Enhancing spatial resolution in MÖNCH for electron microscopy via deep learning

X. Xie,[a,*] **L. Barba Flores,**[b,c] **B. Bejar Haro,**[b,c] **A. Bergamaschi,**[a] **E. Fröjdh,**[a] **E. Müller,**[a] **K.A. Paton,**[a] **E. Poghosyan**[a] **and C. Remlinger**[c]

[a]*Paul Scherrer Institut,*
 5232 Villigen PSI, Switzerland
[b]*Swiss Data Science Center (SDSC Hub at PSI),*
 5232 Villigen PSI, Switzerland
[c]*Swiss Data Science Center,*
 1015 Lausanne, Switzerland

 E-mail: xiangyu.xie@psi.ch

ABSTRACT: Hybrid Pixel Detectors (HPDs) are highly suitable in diffraction-based electron microscopy due to their high frame rates (> 1 kHz), high dynamic range, and good radiation hardness. However, their use in imaging applications has been limited by their relatively large pixel size ($\geq$55 µm) and high-energy (> 80 keV) electrons scattering over multiple pixels in the sensor layer. To realize the full potential of fast, radiation-hard HPDs across electron microscopy modalities, we developed deep learning techniques to precisely localize the impact point of incident electrons in MÖNCH, a charge integrating HPD with 25 µm pixel size. With neural network models trained using labeled data via simulations and experimental measurements, the best spatial resolution obtained, defined in terms of the root mean squared error, was 0.60 pixels for 200 keV electrons, a three-fold improvement over a simple charge centroid method. This article presents the training sample generation, deep learning model design, training results, and imaging outcomes for a sample containing gold nanoparticles.

KEYWORDS: Data processing methods; Hybrid detectors; Pixelated detectors and associated VLSI electronics

---

*Corresponding author.

Contents

1 Introduction

In electron microscopy, the adoption of direct electron detectors (DEDs) has led to the "resolution revolution" [1], enabling new experimental modalities across a range of fields [2]. This revolution is primarily attributed to the improved image quality, efficient image acquisition, and the capability to correct motion during image capture of DEDs over traditional film-based and CCD-based technologies [3]. Among DEDs, back-thinned Monolithic Active Pixel Sensors (MAPS) are the prevalent choice for transmission electron microscopy (TEM) due to their small pixel size, for example 5 μm in the K2 Summit and 14 μm in the Falcon 4 [4]. These sensors excel particularly when used with high-energy (200 keV and 300 keV) electrons, which are transmitted through the ultra-thin sensor with minimal lateral scattering. This effect also reduces the variance in the total energy deposited[1] which facilitates the identification of individual electrons from the detector noise enabling electron counting. However, back-thinned MAPS have drawbacks in terms of radiation hardness and frame rate, restricting their potential applications in electron microscopy.

An alternative type of DEDs, Hybrid Pixel Detectors (HPD), presents a promising solution to these limitations. HPDs consist of a pixelated sensor with pixels bump bonded to a readout chip, offering superior radiation hardness, fast readout, and an extended dynamic range. As a result, HPDs are popular for diffraction-based electron microscopy modalities, as exemplified by detectors like JUNGFRAU [5, 6] and Medipix [7, 8] which are popular for diffraction-based modes of electron microscopy such as microcrystal electron diffraction and scanning transmission electron microscopy (STEM). However, the existing HPDs are less suitable for imaging applications due to their diminished spatial resolution, which arises from their large (≥ 55 μm) pixels as well as the multiple scattering of primary electrons in their relatively thick sensors.

---

[1]The rate of electron energy loss decreases as the electron energy increases within the range of energies used in current-generation microscopes. This decrease is reflected in a reduced variance in the total amount of energy deposited in thin sensors with increasing electron energy.

MÖNCH [9] is a general-purpose, charge-integrating HPD under development which is notable for its small pixels compared with current generation HPDs. With a pitch of 25 µm it is approaching the limitations of bump-bonding technologies. It has a maximum frame rate of 6 kHz, incorporates a 14-bit analog-to-digital converter, and offers various gain modes. Although the pixel size is comparable to that of MAPS, the spatial resolution of MÖNCH remains constrained by the long-range scattering of electrons over 80 keV, as illustrated in figure 1. The random nature of the electron tracks makes it hard to derive a consistent algorithm to find the impact point of the electrons from first principles. However, deep learning methods can exploit information latent in the energy recorded by individual pixels, which depends a) on the energy deposited by electrons as they scatter as well as b) on the diffusion of carriers created by the deposited energy to enhance a detector's spatial resolution.
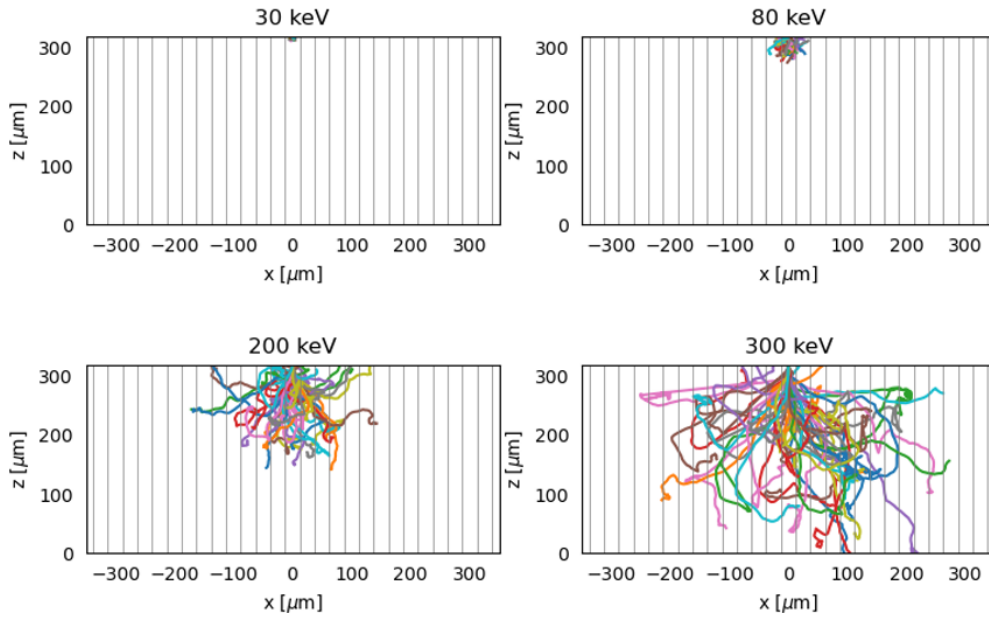

Figure 1. The simulated tracks of electrons with different incident energies entering perpendicularly at $x = 0$ of a 25 µm pixelated silicon sensor. The grid lines indicate pixel boundaries.

In the present study, we focus on 200 keV electrons, a regime where the multiple scattering effect is notably pronounced. Lower energy electrons which are of increasing interest [10] for the increase in the ratio of elastic to inelastic cross-sections (which results in an improved level of image contrast for a given degree of damage to the sample) will be covered by subsequent work. We specifically concentrate on single-electron event analysis using deep learning. The structure and findings of our study are presented as followed: section 2 describes the training sample generation from simulation and, for the first time in similar studies [11, 12], from experimental data. In section 3, we detail the deep learning model, the employed data augmentation techniques, the training setup, and the ensuing training results including the imaging outcomes for a sample containing gold nanoparticles. Finally, section 4 concludes our study and presents an outlook on future research.

2 Training sample generation

Clean and unbiased samples are crucial for training the deep learning model. We started with the simulation sample generation where the ground truth impact points are known. However, simulations can not capture all features of detector response. Therefore, we developed a novel method to obtain ground truth impact points and prepared training samples from experimental data, utilizing a MÖNCH 03 prototype on an electron microscope and limiting the electron beam to a sub-pixel area on the sensor.

One training sample is comprised of a single-electron cluster, cluster features, and the corresponding ground truth impact point. The single-electron cluster, i.e., the detector response in terms of collected charge,[2] is centered in a 15×15 pixel image. Cluster features consist of the cluster charge, the cluster's diameter along both dimensions, its size, the average charge per pixel, the highest pixel charge, and the nominal electron energy.

2.1 Simulation samples

To generate our simulation samples, we used a Geant4-based [13] simulation framework to simulate electron interactions with the sensor, combined with a charge diffusion model [14] for simulating the motion of charge carriers in the sensor. Our simulation configuration closely mirrors the design parameters of the MÖNCH 03 prototype: a 25 μm pixelated, 320 μm-thick silicon sensor containing 400×400 pixels and biased at 90 V. In the simulation, electron impact points span multiple pixels, ensuring a uniform distribution of incident position within one pixel. Figure 2(a) shows a simulation event, including the trajectory of a 200 keV incident electron and the single-electron cluster. We also include the result of the charge centroid method, which calculates the center of gravity of the charge distribution. Figure 2(b) presents a histogram of the residuals in the x direction between the impact and charge centroid points. The root mean square error (RMSE) of 1.73 pixels serves as an indicator of the spatial resolution of MÖNCH as derived from simulation.

2.2 Experimental samples

The electron microscope used in this study is a JEOL JEM-ARM200F NEOARM operating at 200 keV (see figure 3(a)) in STEM mode. No sample was inserted in the microscope. A MÖNCH 03 [9] detector using custom-built mechanics is bottom-mounted on the microscope, as shown in figure 3(b). The detector was operated at a frame rate of 1.6 kHz with an exposure time of 5 μs and a bias voltage of 90 V. MÖNCH's low gain mode was employed so that each pixel had a dynamic range of 100 keV and a noise level of 0.18 keV. Using a custom alignment procedure that we have developed, the electron beam was converged on the surface of the sensor of MÖNCH 03 with a beam diameter of approximately 2 μm. The electron beam was programmed to scan randomly across the sensor, as shown in figure 3(c), with a dwell time of six seconds to acquire 10k frames at each scanning point. By averaging the charge centroid points across all frames at each scanning point, we could obtain an unbiased estimation of the ground truth impact point with negligible uncertainty compared to the pixel size. The electron beam current was set to be low to minimize multiple-electron events. Random scanning and hundreds of scanning points ensured an even distribution of impact points within one pixel, enhancing the generality of our training samples.

---

[2]The collected charge is in the unit of keV, equivalent to the deposited energy of the incident electron.
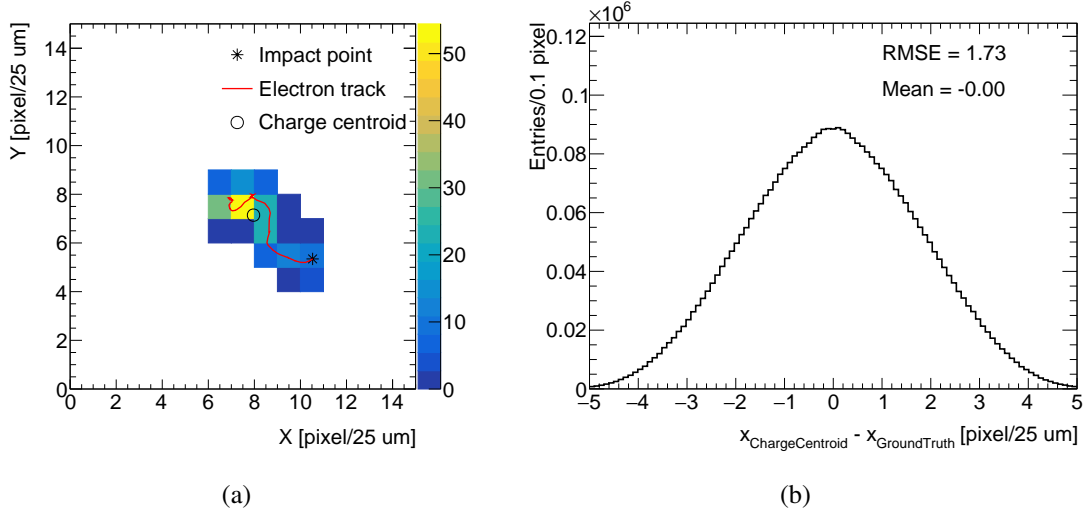
(a) (b)

Figure 2. (a) Simulated 200 keV electron event including the single-electron cluster, the charge centroid and ground truth impact points, and the electron trajectory. z axis in keV stands for the collected charge by each pixel. The charge centroid and ground truth impact points are overlaid. (b) Histogram of the residuals between the impact and charge centroid points in the x direction. The RMSE of 1.73 pixels serves as an indicator of MÖNCH's spatial resolution in the simulation.
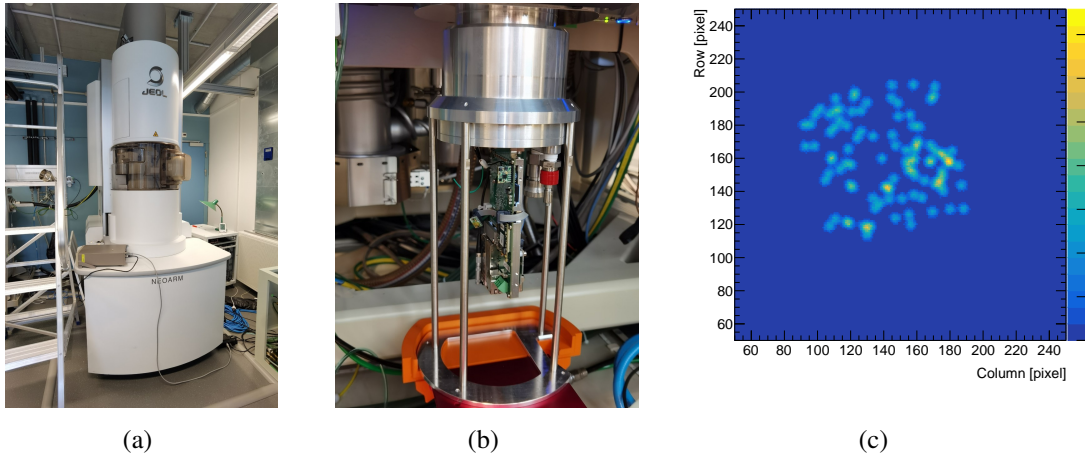

(a) (b) (c)

Figure 3. (a) The JEOL JEM-ARM200F NEOARM electron microscope. (b) MÖNCH 03 with the custom-built mechanics bottom-mounted on the electron microscope. (c) Heat map from scanning points on MÖNCH 03.

To obtain single-electron events, we treated data in the following way: we estimated the pedestal (mean) and noise (RMSE) from frames acquired when the detector was not exposed to any sources of illumination before the actual measurements. Each measurement frame underwent pedestal subtraction, and the raw data was then converted into energy (in terms of keV) using a linear gain calibration. Clusters were identified by grouping adjacent pixels with readout values exceeding five times the noise level (5×0.18 keV ≈ 1 keV). We selected single-electron clusters by ensuring the cluster energy falls within the 20 to 205 keV range to cover backscattered events. To further improve the quality of the training samples, we filtered out those clusters for which the distance between the charge centroid and true impact points is greater than $5\sqrt{2} \times 1.83$ pixels (5σ limit in the 2D plane) and those clusters that

do not cover the impact point (non-primary incident electrons). Figure 4(a) shows an experimental sample, while figure 4(b) displays a histogram indicating a spatial resolution of 1.83 pixels via the charge centroid method, comparable to the simulation result in figure 2(b).
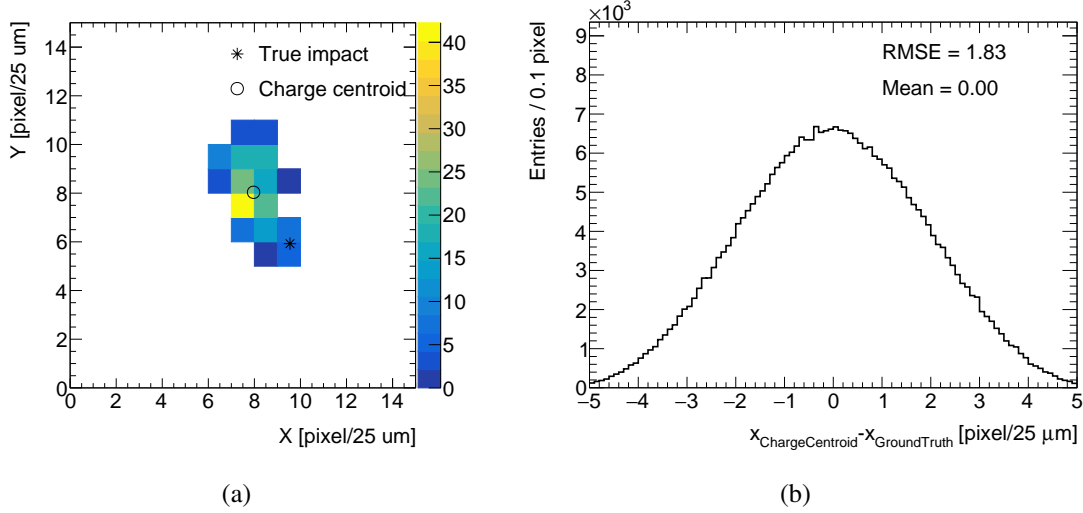

Figure 4. (a) Measured 200 keV electron event including the single-electron cluster, the charge centroid and ground truth impact points. z axis in keV stands for the collected charge by each pixel. (b) Histogram of the residuals between the actual impact and charge centroid points in the x direction. The RMSE of 1.83 pixels serves as an indicator of MÖNCH 03's spatial resolution in the measurement.

3 Deep learning training

To implement and train the deep learning model we used PyTorch [15]. There are 4M samples from simulations and 1.4M samples from experimental data. For the purposes of model training and evaluation, we randomly divided these samples into training and testing sets at a 7:3 ratio.

3.1 Neural network model

Here, we present the model used to produce the results presented in the present work (though it should be noted that this is one of several candidate models being investigated for the purposes of improving MÖNCH's imaging performance for electron microscopy). This comprised a backbone and a feature model. The backbone model is implemented as a convolutional neural network (CNN) to process the detector readout. It consists of seven CNN layers with a fixed kernel size of 3 and 64 output channels, with the rectified linear unit as the activation function between each two CNN layers. The feature model, composed of one fully connected layer of 7 input and 32 output features and a tanh activation function, is designed to handle the extracted features listed in section 2. The outputs of these two sub-models are concatenated and then passed through another fully connected layer of 96 input features and 2 output features to predict the impact point.

3.2 Training setup and data augmentation

We use the stochastic gradient descent optimizer implemented in PyTorch with specific hyperparameters: a learning rate of 0.001, a momentum of 0.9, and a weight decay of 0.0001. Our batch size is set to 128, and the training process spans 100 epochs. The loss function is the mean squared error between predictions and ground truth over two dimensions.

To enhance the model's generalization capabilities, we employ several data augmentation techniques. We apply flip and rotation operations to clusters, effectively increasing the size of our training datasets by a factor of eight. Additional Gaussian noise is introduced to the ground truth of training sets with a standard deviation of 1.0 as label smoothing to prevent overconfidence of the model. Lastly, for the simulated clusters, we introduce extra noise with a standard deviation of 0.18 keV, the noise level of the mode in which MÖNCH was operated during experimental measurements. This step aims to mimic a more realistic detector response. The computational load is handled by a single NVIDIA RTX 4090 GPU and a 16-core CPU, with training times of 9 hours for the 4M simulation samples and 3 hours for the 1.4M experimental samples.

3.3 Deep learning results

For the model trained with simulated data, we observe that the loss curves of both the training and testing sets (figure 5(a)) converge at the end of the training process. In figure 5(b), the histogram of the residual in the x direction between the true and predicted impact positions indicates an RMSE of 0.47 pixels. A mean centering zero indicates an unbiased prediction. Notably, the RMSE is 3.6 times better than the result obtained via the charge centroid method. The histogram is further divided based on whether electrons fully deposit their energy in the sensor or not.[3] MÖNCH displays a better spatial resolution for fully absorbed electrons at 200 keV than backscattered ones, of which the impact points are difficult to distinguish from exit points. Results in the y direction are consistent.

Figures 5(c) and 5(d) represent the results of the experiment-based model training in the same form. The spatial resolution is 0.60 pixels, three times better than the charge centroid method. The mean is also centering zero. On the other hand, using the simulation-based model to predict impact points for the experimental samples gives a spatial resolution of 0.70 pixels. This disparity in performance can be attributed to the imperfect detector calibration and detector simulation. For experimental results, the distinction between fully absorbed and backscattered electrons is less pronounced because of the imperfect calibration. It fits our expectation that the spatial resolution of the experiment-based model is better than that of the simulation-based model for the experimental samples, while the simulation-based model also brought a notable improvement over the charge centroid method.

3.4 First experimental proof of principle

Operating the microscope in TEM mode with the MÖNCH 03 prototype, we conducted imaging of a sample consisting of gold nanoparticles on a continuous carbon membrane with 200 keV electrons.

---

[3]Electrons may backscatter from the sensor, with this occurring for 15 % of the incident electrons when they have an energy of 200 keV. The probability that a 200 keV electron penetrates the 320 um silicon sensor is vanishingly small, however.
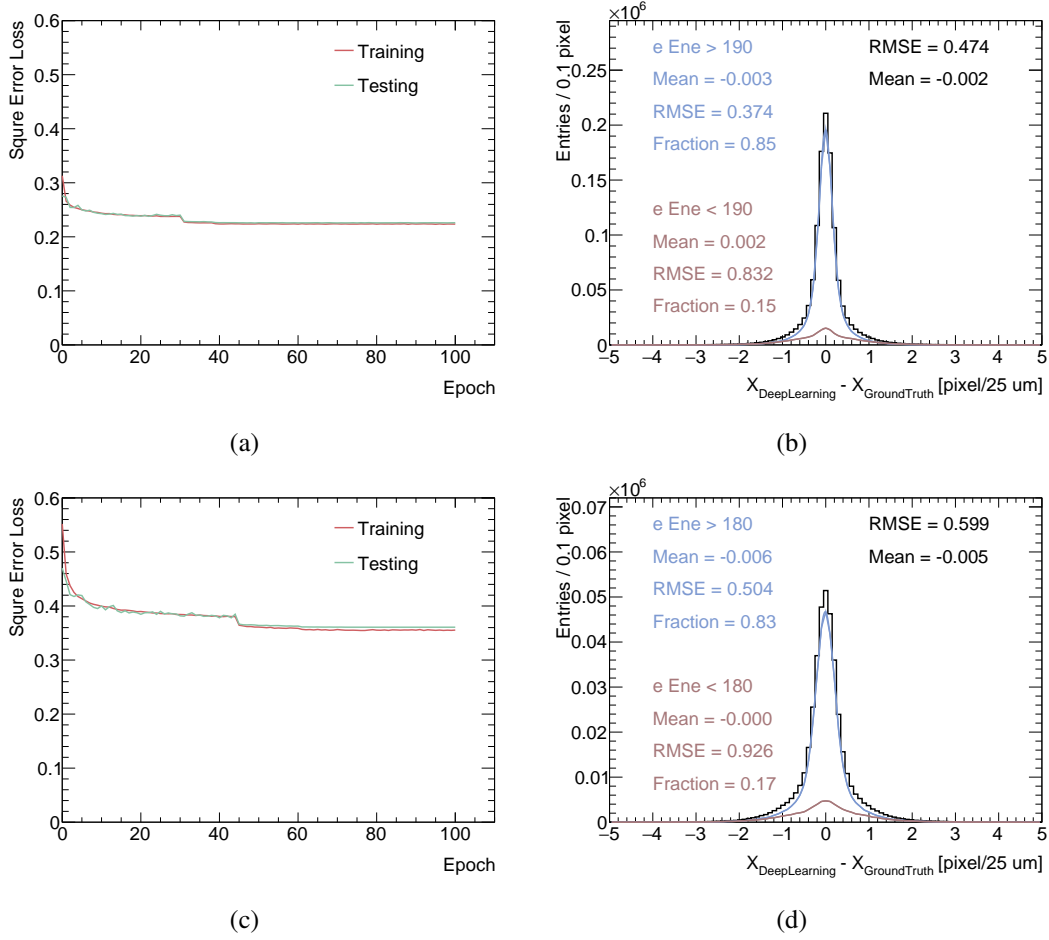
Figure 5. (a) Loss curves for the training and testing sets for the simulation-based model training. (b) Residual between the ground truth impact position and the predicted position via the simulation-based model. Sub-histograms of fully deposited and backscattered electrons are overlaid with the mean, RMSE, and fraction of electrons noted in the figure. The energy thresholds to identify backscattered electrons are 95% and 90% of the incident energy for simulations and measurements respectively. (c) Loss curves for the training and testing sets for the experiment-based model training. (d) Residual between the ground truth impact position and the predicted position via the experiment-based model.

The microscope's magnification was set to 150 kX so that one MÖNCH pixel corresponded to approximately 1 Å in the sample plane. The data analysis procedures remained the same as those described above, followed by the charge centroid method and deep learning inference using the experiment-based model. We acquired super-resolution and dose-fractionated images and applied motion correction using cross-correlation in frequency space. Figures 6(a) and 6(b) present mean-normalized images formed via the charge centroid method and deep learning inference respectively. Figures 6(c) and (d) are the corresponding power spectrum after mean-normalization. With the deep learning method, we obtain the diffraction ring corresponding to 2.35 Å, the spacing of the {1 1 1} gold atomic planes.
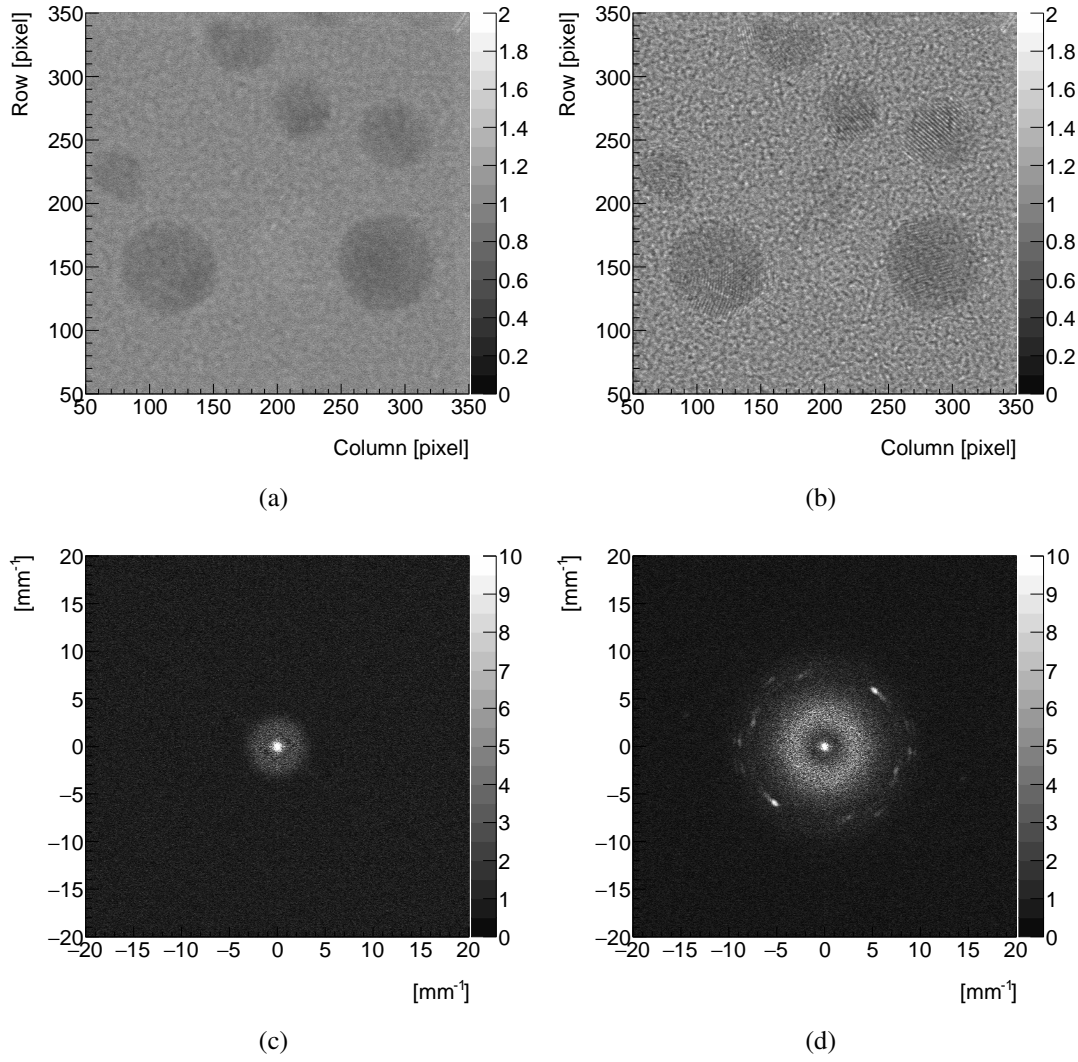
(a)

(b)

(c)

(d)

Figure 6. (a) and (b): images of gold nanoparticles on continuous carbon membrane captured by the MÖNCH 03 at 200 keV reconstructed after mean-normalization, using the charge centroid method and the experiment-based model respectively. Motion correction was performed using cross-correlation in Fourier space. The scratch patterns in the center and on the top right are due to the tracks of defective pixels generated from the motion correction. (c) and (d): power spectrum of images (a) and (b) after mean-normalization. The first diffraction ring, corresponding to 2.35 Å resolution, is prominently visible with deep learning.

4 Conclusion and outlook

In this work, we have successfully developed a deep learning approach for localizing 200 keV electron impacts on MÖNCH. Through experiment-based training enabled by the novel acquisition setup we have developed, we have achieved a sub-pixel spatial resolution of 0.60 pixels for 200 keV electrons, a threefold enhancement compared to the conventional charge centroid method, exploiting the lateral spread of signal in the sensor due to the stochastic way in which incident electrons deposit their energy as well as the diffusion of charge carriers. We have demonstrated the potential of the 25 μm MÖNCH detector for high-resolution, low-dose imaging applications in TEM by enhancing the quality of the images of the gold nanoparticles.

In our ongoing research, we are committed to extending our deep learning approach to lower-energy electrons, which are of increasing interest in the field of cryo-electron microscopy [10]. Lower-energy electrons offer advantages such as a better image contrast for a given electron dose and lower operational costs. For our MÖNCH detector, smaller cluster sizes due to the reduced multiple scattering ranges hold promise for an improved spatial resolution and a higher practical flux rate. We plan to measure the detector's modulation transfer function and detective quantum efficiency across a range of primary electron energies to characterize the detector more fully. More meticulous determination of the detector's gain and pedestal will improve the data quality and the performance of deep learning models. Developing the detector simulations and narrowing the gap between experiment-based and simulation-based models will enable us to disseminate the simulation-based models for different scenarios and explore different detector designs to further improve spatial resolutions achievable with HPDs. With the final design specification of MÖNCH including a frame rate of 6 kHz, a 1200×800 pixel matrix, capacity for tiling on 3 sides, we envision that MÖNCH, combined with super-resolution, will be a highly promising detector for electron microscopy studies using electrons with energies up to 200 keV.

Acknowledgments

Author contributions. Conceptualization: E.M., A.B., E.F., B.B.H.; Simulation: E.F., X.X.; Data acquisition: E.P., E.M., E.F., X.X., K.A.P.; Data analysis and training samples: X.X.; Deep learning model development: X.X., L.B.F., C.R., B.B.H.; Results interpretation: X.X., E.F., E.P., K.A.P.; Writing — original draft: X.X.; Writing — review & editing: K.A.P., E.F., A.B., E.P., E.M., B.B.H., L.B.F., C.R., X.X.; Supervision: B.B.H., E.M., E.F., E.P., A.B.; Funding acquisition: E.M., E.P., E.F.;

References

[1] W. Kühlbrandt, *The Resolution Revolution*, *Science* **343** (2014) 1443.

[2] A.R. Faruqi and G. McMullan, *Direct imaging detectors for electron microscopy*, *Nucl. Instrum. Meth. A* **878** (2018) 180.

[3] A.R. Faruqi and S. Subramaniam, *CCD detectors in high-resolution biological electron microscopy*, *Q. Rev. Biophys.* **33** (2000) 1.

[4] L. Fréchin et al., *High-resolution cryo-EM performance comparison of two latest-generation cryo electron microscopes on the human ribosome*, *J. Struct. Biol.* **215** (2023) 107905.

[5] A. Mozzanica et al., *Characterization results of the JUNGFRAU full scale readout ASIC*, 2016 *JINST* **11** C02047 [INSPIRE].

[6] E. Fröjdh et al., *Discrimination of Aluminum from Silicon by Electron Crystallography with the JUNGFRAU Detector*, *Crystals* **10** (2020) 1148.

[7] R. Ballabriga, M. Campbell and X. Llopart, *An introduction to the Medipix family ASICs*, *Radiat. Meas.* **136** (2020) 106271 [InSPIRE].

[8] D. Johnstone *Low-Dose Scanning Electron Diffraction Microscopy of Mechanochemically Nanostructured Pharmaceuticals*, *Microscopy Microanal.* **25** (2019) 1746.

[9] M. Ramilli et al., *Measurements with MÖNCH, a 25 µm pixel pitch hybrid pixel detector*, 2017 *JINST* **12** C01071 [InSPIRE].

[10] K. Naydenova et al., *CryoEM at 100 keV: a demonstration and prospects*, *IUCrJ* **6** (2019) 1086.

[11] J.P. van Schayck et al., *Sub-pixel electron detection using a convolutional neural network*, *Ultramicroscopy* **218** (2020) 113091.

[12] B. Eckert et al., *Electron Imaging Reconstruction for Pixelated Semiconductor Tracking Detectors in Transmission Electron Microscopes Using the Approach of Convolutional Neural Networks*, *IEEE Trans. Nucl. Sci.* **69** (2022) 1014.

[13] GEANT4 collaboration, *GEANT4 — a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250 [InSPIRE].

[14] A. Schübel et al., *A Geant4 based framework for pixel detector simulation*, 2014 *JINST* **9** C12018 [InSPIRE].

[15] A. Paszke et al., *PyTorch: An Imperative Style, High-Performance Deep Learning Library*, *Adv. Neural Inf. Process. Syst.* **32** (2019) 8024 [`arXiv:1912.01703`] [InSPIRE].