Learning CT Scatter Estimation Without Labeled Data – A Feasibility Study

Joscha Maier, Luca Jordan, Elias Eulig, Fabian Jäger, Stefan Sawall, Michael Knaup, and Marc Kachelrieß

***Abstract*—Since x-ray scattering is a major cause of artifacts, its correction is a crucial step in almost any CT application. Most existing approaches, however, are based on complex theoretical models that need to be tailored to that particular application. To perform scatter estimation in absence of such models, we propose the unsupervised deep scatter estimation (uDSE). Here, uDSE combines a scatter estimation network that operates in projection domain with a scatter correction layer and CT reconstruction layer. In that way scatter estimation can be trained using an unsupervised Wassersten GAN (WGAN) setup in which the parameters of the scatter estimation network are optimized such that the resulting scatter corrected reconstructions cannot be distinguished from samples of a true artifact-free reference set. To demonstrate the feasibility of the proposed approach, uDSE is evaluated for simulated CBCT scans. Applied to the corresponding test data, uDSE is able to remove most of the present scatter artifacts and yields similar CT value accuracy (mean error of 27.9 HU vs. 24.7 HU) as a state-of-the-art supervised scatter estimation approach. Thus, uDSE may be used in the future to learn scatter estimation in cases where labels are not available or cannot be generated with sufficient accuracy.**

***Index Terms*—CT, scatter estimation, deep learning, unsupervised learning.**

I. Introduction

THE contribution of scattered x-rays to the acquired projection data leads to a violation of CT reconstruction criteria, and thus, to the introduction of CT artifacts. In particular this holds true for cone-beam CT (CBCT), where scatter-to-primary ratios may easily be in the order of 1 and above. Therefore, scatter correction is a crucial preprocessing step to achieve diagnostic image quality. Typically, existing approaches implement such a correction by deriving an estimate of the present scatter distribution and by subtracting it subsequently from the acquired raw data. Here, the scatter distribution can either be estimated using dedicated hardware such as beam blockers or primary modulation grids [1]–[5],or using software-based approaches. The latter rely on physical, empirical, or consistency-based models to predict x-ray scattering [6]–[13],or more recently, on neural networks that make use of such models during training, e.g. by being trained to reproduce the output of Monte Carlo simulations [14]–[18]. While these approaches have proven great potential in terms of accuracy and computation time, their performance highly depends on the quality and the availability of labeled data. Since several applications may lack such data, this study proposes the unsupervised deep scatter estimation (uDSE) which is able to overcome this limitation. Instead of using a supervised setup, uDSE relies on a Wasserstein GAN (WGAN) setup that can be trained without labeled data or prior knowledge about the CT scanner's x-ray and scatter properties [19]. Here we demonstrate the feasibility of the proposed approach using simulated CBCT data and compare the results against DSE, our supervised scatter estimation approach [14], [15].

Dr. J. Maier, L. Jordan, and Dr. M. Knaup are with the German Cancer Research Center (DKFZ), Heidelberg, Germany. E. Eulig, F. Jäger, Dr. S. Sawall and Prof. Dr. M. Kachelrieß are with the German Cancer Research Center (DKFZ), Heidelberg, Germany and with the Heidelberg University, Heidelberg, Germany

Corresponding author: Dr. Joscha Maier (joscha.maier@dkfz.de)

II. Material and Methods

A. Deep Scatter Estimation (DSE)

The basic idea of the DSE approach is illustrated in figure 1. Here, DSE uses a U-net-like architecture to predict scatter as a function of the acquired projection data. To learn the corresponding mapping, DSE is trained to reproduce Monte Carlo simulations, i.e. the U-net's weights are determined by minimizing the following loss function:

$$L_{\mathrm{DSE}}(\theta) = \sum_{n}^{B} \left| \frac{\mathrm{DSE}_{\theta}(I_n) - S_n}{S_n} \right|, \qquad (1)$$

where θ denotes the parameter vector, n the sample number within a batch of size B, I_n the flat field-corrected intensities, and S_n the Monte Carlo scatter

7th International Conference on Image Formation in X-Ray Computed Tomography, edited by Joseph Webster Stayman, Proc. of SPIE Vol. 12304, 123042Z
 · 0277-786X · doi: 10.1117/12.2646588

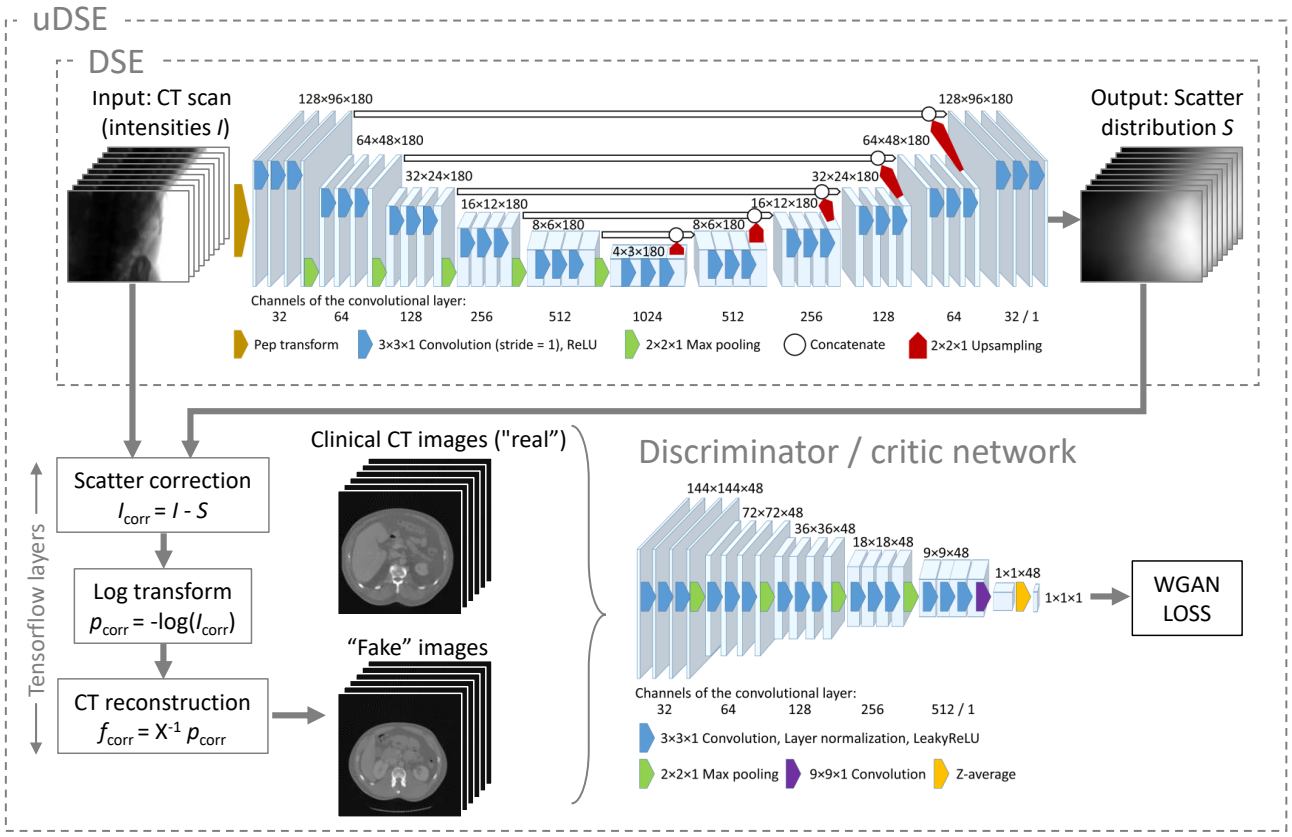

Fig. 1. Schematic of the DSE and the proposed uDSE approach. DSE uses a U-net-like architecture to predict scatter as a function of the acquired projection data. In a supervised setup DSE is trained to reproduce a ground truth Monte Carlos scatter distribution. uDSE, in contrast, uses a Wasserstein GAN (WGAN) setup. Here, a scatter correction followed by a CT reconstruction is performed. Subsequently, the correction is evaluated by a critic network that is trained simultaneously to recognize scatter artifact-free images. By optimizing the weights of the scatter estimation network to fool the critic network, scatter estimation can be learned without labels.

estimate. It has to be noted that the first layer of the DSE network performs a "pep"-transform

$$T_{\text{pep}} : I \rightarrow -I \cdot \ln(I), \tag{2}$$

to be consistent with our DSE publication [14].

B. *Unsupervised Deep Scatter Estimation (uDSE)*

The proposed uDSE approach, shown in figure 1, extends the concept of DSE to cases where labeled data are not available. To do so, it is composed of a generator network and a critic network. Here, the generator combines the DSE network with a scatter correction layer and a Feldkamp reconstruction layer, such that it is able to map acquired intensities I of a CT scan to scatter-corrected CT reconstructions. The critic network, in turn, is designed to distinguish between the generator's output and true scatter-free CT reconstructions. Thus, letting the critic network act as loss function for the generator allows to learn CT scatter estimation without labeled or paired data, respectively.

Here, the corresponding optimization is performed using a WGAN setup in which the generator netork $G_{\theta_g}(I)$ and the critic network $C_{\theta_c}(f)$ are optimized in an alternating manner according to the following loss functions:

$$L_{\text{critic}}(\theta_c) = \sum_{n}^{B} C_{\theta_c}(G_{\theta_g}(I_n)) - C_{\theta_c}(f_{\text{real, n}}), \tag{3}$$

$$L_{\text{gen}}(\theta_g) = -\sum_{n}^{B} C_{\theta_c}(G_{\theta_g}(I_n)), \tag{4}$$

where θ_c and θ_g denote the parameter vectors of the generator and the critic network, n is the sample number within a batch of size B, and f_{real} corresponds to a sample from a set of almost scatter-free clinical CT reconstructions.

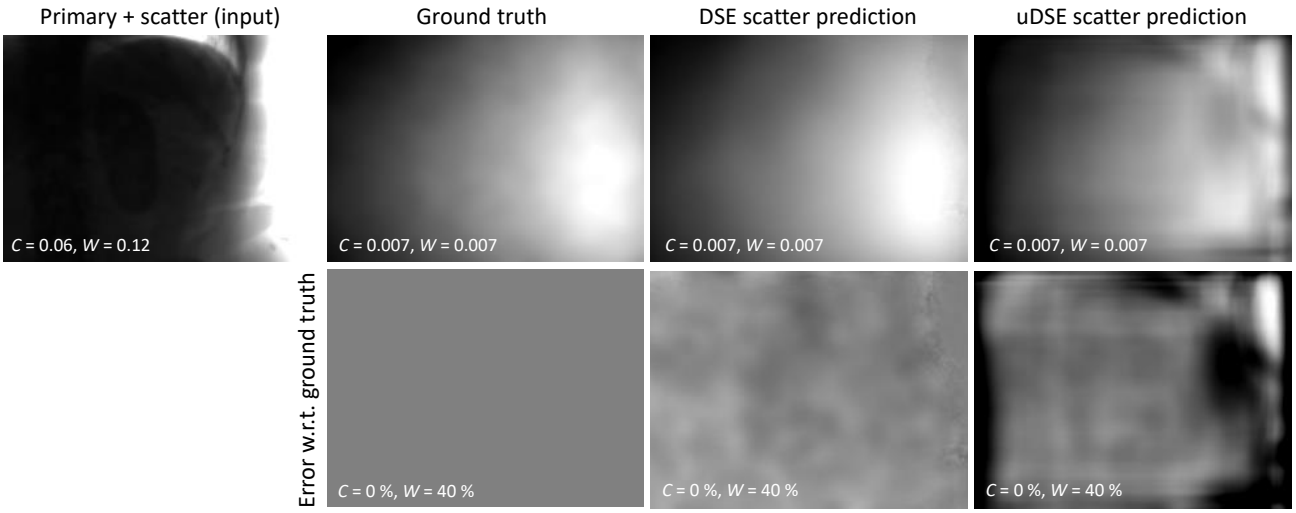

Fig. 2. DSE and uDSE scatter estimates for an exemplary shifted detector projection of the test dataset (top row), as well as the corresponding error with respect to the ground truth Monte Carlo scatter distribution (bottom row).

C. Datasets

In the present study, DSE as well as uDSE were trained and tested on simulated CBCT data. Therefore, clinical CT reconstructions of 65 patients were used as prior. Based on the corresponding voxel volumes f_{prior}, CBCT scans with 360 views and an angular coverage of 360° were simulated at five different z-positions within the abdomen region using a tube voltage of 120 kV, a source-to-isocenter distance of 700 mm, a source-to-detector distance of 1100 mm, and a 1024×768 flat detector with an isotropic pixel spacing of 0.39 mm. Furthermore, a shifted detector was used to increase the field of measurement to about 380 mm. Using this setup, three datasets were generated using different patients: one scatter-corrupted dataset (30 patients) from which input data for the generator network were sampled during training, one scatter-free dataset (30 patients) that was used to provide ideal reference CT reconstructions for the critic network, and a scatter-corrupted dataset (5 remaining patients) for testing. In any case, the scatter-corrupted data were simulated as $I = I_{\text{p}} + S_{\text{MC}}$, where I_p corresponds to a polychromatic forward projection of the prior volume and S_{MC} is a scatter distribution that was generated using our in-house Monte Carlo simulation. The ideal reference data, on the other hand, correspond to a CBCT reconstruction of only the primary intensities, i.e. $f_{\text{real}} = X^{-1}(-\ln(I_{\text{p}}))$, with X^{-1} being the CBCT reconstruction operator. It has to be noted that f_{real} is not sampled directly from the set of clinical CT reconstructions, but is generated via forward and back-projection, to have the same spatial resolution and the same field of measurement as the reconstructions provided by the generator network.

D. Training and Evaluation

The uDSE approach was trained using the datasets described in section II-C. However, to avoid memory issues as well as to increase the computational performance, a 2-fold angular downsampling to 180 views and an 8-fold spatial downsampling to a detector size of 128×96 was performed in advance. This downsampling can be justified by the fact that scatter distributions are known to be of low frequency. Therefore, we only expect a minor degradation of accuracy compared to a training that uses the full size projection data. Given the small detector size, the internal reconstruction operations were also performed on a low resolution grid with $144 \times 144 \times 48$ voxels and an isotropic spacing of 2.7 mm.

The corresponding optimization was implemented according to section II-B using the Tensorflow framework. Here, all hyperparameters except for the batch size (this study: $B = 16$) and the number of iterations (this study: $N_{\text{itr}} = 5000$) were chosen following to the original WGAN publication, i.e. five updates of the critic network are followed by one update of the generator network using an RMSProp optimizer with a learning rate of 0.00005 and a weight clipping to $[-0.01, 0.01]$ in the critic network.

Finally, uDSE was applied to the test data and compared against DSE which was trained in a supervised setting using the Monte Carlo scatter distributions a

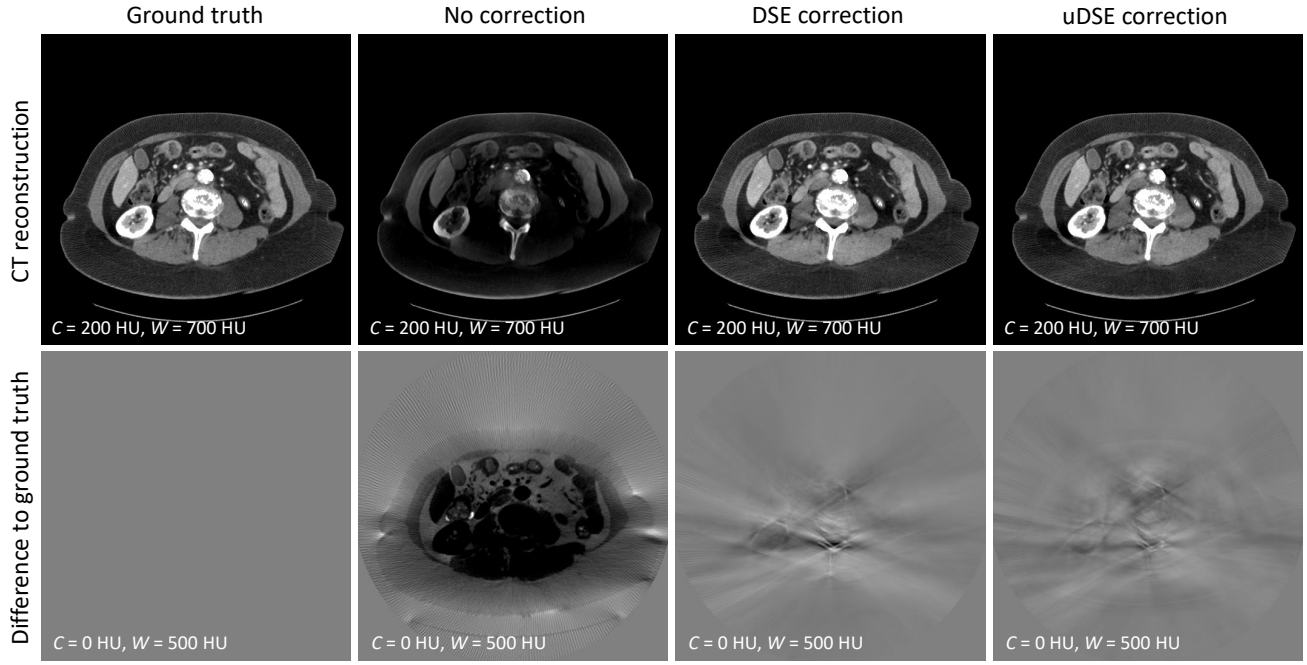

Fig. 3. CT reconstruction without and with DSE and uDSE scatter correction (top row), as well as difference images to the scatter free ground truth (bottom row).

labels. In contrast to the training data, the scatter correction was performed on the full size projection data. Therefore, all scatter predictions were upsampled to the original detector size of 1024×768 pixels prior to scatter correction.

III. Results

To estimate scatter, DSE and uDSE were trained as described in section II-D and evaluated for the five patients of the test dataset. Exemplary scatter predictions are shown in figure 2. Here, the DSE approach yields scatter estimates that are almost equal to the Monte Carlo ground truth while uDSE shows slightly higher deviations. This can be attributed to the fact the uDSE loss function is evaluated in image domain. Therefore, the scatter estimate is less reliable in regions where it has a low impact on image quality, i.e. in regions with low scatter-to-primary ratio. In particular this explains the poor accuracy of uDSE scatter estimates in air regions without patient intersection.

A quantitative evaluation of the scatter estimates in terms of the mean absolute percentage error with respect to the ground truth yields similar trends. Here, the average error of DSE for all 9000 projections of the test dataset is 3.5 % while the average error of uDSE is 9.8 %.

CT reconstructions with and without scatter correction are shown in figure 3. Here, DSE as well as uDSE are able to remove most of the artifacts that are present in the uncorrected reconstruction and provide CT images that are almost equal to the scatter-free ground truth. Quantitatively, the application of the scatter correction improves the mean absolute error of the CT values from 160.7 HU (no correction) to 24.7 HU (DSE) or 27.9 HU (uDSE), respectively.

IV. Discussion and Conclusion

This study introduces a novel approach to learn scatter estimation without labeled data. To do so, the proposed uDSE makes use of a WGAN setup in which the generator network is optimized such that its output, i.e. scatter corrected reconstructions, cannot be distinguished from samples of an artifact-free reference set. Here, we demonstrate the feasibility of uDSE using CBCT simulations as input and clinical CT reconstructions as reference. However, it has to be noted that uDSE is not restricted to this particular choice but can be trained with any tomographic input and any scatter-free reference as long as both distributions are sufficiently equal after scatter correction. The fact that this condition is perfectly met here, i.e. input and reference distributions only differ by scatter, can be considered as a limitation of this study. Practically,

both distribution may additionally differ by the amount of beam hardening, the dynamic range, or the contrast media distribution for instance. Investigating uDSE's performance in such cases is subject to further research and may require the incorporation of additional constraints and correction layers. In the current setup, however, uDSE is able to remove most of the present scatter artifacts and yields similar CT value accuracy (mean error of 27.9 HU vs. 24.7 HU) as a state-of-the-art supervised scatter estimation approach. Thus, uDSE has the potential to extend the concept of neural network-based scatter estimation and correction to scenarios where labels are not available or cannot be generated with sufficient accuracy.

Acknowledgment

Parts of the reconstruction software RayConStruct IR and of the Monte Carlo software RayConStruct MC were provided by RayConStruct® GmbH, Nürnberg, Germany.

References

[1] R. Ning, X. Tang, and D. Conover, "X-ray scatter correction algorithm for cone beam CT imaging," *Medical Physics*, vol. 31, no. 5, pp. 1195–1202, apr 2004.

[2] L. Zhu, N. R. Bennett, and R. Fahrig, "Scatter Correction Method for X-Ray CT Using Primary Modulation: Theory and Preliminary Results," *IEEE Transactions on Medical Imaging*, vol. 25, no. 12, pp. 1573–1587, dec 2006.

[3] R. Grimmer, R. Fahrig, W. Hinshaw, H. Gao, and M. Kachelrieß, "Empirical Cupping Correction for CT Scanners with Primary Modulation (ECCP)," *Medical Physics*, vol. 39, no. 2, pp. 825–831, jan 2012.

[4] L. Ritschl, R. Fahrig, M. Knaup, J. Maier, and M. Kachelrieß, "Robust primary modulation-based scatter estimation for cone-beam CT," *Medical Physics*, vol. 42, no. 1, pp. 469–478, jan 2015.

[5] S. Hsieh, "Estimating scatter in cone beam CT with striped ratio grids: A preliminary investigation," *Medical Physics*, vol. 43, no. 9, pp. 5084–5092, aug 2016.

[6] W. Swindell and P. M. Evans, "Scattered radiation in portal images: A Monte Carlo simulation and a simple physical model," *Medical Physics*, vol. 23, no. 1, pp. 63–73, jan 1996.

[7] B. Ohnesorge, T. Flohr, and K. Klingenbeck-Regn, "Efficient object scatter correction algorithm for third and fourth generation CT scanners," *European Radiology*, vol. 9, no. 3, pp. 563–569, mar 1999.

[8] G. Poludniowski, P. M. Evans, V. N. Hansen, and S. Webb, "An efficient Monte Carlo-based algorithm for scatter correction in keV cone-beam CT," *Physics in Medicine and Biology*, vol. 54, no. 12, pp. 3847–3864, jun 2009.

[9] J. Star-Lack, M. Sun, A. Kaestner, R. Hassanein, G. Virshup, T. Berkus, and M. Oelhafen, "Efficient scatter correction using asymmetric kernels," in *SPIE Medical Imaging Conference*, E. Samei and J. Hsieh, Eds., vol. 7258, feb 2009, p. 72581Z.

[10] M. Sun and J. M. Star-Lack, "Improved scatter correction using adaptive scatter kernel superposition," *Physics in Medicine and Biology*, vol. 55, no. 22, pp. 6695–6720, nov 2010.

[11] E. Meyer, C. Maaß, M. Baer, R. Raupach, B. Schmidt, and M. Kachelrieß, "Empirical scatter correction (ESC): A new CT scatter correction method and its application to metal artifact reduction," in *IEEE Medical Imaging Conference*. IEEE, oct 2010, pp. 2036–2041.

[12] M. Baer and M. Kachelrieß, "Hybrid scatter correction for CT imaging," *Physics in Medicine and Biology*, vol. 57, no. 21, pp. 6849–6867, nov 2012.

[13] A. Maslowski, A. Wang, M. Sun, T. Wareing, I. Davis, and J. Star-Lack, "Acuros CTS: A fast, linear Boltzmann transport equation solver for computed tomography scatter - Part I: Core algorithms and validation," *Medical Physics*, vol. 45, no. 5, pp. 1899–1913, may 2018.

[14] J. Maier, E. Eulig, T. Vöth, M. Knaup, J. Kuntz, S. Sawall, and M. Kachelrieß, "Real-time scatter estimation for medical CT using the deep scatter estimation: Method and robustness analysis with respect to different anatomies, dose levels, tube voltages, and data truncation." *Medical Physics*, vol. 46, no. 1, pp. 238–249, jan 2019.

[15] J. Maier, S. Sawall, M. Knaup, and M. Kachelrieß, "Deep Scatter Estimation (DSE): Accurate Real-Time Scatter Estimation for X-Ray CT Using a Deep Convolutional Neural Network," *Journal of Nondestructive Evaluation*, vol. 37, no. 3, pp. 57:1–9, sep 2018.

[16] Y. Nomura, Q. Xu, H. Shirato, S. Shimizu, and L. Xing, "Projection-domain scatter correction for cone beam computed tomography using a residual convolutional neural network," *Medical Physics*, p. mp.13583, jun 2019.

[17] Y. Jiang, C. Yang, P. Yang, X. Hu, C. Luo, Y. Xue, L. Xu, X. Hu, L. Zhang, J. Wang, K. Sheng, and T. Niu, "Scatter correction of cone-beam CT using a deep residual convolution neural network (DRCNN)," *Physics in Medicine and Biology*, vol. 64, no. 14, p. 145003, jul 2019.

[18] J. Erath, T. Vöth, J. Maier, E. Fournié, M. Petersilka, K. Stierstorfer, and M. Kachelrieß, "Deep learning-based forward and cross-scatter correction in dual-source CT," *Medical Physics*, vol. 48, no. 9, pp. 4824–4842, sep 2021.

[19] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein Generative Adversarial Networks," in *Proceedings of the 34th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, D. Precup and Y. W. Teh, Eds., vol. 70. PMLR, 2017, pp. 214–223.