Angular Sparsity Invariant Tilt Series Generation in Scanning/Transmission Electron Microscopy

Dikshant Sagar
Department of Computer Science
University of California, Irvine
dikshans@uci.edu

Zhaoxu Li
Department of Materials Science and Engineering
University of California, Irvine
zhaoxu.li@uci.edu

Sherif Abdelkarim
Department of Computer Science
University of California, Irvine
abdelkas@uci.edu

Pierre Baldi
Department of Computer Science
University of California, Irvine
pfbaldi@ics.uci.edu

Joseph. P. Patterson
Department of Materials Science and Engineering
University of California, Irvine
patters3@uci.edu

Abstract

Electron tomography is essential for probing three-dimensional material structures at the nanoscale. While acquiring a sufficient number of images across multiple tilt angles is critical, doing so is challenging for beam-sensitive samples. Deep convolutional neural networks have emerged as promising tools for overcoming the limitations of sparse tilt-angle sampling. In this work, we introduce a large-scale simulated electron tomography dataset and propose an efficient encoder–decoder DCNN that predicts intermediate tilt frames to enhance sparse tilt series. Our approach improves 3D reconstruction resolution and scalability, achieving an average performance of 5.77e-9 MSE, 31.63 PSNR, and 0.93 SSIM on the held-out aberrated test set across multiple sparsity levels. To support reproducibility, the code used for training and experiments in this study has been released at https://github.com/dikshantsagar/TSGNet.

1 Introduction

Electron tomography (ET) emerged as a powerful approach for three-dimensional analysis of nanostructures. It reconstructs three-dimensional structures from a series of two-dimensional (S)TEM images acquired over a range of specimen tilt angles [21]. It has been widely used to study the structures of catalysts[7], alloys[32], biomaterials[9], and nanoparticles[26]. For beam-sensitive materials, the total number of tilt images that can be acquired is strictly limited by radiation tolerance[14]. This sparse angular sampling leads to limited projection angles, which manifests as artifacts and reduced resolution of the reconstructed structure[11]. Overcoming this limitation remains challenging, as any additional data acquisition risks compromising the structural integrity of the specimen while maintaining an adequate dose for each tilt image.

Deep convolutional neural networks (DCNNs) have transformed image processing across diverse scientific and engineering domains [1; 15]. By leveraging hierarchical feature extraction and nonlinear representations, DCNNs excel at capturing both fine-grained spatial details and high-level structural

patterns [16]. They have been successfully applied to tasks such as denoising, super-resolution, inpainting, and frame interpolation, where conventional model-based approaches often struggle [31; 28; 5; 24]. In microscopy, DCNNs have shown promise in restoring low-dose images, segmenting complex structures, and enhancing temporal and spatial resolution, underscoring their suitability for addressing the challenges of limited or noisy data [17; 13; 27]. These advances make DCNNs an attractive solution for overcoming the limitations imposed by sparse angular sampling in electron tomography.

In this work, we utilize a new large-scale simulated S/TEM dataset and propose a novel DCNN-based encoder-decoder framework, named TSGNet, to interpolate missing tilt images in sparsely sampled series. By recovering intermediate projections, our method reduces artifacts caused by limited angular sampling while preserving the structural integrity of radiation-sensitive specimens. We validate the approach on simulated STEM tilt series of metal nanoparticles, demonstrating significant improvements over conventional sparse reconstructions and performance comparable to tilt series with twice the angular sampling density.

2 Methods

2.1 Dataset Generation

The datasets were generated in two steps: structure generation and image generation. A total of 101 icosahedral, decahedral, and octahedral structures (Pt, Ag, Au, Cu) were generated for the training set, and 99 face-centered cubic structures (Yb, Ir, Xe, Al) were assigned to the test set; all structures were created with the ASE Python package [12]. Then the generated structures were randomly rotated in the x, y, and z axis to represent random orientations before tilting. Subsequently, these generated structures were input into PyQSTEM[19] to simulate the scanning transmission electron microscopy (STEM) high-angle annular dark-field (HAADF) images. The accelerating voltage is set to 200 keV. The annular detector semi-angle is from 0 to 45 mrad. Defocus value is set to 0. The beam convergence semi-angle is set to 35 mrad for non-aberrated group and 25 mrad for the aberrated group. The spherical aberration is set to 100 μm. The tilt series of the structures were generated by tilting the structure around the y-axis from -90 degrees to 90 degrees with a 1-degree step, producing 181 non-aberrated and 181 aberrated STEM HADDF images for each structure. Thus, both the aberrated and non-aberrated datasets contained 18,281 training images and 17,919 testing images each.

2.1.1 Tilt Angle Sparsity Subsampling

Sparse frame stacks were generated from the original tilt stacks ($512 \times 512 \times 181$) by doing offset-based sampling of the tilt projection. Let each tilt series be represented as a sequence of $N = 181$ frames, $\mathcal{F} = \{f_0, f_1, f_2, ..., f_{180}\}$, where f_i denotes the frame acquired at tilt angle $\theta_i \in \{-90°, -89°, -87°, ..., +90°\}$, with uniform angular spacing. To simulate angular sparsity, we subsample the tilt series using a fixed frame gap $g \in \{1, 2, 3, ..., 12\}$ and an offset $o \in \{0, 1, 2, ..., g-1\}$. Then, the sparse subset with gap g and offset o is given by: $\mathcal{F}_g^{(o)} = \{f_i \in \mathcal{F} | i \equiv o(mod g)\}$

This allows us to generate multiple sparse patterns for each structure, such as $\{0, 3, 6, ..\}$ for $o = 0$ and $\{1, 4, 7, ...\}$ for $o = 1$ when $g = 3$.

2.2 Deep Learning Framework for Tilt Series Generation

Given two input frames f_α and f_β, our model (Fig. 1) predicts an intermediate frame $\hat{f}_\gamma$, where $\theta_\alpha < \theta_\gamma < \theta_\beta$. A shared pyramid encoder extracts multi-scale features $\varphi_\alpha^k, \varphi_\beta^k$ ($k = 1..4$), which are processed by four coarse-to-fine decoders D_k. Each decoder aggregates features from both inputs and refines intermediate flow fields through lightweight residual blocks and transposed convolutions, enabling effective feature integration across scales. The model is trained end-to-end with a reconstruction loss [23] $\mathcal{L}_r = \rho(\hat{f}_\gamma - f_\gamma^{GT}) + \mathcal{L}_{cen}(\hat{f}_\gamma, f_\gamma^{GT})$, where ρ is the Charbonnier loss [4]($\alpha = 0.5$, $\epsilon = 10^{-3}$) and $\mathcal{L}_{cen}$ is the census loss using 7×7 patches [20]. To further enforce geometric consistency, we add $\mathcal{L}_g = \sum_{k=1}^3 \mathcal{L}_{cen}(\hat{\varphi}_{GT}^k, \varphi_{GT}^k)$, which aligns encoder features with ground-truth decoder features. The final objective is $\mathcal{L} = \mathcal{L}_r + \eta\mathcal{L}_g$, $\eta = 0.01$.
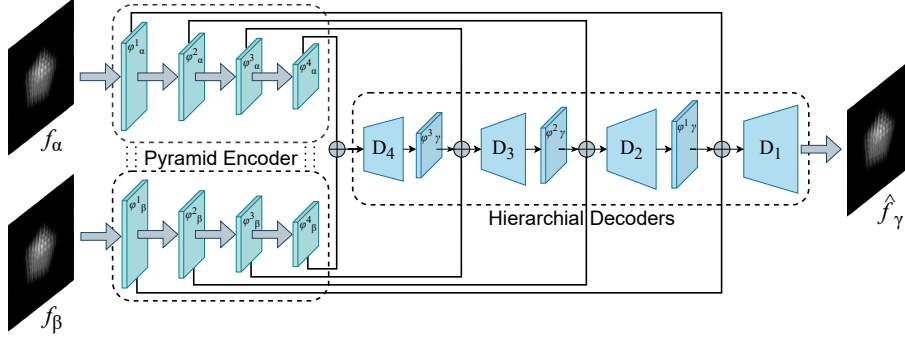
Figure 1: Overview of the proposed TSGNet for tilt frame interpolation. Multi-scale features extracted by a shared pyramid encoder are processed through the decoders to predict intermediate tilt images.

Training was performed for 50 epochs on two NVIDIA A5000 GPUs using AdamW [18] (weight decay 0.02) with a cosine-annealed learning rate from 10^{-4} to 10^{-5} and a batch size of 24.

2.3 3D Reconstruction and Evaluation

Three reconstruction groups were generated. The Reference group used all 181 tilt images and serves as the highest-fidelity baseline. The Sim group used a subset of the 181 tilts defined by the downsampling rate. The Mix group used the same subset as Sim at each downsampling rate, augmented with model-predicted images to fill the missing tilts. The reconstruction from the corresponding tilting series was performed using tomopy python package [10]. Simultaneous iterative reconstruction technique (SIRT) algorithm[8; 29; 25; 29] with 100 iterations was used. After the reconstruction, the reconstructed tensors were segmented into objects and background by the Otsu thresholding method [22]. And only the objects were analyzed with corresponding references using Chamfer distance [2; 3; 6] and RMS error with 1000000 points, providing a quantitative performance comparison.

3 Results

We evaluated the proposed deep encoder–decoder interpolation framework against plain linear interpolation on the synthetic metallic nanocluster tilt series held-out test set. Performance was quantified using mean squared error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity index measure (SSIM), computed between the predicted intermediate frames and their corresponding ground truth images [30] for sparsity subsampling rates 1 to 12 for both aberrated and non-aberrated datasets (See Fig.2). Table 1 summarizes the results averaged across the test set and sparsity levels. Overall, TSGNet outperformed linear interpolation across all metrics in both aberrated and non-aberrated cases, achieving lower reconstruction error and higher perceptual similarity. These gains indicate that the proposed model captures finer structural details and reduces artifacts more effectively than the baseline.

Table 1: Average quantitative comparison between the proposed model and linear interpolation for all sparsity subsampling rates (1-12) on the test set. Lower MSE and higher PSNR/SSIM indicate better performance.

Method	Abberated			Non-Abberated		
	MSE↓	PSNR↑	SSIM↑	MSE↓	PSNR↑	SSIM↑
Linear Interpolation	7.48e-9	30.48	0.90	1.13e-8	28.57	0.86
TSGNet (Ours)	**5.77e-9**	**31.63**	**0.93**	**7.13e-9**	**29.43**	**0.89**

The quality of reconstructions degraded with increasing downsampling in both aberrated and non-aberrated cases. Non-aberrated reconstructions appeared smoother with fewer structures, whereas
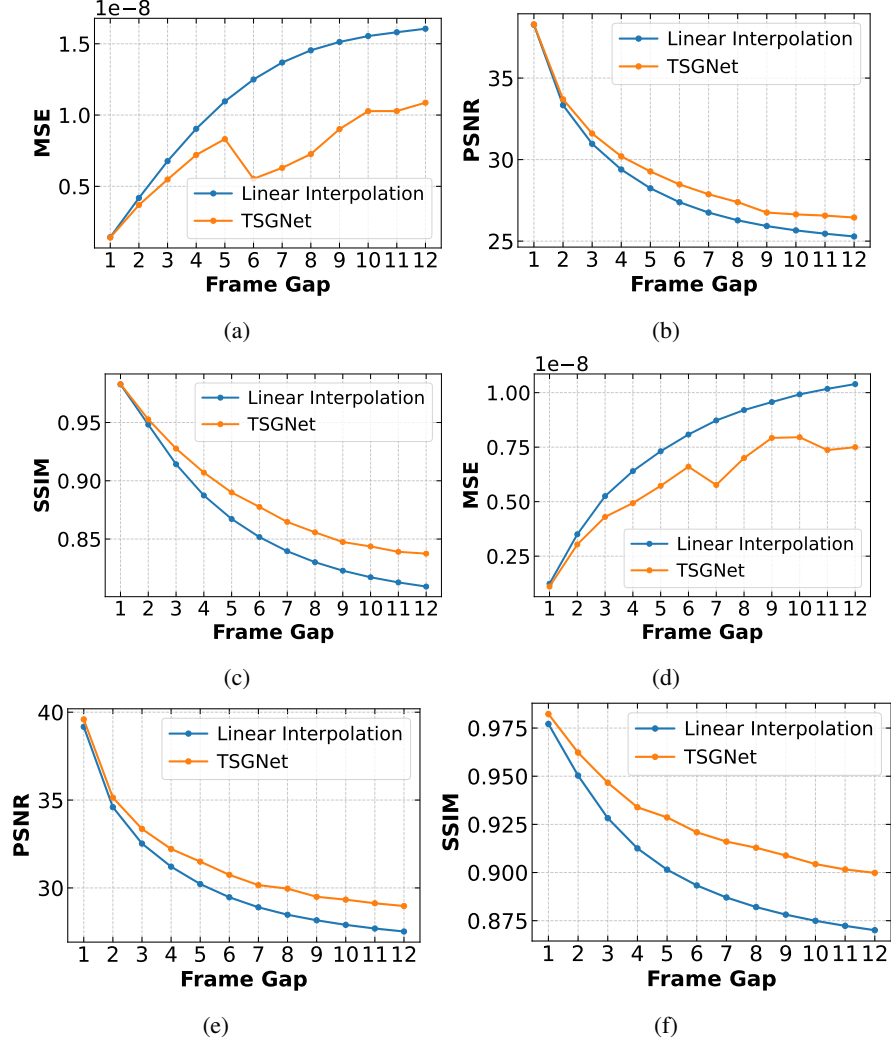
Figure 2: Quantitative comparison of linear interpolation and TSGNet across sparsity levels. Plots show (a) MSE, (b) PSNR, and (c) SSIM for the non-aberrated test set, and (d) MSE, (e) PSNR, and (f) SSIM for the aberrated test set.

aberrated reconstructions revealed finer features but also more background artifacts, likely due to the difficulty of suppressing sharper details. In all settings, the reconstructions in the mixed group (with TSGNet-generated frames) contained fewer artifacts than those in the purely simulated group, indicating that TSGNet-generated images help suppress noise and improve detail recovery (see Fig. 4).

Visual inspection of interpolated frames supports these findings. Linear interpolation produced blurred edges and failed to capture atomic-scale contrast, while our model reconstructed sharp high-frequency details such as atomic columns and defects, even under severe angular sparsity. This fidelity directly improved downstream 3D reconstructions, enabling higher-quality volumes with fewer tilts, thereby reducing acquisition time and dose without sacrificing structural integrity.

Quantitative evaluation with Chamfer distance and RMS error (Fig. 3) confirmed these trends. Similarity to the reference decreased with higher downsampling, stabilizing beyond a gap of six, likely due to loss of fine structural cues. In both aberrated and non-aberrated cases, the mixed group consistently outperformed the simulation group. Notably, when both groups contained the same number of tilt images, the mixed group achieved equal or better similarity, demonstrating that ML-generated intermediate frames effectively enhanced reconstruction quality.
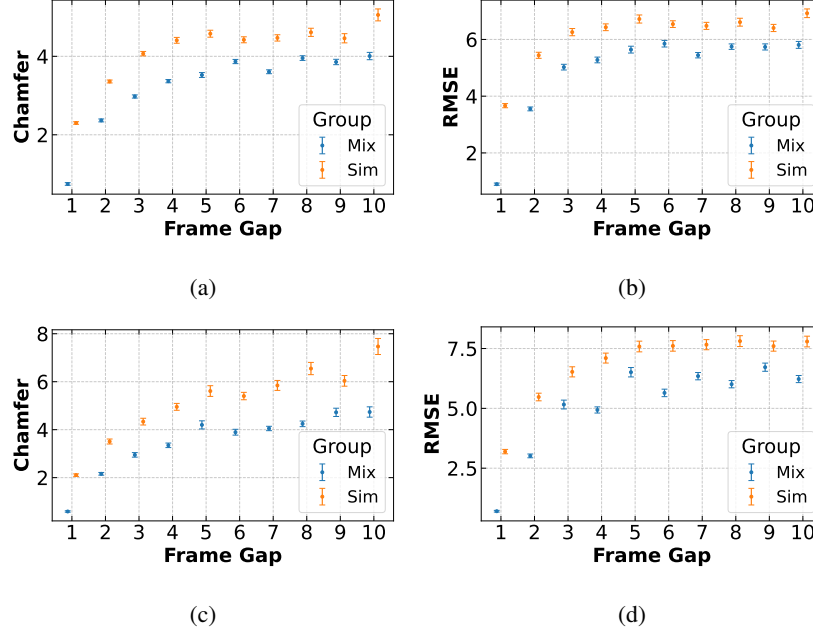
4

Figure 3: Reconstruction metrics for the simulation and mixed groups. (A) Chamfer distances in the non-aberrated case. (B) Root mean square (RMS) errors in the non-aberrated case. (C) Chamfer distances in the aberrated case. (D) RMS errors in the aberrated case.
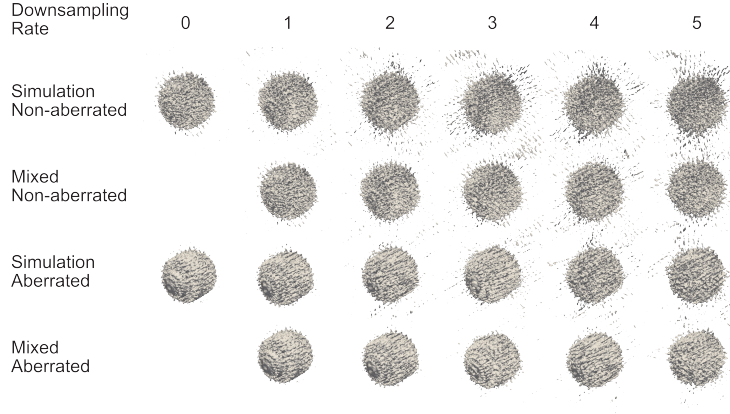

Figure 4: Example reconstructions of the same structure at different downsampling rates (1 - 5) for the simulation and mixed groups, shown for both non-aberrated and aberrated cases. Downsampling rate at 0 indicates reference reconstruction in both cases.

4 Conclusion

In this work, we developed and applied a DCNN encoder-decoder framework named TSGNet to predict missing tilt images in simulated STEM tilt series of nanoclusters. The generated frames closely matched ground truth projections and led to more accurate 3D reconstructions in both aberrated and non-aberrated cases. While our evaluation focused on simulated nanoclusters, the results highlight the promise of deep learning for mitigating angular sparsity and improving tomographic reconstructions of beam-sensitive materials. Looking ahead, we aim to extend this approach to diverse nanostructures, experimental datasets, and advanced learning architectures to further advance high-resolution, low-dose electron tomography.

References

[1] BALDI, P. *Deep learning in science*. Cambridge University Press, 2021.

[2] BARROW, H. G., TENENBAUM, J. M., BOLLES, R. C., AND WOLF, H. C. Parametric correspondence and chamfer matching: Two new techniques for image matching. In *Proceedings of IJCAI* (1977).

[3] BORGEFORS, G. Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6 (1988), 849–865.

[4] CHARBONNIER, P., BLANC-FERAUD, L., AUBERT, G., AND BARLAUD, M. Two deterministic half-quadratic regularization algorithms for computed imaging. In *Proceedings of 1st international conference on image processing* (1994), vol. 2, IEEE, pp. 168–172.

[5] ELHARROUSS, O., ALMAADEED, N., AL-MAADEED, S., AND AKBARI, Y. Image inpainting: A review. *Neural Processing Letters 51*, 2 (2020), 2007–2028.

[6] FAN, H., SU, H., AND GUIBAS, L. J. A point set generation network for 3d object reconstruction from a single image. In *CVPR* (2017).

[7] FRIEDRICH, H., DE JONGH, P. E., VERKLEIJ, A. J., AND DE JONG, K. P. Electron tomography for heterogeneous catalysts and related nanostructured materials. *Chemical Reviews 109*, 5 (2009), 1613–1629. PMID: 19301813.

[8] GILBERT, P. Iterative methods for the three-dimensional reconstruction of an object from projections. *Journal of Theoretical Biology 36*, 1 (1972), 105–117.

[9] GRANDFIELD, K., PALMQUIST, A., AND ENGQVIST, H. High-resolution three-dimensional probes of biomaterials and their interfaces. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 370*, 1963 (2012), 1337–1351.

[10] GÜRSOY, D., DE CARLO, F., XIAO, X., AND JACOBSEN, C. TomoPy: a framework for the analysis of synchrotron tomographic data. *Journal of Synchrotron Radiation 21*, 5 (Sep 2014), 1188–1193.

[11] HAGEN, W. J., WAN, W., AND BRIGGS, J. A. Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. *Journal of Structural Biology 197*, 2 (2017), 191–198. Electron Tomography.

[12] HJORTH LARSEN, A., JØRGEN MORTENSEN, J., BLOMQVIST, J., CASTELLI, I. E., CHRISTENSEN, R., DUŁAK, M., FRIIS, J., GROVES, M. N., HAMMER, B., HARGUS, C., HERMES, E. D., JENNINGS, P. C., BJERRE JENSEN, P., KERMODE, J., KITCHIN, J. R., LEONHARD KOLSBJERG, E., KUBAL, J., KAASBJERG, K., LYSGAARD, S., BERGMANN MARONSSON, J., MAXSON, T., OLSEN, T., PASTEWKA, L., PETERSON, A., ROSTGAARD, C., SCHIØTZ, J., SCHÜTT, O., STRANGE, M., THYGESEN, K. S., VEGGE, T., VILHELMSEN, L., WALTER, M., ZENG, Z., AND JACOBSEN, K. W. The atomic simulation environment—a python library for working with atoms. *Journal of Physics: Condensed Matter 29*, 27 (jun 2017), 273002.

[13] HORWATH, J. P., ZAKHAROV, D. N., MÉGRET, R., AND STACH, E. A. Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images. *npj Computational Materials 6*, 1 (2020), 108.

[14] KONETI, S., ROIBAN, L., DALMAS, F., LANGLOIS, C., GAY, A.-S., CABIAC, A., GRENIER, T., BANJAK, H., MAXIM, V., AND EPICIER, T. Fast electron tomography: Applications to beam sensitive samples and in situ tem or operando environmental tem studies. *Materials Characterization 151* (2019), 480–495.

[15] LECUN, Y., BENGIO, Y., AND HINTON, G. Deep learning. *nature 521*, 7553 (2015), 436–444.

[16] LIN, T.-Y., DOLLÁR, P., GIRSHICK, R., HE, K., HARIHARAN, B., AND BELONGIE, S. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017), pp. 2117–2125.

[17] LOBATO, I., FRIEDRICH, T., AND VAN AERT, S. Deep convolutional neural networks to restore single-shot electron microscopy images. *npj Computational Materials 10*, 1 (2024), 10.

[18] LOSHCHILOV, I., AND HUTTER, F. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101* (2017).

[19] MADSEN, J., LIU, P., KLING, J., WAGNER, J. B., HANSEN, T. W., WINTHER, O., AND SCHIØTZ, J. A deep learning approach to identify local structures in atomic-resolution transmission electron microscopy images. *Advanced Theory and Simulations 1*, 8 (2018), 1800037.

[20] MEISTER, S., HUR, J., AND ROTH, S. Unflow: Unsupervised learning of optical flow with a bidirectional census loss. In *Proceedings of the AAAI conference on artificial intelligence* (2018), vol. 32.

[21] MIDGLEY, P. A., AND DUNIN-BORKOWSKI, R. E. Electron tomography and holography in materials science. *Nature Materials 8*, 4 (Apr. 2009), 271–280.

[22] OTSU, N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics 9*, 1 (1979), 62–66.

[23] PARK, J., LEE, C., AND KIM, C.-S. Asymmetric bilateral motion estimation for video frame interpolation. In *Proceedings of the IEEE/CVF international conference on computer vision* (2021), pp. 14539–14548.

[24] PELEG, T., SZEKELY, P., SABO, D., AND SENDIK, O. Im-net for high resolution video frame interpolation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern Recognition* (2019), pp. 2398–2407.

[25] PELT, D. M., PALENSTIJN, W. J., AND BATENBURG, K. J. Integration of tomopy and the astra toolbox for advanced processing and reconstruction of tomographic synchrotron data. *Journal of Synchrotron Radiation 23*, 3 (2016), 842–849.

[26] STEWART, P. L. Cryo-electron microscopy and cryo-electron tomography of nanoparticles. *WIREs Nanomedicine and Nanobiotechnology 9*, 2 (2017), e1417.

[27] SU, Y.-T., LU, Y., CHEN, M., AND LIU, A.-A. Spatiotemporal joint mitosis detection using cnn-lstm network in time-lapse phase contrast microscopy images. *IEEE Access 5* (2017), 18033–18041.

[28] TIAN, C., XU, Y., ZUO, W., ZHANG, B., FEI, L., AND LIN, C.-W. Coarse-to-fine cnn for image super-resolution. *IEEE Transactions on Multimedia 23* (2020), 1489–1502.

[29] VAN AARLE, W., PALENSTIJN, W. J., DE BEENHOUWER, J., ALTANTZIS, T., BALS, S., BATENBURG, K. J., AND SIJBERS, J. The astra toolbox: A platform for advanced algorithm development in electron tomography. *Ultramicroscopy 157* (2015), 35–47.

[30] WANG, Z., BOVIK, A. C., SHEIKH, H. R., AND SIMONCELLI, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing 13*, 4 (2004), 600–612.

[31] ZHANG, K., ZUO, W., CHEN, Y., MENG, D., AND ZHANG, L. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing 26*, 7 (2017), 3142–3155.

[32] ZHOU, X., LUO, C., HASHIMOTO, T., HUGHES, A., AND THOMPSON, G. Study of localized corrosion in aa2024 aluminium alloy using electron tomography. *Corrosion Science 58* (2012), 299–306.