieh, Y. Gonzalez, B. Li, X. Jia, S. Rit, C. Mory, M. Riblett, G. Hugo, Y. Zhang, Z. Jiang, X. Liu, L. Ren, and P. Keall, “SPARE: Sparse-view reconstruction challenge for 4D cone-beam CT from a 1-min scan,” *Medical Physics*, volume 46, number 9, pages 3799–3811, 2019.

- [100] Z. Zhang, X. Liang, X. Dong, Y. Xie, and G. Cao, "A Sparse-View CT Reconstruction Method Based on Combination of DenseNet and Deconvolution," *IEEE Transactions on Medical Imaging*, volume 37, number 6, pages 1407–1417, 2018.
- [101] D. H. Ye, G. T. Buzzard, M. Ruby, and C. A. Bouman. "Deep Back Projection for Sparse-View CT Reconstruction." (2018), [Online]. Available: <http://arxiv.org/abs/1807.02370> (visited on 03/26/2024), preprint.
- [102] S. Xie, X. Zheng, Y. Chen, L. Xie, J. Liu, Y. Zhang, J. Yan, H. Zhu, and Y. Hu, "Artifact Removal using Improved GoogLeNet for Sparse-view CT Reconstruction," *Scientific Reports*, volume 8, number 1, page 6700, 2018.
- [103] F. Thaler, K. Hammernik, C. Payer, M. Urschler, and D. Štern, "Sparse-View CT Reconstruction Using Wasserstein GANs," in *Machine Learning for Medical Image Reconstruction*, F. Knoll, A. Maier, and D. Rueckert, Eds., volume 11074, Cham: Springer International Publishing, 2018, pages 75–82.
- [104] Y. Han and J. C. Ye. "Framing U-Net via Deep Convolutional Framelets: Application to Sparse-view CT." (2018), [Online]. Available: <http://arxiv.org/abs/1708.08333> (visited on 03/26/2024), preprint.
- [105] J. Hsieh, *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*, 2nd ed. Bellingham: SPIE press, 2009.
- [106] N. Waltrich, S. Sawall, J. Maier, J. Kuntz, K. Stannigel, K. Lindenberg, and M. Kachelrieß, "Effect of truncation on the accuracy of Monte Carlo-based scatter estimation in truncated CBCT," *Medical Physics*, volume 45, number 8, pages 3574–3590, 2018.
- [107] Y. Xia, A. Maier, H. G. Hofmann, F. Dennerlein, K. Mueller, and J. Hornegger, "Reconstruction from truncated projections in cone-beam CT using an efficient 1D filtering," presented at the SPIE Medical Imaging, R. M. Nishikawa and B. R. Whiting, Eds., Lake Buena Vista (Orlando Area), Florida, USA, 2013, page 86681C.
- [108] Wang Xian-Chao, Yan Bin, Liu Hong-Kui, Li Lei, Wei Xing, Hu Guo-En, National Digital Switching System Engineering and Technological Research Center, Zhengzhou 450002, China; and Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian 116000, China, "Efficient reconstruction from truncated data in circular cone-beam CT," *Acta Physica Sinica*, volume 62, number 9, page 098 702, 2013.
- [109] J. S. Maltz, S. Bose, H. P. Shukla, and A. R. Bani-Hashemi, "CT Truncation artifact removal using water-equivalent thicknesses derived from truncated projection data," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Lyon, France: IEEE, 2007, pages 2907–2911.
- [110] K. Sourbelle, M. Kachelrieß, and W. A. Kalender, "Reconstruction from truncated projections in CT using adaptive detruncation," *European Radiology*, volume 15, number 5, pages 1008–1014, 2005.

-
- [111] J. Hsieh, E. Chao, J. Thibault, B. Grekowicz, A. Horst, S. McOlash, and T. J. Myers, “A novel reconstruction algorithm to extend the CT scan field-of-view,” *Medical Physics*, volume 31, number 9, pages 2385–2391, 2004.
- [112] R. Clackdoyle, F. Noo, Junyu Guo, and J. Roberts, “Quantitative reconstruction from truncated projections in classical tomography,” *IEEE Transactions on Nuclear Science*, volume 51, number 5, pages 2570–2578, 2004.
- [113] D. Frenkel, N. Six, J. De Beenhouwer, and J. Sijbers, “Tabu-DART: A dynamic update strategy for efficient discrete algebraic reconstruction,” *The Visual Computer*, volume 39, number 10, pages 4671–4683, 2023.
- [114] Y. He, W. Ming, R. Shen, and J. Chen, “IDART: An Improved Discrete Tomography Algorithm for Reconstructing Images With Multiple Gray Levels,” *IEEE Transactions on Image Processing*, volume 31, pages 2608–2619, 2022.
- [115] R. Pua, M. Park, S. Wi, and S. Cho, “A pseudo-discrete algebraic reconstruction technique (PDART) prior image-based suppression of high density artifacts in computed tomography,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, volume 840, pages 42–50, 2016.
- [116] X. Zhuge, W. J. Palenstijn, and K. J. Batenburg, “TVR-DART: A More Robust Algorithm for Discrete Tomography From Limited Projection Data With Automated Gray Value Estimation,” *IEEE Transactions on Image Processing*, volume 25, number 1, pages 455–468, 2016.
- [117] K. Batenburg, S. Bals, J. Sijbers, C. Kübel, P. Midgley, J. Hernandez, U. Kaiser, E. Encina, E. Coronado, and G. Van Tendeloo, “3D imaging of nanomaterials by discrete tomography,” *Ultramicroscopy*, volume 109, number 6, pages 730–740, 2009.
- [118] K. Batenburg and J. Sijbers, “Dart: A Fast Heuristic Algebraic Reconstruction Algorithm for Discrete Tomography,” in *2007 IEEE International Conference on Image Processing*, San Antonio, TX, USA: IEEE, 2007, pages IV –133–IV –136.
- [119] K. J. Batenburg and J. Sijbers, “DART: A Practical Reconstruction Algorithm for Discrete Tomography,” *IEEE Transactions on Image Processing*, volume 20, number 9, pages 2542–2553, 2011.
- [120] K. J. Batenburg, J. Sijbers, H. F. Poulsen, and E. Knudsen, “DART: A robust algorithm for fast reconstruction of three-dimensional grain maps,” *Journal of Applied Crystallography*, volume 43, number 6, pages 1464–1473, 2010.
- [121] N. Six, J. D. Beenhouwer, and J. Sijbers, “pDART: Discrete algebraic reconstruction using a polychromatic forward model,”
- [122] A. Dabravolski, K. J. Batenburg, and J. Sijbers, “A Multiresolution Approach to Discrete Tomography Using DART,” *PLoS ONE*, volume 9, number 9, K. Chen, Ed., e106090, 2014.

- [123] H. Banjak, M. Costin, C. Vienne, R. Guillaumet, and V. Kaftandjian, “Iterative CT reconstruction on limited angle trajectories applied to robotic inspection,” in *AIP Conference Proceedings*, AIP Publishing LLC, volume 1806, 2017, page 020 009.
- [124] C. Zhang and G. Chen, “Deep-Interior: A new pathway to interior tomographic image reconstruction via a weighted backprojection and deep learning,” *Medical Physics*, volume 51, number 2, pages 946–963, 2024.
- [125] G. Belotti, G. Fattori, G. Baroni, and S. Rit, “Extension of the cone-beam CT field-of-view using two complementary short scans,” *Medical Physics*, mp.16869, 2023.
- [126] B. S. Khural, M. Baer-Beck, E. Fournié, K. Stierstorfer, Y. Huang, and A. Maier, “Deep learning-based extended field of view computed tomography image reconstruction: Influence of network design on image estimation outside the scan field of view,” *Biomedical Physics & Engineering Express*, volume 8, number 2, page 025 021, 2022.
- [127] M. Rossi, G. Belotti, C. Paganelli, A. Pella, A. Barcellini, P. Cerveri, and G. Baroni, “Image-based shading correction for narrow-FOV truncated pelvic CBCT with deep convolutional neural networks and transfer learning,” *Medical Physics*, volume 48, number 11, pages 7112–7126, 2021.
- [128] J. H. Ketola, H. Heino, M. A. K. Juntunen, M. T. Nieminen, and S. I. Inkinen, “Deep learning-based sinogram extension method for interior computed tomography,” in *Medical Imaging 2021: Physics of Medical Imaging*, H. Bosmans, W. Zhao, and L. Yu, Eds., Online Only, United States: SPIE, 2021, page 123.
- [129] Y. Huang, A. Preuhs, M. Manhart, G. Lauritsch, and A. Maier, “Data Extrapolation From Learned Prior Images for Truncation Correction in Computed Tomography,” *IEEE Transactions on Medical Imaging*, volume 40, number 11, pages 3042–3053, 2021.
- [130] G. P. Fonseca, M. Baer-Beck, E. Fournie, C. Hofmann, I. Rinaldi, M. C. Ollers, W. J. Van Elmpt, and F. Verhaegen, “Evaluation of novel AI-based extended field-of-view CT reconstructions,” *Medical Physics*, volume 48, number 7, pages 3583–3594, 2021.
- [131] P. Bao, W. Xia, K. Yang, W. Chen, M. Chen, Y. Xi, S. Niu, J. Zhou, H. Zhang, H. Sun, Z. Wang, and Y. Zhang, “Convolutional Sparse Coding for Compressed Sensing CT Reconstruction,” *IEEE Transactions on Medical Imaging*, volume 38, number 11, pages 2607–2619, 2019.
- [132] M. Nielsen, “Neural Networks and Deep Learning,”
- [133] D. Gomez-Cardona, C. P. Favazza, S. Leng, B. A. Schueler, and K. A. Fetterly, “Adaptation of a channelized Hotelling observer model to accommodate anthropomorphic backgrounds and moving test objects in X-ray angiography,” *Medical Physics*, volume 50, number 11, pages 6737–6747, 2023.

-
- [134] A. Ba, C. Abbey, D. Racine, A. Viry, F. R. Verdun, S. Schmidt, and F. Bochud, “Channelized Hotelling observer correlation with human observers for low-contrast detection in liver CT images,” *Journal of Medical Imaging*, volume 6, number 02, 2019.
- [135] D. Racine, N. Ryckx, A. Ba, F. Becce, A. Viry, F. R. Verdun, and S. Schmidt, “Task-based quantification of image quality using a model observer in abdominal CT: A multicentre study,” *European Radiology*, volume 28, number 12, pages 5203–5210, 2018.
- [136] C. Ma, L. Yu, B. Chen, C. Favazza, S. Leng, and C. McCollough, “Impact of number of repeated scans on model observer performance for a low-contrast detection task in computed tomography,” *Journal of Medical Imaging*, volume 3, number 2, page 023 504, 2016.
- [137] A. Wunderlich, F. Noo, B. D. Gallas, and M. E. Heilbrun, “Exact Confidence Intervals for Channelized Hotelling Observer Performance in Image Quality Studies,” *IEEE Transactions on Medical Imaging*, volume 34, number 2, pages 453–464, 2015.
- [138] Y. Zhang, S. Leng, L. Yu, R. E. Carter, and C. H. McCollough, “Correlation between human and model observer performance for discrimination task in CT,” *Physics in Medicine and Biology*, volume 59, number 13, pages 3389–3404, 2014.
- [139] L. Zhang, C. Cavaro-Ménard, and P. Le Callet, “An overview of model observers,” *IRBM*, volume 35, number 4, pages 214–224, 2014.
- [140] L. Zhang, B. Goossens, C. Cavaro-Ménard, P. L. Callet, and D. Ge, “Channelized model observer for the detection and estimation of signals with unknown amplitude, orientation, and size,” *Journal of the Optical Society of America A*, volume 30, number 11, page 2422, 2013.
- [141] X. He and S. Park, “Model Observers in Medical Imaging Research,” *Theranostics*, volume 3, number 10, pages 774–786, 2013.
- [142] W. C. Röntgen, “Über eine neue Art von Strahlen,” *Sitzungsbericht Phys.-Med. Gesellschaft*, volume 137, pages 132–141, 1895.
- [143] J. T. Bushberg, Ed., *The Essential Physics of Medical Imaging*, 2nd ed. Philadelphia: Lippincott Williams & Wilkins, 2002, 933 pages.
- [144] H. Alkadhi, Ed., *Wie funktioniert CT? eine Einführung in Physik, Funktionsweise und klinische Anwendungen der Computertomographie*. Berlin Heidelberg: Springer, 2011, 272 pages.
- [145] S. Carmignato, W. Dewulf, and R. Leach, Eds., *Industrial X-Ray Computed Tomography*. Cham: Springer International Publishing, 2018.
- [146] V. Nair and G. E. Hinton, “Rectified Linear Units Improve Restricted Boltzmann Machines,”

- [147] K. Hara, D. Saito, and H. Shouno, "Analysis of function of rectified linear unit used in deep learning," in *2015 International Joint Conference on Neural Networks (IJCNN)*, Killarney, Ireland: IEEE, 2015, pages 1–8.
- [148] D. P. Kingma and J. Ba. "Adam: A Method for Stochastic Optimization." (2017), [Online]. Available: <http://arxiv.org/abs/1412.6980> (visited on 03/27/2024), preprint.
- [149] Z. Zhang, "Improved Adam Optimizer for Deep Neural Networks," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, Banff, AB, Canada: IEEE, 2018, pages 1–2.
- [150] Y. Zhang, N. Yue, M. Su, B. Liu, Y. Ding, Y. Zhou, H. Wang, Y. Kuang, and K. Nie, "Improving CBCT quality to CT level using deep learning with generative adversarial network," *Medical Physics*, volume 48, number 6, pages 2816–2826, 2021.
- [151] J. Lee, J. Gu, and J. C. Ye, "Unsupervised CT Metal Artifact Learning Using Attention-Guided -CycleGAN," *IEEE Transactions on Medical Imaging*, volume 40, number 12, pages 3932–3944, 2021.
- [152] O. Ronneberger, P. Fischer, and T. Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation." (2015), [Online]. Available: <http://arxiv.org/abs/1505.04597> (visited on 03/24/2024), preprint.
- [153] K. Simonyan and A. Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition." (2015), [Online]. Available: <http://arxiv.org/abs/1409.1556> (visited on 03/27/2024), preprint.
- [154] F. Yang, D. Zhang, K. Huang, Z. Gao, and Y. Yang, "Incomplete projection reconstruction of computed tomography based on the modified discrete algebraic reconstruction technique," *Measurement Science and Technology*, volume 29, number 2, page 025 405, 2018.
- [155] M. Du, K. Liang, L. Zhang, H. Gao, Y. Liu, and Y. Xing, "Deep-Learning-Based Metal Artefact Reduction With Unsupervised Domain Adaptation Regularization for Practical CT Images," *IEEE Transactions on Medical Imaging*, volume 42, number 8, pages 2133–2145, 2023.
- [156] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, volume 13, number 4, pages 600–612, 2004.
- [157] T. Heimann, B. Van Ginneken, M. Styner, *et al.*, "Comparison and Evaluation of Methods for Liver Segmentation From CT Datasets," *IEEE Transactions on Medical Imaging*, volume 28, number 8, pages 1251–1265, 2009.
- [158] J. Solomon and E. Samei, "A generic framework to simulate realistic lung, liver and renal pathologies in CT imaging," *Physics in Medicine and Biology*, volume 59, number 21, pages 6637–6657, 2014.

-
- [159] M. Eckstein, J. Bartroff, C. Abbey, J. Whiting, and F. Bochud, “Automated computer evaluation and optimization of image compression of x-ray coronary angiograms for signal known exactly detection tasks,” *Optics Express*, volume 11, number 5, page 460, 2003.
- [160] Y. Li, J. Chen, J. L. Brown, S. T. Treves, X. Cao, F. H. Fahey, G. Sgouros, W. E. Bolch, and E. C. Frey, “DeepAMO: A multi-slice, multi-view anthropomorphic model observer for visual detection tasks performed on volume images,” *Journal of Medical Imaging*, volume 8, number 04, 2021.
- [161] A. Wunderlich and F. Noo, “Evaluation of the impact of tube current modulation on lesion detectability using model observers,” in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vancouver, BC: IEEE, 2008, pages 2705–2708.
- [162] A. H. Andersen and A. C. Kak, “Simultaneous Algebraic Reconstruction Technique (SART): A Superior Implementation of the Art Algorithm,” *Ultrasonic Imaging*, volume 6, number 1, pages 81–94, 1984.
- [163] A. Byl, S. Sawall, M. Rafecas, C. Hoeschen, and M. Kachelrieß, “Low-dose CT: Reducing tube current, number of projections, or both?” In *2023 IEEE Medical Imaging Conference Abstract Booklet*, IEEE, 2023, pages 1009–1010.
- [164] A. Byl, M. Knaup, M. Rafecas, C. Hoeschen, and M. Kachelrieß, “Detruncation of clinical CT scans using a discrete algebraic reconstruction technique prior,” in *7th International Conference on Image Formation in X-Ray Computed Tomography*, J. W. Stayman, Ed., Baltimore, United States: SPIE, 2022, page 74.
- [165] N. M. Sheth, W. Zbijewski, M. W. Jacobson, G. Abiola, G. Kleinszig, S. Vogt, S. Soellradl, J. Bialkowski, W. S. Anderson, C. R. Weiss, G. M. Osgood, and J. H. Siewerdsen, “Mobile C-Arm with a CMOS detector: Technical assessment of fluoroscopy and Cone-Beam CT imaging performance,” *Medical Physics*, volume 45, number 12, pages 5420–5436, 2018.

List of Publications

Parts of this thesis have been published in the following journal articles and conference contributions:

Journal Articles

- [1] **A. Byl**, L. Klein, S. Sawall, S. Heinze, H.-P. Schlemmer, and M. Kachelrieß, “Photon-counting normalized metal artifact reduction (NMAR) in diagnostic CT,” *Medical Physics*, volume 48, number 7, pages 3572–3582, 2021.
- [2] **A. Byl**, “A semi-analytical solution to the Schrödinger equation with Gaussian well,” *Student Undergraduate Research E-journal!*, volume 4, 2018.

Conference Contributions

- [1] **A. Byl**, S. Sawall, M. Rafecas, C. Hoeschen, and M. Kachelrieß, “Low-dose CT: Reducing tube current, number of projections, or both?” In *2023 IEEE Medical Imaging Conference Abstract Booklet*, IEEE, 2023, pages 1009–1010.
- [2] **A. Byl**, M. Knaup, M. Rafecas, C. Hoeschen, and M. Kachelrieß, “Detruncation of clinical CT scans using a discrete algebraic reconstruction technique prior,” in *7th International Conference on Image Formation in X-Ray Computed Tomography*, SPIE, volume 12304, 2022, pages 353–359.
- [3] **A. Byl**, L. Klein, J. Hardt, S. Sawall, H.-P. Schlemmer, S. Heinze, and M. Kachelrieß, “Metal artifact reduction in photon counting CT using pseudo-monochromatic images,” in *6th International Conference on Image Formation in X-Ray Computed Tomography*, 2020, pages 256–259.
- [4] **A. Byl**, L. Klein, J. Hardt, E. Wehrse, H.-P. Schlemmer, S. Heinze, M. Uhrig, S. Sawall, and M. Kachelrieß, “Dedicated metal artefact reduction for photon counting CT,” in *ECR 2020 Book of Abstracts*, European Society of Radiology, 2020, page 138.

-
- [5] L. Klein, L. Enzmann, **A. Byl**, C. Liu, S. Sawall, A. Maier, J. Maier, M. Lell, and M. Kachelrieß, “Organ-specific vs. patient risk-specific tube current modulation in thorax CT scans covering the female breast,” in *7th International Conference on Image Formation in X-Ray Computed Tomography*, SPIE, volume 12304, 2022, pages 329–335.
- [6] L. Klein, L. Enzmann, **A. Byl**, C. Liu, S. Sawall, A. Maier, J. Maier, M. Lell, and M. Kachelrieß, “Potential CT radiation dose reduction to the female breast by a novel risk-minimizing tube current modulation,” in *Proceedings of the 108th Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA)*, 2022.
- [7] P. Trapp, **A. Byl**, L. Klein, S. Heinze, H.-P. Schlemmer, S. Sawall, and M. Kachelrieß, “Bin-combination-based noise reduction for metal artifact reduction in photon counting CT,” in *ECR 2021 Book of Abstracts*, European Society of Radiology, 2021.
- [8] S. Heinze, K. Yen, A. Tsaklakis, L. Klein, **A. Byl**, M. Kachelrieß, H.-P. Schlemmer, and S. Sawall, “Forensische bildgebung 2.0,” in *Tagungsband der 99. Jahrestagung der Deutschen Gesellschaft für Rechtsmedizin*, 2020.
- [9] L. Klein, L. Hardt, **A. Byl**, E. Wehrse, S. Heinze, M. Uhrig, H.-P. Schlemmer, C. Ziener, S. Heinze, and M. Kachelrieß, “Kernel considerations for high resolution photon-counting CT: Dose reduction versus spatial resolution,” in *ECR 2020 Book of Abstracts*, European Society of Radiology, 2020, page 137.

Acknowledgments

I would like to express my gratitude to everybody who has helped and supported my research:

Prof. Dr. Joao Seco for his role as first examiner, and his guidance in the context of the thesis advisory committee.

Prof. Dr. Marc Kachelrieß, for the opportunity to do my doctoral studies in his department and supervising my research, as well as his invaluable insights, expertise and suggestions. Thank you for always finding the time to discuss my work.

Prof. Dr. Jürgen Hesser, for his input and encouragement during the thesis advisory committee meetings.

The external partners of the KI-INSPIRE project, especially Prof. Dr. Christoph Hoeschen, Prof. Dr. Magdalena Rafecas, and Dr.-Ing. Ali Pashazadeh, for the successful collaboration.

Dr. Joscha Maier and Dr. Stefan Sawall for their assistance and constructive comments. All current and former colleagues in E025, with whom I have been more than glad to spend time inside and outside the office. My special thanks go to Fabian (for both his friendship and coffee), Carlo, Edith, Elias, Markus, Mishal, Philip, and Tim.

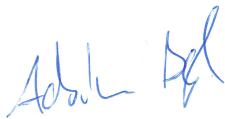
Prof. Dr. André Schöning and Priv.-Doz. Dr. Zoltán Harman for agreeing to serve as examiners for the oral defense.

I am sincerely grateful to my family for their encouragement and support during my studies.

Finally, I express my deep gratitude to Emmy for her unwavering support and faith in me. Without you, this work may not have been possible.

This work was supported by the German Federal Ministry for the Environment, Nature Conservation, Nuclear Safety and Consumer Protection (BMUV) under grant 67KI2036B.

I hereby assure, that I composed this work by myself and did not use any other than the listed resources.

A handwritten signature in blue ink, appearing to read 'Adrian B. B.', is positioned below the text.

Heidelberg, March 28, 2024

