Combining a Convolutional Neural Network and Watershed Segmentation for Identifying U-Bearing Particles in Secondary Ion Mass Spectrometry Images

Jay G. Tarolli[1], Benjamin E. Naes[2] and David Willingham[3]

[1.] Pacific Northwest National Laboratory, Richland, WA, USA.
[2.] Pacific Northwest National Laboratory, Richland, WA, USA.
[3.] Pacific Northwest National Laboratory, Richland, WA, USA.

Secondary ion mass spectrometry (SIMS) has been employed as a technique for screening U-bearing particles within a complex matrix of other components. Typically performed as a two-step process, an initial pre-screening pass acquires images of up to 2500 fields of view (FOV) across the sample. These FOVs are then analyzed to identify U-bearing particles within the matrix such that a second, more accurate, spot mode analysis can be performed to determine isotopic ratios. Identification of the particles between the two analysis stages, however, is often hindered by considerable background signal and variable intensity across images in a data set.

Until now, approaches for particle identification have incorporated a thresholding mechanism in order segment the images. However, the approaches previously used utilize information from the entire set of images in order to determine appropriate threshold values. What has been observed is that a suitable threshold value is not widely applicable to images in the data set due to variable levels of background signal in each image. The automated particle measurement (APM) software included with CAMECA SIMS instruments utilizes up to 11 user-specified parameters to threshold FOVs. However, those values are set based on a single image and, again, are not widely applicable to all FOVs in a data set. In addition, that many user specified variables introduces room for bias in the results. Watershed segmentation that has utilized an adaptive threshold technique performed slightly better, however was still limited by certain cases, such as an image which contained all background signal and no particles. [1]

In order to threshold individual SIMS particle images effectively, reproducibly, and with minimal user bias, a fully convolutional neural network (FCN) based on that by Long, et. al., was designed and implemented with TensorFlow 0.10 using Python 3.5. [2] The network was designed and trained to classify each pixel in a FOV as belonging to either the foreground or background. In terms of the intended application, foreground pixels are those which are a part of a U-bearing particle and background pixels are the remainder. A training data set was compiled for this network from 28 previous analyses of U-bearing particles in which at least $^{235}U^+$ and $^{238}U^+$ were interrogated. For some data sets, signal for $^{234}U^+$, $^{236}U^+$, and/or $^{238}U^1H^+$ was also available. To generate the images for the training data set, all available U signals for a FOV were summed into a single image and a ground truth image was created by manually highlighting particles using a custom application.

Training of the network was performed using 8,000 of the image pairs, reserving 380 for periodic testing of the network. In total, the network was allowed to train for 600 epochs while monitoring loss. After the final training epoch, the FCN was able to classify pixels belonging to a particle with nearly 99% accuracy. Figure 1 shows a selection of FOVs from the reserved testing subset and the identified images after 600 training epochs. It can be observed from Figure 1 that, by the end of the training procedure, the FCN was able to accurately identify particles of different sizes and shapes. In addition, no particles were identified in a FOV that contained only background signal, such as the pair in Figures 1C and 1F.

The identified images output from the FCN were then used as a binary mask to remove signal in an input image which does not belong to a particle. While one goal of combining a FCN and watershed segmentation for identifying particles is to eliminate user specified parameters, two still need to be specified, the softmax threshold value and watershed level. A simulated data set with known U isotopic distributions was generated and segmented while varying these two parameters. By comparing the resulting isotopic distribution plots to the known distributions, as well as the number of particles identified, appropriate values for each parameter were determined.

A final data set of U-bearing particles, created by mixing 5 NIST certified reference materials, was analyzed with a CAMECA IMS-1280 HR and identified and segmented using the FCN and watershed combination. While the distributions are not as resolved as was observed with the simulated data set, the overall counts in each particle are on the same order of magnitude as the simulated results. This is important as the CAMECA APM software over thresholds and over segments particle images, thus reducing the signal available to calculate isotopic ratios in segmented particles. By increasing the signal available an order of magnitude, these ratios could possibly be determined directly from the output of the FCN and watershed approach without the need for a second reanalysis step.

The work presented here demonstrates the first use of a fully convolutional neural network as a replacement for manual and automatic threshold steps of noisy and variable SIMS images for the purpose of identifying U-bearing particles. Incorporation of a neural network allows the necessary preprocessing step to threshold and remove signal from pixels not belonging to a particle based upon image features and spatial information rather than global calculations, or even worse, parameters which can incorporate bias from a user. Comparison of isotopic distribution plots created by identifying a data set using the FCN followed by watershed segmentation with software currently being used for this type of analysis showed the approach based on machine learning produced tighter, more resolved distributions with overall counts an order of magnitude higher. [3]

References:

[1] D Willingham *et al*, Journal of Vacuum Science & Technology B **34** (2016), p. 03H106.
[2] E Shellhamer, J Jong, T Darrell, ArXiv e-prints (2016).
[3] The research described in this paper was conducted under the Laboratory Directed Research and Development Program at Pacific Northwest National Laboratory, a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy.
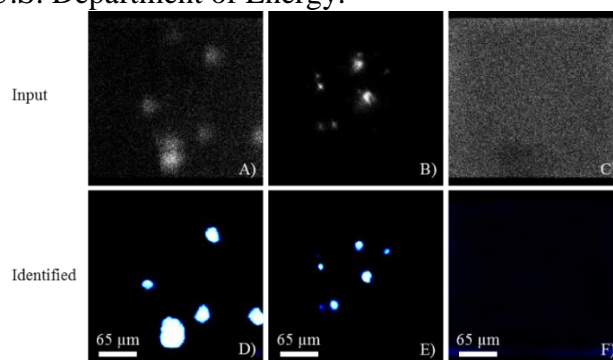
Figure 1. Three FOVs tested during network training. Top row is the input image and the bottom row contains the corresponding identified images after 600 training epochs.