Balance Between the Number of Projections and Exposure Time in Photon Counting CT with a Data-Driven Approach

Raziye Kubra Kumrular [1,*] and Thomas Blumensath [1]

[1]Institute of Sound and Vibration Research, The University of Southampton, SO17 1BJ, U.K.,
* corresponding author, e-mail: r.k.kumrular@soton.ac.uk

ABSTRACT

Photon Counting Computed Tomography (CT) Imaging often requires high X-ray exposure due to low photon counts per channel, leading to prolonged scanning times, which may not be practical. Here, we explored a trade-off between the number of projections and exposure time per projection to optimize scanning efficiency and image quality. By increasing projections and reducing exposure, we initially generated noisier datasets, which we denoised using unsupervised, data-driven techniques. Extending our previous research, we applied unsupervised denoising to synthetic spectral CT datasets with a distinct K-edge in the X-ray absorption spectrum. We compared our results with an iterative reconstruction algorithm that uses a total variation constraint in the spatial and a total generalised variation constraint in the spectral dimension, which uses fewer projections and higher doses for each projection. Although this algorithm employs fewer projections and higher doses, it matches our method in scanning time, allowing a direct comparison to our methods. Our approach significantly reduced scanning time by 36-fold compared to traditional full-dose methods, without compromising image quality. It also eliminates the need for meticulous parameter tuning, simplifying the operational process and enhancing usability.

Keywords: Unsupervised Denoising, Photon Counting CT Imaging, Low Dose

1. INTRODUCTION

Technological advances in Photon Counting Detectors (PCD) have significantly facilitated energy-sensitive imaging applications by counting individual X-ray photons and resolving their energies.[1] This innovation has expanded the boundaries of imaging technology, making it possible for higher spatial resolution,[1–3] accurate material characterization,[4] radiation dose reduction,[1,4] reduced beam hardening[5] and contrast agent enhancement.[1] As a natural extension of these developments, PCD-Computed Tomography (CT) is expected to go beyond current Dual-Energy CT as the next generation of X-ray imaging.[2]

The trade-off between resolution, noise, and the number of projections is a key consideration in tomographic imaging. Increasing the number of projections with shorter exposure times can enhance spatial resolution and detail, albeit may increase the noise in each projection.[6] Conversely, fewer projections with longer exposure times could offer lower noise for projections but may compromise spatial resolution and detail, and make the process more time-consuming.[7,8] However, modelling the noise and reducing scanning time in the first scenario is found to be more efficient and cost-effective than in the second scenario. In this paper, we seek the optimization of this balance in PCD-CT with data-driven and iterative methods.

We introduce a self-supervised deep learning method Noise2Inverse[9] approach with block-based training[10–12] for PCD-CT Imaging, which can be trained on noisy data alone. This paper examines the denoising capability of data-driven methods to produce high-resolution images from 4D noisy spectral tomography datasets, particularly focusing on materials that are highly attenuated and have unique K-edge discontinuities in the energy spectrum (as accurate detection of K-edges is critical for material decomposition). Findings on numerical datasets demonstrate that our approach may effectively reduce measurement noise and save a significant amount of acquisition time without compromising image quality.

2. METHODOLOGY

2.1 Photon Noise

The fundamental uncertainty in the measurement of light is described by photon noise, also called Poisson noise-quantum noise. It is caused by the quantized nature of light and independent photon detection's.[13] In short, photon counting follows a Poisson process, with independent photon detection described by the probability distribution:

$$P(N=k)=\frac{e^{-\lambda t}(\lambda t)^{k}}{k!}, \tag{1}$$

where N denotes the number of photons measured by the detector and λ represents the expected photons per pixel in the detector over the time interval t. As the value of N increases, σ increases by the root square of the signal, so the signal-to-noise ratio increases in proportion to the N factor.[7,8,13] Increasing the X-ray flux or extending the exposure time (and dose) is one of the methods used to enhance image quality.[7,8]

2.2 Measurement Model in PCD-CT

The Beer-Lambert law describes X-ray interaction with materials, which depends on the energy of the X-ray beam and the attenuation coefficient of the material in the scanned object. The general form of Beer-Lambert law for a poly-chromatic X-ray beam is given by:

$$I_i(\mathcal{N})=\int_{\mathcal{N}} I_0(E)e^{-\int_L \mu(E,r)dr}dE \tag{2}$$

where $I_i(E)$ is measured X-ray intensity (number of photons through the object) recorded by a photon counting detector in the i^{th} energy bin and $I_0(E)$ denotes the incident photon fluence from the X-ray source, at energy level E. Term $\mu\left(E,r\right)$ is the linear attenuation coefficient of the materials at energy E and $\int_L \mu\left(E,r\right)dr$ represents the line integral of attenuation along one ray path from the source to one detector element at one rotation angle and $\mathcal{N}$ is the energy range. Converting the Beer-Lambert law from intensity to absorption and discretizing the integral, we have a linear system as below:

$$\tilde{y_i}=Ax_i+\epsilon_i, \tag{3}$$

where A denotes the projection matrix. Term ϵ_i is the noise and $\tilde{y_i}$ is the noisy measurement in the i^{th} bin. The noise level is related to the square root of the total number of detected photons as discussed above. Specifically, PCD-CT suffers from stronger Poisson noise for narrow bins, as the number of detected photons is less than the total number of photons.[1,2,13]

2.3 Self-Supervised Denoising Method

Since PCD-CT suffers from strong noise, self-supervised methods that do not require prior knowledge of the signal and noise become particularly attractive. These methods first train a model using the noisy images, and subsequently, this trained model is used to predict clean images from the noisy ones. Most importantly, the main difference between these models is based on the image sets in the training phase. For example, methods like Noise2Noise[14] leverage multiple images of the same target when clean pairs are unavailable, although this is not applicable in PCD-CT. For a single noisy image with pixel-wise-independent noise, Noise2Self[15] is a viable option. More recently, Noise2Inverse[9] (a version of Noise2Noise and Noise2Self) has been proposed for tomographic imaging. In this approach, noise is independent and zero-mean in the measurement domain, and by using reconstructions from independent subsets, noise becomes independent in the image domain.

2.4 Regularized Image Reconstruction

A traditional alternative approach is to use regularised iterative algorithms such as:

$$\arg\min_x \left\{\frac{1}{2}\|Ax-\tilde{y}\|_2^2+\alpha\mathrm{TV}(x)+\beta_{1,2}\mathrm{TGV}(x)\right\} \tag{4}$$

where $f(x)=\frac{1}{2}\|Ax-\tilde{y}\|_2^2$ is a data fidelity term that measures the distance between the acquired data and the forward-projected reconstructed image, and $g(x)=\alpha\mathrm{TV}(x)+\beta_{1,2}\mathrm{TGV}(x)$ represent the regularization terms that encode smoothness properties for x. Iterative algorithms can solve problem $f(x)$, and the Core Imaging Library (CIL)[16] provides the primal-dual hybrid gradient for this purpose. Here we use the total variation (TV) and total generalised variation (TGV) regularization for $g(x)$, with TV employed for spatial regularization with parameter α, and TGV used for energy regularization with parameters $\beta_{1,2}$.[16] Iterative methods reduce noise and improve image quality, but they are computationally expensive and require dataset-specific parameter tuning.

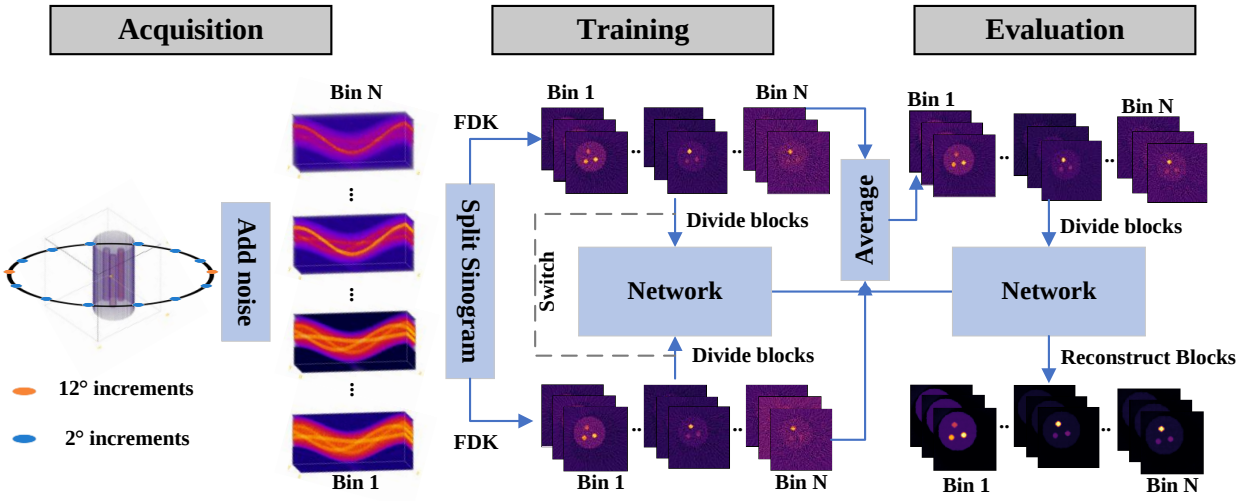

Figure 1: Data-driven approach: The spectral sinogram is obtained over 360 degrees (with 2∘ and 12°) and split into 2 mutually exclusive sets, which are reconstructed independently for each energy channel using FDK. During training, the subsets alternated as input and target to increase training data diversity. The final denoised images were generated by averaging the inputs used during training and applying the trained network.

2.5 Our approach

The Noise2Inverse requires us to generate input and target images with independent noise. To achieve these requirements, we followed the approach illustrated in Fig1. First, we acquired sets of projections over an angular range of 360° with 2° and 12° increments, which were then corrupted with Poisson noise using the relevant source spectrum. After adding the noise, these noisy projections were divided into two subsets, $\tilde{y}_{E_1,1}$ and $\tilde{y}_{E_N,2}$, each containing mutually exclusive projections at equally spaced angles for the same energy channels. After splitting these noisy sinograms, each subset was reconstructed using energy channel-wise FDK,[17] $\tilde{x}_{E_1,1},\ldots,\tilde{x}_{E_N,1}$ and $\tilde{x}_{E_1,2},\ldots,\tilde{x}_{E_N,2}$. In the training stage, reconstructions of the first subset served as the input, and the corresponding target was the image from the second subset. Then, we switched roles: reconstructions of the second subset became the input, and the target was the image from the first subset. This strategy allows robust training by increasing the training data for the model. To generate the final denoised images $x^*_{E_1},\ldots,x^*_{E_N}$, all inputs used in training are averaged and used as input for the trained network.

3. EXPERIMENTS

3.1 Spectral Data Preparation, Acquisition and Reconstruction

A synthetic phantom (Fig.2) was prepared to mimic properties of the real-world datasets,[18,19] closely replicating the physical phantom. The preparation process is detailed in,[10,11] initially involved 2D Spatial and 1D Energy dimensions, but here it was modified to 4D (3D Spatial + 1D Energy) with materials (Fe, Al, ZnO, CeO_2). The left image in Fig.2 illustrates the 3D spatial structure while the right one shows a 2D slice of this 3D structure (in x, y coordinates), enriched by adding the energy (z) dimension as the 3^{rd}. The impact of the K-edge (as a unique marker for material identification) in the energy spectrum is clearly shown.

The synthetic data used in the experiments were acquired with exposure times of 30 and 5 seconds per projection. The process of generating a poly-chromatic X-ray source is detailed in,[10,11] with the main difference being changes in tube voltage and energy bins. To closely mimic real experimental conditions,[18,19] the generated X-ray source spectrum (using SpekPy[20]), produced with a tube voltage of 60 kVp, was interpolated into 200 energy bins. A Tungsten target at a 12° angle was employed, and the X-ray beam was filtered through 0.4 mm Aluminum, 1.5 mm Beryllium, and 1000 mm Air. These parameters were chosen empirically based on visual comparisons with real X-ray source Spectrum.[19] The second column in Fig.3 shows the number of photons detected in a random pixel location over all the energy channels, giving an idea of the source spectrum used.

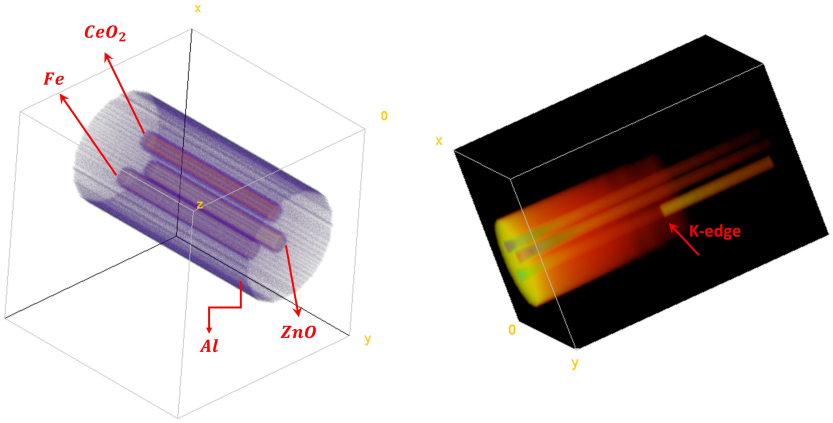

Figure 2: Spatial (left) and Spectral (right) visualization. 3D spatial structure of the simulated phantom shows 4 materials which are CeO_2 (cerium oxide), ZnO (zinc oxide), Fe (iron) and Al (aluminium). K-edge is shown in the red arrow in the energy profile over the z-axis.

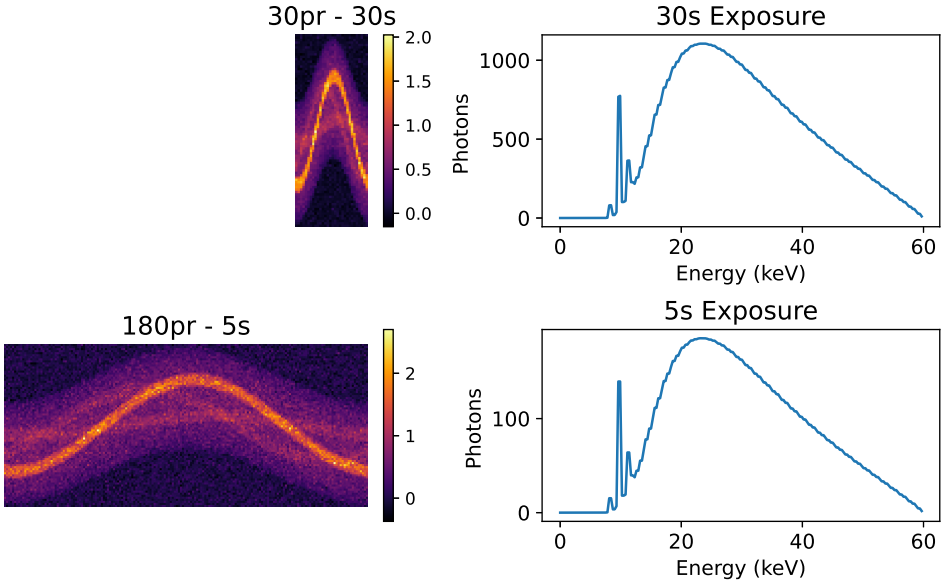

Figure 3: Comparison of exposure time and projection numbers. First column: projections (30–projections and 180–projections). Second column: photons detected in a random pixel (30–second exposure and 5–second exposure per projection. Scanning time is equal in rows.

To acquire the datasets, we created 4D sinograms $(y_{E_1}, y_{E_2}, ..., y_{E_N})$ using 4D spectral phantoms and the source spectrum we generated, employing the same acquisition geometry as in the real data.[18] The full angular range of 360° was covered with both 2° and 12° increments (the first column in Fig.3), and the sinograms were then corrupted with Poisson noise using the relevant X-ray source spectrum $I_0(E)$ discussed above. If $y^p(E)$ is the simulated clean X-ray attenuation value for one pixel, then the noisy pixel $\tilde{y}^p(E)$ for that energy is distributed as:

$$I_0(E)e^{-\tilde{y}^p(E)} \sim \text{Poisson}\left(I_0(E)e^{-y^p(E)}\right), \tag{5}$$

All noisy projections were split into 2 sets and each of them was reconstructed with the FDK for our training strategy.[17] For comparison of our method, we also employed a traditional iterative reconstruction method that imposes a

TV constraint in the spatial and TGV constraint in the spectral dimension and used all projections for both acquisition scenarios. Regularization parameters α, and $\beta_{1,2}$ were set to 0.002, 0.18 and 0.25, respectively, based on previous studies that identified these values as yielding the best reconstruction performance in real experimental data.[19] We run the method for 500 iterations per acquisition.

3.2 Network Implementation and Quality Assessment

We utilized the 3D U-Net architecture from,[21] implemented using PyTorch, which remains state of the art in many biomedical image-denoising applications. This multi-channel input layer is tailored for the specific requirements of spectral image denoising. The network consists of 5 layers and takes $4x16x16x16$, where 4 is the energy dimension, as input and target. Overlap between blocks is 50% in spatial dimension and 75% in spectral dimension. Adam[22] optimizer with a learning rate of 10^{-4} was used to update the network weights. The mini-batch size was 64 and the employed loss function is the mean squared error.

The quality of the denoised images was assessed against the ground truth using the structural similarity index measure (SSIM)[23] and peak signal-to-noise ratio (PSNR)[24] metrics applied channel-wise to the entire image. We further analyzed the overall image quality by computing the mean and standard deviation of SSIM and PSNR metrics in the energy direction (see Table 1). This will help quantify the impact of our experiments across energy channels.

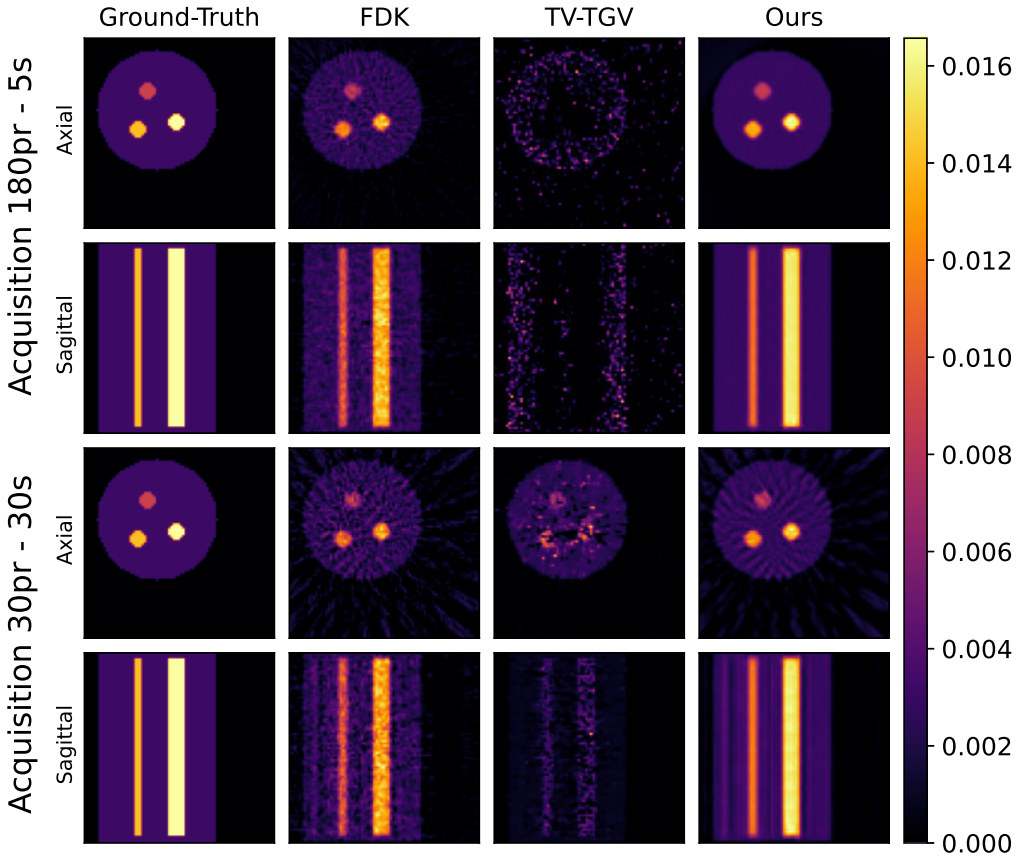

Figure 4: Comparison of methods at 28 keV for different acquisition strategies.

4. RESULTS AND DISCUSSION

Figures 4-5 show axial and sagittal slices of the images for different methods and different acquisition protocols at 28 keV and 42 keV. Each column in the figures represents a different method, including Ground-Truth in the first column, FDK in the second column, TV-TGV in the third column, and Ours in the fourth column. The colour bar on the right indicates the intensity values. Our method had a better performance, especially for high noise (i.e. low photon count) energy channels (the low and high energy channels, where the source spectrum has limited flux), though the average

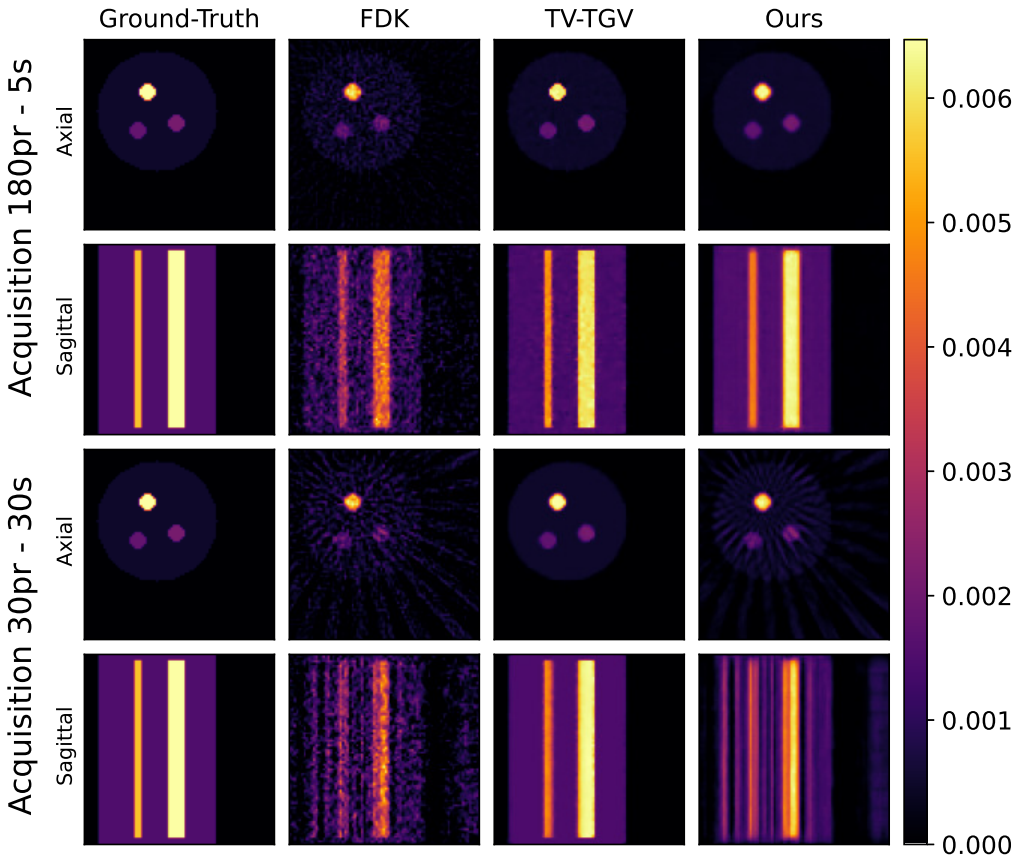

Figure 5: Comparison of methods at 42 keV for different acquisition strategies.

PSNR performance was found to be still better for the TV-TGV for both acquisition setup. Moreover, even though we employed the same regularization parameters, they did not perform consistently for both acquisitions. This inconsistency highlights a well-known drawback of iterative methods: the time-consuming process of selecting optimal regularization parameters for each specific dataset.

Crucially, when the number of projections is decreased and the exposure time is increased, our method struggles to effectively remove the artefacts resulting from under-sampling (the last two rows in Figs.4-5). While our approach is efficient the dealing with measurement noise, it is less effective against under-sampling artefacts. Therefore, it is well suited to optimize the trade-off between noise and artefacts by increasing the number of projections and reducing the exposure time, which is possible in tomographic imaging.[6]

Table 1: PSNR and SSIM results (Mean ± SD).

Method	PSNR (dB)	SSIM
Acquisition: 180pr - 5s		
FDK	23.32 ± 3.98	0.37 ± 0.15
TV-TGV	**33.01 ± 11.75**	0.78 ± 0.27
Ours	30.37 ± 3.67	**0.88 ± 0.07**
Acquisition: 30pr - 30s		
FDK	19.85 ± 3.04	0.22 ± 0.08
TV-TGV	**31.65 ± 8.72**	**0.84 ± 0.24**
Ours	26.98 ± 1.61	0.46 ± 0.04

To evaluate the denoising performance across the energy channels, the regions of interest (ROIs) were selected within the objects of interest and analysis was performed on two ROIs, one of them containing K-edge for each acquisition setup. The attenuation profiles of two different materials were evaluated as illustrated in Fig6. The attenuation profile was better preserved over the energy channels with our approach compared to the TV-TGV method in terms of noise. Interestingly, our approach yielded superior performance in terms of average SSIM for the first acquisition scenario but the average PSNR of our method was inferior to that achieved by the TV-TGV. However, the high standard deviation of 11.75 in the PSNR values for the TV-TGV method indicates significant variability across the energy axis, which may contribute to the less consistent performance compared to our approach (as seen in Tab.1).

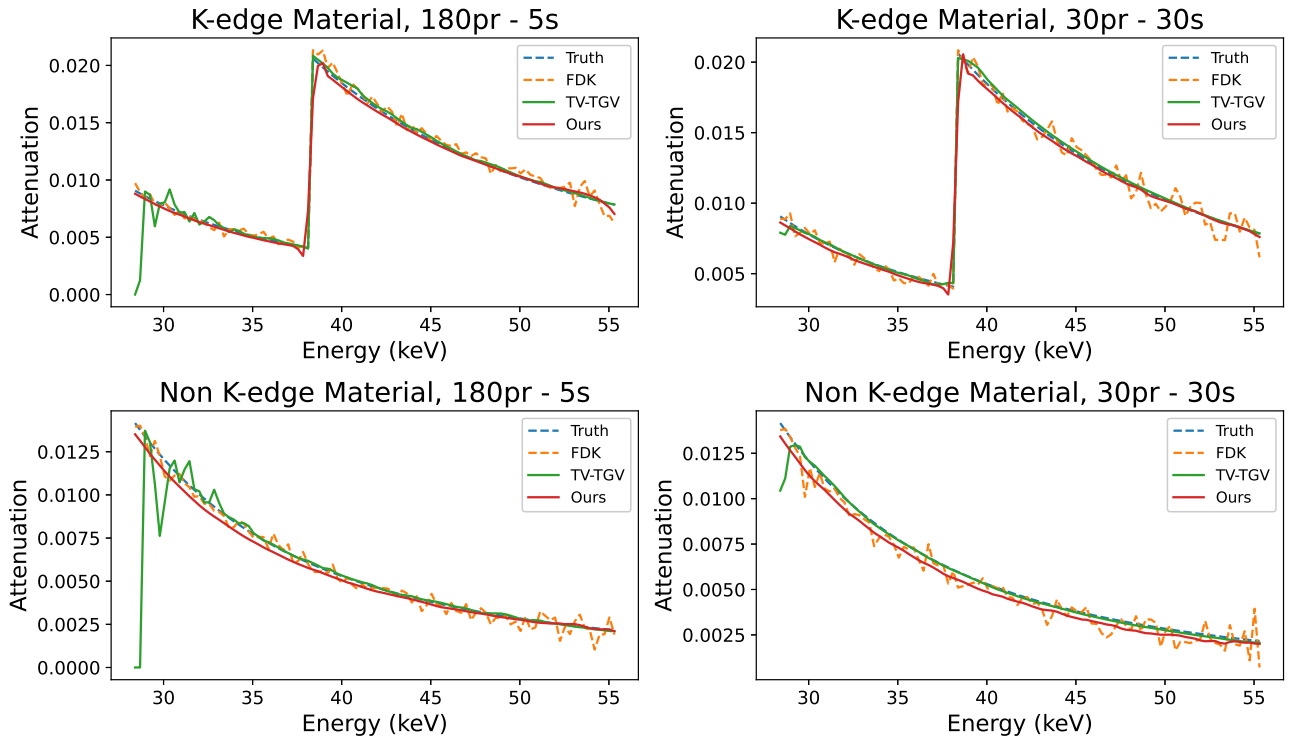

Figure 6: Attenuation profiles of the selected ROIs for both acquisition strategies. The upper line represents K-edge materials, while the bottom line corresponds to non-K-edge materials.

5. CONCLUSION

In this study, we have demonstrated the application of Noise2Inverse[9] to spectral imaging using a block-based training approach[10–12] with a 3D U-Net,[21] showing that our approach in spectral imaging does offer a significant improvement in image quality. This was achieved without needing to fine-tune regularisation parameters which is a drawback of traditional iterative approaches. A 36–fold reduction in the scanning procedure compared to full-dose methods has been demonstrated for the synthetic phantom. However, this reduction introduces challenges, as faster scanning increases Poisson noise or streak artefacts depending on the acquisition strategy. Future work will focus on validating the learned spectral denoising techniques on real-world experimental data and assessing their performance in faster acquisition strategies by comparing them with other advanced denoising methods.

6. ACKNOWLEDGEMENTS

This research was funded by the Republic of Turkiye Ministry of National Education.

REFERENCES

[1] Hsieh, S. S., Leng, S., Rajendran, K., Tao, S., and McCollough, C. H., "Photon counting ct: clinical applications and future developments," *IEEE Transactions on Radiation and Plasma Medical Sciences* **5**(4), 441–452 (2020).

[2] Taguchi, K., Blevis, I., and Iniewski, K., [*Spectral, Photon Counting Computed Tomography: Technology and Applications*], CRC Press, London, UK (2022).

[3] van der Bie, J., van Straten, M., Booij, R., Bos, D., Dijkshoorn, M. L., Hirsch, A., Sharma, S. P., Oei, E. H., and Budde, R. P., "Photon-counting ct: review of initial clinical results," *European Journal of Radiology* , 110829 (2023).

[4] Long, Y. and Fessler, J. A., "Multi-material decomposition using statistical image reconstruction for spectral ct," *IEEE Transactions on medical imaging* **33**(8), 1614–1626 (2014).

[5] Wang, A. S. and Pelc, N. J., "Spectral photon counting ct: Imaging algorithms and performance assessment," *IEEE transactions on radiation and plasma medical sciences* **5**(4), 453–464 (2020).

[6] McCollough, C. H., Bartley, A. C., Carter, R. E., Chen, B., Drees, T. A., Edwards, P., Holmes III, D. R., Huang, A. E., Khan, F., Leng, S., et al., "Low-dose ct for the detection and classification of metastatic liver lesions: results of the 2016 low dose ct grand challenge," *Medical physics* **44**(10), e339–e352 (2017).

[7] Hendee, W. R. and Ritenour, E. R., [*Medical Imaging Physics*], John Wiley & Sons, New York, NY, USA (2003).

[8] Smith, N. B. and Webb, A., [*Introduction to medical imaging: physics, engineering and clinical applications*], Cambridge university press (2010).

[9] Hendriksen, A. A., Pelt, D. M., and Batenburg, K. J., "Noise2inverse: Self-supervised deep convolutional denoising for tomography," *IEEE Transactions on Computational Imaging* **6**, 1320–1335 (2020).

[10] Kumrular, R. K. and Blumensath, T., "Unsupervised denoising for spectral ct images using a u-net with block-based training," in [*Anomaly Detection and Imaging with X-Rays (ADIX) IX*], **13043**, 23–32, SPIE (2024).

[11] Kumrular, R. K. and Blumensath, T., "Unsupervised denoising in spectral ct: Multi-dimensional u-net for energy channel regularisation," *Sensors* **24**(20), 6654 (2024).

[12] Blumensath, T. and Harrane, I., "Machine learning based 3d xct image enhancement for the inspection of am metal components using limited x-ray measurements," *Additive Manufacturing* (2023).

[13] Hasinoff, S. W., "Photon, poisson noise," in [*Computer vision: a reference guide*], 980–982, Springer (2021).

[14] Lehtinen, J., Munkberg, J., Hasselgren, J., Laine, S., Karras, T., Aittala, M., and Aila, T., "Noise2noise: Learning image restoration without clean data," *arXiv preprint arXiv:1803.04189* (2018).

[15] Batson, J. and Royer, L., "Noise2self: Blind denoising by self-supervision," in [*International Conference on Machine Learning*], 524–533, PMLR (2019).

[16] Papoutsellis, E., Ametova, E., Delplancke, C., Fardell, G., Jørgensen, J., Pasca, E., Turner, M., Warr, R., Lionheart, W., and Withers, P., "Core imaging library-part ii: multichannel reconstruction for dynamic and spectral tomography," *Philosophical Transactions of the Royal Society A* **379**(2204), 20200193 (2021).

[17] Jørgensen, J. S., Ametova, E., Burca, G., Fardell, G., Papoutsellis, E., Pasca, E., Thielemans, K., Turner, M., Warr, R., Lionheart, W. R., et al., "Core imaging library-part i: a versatile python framework for tomographic imaging," *Philosophical Transactions of the Royal Society A* **379**(2204), 20200192 (2021).

[18] Warr, R., Jørgensen, J., Papoutsellis, E., Ametova, E., Cernik, R., and Withers, P., "Hyperspectral x-ray ct datasets of an aluminium phantom containing three metal-based powders." https://doi.org/10.5281/zenodo.5825464 (2022). Version 3.

[19] Warr, R., Ametova, E., Cernik, R. J., Fardell, G., Handschuh, S., Jørgensen, J. S., Papoutsellis, E., Pasca, E., and Withers, P. J., "Enhanced hyperspectral tomography for bioimaging by spatiospectral reconstruction," *Scientific Reports* **11**(1), 20818 (2021).

[20] Bujila, R., Omar, A., and Poludniowski, G., "A validation of spekpy: A software toolkit for modelling x-ray tube spectra," *Physica Medica* **75**, 44–54 (2020).

[21] Ronneberger, O., Fischer, P., and Brox, T., "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*], 234–241, Springer (2015).

[22] Kingma, D. and Ba, J., "Adam: A method for stochastic optimization," in [*Proceedings of the 3rd International Conference on Learning Representations (ICLR)*],

[23] Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E., "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.* **13**, 600–612 (2004).

[24] Hore, A. and Ziou, D., "Image quality metrics: Psnr vs. ssim," in [*Proceedings of International Conference on Pattern Recognition*], 2366–2369.