

OPEN

Partial Scanning Transmission Electron Microscopy with Deep Learning

Jeffrey M. Ede[✉] & Richard Beanland

Compressed sensing algorithms are used to decrease electron microscope scan time and electron beam exposure with minimal information loss. Following successful applications of deep learning to compressed sensing, we have developed a two-stage multiscale generative adversarial neural network to complete realistic 512×512 scanning transmission electron micrographs from spiral, jittered gridlike, and other partial scans. For spiral scans and mean squared error based pre-training, this enables electron beam coverage to be decreased by $17.9 \times$ with a 3.8% test set root mean squared intensity error, and by $87.0 \times$ with a 6.2% error. Our generator networks are trained on partial scans created from a new dataset of 16227 scanning transmission electron micrographs. High performance is achieved with adaptive learning rate clipping of loss spikes and an auxiliary trainer network. Our source code, new dataset, and pre-trained models are publicly available.

Aberration corrected scanning transmission electron microscopy (STEM) can achieve imaging resolutions below 0.1 nm, and locate atom columns with pm precision^{1,2}. Nonetheless, the high current density of electron probes produces radiation damage in many materials, limiting the range and type of investigations that can be performed^{3,4}. A number of strategies to minimize beam damage have been proposed, including dose fractionation⁵ and a variety of sparse data collection methods⁶. Perhaps the most intensively investigated approach to the latter is sampling a random subset of pixels, followed by reconstruction using an inpainting algorithm^{3,6–10}. Poisson random sampling of pixels is optimal for reconstruction by compressed sensing algorithms¹¹. However, random sampling exceeds the design parameters of standard electron beam deflection systems, and can only be performed by collecting data slowly^{12,13}, or with the addition of a fast deflection or blanking system^{3,14}.

Sparse data collection methods that are more compatible with conventional beam deflection systems have also been investigated. For example, maintaining a linear fast scan deflection whilst using a widely-spaced slow scan axis with some small random ‘jitter’^{9,12}. However, even small jumps in electron beam position can lead to a significant difference between nominal and actual beam positions in a fast scan. Such jumps can be avoided by driving functions with continuous derivatives, such as those for spiral and Lissajous scan paths^{3,13,15,16}. Sang^{13,16} considered a variety of scans including Archimedes and Fermat spirals, and scans with constant angular or linear displacements, by driving electron beam deflectors with a field-programmable gate array (FPGA) based system. Spirals with constant angular velocity place the least demand on electron beam deflectors. However, dwell times, and therefore electron dose, decreases with radius. Conversely, spirals created with constant spatial speeds are prone to systematic image distortions due to lags in deflector responses. In practice, fixed doses are preferable as they simplify visual inspection and limit the dose dependence of STEM noise¹⁷.

Deep learning has a history of successful applications to image infilling, including image completion¹⁸, irregular gap infilling¹⁹ and supersampling²⁰. This has motivated applications of deep learning to the completion of sparse, or ‘partial’ scans, including supersampling of scanning electron microscopy²¹ (SEM) and STEM images^{22,23}. Where pre-trained models are unavailable for transfer learning²⁴, artificial neural networks (ANNs) are typically trained, validated and tested with large, carefully partitioned machine learning datasets^{25,26} so that they are robust to general use. In practice, this often requires at least a few thousand examples. Indeed, standard machine learning datasets such as CIFAR-10^{27,28}, MNIST²⁹, and ImageNet³⁰ contain tens of thousands or millions of examples. To train an ANN to complete STEM images from partial scans, an ideal dataset might consist of a large number of pairs of partial scans and corresponding high-quality, low noise images, taken with an aberration-corrected STEM. To our knowledge, such a dataset does not exist. As a result, we have collated a new dataset of STEM raster scans from which partial scans can be selected. Selecting partial scans from full scans is

University of Warwick, Department of Physics, Coventry, CV4 7AL, UK. [✉]e-mail: j.m.ede@warwick.ac.uk

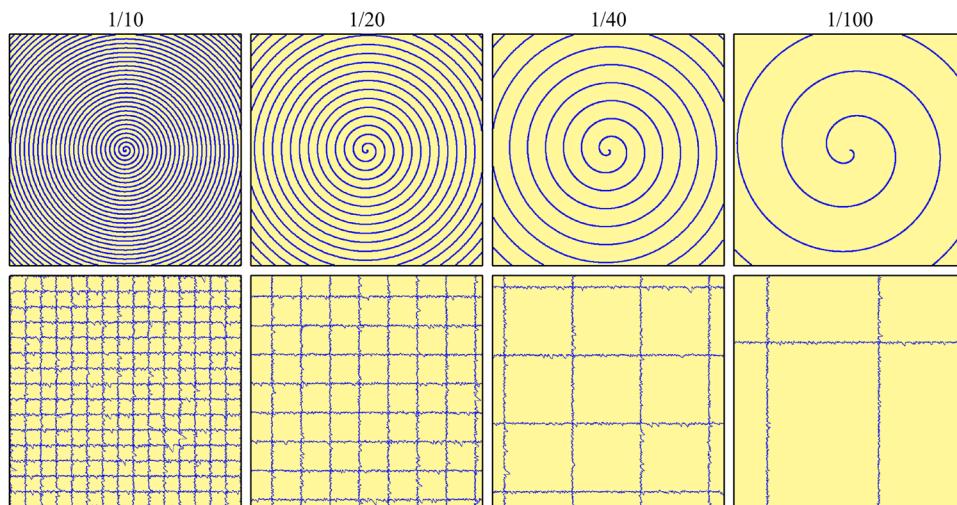


Figure 1. Examples of Archimedes spiral (top) and jittered gridlike (bottom) 512×512 partial scan paths for 1/10, 1/20, 1/40, and 1/100 px coverage.

less expensive than collecting image pairs, and individual pixels selected from experimental images have realistic noise characteristics.

Examples of spiral and jittered gridlike partial scans investigated in this paper are shown in Fig. 1. Continuous spiral scan paths that extend to image corners cannot be created by conventional scan systems without going over image edges. However, such a spiral can be cropped from a spiral with radius at least $2^{-1/2}$ times the minimum image side, at the cost of increased scan time and electron beam damage to the surrounding material. We use Archimedes spirals, where $r \propto \theta$, and r and θ are polar radius and angle coordinates, as these spirals have the most uniform spatial coverage. Jittered gridlike scans would also be difficult to produce with a conventional system, which would suffer variations in dose and distortions due to limited beam deflector response. Nevertheless, these idealized scan paths serve as useful inputs to demonstrate the capabilities of our approach. We expect that other scan paths could be used with similar results.

We fine-tune our ANNs as part of generative adversarial networks³¹ (GANs) to complete realistic images from partial scans. A GAN consists of sets of generators and discriminators that play an adversarial game. Generators learn to produce outputs that look realistic to discriminators, while discriminators learn to distinguish between real and generated examples. Limitedly, discriminators only assess whether outputs look realistic; not if they are correct. This can result in a neural network only generating a subset of outputs, referred to as mode collapse³². To counter this issue, generator learning can be conditioned on an additional distance between generated and true images³³. Meaningful distances can be hand-crafted or learned automatically by considering differences between features imagined by discriminators for real and generated images^{34,35}.

Training

In this section we introduce a new STEM images dataset for machine learning, describe how partial scans were selected from images in our data pipeline, and outline ANN architecture and learning policy. Detailed ANN architecture, learning policy, and experiments are provided as Supplementary Information, and source code is available³⁶.

Data pipeline. To create partial scan examples, we collated a new dataset containing 16227 32-bit floating point STEM images collected with a JEOL ARM200F atomic resolution electron microscope. Individual micrographs were saved to University of Warwick data servers by dozens of scientists working on hundreds of projects as Gatan Microscopy Suite³⁷ generated dm3 or dm4 files. As a result, our dataset has a diverse constitution. Atom columns are visible in two-thirds of STEM images, with most signals imaged at several times their Nyquist rates³⁸, and similar proportions of images are bright and dark field. The other third of images are at magnifications too low for atomic resolution, or are of amorphous materials. Importantly, our dataset contains noisy images, incomplete scans and other low-quality images that would not normally be published. This ensures that ANNs trained on our dataset are robust to general use. The Digital Micrograph image format is rarely used outside the microscopy community. As a result, data has been transferred to the widely supported TIFF³⁹ file format in our publicly available dataset^{40,41}.

Micrographs were split into 12170 training, 1622 validation, and 2435 test set examples. Each subset was collected by a different subset of scientists and has different characteristics. As a result, unseen validation and test sets can be used to quantify the ability of a trained network to generalize. To reduce data read times, each micrograph was split into non-overlapping 512×512 sub-images, referred to as ‘crops’, producing 110933 training, 21259 validation and 28877 test set crops. For convenience, our crops dataset is also available^{40,41}. Each crop, I , was processed in our data pipeline by replacing non-finite electron counts, i.e. NaN and $\pm\infty$, with zeros. Crops were then linearly transformed to have intensities $I_N \in [-1, 1]$, except for uniform crops satisfying

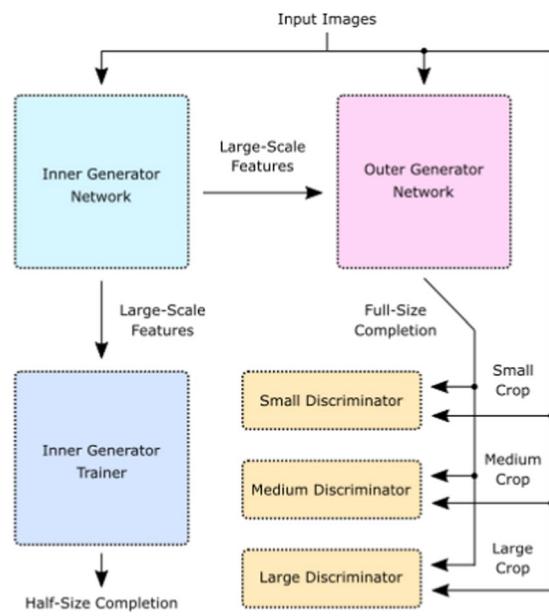


Figure 2. Simplified multiscale generative adversarial network. An inner generator produces large-scale features from inputs. These are mapped to half-size completions by a trainer network and recombined with the input to generate full-size completions by an outer generator. Multiple discriminators assess multiscale crops from input images and full-size completions. This figure was created with Inkscape⁸³.

$\max(I) - \min(I) < 10^{-6}$ where we set $I_N = 0$ everywhere. Finally, each crop was subject to a random combination of flips and 90° rotations to augment the dataset by a factor of eight.

Partial scans, I_{scan} , were selected from raster scan crops, I_N , by multiplication with a binary mask Φ_{path} ,

$$I_{\text{scan}} = \Phi_{\text{path}} I_N, \quad (1)$$

where $\Phi_{\text{path}} = 1$ on a scan path, and $\Phi_{\text{path}} = 0$ otherwise. Raster scans are sampled at a rectangular lattice of discrete locations, so a subset of raster scan pixels are experimental measurements. In addition, although electron probe position error characteristics may differ for partial and raster scans, typical position errors are small^{42,43}. As a result, we expect that partial scans selected from raster scans with binary masks are realistic.

We also selected partial scans with blurred masks to simulate varying dwell times and noise characteristics. These difficulties are encountered in incoherent STEM^{44,45}, where STEM illumination is detected by a transmission electron microscopy (TEM) camera. For simplicity, we created non-physical noise by multiplying I_{scan} with $\eta(\Phi_{\text{path}}) = \Phi_{\text{path}} + (1 - \Phi_{\text{path}})U$, where U is a uniform random variate distributed in $[0, 2]$. ANNs are able to generalize^{46,47}, so we expect similar results for other noise characteristics. A binary mask, with values in $\{0, 1\}$, is a special case where no noise is applied i.e. $\eta(1) = 1$, and $\Phi_{\text{path}} = 0$ is not traversed. Performance is reported for both binary and blurred masks.

The noise characteristics in our new STEM images dataset vary. This is problematic for mean squared error (MSE) based ANN training losses, as differences are higher for crops with higher noise. In effect, this would increase the importance of noisy images in the dataset, even if they are not more representative. Although adaptive ANN optimizers that divide parameter learning rates by gradient sizes⁴⁸ can partially mitigate weighting by varying noise levels, this restricts training to a batch size of 1 and limits momentum. Consequently, we low-passed filtered ground truth images, I_N , to I_{blur} by a 5×5 symmetric Gaussian kernel with a 2.5 px standard deviation, to calculate MSEs for ANN outputs.

Network architecture. To generate realistic images, we developed a multiscale conditional GAN with TensorFlow⁴⁹. Our network can be partitioned into the six convolutional^{50,51} subnetworks shown in Fig. 2: an inner generator, G_{inner} , outer generator, G_{outer} , inner generator trainer, T , and small, medium and large scale discriminators, D_1 , D_2 and D_3 . We refer to the compound network $G(I_{\text{scan}}) = G_{\text{outer}}(G_{\text{inner}}(I_{\text{scan}}), I_{\text{scan}})$ as the generator, and to $D = \{D_1, D_2, D_3\}$ as the multiscale discriminator. The generator is the only network needed for inference.

Following recent work on high-resolution conditional GANs³⁴, we use two generator subnetworks. The inner generator produces large scale features from partial scans bilinearly downsampled from 512×512 to 256×256 . These features are then combined with inputs embedded by the outer generator to output full-size completions. Following Inception^{52,53}, we introduce an auxiliary trainer network that cooperates with the inner generator to output 256×256 completions. This acts as a regularization mechanism, and provides a more direct path for

gradients to backpropagate to the inner generator. To more efficiently utilize initial generator convolutions, partial scans selected with a binary mask are nearest neighbour infilled before being input to the generator.

Multiscale discriminators examine real and generated STEM images to predict whether they are real or generated, adapting to the generator as it learns. Each discriminator assesses different-sized crops selected from 512×512 images, with sizes 70×70 , 140×140 or 280×280 . After selection, crops are bilinearly downsampled to 70×70 before discriminator convolutions. Typically, discriminators are applied at fractions of the full image size³⁴ e.g. $512/2^2$, $512/2^1$ and $512/2^0$. However, we found that discriminators that downsample large fields of view to 70×70 are less sensitive to high-frequency STEM noise characteristics. Processing fixed size image regions with multiple discriminators has been proposed³⁴ to decrease computation for large images, and extended to multiple region sizes³⁴. However, applying discriminators to arrays of non-overlapping image patches⁵⁵ results in periodic artefacts³⁴ that are often corrected by larger-scale discriminators. To avoid these artefacts and reduce computation, we apply discriminators to randomly selected regions at each spatial scale.

Learning policy. Training has two halves. In the non-adversarial first half, the generator and auxiliary trainer cooperate to minimize mean squared errors (MSEs). This is followed by an optional second half of training, where the generator is fine-tuned as part of a GAN to produce realistic images. Our ANNs are trained by ADAM⁵⁶ optimized stochastic gradient descent^{48,57} for up to 2×10^6 iterations, which takes a few days with an Nvidia GTX 1080 Ti GPU and an i7-6700 CPU. The objectives of each ANN are codified by their loss functions.

In the non-adversarial first half of training, the generator, G , learns to minimize the MSE based loss

$$L_{\text{MSE}} = \text{ALRC}(\lambda_{\text{cond}} \text{MSE}(G(I_{\text{scan}}), I_{\text{blur}})), \quad (2)$$

where $\lambda_{\text{cond}} = 200$, and adaptive learning rate clipping⁵⁸ (ALRC) is important to prevent high loss spikes from destabilizing learning. Experiments with and without ALRC are in Supplementary Information. To compensate for varying noise levels, ground truth images were blurred by a 5×5 symmetric Gaussian kernel with a 2.5 px standard deviation. In addition, the inner generator, G_{inner} , cooperates with the auxiliary trainer, T , to minimize

$$L_{\text{aux}} = \text{ALRC}(\lambda_{\text{trainer}} \text{MSE}(T(G_{\text{inner}}(I_{\text{scan}}^{\text{half}}))), I_{\text{blur}}^{\text{half}}), \quad (3)$$

where $\lambda_{\text{trainer}} = 200$, and $I_{\text{scan}}^{\text{half}}$ and $I_{\text{blur}}^{\text{half}}$ are 256×256 inputs bilinearly downsampled from I_{scan} and I_{blur} , respectively.

In the optional adversarial second half of training, we use $N = 3$ discriminator scales with numbers, N_1 , N_2 and N_3 , of discriminators, D_1 , D_2 and D_3 , respectively. There many popular GAN loss functions and regularization mechanisms^{59,60}. In this paper, we use spectral normalization⁶¹ with squared difference losses⁶² for the discriminators,

$$L_D = \frac{1}{N} \sum_{i=1}^N \frac{1}{N_i} [D_i(G(I_{\text{scan}}))^2 + (D_i(I_N) - 1)^2], \quad (4)$$

where discriminators try to predict 1 for real images and 0 for generated images. We found that $N_1 = N_2 = N_3 = 1$ is sufficient to train the generator to produce realistic images. However, higher performance might be achieved with more discriminators e.g. 2 large, 8 medium and 32 small discriminators. The generator learns to minimize the adversarial squared difference loss,

$$L_{\text{adv}} = \frac{1}{N} \sum_{i=1}^N \frac{1}{N_i} D_i(G(I_{\text{scan}}) - 1)^2, \quad (5)$$

by outputting completions that look realistic to discriminators.

Discriminators only assess the realism of generated images; not if they are correct. To lift degeneracy and prevent mode collapse, we condition adversarial training on non-adversarial losses. The total generator loss is

$$L_G = \lambda_{\text{adv}} L_{\text{adv}} + L_{\text{MSE}} + \lambda_{\text{aux}} L_{\text{aux}}, \quad (6)$$

where we found that $\lambda_{\text{aux}} = 1$ and $\lambda_{\text{adv}} = 5$ is effective. We also tried conditioning the second half of training on differences between discriminator imagination^{34,35}. However, we found that MSE guidance converges to slightly lower MSEs and similar structural similarity indexes⁶³ for STEM images.

Performance

To showcase ANN performance, example applications of adversarial and non-adversarial generators to 1/20 px coverage partial STEM completion are shown in Fig. 3. Adversarial completions have more realistic high-frequency spatial information and structure, and are less blurry than non-adversarial completions. Systematic spatial variation is also less noticeable for adversarial completions. For example, higher detail along spiral paths, where errors are lower, can be seen in the bottom two rows of Fig. 3 for non-adversarial completions. Inference only requires a generator, so inference times are the same for adversarial and non-adversarial completions. Single image inference time during training is 45 ms with an Nvidia GTX 1080 Ti GPU, which is fast enough for live partial scan completion.

In practice, 1/20 px scan coverage is sufficient to complete most spiral scans. However, generators cannot reliably complete micrographs with unpredictable structure in regions where there is no coverage. This is demonstrated by example applications of non-adversarial generators to 1/20 px coverage spiral and gridlike partial scans

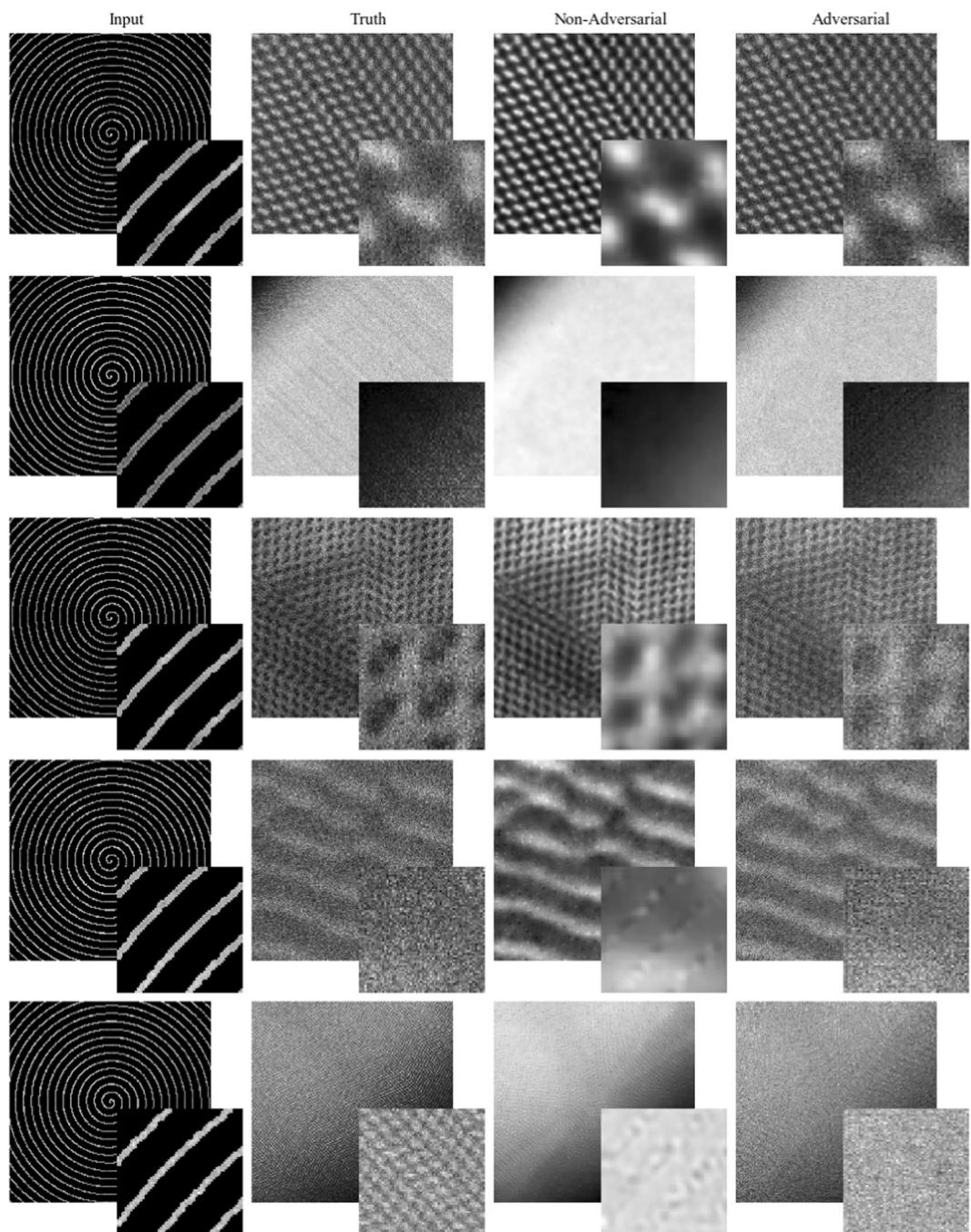


Figure 3. Adversarial and non-adversarial completions for 512×512 test set 1/20 px coverage blurred spiral scan inputs. Adversarial completions have realistic noise characteristics and structure whereas non-adversarial completions are blurry. The bottom row shows a failure case where detail is too fine for the generator to resolve. Enlarged 64×64 regions from the top left of each image are inset to ease comparison, and the bottom two rows show non-adversarial generators outputting more detailed features nearer scan paths.

in Fig. 4. Most noticeably, a generator invents a missing atom at a gap in gridlike scan coverage. Spiral scans have lower errors than gridlike scans as spirals have smaller gaps between coverage. Additional sheets of examples for spiral scans selected with binary masks are provided for scan coverages between 1/17.9 px and 1/87.0 px as Supplementary Information.

To characterize generator performance, MSEs for output pixels are shown in Fig. 5. Errors were calculated for 20000 test set 1/20 px coverage spiral scans selected with blurred masks. Errors systematically increase with increasing distance from paths for non-adversarial training, and are less structured for adversarial training. Similar to other generators^{23,64}, errors are also higher near the edges of non-adversarial outputs where there is less information. We tried various approaches to decrease non-adversarial systematic error variation by modifying loss functions. For examples: by ALRC; multiplying pixel losses by their running means; by ALRC and

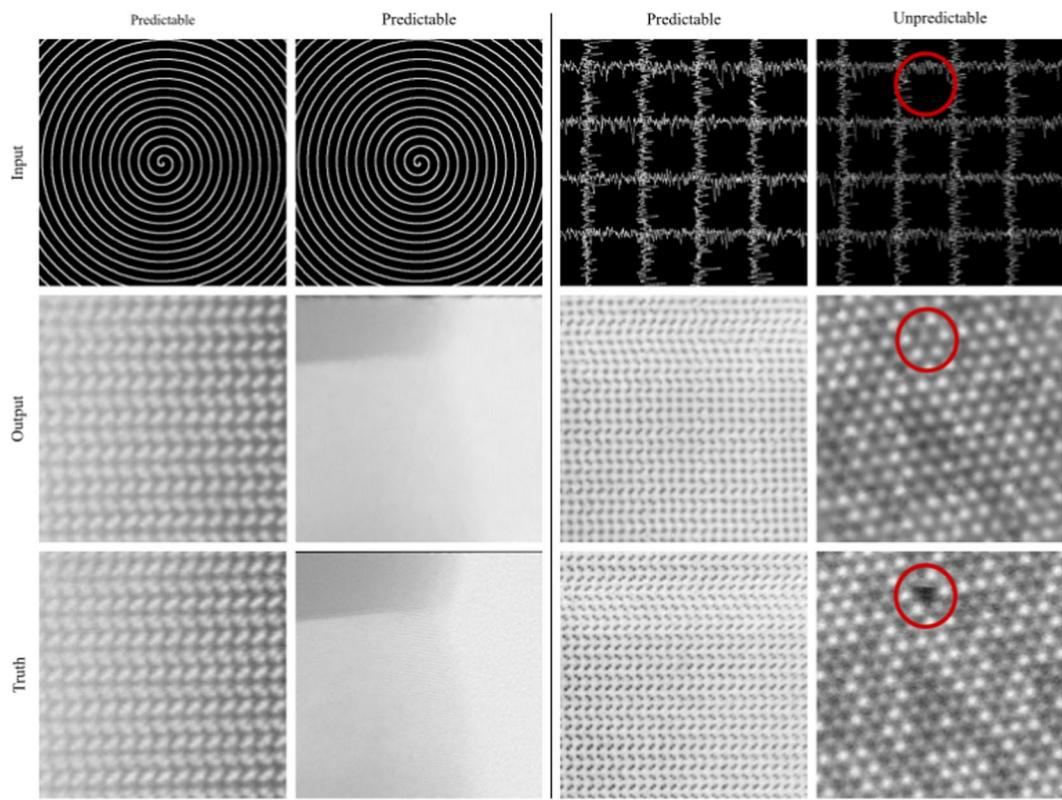


Figure 4. Non-adversarial generator outputs for 512×512 1/20 px coverage blurred spiral and gridlike scan inputs. Images with predictable patterns or structure are accurately completed. Circles accentuate that generators cannot reliably complete unpredictable images where there is no information. This figure was created with Inkscape⁸³.

multiplying pixel losses by their running means; and by ALRC and multiplying pixel losses by final mean losses of a trained network. However, we found that systematic errors are similar for all variants. This is a limitation of partial STEM as information decreases with increasing distance from scan paths. Adversarial completions also exhibit systematic errors that vary with distance from spiral paths. However, spiral variation is dominated by other, less structured, spatial error variation. Errors are higher for adversarial training than for non-adversarial training as GANs complete images with realistic noise characteristics.

Spiral path test set intensity errors are shown in Fig. 6a, and decrease with increasing coverage for binary masks. Test set errors are also presented for deep learning supersampling²³ (DLSS) as they are the only results that are directly comparable. DLSS is an alternative approach to compressed sensing where STEM images are completed from a sublattice of probing locations. Both DLSS and partial STEM results are for the same neural network architecture, learning policy and training dataset. Results depend on datasets, so using the same dataset is essential for quantitative comparison. We find that DLSS errors are lower than spiral errors at all coverages. In addition, spiral errors exponentially increase above DLSS errors at low coverages where minimum distances from spiral paths increase. Although this comparison may appear unfavourable for partial STEM, we expect that this is a limitation of training signals being imaged at several times their Nyquist rates.

Distributions of 20000 spiral path test set root mean squared (RMS) intensity errors for spiral data in Fig. 6a are shown in Fig. 6b. The coverages listed in Fig. 6 are for infinite spiral paths with 1/16, 1/25, 1/36, 1/49, 1/64, 1/81, and 1/100 px coverage after paths are cut by image boundaries; changing coverage. All distributions have a similar peak near an RMS error of 0.04, suggesting that generator performance remains similar for a portion of images as coverage is varied. As coverage decreases, the portion of errors above the peak increases as generators have difficulty with more images. In addition, there is a small peak close to zero for blank or otherwise trivial completions.

Discussion

Partial STEM can decrease scan coverage and total electron dose by $10\text{--}100\times$ with 3–6% test set RMS errors. These errors are small compared to typical STEM noise. Decreased electron dose will enable new STEM applications to beam-sensitive materials, including organic crystals⁶⁵, metal-organic frameworks⁶⁶, nanotubes⁶⁷, and nanoparticle dispersions⁶⁸. Partial STEM can also decrease scan times in proportion to decreased coverage. This will enable increased temporal resolution of dynamic materials, including polar nanoregions in relaxor ferroelectrics^{69,70}, atom motion⁷¹, nanoparticle nucleation⁷², and material interface dynamics⁷³. In addition, faster scans can reduce delay for experimenters, decreasing microscope time. Partial STEM can also be a starting point for algorithms that process STEM images e.g. to find and interpret atomic positions⁷⁴.

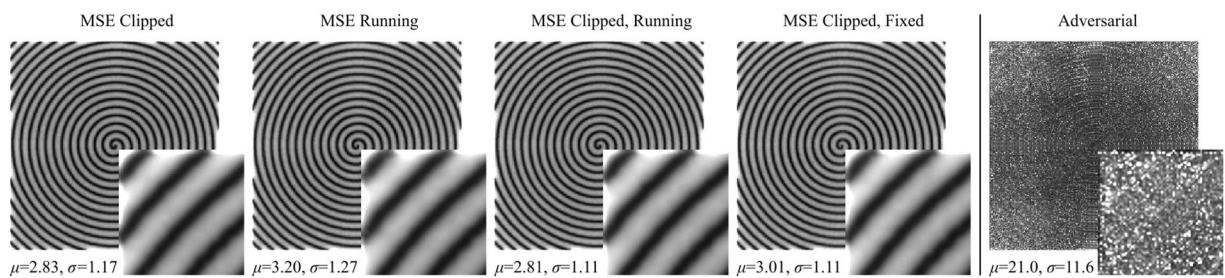


Figure 5. Generator mean squared errors (MSEs) at each output pixel for 20000 512 \times 512 1/20 px coverage test set images. Systematic errors are lower near spiral paths for variants of MSE training, and are less structured for adversarial training. Means, μ , and standard deviations, σ , of all pixels in each image are much higher for adversarial outputs. Enlarged 64 \times 64 regions from the top left of each image are inset to ease comparison, and to show that systematic errors for MSE training are higher near output edges.

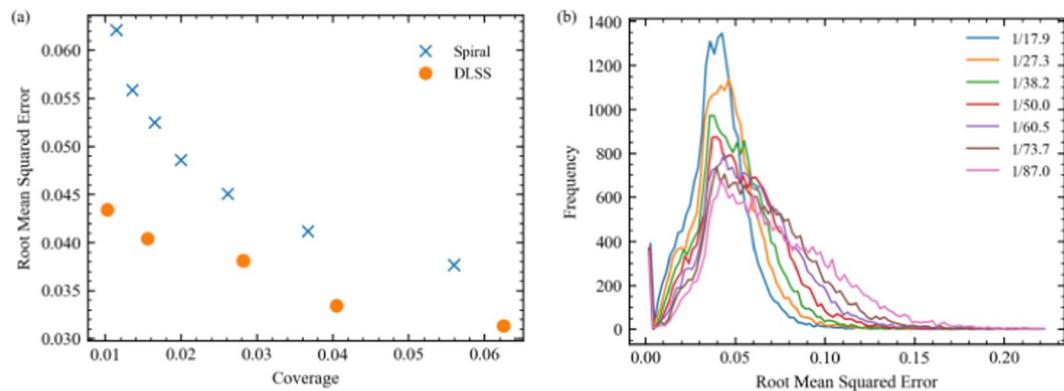


Figure 6. Test set root mean squared (RMS) intensity errors for spiral scans in [0, 1] selected with binary masks. (a) RMS errors decrease with increasing electron probe coverage, and are higher than deep learning supersampling²³ (DLSS) errors. (b) Frequency distributions of 20000 test set RMS errors for 100 bins in [0, 0.224] and scan coverages in the legend.

Our generators are trained for fixed coverages and 512 \times 512 inputs. However, recent research has introduced loss function modifications that can be used to train a single generator for multiple coverages with minimal performance loss²³. Using a single GAN improves portability as each of our GANs requires 1.3 GB of storage space with 32 bit model parameters, and limits technical debt that may accompany a large number of models. Although our generator input sizes are fixed, they can be tiled across larger images; potentially processing tiles in a single batch for computational efficiency. To reduce higher errors at the edge of generator outputs, tiles can be overlapped so that edges may be discarded⁶⁴. Smaller images could be padded. Alternatively, dedicated generators can be trained for other output sizes.

There is an effectively infinite number of possible partial scan paths for 512 \times 512 STEM images. In this paper, we focus on spiral and gridlike partial scans. For a fixed coverage, we find that the most effective method to decrease errors is to minimize maximum distances from input information. The less information there is about an output region, the more information that needs to be extrapolated, and the higher the error. For example, we find that errors are lower for spiral scans than gridlike scans as maximum distances from input information are lower. Really, the optimal scan shape is not static: It is specific to a given image and generator architecture. As a result, we are actively developing an intelligent partial scan system that adapts to inputs as they are scanned.

Partial STEM has a number of limitations relative to DLSS. For a start, partial STEM may require a custom scan system. Even if a scan system supports or can be reprogrammed to support custom scan paths, it may be insufficiently responsive. In contrast, DLSS can be applied as a postprocessing step without hardware modification. Another limitation of partial STEM is that errors increase with increasing distance from scan paths. Distances from continuous scan paths cannot be decreased without increasing coverage. Finally, most features in our new STEM crops dataset are sampled at several times their Nyquist rates. Electron microscopists often record images above minimum sufficient resolutions and intensities to ease visual inspection and limit the effects of drift⁷⁵, shot¹⁷, and other noise. This means that a DLSS lattice can still access most high frequency information in our dataset.

Test set DLSS errors are lower than partial STEM errors for the same architecture and learning policy. However, this is not conclusive as generators were trained for a few days; rather than until validation errors diverged from training errors. For example, we expect that spirals need more training iterations than DLSS as

nearest neighbour infilled spiral regions have varying shapes, whereas infilled regions of DLSS grids are square. In addition, limited high frequency information in training data limits one of the key strengths of partial STEM that DLSS lacks: access to high-frequency information from neighbouring pixels. As a result, we expect that partial STEM performance would be higher for signals imaged closer to their Nyquist rates.

To generate realistic images, we fine-tuned partial STEM generators as part of GANs. GANs generate images with more realistic high-frequency spatial components and structure than MSE training. However, GANs focus on semantics; rather than intensity differences. This means that although adversarial completions have realistic characteristics, such as high-frequency noise, individual pixel values differ from true values. GANs can also be difficult to train^{76,77}, and training requires additional computation. Nevertheless, inference time is the same for adversarial and non-adversarial generators after training.

Encouragingly, ANNs are universal approximators⁷⁸ that can represent⁷⁹ the optimal mapping from partial scans with arbitrary accuracy. This overcomes the limitations of traditional algorithms where performance is fixed. If ANN performance is insufficient or surpassed by another method, training or development can be continued to achieve higher performance. Indeed, validation errors did not diverge from training errors during our experiments, so we are presenting lower bounds for performance. In this paper, we compare spiral STEM performance against DLSS. It is the only method that we can rigorously and quantitatively compare against as it used the same test set data. This yielded a new insight into how signals being imaged above their Nyquist rates may affect performance discussed two paragraphs earlier, and highlights the importance of standardized datasets like our new STEM images dataset. As machine learning becomes more established in the electron microscopy community, we hope that standardized datasets will also become established to standardize performance benchmarks.

Detailed neural network architecture, learning policy, experiments, and additional sheets of examples are provided as Supplementary Information. Further improvements might be made with AdaNet⁸⁰, Ludwig⁸¹, or other automatic machine learning⁸² algorithms, and we encourage further development. In this spirit, we have made our source code³⁶, a new dataset containing 16227 STEM images^{40,41}, and pre-trained models publicly available. For convenience, new datasets containing 161069 non-overlapping 512×512 crops from STEM images used for training, and 19769 antialiased 96×96 area downsampled STEM images created for faster ANN development, are also available.

Conclusions

Partial STEM with deep learning can decrease electron dose and scan time by over an order of magnitude with minimal information loss. In addition, realistic STEM images can be completed by fine-tuning generators as part of a GAN. Detailed MSE characteristics are provided for multiple coverages, including MSEs per output pixel for 1/20 px coverage spiral scans. Partial STEM will enable new beam sensitive applications, so we have made our source code, new STEM dataset, pre-trained models, and details of experiments available to encourage further investigation. High performance is achieved by the introduction of an auxiliary trainer network, and adaptive learning rate clipping of high losses. We expect our results to be generalizable to SEM and other scan systems.

Data availability

New STEM datasets are available on our publicly accessible datarserver^{40,41}. Source code for ANNs and to create images is in a GitHub repository with links to pre-trained models³⁶. For additional information contact the corresponding author (J.M.E.).

Received: 12 February 2020; Accepted: 28 April 2020;

Published online: 20 May 2020

References

- Yankovich, A. B., Berkels, B., Dahmen, W., Binev, P. & Voyles, P. M. High-Precision Scanning Transmission Electron Microscopy at Coarse Pixel Sampling for Reduced Electron Dose. *Adv. Struct. Chem. Imaging* **1**, 2 (2015).
- Peters, J. J. P., Apachitei, G., Beanland, R., Alexe, M. & Sanchez, A. M. Polarization Curling and Flux Closures in Multiferroic Tunnel Junctions. *Nat. Commun.* **7**, 13484 (2016).
- Hujsak, K., Myers, B. D., Roth, E., Li, Y. & Dravid, V. P. Suppressing Electron Exposure Artifacts: An Electron Scanning Paradigm with Bayesian Machine Learning. *Microsc. Microanal.* **22**, 778–788 (2016).
- Egerton, R. F., Li, P. & Malac, M. Radiation Damage in the TEM and SEM. *Micron* **35**, 399–409 (2004).
- Jones, L. *et al.* Managing Dose-, Damage- and Data-Rates in Multi-Frame Spectrum-Imaging. *Microscopy* **67**, i98–i113 (2018).
- Trampert, P. *et al.* How Should a Fixed Budget of Dwell Time be Spent in Scanning Electron Microscopy to Optimize Image Quality? *Ultramicroscopy* **191**, 11–17 (2018).
- Anderson, H. S., Ilic-Helms, J., Rohrer, B., Wheeler, J. & Larson, K. Sparse Imaging for Fast Electron Microscopy. In *Computational Imaging XI*, vol. 8657, 86570C (International Society for Optics and Photonics, 2013).
- Stevens, A., Yang, H., Carin, L., Arslan, I. & Browning, N. D. The Potential for Bayesian Compressive Sensing to Significantly Reduce Electron Dose in High-Resolution STEM Images. *Microscopy* **63**, 41–51 (2013).
- Stevens, A. *et al.* A Sub-Sampled Approach to Extremely Low-Dose STEM. *Appl. Phys. Lett.* **112**, 043104 (2018).
- Hwang, S., Han, C. W., Venkatakrishnan, S. V., Bouman, C. A. & Ortalan, V. Towards the Low-Dose Characterization of Beam Sensitive Nanostructures via Implementation of Sparse Image Acquisition in Scanning Transmission Electron Microscopy. *Meas. Sci. Technol.* **28**, 045402 (2017).
- Candes, E. & Romberg, J. Sparsity and Incoherence in Compressive Sampling. *Inverse Probl.* **23**, 969 (2007).
- Kovarik, L., Stevens, A., Liyu, A. & Browning, N. D. Implementing an Accurate and Rapid Sparse Sampling Approach for Low-Dose Atomic Resolution STEM Imaging. *Appl. Phys. Lett.* **109**, 164102 (2016).
- Sang, X. *et al.* Dynamic Scan Control in STEM: Spiral Scans. *Adv. Struct. Chem. Imaging* **2**, 6 (2017).
- Béché, A., Goris, B., Freitag, B. & Verbeeck, J. Development of a Fast Electromagnetic Beam Blanker for Compressed Sensing in Scanning Transmission Electron Microscopy. *Appl. Phys. Lett.* **108**, 093103 (2016).
- Li, X., Dyck, O., Kalinin, S. V. & Jesse, S. Compressed Sensing of Scanning Transmission Electron Microscopy (STEM) with Nonrectangular Scans. *Microsc. Microanal.* **24**, 623–633 (2018).

16. Sang, X. *et al.* Precision Controlled Atomic Resolution Scanning Transmission Electron Microscopy using Spiral Scan Pathways. *Sci. Reports* **7**, 43585 (2017).
17. Seki, T., Ikuhara, Y. & Shibata, N. Theoretical Framework of Statistical Noise in Scanning Transmission Electron Microscopy. *Ultramicroscopy* **193**, 118–125 (2018).
18. Wu, X. *et al.* Deep Portrait Image Completion and Extrapolation. *IEEE Transactions on Image Process.* (2019).
19. Liu, G. *et al.* Image Inpainting for Irregular Holes using Partial Convolutions. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 85–100 (2018).
20. Yang, W. *et al.* Deep Learning for Single Image Super-Resolution: A Brief Review. *IEEE Transactions on Multimed.* (2019).
21. Fang, L. *et al.* Deep Learning-Based Point-Scanning Super-Resolution Imaging. *bioRxiv* 740548 (2019).
22. de Haan, K., Ballard, Z. S., Rivenson, Y., Wu, Y. & Ozcan, A. Resolution Enhancement in Scanning Electron Microscopy using Deep Learning. *Sci. Reports* **9**, 12050, <https://doi.org/10.1038/s41598-019-48444-2> (2019).
23. Ede, J. M. Deep Learning Supersampled Scanning Transmission Electron Microscopy. *arXiv preprint arXiv:1910.10467* (2019).
24. Tan, C. *et al.* *A Survey on Deep Transfer Learning.* In *International Conference on Artificial Neural Networks*, 270–279 (Springer, 2018).
25. Raschka, S. Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning. *arXiv preprint arXiv:1811.12808* (2018).
26. Roh, Y., Heo, G. & Whang, S. E. A Survey on Data Collection for Machine Learning: A Big Data-AI Integration Perspective. *IEEE Transactions on Knowl. Data Eng.* (2019).
27. Krizhevsky, A., Nair, V. & Hinton, G. The CIFAR-10 Dataset. Online: <http://www.cs.toronto.edu/~kriz/cifar.html> (2014).
28. Krizhevsky, A. & Hinton, G. Learning Multiple Layers of Features from Tiny Images. Tech. Rep., Citeseer (2009).
29. LeCun, Y., Cortes, C. & Burges, C. MNIST Handwritten Digit Database. AT&T Labs, online: <http://yann.lecun.com/exdb/mnist> (2010).
30. Russakovsky, O. *et al.* ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015).
31. Goodfellow, I. *et al.* Generative Adversarial Nets. In *Advances in Neural Information Processing Systems*, 2672–2680 (2014).
32. Bang, D. & Shim, H. MGGAN: Solving Mode Collapse using Manifold Guided Training. *arXiv preprint arXiv:1804.04391* (2018).
33. Mirza, M. & Osindero, S. Conditional Generative Adversarial Nets. *arXiv preprint arXiv:1411.1784* (2014).
34. Wang, T.-C. *et al.* High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 8798–8807 (2018).
35. Larsen, A. B. L., Sønderby, S. K., Larochelle, H. & Winther, O. Autoencoding Beyond Pixels using a Learned Similarity Metric. *arXiv preprint arXiv:1512.09300* (2015).
36. Ede, J. M. Partial STEM Repository. Online: <https://github.com/Jeffrey-Ede/partial-STEM>, <https://doi.org/10.5281/zenodo.3662481> (2019).
37. Gatan. Gatan Microscopy Suite. Online: www.gatan.com/products/tem-analysis/gatan-microscopy-suite-software (2019).
38. Landau, H. Sampling, Data Transmission, and the Nyquist Rate. *Proc. IEEE* **55**, 1701–1706 (1967).
39. Adobe Developers Association *et al.* TIFF Revision 6.0. Online: www.adobe.io/content/dam/udp/en/open/standards/tiff/TIFF6.pdf (1992).
40. Ede, J. M. STEM Datasets. Online: <https://github.com/Jeffrey-Ede/datasets/wiki> (2019).
41. Ede, J. M. Warwick Electron Microscopy Datasets. *arXiv preprint arXiv:2003.01113* (2020).
42. Ophus, C., Ciston, J. & Nelson, C. T. Correcting Nonlinear Drift Distortion of Scanning Probe and Scanning Transmission Electron Microscopies from Image Pairs with Orthogonal Scan Directions. *Ultramicroscopy* **162**, 1–9 (2016).
43. Sang, X. & LeBeau, J. M. Revolving Scanning Transmission Electron Microscopy: Correcting Sample Drift Distortion Without Prior Knowledge. *Ultramicroscopy* **138**, 28–35 (2014).
44. Krause, F. F. *et al.* ISTEM: A Realisation of Incoherent Imaging for Ultra-High Resolution TEM Beyond the Classical Information Limit. In *European Microscopy Congress 2016: Proceedings*, 501–502 (Wiley Online Library, 2016).
45. Hartel, P., Rose, H. & Dinges, C. Conditions and Reasons for Incoherent Imaging in STEM. *Ultramicroscopy* **63**, 93–114 (1996).
46. Neyshabur, B., Bhojanapalli, S., McAllester, D. & Srebro, N. Exploring Generalization in Deep Learning. In *Advances in Neural Information Processing Systems*, 5947–5956 (2017).
47. Kawaguchi, K., Kaelbling, L. P. & Bengio, Y. Generalization in Deep Learning. *arXiv preprint arXiv:1710.05468* (2017).
48. Ruder, S. An Overview of Gradient Descent Optimization Algorithms. *arXiv preprint arXiv:1609.04747* (2016).
49. Abadi, M. *et al.* TensorFlow: A System for Large-Scale Machine Learning. In *OSDI*, vol. 16, 265–283 (2016).
50. McCann, M. T., Jin, K. H. & Unser, M. Convolutional Neural Networks for Inverse Problems in Imaging: A Review. *IEEE Signal Process. Mag.* **34**, 85–95 (2017).
51. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems*, 1097–1105 (2012).
52. Szegedy, C. *et al.* Going Deeper with Convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–9 (2015).
53. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2818–2826 (2016).
54. Durugkar, I., Gemp, I. & Mahadevan, S. Generative Multi-Adversarial Networks. *arXiv preprint arXiv:1611.01673* (2016).
55. Isola, P., Zhu, J.-Y., Zhou, T. & Efros, A. A. Image-to-Image Translation with Conditional Adversarial Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1125–1134 (2017).
56. Kingma, D. P. & Ba, J. ADAM: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980* (2014).
57. Zou, D., Cao, Y., Zhou, D. & Gu, Q. Stochastic Gradient Descent Optimizes Over-Parameterized Deep ReLU Networks. *arXiv preprint arXiv:1811.08888* (2018).
58. Ede, J. M. & Beanland, R. Adaptive Learning Rate Clipping Stabilizes Learning. *Mach. Learn. Sci. Technol.* (2020).
59. Wang, Z., She, Q. & Ward, T. E. Generative Adversarial Networks: A Survey and Taxonomy. *arXiv preprint arXiv:1906.01529* (2019).
60. Dong, H.-W. & Yang, Y.-H. Towards a Deeper Understanding of Adversarial Losses. *arXiv preprint arXiv:1901.08753* (2019).
61. Miyato, T., Kataoka, T., Koyama, M. & Yoshida, Y. Spectral Normalization for Generative Adversarial Networks. *arXiv preprint arXiv:1802.05957* (2018).
62. Mao, X. *et al.* Least Squares Generative Adversarial Networks. In *Proceedings of the IEEE International Conference on Computer Vision*, 2794–2802 (2017).
63. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Process.* **13**, 600–612 (2004).
64. Ede, J. M. & Beanland, R. Improving Electron Micrograph Signal-to-Noise with an Atrous Convolutional Encoder-Decoder. *Ultramicroscopy* **202**, 18–25 (2019).
65. Sari, M., Cattle, J., Hondow, N., Brydson, R. & Brown, A. Low Dose Scanning Transmission Electron Microscopy of Organic Crystals by Scanning Moiré Fringes. *Micron* **120**, 1–9 (2019).
66. Mayoral, A., Mahugo, R., Sánchez-Sánchez, M. & Díaz, I. Cs-Corrected STEM Imaging of Both Pure and Silver-Supported Metal-Organic Framework MIL-100 (Fe). *ChemCatChem* **9**, 3497–3502 (2017).
67. Gnanasekaran, K., de With, G. & Friedrich, H. Quantification and Optimization of ADF-STEM Image Contrast for Beam-Sensitive Materials. *Royal Soc. Open Sci.* **5**, 171838 (2018).

68. Ilett, M., Brydson, R., Brown, A. & Hondow, N. Cryo-Analytical STEM of Frozen, Aqueous Dispersions of Nanoparticles. *Micron* **120**, 35–42 (2019).
69. Kumar, A., Dhall, R. & LeBeau, J. M. *In Situ* Ferroelectric Domain Dynamics Probed with Differential Phase Contrast Imaging. *Microsc. Microanal.* **25**, 1838–1839 (2019).
70. Xie, L. *et al.* Static and Dynamic Polar Nanoregions in Relaxor Ferroelectric $\text{Ba}(\text{Ti}_{1-x}\text{Sn}_x)\text{O}_3$ System at High Temperature. *Phys. Rev. B* **85**, 014118 (2012).
71. Aydin, C. *et al.* Tracking Iridium Atoms with Electron Microscopy: First Steps of Metal Nanocluster Formation in One-Dimensional Zeolite Channels. *Nano Lett.* **11**, 5537–5541 (2011).
72. Hussein, H. E. *et al.* Tracking Metal Electrodeposition Dynamics from Nucleation and Growth of a Single Atom to a Crystalline Nanoparticle. *ACS Nano* **12**, 7388–7396 (2018).
73. Chen, S. *et al.* Atomic Structure and Migration Dynamics of $\text{MoS}_2/\text{Li}_x\text{MoS}_2$ Interface. *Nano Energy* **48**, 560–568 (2018).
74. Ziatdinov, M. *et al.* Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations. *ACS Nano* **11**, 12742–12752 (2017).
75. Jones, L. & Nellist, P. D. Identifying and Correcting Scan Noise and Drift in the Scanning Transmission Electron Microscope. *Microsc. Microanal.* **19**, 1050–1060 (2013).
76. Salimans, T. *et al.* Improved Techniques for Training GANs. In *Advances in Neural Information Processing Systems*, 2234–2242 (2016).
77. Liang, K. J., Li, C., Wang, G. & Carin, L. Generative Adversarial Network Training is a Continual Learning Problem. *arXiv preprint arXiv:1811.11083* (2018).
78. Hornik, K., Stinchcombe, M. & White, H. Multilayer Feedforward Networks are Universal Approximators. *Neural Networks* **2**, 359–366 (1989).
79. Lin, H. W., Tegmark, M. & Rolnick, D. Why does Deep and Cheap Learning Work so Well? *J. Stat. Phys.* **168**, 1223–1247 (2017).
80. Weill, C. *et al.* AdaNet: A Scalable and Flexible Framework for Automatically Learning Ensembles. *arXiv preprint arXiv:1905.00080* (2019).
81. Molino, P., Dudin, Y. & Miryala, S. S. Ludwig: A Type-Based Declarative Deep Learning Toolbox. *arXiv preprint arXiv:1909.07930* (2019).
82. He, X., Zhao, K. & Chu, X. AutoML: A Survey of the State-of-the-Art. *arXiv preprint arXiv:1908.00709* (2019).
83. Harrington, B. *et al.* Inkscape 0.92, Online: <http://www.inkscape.org/> (2020).

Acknowledgements

Thanks go to Julie Robinson for advice on finding publication venues and to Marin Alexe for helpful discussion. J.M.E. and R.B. acknowledge EPSRC grant EP/N035437/1 for financial support. In addition, J.M.E. acknowledges EPSRC Studentship 1917382.

Author contributions

J.M.E. proposed this research, wrote the code, collated training data, performed experiments and analysis, created repositories, and co-wrote this paper. R.B. supervised and co-wrote this paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-65261-0>.

Correspondence and requests for materials should be addressed to J.M.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020