MATERIALS SCIENCE

Self-supervised machine learning framework for high-throughput electron microscopy

Joodeok Kim[1,2]†, Jinho Rhee[1,2]†, Sungsu Kang[1,2], Mingyu Jung[3], Jihoon Kim[1,2], Miji Jeon[3], Junsun Park[3], Jimin Ham[4], Byung Hyo Kim[5], Won Chul Lee[4]*, Soung-Hun Roh[3]*, Jungwon Park[1,2,6,7,8]*

Transmission electron microscopy (TEM) is a crucial analysis method in materials science and structural biology, as it offers a high spatiotemporal resolution for structural characterization and reveals structure-property relationships and structural dynamics at atomic and molecular levels. Despite technical advancements in EM, the nature of the electron beam makes the EM imaging inherently detrimental to materials even in low-dose applications. We introduce SHINE, the Self-supervised High-throughput Image denoising Neural network for Electron microscopy, accelerating minimally invasive low-dose EM of diverse material systems. SHINE uses only a single raw image dataset with intrinsic noise, which makes it suitable for limited-size datasets and eliminates the need for expensive ground-truth training datasets. We quantitatively demonstrate that SHINE overcomes the information limit in the current high-resolution TEM, in situ liquid phase TEM, time-series scanning TEM, and cryo-TEM, facilitating unambiguous high-throughput structure analysis across a broad spectrum of materials.

INTRODUCTION

Transmission electron microscopy (TEM) is a key analysis method in materials science and structural biology. It offers high spatiotemporal resolution in structural characterization, which helps uncover the structure-property relationship and structural dynamics at the atomic and molecular levels. With a resolution approaching the scale of individual atoms, structures of materials can be understood at a level where constituent atoms are thoroughly located. For example, high-resolution TEM and scanning TEM (STEM) can be used to locate catalytic single atoms on support materials (*1*, *2*). Further, (S)TEM can be applied to the structure analysis of two-dimensional (2D) materials (*3*–*6*) and thin-film semiconductors (*7*–*10*) to realize atomic-level precision. Beyond the conventional 2D microscopic analysis, serial imaging based on (S)TEM combined with computed structure reconstruction allows characterizing material structures in three dimensions. The positions of constituent atoms in single nanoparticles, exposed to vacuum or immersed in solution, can now be determined by electron tomography (*11*, *12*) and Brownian one-particle reconstruction (*13*, *14*). The conformational degeneracy of membrane proteins can indeed be addressed without crystallization, which enables a profound understanding of protein dynamics (*15*, *16*) and drug-target interactions (*17*, *18*). Cryo–electron tomography (cryo-ET) with computational reconstruction is widely exploited for tomography wherein 3D structures and positions of individual protein molecules are resolved inside a single cell (*11*, *18*–*20*).

For further pushing the information limit that can be gained by EM, recent efforts have focused on developing microscopes with enhanced beam control and detector efficiency. The incorporation of spherical (*Cs*) and chromatic (*Cc*) aberration correctors, which can achieve *Cs* and *Cc* less than 0.01 mm, has helped improve the coherence of the electron beam (*21*). Moreover, the implementation of a direct electron detector (DED) has replaced conventional charge-coupled devices (CCDs), elevating the quantum efficiency of electron detection to nearly 0.9 with less than 10% coincident loss. This enables nonlethal low-dose imaging of beam-sensitive materials through electron counting mode (*22*–*24*). Despite these advancements, imaging using EM suffers from intrinsic limitations caused by the inherent incoherence in the irradiating electron beam and the process of generating microscopic images as electrons pass through the camera and are detected. Low-dose imaging produces Poisson noise from the limited number of electrons that are incident to the detector. In addition, pixelated detectors limit pixel size, coincident loss, and correction of sample drift that occurs during acquisition time, which results in blur and noise in the acquired image. Discarding such inherent information-limiting sources is a critical requirement for overcoming the current information limit of (S)TEM and accomplishing damage-free EM imaging. Given the state-of-the-art instrumental development of the modern EM, such a goal can be achieved by processing image data to deconvolute object-oriented information from intrinsic noise using a software-based approach.

Thus far, diverse denoising methods have been developed to address both instrumental and physical noise sources embedded in TEM images and to isolate meaningful information (*25*, *26*). Traditional denoising techniques such as Gaussian filtering, bandpass filtering, Wiener filtering, and block-matching and 3D filtering (BM3D) denoiser often compromise on the details and result in a loss of information in the images (*27*). More recently, the use of neural network–based techniques for image denoising has shown a promising potential in producing valuable information even when images are corrupted by noise (*26*, *28*, *29*). However, most of these developments use training

[1]School of Chemical and Biological Engineering, Institute of Chemical Processes, Seoul National University, Seoul 08826, Republic of Korea. [2]Center for Nanoparticle Research, Institute for Basic Science (IBS), Seoul 08826, Republic of Korea. [3]School of Biological Sciences, Institute of Molecular Biology and Genetics, Seoul National University, Seoul 08826, Republic of Korea. [4]Department of Mechanical Engineering, BK21FOUR ERICA-ACE Center, Hanyang University, Ansan, Gyeonggi 15588, Republic of Korea. [5]Department of Materials Science and Engineering, Soongsil University, Seoul 06978, Republic of Korea. [6]Institute of Engineering Research, College of Engineering, Seoul National University, Seoul 08826, Republic of Korea. [7]Advanced Institute of Convergence Technology, Seoul National University, Suwon 16229, Republic of Korea. [8]Hyundai Motor Group-Seoul National University (HMG-SNU) Joint Battery Research Center (JBRC), Seoul National University, Seoul 08826, Republic of Korea.
*Corresponding author. Email: wonchullee@hanyang.ac.kr (W.C.L.); shroh@snu.ac.kr (S.-H.R.); jungwonpark@snu.ac.kr (J.P.)
†These authors contributed equally to this work.

processes that require noise-free reference images, which are difficult to obtain in practice because of the reduced signal from low dose rates and sample drift that occurs during image acquisition. Owing to the case-dependent noise characteristics of training datasets, it is also impractically difficult to acquire a substantial number of reference images for each distinct imaging condition for the training network. Therefore, a method that can denoise TEM images in the absence of reference images and make the whole process self-consistent from raw data acquisition to the extraction of information is highly desired.

In this paper, we introduce SHINE (Self-supervised High-throughput Image denoising Neural network for Electron microscopy), a self-supervised denoising method that requires only a single raw image dataset with intrinsic noise, making it suitable for limited-size datasets and removing the extra effort to collect training data. Our approach aims to remove the noise patterns that occur in real-world TEM observations, such as variations in magnification, signal dispersion across pixel boundaries, and sample drift during imaging. We address the issue of variance in noise patterns across different TEM observations by calculating the correlation of noisy images and exploiting the calculated noise pattern for denoising. We evaluated the network through simulated datasets, and the results confirm its superior performance in terms of peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) compared to that of previous self-supervised denoising techniques. Subsequently, we applied SHINE to various types of advanced time-series TEM observations, such as liquid cell in situ TEM and time-series STEM, and the results indicate that it can generate high-quality images with improved contrast. Further, SHINE is applied to cryo-ET of a whole cell, where the quality of the tilt-series micrographs and tomographic reconstruction of the cell are enhanced. A series of demonstrations confirms that SHINE eliminates the need for expensive ground-truth training datasets, enabling high-throughput EM, and has a great potential for accelerating damage-free low-dose (S)TEM.

RESULTS

Blind-spot video denoiser for enhancing the quality of TEM images

We aim to develop a denoising approach that can be used over a wide range of in situ EM including liquid cell in situ TEM, time-series STEM, and cryo-ET (Fig. 1A). Existing denoising techniques based on neural networks focused on creating a denoising network for a variety of images (*26*, *28*–*30*). In most of these techniques, a large dataset containing both clean and noisy images is required to train the denoising network for handling various forms of noise. Consequently, the primary difficulty in designing a real-world image denoising model is the associated high cost and challenge involved in acquiring a noisy reference image dataset for training. This problem becomes more serious in in situ TEM. In in situ TEM observation, obtaining a clean, noise-free image for training is difficult because of the limited dose rates and reactions of objects with incident electrons.

Understanding the dependency between the noise patterns of nearby pixels in each image can help improve the performance of a neural network used for denoising. When using a CCD or DED in the linear acquisition mode, the received electron signal has a diffusive distribution, forming a signal that is larger than the physical size of the sensor pixels. Each electron signal not only affects the pixel that receives the incident electron but also affects nearby pixels, introducing the dependency of noise statistics between these pixels. In addition, in STEM imaging, raster scanning and sample drift can perturb noise-pixel dependency generated during imaging (Fig. 1A).

We design our neural network to reflect the basic principles of how incoming electrons are sensed by the detector to generate (S) TEM images [see Fig. 1 (B and C), Materials and Methods, and fig. S1]. In addition, it does not require clean images for training, but rather uses raw images, which have intrinsic noise for both training and denoising. A blind-spot network is adopted for realizing a self-supervised denoising network through dilated convolution, as demonstrated in Fig. 1C (*31*). Although the blind-spot denoiser was originally developed for single-image denoising with known noise statistics, they are typically unknown in most in situ TEM experiments (*32*). To deal with time-series TEM images, we use L2 loss, or mean squared error (MSE) loss, during neural network training and add temporal information from four nearby frames to the network inputs (see Fig. 1B and Materials and Methods). To account for the correlation between neighboring pixels, dependent on the choice of imaging mode and condition, we estimate the blind-spot size for images with unknown noise statistics by calculating the correlation of raw images (Materials and Methods). SHINE with the optimized size of blind spot for different imaging is referred with respect to the size of the blind spot such as 1×1, 3×3, and 5×5. To consider the realistic image detection process used for breaking the correlation between noise and pixels, we create a larger blind-spot version by altering the dilation of the convolutions. For example, the network with a blind-spot size of 3×3 cannot infer the exact pixel value along with the surrounding eight pixels, causing it to remove the nine-pixel level structured noise (referred to as the 3×3 blind spot; Fig. 1C). In every training process, we generate a unique denoising model for each dataset using the same dataset for both training and inference (see Fig. 1C and Materials and Methods). All training and validation losses are shown in fig. S2.

Validation of the denoising framework based on the simulated dataset

We generated simulated TEM datasets for two types of liquid-phase TEM (LPTEM), i.e., silicon nitride and graphene liquid cell TEM, to evaluate the effectiveness of our denoising technique. We created noisy in situ TEM time-series images that represent the growth of Au nanoparticles, showing Brownian motion, in a silicon nitride liquid cell using the LPTEMsimulator (Fig. 1D, top, and Materials and Methods) (*25*). We then calculated the 2D spatial correlation of noisy simulation to estimate the unknown noise statistics of raw datasets (Materials and Methods). The resulting 2D heatmap of noise statistics indicates that the noise in the noisy simulation exhibits a high correlation within a four-pixel distance in each direction (Fig. 1D, bottom). This suggests that, in this dataset, noise would affect nearby 9×9 pixels, implying that using a 9×9 blind spot would yield the optimal denoising. Denoising simulated images allowed us to compare the results from SHINE (including 1×1, 3×3, and 9×9 blind-spot version) with the ground-truth images and the ones obtained from several other denoising methods, including the conventional BM3D denoiser (*27*), deep learning–based self-supervised single image denoising method Noise2Void (N2V) (*32*), deep learning–based self-supervised video denoiser (UDVD) (*33*), and modified UDVD incorporating statistics of noise (UDVD*). Visually, our 9×9

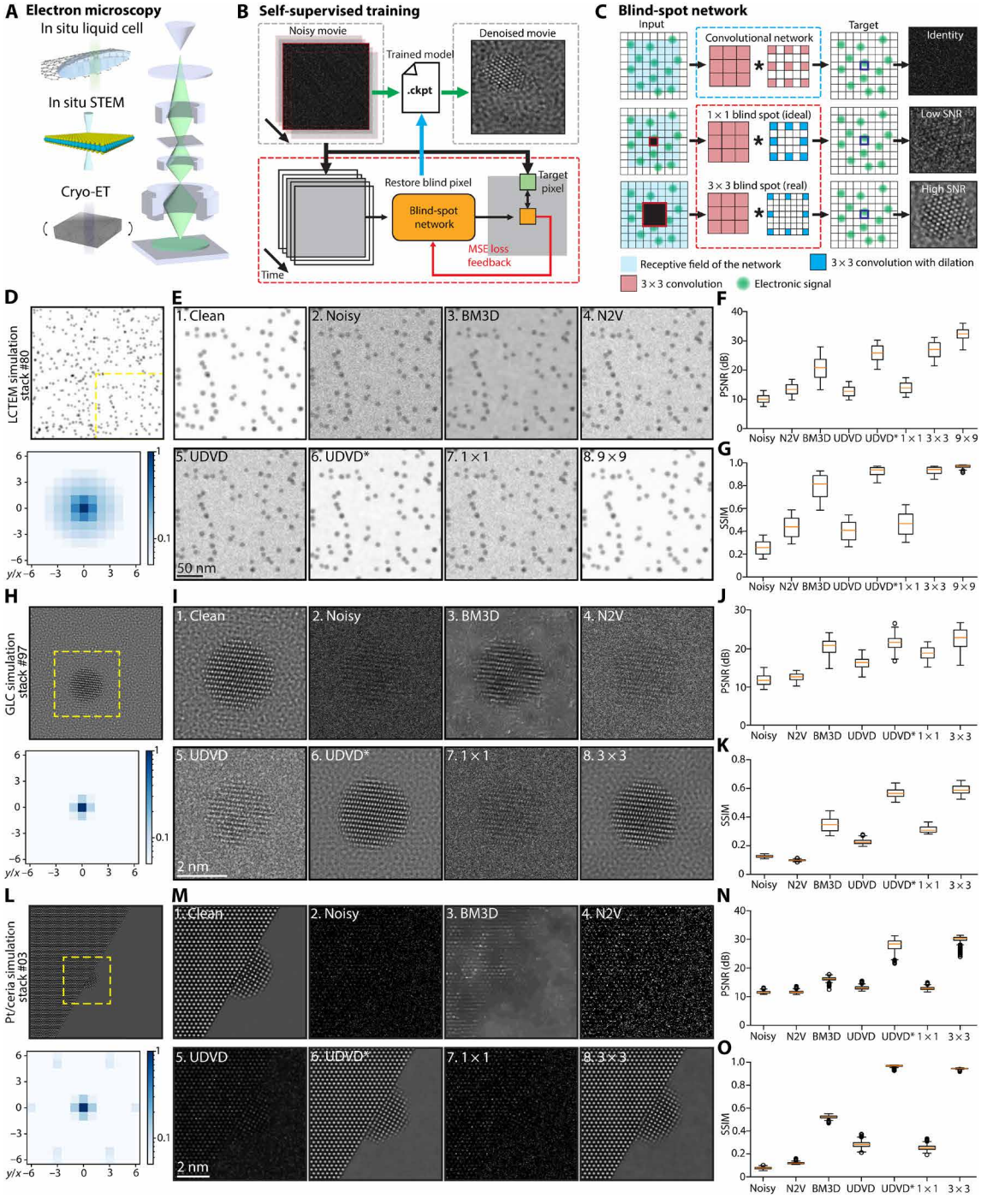

Fig. 1. Structure of SHINE and validation with the simulated TEM datasets. (**A**) Schematic of various TEM-based observations. (**B**) Schematic of a single-data video denoising procedure used throughout this work. SHINE contains a blind spot and cannot see the exact pixel value of noisy images for interference. (**C**) Simple diagram of generating a blind spot in the network using a flexible blind-spot denoising network with dilated convolution. The red-highlighted box shows a blind spot area that cannot be seen by the network. (**D**) Clean TEM simulation of Au nanoparticles in the silicon nitride liquid cell (top). 2D spatial correlation heatmap showing noise statistics of a noise-simulated TEM image of Au nanoparticles (bottom). LCTEM, liquid cell TEM. (**E**) Comparison of denoising performance using various methods. (**F** and **G**) Comparison of image quality metrics with PSNR and SSIM between the denoised results and the original image. (**H**) Clean TEM simulation of Pt nanoparticle in the graphene liquid cell (top). 2D spatial correlation heatmap showing noise statistics of a noise-simulated TEM image of Pt nanoparticle (bottom). (**I**) Comparison of the denoising performance using various methods. (**J** and **K**) Comparison of image quality metrics with PSNR and SSIM between denoised results and the original image. (**L**) Clean TEM simulation of Pt nanoparticle on ceria support (top). 2D spatial correlation heatmap showing noise statistics of a noise-simulated TEM image of Pt nanoparticle (bottom). (**M**) Comparison of the denoising performance using various methods. (**N** and **O**) Comparison of image quality metrics with PSNR and SSIM between the denoised results and the original image.

blind-spot denoiser outperforms all other denoising methods (Fig. 1E and movie S1), whereas the N2V, UDVD, and 1 × 1 blind-spot denoisers do not show substantial visual improvement because of the overfitting of noise. This confirms the superiority of SHINE over previously reported methods. Further, we also conducted a quantitative evaluation of image quality metrics based on PSNR and SSIM. The results showed that our 9 × 9 blind-spot network provided better PSNR and SSIM values than those of the other methods, with an almost 20-dB improvement in PSNR and a 0.6 difference in SSIM value (Fig. 1, F and G). These results confirm the successful restoration of the original image, especially for images with light noise, where the conventional BM3D denoiser and modified UDVD* performed better than the N2V, UDVD, and 1 × 1 blind-spot denoisers, highlighting the importance of incorporating a large blind spot to avoid noise overfitting.

In another validation process, we used a simulation to model the Brownian movement of a Pt nanoparticle, including both rotational and transitional motions, within a graphene liquid cell (Fig. 1H and movie S2). With a dose rate of 50 $e^{-}/Å^2$ and a pixel size of 0.217 Å per pixel, we generated TEM simulations using clTEM (*34*), as described in Materials and Methods. After imaging, we estimated the blind-spot size for noisy simulation from a 2D correlation heatmap and chose to use a 3 × 3 blind spot (Fig. 1H). We demonstrated that SHINE was suitable for highly noisy images, surpassing the N2V and UDVD methods in terms of PSNR and SSIM. Our network recovers atomic columns previously affected by noise, which results in an image almost indistinguishable from the ground truth. In addition, our 3 × 3 blind-spot network outperforms other methods, achieving an almost 10-dB enhancement in the PSNR and a 0.4 enhancement in SSIM value (Fig. 1, J and K).

Further, this pattern is reflected in the simulated image of Pt nanoparticle on ceria support, wherein our framework successfully restores the shape of atoms in the noisy simulation (Fig. 1, L to O, and movie S3). Our denoising approach can achieve better visual quality and high-quality metric values, which confirms that the size of the blind spot is key to producing a high-quality image using the self-supervised denoising framework. SHINE provides significant enhancements in both visual quality and SNRs, performing well on simulated datasets. Our neural network–based denoising approach showed notable improvements in both PSNR and SSIM in the case of Pt nanoparticle on ceria support, where it achieved a 20-dB improvement in PSNR and a 0.9 improvement in SSIM value.

Applications of blind-spot video denoiser for in situ (S)TEM

We applied SHINE for multiple in situ TEM observations of Au nanoparticles dissolving in a graphene liquid cell (Fig. 2A, fig. S3, and movies S4 and S5). For real dataset applications, we adhered to the same procedure as for simulated datasets. We first estimated noise statistics of a real noisy image, indicating a blind-spot size of 3× 3 (Fig. 2B). We then trained the network using only 100 frames from the full dataset and compared the results with networks trained and inferred using N2V, the 1 × 1 blind spot, and the 3 × 3 blind spot to demonstrate the capability of our framework with a limited data size (Fig. 2C). Our 3 × 3 blind-spot network outperforms other denoising methods by removing noise effectively, which results in an image with significantly enhanced visual quality. The denoised image exhibits high contrast and is practically noise-free, containing similar information to the average of five frames (Fig. 2C). Our approach enhances the intensities of the lattice peaks in the fast Fourier transform (FFT) image, unlike the denoised images produced by N2V and 1 × 1 blind-spot methods. It exhibits artifact rings at the single-pixel level, which affect the visual quality and frequency peaks of the nanoparticles. However, an image denoised by the 3 × 3 blind spot shows only aliasing patterns in the FFT image (fig. S4). The boundary and lattice structures of nanoparticles are visible in the denoised image, which enables the tracking of the nanoparticle size and crystal structure changes under a low dose with fast scan rate imaging (Fig. 2D and fig. S4). The improved contrast in both normal and FFT images allows the successful measurement of the particle size and angular orientation of (111) lattice peaks in every single frame, as captured in Fig. 2E. Nanoparticles rapidly change in size and angle after 175 ms, which indicates a stepwise process of dissolution. The denoising framework enhances the clarity and reduces noise in the output, potentially boosting the temporal resolution of the analysis. Our denoised images show a significant improvement in SNR, with an almost 15-dB increase compared to that of the original TEM images, as assessed by our SNR analysis (Materials and Methods).

We conducted an in situ experiment using an electrochemical silicon nitride liquid cell to observe the hydrogen evolution reaction (HER) from MoS_2 monolayer catalysts (Fig. 3A, Materials and Methods, and movie S6) (*35*). We followed a similar training approach and trained our network with N2V, the 1 × 1 blind spot, and the 3 × 3 blind spot, the latter of which is determined from the calculated noise pattern, to denoise noisy time-series images obtained from in situ TEM imaging (Fig. 3, B and C, and Materials and Methods). Our denoising approach consistently outperforms previous techniques, achieving improved contrast and better performance with the real dataset. The enhanced contrast facilitates the differentiation between the regions of generating hydrogen bubbles and MoS_2 and the identification of bubbles located beneath the MoS_2 layers. This improved clarity allows us to detect morphological changes without averaging multiple frames (Fig. 3D). To demonstrate this contrast difference, we conducted line intensity analysis in areas 1 and 2 in Fig. 3C, corresponding to a region including carbon, bubbles, and MoS_2 films and a region including MoS_2 films and silicon nitride (Fig. 3, E and F). The 3 × 3 blind-spot denoised images show significant improvements that can resolve differences in MoS_2 films, bubbles, and silicon nitride, whereas the other images cannot.

We evaluate the performance of the denoising network for image datasets collected from a carbon film liquid cell TEM wherein the contrast is affected heavily by noise from the thick window material and liquid molecules (Fig. 4A and movie S7) (*36*). We encapsulated a $Pd(acac)_2$ precursor into the carbon film liquid cell and used an electron beam to initiate nanoparticle nucleation and growth during in situ TEM imaging. To train and infer the clean image, we used a dataset of only 100 frames subsampled with the same frame interval. Similar to the previous experiment, we first computed spatial correlation to determine the blind-spot size (3 × 3), and then we used N2V, the 1 × 1 blind spot, and the 3 × 3 blind spot to train the network (Fig. 4, B and C). Our denoising network produced improved visual quality and enhanced the contrast of Pd nanoparticles in TEM images. We applied a simple binarization filter for separating nanoparticles from the background, to validate whether particle-oriented information can be separated. The 3 × 3 blind-spot denoised image shows an almost perfect separation of the nanoparticles and the background, which makes it easier to analyze nanoparticle growth. This clear separation is evident in the histogram. The noisy,

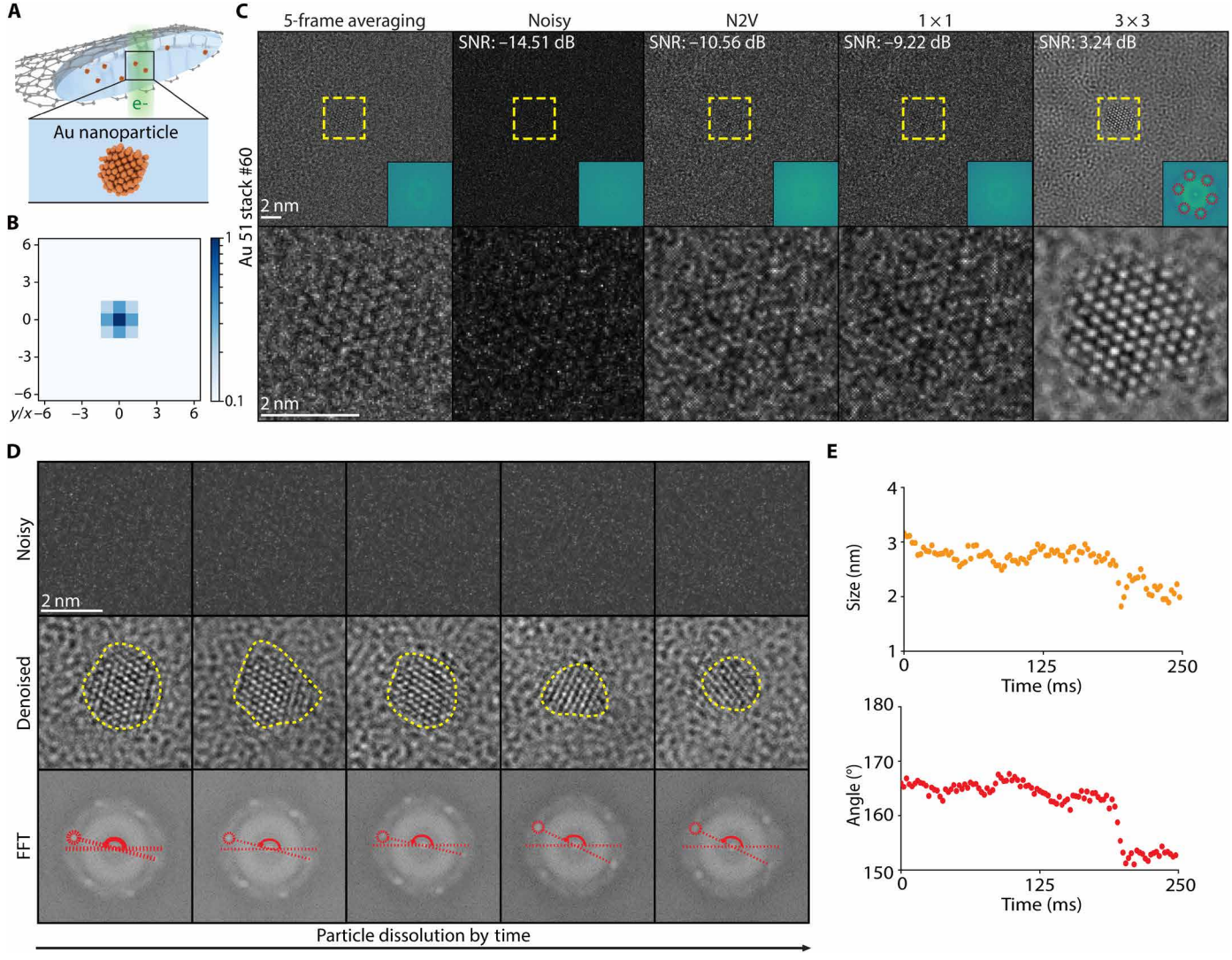

Fig. 2. Denoising of nanoparticle dynamics in liquid cell TEM using SHINE. (**A**) Schematic of the graphene liquid cell with the encapsulated Au nanoparticles. (**B**) 2D spatial correlation heatmap showing noise statistics of a noisy TEM image of Au nanoparticles in graphene liquid cell. (**C**) Comparison of visual quality between multiple denoising methods (top). Magnified atomic column of a nanoparticle (bottom) and FFT images of results (insets). (**D**) Time-series images of Au nanoparticle dissolution in the graphene liquid cell. The yellow line shows the measured projected area of the Au nanoparticles. Red markers show the angular orientation of the (111) lattice of the nanoparticle. (**E**) Time-dependent change of the nanoparticle size and orientation, which shows that the nanoparticle is rotating and dissolving in the liquid medium.

N2V, and 1 × 1 blind-spot denoised images displayed a Gaussian distribution in the intensity histogram, which leads to the difficulty in separating the particles and background (Fig. 4D). However, the 3 × 3 blind-spot denoised image exhibits two peaks that can be deconvoluted into the nanoparticles and the background of the carbon film and encapsulated liquid. This binarized image enables the facile measurements of particle size, SD, and number of nanoparticles in the field of view (Fig. 4, E to G).

The method presented in this study can also be applied to sequential STEM observations. For example, we captured the time-series STEM images of Pt atoms deposited on a MoS_2 film with two different pixel dwell times, i.e., 5 and 10 μs (Fig. 5A, Materials and Methods, and movies S8 and S9). The asymmetric 2D noise heatmap, caused by specimen drift, indicates the use of a 5 × 5 blind spot for the denoising network to fully address the noise pattern (Fig. 5B). We successfully trained SHINE with only 20 frames. The 5 × 5 denoised image demonstrated superior quality compared to that of the 3 × 3 denoised image (Fig. 5C) because the drift influenced both the immediate neighboring pixels and the second nearest pixels, although the two versions of the denoised image show the temporal movement of Pt atoms over the MoS_2 film over time and the detailed atomic positions in the film. Despite the short illumination time, it enables the automatic selection of atomic positions in all frames and tracking time-resolved movements of Pt atoms through ImageJ's automatic atomic selection with local maxima picking (Fig. 5, D and E, and Materials and Methods) (*37*).

Application of blind-spot video denoiser for cryo-ET

We prepared a cryo–electron tomogram dataset using whole-cell tomography techniques (Fig. 6A, Materials and Methods, and fig. S5).

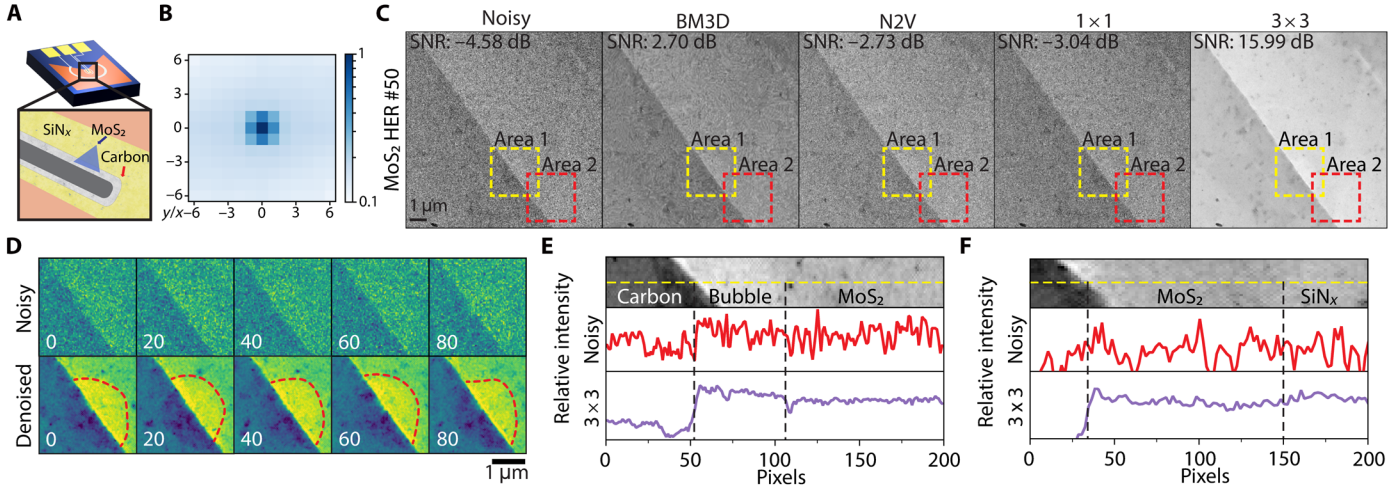

Fig. 3. Denoising of electrochemical reaction in liquid cell TEM using SHINE. (**A**) Schematic of hydrogen evolution reaction (HER) in the liquid cell TEM. (**B**) 2D spatial correlation heatmap showing noise statistics of a noisy TEM image of HER on MoS_2 film. (**C**) Comparison of visual quality between multiple denoising results. (**D**) Time-series images of bubble underneath the MoS_2 film. (**E**) Line intensity profile of area 1, the area including the carbon, bubble, and MoS_2 film. (**F**) Line intensity profile of area 2, the area including the MoS_2 film and silicon nitride window.

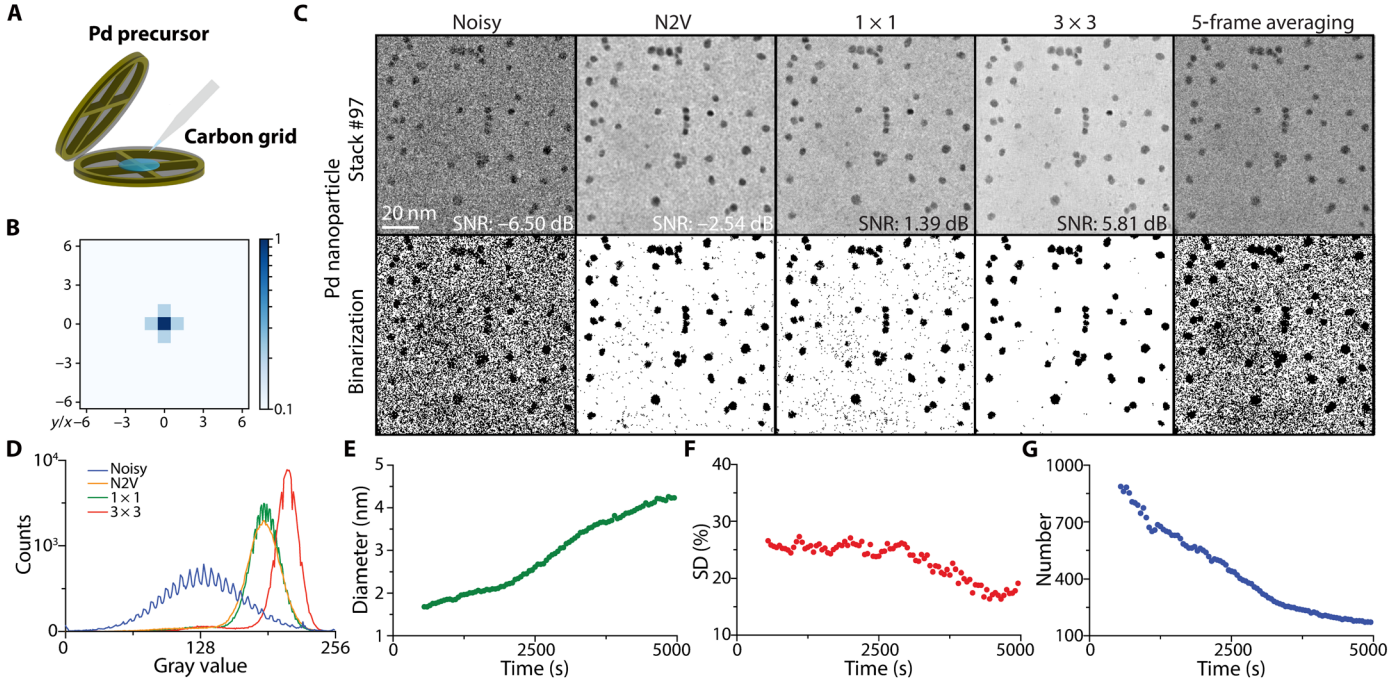

Fig. 4. Improving image quality for time-series TEM using SHINE. (**A**) Schematic of the preparation of the carbon film liquid cell with encapsulated Pd nanoparticles. (**B**) 2D spatial correlation heatmap showing noise statistics of a noisy TEM image of Pd nanoparticles in carbon film liquid cell. (**C**) Comparison of the visual result of the original and denoised images. Example of input noisy image, denoised image using N2V, our 1 × 1, and 3 × 3 blind-spot method, and five frame–averaged image (top). Simple binarization-filtered image to separate particle and background (bottom). (**D**) Histogram of the intensity values of the image in (C). (**E**) Time-dependent average size of the nanoparticles. (**F**) Time-dependent SD changes in the nanoparticles in images. (**G**) The number of nanoparticles in the field of view.

We first estimated the statistics of noise and decided to use the 5 × 5 blind spot for denoising with SHINE (Fig. 6B). For comparison, we denoised the same dataset using Topaz-Denoise AI (*38*). We observe a substantial enhancement in both the quantitative and the perceptual quality of the tilt-series micrographs after denoising with our network (Fig. 6, C and D, and movie S10). It was possible to detect detailed features of the cell membrane and granular structures within the cell from only a single fraction of our denoised images, whereas Topaz-Denoise AI also improves the contrast but preserves fewer features (Fig. 6D).

Then, we conducted a tomographic reconstruction of the cell for comparison, as described in Materials and Methods. Our blind-spot denoiser enhanced the visibility of key structural features that include the cell membrane and inner structures of the mitochondria

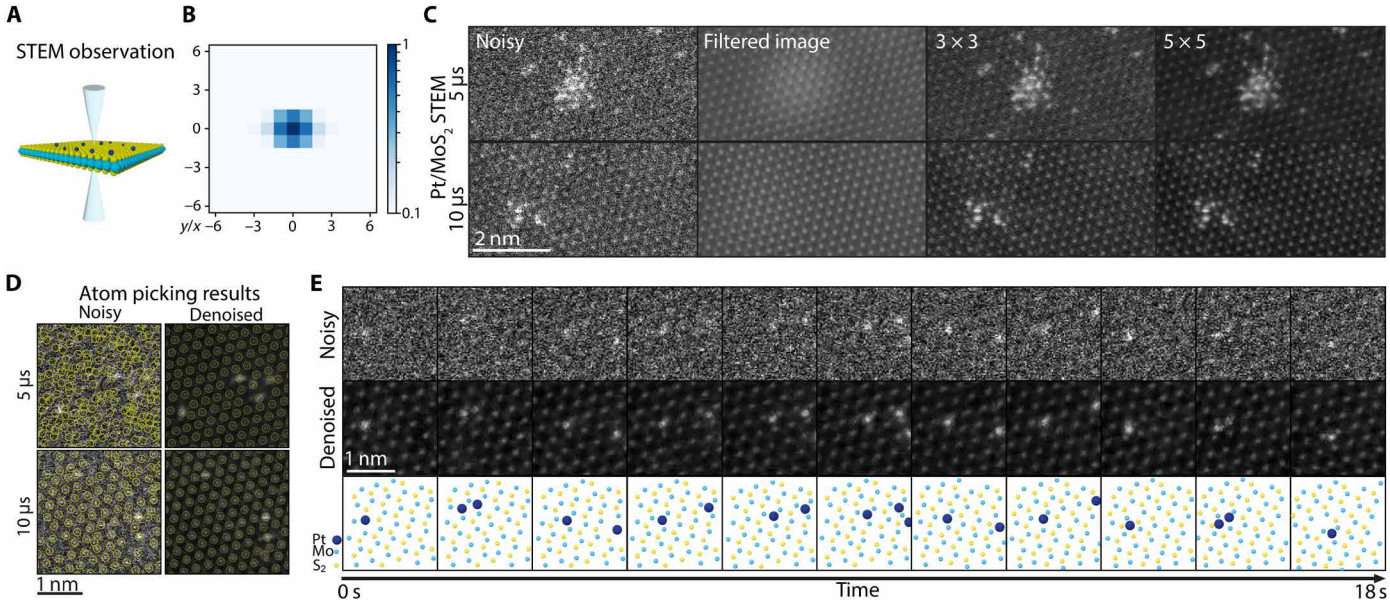

Fig. 5. Improving image quality for time-series STEM using SHINE. (**A**) Schematic of the in situ STEM observation of MoS_2. (**B**) 2D spatial correlation heatmap showing noise statistics of a noisy STEM image of Pt single atom over MoS_2 film. (**C**) Comparison of the original visual result and denoised results of multiple methods (spot filter, 3 × 3 blind spot, and 5 × 5 blind spot). (**D**) Comparison of the atomic picking result of the original and denoised image using local maxima. (**E**) Time-dependent position change of Pt atoms, which shows that the Pt atoms move over the MoS_2 film.

(Fig. 6E and movie S11). In addition, our tomogram from the first fraction image also displays improved contrast compared with the tomogram from the first fraction image denoised with Topaz or raw tomogram using full-dose imaging, demonstrating promising results for low-dose reconstruction.

Furthermore, we additionally denoised low-dose datasets with different characteristics to verify the network performance on damage-free imaging. We prepared another cryo–electron tomogram dataset with an electron dose of 3 $e^-/Å^2$ for each fractionated frame (details described in section S1). Denoised results demonstrate that our network performs effectively for low-dose imaging down to 0.1 $e^-/Å^2$ (fig. S6).

DISCUSSION

We introduce a high-throughput EM framework that uses a neural network with flexible blind-spot architecture. SHINE overcomes the negative effects of signal perturbations that occur commonly during real signal acquisition by nearby pixels. We demonstrated the efficacy of SHINE using both real-world and simulated TEM datasets and compared it to previous denoising methods using visual quality and image quality metrics. We achieved significant improvements in image quality by adjusting the size of the blind spot to match the noise characteristics of the images. Our quantitative analysis showed that SHINE outperformed the N2V and UDVD methods in PSNR and SSIM metrics. These results indicate that SHINE can be applied to different in situ TEM approaches, including liquid cell TEM, time-series STEM, and cryo-ET with minimal training data. SHINE enhances the accuracy of atomic column identification and nanoparticle size measurement in high-resolution TEM images while also restoring contrast affected by noise. In the field of STEM imaging, the technique substantially augments image quality, allowing for meticulous atomic-level characterization even with brief pixel dwell times. We observed that segmented images at substantially lower dose rates compared to the original still adequately discern detailed features of tomograms in whole-cell tomography. The improvements achieved using this approach provide increased confidence in measurements taken at low dose rates and have the potential to enhance temporal and spatial resolutions.

MATERIALS AND METHODS

Model architectures

In this study, we used our network with dilated convolutions based on an efficient blind-spot network (fig. S1) (*31*). The architecture was designed to expect input map features with five dimensions, which includes the exact input frame and four nearby frames. The exact input frame was processed with a 1 × 1 convolutional layer and 48 channels, whereas the nearby frames were treated with 3 × 3 convolutional layers and 16 channels. Our model included eight layers of convolution blocks, each containing two 3 × 3 convolutional layers. After each convolution block, the collected features were processed with 3 × 3 dilated convolutional layers for preserving the blind-spot constraint of the network and collected as the output feature.

Blind-spot self-supervised denoising framework and loss function

We implemented a mapping from $\boldsymbol{x}$ (time-series image) to $\boldsymbol{x}'$ (denoised image) in training our blind-spot CNN (convolutional neural network). In SHINE, we implemented the blind-spot receptive field for image denoising (*32*) while allowing the blind-spot network to have flexible blind-spot size, estimated from noise statistics in the

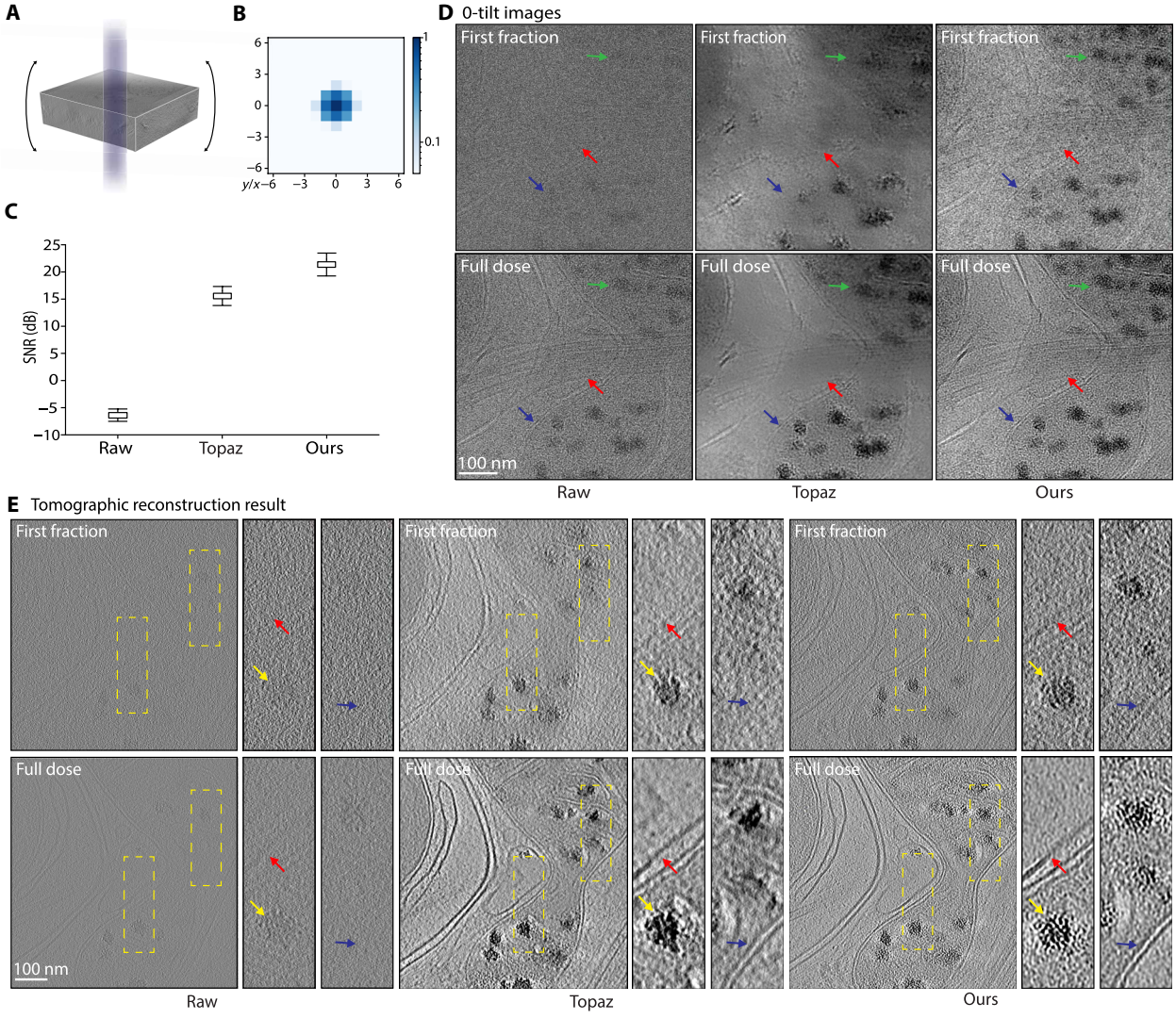

Fig. 6. Application of SHINE for cryo-ET. (**A**) Schematic of the cryo-ET tilt-series collection of the cell. (**B**) 2D spatial correlation heatmap showing noise statistics of a noisy cryo-ET tilt-series image of the cell. (**C**) Comparison of image quality metrics with SNR between denoised results and the original image. (**D**) 0°-tilt micrograph of the cryo-ET tilt-series image before denoising (raw), after denoising with Topaz-Denoise AI (Topaz), and with SHINE (ours). (**E**) Comparison of tomogram reconstruction results of the original and denoised images with a single fraction and full dose rate. SHINE reveals the detailed cell membrane and inner cell structure by reducing background noise.

noisy image. For a single pixel i in the input of each prediction $\boldsymbol{x}$, we made a blind spot by excluding the center pixel i and its nearby pixels in the receptive field $\boldsymbol{x}_{RF(i)}$ of our network to account for the noise patterns occurring in real-world TEM observations. Therefore, we can define our blind-spot CNN as the function

$$f\left[\boldsymbol{x}_{RF(i)};\theta\right]=\boldsymbol{x}_i' \tag{1}$$

where θ denotes the parameters of blind-spot CNN we want to train.

In our training scheme, we train a blind-spot network by minimizing the same empirical risk as in N2V

$$\arg\min_{\theta}\sum_i\sum_j L\left\{f\left[\boldsymbol{x}^j_{RF(i)};\theta\right],\boldsymbol{x}^j_i\right\} \tag{2}$$

$\boldsymbol{x}^j_i$ is used as the target and the input for the network; the blind-spot network effectively erases the information of the original pixel, which prevents the network from learning the identity. For TEM observations, we use the L2 loss function to minimize the difference between the original and generated images

$$L\left(\boldsymbol{x}^j_i,\boldsymbol{x}'^j_i\right)=\left(\boldsymbol{x}^j_i-\boldsymbol{x}'^j_i\right)^2 \tag{3}$$

Noise statistics estimation and blind-spot size determination

Knowing the noise statistics of the image is important because we can enhance the prediction quality of the blind-spot network by selecting a blind-spot size that matches the influence of the noise on nearby pixels. However, the statistics of noise generated through imaging are unknown, and the noise in different images cannot be regarded as identical. For each dataset, therefore, we estimated the noise statistics by calculating the 2D correlation heatmap between two raw images to determine the blind-spot size before denoising the images. Correlation of two image tensors, x and y, was calculated using the equation (*39*)

$$\text{Correlation} = \frac{\sum_i \left(x_i - \langle x_i \rangle\right)\left(y_i - \langle y_i \rangle\right)}{\left[\sum_i \left(x_i - \langle x_i \rangle\right)^2 \sum_i \left(y_i - \langle y_i \rangle\right)^2\right]^{1/2}} \quad (4)$$

where the angle brackets indicate averaging.

Before computing the 2D heatmap of correlation, we first selected a single raw image from the dataset and standardized the image tensor for further calculation. To understand how noise is spatially correlated, we shifted the image tensor by one pixel at a time in both the x and y directions. We then calculated the correlation between the original and the shifted tensor to create the 2D spatial correlation heatmap, which indicates how noise affects the neighboring pixels.

In addition, to validate our fundamental assumption that noise is spatially correlated across the image, we divided noisy images from different imaging modes in real-world TEM observations into 16 subsections. We then compared the estimated noise statistics of these subregions with the noise correlation of the whole frame used for blind-spot size determination (see section S2 and fig. S7). Homogeneous noise correlation across each image demonstrated that the noise generated during the TEM imaging is uniformly spatially correlated. This result also supports our determination of the blind-spot size based on the spatial correlation of noise in the entire image, as the noise shows a nearly invariable correlation throughout the image.

Ablation study on the number of input frames for denoising

In this study, we took five frames as an input for all datasets, using various blind-spot sizes determined by the noise statistics of each dataset. By incorporating adjacent frames to estimate the output pixel in the blind-spot network, we accounted for spatial-temporal neighborhoods to improve prediction accuracy. Previous video denoising networks have also demonstrated improved performance when incorporating temporal information from neighboring frames, suggesting the use of five input frames (*33*, *40*). To examine the effect of the number of input frames in our network, we conducted experiments using different numbers of input frames for denoising simulated datasets (fig. S8).

PSNR and SSIM of the denoised dataset in fig. S8 (A and B) with five input frames exhibit the most improvement, whereas PSNR and SSIM values for the denoised datasets in fig. S8 (C and D) reveal that using three or more input frames, thereby including temporal information, yields better performance than using a single frame, which is consistent with previous studies. However, no considerable improvement was observed for denoised datasets in fig. S8 (E and F). These results demonstrate that although the performance of the network increases when using the temporal information, the optimal number for input frames may vary depending on the datasets. In addition, the decrease in training efficiency caused by too many input frames must be taken into consideration. Therefore, with the consideration of performance and training efficiency, we used five input frames for denoising our datasets.

Training details

We used the noisy training dataset as a testing dataset without any preprocessing for evaluating the denoising framework. During training, we extract a randomly cropped image with 256 × 256 pixels in every frame to prevent the network from fully remembering noise patterns. Furthermore, we used data augmentation while training, and we randomly rotated each image by 90° and mirrored it vertically and horizontally. We use different models trained with each dataset for denoising the datasets.

We used a batch size of 16 for our blind-spot network training with an RAdam (*41*) optimizer with a learning rate of 0.001. We used a warm-up and cosine annealing learning rate scheduler and adjusted the number of epochs depending on the training dataset size (2000 epochs for all datasets, except 100 epochs for the cryo-ET datasets). Our code is implemented with PyTorch Lightning (*42*, *43*), and it is trained with a single NVIDIA RTX 3090. We used the PyTorch native automatic mixed precision for training and inferencing to improve speed. Processing time including training time for 100 frames with 512 × 512 pixels took 1.5 hours to obtain denoised results for every model. Most training with small datasets rapidly fits and starts overfitting after several epochs. We excluded two images from the dataset as validation datasets, and we used the best validation loss model for denoising images.

For cryo-ET images, we generated preprocessed patches cropped with 1024 × 1024 pixels with overlapping 256 pixels with patches for the tilt images that are 4096 × 4096 pixels. A total of 50% of preprocessed patches are used for the training model. We used a batch size of 8, and the training took about 2.5 and 1.5 hours for inference.

For N2V training, we used CSBDeep implementation as in the original paper with the default parameters, i.e., batch size of 128 with 64 × 64 pixels per input patch with a CSBDeep default learning rate schedule.

Estimating the SNR in a real experiment dataset

We quantified the SNR of the real dataset using two independent measurements of pseudosame images. On the basis of this approach

$$\text{SNR} = 10\log_{10}\left(\frac{p}{1-p}\right) \quad (5)$$

where p denotes the cross-correlation coefficient between two measurements (*39*). We used this to measure the SNR of denoising by calculating the correlation between two nearby frames. The image drift and signal difference between the two frames led to a false calculation of the micrograph. However, despite this disadvantage, it was difficult to assert image quality without a reference image, and, therefore, we use this SNR value as a quality metric of images without the ground truth for in situ datasets.

For cryo-ET images, we calculated the correlation between two summed frames from the tilt-series micrograph using the same equation as in the Topaz-Denoise AI paper (*38*). We generated odd and even sums for every 10 frames in the tilt-series stack using "SplitSum" of MotionCor2 (*44*) and used these two frames to calculate the correlation for each tilt angle.

Au nanoparticle growth simulation in the silicon nitride liquid cell

A silicon nitride liquid cell time-series image was simulated using LPTEMsimulator (*25*) using MATLAB. We set the silicon nitride window thickness to 50 nm and the water liquid to a thickness of 50 nm. A total of 350 nanoparticles were simulated in the liquid cell with a diffusivity of 0.2 pixel^2 per frame and an initial diameter of 6 nm with linear growth. Imaging conditions included a 20 $e^-/Å^2$ dose rate per image, and the size of the image was 512 × 512 with a pixel size of 1.045 nm. The noise in the image was generated with a detector conversion factor of 3.47.

Pt nanoparticle rotation simulation in the graphene liquid cell

Graphene liquid cell time-series images of the Pt nanoparticle were simulated using clTEM with Python (*13*, *34*). In the liquid cell, single-layered graphene sheets were located on the top and bottom of the water layer to a thickness of 10 nm. The atomic coordinates of water were obtained by molecular dynamics simulation using the large-scale atomic/molecular massively parallel simulator package. A CHARMM force field and TIP3P water model were used for the simulation. A total of 10,968 water molecules were created randomly inside the 57.28 Å–by–57.28 Å–by 100.0 Å simulation box. After the implementation of energy minimization to stabilize the system, the NVT ensemble with a 1-fs time step and Nosé-Hoover thermostat was performed for 100 ps; water trajectories were recorded every 100 fs. Atom coordinates of the 2.5-nm-sized Pt nanoparticle were generated with known lattice parameters. In every frame, the Pt nanoparticle rotates randomly between a delay of 2 and 10 frames with thermal vibration in both the angle and position. The TEM images of the atomic models including nanoparticle, water, and graphene were obtained via clTEM simulation with a pixel size of 0.217 Å and 512 × 512 sized images. In this case, we simulate an image with a dose rate set at 50 $e^-/Å^2$ per frame with in-focus on a C3 aberration of −0.01 mm, a C5 aberration of 3 mm, and an objective aperture of 30 mrad. The K2 detector parameter was used for the simulation. Each simulated TEM image included a 2.5-nm-sized Pt nanoparticle and a set of geometries of the water molecules.

Pt nanoparticle migration simulation on ceria support

Pt nanoparticle randomly moving on ceria support images was simulated using clTEM with Python (*34*). The Pt nanoparticle loaded on ceria support under vacuum conditions was located on the field of view, and we randomly shifted the atomic coordinates of the Pt nanoparticle, while adding drift in every frame. TEM images of atomic models including Pt nanoparticle and ceria were obtained via clTEM simulation with a pixel size of 0.332 Å with 512 × 512 sized images. In this case, we simulated an image with a dose rate set at 2 $e^-/Å^2$ per frame with 7 nm over focus with a C3 aberration of −0.01 mm, a C5 aberration of 3 mm, and an objective aperture of 30 mrad. The K2 detector parameter was used for the simulation.

Au nanoparticle dissolution data acquisition in the graphene liquid cell

Procedures for synthesizing few-layer graphene by the low-pressure chemical vapor deposition method and preparing graphene-coated TEM grids were described elsewhere (*45*). Tris-HCl solution (10 mM), 20 mM $FeCl_3$ aqueous solution, and Au nanoparticle solution were prepared separately to prepare the nanoparticle solution sample. HCl (18 μl; 37 wt % in water) was added to the 10 ml of 20 mM $FeCl_3$ aqueous solution for suppressing the hydrolysis of $FeCl_3$ (*46*). The Au nanoparticle solution was purchased from Nanoprobes Inc. and contained 75 μM Au nanoparticles with a diameter of approximately 1.4 nm, 50 mM sodium phosphate buffer, and 0.5 M NaCl. $FeCl_3$ solution (50 μl), 75 μl of the tris-HCl solution, and 5 μl of the Au nanoparticle solution were mixed for preparing the graphene liquid cell (GLC); 0.4 μl of the mixed solution was sandwiched between the two graphene-coated TEM grids. The as-prepared GLC was left in an ambient condition for approximately 10 min before imaging to ensure complete adherence between the two graphene sheets. In situ TEM movies of Au nanoparticles dissolving in the GLC were obtained at a temporal resolution of 2.5 ms using the TEAM 1 microscope in the National Center for Electron Microscopy. The microscope was equipped with a postspecimen geometric and chromatic aberration corrector and K2 IS DED (Gatan). Images with 1920 × 1728 pixels and a 0.317-Å pixel resolution were acquired at a dose rate of ~15 $e^-/Å^2$ per frame at an acceleration voltage of 300 kV. Before processing, each Au nanoparticle movie was cropped to focus on a region of interest with 512 × 512 pixels.

Pd nanoparticle growth data acquisition in the carbon film liquid cell

For the in situ TEM experiment, we prepared a liquid cell compatible with normal TEM holders. A precursor solution was prepared by mixing 10 mg of $Pd(acac)_2$ and 0.1 ml of oleylamine into 0.9 ml of dichlorobenzene; 0.2 μl of the solution was loaded over the grid. We created the liquid cell by sandwiching two carbon film copper TEM grids together. The growth of the nanoparticles was monitored using JEM-2100F (JEOL Ltd.) operated at the acceleration voltage of 200 kV, equipped with an UltraScan 1000XP CCD detector (Gatan). The electron dose rate during in situ LPTEM analysis was maintained at 1000 $e^-/Å^2 \cdot s$ for reducing the precursor to induce nanoparticle growth. In situ LPTEM videos were recorded with a frame rate of 2 frames per second (fps) with 0.434-nm pixel resolution, and we subsampled 100 frames for training and denoising from the total of 4000 frames.

Pt single atom over MoS_2 film data acquisition

For the atomic dispersion of Pt on the MoS_2 monolayer film, the chemical vapor deposition–grown MoS_2 film on the sapphire substrate (1 cm^2) was immersed in the 2 ml of tetraammineplatinum(II) nitrate precursor solution [1 mM, dissolved in deionized (DI) water] for 1 hour. After Pt deposition, Pt-MoS_2 film was washed three times with DI water and blown by N_2. For STEM imaging, Pt-MoS_2 film was transferred onto the TEM grid by a polymer-assisted wet transfer process. High-angle annular dark-field–STEM imaging, operated at an accelerating voltage of 200 kV, was performed with a JEM-ARM200F (JEOL Ltd.) equipped with a spherical aberration corrector (probe corrector) and installed at the National Center for Inter-University Research Facility in Seoul National University. The STEM images were collected at a 0.133-Å pixel resolution with a convergence semiangle of 19 mrad and a collection semiangle of 68 to 280 mrad.

HER data acquisition for the MoS_2 film

Polystyrene (PS) solution (9 g of polystyrene beads in 100 ml of toluene) was deposited on MoS_2 monolayer flakes on a sapphire substrate using a spin coater (rotation speed: 1500 rpm, rotation time: 60 s). The substrate deposited by PS was baked for strong adhesion between PS and MoS_2 layers (at 90°C for 5 min). After immersion of the substrate in DI water, the PS and MoS_2 film was peeled off manually from the sapphire substrate. The stripped film was transferred to the working electrode (WE) of the microchip (top chip) and immersed in toluene for 1 hour to dissolve the PS film. After PS dissolution, the microchip with the as-transferred MoS_2 layers was rinsed with acetone and blown by high-purity N_2 gas. The liquid cell for LPTEM analysis was composed of a top chip (6 mm by 4.5 mm by 300 μm) and a bottom chip (2 mm by 2 mm by 300 μm). The top and bottom microchips included a 50-nm-thick electron transparent SiN_x window (window dimensions: 550 μm by

40 μm for the top and 550 μm by 50 μm for the bottom). The WE, counter electrode, and reference electrode were micropatterned on the top chip. Metal circuits were sealed with a spacer (SU-8, thickness: 500 nm). All microchips were purchased from Protochips. H_2SO_4 aqueous solution (0.1 nN) was saturated with Ar gas for a prolonged bubbling time and used as the electrolyte in the liquid cell. The electrolyte was injected into the liquid cell by a liquid flow system (injection speed: 120 μl hour^{-1}). Potentiostat (Reference 600+, Gamry) was connected to the LPTEM holder (Poseidon, Protochips) for electrochemical experiments. In situ LPTEM experiments were performed using JEM-2100F (JEOL Ltd.) operated at an acceleration voltage of 200 kV, equipped with an UltraScan 1000XP CCD detector (Gatan). The electron dose rate during in situ LPTEM analysis was maintained at <0.247 $e^-/Å^2 \cdot s$ to minimize beam-induced damage. In situ LPTEM videos were recorded at a 28.57-nm pixel resolution with a frame rate of 10 fps.

Cryo-ET data acquisition, processing, and tomogram reconstruction

A day before plating neurons, a 200-mesh carbon-coated finder gold grid (Quantifoil R2/1, LF) was ultraviolet sterilized for 1 hour and placed on a 35-mm glass bottom dish (P35G-1.5-14-C, MatTek) and treated with 150 μl of poly-D-lysine (0.1 mg/ml; P6282, Merck) solution overnight at room temperature. The dish was rinsed twice with phosphate-buffered saline before plating neurons. The neurons were harvested from Sprague Dawley rat embryos staged at E18. The cells were suspended in the culture medium comprising neurobasal medium (NbActiv4, BrainBits), 1% penicillin-streptomycin (P4333, Merck), and N-2 supplement (17502048, Thermo Fisher Scientific). They were plated on a 35-mm dish at a density of 400,000 cells/mm^2. The cells were grown at 37°C and 5% CO_2 conditions, and half of the culture medium was exchanged twice a week. Grids were blotted from the back side for 5 s using a sensor blotting option and plunged into liquid ethane with EM GP2 (Leica) and stored in LN_2 before imaging in Cryo-EM.

Cryo-ET data were acquired on a 200-kV Glacios cryo-TEM (Thermo Fisher Scientific) with a Falcon 4 DED detector (Thermo Fisher Scientific) using TOMO 4 (Thermo Fisher Scientific) software. Tilt series were acquired with a dose-symmetric scheme ranging from −40° to 60°, and defocus was set to −5 μm. Nominal magnification was ×28,000, and pixel size was 3.6 Å. The total dose per tilt series was 123 $e^-/Å^2$, with each stack of tilt series consisting of 10 frames.

The raw and denoised micrographs using SHINE were processed with the MotionCor2 program for local motion correction and dose weighting (*44*). The denoising process is performed before the motion correction; however, for the Topaz denoised dataset, we applied the motion correction first and then applied the denoising model, which is the same as the original Topaz AI paper.

Before tomogram reconstruction of processed datasets, the contrast transfer function estimation of the raw tilt series was performed by WARP (*47*). We used the IMOD software package for our methodology in the reconstruction of tomograms (*48*). The procedure for tomogram reconstruction using IMOD comprised several steps. Initially, we preprocessed the tilt-series images by importing them into IMOD, where we performed a coarse alignment using a cross-correlation method. Subsequently, we used patch tracing to generate a fiducial model with a patch size of 1240 for fine alignment. Following the alignment process, we generated an initial 3D tomogram with a binning size of 2 by applying a weighted back-projection algorithm to the aligned tilt series. Other parameters were set by default.

Supplementary Materials

The PDF file includes:
Sections S1 and S2
Figs. S1 to S8
Legends for movies S1 to S11
References

Other Supplementary Material for this manuscript includes the following:
Movies S1 to S11

REFERENCES AND NOTES

1. B. Qiao, A. Wang, X. Yang, L. F. Allard, Z. Jiang, Y. Cui, J. Liu, J. Li, T. Zhang, Single-atom catalysis of CO oxidation using Pt_1/FeO_x. *Nat. Chem.* **3**, 634–641 (2011).
2. N. Clark, D. J. Kelly, M. Zhou, Y. C. Zou, C. W. Myung, D. G. Hopkinson, C. Schran, A. Michaelides, R. Gorbachev, S. J. Haigh, Tracking single adatoms in liquid in a transmission electron microscope. *Nature* **609**, 942–947 (2022).
3. W. Wang, S. W. Chee, H. Yan, I. Erofeev, U. Mirsaidov, Growth dynamics of vertical and lateral layered double hydroxide nanosheets during electrodeposition. *Nano Lett.* **21**, 5977–5983 (2021).
4. J. Yang, Z. Zeng, J. Kang, S. Betzler, C. Czarnik, X. Zhang, C. Ophus, C. Yu, K. Bustillo, M. Pan, J. Qiu, L. W. Wang, H. Zheng, Formation of two-dimensional transition metal oxide nanosheets with nanoparticles as intermediates. *Nat. Mater.* **18**, 970–976 (2019).
5. P. Y. Huang, S. Kurasch, A. Srivastava, V. Skakalova, J. Kotakoski, A. V. Krasheninnikov, R. Hovden, Q. Mao, J. C. Meyer, J. Smet, D. A. Muller, U. Kaiser, Direct imaging of a two-dimensional silica glass on graphene. *Nano Lett.* **12**, 1081–1086 (2012).
6. H. M. Nilsson, L. de Knoop, J. Cumings, E. Olsson, Localized resistance measurements of wrinkled reduced graphene oxide using in-situ transmission electron microscopy. *Carbon* **113**, 340–345 (2017).
7. Y. Cheng, Z. Gao, K. H. Ye, H. W. Park, Y. Zheng, Y. Zheng, J. Gao, M. H. Park, J. H. Choi, K. H. Xue, C. S. Hwang, H. Lyu, Reversible transition between the polar and antipolar phases and its implications for wake-up and fatigue in HfO_2-based ferroelectric thin film. *Nat. Commun.* **13**, 645 (2022).
8. S. Kang, W.-S. Jang, A. N. Morozovska, O. Kwon, Y. Jin, Y.-H. Kim, H. Bae, C. Wang, S.-H. Yang, A. Belianinov, S. Randolph, E. A. Eliseev, L. Collins, Y. Park, S. Jo, M.-H. Jung, K.-J. Go, H. W. Cho, S.-Y. Choi, J. H. Jang, S. Kim, H. Y. Jeong, J. Lee, O. S. Ovchinnikova, J. Heo, S. V. Kalinin, Y.-M. Kim, Y. Kim, Highly enhanced ferroelectricity in HfO_2-based ferroelectric thin film by light ion bombardment. *Science* **376**, 731–738 (2022).
9. J. C. Ondry, J. P. Philbin, M. Lostica, E. Rabani, A. P. Alivisatos, Resilient pathways to atomic attachment of quantum dot dimers and artificial solids from faceted CdSe quantum dot building blocks. *ACS Nano* **13**, 12322–12344 (2019).
10. H. Ma, D. Kim, S. I. Park, B. K. Choi, G. Park, H. Baek, H. Lee, H. Kim, J. S. Yu, W. C. Lee, J. Park, J. Yang, Direct observation of off-stoichiometry-induced phase transformation of 2D CdSe quantum nanosheets. *Adv. Sci.* **10**, 2205690 (2023).
11. M. C. Scott, C. C. Chen, M. Mecklenburg, C. Zhu, R. Xu, P. Ercius, U. Dahmen, B. C. Regan, J. Miao, Electron tomography at 2.4-ångström resolution. *Nature* **483**, 444–447 (2012).
12. Y. Yang, J. Zhou, F. Zhu, Y. Yuan, D. J. Chang, D. S. Kim, M. Pham, A. Rana, X. Tian, Y. Yao, S. J. Osher, A. K. Schmid, L. Hu, P. Ercius, J. Miao, Determining the three-dimensional atomic structure of an amorphous solid. *Nature* **592**, 60–64 (2021).
13. B. H. Kim, J. Heo, S. Kim, C. F. Reboul, H. Chun, D. Kang, H. Bae, H. Hyun, J. Lim, H. Lee, B. Han, T. Hyeon, A. P. Alivisatos, P. Ercius, H. Elmlund, J. Park, Critical differences in 3D atomic structure of individual ligand-protected nanocrystals in solution. *Science* **368**, 60–67 (2020).
14. J. Park, H. Elmlund, P. Ercius, J. M. Yuk, D. T. Limmer, Q. Chen, K. Kim, S. H. Han, D. A. Weitz, A. Zettl, A. P. Alivisatos, 3D structure of individual nanocrystals in solution by electron microscopy. *Science* **349**, 290–295 (2015).
15. J. L. Parker, J. C. Deme, G. Kuteyi, Z. Wu, J. Huo, I. D. Goldman, R. J. Owens, P. C. Biggin, S. M. Lea, S. Newstead, Structural basis of antifolate recognition and transport by PCFT. *Nature* **595**, 130–134 (2021).
16. F. Lauber, J. C. Deme, S. M. Lea, B. C. Berks, Type 9 secretion system structures reveal a new protein transport mechanism. *Nature* **564**, 77–82 (2018).
17. J. P. Renaud, A. Chari, C. Ciferri, W. T. Liu, H. W. Rémigy, H. Stark, C. Wiesmann, Cryo-EM in drug discovery: Achievements, limitations and prospects. *Nat. Rev. Drug Discov.* **17**, 471–492 (2018).
18. J. L. Vilas, J. M. Carazo, C. O. S. Sorzano, Emerging themes in CryoEM—Single particle analysis image processing. *Chem. Rev.* **122**, 13915–13951 (2022).

19. F. Nudelman, K. Pieterse, A. George, P. H. H. Bomans, H. Friedrich, L. J. Brylka, P. A. J. Hilbers, G. De With, N. A. J. M. Sommerdijk, The role of collagen in bone apatite formation in the presence of hydroxyapatite nucleation inhibitors. *Nat. Mater.* **9**, 1004–1009 (2010).
20. E. M. Pouget, P. H. H. Bomans, J. A. C. M. Goos, P. M. Frederik, G. de With, N. A. J. M. Sommerdijk, The initial stages of template-controlled $CaCO_3$ formation revealed by cryo-TEM. *Science* **323**, 1455–1458 (2009).
21. N. de Jonge, L. Houben, R. E. Dunin-Borkowski, F. M. Ross, Resolution and aberration correction in liquid cell transmission electron microscopy. *Nat. Rev. Mater.* **4**, 61–78 (2019).
22. S. Turner, O. I. Lebedev, F. Schröder, D. Esken, R. A. Fischer, G. Van Tendeloo, Direct imaging of loaded metal–organic framework materials (Metal@MOF-5). *Chem. Mater.* **20**, 5622–5627 (2008).
23. Y. Zhu, J. Ciston, B. Zheng, X. Miao, C. Czarnik, Y. Pan, R. Sougrat, Z. Lai, C. E. Hsiung, K. Yao, I. Pinnau, M. Pan, Y. Han, Unravelling surface and interfacial structures of a metal-organic framework by transmission electron microscopy. *Nat. Mater.* **16**, 532–536 (2017).
24. M. U. Rothmann, J. S. Kim, J. Borchert, K. B. Lohmann, C. M. O'Leary, A. A. Sheader, L. Clark, H. J. Snaith, M. B. Johnston, P. D. Nellist, L. M. Herz, Atomic-scale microstructure of metal halide perovskite. *Science* **370**, eabb5940 (2020).
25. L. Yao, Z. Ou, B. Luo, C. Xu, Q. Chen, Machine learning to reveal nanoparticle dynamics from liquid-phase TEM videos. *ACS Cent. Sci.* **6**, 1421–1430 (2020).
26. J. L. Vincent, R. Manzorro, S. Mohan, B. Tang, D. Y. Sheth, E. P. Simoncelli, D. S. Matteson, C. Fernandez-Granda, P. A. Crozier, Developing and evaluating deep neural network-based denoising for nanoparticle TEM images with ultra-low signal-to-noise. *Microsc. Microanal.* **27**, 1431–1447 (2021).
27. K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising with block-matching and 3D filtering, paper presented at the SPIE Electronic Imaging: Algorithms and Systems V, San Jose, CA, 17 February 2006.
28. F. Wang, T. R. Henninen, D. Keller, R. Erni, Noise2Atom: Unsupervised denoising for scanning transmission electron microscopy images. *Appl. Microsc.* **50**, 23 (2020).
29. S. Mohan, R. Manzorro, J. L. Vincent, B. Tang, D. Y. Sheth, E. P. Simoncelli, D. S. Matteson, P. A. Crozier, C. Fernandez-Granda, Deep denoising for scientific discovery: A case study in electron microscopy. *IEEE Trans. Comput. Imaging.* **8**, 585–597 (2022).
30. K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, Beyond a Gaussian Denoiser: Residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.* **26**, 3142–3155 (2017).
31. D. Honzátko, S. A. Bigdeli, E. Türetken, L. A. Dunbar, Efficient blind-spot neural network architecture for image denoising, in *Proceedings of 2020 7th Swiss Conference on Data Science (SDS)* (IEEE, 2020), pp. 59–60.
32. A. Krull, T. Buchholz, F. Jug, Noise2Void - Learning denoising from single noisy images, in *2017 IEEE Conference on Computer Vision and Pattern Recognition* (CVPR, 2017), pp. 2129–2137.
33. D. Sheth, S. Mohan, J. L. Vincent, R. Manzorro, P. A. Crozier, M. M. Khapra, E. P. Simoncelli, C. Fernandez-Granda, Unsupervised deep video denoising, in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (IEEE, 2021), pp. 1759–1768.
34. J. J. P. Peters, A fast frozen phonon algorithm using mixed static potentials. *Ultramicroscopy* **229**, 113364 (2021).
35. J. Kim, A. Park, J. Kim, S. J. Kwak, J. Y. Lee, D. Lee, S. Kim, B. K. Choi, S. Kim, J. Kwag, Y. Kim, S. Jeon, W. C. Lee, T. Hyeon, C. H. Lee, W. B. Lee, J. Park, Observation of H_2 evolution and electrolyte diffusion on MoS_2 monolayer by in situ liquid-phase transmission electron microscopy. *Adv. Mater.* **34**, 2206066 (2022).
36. C. Zhu, S. Liang, E. Song, Y. Zhou, W. Wang, F. Shan, Y. Shi, C. Hao, K. Yin, T. Zhang, J. Liu, H. Zheng, L. Sun, In-situ liquid cell transmission electron microscopy investigation on oriented attachment of gold nanoparticles. *Nat. Commun.* **9**, 421 (2018).
37. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
38. T. Bepler, K. Kelley, A. J. Noble, B. Berger, Topaz-Denoise: General deep denoising models for cryoEM and cryoET. *Nat. Commun.* **11**, 5208 (2020).
39. J. Frank, L. Al-Ali, Signal-to-noise ratio of electron micrographs obtained by cross correlation. *Nature* **256**, 376–379 (1975).
40. M. Tassano, G. France, J. Delon, T. Veit, FastDVDnet: Towards real-time deep video denoising without flow estimation, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (CVPR, 2020), pp. 1351–1360.
41. L. Liu, H. Jiang, P. He, W. Chen, X. Liu, J. Gao, J. Han, On the variance of the adaptive learning rate and beyond, in *8th International Conference on Learning Representations (ICLR)* (ICLR, 2020).
42. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, PyTorch: An imperative style, high-performance deep learning library, in *Advances in Neural Information Processing Systems (NeurIPS)*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. B. Fox, R. Garnett, Eds. (Curran Associates Inc., 2019), pp. 8024–8035.
43. W. Falcon, J. Borovec, A. Wälchli, N. Eggert, J. Schock, J. Jordan, N. Skafte, V. Bereznyuk, E. Harris, T. Murrell, P. Yu, S. Præsius, T. Addair, J. Zhong, D. Lipin, S. Uchida, S. Bapat, H. Schröter, B. Dayma, A. Karnachev, A. Kulkarni, S. Komatsu, J.-B. Schiratti, H. Mary, D. Byrne, C. Eyzaguirre, A. Bakhtin, PyTorchLightning/pytorch-lightning: 0.7.6 release. Zenodo, 10.5281/zenodo.3828935 (2020).
44. S. Q. Zheng, E. Palovcak, J. P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
45. S. Kang, J.-H. Kim, M. Lee, J. W. Yu, J. Kim, D. Kang, H. Baek, Y. Bae, B. H. Kim, S. Kang, S. Shim, S.-J. Park, W. B. Lee, T. Hyeon, J. Sung, J. Park, Real-space imaging of nanoparticle transport and interaction dynamics by graphene liquid cell TEM. *Sci. Adv.* **7**, 5419 (2021).
46. X. Ye, M. R. Jones, L. B. Frechette, Q. Chen, A. S. Powers, P. Ercius, G. Dunn, G. M. Rotskoff, S. C. Nguyen, V. P. Adiga, A. Zettl, E. Rabani, P. L. Geissler, A. P. Alivisatos, Single-particle mapping of nonequilibrium nanocrystal transformations. *Science* **354**, 874–877 (2016).
47. D. Tegunov, P. Cramer, Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* **16**, 1146–1152 (2019).
48. J. R. Kremer, D. N. Mastronarde, J. R. McIntosh, Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).
49. M. G. Campbell, D. Veesler, A. Cheng, C. S. Potter, B. Carragher, 2.8 Å resolution reconstruction of the *Thermoplasma acidophilum* 20S proteasome using cryo-electron microscopy. *eLife* **4**, e06380 (2015).
50. H. Iqbal, PlotNeuralNet: Release v1.0.0. Zenodo, 10.5281/zenodo.2526396 (2018).

Acknowledgments

Funding: Joodeok Kim, J.R., Jihoon Kim., S.K., and Jungwon Park acknowledge the Institute for Basic Science (grants IBS-R006-D1). Jungwon Park acknowledges the National Research Foundation of Korea (NRF) grant, funded by the Korean government (MSIT) (RS-2023-00283902 and RS-2024-00408823). Data acquisition and analysis from liquid-phase EM are supported by the Samsung Research Funding & Incubation Center of Samsung Electronics under project number SSTF-BA2302-06. S.-H.R. acknowledges the financial support from the Creative-Pioneering Researchers Program through Seoul National University and the NRF grant funded by the Korean government (NRF-2020R1A5A1018081 and 2021M3A9I4021220). J.H. and W.C.L. acknowledge financial supports from the NRF grant funded by the Korean government (MSIT) (RS-2024-00341798 and 2022R1A4A3031263). B.H.K. acknowledges the NRF (grant no. NRF-2021R1C1C1014339). **Author contributions:** W.C.L., S.-H.R., and Jungwon Park supervised the research. Jungwon Park and Joodeok Kim conceptualized and designed the research. Joodeok Kim, S.K., M.Ju., Jihoon Kim, and M.Je. performed the experiment. Joodeok Kim and M.Je. analyzed the data. Joodeok Kim wrote the code and performed the TEM simulation. Joodeok Kim, J.R., S.K., M.Ju., Jihoon Kim, M.Je., Junsun Park, J.H., B.H.K., W.C.L., S.-H.R., and Jungwon Park were involved in the discussion. Joodeok Kim, J.R., S.K., M.Ju., B.H.K., W.C.L., S.-H.R., and Jungwon Park wrote the manuscript. All authors read and reviewed the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Our source code and datasets are available at https://zenodo.org/records/14922032 (Zenodo) and https://github.com/jungwonparklab/SHINE (GitHub repository).

Submitted 20 August 2024
Accepted 26 February 2025
Published 2 April 2025
10.1126/sciadv.ads5552