NDT net e-Journal of Nondestructive Testing - ISSN 1435-4934 - www.ndt.net

14th Conference on Industrial Computed Tomography, Antwerp, Belgium (iCT 2025), www.ict2025.org

Simulating X-ray energy and detector signal processing of an industrial CT using implicit neural representations

Edwin Blum[1], Moritz Burmeister[1], Florian Stamer[1], Gisela Lanza[1]

[1]Karlsruhe Institute of Technology, Kaiserstraße 12, 76131 Karlsruhe, Germany, e-mail: edwin.blum@kit.edu, florian.stamer@kit.edu, gisela.lanza@kit.edu

Abstract

Simulating computed tomography (CT) systems offers numerous advantages, including the optimization of scan parameters, training of specialist personnel, and quantification of measurement uncertainties. Current simulation approaches, often referred to as virtual computed tomography (vCT), predominantly rely on analytical models. However, these models require extensive system-specific tuning to produce realistic synthetic measurements, creating a significant barrier to broader adoption and efficiency. To address this challenge, this work explores the potential of implicit neural representation (INR) as an alternative to the analytical models used in vCT. INRs excel at representing complex, high-dimensional data in a continuous and differentiable manner, making them a promising substitute for traditional analytical models. As a first building block, we propose a two-stage approach for simulating the X-ray beam energy and detector signal processing in industrial CT systems. This method is trained and evaluated using real-world data. Results demonstrate that the proposed INR-based architecture can accurately generate synthetic projections for parameter configurations within the training dataset. However, its poor performance on the test dataset highlights limitations in generalization beyond the training data. Potential methods to address these shortcomings are discussed. This study underscores the potential of INRs as a flexible framework for simulating complex CT systems. By capturing subtle system-specific characteristics and reducing dependence on explicitly defined parameterizations, INRs could pave the way for more versatile and efficient simulation models.

Keywords: Machine Learning, Simulation, vCT, Implicit Neural Representations

1 Introduction

Virtual computed tomography (vCT) is a simulation approach used to generate synthetic measurements based on CAD data. In order to achieve this, the physical processes of X-ray beam generation, ray attenuation and signal processing in the detector are approximated by mathematical models which represent the vCT [1]. vCTs are valuable for optimizing scan parameters, training specialist personnel, and quantifying measurement uncertainties in real computed tomography (CT) systems [1–3]. The current generation of vCT relies primarily on analytical models, Monte Carlo simulations, or a combination of these approaches [2, 4]. While these methods provide reasonable approximations of real systems, they often suffer from lengthy computation times and require extensive tuning specific to an individual CT system to produce realistic synthetic measurements, limiting their user-friendliness and practical utility [5]. Despite the availability of initial commercial solutions [2], these challenges remain a barrier to broader adoption and efficiency. This work presents a novel approach to vCT by using implicit neural representatios (INR) to model X-ray beam generation and detector signal processing. INRs parametrize these processes based on real X-ray projections and their corresponding CT parameters (e.g. tube voltage, current, and detector integration time). This eliminates the need for manual tuning of the analytical models of current vCT. Instead of relying on predefined mathematical formulations, INRs learn functions directly from data, enabling them to approximate processes that are analytically intractable. This is, for example, the case with X-ray beam models. Although this approach risks generating physically inaccurate results, it offers advantages when analytical models require oversimplification. Furthermore, real CT systems show system specific characteristics such as X-ray source and detector wear. Figure 1 illustrates such detector wear. Ideally, each detector pixel is equally sensitive when exposed to a specific X-ray beam, resulting in a homogeneous projection. However, in the real world this is not the case with pixel response changing over time due to aging and burn-in effects. The left, predominantly homogeneous, projection was created using a factory new Zeiss Metrotom 1500, where as the right projection was created using a 10 year old Zeiss Metrotom 800. The right projection displays distinct noise, as well as a characteristic pattern of light and dark spots that result from differently stressed pixels in the life cycle of the CT. As an object to be measured is usually placed in the center of the detector, the outer pixel are generally exposed to higher energy X-ray beams, resulting in higher wear and thus the visible vignette effect.

Current vCT mostly ignore such behavior, as they are hard to replicate using analytical models. Usually post processing steps like applying noise to ideal synthetic projections are performed to achieve more realistic results [7]. This again motivates the use of INR within vCT, as such characteristics can be learned from real data.

With this work, we want to explore the potential of INR for use within vCT. If INR are able to approximate good models for simulating CTs based only on real data, a new kind of vCT can be achieved, which not only requires any manual tuning, but is also capable of modeling system specific characteristics. With the long-term goal of a fully functioning INR based vCT, we are focusing this work on an initial performance evaluation of INRs for modeling beam generation and detector signal processing in vCT, while the simulation of attenuation processes will be part of future work.

https://doi.org/10.58286/30752

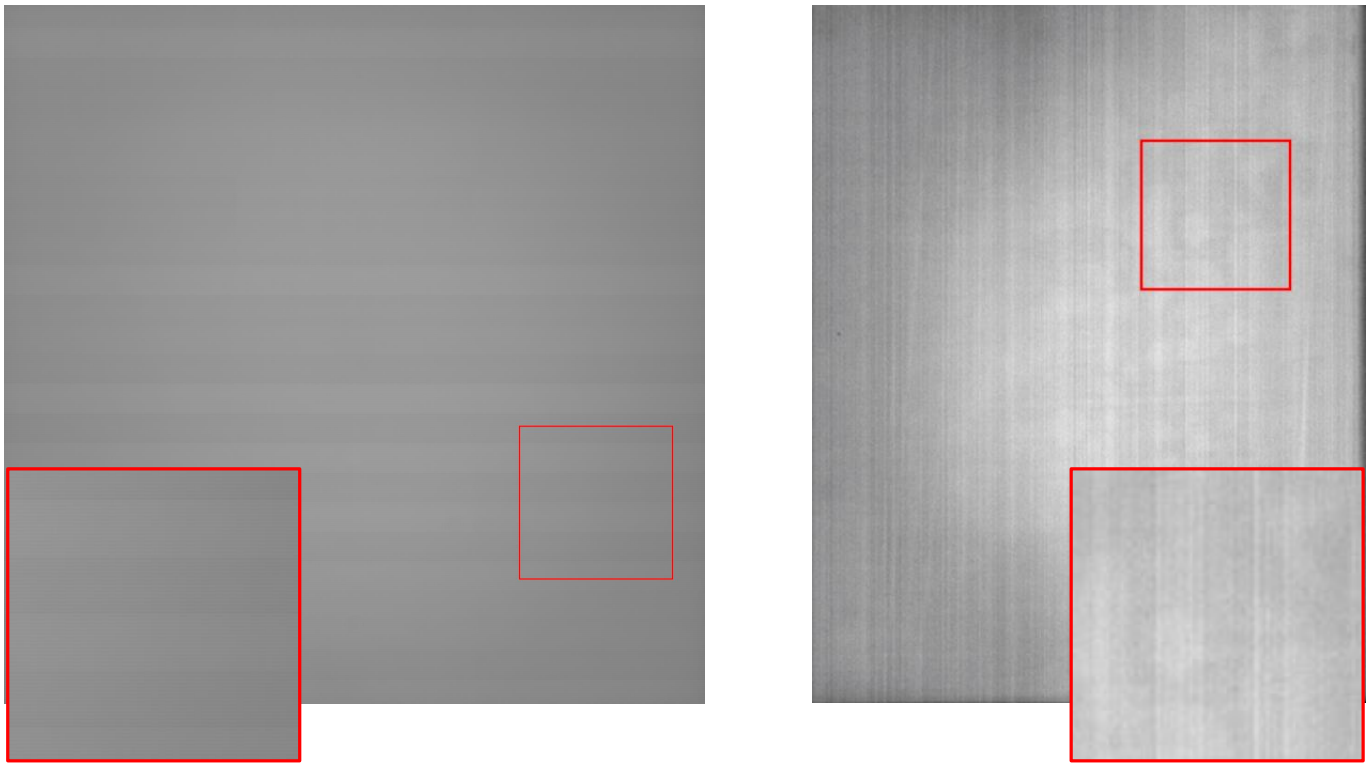

Figure 1: Difference in detector wear between a brand-new Zeiss Metrotom 1500 detector (left) and a 10-year-old Zeiss Metrotom 800 detector (right), the latter displaying pronounced wear and increased noise.

2 Fundamentals of Implicit Neural Representations

INRs are an emerging field in the context of supervised learning, capable of approximating continuous differentiable signals through neural networks. This concept has existed for many years, but the approximation quality of old approaches was severely limited. In particular, very complex signals could not be approximated with sufficient quality. In recent years, new, more powerful INR architectures have been introduced that have led to a new state-of-the-art of INR, which are able to represent very complex signals with a high degree of accuracy. This makes them especially suited for approximating physical problems [8]. INR work on partial differential equations (PDE) in the form of:

$$F(x, \Phi, \nabla_x \Phi, \nabla_x^2 \Phi, \ldots) = 0, \quad \Phi : x \mapsto \Phi(x)$$

where Φ is the function of interest mapping some sample points x to a value $\Phi(x)$. The goal is to parametrize the function Φ by a neural network in a way, to full fill every constraint of the PDE F. Hence, parametrizing Φ implicitly by the PDE. This is achieved by training the neural network on a datatset of valid samples which abide all constraints of the PDE. Recent advancements in INR approximation quality are attributed to the development of new activation functions that are not only continuously differentiable but also possess derivatives that are continuously differentiable to all orders. This not only enables the approximation of a smooth and continuous signal, but also allows the functions to be tuned based on their derivatives, which is particularly advantageous for optimizing physical problems, where quantities such as gradients, velocities, or accelerations — often derived from the signal itself — play a crucial role in accurately modeling and solving these systems [8, 9].
Formulated as a machine learning optimization problem, INRs can be described as follows:

$$\text{Loss}(\theta, \mathscr{D}) = \left\| F\left(x_i, \Phi(x_i), \Phi_\theta(x_i), \nabla_x \Phi_\theta(x_i), \nabla_x^2 \Phi_\theta(x_i), \ldots\right) \right\|_2^2$$

An INR models the function $\Phi(x)$ using a neural network $\Phi_\Theta(x)$ where Θ denotes the network's parameters, including weights and biases. During training, a dataset $\mathscr{D} = \{x_i, \Phi(x_i)\}$, consisting of input-output pairs that satisfy the governing PDE, is used to optimize $\Phi_\Theta(x)$ to accurately capture the relationship between x and $\Phi(x)$. Typically, the L2 norm is employed to minimize the deviation between the predicted and target values. Additionally, constraints involving the derivatives of the function can be incorporated into the loss function to ensure that the neural network not only approximates the target values but also respects specific properties of the function, such as smoothness, continuity, or compliance with physical laws dictated by the PDE. Since an INR represents $\Phi_\Theta(x)$ as a continuous signal, it enables the parameterized function to generalize effectively, making accurate predictions at sample points not included in the training dataset.

3 Simulating X-Ray energy and Detector signal processing

Our goal is to create synthetic X-Ray projections. Like stated in the introduction, we focus solely on the X-ray generation and signal processing in the detector. Essentially, we try to approximate an empty X-Ray projection. In order to achieve this, we propose a two stage INR approach. During a measurement, X-Rays with a certain intensity are created. This intensity depends on multiple factors, among other things, it depends on the applied acceleration voltage, tube current, filter, as well as the anode material and angle [6]. Using the CT detector, a gray scale value is calculated based on this measured intensity. This value is always composed of the processes of X-ray beam generation in the X-ray source and the processing of these beams in the detector. For training the INRs we therefore have to rely on these gray scale pixel values. As mentioned above, the intensity of the X-Rays are dependent on multiple factors. By training the INRs specifically for one CT, the exact structure of the X-ray tube can be disregarded in the training process. In this first approach, we won't use any filter, as there are attenuation processes involved to alter the X-Ray beam spectrum while passing through the filter. Our goal is therefore to predict the resulting gray scale pixel values of a projection based on the defined tube voltage, current and integration time of the detector. The two stage approach is depicted in figure 2.

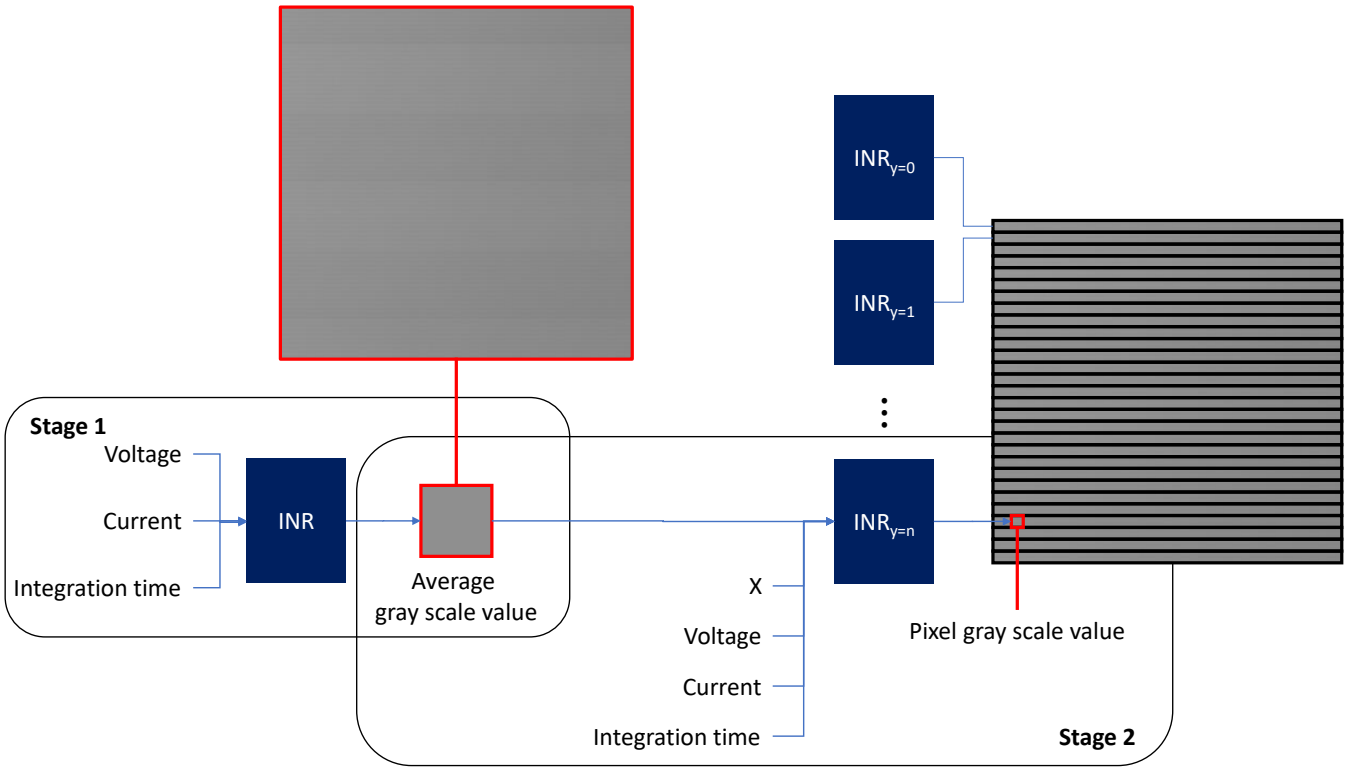

Figure 2: Overview of the two stage INR approach.

The first stage predicts the average gray scale pixel value of a projection for a given set of sample points (current, voltage, integration time). We interpret this prediction as a pixel-independent intensity, emitting uniformly from the X-ray source. This will come in handy when attenuation processes are to be modeled, as this predicted intensity can be altered based on the transmission length through some matter. Stage two maps this predicted intensity back to the pixel coordinate. We use one INR per pixel row. Ideally, one INR per pixel should be used, as every pixel can be interpreted as an individual entity. However, synthesizing full scale projections would lead to creating millions of small INR. E.g. the detector of the used Zeiss Metrotom 1500 consists of 3024x3024 pixels and would therefore need 9.144.576 INRs. By using one INR per pixel row, we reduce the amount of INRs to a easy manageable amount, but increase the complexity of the function to be parametrized. The goal for stage two is to predict the X-ray beam energy dependant pixel response, which leads to the calculated gray scale value. For this, the INR maps the stage one prediction together with the voltage, current and integration time of the current sample point for a given pixel (defined by the x-coordinate in the detector grid) to the measured gray scale pixel value. Essentially learning the unique energy dependant behaviour of every individual detector pixel. By running this two stage approach for every pixel of the detector, synthetic X-ray projections can be created. The sole constraint for the optimization of this approach is the deviation between the real measured value and the prediction, therefore, using the L2 norm as the loss function. For both stages we utilize SIREN activation functions

[8] and every INR consists of two hidden layers with 100 neurons each.

4 Experimental Setup

We evaluate our approach using data from a Zeiss Metrotom 1500. The "train" dataset consists of measurements with varying acceleration voltages $\{100, 125, 150, 175, 200\}kV$, currents $\{500, 600, 700, 800, 900, 1000\}\mu A$ and detector integration times $\{67, 100, 250\}ms$. A projection is created with every combination, resulting in 90 measurements. Furthermore, a "test" dataset is created consisting of the following parameter combinations with respective projections: tube voltages $\{111, 134, 164, 190\}kV$, currents $\{550, 650, 750, 850, 950\}\mu A$ and integration times $\{67, 100, 250\}ms$, resulting in 60 measurements. The detector resolution was set to 1512x1512 pixels. No filter was used and all other parameters of the CT were kept constant. For training and evaluation the data is normalized to lie between 0 and 1. Both stages are trained using ADAM optimizer and every training iteration consists of all data points available. Therefore, no batching is performed. Stage one is evaluated on the difference between the predicted and actual mean gray scale of the projections value for the train and test data set. With stage two creating synthetic X-ray projections, image metrics can be used to quantify the discrepancy between the real and synthetic projection. Therefore, the results for train and test data are evaluated based on the Mean-Square-Error (MSE), Peak-Signal-to-Noise-Ration (PSNR) and Structural Similarity Index (SSIM). MSE measures the average squared difference between corresponding pixels of two images. A lower MSE value indicates a closer match between the images, with zero indicating identical images. However, it does not account for perceptual quality, treating all pixel errors equally. PSNR is derived from MSE and represents the ratio of the maximum possible pixel value to the distortion (error). It's expressed in decibels (dB). Higher PSNR values indicate better image quality. Like MSE, it doesn't always correlate well with human perception. SSIM is a perceptual metric that compares two images based on structural information, luminance, and contrast. It aims to quantify perceived quality more effectively than MSE or PSNR. SSIM values range from -1 to 1, with 1 indicating identical images and higher values reflecting better perceived similarity [10].

5 Results

5.1 Stage 1: X-Ray energy approximation

We evaluate Stage one on the achieved deviation between the normalized prediction and ground truth of the train and test dataset. The results are depicted in figure 3.

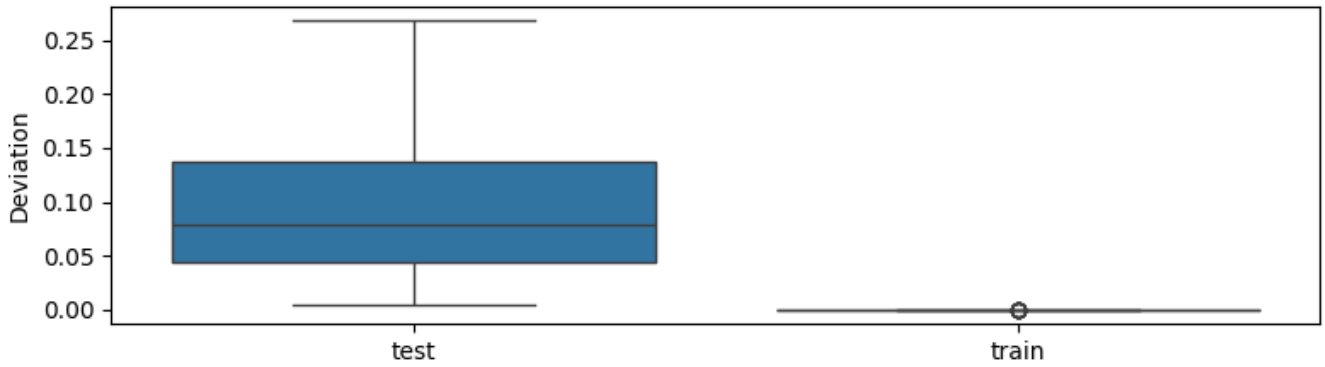

Figure 3: Results of stage 1.

The results on the train dataset show hardly any deviation between the average gray scale value and the prediction. On the other hand, the test data set shows a more inhomogeneous picture, with deviations close to zero up to a maximum of 0.268. The large deviations in the test data set indicates that the trained INR can only generalize with difficulty beyond the train dataset. Usually this can be a sign of neural network overfitting. However, further tests utlizing an early stopping criterion during training highlighted comparable results. Consequently, it is more likely that the function parameterized using the training data fails to accurately capture the physical relationships between voltage, current, and integration time as they occur in reality. Another potential explanation for the deviation between the training and test datasets could be the presence of statistical noise in the projections, which may impede the approximation of the underlying signal.

5.2 Stage 2: Detector signal processing

The MSE, PSNR and SSIM achieved for the train and test data sets are shown in figure 4. Once again, very good results are achieved for the train data set in all metrics. All values of the MSE and SSIM are at the optimum of 0 for the MSE and 1 for the SSIM with almost no scatter. The PSNR shows slightly more scatter, with the mean value achieved being a PSNR of around 50. This is a very good result. As a rule, one speaks of excellent results from a PSNR of 40. Once again, however, the test data set

shows significantly poorer results. This again suggests a lack of generalization across unlearned parameter combinations. Despite this, the promising results on the training dataset indicate the potential of the approach. INRs have shown the capability to model functions with the level of complexity required for applications in vCT. With an optimized INR architecture that supports better generalization, these advantages can be fully realized.

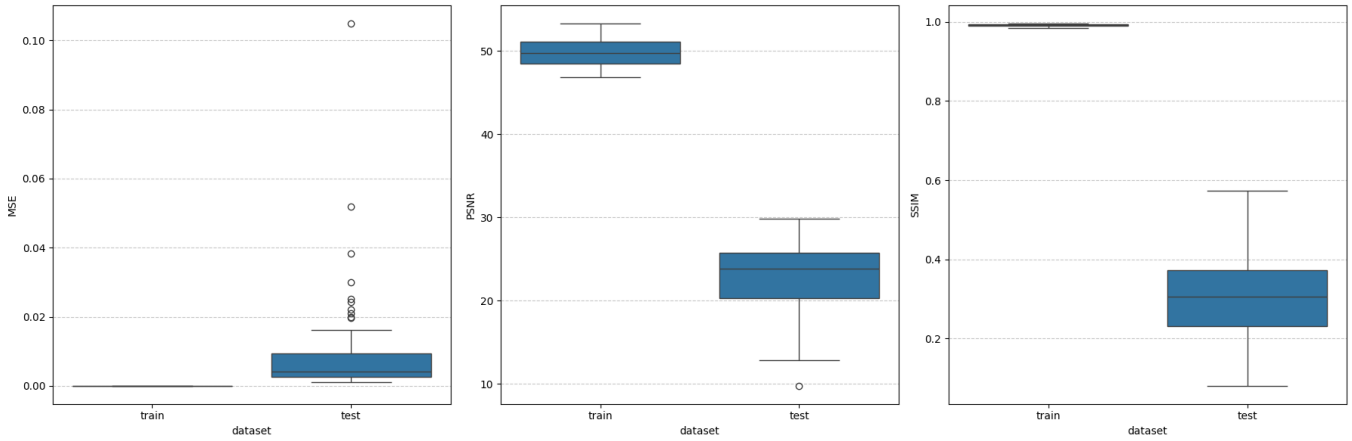

Figure 4: Results of stage 2.

6 Conclusion and future work

In this work, we introduced a novel approach for simulating industrial CT. Unlike traditional methods that rely on analytical models, we employed implicit neural representations to optimize such models using real CT data. As a first building block, we implemented a two stage approach to approximate X-ray beam energy and detector signal processing. Our results demonstrate that the proposed INR architecture can generate synthetic projections with high accuracy for parameter configurations present in the training dataset. These findings suggest that INRs are generally capable of parameterizing functions with the complexity required for vCT. However, the poor performance observed on the test dataset indicates a limited ability to generalize beyond the training data. This lack of generalization suggests that the parameterized function fails to adhere to the underlying physical laws, presenting a significant challenge for the application of INRs in vCT. Since training the approach on all possible parameter configurations is impractical, future work will focus on improving its physical accuracy. One promising direction is the use of Physics-Guided Neural Networks (PGNNs), which combine INRs with physics-based loss functions. PGNNs have been shown to provide more physically accurate predictions when applied to solving PDE [11]. As a next step, we will identify potential physics-based optimization criteria and assess their effectiveness in achieving a more realistic simulation. Additionally, the network architecture offers significant opportunities for refinement. In this initial performance analysis of the INR, the network was designed to predict plain gray scale values at each stage. However, this approach leads to a loss of critical information, particularly in stage two, where voltage, current, and integration time need to be reintroduced to achieve accurate approximations. To address this, stage one's output can be reformulated as an embedding that inherently captures these key parameters. This requires exploring effective methods for embedding continuous signals such as voltage, current, and integration time. One promising technique for this purpose is positional encoding [12]. Future work will focus on investigating how to effectively integrate such embeddings into our approach. Despite these limitations, the introduced two-stage approach successfully generated realistic X-ray projections. This demonstrates the potential of INRs to serve as a flexible framework for simulating complex CT systems, capturing subtle system-specific characteristics, and reducing the reliance on explicitly defined parameterizations. Additionally, this approach could pave the way for integrating data-driven methods into the simulation pipeline, enabling adaptive simulations tailored to specific applications or hardware configurations.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 471687386 and SFB-1574 – 465978608. In addition, the authors would like to thank the Ministry of Science, Research and Arts of the Federal State of Baden-Württemberg for the financial support of the projects within the InnovationCampus Future Mobility.

References

[1] C. Bellon, A. Deresch, C. Gollwitzer, G.-R. Jaenisch, Radiographic simulator aRTist: version 2, 18th World Conference on Nondestructive Testing, 2012.

[2] G.-R. Jaenisch, C. Bellon, U. Ewert, aRTist - Analytical RT Inspection Simulation Tool for Industrial Application, 2008.

[3] M. Reiter, M. Erler, C. Kuhn, C. Gusenbauer, J. Kastner, SimCT: a simulation tool for X-ray imaging, 6th Conference on Industrial Computed Tomography (iCT), vol. 21, no. 2, 2016
[4] Y. Kyriakou, T. Riedel, W.-A. Kalender. Combining deterministic and Monte Carlo calculations for fast estimation of scatter intensities in CT. Physics in Medicine and Biology 51, Nr. 18, 2006.
[5] E. Blum, F. Stamer, G. Lanza. CT-Messungen automatisieren/Automating CT measurements. wt Werkstattstechnik online 113, Nr. 11–12, 2023.
[6] S. Carmignato, W. Dewulf, R. Leach, Industrial X-Ray Computed Tomography. Cham: Springer International Publishing, 2018.
[7] J. Hiller, L.-M. Reindl, A Computer Simulation Platform for the Estimation of Measurement Uncertainties in Dimensional X-Ray Computed Tomography. Measurement 45, Nr. 8, 2012.
[8] V. Sitzmann, J.-N.-P. Martel, A.-W. Bergman, D.-B. Lindell, G. Wetzstein, Implicit Neural Representations with Periodic Activation Functions. arXiv, 2020.
[9] V. Saragadam, D. LeJeune, J. Tan, G. Balakrishnan, A. Veeraraghavan, R.-G. Baraniuk, WIRE: Wavelet Implicit Neural Representations, 2023.
[10] U. Sara, M. Akter, M.-S. Uddin, Image Quality Assessment through FSIM, SSIM, MSE and PSNR — A Comparative Study. Journal of Computer and Communications, 7, 8-18, 2019.
[11] A. Daw, A. Karpatne, W. Watkins, J. Read, V. Kumar, Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling. arXiv, 2021.
[12] B. Mildenhall, P.-P. Srinivasan, M. Tancik, J.-T. Barron, R. Ramamoorthi, R. Ng, NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, 2020.