Denoising STEM Electron Energy Loss Spectra using Convolutional Autoencoders

Mark Oxley[1], Maxim Ziatdinov[2] and Sergei Kalinin[1]

[1]Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee, United States, [2]Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, United States

Core-shell STEM EELS has become a powerful tool for the probing of complex materials, providing not only chemical mapping, but information about local bonding and electronic states. Due to the low cross sections for core-shell ionization the spectra recorded as a function of probe position have an inherently low signal to noise ratio. This problem has been most commonly addressed using methods such as principal component analysis [1]. In this talk we will examine the use of a simple convolutional autoencoders to denoise both simulated and experimental spectra

To provide a test of this approach we will first use simulated O K-shell spectra form strontium titanate calculated using a combination of dynamical scattering theory and density functional theory [2]. This simulated dataset is shown in Fig 1 with the spectrum image shown in Fig. 1 (a), the average spectrum in Fig. 1 (b) and selected spectra from different probe positions shown in Fig. 1 (c). Gaussian noise was added to the spectra and the corresponding results are shown in Fig. 1 (d)-(f). The inset in Fig 1 (d) is the mean squared error (MSE) of the noisy spectrum image compared with the perfect case. The individual spectra in Fig. 1 (f) exhibit a significant amount of noise. It is these datasets that will be used to train the autoencoder. The autoencoder was trained for a total of 750 iterations.

 In Fig. 2 we show the results of using the autoencoder to denoise the spectra shown in Fig. 1. The spectrum image in Fig. 2 (a) has had the MSE reduced by over half, though some speckling is still present. However, the individual spectra shown in Fig. 2 (c) are in excellent visual agreement with the original spectra in Fig. 1 (c). Since the noisy dataset formed part of the training procedure, it is not unusual that the agreement is excellent. We have found that if we attempt to clean the same simulated dataset with different levels of noise applied, the results are not in such good agreement. This could lead to errors in measurements of peak positions and relative peak heights often used in spectral analysis.

There are of course additional issues with applying this method to experimental datasets. For the simulated dataset the spectra had the same shape and intensity distribution as the training set. In addition, the level of noise on the experimental dataset need to be accurately determined. In this presentation we will examine the pitfalls of using a poorly developed training set for denoising spectra and how to avoid these problems using a systematic approach to experimental data. We will demonstrate the application of this standard simulated dataset to experimental EELS data, and how the simulated data can be adjusted to provide a close match to the parameters of the experimental dataset. Both the python notebooks implementing the data processing and autoencoder itself, along with the simulated dataset will be made generally available so other researchers can readily apply this methodology to their own datasets.
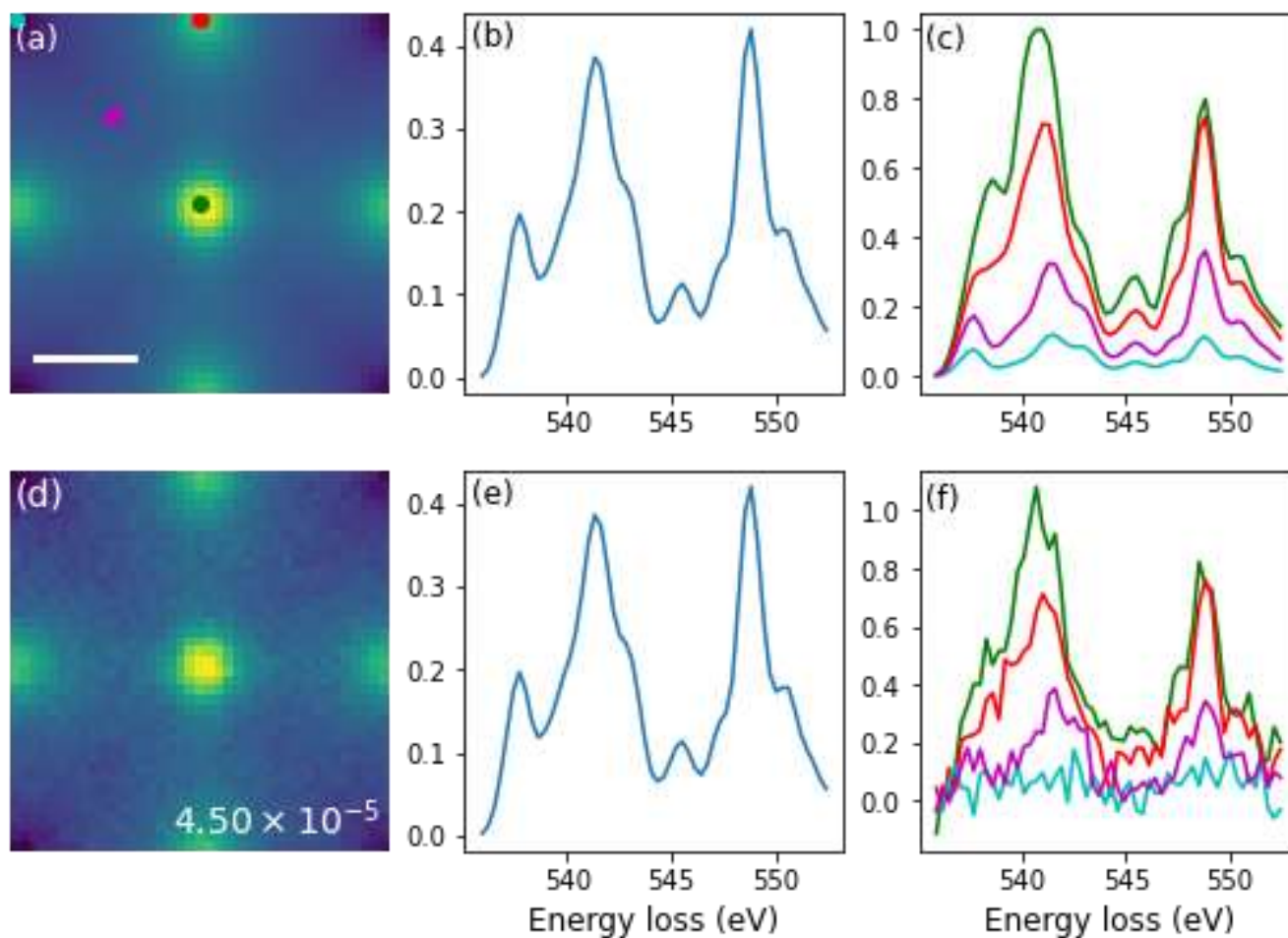
Figure 1. Simulated O K-shell EELS data set for STO. (a) The spectrum image formed by summing over all energies. (b) The average spectrum. (c) Individual spectra taken from the positions indicated on (a). The data set has been normalized between 0 and 1. With random noise added as discussed in the text. The standard deviation of the normal distribution used is 0.05. (d) The noisy spectrum image. (e) The average of the noisy spectra. (f) Individual noisy spectra taken from the positions indicated on (a). The scale bar on (a) is 1 Å and the inset in (d) is the MSE.
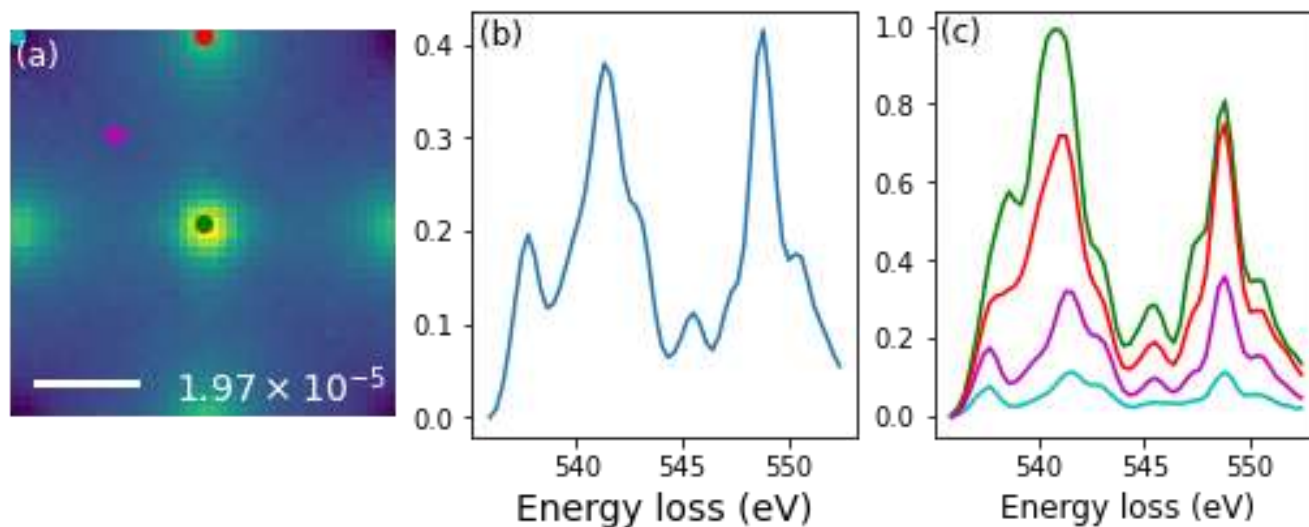
Figure 2. Cleaned O K-shell EELS data set for STO for $\sigma = 0.05$. (a) The spectrum image formed by summing over all energies. (b) The average spectrum. (c) Individual spectra taken from the positions indicated on (a). The scale bar on (a) is 1 Å and the inset is the MSE.

References
[1] Bonnet, N., Journal of Microscopy **190** (1998), p. 2.
[2] Oxley, M. P. *et al.*; Microscopy and Microanalysis **20** (2014), p. 784.