

4 – Machine Learning for Phase Retrieval from 4D-STEM Data

Michael Cao¹, Michael Matty¹, Zhen Chen¹, Li Li², Eun-Ah Kim¹ and David Muller¹

¹Cornell University, Ithaca, New York, United States, ²Google Research, Mountain View, California, United States

Modern direct electron detectors enable recording the full scattering distribution from an electron beam focused to atomic dimensions at imaging speeds, producing 4-dimensional, phase-space data sets. Current data rates are Gb/s, with Tb/s expected shortly as detector speeds increase, following a Moore's-law-like scaling as the underlying silicon technology improves [1,2]. Encoded in these large and growing data sets are both the phase and amplitude of the exit wave, from which the expectation values of physical operators can be reconstructed to learn the structure and fields in the probed sample. However, real space images of atomic structure need to be reconstructed from the scattering data. Applying a linear filter such as an annular dark field (ADF) mask quickly generates an interpretable image at the cost of discarding a large fraction of the data. On the other hand, ptychography takes full advantage of the diffraction data to significantly improve signal-to-noise and resolution but is computationally expensive and cannot be applied to data from thick samples.

We propose using machine learning to create a fast process that can transform large 4D datasets into high-resolution 2D images. As a training dataset, we used experimental diffraction data from monolayer MoS₂ collected on the Electron Microscope Pixel Array Detector (EMPAD) developed at Cornell [1]. The EMPAD's 1,000,000:1 dynamic range and single electron detection allows us to give the neural network the full diffraction data without any electron count loss due to saturation. For the training label, we used a ptychographic reconstruction of the MoS₂ (Fig. 1c) [2]. By using a high-resolution label, we trained a neural network that is capable of outputting high resolution images. For our test data, we used an experimental data set from a thick PrScO₃ cross-section. PrScO₃ is a stronger scatterer and has a different lattice symmetry from MoS₂. This is a test case on the machine learning based reconstruction of data from a thick sample that is outside the reach of ptychography.

As a baseline, we use a linear model with two special dense layers for our network (Fig. 1a). We choose to input a cluster of diffraction patterns rather than a single one to incorporate locality into the model. The first layer weighs the individual patterns before stacking them. The second layer applies a single filter to the stack and sums them. The output is the pixel intensity of the real-space scanning position where the cluster is centered. Both layers are forced to be radially symmetric to avoid overfitting to any particular crystal symmetry. The results of the linear network (Fig. 2b) outputs MoS₂ images that have better contrast than ADF but are not as good as ptychography. The PrScO₃ test output, however, is worse than its respective ADF.

The linear model is more transparently understood from its learned filters but is limited in its ability to create high resolution images. Therefore, we move to implementing a more advanced 4D convolutional network. While its learning is more opaque, it already naturally considers the local structure of the diffraction data across scanning points. We use two 4D convolutional layers with a

pooling layer after each convolution. Then we add a dropout layer to combat overfitting and a final dense layer to output an image intensity (Fig 1b). While more computationally intensive to train compared to the linear model, the results for both training and testing are significantly improved (Fig 2c). The MoS₂ output essentially matches the ptychography training label and the transfer to the PrScO₃ test set shows higher contrast and resolution compared to ADF.

These results show promise in the possibility of generating high-resolution images on the quality of ptychography using machine learning. More importantly, once the initial cost of training is taken, trained neural networks generate images orders of magnitude faster (seconds versus hours) than ptychographic algorithms. The 4D convolutional model has been successful, but further optimization to the hyperparameters as well as testing for its capability on a larger array of different samples can significantly improve the neural network's performance [3].

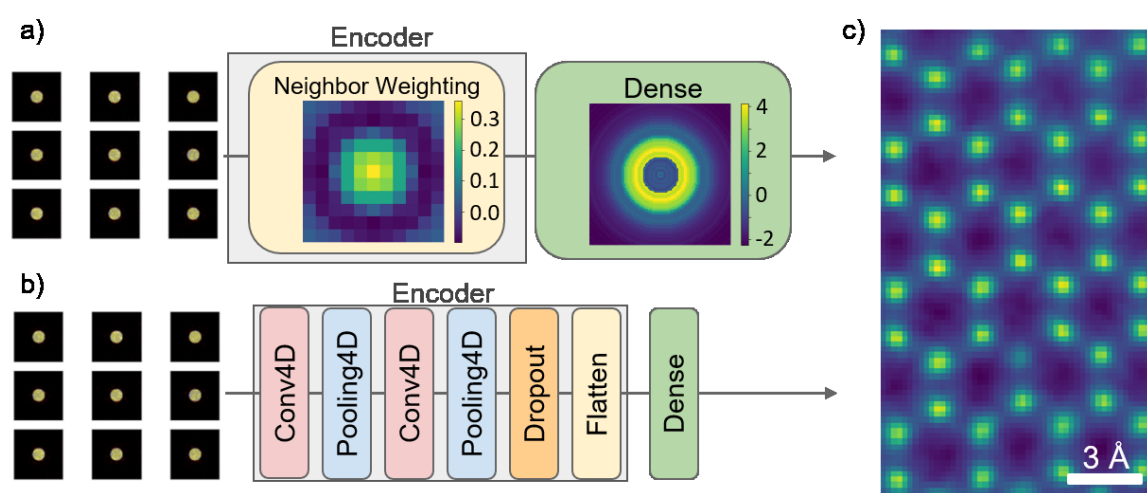


Figure 1. Figure 1: Neural network setup. a) Schematic of the linear model with examples of the trained masks used to generate the output. b) Schematic of the convolutional model. c) An MoS₂ monolayer was used as the diffraction training set with a ptychographic reconstruction as the output label.

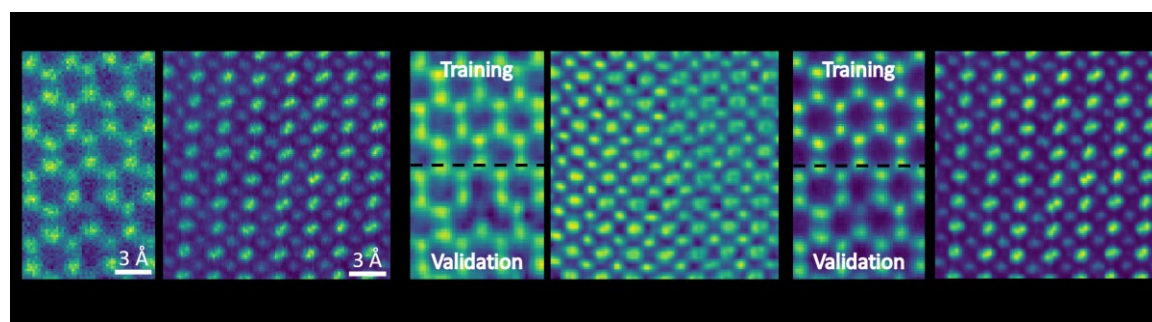


Figure 2. Figure 2: MoS₂ and PrScO₃ images. Column a) shows conventional ADF images as a reference to gauge performance. b) Images output from the linear model. c) Images output from the convolutional model. The training label was the top half of figure 1c and did not include the sulfur monovacancy in the lower half.

References

- [1] M. W. Tate et al., *Microscopy and Microanalysis* 22 (2016), p. 237-249.
- [2] Y. Jiang, Y. et al., *Nature* 559, (2018), 343–349.

[3] This work is supported by the NSF (Platform for the Accelerated Realization, Analysis, and Discovery of Interface Materials (PARADIM)) under Cooperative Agreement No. DMR-1539918 and DOE DE-SC0018946, with facilities support from the Cornell Center for Materials Research, an NSF MRSEC (DMR-1719875).