2976
doi:10.1017/S1431927622011151
Microsc. Microanal. 28 (Suppl 1), 2022

How to Get Something Out of Nothing (Almost!): Extracting Information from Noisy Data

Peter A. Crozier[1],

[1.] School for the Engineering of Matter, Transport, and Energy, Arizona State University, Tempe, Arizona)
* Corresponding author: crozier@asu.edu

The availability of high readout-rate, high sensitivity, direct electron detectors will have a major impact on our ability to understand not only the spatial but also the temporal structure of materials systems. Areas where spatiotemporal information is essential for fundamental understanding of functionality include phase changes, transport and chemical conversion processes. For example, in heterogeneous catalysis, reactant molecules chemisorb onto nanoparticles surfaces and are transformed into product molecules. The breaking and making of chemical bonds is associated with structural dynamics in the particles such as fluxionalities in strain as well as atomic diffusion processes. Understanding this fluxionality and how it relates to functionality will be key to designing new and improved catalysts. However, extracting subtle structural information from time-resolved electron microscopy images is not straightforward. The short exposure times needed for high temporal resolution, usually results in very low signal-to-noise (SNR) in individual frames and new data processing approaches must be developed to extract scientifically useful information.

In our current projects, we perform *in situ* and *operando* experiments on a Thermo Fisher Titan environmental transmission electron microscope (ETEM) fitted with a Gatan K3 direct electron detector. We typically run the detector between 40 and 100 frames per second so our data is often in the form of large movies featuring metal or ceramic nanoparticles. Techniques such as blob detection are reasonably robust in the presence of noise and have been employed to determine atomic column locations and intensities [1]. As the temperature and gas environment are varied, a nanoparticle may undergo changes in shape, orientation and composition. The system is often stochastic with periods of relative stability punctuated with periods of intense structural dynamics. To provide a quantitative description of dynamic evolution, event detection algorithms must be developed that can function with noisy data. Exponential weighted averaging combined with time derivatives is a promising approach to rapidly identify the onset of structural instability (see Gilankar et al in these proceedings).

Machine learning offers more wholistic and diverse approaches for processing large datasets. The convolutional neural networks that we are developing required extensive training. If prior information is available on the images, for example, they show metal atomic columns, supervised networks can provide impressive performance when they work. Training requires a large number of simulated images from likely structures and defects in the system of interest. To denoise images of Pt nanoparticles supported on CeO_2, we have developed a network trained on a series of almost 20,000 images from 900 structures. The image simulations (performed with Dr. Probe) are implemented on HPC systems to provide high throughput. The network performance can be benchmarked against subsets of the simulated structures and statistical metrics (likelihood maps) can be generated to provide quantitative confidence levels in the output from experimental datasets [2].

One advantage of supervised networks is that they do not require large volumes of real data. One weakness is the concern that the simulated training data may not accurately capture the underlying ground truth for the experiment. However, if very large experimental datasets are available, it becomes possible to train networks directly on experimental data. This unsupervised approach to training has the potential to denoise TEM images from any structure provided enough experimental data is available. We are testing the performance of unsupervised networks for denoising large *in situ* image time series. During a typical *in situ* experiment, terabytes of data may be recorded from samples under nominally identical conditions. The potential value of the unsupervised approach is that, with large enough datasets, the network is train directly on the material of interest and no external biases from simulation are present. The preliminary output from such networks appears encouraging although it is necessary to develop criteria to objectively evaluate their performance (see Morales et al these proceedings).

The purpose of microscopy is to solve materials problems and segmentation can be helpful to translate EM signals into higher level models of materials structure, microstructures or compositions. We have an interest in complex oxides where dynamic distributions of oxygen vacancies may evolve in ways that are not easy to predict. For example, during surface oxygen exchange, the oxygen vacancy creation changes at different rates with time according to the location. Interpreting the anion and cation column intensity variations in time can be challenging because they depend on thickness, tilt, and defocus as well as column occupancy. Effects, such as channeling, cause non-linear intensity variations which are in general different for anions and cations. We have been developing a segmentation approach to translate image intensity from cations and anions directly to column occupancy. This is accomplished using a supervised neural network which is trained on structural models with different point defect configurations and corresponding simulated images covering a wide range of thickness, tilt and defocus. Such an approach allows local cation and anion dynamics to be rapidly quantified in long time series on samples of varying thickness. The network is able to manage the moderate variation in tilt and defocus and performs well even when the data is very noisy (see Leibovich et al these proceedings). The advantages and disadvantages of this image segmentation approach will be discussed [3].

References:

[1] R. Manzorro et al., Microscopy and Microanalysis (in press) (2022).
[2] JLVincent et al., Microscopy and Microanalysis (2021), p. 1.
[3] We gratefully acknowledge the support of the following NSF grants (OAC 1940263, OAC 1940097, CBET 1604971 and DMR 184084). We also acknowledge the support from DOE grant BES DE-SC0004954. The author acknowledges HPC resources available through ASU, and NYU as well as the John M. Cowley Center for High Resolution Electron Microscopy at ASU.