

PAPER • OPEN ACCESS

Machine-learning approach for quantified resolvability enhancement of low-dose STEM data

To cite this article: Laura Gambini *et al* 2023 *Mach. Learn.: Sci. Technol.* **4** 015025

View the [article online](#) for updates and enhancements.

You may also like

- [Adaptive learning rate clipping stabilizes learning](#)
Jeffrey M Ede and Richard Beanland
- [Adaptive partial scanning transmission electron microscopy with reinforcement learning](#)
Jeffrey M Ede
- [Deep learning in electron microscopy](#)
Jeffrey M Ede



PAPER

OPEN ACCESS

RECEIVED
11 January 2023ACCEPTED FOR PUBLICATION
10 February 2023PUBLISHED
22 February 2023

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Machine-learning approach for quantified resolvability enhancement of low-dose STEM data

Laura Gambini^{1,*} , Tiarnan Mullarkey^{1,2} , Lewys Jones^{1,2} and Stefano Sanvito¹ ¹ School of Physics, AMBER and CRANN Institute, Trinity College Dublin, Dublin 2, Ireland² Advanced Microscopy Laboratory, Centre for Research on Adaptive Nanostructures and Nanodevices (CRANN), Dublin 2, Ireland

* Author to whom any correspondence should be addressed.

E-mail: gambinil@tcd.ie**Keywords:** scanning transmission electron microscope, image denoising, Poisson noise, autoencoder

Abstract

High-resolution electron microscopy is achievable only when a high electron dose is employed, a practice that may cause damage to the specimen and, in general, affects the observation. This drawback sets some limitations on the range of applications of high-resolution electron microscopy. Our work proposes a strategy, based on machine learning, which enables a significant improvement in the quality of Scanning Transmission Electron Microscope images generated at low electron dose, strongly affected by Poisson noise. In particular, we develop an autoencoder, trained on a large database of images, which is thoroughly tested on both synthetic and actual microscopy data. The algorithm is demonstrated to drastically reduce the noise level and approach ground-truth precision over a broad range of electron beam intensities. Importantly, it does not require human data pre-processing or the explicit knowledge of the dose level employed and can run at a speed compatible with live data acquisition. Furthermore, a quantitative unbiased benchmarking protocol is proposed to compare different denoising workflows.

1. Introduction

Much progress has been made in the field of electron microscopy in the last decades. At present, aberration-corrected scanning transmission electron microscopes (STEMs) provide the highest resolution of all imaging instruments, below 0.1 nm, and allow to investigate the structure and chemical composition of materials at the atomic scale [1]. STEM operates by focusing a convergent electron beam to a small area and scanning it across the sample. The signal can then be detected either directly under the sample or at a wide angle from it, two modes known as bright and dark field. One of the major limitations of STEM is that atomic resolution is achievable only when the specimen is illuminated by a very intense electron beam. This maximizes the signal-to-noise ratio, but may damage the sample and the observation. In fact, the damage is a function of the electron dose, defined as the total number of electrons per unit area hitting the specimen, and it is caused by various energy-loss mechanisms. Among the most common ones, there is the so-called knock-on damage, where the atoms of the sample are displaced from their sites due to a transfer of momentum from the incident electrons [2]. In this condition the specimen under investigation changes in time as the measure progresses.

The sample integrity can be protected by decreasing the electron dose, but this leads to a deterioration of the image quality, thus reducing the chance to extract useful information from the data. The reason behind such loss of resolution can be ultimately identified in the presence of Poisson noise, which increases upon reducing the number of incident electrons [3], according to the relation,

$$f \propto \frac{1}{\sqrt{\rho}}, \quad (1)$$

where f is the noise and ρ the dose. In contrast to other types of signal distortions, such as Gaussian noise, scan noise and drift [4], Poisson noise is related to the quantized nature of the electron beam and therefore cannot be eliminated by improving the instrumentation or changing the working conditions. For instance, it is possible to eliminate Gaussian noise completely by replacing analog with digital data acquisition, a strategy that itself represents one of the latest electron microscopy advancements [5]. Another advantage of digital data acquisition is the ability to generate images, where the pixel intensity corresponds to the number of electrons detected at a given pixel, namely the pixel intensity is a directly observable quantity. In our opinion digital data acquisition represents the ultimate future of electron microscopy, a consideration that motivated our choice of considering only Poisson noise in this work.

Notably, noise removal is a common practice in image processing. However, most of the denoising techniques provide accurate results only in the case of additive noise, such as Gaussian. On the contrary, Poisson noise is signal-dependent and requires more advanced techniques [6].

One of the most commonly used strategies for noise removal in electron microscopy is the application of a smoothing Gaussian filter [7, 8]. However, as it will be shown in this work, such method can sometimes lead to results that are less precise than the original noisy version of the image. A more sophisticated technique is known as ‘block matching and 3D filtering’ [9]. In this case, images are decomposed into fragments, which are grouped by similarity, and then the fragments are passed through filters. Unfortunately, such denoising scheme assumes that the images to process correspond to a 2D periodic structure, an assumption potentially leading to artifacts, such as the inability to identify genuine vacant sites.

An alternative to these methods is provided by deep-learning techniques, which are becoming increasingly popular in the microscopy field across several applications [10]. The majority of the available denoising algorithms involve Gaussian noise only, so that they are useful only for analog data acquisition. State-of-the-art neural networks for Poisson-noise removal have been proposed [11]. These provide significant noise reduction, but their performance rapidly degrades at low doses. The reason behind such low-dose accuracy loss can be identified in the nature of the training set, made of simulated STEM images obtained by using a simple linear imaging model. In fact, in order to have a more realistic dataset, it is advisable to use simulation techniques that implement a multislice algorithm [12] or the Bloch wave method [13]. These quantum mechanical techniques employ a detailed description of the specimen and of the instrument settings to generate the images, meaning that they implement a more faithful simulation of an actual measurement. In general, electron microscope images do not follow a simple linear image model; therefore, any linear method cannot be quantitatively precise [13]. Furthermore, this state-of-the-art technique requires inputs from the users, who should decide whether or not to apply some level of up-sampling/down-sampling before processing the image.

In this manuscript we propose an alternative machine-learning architecture trained on a diverse dataset of synthetic images, which is unbiased against periodic structures and operates across a broad range of dose levels. In particular, we have implemented a denoising convolutional autoencoder, trained over pairs of ideal and corrupted STEM images, in which Poisson noise is applied. The ideal images are referred to as infinite-dose images and have been generated by using the Prismatic simulation software [3, 14]. This is a highly-optimized multi-GPU simulation package, based on the multi-slice algorithm. The advantage of using Prismatic in place of other simulation codes [13] lies in the possibility of using the PRISM (plane-wave reciprocal-space interpolated scattering matrix) algorithm, which delivers a significant acceleration in the simulations. This is a key feature considering the large amount of images needed for the training. Our training set then contains various materials and includes defective structures, for instance incorporating vacancies. The corrupted images are distributed across a broad range of dose settings. The reason for training the model on simulated data is twofold. Firstly, generating an experimental dataset is an extremely expensive and time-consuming endeavor, when compared to images obtained by simulation. Secondly, the training of the autoencoder requires the ground-truth (noiseless) images, which represent the algorithm target. Clearly, such ideal images cannot be obtained under real experimental working conditions.

Our model is then tested on both simulated and real data, obtained at different levels of dose and for various materials, not included in the training dataset. Finally, in order to properly analyze the results, a technique based on atom localization is developed. This allows one to quantify the improvement in the quality of information extracted from the images and to compare results processed with different denoising strategies. Such aspect is completely overlooked in the majority of the works dealing with STEM-image denoising, where the improvement of a method is usually assessed only through a qualitative visual comparison. Our work thus proposes a strategy for quantitative benchmarking different algorithms. Finally, it should be noted that our approach does not require any data pre-processing, since it predicts the actual

signal of a microscopy measurement. This allows the algorithm to be integrated into existing microscopes and be applied live during the data acquisition.

2. Model construction

2.1. Neural Network

The deep-learning method chosen for the denoising process is an artificial neural network known as autoencoder, traditionally used for dimensionality reduction and feature extraction [15]. This consists of two parts: an encoder that maps the input (in this case the noisy images) into a reduced space called the *latent space*, and a decoder that maps the latent space into the output (in this case the infinite-dose images). The architecture proposed here consists of ten layers in total, five for the encoder and five for the decoder, and the input is made of electron-dose 2D maps resolved over a 128×128 -pixel grid. The activation function is *ReLU* for each layer and the optimizer applied during the training is *Adam* [16]. Each convolutional layer is formed by 32 filters, except for the last layer, in which only one filter is used. This is because the output should have the same size as the input. The kernel size measures three pixels for both spatial directions and the *same* padding is imposed. Max pooling and up-sampling layers are exploited in order to reduce and increase the data dimension, respectively. For an input with size $(128, 128, 1)$, the latent space representation size is $(32, 32, 32)$. The model architecture was optimized by performing hyperparameters tuning over a validation set, additional information can be found in appendix A.

The definition of the loss function is particularly significant in constructing the autoencoder. The commonly used mean squared error (MSE) is not suitable for the training of the proposed dataset. In fact, in the standard MSE equal importance is given to both black (low intensity) and non-black (high intensity) pixels, even if the interest sits mainly with the non-black pixels, which indicate the presence of the atoms. Note that the importance of black and white pixels is reversed when the image is taken in the bright field. It should also be noted that Poisson noise varies according to the pixel intensity. Therefore, when a pixel is black no noise is detected and the corrupted and uncorrupted images are identical. For this reason, we employ a customized loss function, which gives more importance to the pixels that are not black in the original images contained in the training set. This loss function can be described as a weighted MSE (WMSE), with a 1000:1 weight ratio for non-black pixels. It can be expressed as

$$\text{WMSE} = \frac{1}{n} \sum_{i=1}^n w_i (\tilde{y}_i - y_i)^2 \quad (2)$$

$$w_i = \begin{cases} 1 & \text{if } i = \min(y); \\ 1000 & \text{otherwise.} \end{cases}$$

where n is the number of pixels in each image, w_i is the weight associated to the i th pixel, \tilde{y}_i the predicted value and y_i the true value.

2.2. Training set

The autoencoder is trained on a dataset made of about 27 000 simulated images, all generated by using the Prismatic software [3, 14].

Different materials have been considered in the creation of the training set, namely, graphene, graphite, GaAs, InAs, MoS₂, SrTiO₃, and Si. These include pristine and defective structures incorporating vacancies, generated across various imaging conditions. Details about the training set construction can be found in appendix B. All images have been cropped into 128×128 -pixel plots, which is the format used to construct the autoencoder. Note that, common real images are usually larger (at least 512×512 pixel), but dealing with reduced size allows one to increase the computational efficiency and to identify more details in the reconstructed data.

The pixel intensity of the images generated by Prismatic corresponds to the fractional intensity of the entire electron beam that is scattered to the specified STEM detector at a given pixel. This takes values comprised between 0 and 1, and usually corresponds to around 10%–15% of the entire beam intensity. Measuring the signal in fractional units does not allow us to directly retain information of the actual physical electron dose used. As such, before applying the Poisson noise, it is necessary to convert the pixel values into integers, representing the physical number of electrons at a given pixel. This can be obtained by multiplying the original pixel value by the total electron dose (in units of electrons per Å²) and then by multiplying by the

pixel size (in \AA^2). Such conversion is important to generate a training set with an intensity distribution compatible to that of typical experimental data. As a consequence, any type of pre-processing of the test data can be avoided, which is a crucial condition to use our machine-learning tool during the actual microscope real-time data acquisition. It is also important to remark that in doing so the pixels value, namely the outcome of our autoencoder, will describe a directly measurable quantity with proper physical meaning. This is not common practice in the case of analog-acquired data, where the pixel value cannot be directly associated to an observable and the images are usually scaled between 0 and 1 before training and testing.

The choice of the dose values included in the training set highly affects the performance of the model and the ability to denoise images taken over a wide range of doses. As such one wishes to keep that range as wide as possible. However, both the time required for training and the computational costs also scale with the size training set and diversity of the data, so that a compromise between data variety and computational effort must be found. For this reason we have selected a dose range going from $500 e^- \text{\AA}^{-2}$ to $10\,000 e^- \text{\AA}^{-2}$. In the construction of the dataset the dose value for each image is randomly selected within the defined range, so that there is no dose bias across the various materials.

3. Results

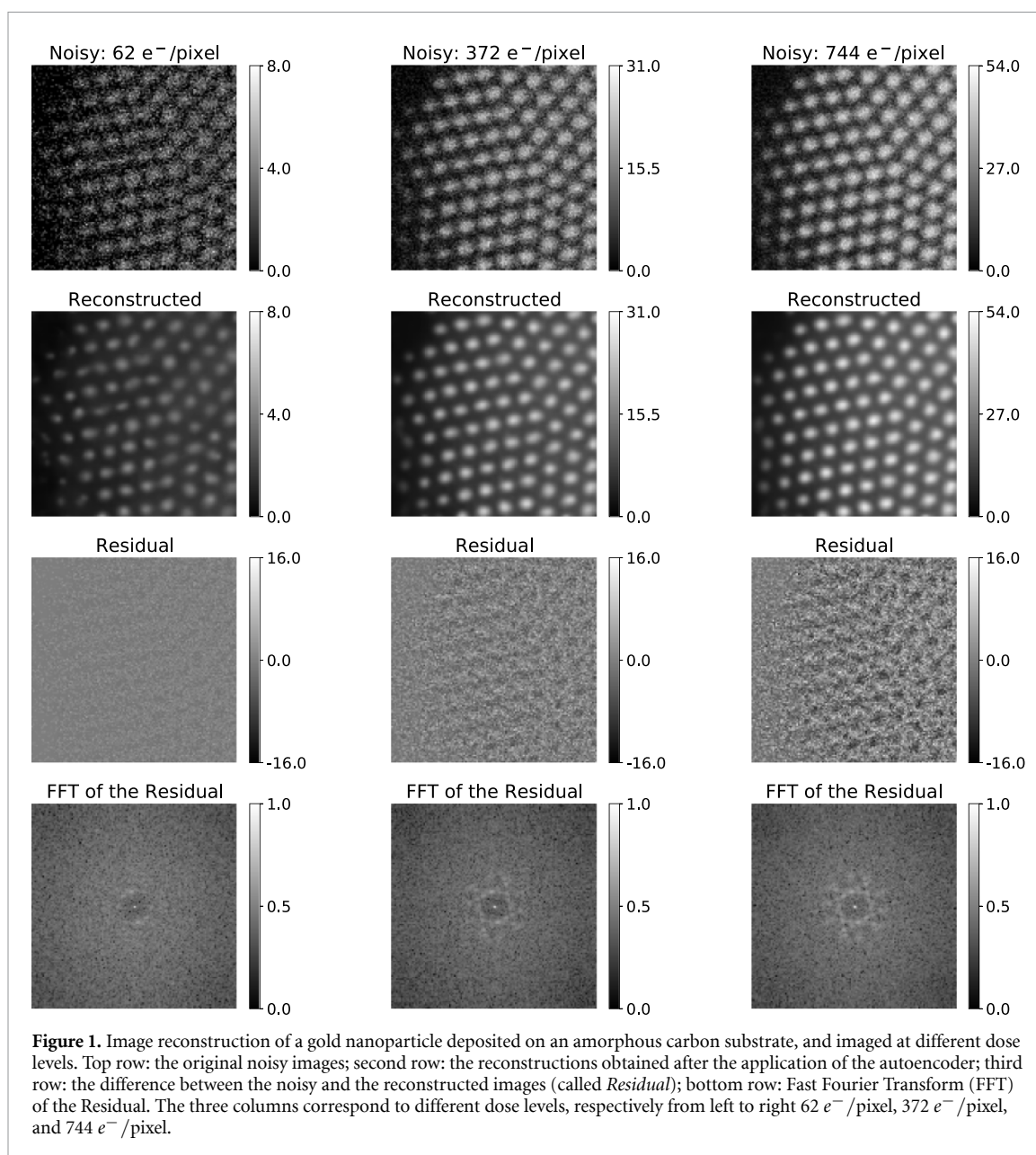
The model has been trained on Quadro RTX 8000 GPUs, for 500 epochs, with a batch size of 64. GPUs, provided by Nvidia, allowed us to significantly speed up the training time. In fact, the time required to train one epoch over the Quadro RTX 8000 GPUs is about 15 s, while with Tesla K40c approximately 120 s are needed. The training of one epoch of the same dataset on a CPU would require 717 s. After the training, during approximately 2 h and 30 min, the model performance has been tested on simulated and experimental images. Once the model is trained, the time required to denoise one image is less than 1 s, namely it is comparable to the acquisition time at an actual STEM.

Running the model on experimental images is ultimately the fundamental test to establish the effectiveness and usefulness of the algorithm over real data. However, the tests performed on simulated images provide us with useful information as well. In fact, the synthetic images include infinite-dose ground-truth data, which then help us in producing a quantitative benchmark for the algorithm. Different techniques can be used to validate the model, at both the qualitative and quantitative level, these are illustrated in what follows.

3.1. Visual comparison

The most straightforward qualitative evaluation consists of a visual comparison between the reconstructed and the corrupted images. The improvement brought by the denoising process is easily recognizable even by researchers, who are not experts in electron microscopy. A visual example is shown in figure 1, which displays the reconstruction of the digital experimental images of a gold nanoparticle deposited on an amorphous carbon substrate, obtained at different dose levels (note that gold is not included in the training set). The data are provided in the form of 20 different frames of the same sample region; by summing up the signals of an increasing number of frames one can obtain multiple images at different doses. In the experimental acquisition the dwell time was $2 \mu\text{s}$ for each frame and the beam current was approximately 5 pA. This means that the dose of a single frame is $62 e^-/\text{pixel}$. The dose of each image is then $62 e^-/\text{pixel}$ multiplied by the number of frames used for the image. For the sake of brevity, only 3 of the 20 consecutive sums are presented here. In particular, 128×128 -pixel portions of the images acquired at $62 e^-/\text{pixel}$, $372 e^-/\text{pixel}$, and $744 e^-/\text{pixel}$ are shown in figure 1. The top row displays the noisy images, the second one corresponds to the reconstructions obtained after the application of the autoencoder, the third row contains the difference between the noisy and the reconstructed images (referred to as *Residual*), and the bottom row shows the fast Fourier transform of the Residual. Although gold is not classified as a beam-sensitive material, the example provided is effective, since it tracks the model performance across different dose levels, a task that remains challenging when dealing with experimental data.

The reconstructed data always display a significant quality improvement over the original noisy images. In fact, from the reconstructions, one can immediately recognize the five crystallite forming the gold nanoparticle, regardless of the noise level of the original image. The most notable difference between the three reconstructions is the shape of the individual atoms, which appear progressively more round as the dose increases. Note that our algorithm is not trained to necessarily return round atoms, but only to denoise the signal. This is why at very low dose, as in the case of $62 e^-/\text{pixel}$, there is significant atomic distortion. It is also worth noting that no further progress is found when adding additional frames to the $744 e^-/\text{pixel}$ case (namely when increasing the dose). This result allows one to conclude that, upon autoencoder



reconstruction, an increase in the dose is not needed to obtain a satisfactory reconstruction. As a consequence, the beam damage to the sample can be reduced.

The two bottom rows of figure 1 show some periodicity in the removed noise. This is an expected feature of the Poisson noise, which scales with the pixel intensity, and does not imply a loss in crystallographic information. Such periodicity in the residual appears to be more evident for high-dose images, a fact that simply demonstrates that the autoencoder can remove the noise at high dose more efficiently than at low dose. It is worth mentioning that crystal structure information in the residual would not be found in images affected by Gaussian noise only. Gaussian noise, in fact, is not intensity-dependent and therefore does not follow the crystal structure. A final important consideration is that the model is able to efficiently denoise the data despite the presence of an amorphous substrate, which generates diffuse scattering and hence additional non-Poisson noise. Substrate scattering was not included in the training set, so that the reconstructed images should be considered as Poisson denoised but still inclusive of substrate scattering.

3.2. Line profile analysis

A quantitative evaluation of the model performance can be achieved through the so-called line profile analysis. This consists in selecting one line of pixels, along the horizontal direction in this case, and by plotting the pixel intensity at each position. One thus obtains an intensity scan profile that can be used to

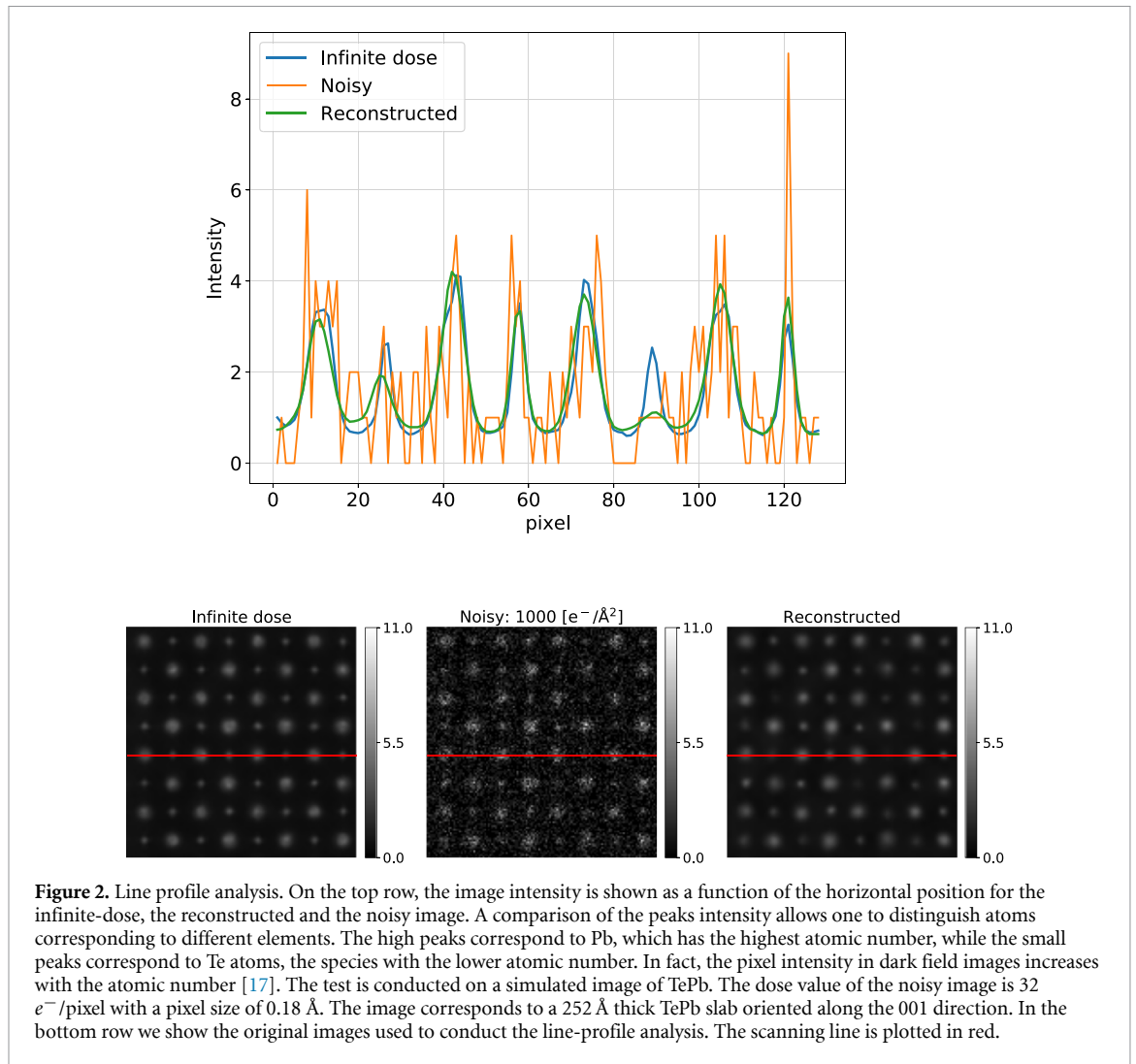


Figure 2. Line profile analysis. On the top row, the image intensity is shown as a function of the horizontal position for the infinite-dose, the reconstructed and the noisy image. A comparison of the peaks intensity allows one to distinguish atoms corresponding to different elements. The high peaks correspond to Pb, which has the highest atomic number, while the small peaks correspond to Te atoms, the species with the lower atomic number. In fact, the pixel intensity in dark field images increases with the atomic number [17]. The test is conducted on a simulated image of TePb. The dose value of the noisy image is $32 e^-/\text{pixel}$ with a pixel size of 0.18 \AA . The image corresponds to a 252 \AA thick TePb slab oriented along the 001 direction. In the bottom row we show the original images used to conduct the line-profile analysis. The scanning line is plotted in red.

distinguish atoms of different elements, as shown in figure 2. The line profile analysis is here conducted on a synthetic image of a 252 \AA thick TePb specimen oriented along the 001 direction, taken at the low dose of $1000 e^- \text{ \AA}^{-2}$, with a pixel size of 0.18 \AA . This corresponds to a pixel dose of $32 e^-/\text{pixel}$. As we can note from the figure, although the intensity profile of the reconstructed image is not identical to that of the infinite-dose one, the reconstruction appears to be accurate enough to localize and distinguish Te and Pb atoms, namely it contains the same content of information of the ground-truth case. In contrast, when the same line profile analysis is conducted over the noisy image the two different species appear indistinguishable so that the chemical information cannot be extracted. The original images used to perform the line profile analysis can be seen in the bottom row of figure 2.

3.3. Precision of atomic column localization

Another technique that can be used to quantitatively validate the model, involves the determination of atomic column localization. This essentially defines the position of the various atomic columns, thus allowing one to extract quantitative structural information from the STEM measurements. By performing atomic column localization, it is possible to quantify possible lattice strain and measure its error. This is a technologically useful information, since strain affects many physical properties of a material [18]. Several computational schemes and associated software are available for this purpose, one of them is the Matlab-based package StatSTEM [19]. StatSTEM is based on the principle that, in STEM images, the intensity peaks are located at the atomic column position and these can be approximated by a Gaussian function [20]. Since several localization methods can be used to determine the position of the Gaussians, it is important that our quantitative analysis considers both the localization method and the denoising algorithm, in order to provide a quantitative benchmark of the various possible image-processing workflows.

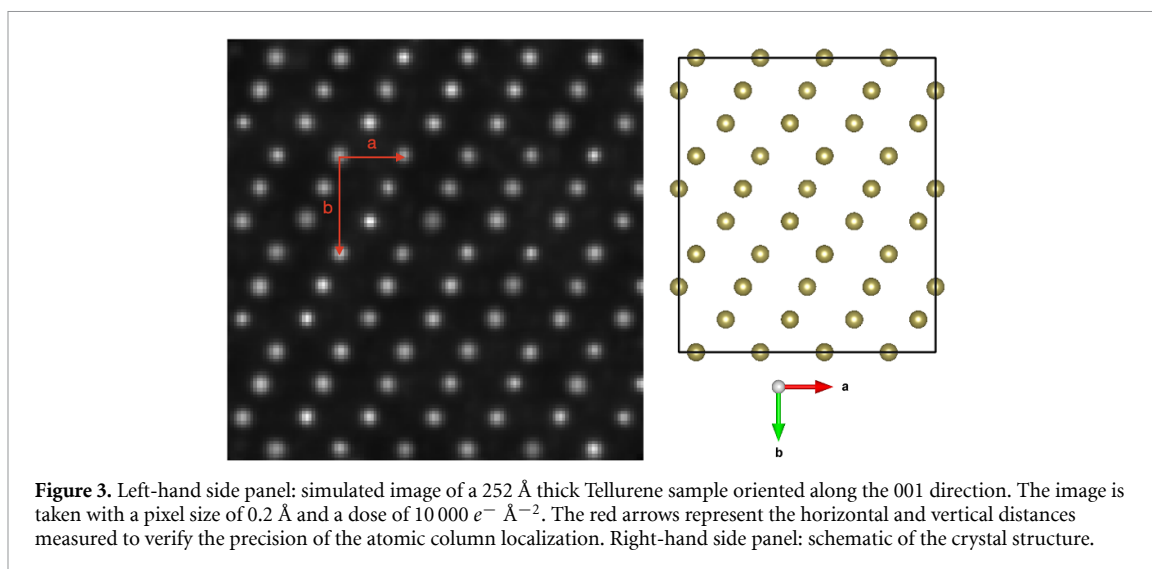


Figure 3. Left-hand side panel: simulated image of a 252 Å thick Tellurene sample oriented along the 001 direction. The image is taken with a pixel size of 0.2 Å and a dose of 10 000 $e^- \text{Å}^{-2}$. The red arrows represent the horizontal and vertical distances measured to verify the precision of the atomic column localization. Right-hand side panel: schematic of the crystal structure.

The precision of the atomic column localization can be estimated by measuring the distances between the various atoms, in both the horizontal and vertical directions. As these are determined solely by the crystalline structure, a statistical distribution of the distances provides a quantitative measure of the accuracy of the combined denoising and localization algorithm. Thus, the standard deviation of the computed distances can be taken as a measure of the accuracy and one can compare results obtained for the infinite-dose, the noisy and the reconstructed images. A reduction in the distance standard deviation corresponds to an enhancement in the image resolvability. The strain error along the horizontal and vertical direction can then be found by dividing the standard deviation in the position by the reference horizontal and vertical distances, respectively.

Figure 3 shows the simulated image used for this investigation. This corresponds to a 252 Å thick Tellurene sample oriented along the 001 direction and imaged with a pixel size of 0.2 Å. The reference horizontal and vertical distances, a and b respectively, are marked by the red arrows. The easiness of identifying these distances makes Tellurene a good candidate for strain error analysis. The plot in figure 3 corresponds to the highest dose considered for this analysis, namely 10 000 $e^- \text{Å}^{-2}$, while denoising is also performed for images taken at 500 $e^- \text{Å}^{-2}$, 1000 $e^- \text{Å}^{-2}$, 2500 $e^- \text{Å}^{-2}$, 5000 $e^- \text{Å}^{-2}$ and 7500 $e^- \text{Å}^{-2}$.

Our denoising autoencoder is then tested against a commonly used algorithm for image processing, namely a Gaussian filter [21]. The procedure to measure the column localization by using StatSTEM is as follows. Firstly, one needs to define the starting coordinates for the atomic column positions, namely the local maxima in the image. This can be achieved by using one of the two available peak-finder routines that may include also filters to smooth the image. Such filtering option is not employed in the present study, where the only simplification is the setting of a threshold value to remove nuisance pixel intensities from the background. This step is necessary to avoid the identification of too many fictitious atoms in the case of images characterized by strong noise. In order to set the same value for each image and to make the analysis coherent, the intensity is normalized to 1; then we find that the minimum threshold value compatible with the algorithm memory requirement is 0.12. The difference between the two peak-finder routines lies in the way the smoothing of the image can be achieved, namely either by adding filters or by altering the estimated radius of the atomic columns. The so-called *Peak-finder routine 2* is used for this analysis; however, the same results can be achieved with the *Peak-finder routine 1* when no filters are imposed.

After this first step, one has the option to manually add atoms positions, an operation that is not considered in this case in order to avoid any human bias in the workflow. Once the starting coordinates are identified, a fitting procedure can be used to model the image as a superposition of Gaussian peaks. In this step, it is possible to specify the width of the atomic columns, by choosing between the *Same* and *Different* options. In the first case all Gaussian are taken with the same width. In contrast, the second more computationally demanding option makes the approach more general, since it does not assume that all the Gaussian peaks have the same width. As such it avoids the introduction of any *a priori* knowledge of the image. The final result of this procedure is a set of coordinates, which correspond to the atomic columns in the image. These values can be used to measure the horizontal and vertical distances a and b between the atomic columns, as shown in figure 3. The values obtained from this operation are displayed as histograms in

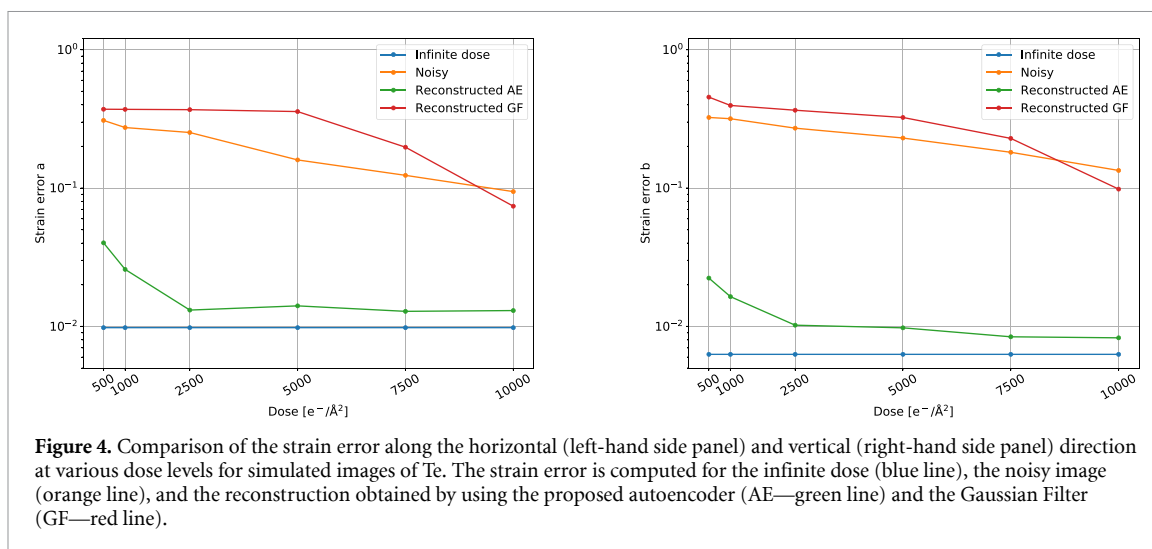


Figure 4. Comparison of the strain error along the horizontal (left-hand side panel) and vertical (right-hand side panel) direction at various dose levels for simulated images of Te. The strain error is computed for the infinite dose (blue line), the noisy image (orange line), and the reconstruction obtained by using the proposed autoencoder (AE—green line) and the Gaussian Filter (GF—red line).

the appendix. From the histogram corresponding to each image, the strain error can be computed and compared at different dose values. This is shown for the distances a and b of Tellurene (see figure 3) in figure 4.

In the figures, the strain error for the infinite-dose case is represented as a blue line and, by definition, it is dose-independent. This, in fact, represents the ultimate theoretical precision achievable by noise-free STEM. In contrast, the noisy images have a strain error that grows with reducing the dose, in an approximate exponential behavior. Our autoencoder is able to drastically improve over the noisy images and returns us a strain error that trails closely that corresponding to the infinite-dose case. In more detail, we find that the denoised images present a strain error, which is approximately dose independent when the electron dose is higher than $2500 \text{ e}^- \text{ \AA}^{-2}$. For lower doses a sharp error increase is reported. Such increase, however, leaves the strain error far below what computed for the noisy images. We then conclude that the autoencoder is less effective at ultra-low dose, but still remains significantly performing across the entire range. Note that one, in principle, can also train different autoencoders for different dose levels. This, however, may not be a good strategy, since often the beam intensity is not accurately measured during the STEM operation. For this reason an autoencoder trained over a broad dose range appears more flexible for real-life situations. Interestingly, simple Gaussian filtering (red lines in the figures) appears unable to improve the column localization of noisy images, with the only exception at very high doses. Therefore, the commonly used Gaussian filter should be avoided when performing accurate quantitative measures of atomic positions, at least when the data-processing workflow remains completely users unbiased. In this situation our results suggest that even the untreated images provide a better estimate of the column positions, unless rather large doses are used.

We believe that the comparison provided here should represent a general benchmarking scheme to compare different denoising workflows in a completely unbiased way.

4. Conclusions

We have proposed a neural network trained on simulated STEM images, which allows one to successfully remove Poisson noise from low-dose digital data. Tests on both simulated and experimental images have been conducted to validate the model. Our results show a clear improvement in the image resolution and in the possibility to extract useful information from the data in a completely unbiased way. The use of this model may allow a drastic reduction of the dose level to be employed in real-life measurements, making it possible to analyze very beam-sensitive compounds. Notably, our proposed denoising algorithm does not require any human input or knowledge of the actual beam intensity, being trained over a broad range of doses. Crucially, the denoising process of a 128×128 -pixel image can be performed within approximately one second, making the scheme amenable to be used in a live mode during the actual STEM measurement.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

L G is supported by Science Foundation Ireland (AMBER Center Grant 12/RC/2278_P2). L J is supported by SFI-Royal Society Fellowship (Grant URF/RI/191637). T M acknowledges the SFI and EPSRC Centre for Doctoral Training in the Advanced Characterisation of Materials (Award References 18/EPSRC-CDT-3581 and EP/S023259/1).

All calculations have been performed on the Boyle cluster maintained by the Trinity Centre for High Performance Computing. The hardware was funded through grants from the European Research Council and Science Foundation Ireland. The calculations were accelerated on Nvidia Quadro RTX 8000 GPUs, provided by Nvidia under the Nvidia Academic Hardware Grant Program, which we kindly acknowledge. Microscopy characterisation and analysis have been performed at the CRANN Advanced Microscopy Laboratory (AML www.tcd.ie/crann/aml/).

Appendix A. Hyperparameters tuning

Hyperparameter tuning was performed to optimize the model architecture. A list of all the hyperparameters of the model and the range of hyperparameter values investigated can be found in table A1.

Appendix B. Training set details

The Prismatic package [3, 14] has been used to simulate the images of the various materials forming the training set. More specifically, images of graphene, graphite, GaAs, InAs, MoS₂, SrTiO₃, and Si have been generated across different imaging conditions: (a) the pixel size varies from 0.08 to 0.3 Å, (b) the sample thickness range is 250–630 Å, (c) the electron beam energy range is 80–200 keV, and (d) the inner and outer detector collection angles are 30 and 60 mrad, 30 and 70 mrad, 50 and 180 mrad, 60 and 180 mrad, 65 and 180 mrad, 75 and 180 mrad, 80 and 180 mrad. These intervals correspond to realistic imaging and materials parameters, commonly encountered in actual STEM experiments. For each material, 10 to 12 images were generated, including pristine and defective structures incorporating vacancies. In order to avoid reconstruction bias, each simulated image has been rotated at two random angles so to increase the variety of the dataset. Poisson noise has been applied according to multiple dose values, within the range 500 to 10 000 e⁻ Å⁻². Going more in detail, for each image 60 dose values have been randomly selected within the defined range, so that there is no dose bias across the dataset. It is worth mentioning that a different dose distribution across the dataset has also been tested. In this case, the number of images N_i at a certain dose value was proportional to the selected dose value, according to the relation:

$$N_i(\text{dose}) \propto \frac{1}{\sqrt{\text{dose}}}, \quad (\text{B.1})$$

which is the same relation that occurs between Poisson noise and dose (see equation (1)). However, this solution did not improve the model performance. Finally, the images have been cropped into 128 × 128-pixel plots. The result is a dataset made of about 30 000 images: 10% was used as validation set and 90% as training set. The different materials are equally distributed in the dataset. The mentioned values have been chosen in agreement to the desired training set size, which has been chosen according to the computational memory availability.

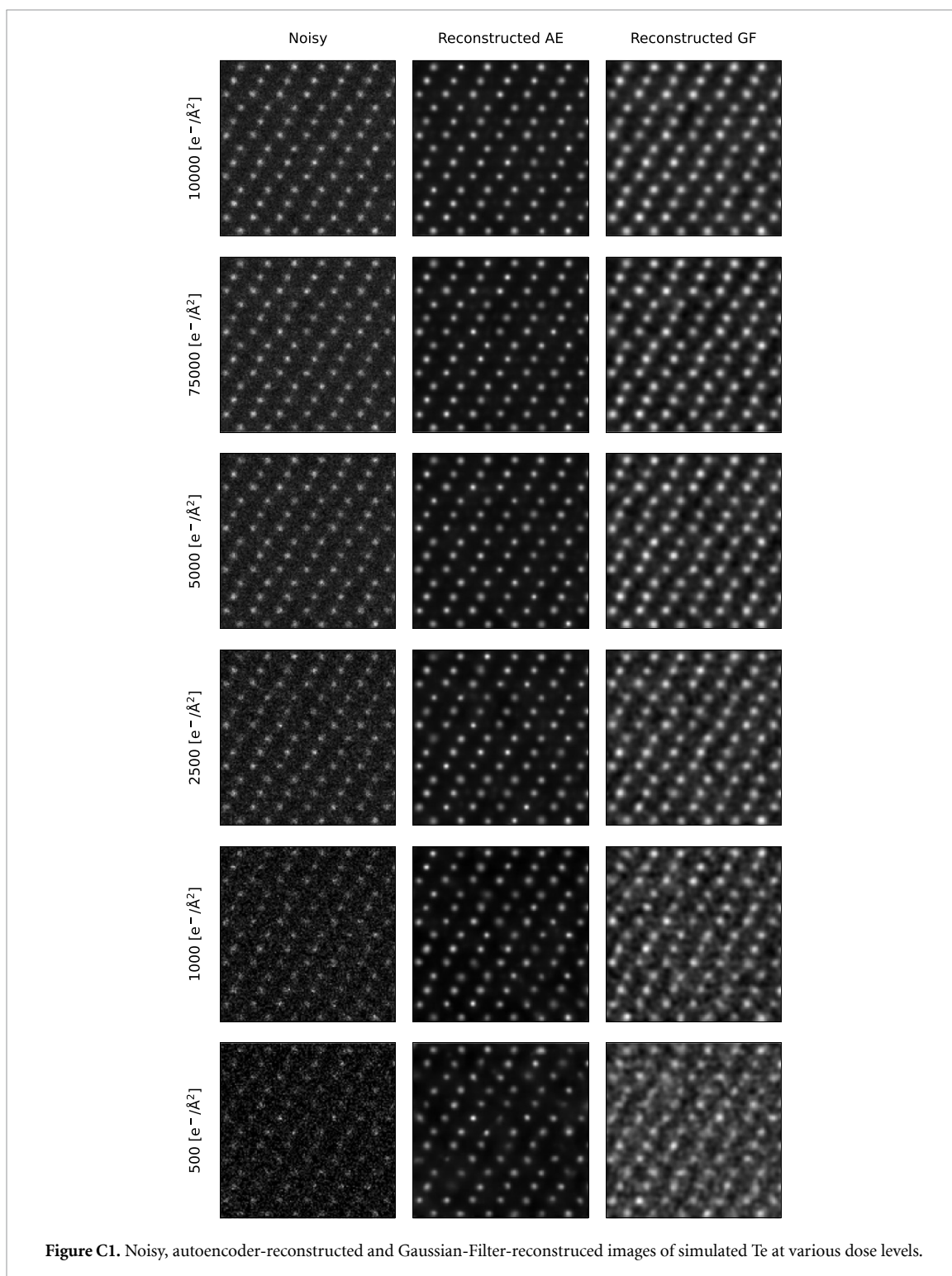
Appendix C. Atomic column localization details

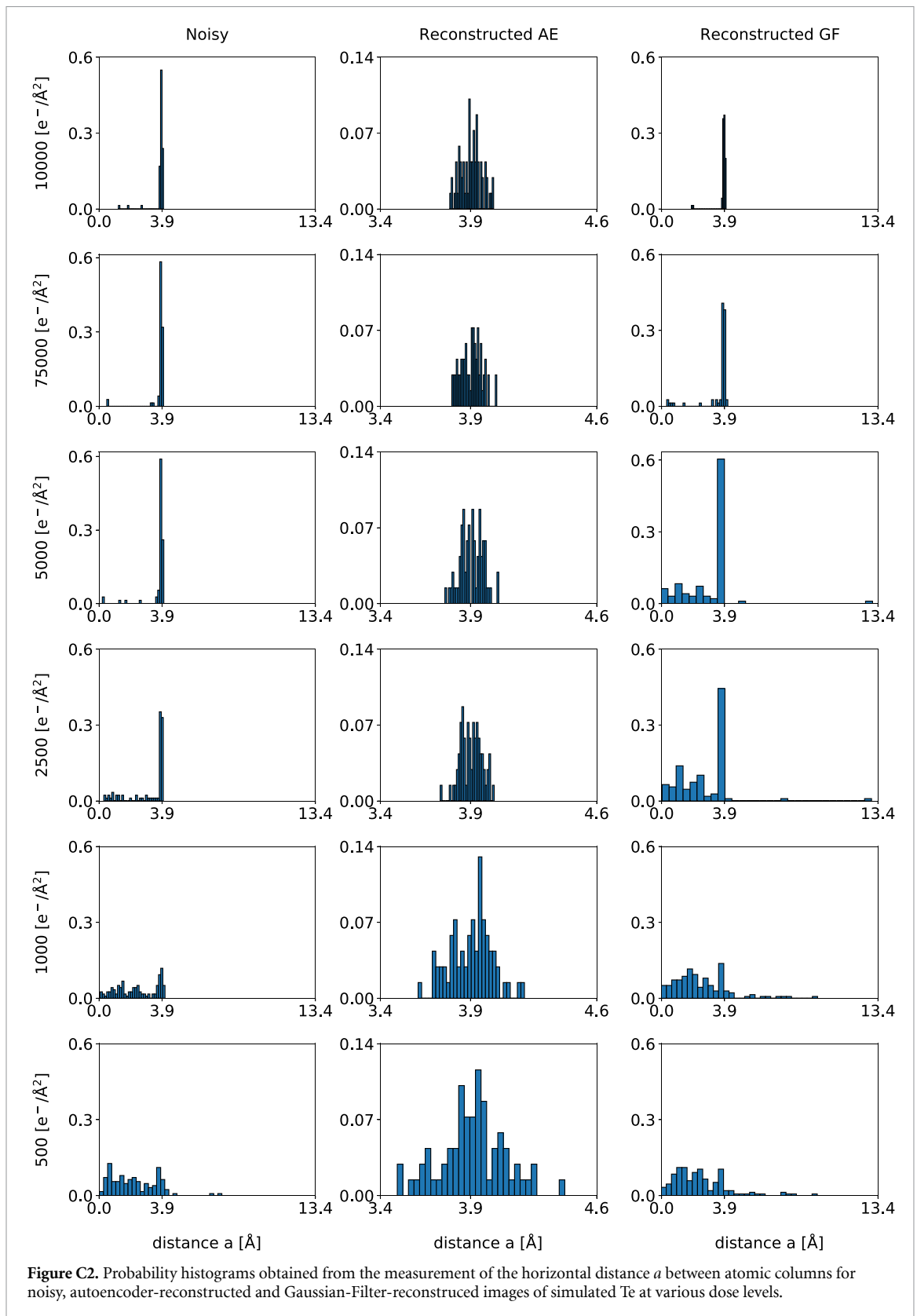
The data used for the analysis of the atomic column localization precision are displayed in figure C1. The first column shows 128 × 128-pixel images of Tellurene simulated at various dose levels, ranging from 10 000 e⁻ Å⁻² to 500 e⁻ Å⁻². The reconstruction obtained by using the proposed autoencoder (*Reconstructed AE*) and that obtained with a Gaussian Filter (*Reconstructed GF*) can be seen in the second and third column, respectively. The resolution enhancement is more evident for higher dose levels, for both reconstruction techniques. However, the autoencoder appears to be significantly more successful in the reconstruction task at low dose, and in general at any dose value.

The StatSTEM software used to localize the atom position in these images, is described in section 3.3 of the main paper. Once the atomic columns are localized, the horizontal and vertical distances, displayed in figure 3, are measured and they can be represented as histograms, for each image. Figure C2 shows the

Table A1. List of all the hyperparameters defining our model together with the corresponding range of values investigated.

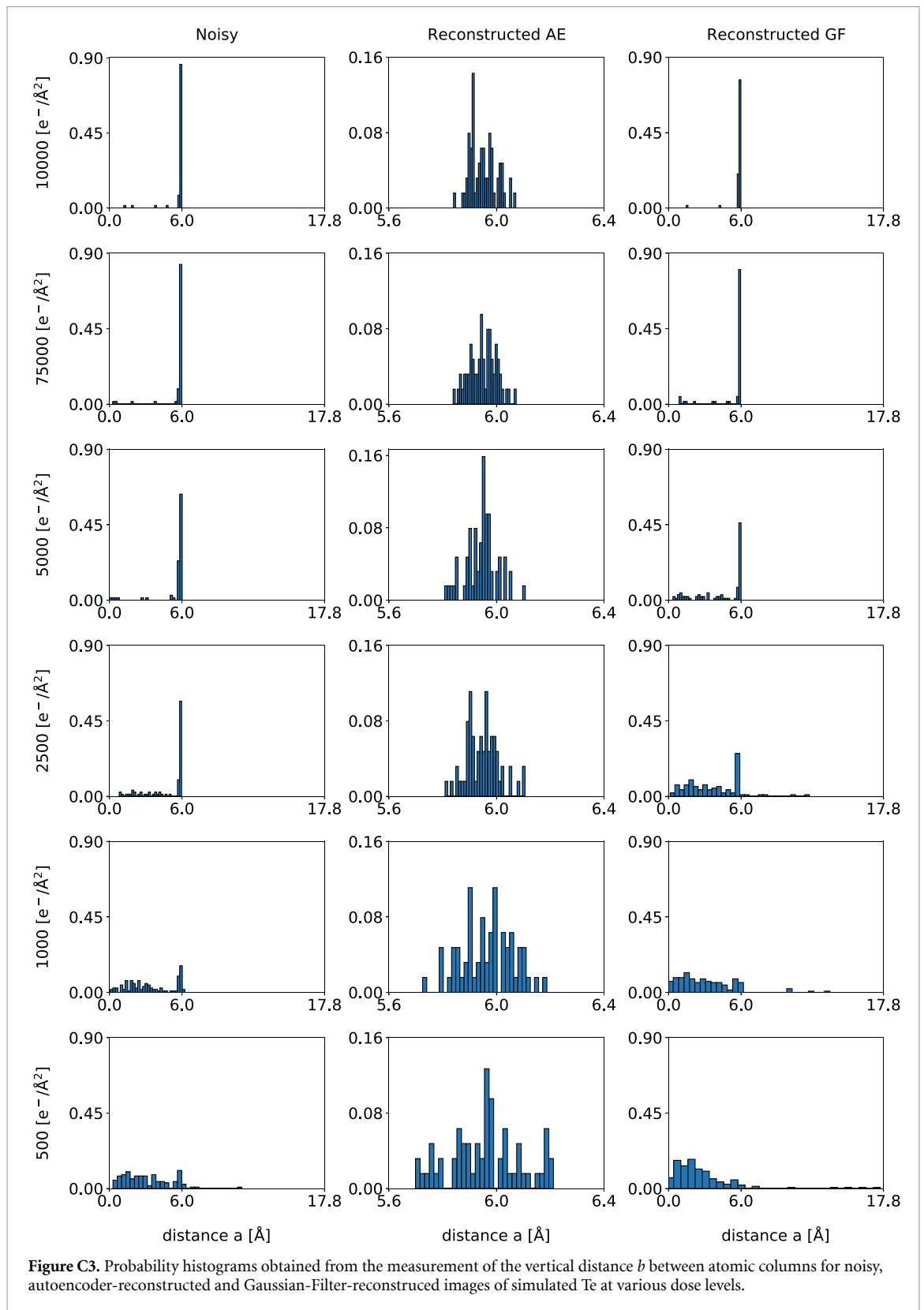
<i>Hyperparameter</i>	<i>Investigated values</i>
Number of layers	6, 10, 14
Number of filters in the Conv2D layers	8, 16, 32, 64
Loss function weight	10, 100, 1000, 10 000
Optimizer	Adam, Adadelta
Batch size	16, 32, 64





histograms for the horizontal distance, a , while figure C3 shows the histograms for the vertical distance, b . The results are here represented as probability histograms, which means that the data are normalized to one. The data associated with the noisy images are placed in the first column, while those associated with the autoencoder and the Gaussian Filter reconstructions are in the second and third column, respectively.

The distribution appears to be more uniform for the autoencoder-reconstructed data (*Reconstructed AE*), compared to the other results. In the noisy and Gaussian-Filter-reconstructed (*Reconstructed GF*) images



some atoms are misplaced and incorrectly localized, therefore the spread in the histograms appears wider. It should be noted that the scale of the x-axis in the case of the autoencoder-reconstructed data is different from that of the other two columns. The distances reported in the central column of both figures C2 and C3 appear to be more localized around a single value with a distribution following an approximately normal distribution. Such feature facilitates the atoms localization.

ORCID iDs

Laura Gambini  <https://orcid.org/0000-0001-6316-7498>
Tiarnan Mullarkey  <https://orcid.org/0000-0002-5209-1712>
Lewys Jones  <https://orcid.org/0000-0002-6907-0731>
Stefano Sanvito  <https://orcid.org/0000-0002-0291-715X>

References

- [1] Brydson R 2011 *Aberration-Corrected Analytical Transmission Electron Microscopy* (West Sussex: Wiley Online Library)
- [2] Egerton R F, Li P and Malac M 2004 Radiation damage in the TEM and SEM *Micron* **35** 399–409
- [3] Pryor A, Ophus C and Miao J 2017 A streaming multi-GPU implementation of image simulation algorithms for scanning transmission electron microscopy *Adv. Struct. Chem. Imaging* **3** 1–14
- [4] Jones L and Nellist P D 2013 Identifying and correcting scan noise and drift in the scanning transmission electron microscope *Microsc. Microanal.* **19** 1050
- [5] Mullarkey T, Downing C and Jones L 2021 Development of a practicable digital pulse read-out for dark-field stem *Microsc. Microanal.* **27** 99–108
- [6] Thakur K V, Damodare O H and Sapkal A M 2016 Poisson noise reducing bilateral filter *Proc. Comput. Sci.* **79** 861–5
- [7] Du H 2015 A nonlinear filtering algorithm for denoising HR(S)TEM micrographs *Ultramicroscopy* **151** 62–67
- [8] Meyer J C et al 2011 Experimental analysis of charge redistribution due to chemical bonding by high-resolution transmission electron microscopy *Nat. Mater.* **10** 209–15
- [9] Mevenkamp N, Binev P, Dahmen W, Voyles P M, Yankovich A B and Berkels B 2015 Poisson noise removal from high-resolution stem images based on periodic block matching *Adv. Struct. Chem. Imaging* **1** 1–19
- [10] Ede J M 2021 Deep learning in electron microscopy *Mach. Learn.: Sci. Technol.* **2** 011004
- [11] Lin R, Zhang R, Wang C, Yang X-Q and Xin H L 2021 Temimagenet training library and atomsegnet deep-learning models for high-precision atom segmentation, localization, denoising and deblurring of atomic-resolution images *Sci. Rep.* **11** 1–15
- [12] Cowley J M and Moodie A F 1957 The scattering of electrons by atoms and crystals. I. A new theoretical approach *Acta Crystallogr.* **10** 609–19
- [13] Kirkland E J 1998 *Advanced Computing in Electron Microscopy* (Berlin: Springer)
- [14] Ophus C 2017 A fast image simulation algorithm for scanning transmission electron microscopy *Adv. Struct. Chem. Imaging* **3** 1–11
- [15] Goodfellow I, Bengio Y and Courville A 2016 *Deep Learning* vol 1 (Cambridge: MIT Press)
- [16] Kingma D P and Ba J 2014 Adam: a method for stochastic optimization (arXiv:1412.6980)
- [17] Krivanek O L et al 2010 Atom-by-atom structural and chemical analysis by annular dark-field electron microscopy *Nature* **464** 571–4
- [18] Han Y, Nguyen K, Cao M, Cueva P, Xie S, Tate M W, Purohit P, Gruner S M, Park J and Muller D A 2018 Strain mapping of two-dimensional heterostructures with subpicometer precision *Nano Lett.* **18** 3746–51
- [19] De Backer A, Van den Bos K H W, Van den Broek W, Sijbers J and Van Aert S 2016 Statstem: an efficient approach for accurate and precise model-based quantification of atomic resolution electron microscopy images *Ultramicroscopy* **171** 104–16
- [20] van Dyck D 2002 High-resolution electron microscopy *Advances in Imaging and Electron Physics* vol 123 (Amsterdam: Elsevier) pp 105–71
- [21] van der Walt S, Schönberger J L, Nunez-Iglesias J, Boulogne F, Warner J D, Yager N, Gouillart E and Yu T (the scikit-image contributors) 2014 Scikit-image: image processing in Python *PeerJ* **2** e453