

Separating physically distinct mechanisms in complex infrared plasmonic nanostructures via machine learning enhanced electron energy loss spectroscopy

Sergei V. Kalinin, Kevin Roccapriore, Shin Hum Cho, Delia J. Milliron, Rama Vasudevan, Maxim Ziatdinov, and Jordan A. Hachtel**

Dr. Sergei V. Kalinin, Dr. Kevin M. Roccapriore, Dr. Rama Vasudevan, Dr. Maxim Ziatdinov, Dr. Jordan A. Hachtel
Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
E-mail: sergei2@ornl.gov, hachtelja@ornl.gov

Dr. Shin Hum Cho, Dr. Delia J. Milliron
McKetta Department of Chemical Engineering, The University of Texas at Austin, Austin, TX 78712, USA

Dr. Shin Hum Cho
Samsung Electronics, Samsung Semiconductor R&D, Hwaseong, Gyeonggi-do 18448, Republic of Korea

Keywords: scanning transmission electron microscopy, electron energy loss spectroscopy, nanoparticle arrays, machine learning, plasmonics, nanophotonics

Abstract

Electron energy loss spectroscopy (EELS) enables direct exploration of plasmonic phenomena at the nanometer level. To isolate individual plasmon modes, linear unmixing methods can be used to separate different physical mechanisms, but in larger and more complex systems the interpretability of the components becomes uncertain. Here, infrared plasmonic resonances in self-assembled heterogeneous monolayer films of doped-semiconductor nanoparticles are examined beyond linear unmixing techniques, and both supervised and unsupervised machine-learning-based analyses of hyperspectral EELS datasets are demonstrated. In the supervised approach, a human operator labels a small number of pixels in the hyperspectral dataset corresponding to features of interest which are then propagated across the entire dataset. In the unsupervised approach, non-linear autoencoders are used to create a highly-reduced latent-space representation of the dataset, within which insight into the relevant physics can be gleaned from straight-forward distance metrics that do not depend on operator input and bias. The advantage of these approaches is that the labeling

separates physical mechanisms without altering the data, enabling robust analyses of the influence of heterogeneities in mesoscale complex systems.

1. Introduction

The field of nanoplasmonics centers on the confinement of optical excitations to the surfaces of nanoscale structures, enabling highly controllable detection and sensing,^[1] enhancement of photovoltaic responses,^[2] novel biomedical applications,^[3] and potentially all-optical circuitry.^[4] As systems become increasingly complex the need for direct nanoscale experimental investigation has made electron energy loss spectroscopy (EELS) within a scanning transmission electron microscope (STEM) a key technique due its ability to sample the true localization of the confinement at its native length scale with hyperspectral imaging.^[5-8] However, the delocalized nature of plasmons results in significant overlap of distinct modes, both spatially and spectrally, often obscuring the true intensities and localizations.

As a result, there is an inherent need within plasmonic STEM-EELS analysis for exploratory data analysis methods that can both reduce the dimensionality of the EELS data sets and isolate relevant physics mechanisms that are tied to specific geometric features such as the corners of a nanoparticle, the edges of a film, etc. For core-loss EELS, the early work by Bosman^[9] and others^[10-13] following the visionary set of publications by Bonnett^[14,15] introduced linear unmixing methods, most notably principal component analysis (PCA). Generally, linear unmixing methods represent the data as a linear combination of components (or endmembers) and corresponding loading maps (or weights). This approach reduces the dimensionality of the data set and allows visualization of regions with dissimilar behaviors. Much of the early criticisms of PCA-EELS was centered at the fact that the second and higher-order PCA components have negative values, i.e., they represented clearly non-

physical behaviors since the EELS signal is formed by detector counts. The subsequent introduction of non-negative matrix factorization (NMF) and Bayesian linear unmixing^[16,17] introduced the constraint of non-negativity, which is a better physical analog to EELS^[16] and resulted in more physical components of the decomposition, especially for the case of plasmonics. More generally, additional physical constraints such as sum to one, spatial smoothness, and sparsity can be introduced,^[18] confining the behavior of the components and loading maps to those that are more likely to be physically possible. It should be noted, however, that similar to other machine learning (ML) techniques, the components are strictly speaking defined only in a statistical sense and therefore, physical interpretation remains preponderantly qualitative unless specific physical constraints are incorporated explicitly.

Linear unmixing methods are generally characterized by two additional limitations. The first is that the spatial structure does not factor into the analysis, meaning that random transposition of the spatial pixels will not affect the component maps (unless explicit spatial regularization is used). The introduction of structured kernel Gaussian process (GP)-based analysis methods, where the interpolation specifically takes into account the spatial grid structure of the measurement^[19-21] or 3D convolutional neural networks can address this issue and explore both energy and spatial domains.^[22] The second limitation is that linear unmixing methods produce highly meaningful results when the physics of the system can, with a high veracity, be approximated as a sum of a number of (unknown) linear components.^[23,24] This assumption holds well for the core-loss EELS signals, where the signal-generation comes from the interaction of the sub-Ångstrom electron beam with the nuclei of individual atomic species and also for simple plasmonic systems, which can be approximated straightforwardly by a small number of plasmon modes. However, for complex systems with a wide variety of modes with highly varying and often overlapping intensity, frequency, and localization, the

unmixing can result in unphysical components, where weaker modes can appear only as modifications of the dominant modes making physical interpretation challenging.

Here, we propose a pathway for the supervised analysis of EELS datasets to separate and classify regions of the films with physically distinct spectral responses. The classifications are shown to be robust, to accurately capture the common spatio-spectral tropes of the complex nanostructures, and to be transferable between different datasets to allow high-throughput analysis of large areas of the sample. We also demonstrate an unsupervised version of the approach using non-linear autoencoders (AE) combined with clustering in the latent space of the system AE to derive the highly reduced representations of the system response that yield insight into the relevant physics, that do not depend on operator input and bias. Most critically, these methodologies represent a significant advance over the traditional PCA/NMF analyses used in low-loss EELS by providing direct access to the spatial distributions of components in a hyperspectral dataset without modification of the spectral information. The combination enables directly interpretable separation of distinct nanoscale physical mechanisms.

The typical problem emerging in the context of the EELS data analysis is classification, i.e., interpretation of spectra at each pixel as belonging to a specific class of materials behaviors, as shown in **Figure 1** (a). Here, the network is trained using known labeled data (e.g., a subset of the image with known pixel-wise labels) and then labels can be propagated over a full data set or data obtained under similar conditions. The initial labels can be generated based on known morphological features, recognized spectral signatures, etc. Label propagation can be further applied to regions where spatial features are unavailable or where spectral images are obtained at lower spatial resolution, providing insight into the local physical mechanisms.

An alternative workflow for EELS data analysis is based on the AE approach where the pixel-wise EELS data is dimensionally reduced via a set of convolutional filters and fully connected (dense) layers into a low-dimensional latent space (bottleneck) then subsequently mapped back (deconvolve) into the original shape. The network training aims to achieve the best match between the input and output, which ideally will coincide with the bottleneck serving as the compression. The latent features can provide insight into characteristic behaviors of EELS data via spatial visualization, clustering, and density-based distance metrics, somewhat analogously to PCA and NMF analyses but occurring in the latent space of learned features. The key assumption of this analysis is that if the full dataset can be represented by a small number of latent variables then the variation of these variables across the initial dataset provides insight into the relevant physics. Note that in addition to the applications explored here, AEs are extremely powerful tools for denoising, hyperspectral image reconstruction, state compression, and other applications.^[25]

2. Results and Discussion

For this work we examine self-assembled arrays of fluorine- and tin-doped indium oxide nanoparticles (FT:IO). These doped-semiconductor nanoparticles (NPs) are a developing class of materials that natively supports infrared plasmon resonances and are highly tunable in terms of their size, shape, and plasmonic response (e.g., localized plasmon frequency) through variations in the fluorine and/or tin doping concentration.^[26,27] Moreover, while the self-assembled arrays have many interesting applications^[28] they are difficult to treat theoretically due to the inherent heterogeneity of particle morphology and spacing that results from the colloidal synthesis and assembly processes.^[29] While, perfect regular arrays can (and have) be treated using numerical simulations[cite the JCP article], accurately treating non-periodic heterogeneity, such as missing particles, variations in spacing, breaks in the film,

and large-scale holes ~~can all dramatically affect array properties, involving~~ can involve hundreds of NPs, ~~and~~ rendering direct simulation or calculation of the plasmonic response untenable. Therefore, the combination of EELS, which can access the ~~real~~ plasmonic response at these nanoscale heterogeneities, and ML, which can classify and separate physical mechanisms in these complex structures, is an ideal pathway toward effective plasmonic analysis in self-assembled arrays.

The arrays of FT:IO NPs are self-assembled via the liquid-air interface method on TEM grids with ultrathin SiN_x membranes (~5 nm). Another key challenge in doped semiconductor plasmonics is access to the native infrared plasmon resonances in the FT:IO particles; ergo we must use monochromation to achieve the necessary spectral resolution with EELS. The monochromated EELS measurements are performed using a Nion HERMES UltraSTEM 100 operated at an accelerating voltage of 60 kV. The monochromator on this instrument has a variable slit allowing us to reduce the absolute energy resolution for increased signal. While at maximum monochromation the instrument can achieve an energy resolution on the scale of 5-6 meV, we find plasmon linewidths in the regime of ~100 meV, so an energy resolution of ~50 meV is chosen to maximize signal without significantly broadening the observed plasmonic features. All EELS acquisitions presented here are hyperspectral datasets, or spectrum images (SI). In an SI, the beam is rastered across the region of interest and an EEL spectrum is acquired at each probe position, resulting in a 3D dataset with two spatial dimensions and one spectral dimension.

There are two critical pre-processing steps for the data: the removal of the zero-loss peak (ZLP) and removal of X-ray outliers. The ZLP is the peak in the EEL spectrum containing all elastic scattering and non-interacting (no energy loss) electrons and for thin samples it is by far the dominant feature in the EEL spectrum (typically 3-5 orders of magnitude larger than the low-loss signal); thus, small variations in the ZLP are dominant

mathematically with respect to the real plasmonic peaks. This can be done in a number of ways,^[30] either by simply cutting the spectral axis off at a channel above the ZLP or by performing a power law fit to the tail of the ZLP and subtracting it from each pixel in the SI. Furthermore, datasets must be cleaned to remove the small clusters of spectrometer channels with signal orders of magnitude higher than the surrounding channels that are caused by stray X-rays hitting the scintillator in the spectrometer. This is straightforwardly achieved by averaging with the surrounding unaffected pixels. Once the spectra are processed, the two spatial dimensions can be flattened into a 1D array to provide the samples for NMF^[31] decomposition.

Here, we use NMF, since the non-negative constraint offers a much better physical match to EELS compared to PCA and hence the components and abundances are more representative of the genuine features of the EELS dataset. The NMF decomposition is shown for two regions in the FT:IO films in **Figure 2**, one from a regular part of the array (referred to as Regular Array – RA) and one from a more heterogeneous part of the array (referred to as Heterogeneous Array – HA). High angle annular dark field (HAADF)-STEM reference images are shown in Figure 2 (a,b) for the RA and HA, respectively, and show that some of the native heterogeneity in the self-assembly process can be seen in the RA image, but the HA dataset has actual missing particles in the array and is acquired at the edge of the film presenting heterogeneity with fundamentally different physical features.

We perform a relatively simplistic 4-component decomposition to demonstrate conventional NMF plasmon decomposition in hyperspectral datasets. The decomposition produces both a spectral endmember and a spatial abundance map for each component, such that the real spectrum at each pixel is approximately a linear combination of the component's spectra weighted by the component's spatial abundance at that pixel, which are shown in Figure 2 (c,d) for the RA and HA decompositions, respectively. The NMF decomposition

effectively highlights different regions of the self-assembled films and associates them with a spectrum. ~~However, while~~ The two datasets are acquired on the same sample, and hence we observe some similarities and some differences ~~one would hope for similar components, for the most part the spectral and spatial components are highly different.~~ For the similarities, in both samples there is a strong peak at around ~ 0.5 eV, this is the array plasmon mode generated by the collective resonances of the individual particles in the array. Since both datasets come from the same sample, the array plasmon mode has the same frequency and linewidth in both, which is observed in the similar peaks in Fig. 2c(1) and 2d(1-3). ~~The only component that is similar is the one~~ The other notable similarity between the two datasets, is that both NMF analyses produce a component localized to the particles, ~~(component 2 of the RA and component 4 of the HA~~ Fig. 2c(2) and 2d(4), which are similar in both the abundance maps and the spectral endmembers) ~~while the others bear little similarity.~~ This The differences between the NMF decompositions occurs because in complex systems, such as this one, multiple different phenomenon can be absorbed into single NMF components. Furthermore, ~~and~~ the addition of heterogeneity in the HA dataset adds new aspects of the spectral response that are folded into the decomposition breaking the comparability of the RA and HA components. Moreover, there may be nuanced behaviors that cannot be accurately accounted for by a small number of components and are omitted entirely from the NMF decomposition. Indeed, by correlating the NMF components to the raw dataset it is observed that energy correlations can still be obtained for over 100 components (Supporting Information **Figure S1**).

Furthermore, for systems such as these where spatial components overlap (note how Fig. 2c(1) and 2d(1) both have non-zero minima in the colorbars), the EELS signal at these overlapped probe positions is split between components. As a result, the spectral components do not accurately represent the spectra in the regions highlighted by the spatial abundance

map. For example, the spatial maps in Fig. 2c(2) and 2d(4) are localized to the nanoparticles (not the gaps/holes/edge/outside regions), however the spectral components are nonphysical because the more dominant components in Fig. 2c(1) and 2d(1) still have significant intensity inside the nanoparticles as well. Thus, we can see that while NMF can produce intuitive and physical components for simple systems, in complex systems with varying degrees of heterogeneity the decomposition is often unphysical and breaks the EELS signal into non-interpretable pieces. This is the critical difference of the ML approach presented in this manuscript from conventional linear unmixing: the labeling only separates mechanisms spatially not spectrally, so the resulting spectral components represent the true ~~plasmonic response~~ EELS signal for the labeled pixels in the hyperspectral dataset of the system.

However, even for complex systems, NMF is an excellent exploratory approach to guide a more rigorous analysis. For instance, we note that a large fraction of the spatial pixels of the NMF decomposition shown for the HA in Figure 2 can be associated with specific physical features. This allows us to form an intuition for what the real physics in the system should be and we can start referring to larger scale features of the dataset with labels guided by NMF. For instance, we can see that the particles and gaps between the particles are highlighted in different NMF components and that the hole in the film and the edge of the film are highlighted in different NMF components, leading us to believe that we can understand the physics of the system by focusing on these regions individually. While the true ~~plasmonic response~~ EELS signal of the array is more nuanced than this discrete labeling, it is of fundamental importance to classify different regions of the film according to the spectral differences (if any) observed between them. However, we need to move beyond NMF to achieve a physical description of the plasmonic response in each of these regions, which drives the desire for the ML-based approach utilized here. The goal is to classify specific pixels in the hyperspectral dataset with distinctly different ~~plasmonic responses~~ EELS signals.

Once the different regions are classified based on the ML labeling, we can average the spectra from these different regions to compare the genuine EELS data from each classification with the knowledge that they represent different physical mechanisms.

There is inherent interest in both supervised and unsupervised ML methodologies in order to make use of the hard-won expertise of researchers in the field and to side-step the biases that come with that expertise for a clear objective analysis of the data. Thus, we will perform separation of the physical mechanisms through two methods: label propagation (supervised) and latent space clustering (unsupervised).

We begin with label propagation. To achieve robust label propagation we adapt the approach for recognition imaging originally proposed for scanning probe microscopy (SPM) by Nikiforov et al.^[32] a decade ago. The operator selects pixels in the hyperspectral dataset they view as representative of a specific classification; these labels are then used to train the network. Subsequently, the trained network propagates those labels through the rest of the dataset to identify the other pixels that are most similar to the operator assigned classification. This presents an excellent use for the NMF decompositions as we can use the fact that specific physical features are highlighted with high intensity in the abundance maps to guide the creation of masks to isolate specific physical features that we have a mathematical reason to believe are fundamentally distinct. For instance, in examining the HA dataset NMF decomposition, one can simply use intensity thresholds on the NMF abundance maps to define six different labels corresponding to different physical features of the sample. It is important to note that the choice of six labels is not objective but is guided by the operator's expertise and the intuition gained from the NMF decomposition. Here, for the HA dataset we define the following six labels: particle centers, particle edges, array gaps, array holes, array edges, and the region outside the film. It is critical to avoid overlap (e.g., the same

hyperspectral pixel assigned to two different labels by the operator) so the masks are chosen and refined such that each pixel has only one label.

We implemented the fully connected multilayer perceptron (MLP) with four hidden layers of 30 neurons each, *tanh* activation for internal layers and *softmax* for the output layer. The dimensionality of the input layer is determined by the number of the NMF components and was chosen to be 24. The dimensionality of the output layer is determined by the number of labels (the six labels identified in the previous paragraph). Adam optimizer for accuracy metrics and categorical cross entropy loss was used as is normal for classification problems. The network was implemented in Keras^[33] with TensorFlow-2 on the back end. While the hyperparameters of the network can in principle be optimized more systematically, the accuracy was found to exceed 99.7% after the initial optimization and was assumed to be adequate for the task. We further note that a similar approach can be implemented using the convolutional networks on spectra (1D input) or local subsets of EELS data (3D convolutional inputs); however, given the volume of EELS data sets these approaches are unnecessary for the present case but can be expected to become important for larger data volumes in, for example, 4D-STEM.

Figure 3 shows MLP label propagation for human-assigned classifications in the HA dataset. To start we use label propagation as a means of refining the human labeled masks. There are many conventions that an operator could use to define the initial pixels for the operator labeling. While one could directly assign each individual pixel to a label based off of the operator's intuition, this would be prohibitive for large hyperspectral datasets and not conducive to tuning and adjusting parameters to try to identify hidden features. Here, we have used the simplistic NMF decompositions shown in Fig. 2. By thresholding the intensity of the different spatial components, we can create Boolean masks of particles in different regions (i.e. take the spatial abundance map shown in Fig. 2d(1) and set a threshold at an intensity of

20, this would result in a mask that was localized entirely in the gaps between particles in the array).

Figure 3 (a) shows the different masks produced when the operator assigns a classification to every single pixel in the dataset based on the NMF decompositions, each color representing one operator classification. The result of the MLP reconstruction is shown in Figure 3 (b). Here, we can see that the two are nearly identical with only a few pixels different in the refined mask. Such a result is not so surprising as the initial masks were chosen based on NMF decompositions in the first place and hence were chosen based on mathematical relevance. Nevertheless, it is encouraging to see MLP verify the human labels as viable classifications of distinct spectral behavior.

The key benefit of the MLP approach is that the total labeling shown in Figure 3 (a,b) is not necessary for a successful label propagation process. We need only a few pixels to successfully drive the separation of the different classifications. Here, we tighten the thresholding used on the NMF abundance maps in the mask creation process to generate human labeling where only a handful of pixels is used to train the MLP and the rest are unlabeled. The partial labels are shown in Figure 3 (c), where the black pixels represent unlabeled data in the hyperspectral dataset. We apply the same MLP label propagation process and observe that the MLP network returns an almost identical labeling of the overall dataset. There are two benefits to the match between Figure 3 (b) and (d): reduction of time spent on operator-labeling and the removal of bias from the operator labeling. In order to achieve the labeling shown in Figure 3 (a) the operator must carefully choose the thresholds to optimize the labels and in the event of overlapping labels determine which pixels go with which label, introducing a large degree of bias in the system. Since only a small number of pixels are required to achieve rigorous labeling, the operator can be less deliberate with the thresholding of the NMF components or potentially only select a few key areas manually for

the initial labels (without using NMF at all). This enables more regions to be examined in the same time period, without the result being assumed in advance and without sacrificing rigor. This is critical for high-throughput analysis of specimens such as those where there may be many different regions that must be analyzed to understand the composite behavior of the mesoscale array. By further restricting the number of labeled pixels we can see that meaningful label propagation is achieved with networks trained on only 0.31% of the total dataset (**Figure S2**). Only a few key representative pixels need to be chosen and the MLP label propagation can identify all other like pixels in the dataset with comparable accuracy to a human operator. This enables the ability to rigorously explore datasets where the operator's intuition only leads to being able to accurately classify a few pixels and allows the operator to experiment with different labels to try and identify the labels that most physically represent the real data.

While the identification of comparable regions through label propagation is important, the critical aspect of this process is to separate different physical mechanisms based on the spectral response. Thus, the important question to ask is 'How physical are the propagated labels?' We address this question in **Figure 4** using the bag-of-features approach (where the network is only trained on a few of the operator-identified labels). By comparing the representative spectra for the propagated labels, we can see how they match the spectrum from the pixels selected by the operator. The better the match, the more accurately the MLP has identified a truly unique physical mechanism and its localization.

We first train the network on only two labels (particle centers and array gaps). The operator-labels and the MLP propagated masks are shown in Figure 4 (a,b). Even with only two labels, MLP propagation results in a reasonably physical result; all gaps are labeled as gaps, all particles are labeled as particles. The only problem area is the outside region, which is unphysically labeled a particle. However, either label would have been a bad classification

of this region, so the unphysical result is unavoidable due to insufficient labels. The importance of this unphysical labeling can be seen in the representative spectra shown in Figure 4 (c) for the particles and 4 (d) for the gaps. Both regions show significant differences between the human-labeled spectrum and the ML-labeled spectrum. In the particle spectra, we see that the shoulder at 0.7 eV from the human-labeled dataset is suppressed in the ML dataset and the 0.9 eV peak is almost absent. In the array spectra, the 0.7 eV shoulder is also suppressed in the MLP spectrum and the dominant peak at 0.45 eV is slightly noticeably redshifted compared to the human-labeled peak. These differences indicate that each of the MLP propagated labels contain pixels exhibiting a different physical response from the pixels in the initial human label and that the propagation is unphysical.

The failure of the two-label propagation in Figure 4 (a-d) follows necessarily from not including sufficient labels in the MLP training. We perform the same analysis including the ‘outside’ and ‘array hole’ labels in the MLP training and the results are shown in Figure 4 (e,f). By even a cursory examination of the MLP labels we observe that a more physical result has been achieved as now the pixels outside the film are labeled as such (as opposed to particle centers) and that the hole in the center of the array is labeled as such (as opposed to an array gap). As a result we see a much better match between the representative spectra for the four-label classification in Figure 4 (g,h). We see a reasonably strong match in the intensity of the 0.7 eV shoulder for both regions and the redshift of the ~ 0.45 eV peak is gone. Both improvements clearly arise from having incorporated physical mechanisms localized to the hole/edge/outside regions into labels corresponding to particle centers and array gaps. With more labels comes a better physical representation of the label in both the representative spectrum and its spatial profile in the dataset. We still see a significant difference in the particle centers at the 0.9 eV peak, so we return to the six-label classification (now including the particle edge and array edge labels used in Figure 3) for Figure 4 (i-l).

Here, we see the reason why including separate labels for the particle center and particle edge is critical, since the 0.9 eV peak for the particle center spectrum matches quantitatively with the human-labeled data. This peak corresponds to the high energy non-radiative breathing mode that is localized to the volume of the particles. The degree of localization to the volume is significant and the mode is only excited strongly at the center of the particle and not at the particle edges and surfaces⁷. The lack of plasmonic signals over the outer edge of the nanoparticle is ~~which is~~ consistent with depletion region effects macroscopically measured for this type of material^[34], but never before been directly observed at the nanoscale. ~~Thus, when we exclude the edges from the particle center label, we achieve a more accurate representation of the physical mechanism and the MLP and human-labeled data match more closely.~~ Similarly, the array gap label is not affected by the inclusion of the array edge label because the array edge region was labeled as outside in the four-label classification; thus, the separation between the outside and the array edge does not influence the spectrum of the array gap. It is important to note that even the six-label classification is incomplete and imperfect and misses some of the complexity and nuance of the ground truth of the system, but the label-propagation approach enables robust repeatable separation of different mechanisms in a complex system without any decomposition into components, which can be unphysical or that cut-out low significance information. However, by comparing the spectra of the initially labeled pixels to the final propagated labeled pixels we have a mechanism to robustly determine if any significant effects are missing. As seen in Fig. 4, without accounting for the depletion region at the edges of the cubes a strong match between the initial/propagated labels is not achieved, demonstrating how we can use the supervised ML approach to extract new insights out of complex materials systems.

We can now use the trained MLP network to separate other hyperspectral datasets into the desired labels without the human operator label identification or the MLP training steps.

We test the transferability of the MLP network between the HA and RA dataset and vice versa. The two datasets are from the same sample so comparable physical mechanisms should result in comparable spectral features that could be identified by the MLP network. First, we train the network only using the particle center and array gap labels on the HA dataset, the same as in Figure 4 (a-d). We then use this network to classify the particle centers and array gaps in the RA dataset, as shown in **Figure 5** (a,b) and we observe the critical result that the MLP network can label particle centers and array gaps accurately in an entirely different dataset.

To further test the transferability, we include the other labels from the six-label classification of the HA dataset. This is a critical test, as the region has no visible array holes and no array edge or outside region, so we can test how accurately the MLP can find features even when there are labels that are not represented in the dataset. The result is shown in Figure 5 (c,d) where it is seen that the MLP network finds the particle edges and centers effectively and divides the space between the particles between the array gap and array hole labels. While initially it can seem unphysical for the gaps in the RA dataset to be labeled as holes, it is important to note that the spacing between particles is actually greater in the areas labeled holes (at the top of the dataset) and possesses some misoriented particles while the areas labeled gaps (towards the bottom) are more regular. This indicates that the spectral response of areas with larger spacings and less regularity in the array are more physically connected to holes in the array than they are to the standard gaps of a regular array. The transfer also shows that labels that are inarguably not present in the RA (the array edge and outside labels) are not applied to any pixels in the MLP labeling. This is extremely encouraging since the difference between a hole and a gap is partly subjective (how big does a gap have to be before it becomes a hole?) but the end of the film is more tangible. The fact

that no significant weight is imparted to labels that are clearly not present in the dataset validates the transferability of the MLP training between datasets.

The process is also reversible. Figure 5 (e,f) show the two-label training from the RA dataset (particle center and array gap) applied to the HA dataset and achieves extremely similar results to the two-label classification of the HA dataset shown in Figure 4 (a,b), where the MLP is trained directly on the HA. While it is not possible to train the MLP using all six labels from Figure 3 and Figure 4 on the RA dataset (which does not have the visible heterogeneities like the film edge and holes), we perform a three-label training (particle center, particle edge, and array gap) and see that the MLP-labeled pixels are the most accurate classification of the labels available and successfully distinguishes between the particle centers and particle edges based on the intensity of the high-energy breathing mode. It is important to note that there is significantly more noise in the RA-labeling performed with HA-trained network (Fig. 5a-d) than in when the HA-labeling with the RA-trained network (5e-h). This is likely due to the fact that the HA-trained network has more complexity due to the additional labels that do not really exist in the RA dataset. Thus, when assigning a pixel as particle edge/center or attempting to distinguish between ill-defined concepts like an array gap vs. an array hole, there is more uncertainty in the assignment (and hence more noise). When we label the HA dataset with the RA-trained network only the region outside of the array, where the EELS signal cannot be reproduced by spectral components from within the array, shows any significant noise. These results show the limitation of the technique and imply that increased complexity in the trained network and the absence of sufficient labels can negatively impact transferability.

The transferability of the trained network to other samples shows that these heterogeneous regions are repeatably distinct from one another, and that the spectrum at the edge of the array or in a hole in the array can be determined without examining a spatial map.

This demonstrates the possibility of training a ‘master MLP’ that could be applied to any dataset with any variety of heterogeneity and have those heterogeneities accurately classified and visualized for high-throughput characterization of self-assembled films.

We note that while the MLP label propagation within a single dataset is robust, the transferability of the network to other datasets is significantly less so. An example is shown in **Figure S3**, where we perform two different types of background subtraction on the RA dataset before using the MLP network trained on the HA dataset for classification. Here, if we use the same method of background subtraction where we fit a power law to the ZLP tail and subtract it from each spectrum between the HA and the RA dataset, we return a satisfactory classification of the RA dataset, but if we use a different method (cutting the spectrum off and leaving the tail) we achieve an entirely unphysical classification. Additionally, the transferability is highly sensitive to the number of NMF components used in the initial MLP training and significant variability is observed in the transferability if only a low number are used (**Figure S4**).

The alternative to a master MLP that can classify without human-operator input is a purely unsupervised approach. Here, we have adopted the convolutional AE approach, as shown in Figure 1 (b). In the encoding stage, the signal (i.e., individual spectra) is simplified via the set of 1D convolutional filters and further translated to a fully connected latent layer. In the decoding stage, the decoder aims to reconstruct the signal from the reduced lattice representation. The encoder and decoder are trained simultaneously and jointly act to find the most efficient representation of the signal in terms of the small number of the (non-linear) latent variables. Note that a linear encoder will be similar to PCA in terms of data it provides.^[35]

Here, the AE architecture is implemented using five 1D convolutional layers with the *tanh* activation function and (8, 16, 32, 64, 64) convolutional filters (kernels) with a 2D latent space. The decoder was based on the set of 1D convolutional layers with *tanh* activation and upsampling with nearest-neighbor interpolation. An additional 1D convolutional layer with the ReLU activation function was added to ensure smoothness of the output and the final layer used the sigmoid function activation. The AE was implemented with the mean squared error loss and Adam optimizer with default learning rate (=0.01).

The AE training generates a manifold with an assignable number of dimensions where each coordinate on the manifold is a spectrum and the range of potential spectra is mathematically derived to represent all latent mechanisms and distributions active in the system. As a result, we can represent each pixel in our hyperspectral dataset in terms of a Cartesian coordinate on this manifold and separate physically distinct mechanisms via distance metrics in latent space without any prior knowledge. While a number of metrics and methodologies can be applied for this classification (e.g., distance-based k-means clustering or density-based DBSCAN), we utilize Gaussian mixture modeling (GMM),^[36] in which a selectable number of clusters in latent space are formed based on unsupervised decomposition into Gaussian-like groups. This clustering method is naturally applicable to this case since the distributions in latent spaces are non-linear manifolds and as such, simple distance-based methods are prone to misclassification. An example of this 2D spectral manifold and how we use it to classify mechanisms in the dataset are shown in Supporting Information, **Figure S5**.

We use two latent variables so the latent space can be fully visualized in the 2D plane (for higher dimensional latent spaces only marginal distributions can be visualized straightforwardly and the structure is less apparent). We also use many different numbers of GMM classifications to test the robustness of the method for an insufficient number of classifications as we did with the supervised approach shown in Figure 4. **Figure 6** shows

both the latent-space clustering and the real-space labeling from the AE with (a) two GMM classifications, (b) four GMM classifications, and (c) six GMM classifications. The results for two classifications (Figure 6a) are extremely promising when compared to the supervised approach in Figure 4b. We see that unsupervised labeling easily identifies the particles and the gaps, except in this example we do not achieve the unphysical result of pixels far away from the film being classified as particles (as shown in Figure 4a). This can also be used to demonstrate the power of using advanced distance metrics in latent space, as is shown in the Supporting Information **Figure S6** where we show that k-means clustering produces the same unphysical result shown in Figure 4b, while the GMM successfully avoids this mislabeling. The four classification labeling is also a pure improvement on the supervised approach since the AE produces much smoother labeling of the dataset than the MLP, especially in labeling the hole, which is very much off-center in Figure 4f but is extremely symmetric with respect to the real geometry of the sample in Figure 6b.

Once we reach six classifications, a significant deviation between the supervised and unsupervised approach was observed. In the supervised labeling, a specific label was assigned to the particle edges to distinguish from the particle center that exhibit the breathing mode strongly, but unsupervised labeling does not return this label, instead choosing to distinguish between particles close to the edge and particles further from the edge. This difference can be understood by examining the representative spectra from the regions assigned to each label in MLP label propagation and in AE latent-space clustering.

In **Figure 7** we compare supervised MLP labeling to unsupervised AE labeling for six labels/GMM classifications. The labeling for each is shown in Figure 7 (a,b) and show that most labels are almost identical except for the two labels associated with the particles themselves. The representative spectra from these two particle labels for MLP and AE are

shown in Figure 7 (c,d), respectively, and the other labels (array gap, array edge, array hole, and outside) are shown in Figure 7 (e-h), respectively.

In the particle labels, we note the same effect as was observed in the two- and four-label MLP classifications shown in Figure 4, where the 0.9 eV breathing mode peak is significantly reduced in the AE labeling compared the MLP labeling, once again due to the lack of separation between the particle edges and centers in the final labels. However, we can understand what the AE classification selected instead of the particle center/edge distinction by examining the energy of the dominant ~ 0.5 eV peak in each of the labels. In the operator-labeled spectra both peaks occur at the same frequency, 0.497 eV; however, in the AE the major distinction between the two labels is the frequency of this peak (0.511 eV for the first label and 0.467 eV for the second). This demonstrates that the dominant plasmon peak is significantly redshifted near the edge of the film compared with the center, which is why the AE selects the particles at the edge of the film for the second label (as opposed to the edges of the actual particles).

Indeed, by looking at the frequency comparing the array gap and array edge labels we can see the same trend, even in the operator-assisted classification where the more central regions have a frequency near 0.5 eV (array gap) and the outer regions have a significantly lower frequency around 0.46 eV (array edge). This is consistent with other EELS studies on metamaterials, which found that the frequency of the dominant plasmon modes near the edge of the array were red-shifted,^[37] demonstrating that this is a real physical mechanism that is selected by the AE. We also note that by using the knowledge of the redshift toward the edge of the film and choosing the initial human labeling based on proximity to the edge, the MLP label propagation returns nearly identical classifications as the AE (**Figure S7**), confirming the validity of the result.

Figure S7 shows that we can use operator intuition to pick the correct pixels to allow for label propagation to reproduce the AE component separation, but the reverse is not true. The AE does not select the distinction between the particle edge/center due to the depletion region as significant. The reason is that the spectra in these array modes is dominated by the array plasmon mode, and subtle changes in the frequency or intensity of this peak constitute greater mathematical differences than the subtle difference in the breathing mode peak (which is an order of magnitude lower in intensity than the array mode). One can increase the number of GMM classifications, however even with 10 GMM classifications the dominant factor is still the array mode frequency and intensity (**Figure S8**). While the AE labeling is effective and robust, these results highlight the need for both supervised and unsupervised ML approaches.

3. Conclusion

To summarize, we explored the ML workflow for identification of regions exhibiting unique physical behaviors in complex plasmonic structures. While it is established that classical linear unmixing methods are extremely useful for the initial exploratory data analysis, the high environmental dependencies of complex systems leads to large numbers of components capturing multiple degenerate aspects of the plasmonic response and become increasingly difficult to interpret. Here, we showed that recognition imaging can be used to propagate partially known labels across the data set, identifying spatial regions with a common spectral response to elucidate relevant physics across the entire sampled region. This is an excellent alternative to the unmixing approach as it involves classification of unaltered data, ensuring that no information is lost (as is the case in an NMF decomposition). This approach can be further used to establish universal networks that can extrapolate across multiple data sets and enable transition toward *ex situ* analysis of high-volume EELS datasets

and automated *in situ* experimentation. We further show that such labeling can be done without the operator step by training AEs to represent an entire hyperspectral dataset in latent space such that we can effectively cluster and classify without the operator determining what they believe are the most relevant phenomena at play.

Moreover, these techniques can be universally applied for hyperspectral imaging techniques including cathodoluminescence, scanning tunneling spectroscopy, and others. We note that if the high-veracity models for the formation of the spectra are available, the networks can be trained on theoretical models and used to derive the materials parameters from experimental data.^[38-40] Furthermore, conditional AEs can be used to provide unsupervised exploration of variability within known classes^[41] and recent advances in the space of semi-supervised learning can be leveraged to propagate labels in diverse datasets with very sparse labels.^[42-44]

This distinction observed in Figure 7 demonstrates both the benefits and limitations of supervised and unsupervised approaches. Supervised approaches are subject to the operator's bias: if the operator believes that the particle edge/center distinction is more relevant than the array edge/center distinction, the ML-driven label propagation confirms this bias. However, while the unsupervised approach is subject to no such bias, it will necessarily favor the most mathematically relevant changes and can ignore and miss subtle yet important changes that a domain expert will not. Regardless, each methodology successfully separates, and isolates regions of the sample based on real physical mechanisms that influence the spectral response and the tandem approach allows an extremely rigorous exploration and visualization of such processes in hyperspectral data.

4. Experimental Section/Methods

Data and materials availability: The codes used in this work are available as a Jupyter notebook at <https://git.io/JURFs>

Materials: Indium (III) acetate ($\text{In}(\text{ac})_3$, 99.99%), Tin (IV) acetate ($\text{Sn}(\text{ac})_4$), Oleic acid (OA, 90%, technical grade), Oleyl alcohol (OIA1, 85%, technical grade) were purchased from Sigma-Aldrich. Tin (IV) fluoride (SnF_4 , 99%) was purchased from Alfa Aesar, Hexane (99.9%), Isopropyl alcohol (99.5%, Certified ACS), were purchased from Fisher Chemical. All chemicals were used as received without any further purification.

Fluorine Doped Indium Tin Oxide (FT:IO, F,Sn:In₂O₃) Nanoparticle Synthesis: All synthesis procedures are undertaken by employing standard Schlenk line techniques using a modification of previously reported methods for continuous slow injection synthesis of indium oxide nanoparticles²⁷. 29.46 $\text{In}(\text{ac})_3$ 1342.97 mg (4.6 mmol), SnF_4 48.68 mg (5%, 0.25 mmol), $\text{Sn}(\text{ac})_4$ 53.23 mg (3%, 0.15 mol), and oleic acid (10 ml) are loaded in a three-neck round-bottom flask in a N₂-filled glove box. The precursors are stirred with a magnetic bar at 600 rpm and degassed under vacuum at 120 °C for 15 min. The injection solution is added at rate of 0.2 ml/min, into 13 ml of oleyl alcohol maintained at 290 °C vented with a 19-gauge needle under inert N₂ gas flow. The reaction mixture turns blue a few minutes into the injection. Subsequently, growth is terminated by removal of the heating mantle and cooled by blowing air on the three-neck flask vessel. The nanoparticles are dispersed in hexane, then isopropyl alcohol antisolvent is added and the mixture is centrifuged at 7500 rpm for 10 min. The washing procedure is repeated 3 times and the nanoparticles are redispersed in 10 ml of hexane. The resultant nanoparticle dispersion is centrifuged at 2000 rpm for 3 min to remove non-dispersible aggregates and the supernatant is collected as the nanoparticle stock sample. A concentration series (0 and 10 % $\text{Sn}(\text{ac})_4$) of doped F,Sn:In₂O₃ nanoparticles was synthesized by controlling the $\text{Sn}(\text{ac})_4$ to $\text{In}(\text{ac})_3$

molar precursor ratio, while SnF_4 was maintained at 5% molar ratio, keeping other reaction parameters identical. All data here acquired from the 0% $\text{Sn}(\text{ac})_4$ samples.

STEM Characterization: A Nion aberration-corrected ultrahigh-energy resolution monochromated EELS-STEM (HERMESTM)^[45] operated at 60 kV, equipped with a Nion Iris spectrometer. The variable monochromator slit is set to allow an energy resolution of ~30-50 meV to achieve an optimal balance of energy resolution and beam current. For these experiments, a probe current of ~20 pA with a convergence angle of 30 mrad, a collection angle of 25 mrad in the EEL spectrometer, were used.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This effort (electron microscopy, feature extraction) is based upon work supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), Materials Sciences and Engineering Division (K.M.R., S.V.K.) and was performed and partially supported (J.A.H., M.Z., R. K. V.) at the Oak Ridge National Laboratory's Center for Nanophase Materials Sciences (CNMS), a U.S. Department of Energy, Office of Science User Facility. S.H.C and D.J.M. acknowledge (NSF CHE-1905263, and CDCM, an NSF MRSEC DMR-1720595), the Welch Foundation (F-1848), and the Fulbright Program (IIE-15151071). Electron microscopy was performed using instrumentation within ORNL's Materials Characterization Core provided by UT-Battelle, LLC, under Contract No. DE-AC05-00OR22725 with the DOE and sponsored by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

The manuscript was written through contributions of all authors. (S.V.K.) devised the experiment, set up the ML architectures, and wrote the manuscript, (K.M.R.) and (J.A.H.) performed the microscopy and analysis and helped write the manuscript, (J.A.H.) analyzed the data via the ML methods in the text and supervised the experiment, (M.Z.) and (R.V.) created the ML models used in this work and helped write the manuscript, (S.H.C.) and (D.M.) synthesized the samples and assisted in writing the manuscript. All authors have given approval to the final version of the manuscript.

Received: ((will be filled in by the editorial staff))

Revised: ((will be filled in by the editorial staff))

Published online: ((will be filled in by the editorial staff))

References

- [1] Zijlstra, P., Paulo, P. M. R. & Orrit, M. Optical detection of single non-absorbing molecules using the surface plasmon resonance of a gold nanorod. *Nat. Nanotechnol.* **7**, 379-382, doi:10.1038/nnano.2012.51 (2012).
- [2] Atwater, H. A. & Polman, A. Plasmonics for improved photovoltaic devices. *Nat. Mater.* **9**, 205-213, doi:10.1038/nmat2629 (2010).
- [3] Anker, J. N. *et al.* Biosensing with plasmonic nanosensors. *Nat. Mater.* **7**, 442-453, doi:10.1038/nmat2162 (2008).
- [4] Ebbesen, T. W., Genet, C. & Bozhevolnyi, S. I. Surface-plasmon circuitry. *Physics Today* **61**, 44-50, doi:10.1063/1.2930735 (2008).
- [5] Koh, A. L. *et al.* Electron Energy-Loss Spectroscopy (EELS) of Surface Plasmons in Single Silver Nanoparticles and Dimers: Influence of Beam Damage and Mapping of Dark Modes. *ACS Nano* **3**, 3015-3022, doi:10.1021/nn900922z (2009).
- [6] Rossouw, D., Couillard, M., Vickery, J., Kumacheva, E. & Botton, G. A. Multipolar Plasmonic Resonances in Silver Nanowire Antennas Imaged with a Subnanometer Electron Probe. *Nano Letters* **11**, 1499-1504, doi:10.1021/nl200634w (2011).
- [7] Kociak, M. & Stéphan, O. Mapping plasmons at the nanometer scale in an electron microscope. *Chemical Society Reviews* **43**, 3865-3883, doi:10.1039/C3CS60478K (2014).
- [8] Smith, K. C. *et al.* Direct Observation of Infrared Plasmonic Fano Antiresonances by a Nanoscale Electron Probe. *Phys. Rev. Lett.* **123**, 177401, doi:10.1103/PhysRevLett.123.177401 (2019).

- [9] Bosman, M., Watanabe, M., Alexander, D. T. L. & Keast, V. J. Mapping chemical and bonding information using multivariate analysis of electron energy-loss spectrum images. *Ultramicroscopy* **106**, 1024-1032, doi:10.1016/j.ultramic.2006.04.016 (2006).
- [10] Serin, V. *et al.* TEM and EELS measurements of interface roughness in epitaxial Fe/MgO/Fe magnetic tunnel junctions. *Phys. Rev. B* **79** doi:10.1103/PhysRevB.79.144413 (2009).
- [11] Kotula, P. G. & Keenan, M. R. Application of multivariate statistical analysis to STEM X-ray spectral images: Interfacial analysis in microelectronics. *Microsc. microanal.* **12**, 538-544, doi:10.1017/s1431927606060636 (2006).
- [12] Kotula, P. G., Keenan, M. R. & Michael, J. R. Tomographic spectral imaging with focused ion beam/scanning electron microscopy/energy-dispersive spectroscopy and multivariate statistical analysis: Comprehensive three-dimensional microanalysis. *Scanning* **28**, 63-63 (2006).
- [13] Samet, L. *et al.* EELS study of interfaces in magnetoresistive LSMO/STO/LSMO tunnel junctions. *European Physical Journal B* **34**, 179-192, doi:10.1140/epjb/e2003-00210-8 (2003).
- [14] Bonnet, N. Multivariate statistical methods for the analysis of microscope image series: applications in materials science. *Journal of Microscopy-Oxford* **190**, 2-18, doi:10.1046/j.1365-2818.1998.3250876.x (1998).
- [15] Bonnet, N. in *Advances in Imaging and Electron Physics, Vol 114* Vol. 114 *Advances in Imaging and Electron Physics* (ed P. W. Hawkes) 1-77 (Elsevier Academic Press Inc, 2000).

- [16] Dobigeon, N. & Brun, N. Spectral mixture analysis of EELS spectrum-images. *Ultramicroscopy* **120**, 25-34, doi:10.1016/j.ultramic.2012.05.006 (2012).
- [17] Altmann, Y., Dobigeon, N., McLaughlin, S. & Tourneret, J. Y. Nonlinear Spectral Unmixing of Hyperspectral Images Using Gaussian Processes. *IEEE Trans. Signal Process.* **61**, 2442-2453, doi:10.1109/tsp.2013.2245127 (2013).
- [18] Kannan, R. *et al.* Deep data analysis via physically constrained linear unmixing: universal framework, domain examples, and a community-wide platform. *Advanced Structural and Chemical Imaging* **4**, 20, doi:10.1186/s40679-018-0055-8 (2018).
- [19] Wilson, A. G., Hu, Z. T., Salakhutdinov, R. & Xing, E. P. in *Artificial Intelligence and Statistics, Vol 51* Vol. 51 *JMLR Workshop and Conference Proceedings* (eds A. Gretton & C. C. Robert) 370-378 (Microtome Publishing, 2016).
- [20] Gardner, J. R., Pleiss, G., Wu, R. H., Weinberger, K. Q. & Wilson, A. G. in *International Conference on Artificial Intelligence and Statistics, Vol 84* Vol. 84 *Proceedings of Machine Learning Research* (eds A. Storkey & F. PerezCruz) (Microtome Publishing, 2018).
- [21] Kalinin, S. V., Lupini, A. R., Vasudevan, R. K. & Ziatdinov, M. Gaussian process analysis of Electron Energy Loss Spectroscopy (EELS) data: parallel reconstruction and kernel control. *arXiv:2005.10507 [cond-mat, physics:physics]* (2020).
- [22] Mei, S. H. *et al.* Unsupervised Spatial-Spectral Feature Learning by 3D Convolutional Autoencoder for Hyperspectral Classification. *IEEE Trans. Geosci. Remote Sensing* **57**, 6808-6820, doi:10.1109/tgrs.2019.2908756 (2019).

- [23] Jesse, S. & Kalinin, S. V. Principal component and spatial correlation analysis of spectroscopic-imaging data in scanning probe microscopy. *Nanotechnology* **20**, 085714, doi:10.1088/0957-4484/20/8/085714 (2009).
- [24] Strelcov, E. *et al.* Deep Data Analysis of Conductive Phenomena on Complex Oxide Interfaces: Physics from Data Mining. *ACS Nano* **8**, 6449-6457, doi:10.1021/nn502029b (2014).
- [25] Tschannen, M., Bachem, O. & Lucic, M. J. a. p. a. Recent advances in autoencoder-based representation learning. (2018).
- [26] Cho, S. H. *et al.* Syntheses of Colloidal F:In₂O₃ Cubes: Fluorine-Induced Faceting and Infrared Plasmonic Response. *Chem. Mat.* **31**, 2661-2676, doi:10.1021/acs.chemmater.9b00906 (2019).
- [27] Cho, S. H. *et al.* Spectrally tunable infrared plasmonic F,Sn:In₂O₃ nanocrystal cubes. *The Journal of Chemical Physics* **152**, 014709, doi:10.1063/1.5139050 (2020).
- [28] Klinkova, A., Choueiri, R. M. & Kumacheva, E. Self-assembled plasmonic nanostructures. *Chemical Society Reviews* **43**, 3976-3991, doi:10.1039/C3CS60341E (2014).
- [29] Johns, R. W. *et al.* Direct observation of narrow mid-infrared plasmon linewidths of single metal oxide nanocrystals. *Nat. Commun.* **7**, 6, doi:10.1038/ncomms11583 (2016).
- [30] Fung, K. L. Y. *et al.* Accurate EELS background subtraction – an adaptable method in MATLAB. *Ultramicroscopy* **217**, 113052, doi:10.1016/j.ultramic.2020.113052 (2020).
- [31] Lee, D. D. & Seung, H. S. Algorithms for Non-negative Matrix Factorization. *Advances in neural information processing systems*. 556-562 (2001).

- [32] Nikiforov, M. P. *et al.* Functional recognition imaging using artificial neural networks: applications to rapid cellular identification via broadband electromechanical response. *Nanotechnology* **20**, 405708, doi:10.1088/0957-4484/20/40/405708 (2009).
- [33] Chollet, F. Keras: The Python Deep Learning library. *Astrophysics Source Code Library*, (2015). <<https://keras.io>>
- [34] Tandon, B., Agrawal, A., Heo, S. & Milliron, D. J. Competition between Depletion Effects and Coupling in the Plasmon Modulation of Doped Metal Oxide Nanocrystals. *Nano Letters* **19**, 2012-2019, doi:10.1021/acs.nanolett.9b00079 (2019).
- [35] Bao, X., Lucas, J., Sachdeva, S. & Grosse, R. Regularized linear autoencoders recover the principal components, eventually. *arXiv:2007.06731 [cs, stat]* (2020).
- [36] Yu, G., Sapiro, G. & Mallat, S. Solving Inverse Problems With Piecewise Linear Estimators: From Gaussian Mixture Models to Structured Sparsity. *IEEE Transactions on Image Processing* **21**, 2481-2499, doi:10.1109/TIP.2011.2176743 (2012).
- [37] von Cube, F. *et al.* From Isolated Metaatoms to Photonic Metamaterials: Evolution of the Plasmonic Near-Field. *Nano Letters* **13**, 703-708, doi:10.1021/nl3043757 (2013).
- [38] Ovchinnikov, O. S., Jesse, S., Binctacchit, P., Trolier-McKinstry, S. & Kalinin, S. V. Disorder Identification in Hysteresis Data: Recognition Analysis of the Random-Bond-Random-Field Ising Model. *Physical Review Letters* **103**, doi:10.1103/PhysRevLett.103.157203 (2009).
- [39] Kalinin, S. V. *et al.* Direct evidence of mesoscopic dynamic heterogeneities at the surfaces of ergodic ferroelectric relaxors. *Phys. Rev. B* **81**, 064107, doi:10.1103/PhysRevB.81.064107 (2010).

- [40] Kumar, A. *et al.* Spatially resolved mapping of disorder type and distribution in random systems using artificial neural network recognition. *Phys. Rev. B* **84**, 024203, doi:10.1103/PhysRevB.84.024203 (2011).
- [41] Kingma, D. P., Mohamed, S., Rezende, D. J. & Welling, M. Semi-supervised learning with deep generative models. *Advances in neural information processing systems*, 3581-3589 (2014).
- [42] Wu, H. & Prasad, S. Semi-Supervised Deep Learning Using Pseudo Labels for Hyperspectral Image Classification. *IEEE Trans. Image Process.* **27**, 1259-1270, doi:10.1109/tip.2017.2772836 (2018).
- [43] Dai, D., and Luc Van Gool. Unsupervised high-level feature learning by ensemble projection for semi-supervised image classification and image clustering. *arXiv preprint arXiv:1602.00955* (2016).
- [44] Yalniz, I. Z., et al. Billion-scale semi-supervised learning for image classification. *arXiv preprint arXiv:1905.00546* (2019).
- [45] Krivanek, O. L. *et al.* Monochromated STEM with a 30 meV-wide atom-sized electron probe. *Microscopy* **62**(1), doi:10.1093/jmicro/dfs089 (2013).

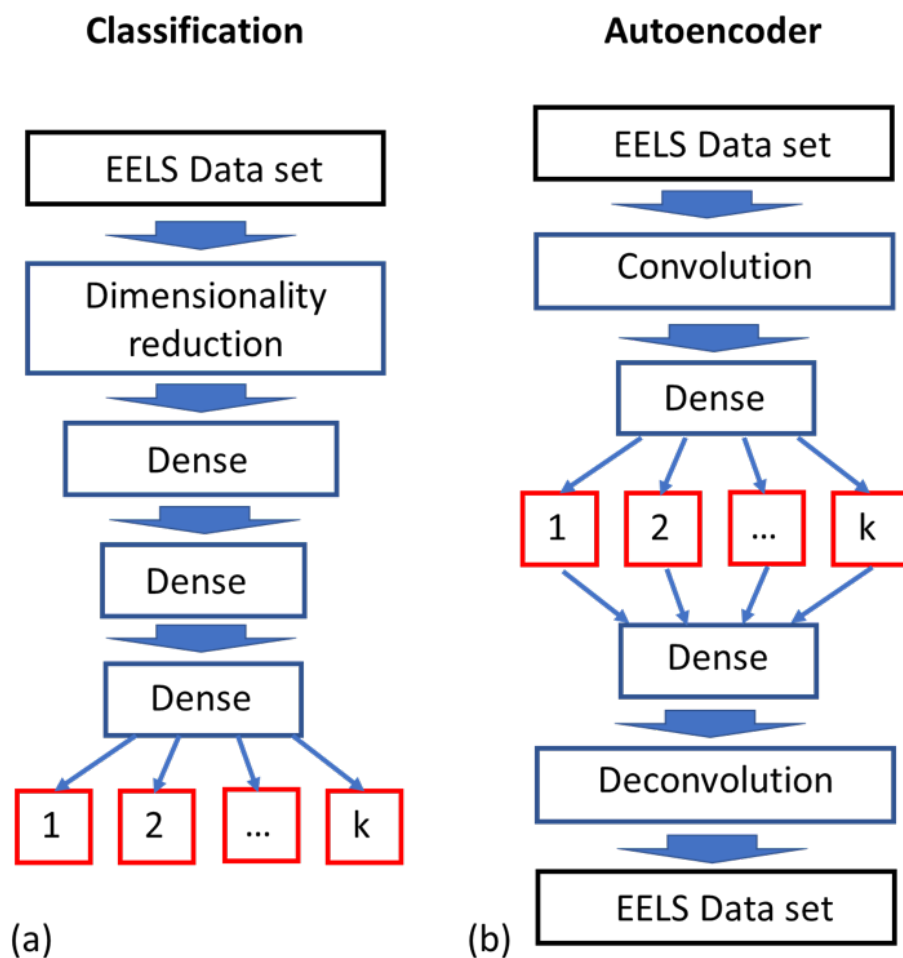


Figure 1. Deep neural networks-based workflows for EELS data analysis. (a) Classification of the spectral data set associates each spectrum with a specific label. Multiple variants of the network are possible, including the combination of dimensionality reduction and dense (i.e., fully connected) neural network, convolutional neural network, etc. (b) Autoencoder network compresses the data set to a small number of latent variables and subsequently expands it into original data set via a set of up-sampling and convolution or transposed convolution operations (deconvolution). Distribution of observed behaviors in latent space provide insight into the types of materials behavior in the unsupervised manner.

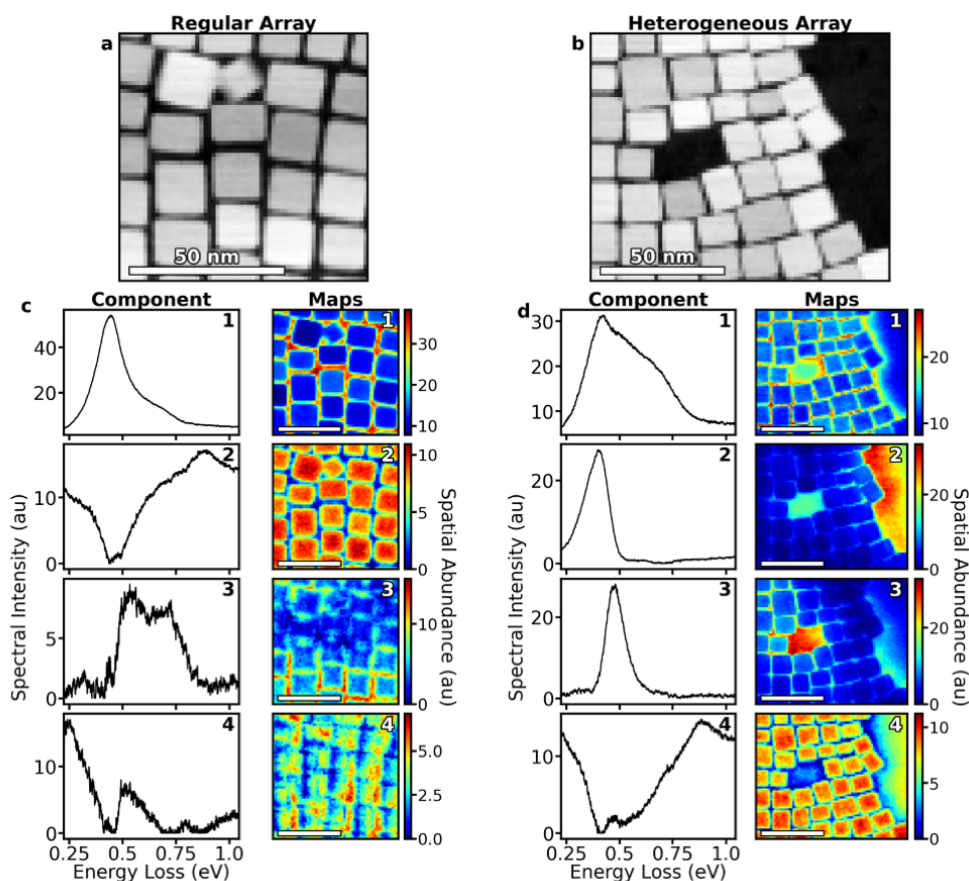


Figure 2. Sample NMF decompositions for FT:IO self-assembled arrays with differing degree of heterogeneity. (a,b) HAADF-STEM reference images show structure of the self-assembled film for both a regular array (RA) and a heterogeneous array (HA), respectively. (c,d) 4-component NMF decompositions for the RA (c) and HA (d) datasets. The NMF decomposition can isolate heterogeneity within the samples but components are not physical and hence cannot be directly used to understand the plasmonic response in the sample. Scale bars = 50 nm.

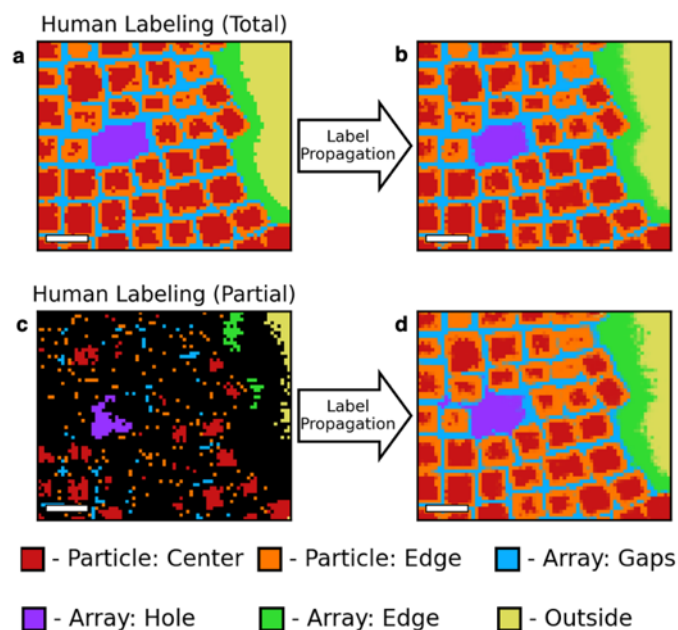


Figure 3. Multilayer perceptron (MLP) label propagation in complex plasmonic nanoarrays.

(a) Human operator generated masks that label every pixel in an EELS hyperspectral dataset. (b) MLP propagation of labels showing little change and validating label choices as physically significant. (c,d) Human operator generated masks where majority of data is unlabeled and MLP propagation of the labels showing the same classification output as the total-human labeling dataset, demonstrating robustness of the approach. Scale bars = 20 nm.

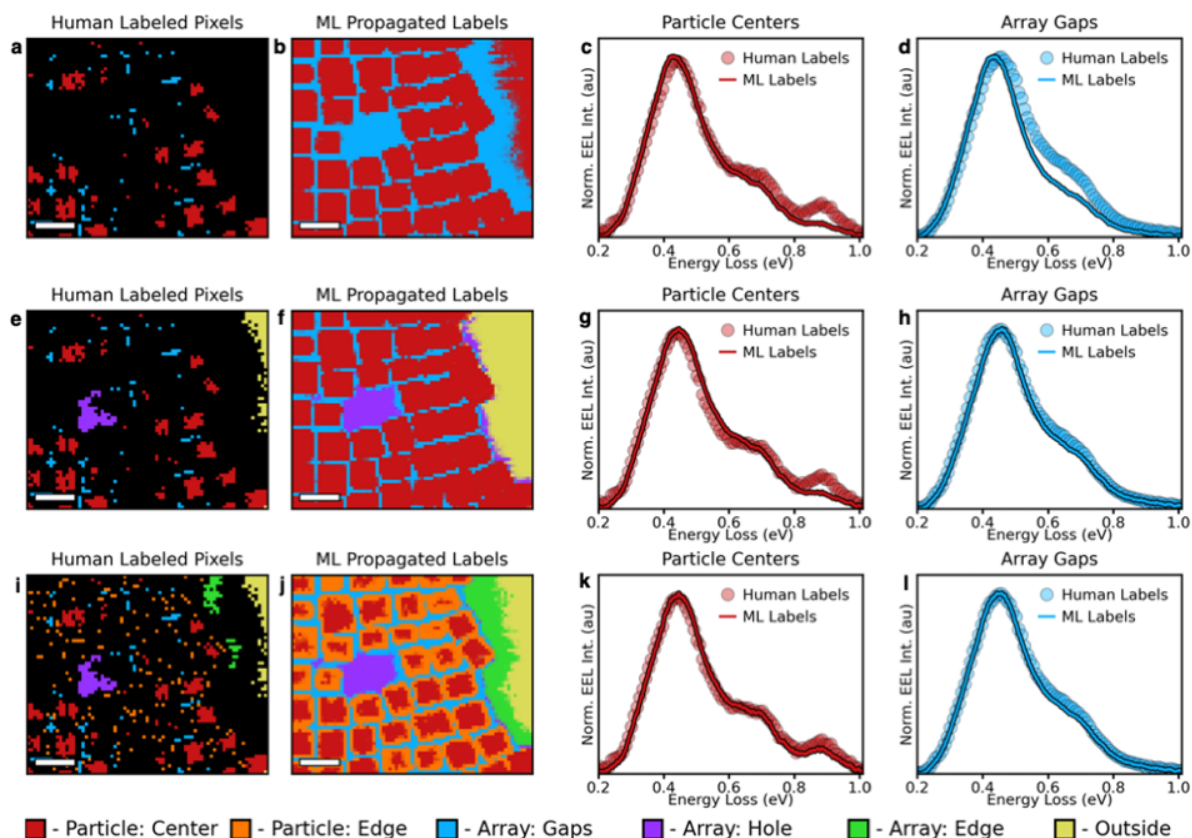


Figure 4. Physical accuracy of classification in bag-of-features MLP labeling. (a) The human-assigned labels for only two of the six classes identified in Figure 3 (particle centers and array gaps) and the MLP propagated labels across the entire dataset showing how the MLP process assigns all pixels to the most valid label. (c,d) Comparison of spectra from human-labeled pixels and MLP labeled pixels for particle centers (c) and array gaps (d) showing quantitative and qualitative differences. (e-h) Same bag-of-features approach as in (a-d) except now with ‘outside’ and ‘array hole’ labels included in the MLP training showing more accurate label propagation and representative spectra. (i-l) MLP process for all six labels showing quantitatively and qualitatively accurate classification according to the spectra. Scale bars = 20 nm.

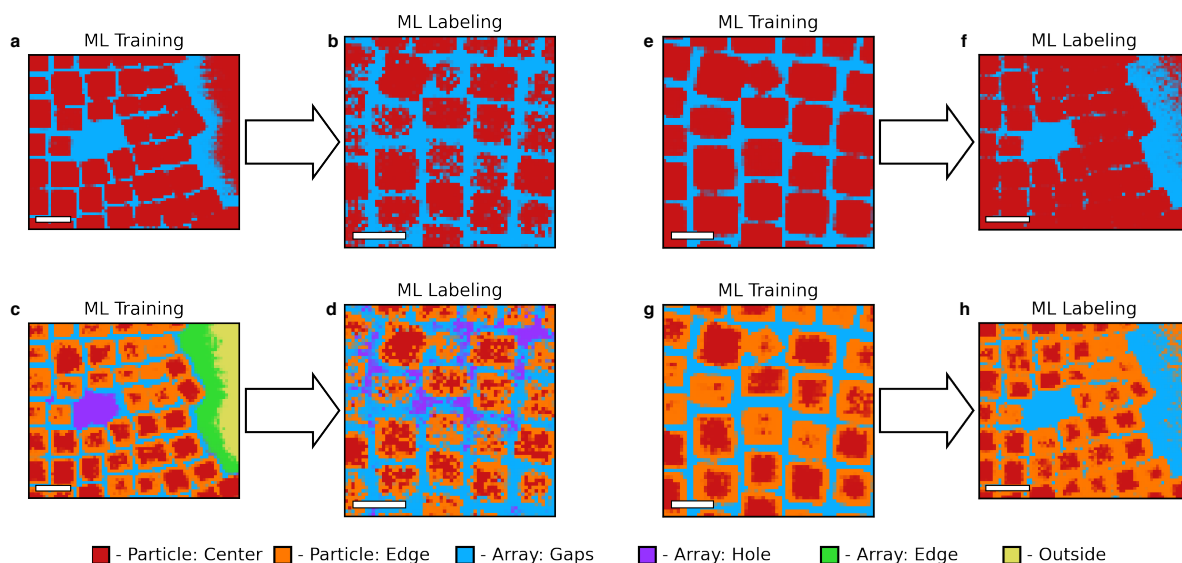


Figure 5. Transferability of training approach between different datasets containing comparable features. (a) MLP network is trained on the HA dataset using only two labels (particles and gaps). (b) Trained network is then applied to classify the RA dataset. (c,d) Network is trained on the full six-labels (as shown in Figure 4) and applied to the RA dataset. (e,f) Transferability of a two-label MLP training on the RA dataset applied to the HA dataset, opposite of HA-RA transferability shown in (a,b). (g,h) Transferability of a three-label MLP training on the RA dataset applied to the HA dataset (particle centers, particle edges, array gaps). Transferability provides meaningful labels in both directions. Scale bars = 20 nm.

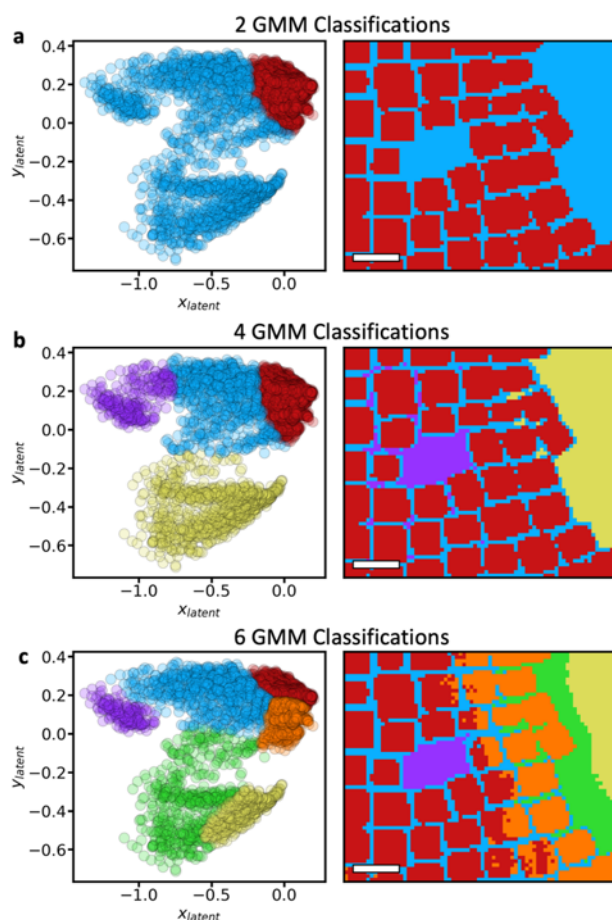


Figure 6. Latent space classification for unsupervised labeling. (a) GMM clustering of the 2D latent space and spatial label assignment in the HA dataset for two GMM classifications, (b) four GMM classifications, and (c) six GMM classifications showing comparable results in labeling to the operator-assisted label propagation. Scale bars = 20 nm.

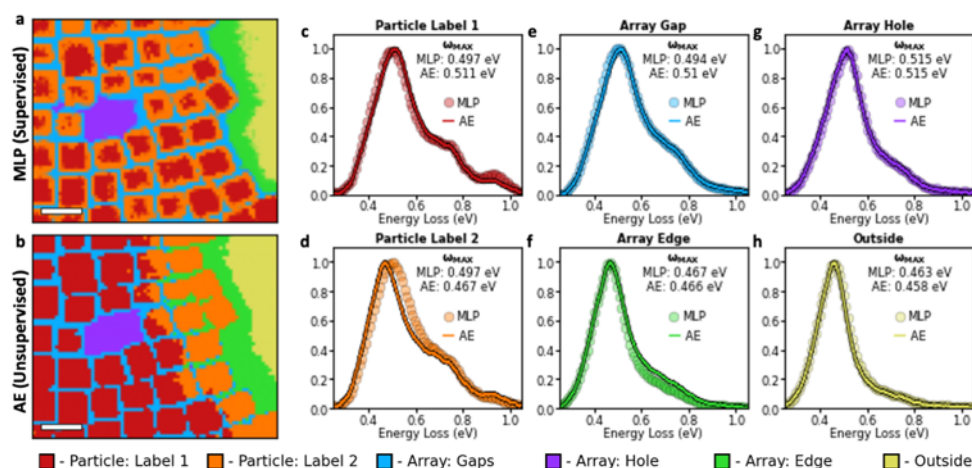


Figure 7. Isolating different physical mechanisms with supervised and unsupervised learning. (a) Six-label operator-assisted MLP label propagation for HA dataset and (b) six-GMM classification AE labeling for HA dataset showing similar labels for all aspects other than the two particle labels. (c-h) Comparison of representative spectra between the MLP and AE labeling: (c) Particle Label 1, (d) Particle Label 2, (e) Array Gap, (f) Array Edge, (g) Array Hole, (h) Outside regions. AE detects redshift at the edge of the array, while operator assisted MLP labels focus on localization of the breathing mode to particle centers. Scale bars = 20 nm.