Journal of Physics: Conference Series **3010** (2025) 012172 doi:10.1088/1742-6596/3010/1/012172

Improving Image Reconstruction for Ultra-Fast Ptychographic Acquisitions via Deep Learning Denoising

E. Erin[1 a], L. Fardin[1], D. Batey[2], M. Burian[3], S. Vogel[3], S. Grimm[3], M. Fratini[4,5], M. Palombo[6], Fenglei Zhou[1,7], Geoff J. M. Parker[1,8], A. Olivo[1], S. Cipiccia[1,2]

[1] Department of Medical Physics and Biomedical Engineering, University College London, Malet Place Engineering, Gower St, London, WC1E 6BT, UK.
[2] Diamond Light Source, Harwell Science and Innovation Campus, Fermi Avenue, Didcot, OX11 0DE, UK.
[3] DECTRIS Ltd., Taefernweg 1, Baden-Daettwil, 5405, Switzerland.
[4] Institute of Nanotechnology, CNR-Nanotec, Piazzale Aldo Moro 5, 00185 Rome, Italy.
[5] Santa Lucia Foundation, via Ardeatina 356, 00179 Rome, Italy.
[6] Cardiff University Brain Research Imaging Centre, Cardiff University, Maindy Road, Cardiff CF24 4HQ, UK.
[7] BioPhantoms Ltd., Watson House, 4 Mill Park, Cambridge CB1 2FH, UK.
[8] Bioxydyn Limited, St James Tower, 7 Charlotte Street, Manchester M1 4DZ, UK.

Abstract. X-ray ptychography is a scanning coherent diffraction imaging technique which combines nanometer-scale resolution with high penetration depth. This method has been proven to be suitable for scanning weakly absorbing samples and therefore potentially very valuable for medical applications such as brain imaging. However, currently employed scanning techniques present challenges: step-scanning is too slow and inefficient, while fly-scanning introduces blurring and noise into reconstructions due to the motion and reduced photon counts per pixel. To date, only a few methods have been proposed to denoise reconstructions, most of which rely on traditional approaches and are limited in addressing the challenges posed by noise and blurring. To overcome these limitations, we investigate the possibility of using a deep learning-based denoising method combined with position binning. The deep learning-based denoising method, Deep Image Prior (DIP), denoises the reconstructions while position binning increases the photon count statistics per pixel. The method can be integrated within the existing iterative phase retrieval algorithms to denoise the object or probe in between iterations. The method is tested in far-field geometry on two different samples: a Siemens star resolution target and a polymer-based phantom mimicking the white matter of the brain. By assessing the resolution via Fourier ring correlation, we measure up to a 14 % increase in the resolution. However, depending on the architecture used, artifacts due to machine hallucination appear in the denoised images which could be affecting the observed enhancement in resolution. This will be the subject of further investigation.

Corresponding author:
[a]ecem.erin.23@ucl.ac.uk

Journal of Physics: Conference Series **3010** (2025) 012172 doi:10.1088/1742-6596/3010/1/012172

1 Introduction

For many years, imaging techniques such as electron microscopy or visible microscopy were the go-to methods for achieving a high magnification of objects. Such microscopy techniques require the use of an optical element, such as a lens, to form an image. The quality of the image is limited by the performance of the lens which is determined by multiple factors, including efficiency, numerical aperture, and aberrations. To manufacture high performance lenses is increasingly difficult when working in the hard X-ray regime.

Coherent Diffraction Imaging (CDI) overcomes the lens-induced limitations of standard microscopy methods by replacing the lens with computational algorithms (1; 2; 3; 4; 5). The detector is placed where the lens would be to collect diffraction patterns, and computational algorithms form an image starting from the recorded diffraction patterns. The resulting image of the object is aberration-free and theoretically diffraction-limited in resolution.

Ptychography, a scanning form of CDI, allows imaging of extended objects by shifting the illumination with respect to the sample at overlapping steps (3). In conventional ptychography, the scan is done in step-scan mode which is inefficient as it is affected by the overhead due to the move-settle waiting time. To speed up the scanning process, the fly-scanning approach has been introduced (6), (7), where the illumination moves at a constant velocity across the object. Although fly-scanning enables fast ptychography, the motion induces blur and, when also combined with low photon statistics, results in noisy reconstructions. This poses a significant challenge especially in medical imaging as some useful features might be lost during the denoising process.

Traditional denoising methods, such as filtering or statistical methods are often applied to denoise medical images (8). However, they are generally limited to specific imaging methods or noise types (9). Deep learning-based denoising methods offer an alternative (10) but they require extensive training data and ground truth images as labels which are difficult to obtain in ptychography except through simulations. Considering the lack of ground truth data and the limited amount of available training data, we propose the use of a deep learning denoising method called Deep Image Prior (DIP), to remove the noise from reconstructions. This approach utilizes the inherent structure of the imaging data without requiring extensive training datasets, thus offering a promising solution for improving the resolution of the ultra-fast X-ray ptychography reconstruction.

2 Methods

2.1 Experimental Setup

Experimental data were acquired at beamline I13-1 at the Diamond Light Source. Using a double crystal Si-111 monochromator, X-rays with the energy of 11.1 keV are selected. A Fresnel Zone Plate (FZP) with a diameter of 400 µm, outer width of 200 nm is used as the pre-sample focusing optics. An order sorting aperture, with a diameter of 10 µm, is placed close to the focus of the FZP to select the first diffraction order and filter out the higher orders while a central stop is placed right after the FZP, to block the zeroth order. The sample is positioned 8 mm downstream of the focal plane and 10 m downstream of the sample. A vacuum pipe filled with He, placed in between the sample and the detector, minimizes air scattering.

2.2 Detector

The diffraction patterns in the far-field were acquired with an ultra-fast, hybrid single photon counting prototype detector, SELUN, a novel development by DECTRIS AG. Its ASIC allows a continuous frame rate of up to 120 kHz. This high frame rate is enabled by fast front-end electronics and high data readout speed which can reach up to 10 GB s^{-1} in combination with on-chip compression and the possibility for 2×2 binning. This prototype can count more than 45 Mcounts/pixel/s unbinned at 11.1 keV photons and 50 % energy threshold. The detector features a silicon sensor with a thickness of 450 µm on a single chip, 192×192 pixels with a size of 100 µm × 100 µm, covering an active area of 19.2 mm × 19.2 mm (11). Thus, this detector offers a high frame rate, high dynamic range, and measurements without readout noise nor dark current, providing a well-suited solution for fast-scanning ptychography.

2.3 Tool for Reconstruction and Analysis of Reconstruction Quality

All the data presented in this paper were processed using the ePIE algorithm implemented within the PtyREX code (12). The spatial resolution of the reconstructions was assessed using the Fourier ring correlation method (FRC) (13), which is a commonly used image resolution assessment metric (13), (14). It determines the resolution, which is the intersection point of the FRC curve with a predetermined threshold, in Fourier space from a single image as explained in reference (15) and implemented in a Python code called siFRC (16).

2.4 Position Binning

Ptychography allows the imaging of extended objects. However, as the object size increases, very large datasets are produced, creating a substantial computational burden. During the reported experiment, data were acquired with very fine spatial steps with a step size of 0.025 µm and position binning can be applied to reduce computational time and increase statistics. Position binning involves summing counts of multiple diffractograms from a few neighbouring scanning positions up to a certain bin size. This approach increases the photon counts per pixel and consequently improves the Signal-to-Noise Ratio (SNR), which is critical for more accurate reconstructions.

2.5 A Deep Learning Denoising Approach: Deep Image Prior

Deep learning methods have been applied in ptychography to replace phase retrieval and to prevent crosstalk between slices in multislice ptychography (17; 18; 19; 20). In this work we aim to apply DIP for denoising purposes. DIP suggests that the network's architecture inherently captures a significant portion of the statistics of the image without requiring any training or trained parameters (21). To find optimal parameters that will result in the restored/clean image, DIP searches for the parameters space in each iteration. At early iterations, the architecture learns the clean/noiseless image, and noise is added to the clean image after some iteration, meaning the convergence to the clean image is reached before the end of the training process. Beyond this point, the network learns the noise along with the clean image, leading to a decrease in the network's performance and metrics. The denoising experiments were performed on different image types: a polymer-based high-resolution brain white matter phantom and a more geometrical and regular Siemens star resolution target.

For all denoising experiments, we employed a U-Net architecture with a feature map configuration of $[8, 16, 32, 64, 128]$ (**Arch 1**) as recommended in the work by Ulyanov and collaborators (21). Additionally, for the brain sample, we tested an alternative architecture with a feature map configuration of $[128, 128, 128, 128, 128]$ (**Arch 2**). Both architectures included two residual connections, five downsampling and five upsampling layers and were implemented using PyTorch (22) and Monai (23). The Adam optimizer with a learning rate 3×10^{-5} was used and the Mean Squared Error (MSE) was utilized as a loss function. An early stopping criterion, based on the loss function, was employed to prevent overfitting, and the architectures were trained for 4000 epochs if the early stopping criteria were not met. The performance of the architecture was monitored by checking the PSNR score. The network was trained on an NVIDIA Corporation GA104GL [RTX A4000] GPU with 125 GB of RAM resulting in 2-3 minutes per image to complete the denoising for the Siemens star. For all experiments, a fixed random image was given as an input, and all architectural parameters were set as specified before for both **Arch1** and **Arch2**.

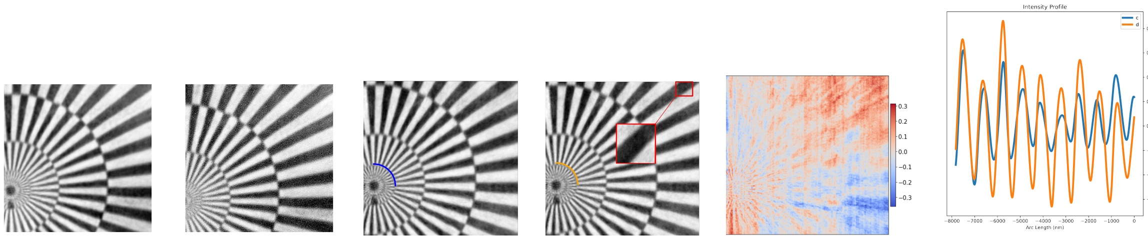
3 Results

3.1 Denoising of The Siemens Star

The reconstructed Siemens star test sample, which contained noise and blurring artifacts, was post-processed using **Arch1**. During the training, the model parameters were iteratively optimized by gradient search to minimize the difference between the noisy reconstruction and the desired clean image. The FRC was calculated to assess the resolution, with the results presented in Figure 1 and Table 1. Figure 1 shows the reconstructed images obtained without, Figure 1a, and with, Figure 1c, position binning. We observed that the denoised version of the unbinned measurement, Figure 1b, appeared noisier than its original counterpart, Figure 1a. In contrast, Figure 1d shows the denoised and binned reconstruction and demonstrates that the DIP denoising step was more effective and improved the resolution when applied to the binned reconstruction, Figure 1c. However, upon closer inspection of the denoised image, we observed regular grid artifact in the denoised reconstruction of Siemens star object, as shown in Figure 1d.

Data type	Label	Denoised	Intersection point	Resolution
Original recons.	a	-	0.351	126 nm
Original recons.	b	✓	0.331	134 nm
Binned recons.	c	-	0.365	121 nm
Binned recons.	d	✓	0.409	108 nm

Table 1: Intersection points and resolution values for different reconstructions with a reconstructed pixel size of 44 nm.

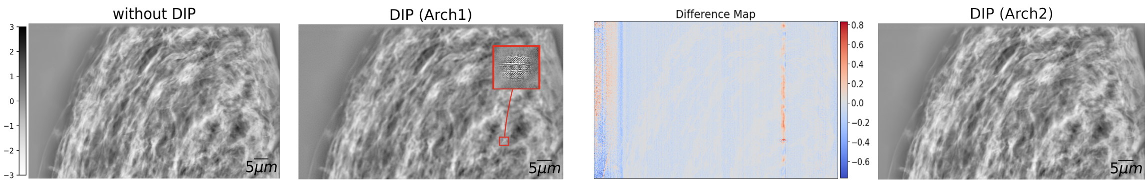


(a) (b) (c) (d) (e) (f)

Figure 1: **The effect of binning and denoising with Arch1 on the resolution of the reconstructions of a Siemens star phantom.** (a) Original reconstruction, where the data is acquired at 28 kHz. (b) The denoised version of the original reconstruction (a). (c) The binned version of the original data acquired at 28 kHz. (d) The denoised version of the binned data (c). (e) The difference map, highlighting the difference between the denoised-binned reconstruction (d) and the original noisy reconstruction (a). (f) Radial profile intensity comparison of c and d.

3.2 Denoising of The Brain Phantom

The brain phantom is made of polymer microcylinders of polycaprolactone (PCL) to mimic the axons in the white matter of the brain. This phantom, by being thoroughly characterized and with closely matching the density to the brain tissue, has demonstrated its representativeness of white matter tissue (24), (25). We applied **Arch1** to the unbinned noisy reconstruction of the brain phantom results of which are shown Figure 2 (b). As for the Siemens star, a regular grid artifact appears in the denoised image, (red square in Figure 2 (b)) but this time more visible than the Siemens star, possibly due to the less regular nature of the sample. To understand if the hallucinations are architecture-dependent, we performed the denoising using another architecture, **Arch2**, on the brain sample, as it is easier to detect the artifact on a less regular sample. The results are reported in Figure 2(d), where no artifacts are observed which suggests that the artifacts might be architecture dependent.



(a) (b) (c) (d)

Figure 2: **Reconstructions of a brain white matter phantom.** Non-denoised and denoised reconstructions, with **Arch1** and **Arch2**, of the brain white matter phantom are shown in a, b, and d respectively. The artifact generated by the **Arch1** and the difference map between non-denoised and denoised reconstructions using **Arch 1** are shown in the b and d respectively.

To evaluate any change in resolution after the denoising, FRC analysis was performed. Both denoised and original reconstructions were cropped and Hanning window applied to minimize boundary effects. The measured resolutions are 179 nm, 170 nm and 240 nm respectively. However, FRC analysis of the denoised brain reconstruction with **Arch1** is not reliable due to the artifact while the denoising with **Arch2** worsens the resolution.

4 Discussion

There were previous attempts to utilize conventional methods to denoise ptychography reconstructions (26) (27). This work investigates a strategy for noise removal in ultra-fast X-ray ptychography by using a deep learning method, DIP.

When measurements are acquired at very fine steps, binning can help with the denoising procedure since it increases the statistics, resulting in better reconstructions with higher resolution and SNR, as shown in Figure 1 (a,c) and Table 1. Consequently, the Siemens star sample was binned prior to denoising. However, binning too many positions would introduce blurring (de-coherence), which could be mitigated by increasing the number of spatial modes in the reconstruction (28) which will be explored in future work.

To test the effectiveness of the proposed deep learning denoising method, we applied the method to different object types. We found that, regardless of the object type, images denoised with **Arch1** exhibited machine hallucinations. Furthermore, we suspect that the grid artifact could contribute to the

Journal of Physics: Conference Series **3010** (2025) 012172 doi:10.1088/1742-6596/3010/1/012172

higher resolution observed in the Fourier ring correlation analysis of the reconstruction denoised with **Arch1** compared to the original image.

Hallucinations of a deep learning model occur when an algorithm generates features that do not correspond to the actual sample structures. To investigate whether the artifact/hallucination is architecture-dependent, we denoised the brain phantom reconstruction using **Arch2**. Unlike **Arch1**, **Arch2** did not generate any hallucinations. The reason why these artifacts only appears in **Arch1** is unknown and requires further investigation. Moreover, while the artifacts were easily detected in the irregular biological tissue, it was harder to detect them in the regular Siemens star sample. Therefore, this highlights a potential risk: hallucinations may go undetected when the method is applied to regular structures, such as integrated circuits.

Summarising, DIP shows potential for applications in ptychographic reconstruction, both as a denoiser and as a prior generator in post-processing. DIP can be integrated into phase retrieval algorithms between iterations to remove noise and prevent its propagation across iterations. Additionally, a trained DIP network can be used to generate reconstruction priors for both the object and the probe, which can then be fed into iterative reconstruction algorithms. However, our experiments have shown architecture dependent weaknesses of the method, which require further investigation as well as comparison with standard non-ML-powered denoising methods.

5 Acknowledgements

This work is supported by EPSRC New Investigator Award EP/X020657/1. MP is supported by UKRI Future Leaders Fellowship (MR/T020296/2). We thank the Diamond Light Source for the beamtime MG34189-1 and MG32306-2.

We would like to thank all beamline staff of I13-1 for their help during the experiment, the team of DECTRIS namely, Sebastian Commichau, and Danial Mayani, for providing a state-of-art detector and Dr. Carole Sudre from UCL for her suggestions on testing the robustness of the proposed methods.

References

[1] Marchesini S, Krishnan H, Daurer BJ, Shapiro DA, Perciano T, Sethian JA, et al. SHARP: a distributed GPU-based ptychographic solver. Journal of Applied Crystallography. 2016;49(4):1245-52. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1107/S1600576716008074.

[2] Enders B, Thibault P. A computational framework for ptychographic reconstructions. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2016;472(2196):20160640. Available from: https://royalsocietypublishing.org/doi/abs/10.1098/rspa.2016.0640.

[3] Maiden A, Rodenburg JM. An improved ptychographical phase retrieval algorithm for diffractive imaging. Ultramicroscopy. 2009;109 10:1256-62. Available from: https://api.semanticscholar.org/CorpusID:8742666.

[4] Guizar-Sicairos M, Fienup J. Phase retrieval with transverse translation diversity: A nonlinear optimization approach. Optics express. 2008 06;16:7264-78.

[5] Wakonig K, Stadler HC, Odstrčil M, Tsai EHR, Diaz A, Holler M, et al. It PtychoShelves, a versatile high-level framework for high-performance analysis of ptychographic data. Journal of Applied Crystallography. 2020 Apr;53(2):574–586.

[6] Jiang Y, Deng J, Yao Y, Klug JA, Mashrafi T, Roehrig Cea. Achieving high spatial resolution in a large field-of-view using lensless x-ray imaging. Applied Physics Letters. 2021;119(12):124101.

[7] Batey D, Rau C, Cipiccia S. High-speed X-ray ptychographic tomography. Scientific Reports. 2022;12:7846.

[8] Patil R, Bhosale S. Medical Image Denoising Techniques: A Review. 2022 Jan;4:21–33.

[9] Zhang K, Ren W, Luo W, Lai WS, Stenger B, Yang MH, et al. Deep Image Deblurring: A Survey. 2022 May;(arXiv:2201.10700). ArXiv:2201.10700 [cs]. Available from: http://arxiv.org/abs/2201.10700.

[10] Kaur A, Dong G. A Complete Review on Image Denoising Techniques for Medical Images. Neural Processing Letters. 2023 Dec;55(6):7807–7850.

[11] Zambon P, Vávra J, Montemurro G, Bottinelli S, Dudina A, Schnyder R, et al. High-frame rate and high-count rate hybrid pixel detector for 4D STEM applications. Frontiers in Physics. 2023 12;11.

[12] Batey DJ. Ptychographic Imaging of Mixed States; 2014. Available from: `https://etheses.whiterose.ac.uk/8524/`.

[13] Banterle N, Bui KH, Lemke EA, Beck M. Fourier ring correlation as a resolution criterion for super-resolution microscopy. Journal of Structural Biology. 2013;183(3):363–367.

[14] Zhao W, Huang X, Yang J, Qu L, Qiu G, Zhao Y, et al. Quantitatively mapping local quality of super-resolution microscopy by rolling Fourier ring correlation. Light: Science & Applications. 2023 Dec;12(1):298.

[15] Koho S, Tortarolo G, Castello M, Deguchi T, Diaspro A, Vicidomini G. Fourier ring correlation simplifies image restoration in fluorescence microscopy. Nature Communications. 2019 Jul;10(1):3103.

[16] Kc P and De Andrade V . siFRC2024; 2024. Available from: `https://github.com/prabhatkc/siFRC/tree/master`.

[17] Shamshad F, Ahmed A. Robust Compressive Phase Retrieval via Deep Generative Priors. 2018 Aug;(arXiv:1808.05854). ArXiv:1808.05854 [cs, stat]. Available from: `http://arxiv.org/abs/1808.05854`.

[18] Wang F, Bian Y, Wang H, Lyu M, Pedrini G, Osten W, et al. Phase imaging with an untrained neural network. Light: Science Applications. 2020 May;9(1):77.

[19] Du M, Huang X, Jacobsen C. Using a modified double deep image prior for crosstalk mitigation in multislice ptychography. Journal of Synchrotron Radiation. 2021 Jul;28(4):1137–1145. ArXiv:2102.00869 [eess].

[20] Guzzi F, Kourousias G, Billè F, Pugliese R, Gianoncelli A, Carrato S. A Deep Prior Method for Fourier Ptychography Microscopy. In: 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO); 2021. p. 1781-6.

[21] Ulyanov D, Vedaldi A, Lempitsky V. Deep Image Prior. International Journal of Computer Vision. 2020 Jul;128(7):1867–1888. ArXiv:1711.10925 [cs, stat].

[22] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al.. PyTorch: An Imperative Style, High-Performance Deep Learning Library; 2019. Available from: `https://arxiv.org/abs/1912.01703`.

[23] Cardoso MJ, Li W, Brown R, Ma N, Kerfoot E, Wang Y, et al.. MONAI: An open-source framework for deep learning in healthcare; 2022. Available from: `https://arxiv.org/abs/2211.02701`.

[24] Zhou FL, Hubbard PL, Eichhorn SJ, Parker GJM. Coaxially Electrospun Axon-Mimicking Fibers for Diffusion Magnetic Resonance Imaging. ACS Applied Materials Interfaces. 2012;4(11):6311–6316.

[25] Cipiccia S, Fratini M, Erin E, Palombo M, Vogel S, Burian M, et al. Fast X-ray ptychography: towards nanoscale imaging of large volume of brain. The European Physical Journal Plus. 2024 May;139(55):434.

[26] Chang H, Enfedaque P, Zhang J, Reinhardt J, Enders B, Yu YS, et al. Advanced denoising for X-ray ptychography. Optics Express. 2019 Apr;27(8):10395.

[27] Melnyk O, Römer P. Background Denoising for Ptychography via Wigner Distribution Deconvolution; 2024. Available from: `https://arxiv.org/abs/2402.15353`.

[28] Thibault P, Menzel A. Reconstructing state mixtures from diffraction measurements. Nature. 2013 Feb;494(7435):68–71.