

# LoDoInd: Introducing A Benchmark Low-dose Industrial CT Dataset and Enhancing Denoising with 2.5D Deep Learning Techniques

Jiayang Shi<sup>1</sup>, Omar Elkilany<sup>2</sup>, Andreas Fischer<sup>2</sup>, Alexander Suppes<sup>2</sup>, Daniël M. Pelt<sup>1</sup>, K. Joost Batenburg<sup>1</sup>

<sup>1</sup>LIACS, Leiden University, Niels Bohrweg 1, 2333 CA Leiden, the Netherlands, e-mail: j.shi@liacs.leidenuniv.nl, d.m.pelt@liacs.leidenuniv.nl, k.j.batenburg@liacs.leidenuniv.nl

<sup>2</sup>Waygate Technologies, Albert-Einstein-Straße 22, 31515 Wunstorf, Germany, e-mail:

OmarMohamedYoussef.Elkilany@bakerhughes.com, Andreas.Fischer@bakerhughes.com, alexander.suppes@bakerhughes.com

## Abstract

Computed Tomography (CT) is a widely employed non-destructive testing tool. In industrial applications, minimizing scanning time is crucial for efficient in-line inspection. One approach to achieve faster scanning is through low-dose CT. However, the reduction in radiation dose results in increased noise levels in the reconstructed CT images. Deep learning-based post-processing methods have shown promise in mitigating this noise, but their effectiveness relies on access to datasets with a substantial amount of training data.

Existing low-dose CT datasets either are not specifically tailored for industrial applications or are based on simulated image formation. In this study, we present a new benchmark low-dose CT dataset, LoDoInd, which consists of experimental low-dose CT images explicitly designed for industrial purposes. LoDoInd incorporates complex and diverse secondary filling objects within the same testing object, simulating real-world scenarios encountered in industrial settings. The dataset can be accessed at this Zenodo repository.

Building upon the foundation set by LoDoInd, we further investigate the efficacy of various post-processing methods in denoising tasks. Through a detailed comparative analysis of 2D, 2.5D, and 3D training, we demonstrate that 2.5D training strikes an optimal balance between performance and computational efficiency. This analysis showcases the potential of deep learning in improving the quality of low-dose CT images and also offers valuable insights for enhancing industrial CT applications with practical, efficient AI solutions. The corresponding code is available at this GitHub repository.

**Keywords:** industrial CT dataset, low-dose CT, deep learning, 2.5D training

## 1 Introduction

CT is a crucial imaging technique employed in a wide range of industrial applications, including defect analysis and material characterization [1]. To enable efficient in-line inspection, low-dose CT has emerged as a potential solution due to its reduced scanning time compared to standard CT dose settings. However, low-dose CT often introduces significant noise into the reconstructed images, compromising the accuracy of industrial analysis.

Recent advancements in deep learning techniques have shown promise in reducing noise for low-dose CT [2–6]. However, these methods require datasets that include low-dose CT images alongside corresponding reference images. Existing CT datasets for this purpose fall into two categories. Firstly, some datasets [7–9] simulate low-dose images without involving actual low-dose scanning. Secondly, other datasets [10, 11] are available but are not suitable for industrial CT applications as they primarily focus on medical or non-industrial objects. Consequently, there is a need for a dedicated low-dose CT dataset designed specifically for industrial CT applications. Table 1 presents a comparison between the existing datasets and our proposed dataset.

Dataset	Scanned Low-dose	Multiple Dose Levels	Industrial CT
LowDoChallenge [7], LoDoPaB [8], AppleCT [9]	✗	✗	✗
2DeteCT [10]	✓	✗	✗
ChickenBoneCT [11]	✓	✓	✗
LoDoInd (ours)	✓	✓	✓

Table 1: Comparison of differences between existing datasets and our proposed dataset.

To address these limitations, we present a new low-dose CT dataset specifically tailored for industrial CT. The dataset includes five distinct dose levels and features a pipe object filled with 15 diverse secondary filling objects, such as lentils, pine nuts, and asphalt gravels. This variation enables comprehensive benchmarking. By manipulating the X-ray tube current, we performed scans of the pipe object under different dose levels, simulating realistic scenarios encountered in industrial CT practice.

To demonstrate the utility and versatility of our newly presented low-dose CT dataset, particularly in advancing industrial CT

	Averaged Projs	Exposure Time/ms	Scan Time/min	Voltage/kV	Current/ $\mu$ A
Reference	6	333	59.3	140	180
Noise Level 1	1	333	9.9	140	180
Noise Level 2	1	333	9.9	140	90
Noise Level 3	1	333	9.9	140	45
Noise Level 4	1	333	9.9	140	23
Noise Level 5	1	333	9.9	140	12

Table 2: The acquisition parameters of different noise levels. The main changing acquisition parameters to introduce different noise levels are highlighted.

analysis, we have conducted an illustrative case study focused on deep learning-based noise reduction. This case study, leveraging the unique characteristics of the LoDoInd dataset, explores and compares the effectiveness of 2D, 2.5D, and 3D training modalities as post-processing techniques in low-dose CT image denoising. By employing this dataset, we aim to showcase its potential in facilitating comprehensive analyses of AI-based methods for industrial CT, thus bridging the gap between theoretical advancements and practical industrial applications. Our evaluation includes both quantitative metrics and visual assessments, providing a thorough analysis of the performance of each training modality. This approach not only highlights the capabilities of the dataset but also contributes to the broader understanding of noise reduction techniques in low-dose industrial CT imaging.

## 2 Dataset

### Sample.

For our dataset, a sample was prepared by combining 15 distinct materials, which were carefully mixed and placed inside a plastic tube. The materials included coriander, pine nuts, black cumin, rosemary, lentils, raisins, rice, pepper, flower soil, clay granulate, marble gravel, sand, mulch, and asphalt gravel. While these materials themselves are not typical objects found in industrial CT, they were strategically chosen to reflect the complexity and multi-material nature of items commonly encountered in various industrial CT applications, such as manufacturing, electronic devices, and the food industry [1]. Importantly, this selection focused on introducing noise artifacts while deliberately reducing the potential for other artifacts, such as beam hardening or metal artifacts, thereby ensuring a controlled environment for the study of noise reduction techniques. This approach ensures a diverse range of variations, providing a comprehensive benchmarking capability for our dataset and making it suitable for evaluating deep learning techniques in complex, real-world industrial CT scenarios. An image of all filling material and holding pipe is shown in Figure 1.



Figure 1: The used materials and pipe for the dataset.

facilitate deep learning applications in industrial CT, particularly for adopting low-dose CT techniques to achieve high-throughput analysis. To this end, we are releasing reconstructed volumes for all five noise levels, together with the reference data, each with dimensions of  $4000 \times 1250 \times 1250$ . These volumes have been pre-registered to enable immediate use in benchmarking deep learning methods, thereby assisting in evaluating their practical applicability in industrial settings. Each volume comprises 4000 slices, with individual slices presented as  $1250 \times 1250$  floating point images in Tag Image File Format (TIFF) format. The dataset

**Scanning System.** The used scanning system was a Phoenix VtomeX M300 CT machine with 300 kV microfocus X-ray tube.

**Acquisition Parameters.** For each noise level and reference scan, we obtained a total of 1780 projections using a fixed detector timing of 333 ms per projection. The reference scan involved an average of 6 projection images per scanning angle, whereas for all noise levels, only a single projection image was acquired. To introduce different noise levels, the tube voltage was fixed at 140 kV, and the tube current was varied to create various levels of noise.

**Reconstruction.** The acquired data was reconstructed using Phoenix datoslx software [12]. The resulting volumes were  $4000 \times 1250 \times 1250$  pixels in size. In order to minimize potential positional variations, we performed registration of the reconstructed volumes for all noise levels to the reference volume using Elastix [13]. The registered reconstructed slices for each noise level are illustrated in Figure 2.

**Data Availability.** The LoDoInd dataset is specifically designed to facilitate deep learning applications in industrial CT, particularly for adopting low-dose CT techniques to achieve high-throughput analysis. To this end, we are releasing reconstructed volumes for all five noise levels, together with the reference data, each with dimensions of  $4000 \times 1250 \times 1250$ . These volumes have been pre-registered to enable immediate use in benchmarking deep learning methods, thereby assisting in evaluating their practical applicability in industrial settings. Each volume comprises 4000 slices, with individual slices presented as  $1250 \times 1250$  floating point images in Tag Image File Format (TIFF) format. The dataset

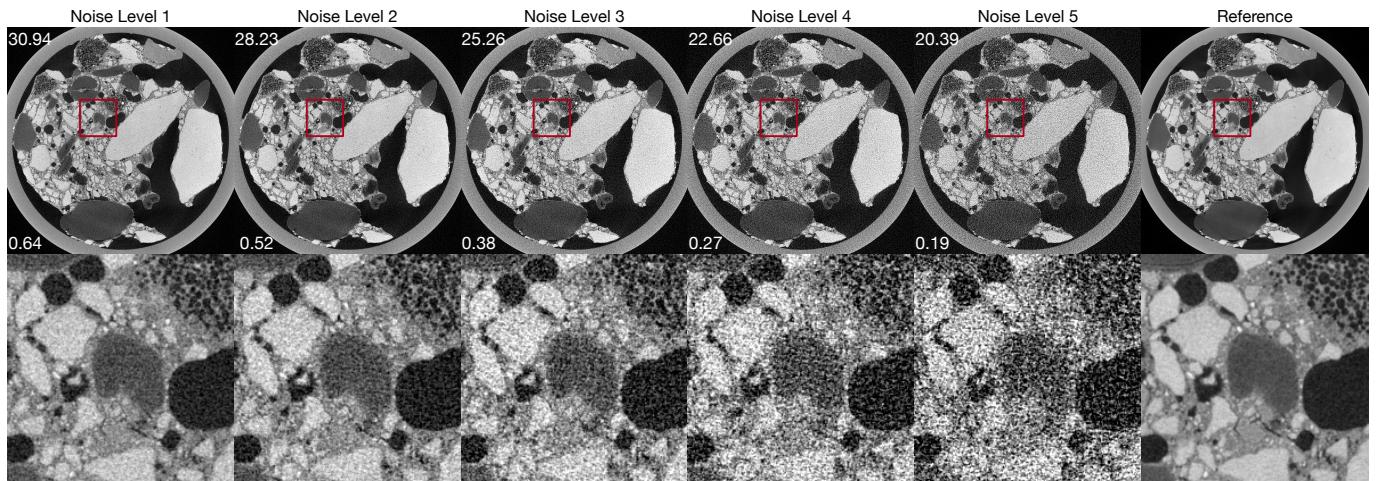


Figure 2: The reconstructed slices of five different noise levels along with the reference slice scanned under normal CT dose. The PSNR and SSIM numbers are given in the upper-left and lower-left corners. The red insets indicate enlarged views.

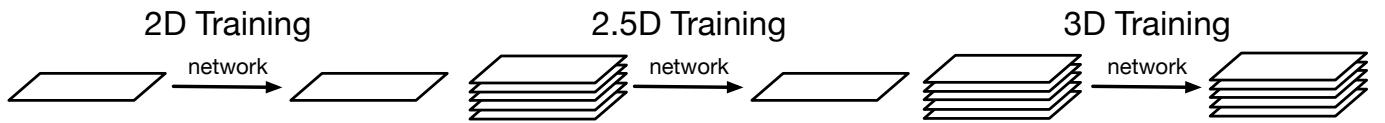


Figure 3: Illustration of the differences between 2D, 2.5D, and 3D training paradigms.

is accessible at [here](#).

### 3 Results

**Preliminary: 2D, 2.5D, and 3D Training.** Denoising low-dose CT images has conventionally been addressed using 2D techniques. In such methods, a 2D neural network processes a noisy slice and produces a corresponding denoised slice [2, 3, 5]. However, given that CT images inherently represent 3D volumes, the use of 3D neural networks, which process and output entire volumes, can harness inter-slice information and consequently enhance image quality [4, 14]. Despite their potential benefits, the primary limitation of 3D approaches lies in their computational intensity [15], which can constrain their practical implementation. To bridge the gap between the 2D simplicity and 3D information richness, 2.5D training was proposed. In this approach, the neural network takes a stack of slices as input but outputs only a single denoised slice. The 2.5D method offers a balance, achieving performance levels comparable to 3D training but with substantially reduced computational requirements [16, 17]. Figure 3 schematically illustrates the distinctions between 2D, 2.5D, and 3D training methods.

To systematically assess the performance of 2D, 2.5D, and 3D training approaches in real-world settings, and to understand the influence of different noise levels on each method's efficacy, we carried out a series of experiments to assess the performance of these three paradigms. The following sections detail the results of these experiments.

**Experiments Details.** Our experiments focused on noise reduction for low-dose CT images, employing supervised learning techniques to map noisy images to their corresponding reference images. For each designated noise level, the dataset was divided such that the top 2000 slices were used for training and the bottom 2000 slices for testing. 25% of the training data was randomly selected as the validation set in each experiment. Consistency across experiments was maintained by setting the number of training epochs to 100 and adopting the same stopping criterion, that if validation loss is not improved from 10 consecutive epochs, then the training stops early.

In these experiments, we opted for two benchmark neural networks: the UNet [18] and the Mixed-Scale Dense (MS-D) network [19]. For the 2D and 2.5D training paradigms, the UNet architecture maintained its originally proposed starting intermediate channel count of 64. However, when adapting to the 3D training framework, we had to reduce the starting intermediate channel count to 16, primarily due to memory constraints. Meanwhile, the MS-D network had a depth of 100 for 2D and 2.5D trainings and depth of 50 for 3D trainings.

For 2.5D training, we adopted two different stack sizes, 3 and 5. We conducted the 3D training both in cubic-like size (400, 320, 320) and in shallow volumes of (16, 1250, 1250). The configurations of these experiments are detailed in Table 3.

#### 3.1 Denoising Performance

**Evaluating Training Modalities: 2D, 2.5D, and 3D.** As shown in Table 4, we assess the denoising effectiveness of various training configurations across different noise levels. In the same neural network architecture, 2.5D training consistently yields

Network Architecture	Type	Config	Trainable Parameters ↓	Mapping
UNet	2D	ch=64	31042369	(1250, 1250) → (1250, 1250)
	3D	ch=16, small	5647857	(400, 320, 320) → (400, 320, 320)
	3D	ch=16, shallow	5647857	(16, 1250, 1250) → (16, 1250, 1250)
	2.5D	ch=64, stack=3	31043521	(3, 1250, 1250) → (1250, 1250)
	2.5D	ch=64, stack=5	31044673	(5, 1250, 1250) → (1250, 1250)
MS-D	2D	dp=100	45652	(1250, 1250) → (1250, 1250)
	3D	dp=50, small	34527	(400, 320, 320) → (400, 320, 320)
	3D	dp=50, shallow	34527	(16, 1250, 1250) → (16, 1250, 1250)
	2.5D	dp=100, stack=3	47454	(3, 1250, 1250) → (1250, 1250)
	2.5D	dp=100, stack=5	49256	(5, 1250, 1250) → (1250, 1250)

Table 3: Training configurations: trainable parameters of neural networks and training setups.

the highest SSIM values, indicating superior structural preservation in denoised images. From noise level 2 through 5, 2.5D training consistently demonstrates improved performance compared to 2D training for both UNet and MS-D networks.

For the MS-D network, 2.5D training achieves the best denoising performance in terms of PSNR and SSIM values over 2D and 3D trainings. Figure 4 compares 2D, 2.5D, and 3D trainings using the MS-D network at noise levels 4 and 5. Consistent with our tabular findings, the 2.5D modality exhibits the highest PSNR and SSIM values, demonstrating its effectiveness in maintaining image fidelity.

Both 2D and 2.5D training modalities reveal sharper image details in the denoised outputs compared to their 3D counterparts. The substantial computational demands of 3D training necessitate a compromise in neural network size, potentially limiting its effectiveness at higher noise levels. This finding underscores the practical limitations of 3D training. Conversely, 2.5D training, which utilizes information from adjacent slices, provides a balanced approach. It captures the benefits of 3D contextual information without the extensive computational cost, proving to be a viable alternative that mitigates the trade-offs associated with full 3D training.

**Contrasting Different Neural Network Architectures.** The analysis reveals a notable performance disparity between the MS-D and UNet networks across all training configurations. Remarkably, the MS-D network outperforms UNet despite having significantly fewer trainable parameters — over 100 times fewer. This comparison is visually represented in Figure 5, where we compare the optimally performing MS-D network (2.5D training with a stack of 5) against UNet (3D training with shallow volumes).

The comparative visualization clearly demonstrates that the MS-D network produces smoother images with enhanced detail clarity, particularly in scenarios with high noise levels. This qualitative assessment is corroborated by quantitative metrics, as evidenced by superior PSNR and SSIM values achieved by the MS-D network. The results underscore the MS-D network's efficiency in handling noise while maintaining image fidelity, thus highlighting its potential for effective denoising applications in challenging conditions.

### 3.2 Computation Cost

This section evaluates the computation costs associated with different neural networks and training modalities. Detailed data is presented in Table 5. To ensure a fair comparison, all evaluations were conducted on the same machine equipped with an RTX 3090 GPU, under identical conditions.

The analysis indicates that 2.5D training incurs a modest increase in memory usage and computation time compared to 2D training. In contrast, 3D training configurations require significantly more memory, even though smaller model sizes are employed. The use of these smaller models in 3D training leads to a reduction in computation time.

From these observations, we infer that while 2.5D training presents a slight increase in computational demands, it offers improved denoising performance, striking a favorable balance between efficiency and effectiveness. On the other hand, fully 3D training poses practical challenges; it often necessitates a trade-off between computational demands and model size. The findings suggest that leveraging 3D information is advantageous for denoising tasks, but this benefit is more efficiently captured through 2.5D training rather than full 3D training. This approach efficiently utilizes the depth information without incurring the significant computational overhead associated with 3D models.

Noise Level 1							
Network	Config	PSNR ↑	SSIM ↑	Network	Config	PSNR ↑	SSIM ↑
UNet	2D	34.96	0.908	UNet	2D	33.69	0.900
	3D-small	30.39	0.876		3D-small	31.42	0.864
	3D-shallow	<b>35.60</b>	0.903		3D-shallow	<b>35.25</b>	0.892
	2.5D-stack=3	34.56	<b>0.910</b>		2.5D-stack=3	34.99	<b>0.904</b>
	2.5D-stack=5	34.44	<b>0.910</b>		2.5D-stack=5	34.67	<b>0.904</b>
MS-D	2D	38.46	0.910	MS-D	2D	37.60	0.902
	3D-small	37.79	0.900		3D-small	36.82	0.892
	3D-shallow	38.13	0.905		3D-shallow	37.17	0.897
	2.5D-stack=3	<b>38.64</b>	<b>0.911</b>		2.5D-stack=3	37.68	<b>0.904</b>
	2.5D-stack=5	38.63	<b>0.911</b>		2.5D-stack=5	<b>37.74</b>	<b>0.904</b>
Noise Level 3							
Network	Config	PSNR ↑	SSIM ↑	Network	Config	PSNR ↑	SSIM ↑
UNet	2D	34.28	0.891	UNet	2D	33.27	0.876
	3D-small	30.75	0.852		3D-small	30.34	0.839
	3D-shallow	<b>35.15</b>	0.882		3D-shallow	<b>34.39</b>	0.873
	2.5D-stack=3	34.99	<b>0.904</b>		2.5D-stack=3	33.67	0.882
	2.5D-stack=5	34.67	<b>0.904</b>		2.5D-stack=5	34.12	<b>0.884</b>
MS-D	2D	36.64	0.891	MS-D	2D	35.65	0.877
	3D-small	36.16	0.881		3D-small	35.12	0.866
	3D-shallow	36.30	0.884		3D-shallow	35.34	0.873
	2.5D-stack=3	36.85	0.893		2.5D-stack=3	35.97	0.881
	2.5D-stack=5	<b>36.90</b>	<b>0.894</b>		2.5D-stack=5	<b>35.98</b>	0.882
Noise Level 5							
Network	Config	PSNR ↑	SSIM ↑				
UNet	2D	32.78	0.860				
	3D-small	30.46	0.821				
	3D-shallow	<b>33.46</b>	0.860				
	2.5D-stack=3	32.43	<b>0.863</b>				
	2.5D-stack=5	32.44	0.862				
MS-D	2D	34.84	0.864				
	3D-small	34.23	0.852				
	3D-shallow	34.61	0.861				
	2.5D-stack=3	35.04	0.868				
	2.5D-stack=5	<b>35.17</b>	<b>0.870</b>				

Table 4: Comparative Analysis of Denoising Performance Across Training Configurations. The highest PSNR/SSIM values within each neural network architecture are highlighted, while the top PSNR/SSIM values across all architectures for the same noise level are bolded.

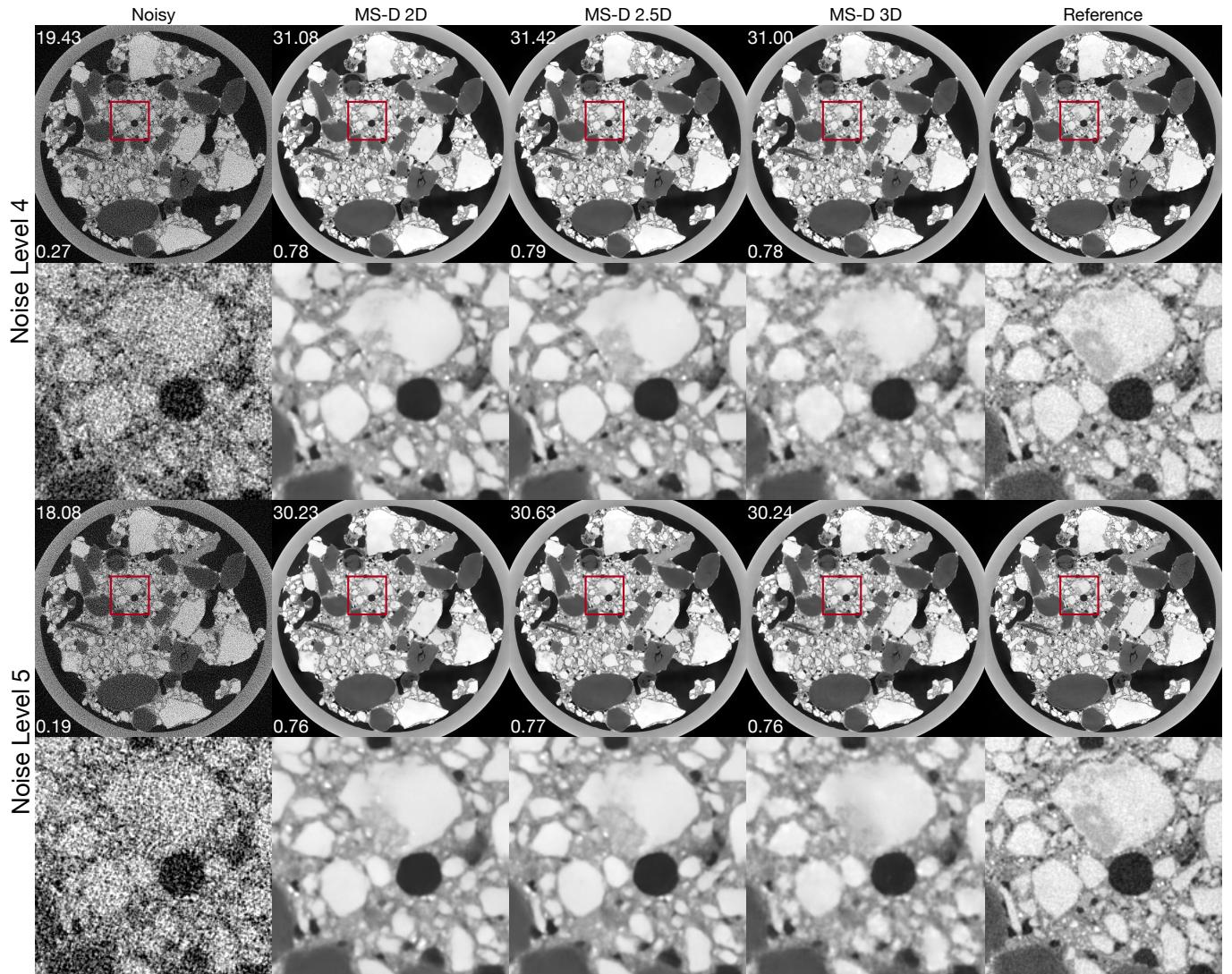


Figure 4: The results of deep learning-based post-processing applied to our dataset using 2D, 2.5D and 3D MS-D network. The PSNR and SSIM numbers are given in the upper-left and lower-left corners. The red insets indicate enlarged views.

Network Architecture	Type	Config	Memory/MiB ↓	Time/hrs:mins ↓
UNet	2D	ch=64	6016	08:49
	3D	ch=16, small	27454*	—*
	3D	ch=16, shallow	18382	02:49
	2.5D	ch=64, stack=3	6036	10:29
	2.5D	ch=64, stack=5	6060	10:30
MS-D	2D	dp=100	2608	06:54
	3D	dp=50, small	17772	03:16
	3D	dp=50, shallow	11856	03:00
	2.5D	dp=100, stack=3	2672	07:31
	2.5D	dp=100, stack=5	2674	07:30

Table 5: Computation Costs for Various Training Approaches: Memory Usage and Computation Time Analysis. These metrics were evaluated on an RTX 3090 GPU with 24,576 MiB memory. The computation time reflects the duration required to process an entire volume of dimensions (2000, 1250, 1250). \*Note that the 3D UNet training with (400, 320, 320) volumes exceeded the memory capacity of the RTX 3090 GPU; thus, measurements for this setup were conducted on a cluster equipped with an A100 GPU. However, isolated computation time analysis on the A100 platform was not feasible.

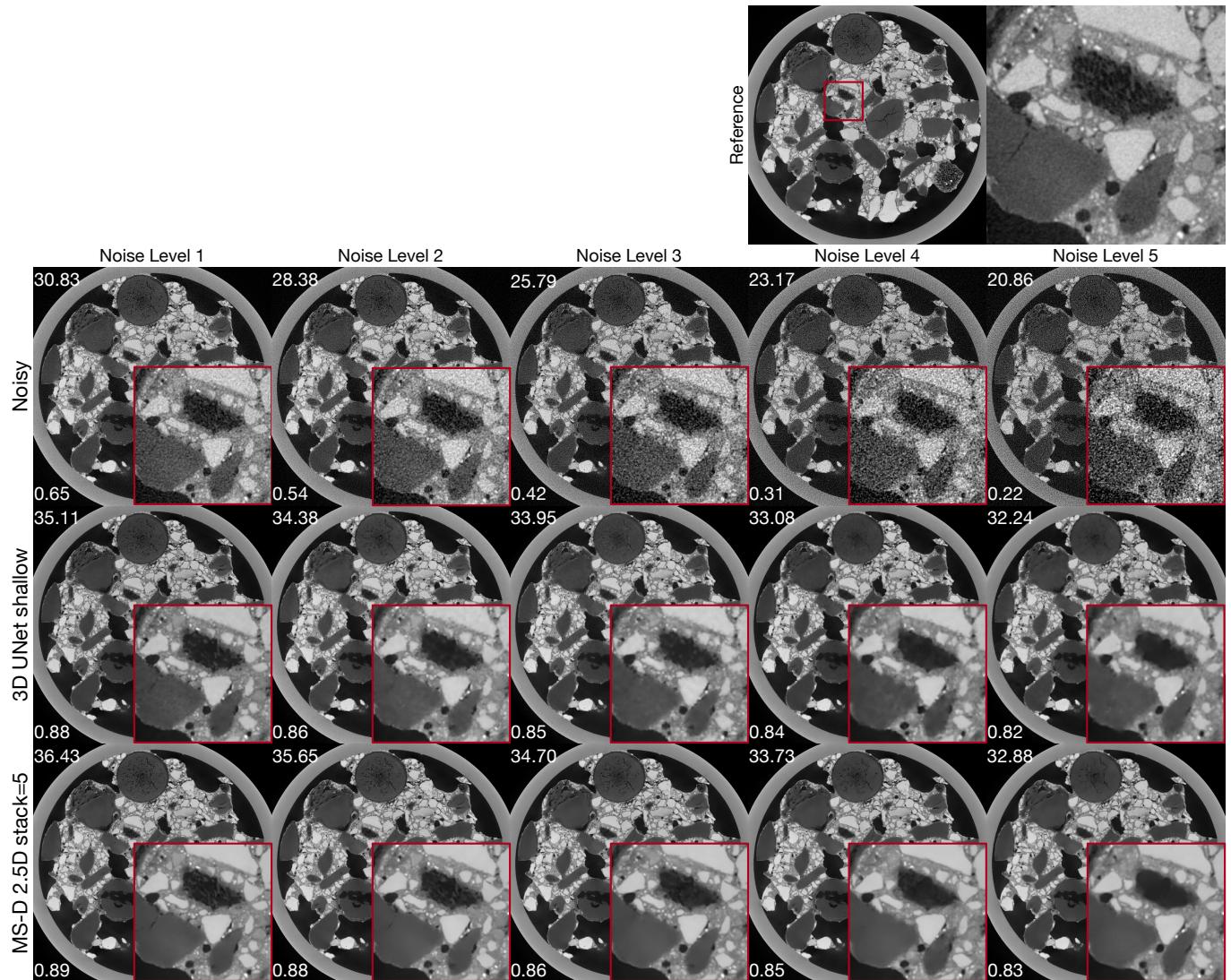


Figure 5: Comparative Analysis of UNet and MS-D Networks: Highlighting the Optimal Performances of 3D Training with UNet on Shallow Volumes and 2.5D Training with a Stack of 5 Slices. Enlarged details are shown in the red insets at the bottom-right corners, with PSNR and SSIM values indicated in the top-left and bottom-left corners, respectively.

## 4 Conclusion

In this study, we introduced LoDoInd, a benchmark dataset tailored for low-dose CT imaging in industrial applications. This dataset represents a significant step forward in addressing the specific challenges of industrial CT analysis. Additionally, our research encompassed a comprehensive case study to evaluate various training modalities and neural network architectures in the context of denoising.

Our findings reveal that incorporating 3D information is beneficial for denoising tasks, as it adds valuable depth details. However, we observed that traditional 3D training is often constrained by a trade-off between computational costs and network capability. This limitation is particularly pronounced when dealing with high-resolution industrial CT data. In contrast, our results demonstrate that 2.5D training strikes a balance, offering enhanced denoising performance without the substantial computational burden associated with full 3D training.

This study not only provides a valuable dataset for future research in industrial CT image analysis but also establishes a clear direction for the development of efficient and effective denoising techniques, highlighting the potential of 2.5D training as a viable solution for handling complex, real-world industrial imaging challenges.

## Acknowledgements

This research was co-financed by the European Union H2020-MSCA-ITN-2020 under grant agreement no. 956172 (xCTing).

## References

- [1] L. De Chiffre, S. Carmignato, J.-P. Kruth, R. Schmitt, A. Weckenmann, Industrial applications of computed tomography, *CIRP annals* 63 (2) (2014) 655–677.
- [2] H. Chen, Y. Zhang, W. Zhang, P. Liao, K. Li, J. Zhou, G. Wang, Low-dose ct via convolutional neural network, *Biomedical optics express* 8 (2) (2017) 679–694.
- [3] D. M. Pelt, K. J. Batenburg, J. A. Sethian, Improving tomographic reconstruction from limited data using mixed-scale dense convolutional neural networks, *Journal of Imaging* 4 (11) (2018) 128.
- [4] M. Li, W. Hsu, X. Xie, J. Cong, W. Gao, Sacnn: Self-attention convolutional neural network for low-dose ct denoising with self-supervised perceptual loss network, *IEEE transactions on medical imaging* 39 (7) (2020) 2289–2301.
- [5] J. Gu, J. C. Ye, Adain-based tunable cyclegan for efficient unsupervised low-dose ct denoising, *IEEE Transactions on Computational Imaging* 7 (2021) 73–85.
- [6] J. Shi, D. M. Pelt, K. J. Batenburg, Multi-stage deep learning artifact reduction for computed tomography, *arXiv preprint arXiv:2309.00494* (2023).
- [7] C. McCollough, Tu-fg-207a-04: overview of the low dose ct grand challenge, *Medical physics* 43 (6Part35) (2016) 3759–3760.
- [8] J. Leuschner, M. Schmidt, D. O. Baguer, P. Maass, Lodopab-ct, a benchmark dataset for low-dose computed tomography reconstruction, *Scientific Data* 8 (1) (2021) 109.
- [9] S. B. Coban, V. Andriashen, P. S. Ganguly, M. van Eijnatten, K. J. Batenburg, Parallel-beam x-ray ct datasets of apples with internal defects and label balancing for machine learning, *arXiv preprint arXiv:2012.13346* (2020).
- [10] M. B. Kiss, S. B. Coban, K. J. Batenburg, T. van Leeuwen, F. Lucka, 2detect—a large 2d expandable, trainable, experimental computed tomography dataset for machine learning, *arXiv preprint arXiv:2306.05907* (2023).
- [11] A. Meaney, Cone-beam computed tomography dataset of a chicken bone imaged at 4 different dose levels (Aug. 2022). doi:10.5281/zenodo.6990764.  
URL <https://doi.org/10.5281/zenodo.6990764>
- [12] Phoenix datoslx software, <https://www.bakerhughes.com/waygate-technologies/ndt-software/phoenix-datosx-industrial-ct-scanning-software>, accessed: 2023-07-12.
- [13] S. Klein, M. Staring, K. Murphy, M. A. Viergever, J. P. Pluim, Elastix: a toolbox for intensity-based medical image registration, *IEEE transactions on medical imaging* 29 (1) (2009) 196–205.
- [14] H. Shan, Y. Zhang, Q. Yang, U. Kruger, M. K. Kalra, L. Sun, W. Cong, G. Wang, 3-d convolutional encoder-decoder network for low-dose ct via transfer learning from a 2-d trained network, *IEEE transactions on medical imaging* 37 (6) (2018) 1522–1534.
- [15] J. Leuschner, M. Schmidt, P. S. Ganguly, V. Andriashen, S. B. Coban, A. Denker, D. Bauer, A. Hadjifaradji, K. J. Batenburg, P. Maass, et al., Quantitative comparison of deep learning-based image reconstruction methods for low-dose and sparse-angle ct applications, *Journal of Imaging* 7 (3) (2021) 44.
- [16] A. Ziabari, D. H. Ye, S. Srivastava, K. D. Sauer, J.-B. Thibault, C. A. Bouman, 2.5 d deep learning for ct image reconstruction using a multi-gpu implementation, in: 2018 52nd Asilomar Conference on Signals, Systems, and Computers, IEEE, 2018, pp. 2044–2049.
- [17] A. A. Hendriksen, M. Bührer, L. Leone, M. Merlini, N. Vigano, D. M. Pelt, F. Marone, M. Di Michiel, K. J. Batenburg, Deep denoising for multi-dimensional synchrotron x-ray tomography without high-quality reference data, *Scientific reports* 11 (1) (2021) 11895.

- [18] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, Springer, 2015, pp. 234–241.
- [19] D. M. Pelt, J. A. Sethian, A mixed-scale dense convolutional neural network for image analysis, *Proceedings of the National Academy of Sciences* 115 (2) (2018) 254–259.