Scanning Electron Microscope Image Segmentation with Foundation AI Vision Model for Nanoparticles in Autonomous Materials Explorations

Timothy B. Gaines
Elect. Engineering & Computer Science
University of Missouri
Columbia, Missouri, USA

Camden Boyle
Mech. & Aerospace Engineering
University of Missouri
Columbia, Missouri, USA

James M. Keller
Elect. Engineering & Computer Science
University of Missouri
Columbia, Missouri, USA

Matthew R. Maschmann
Mech. & Aerospace Engineering
University of Missouri
Columbia, Missouri, USA

Stanton Price
Geotechnical and Structures Lab
Engr. Research & Dev. Center
Vicksburg, Mississippi, USA

Grant J. Scott
Elect. Engineering & Computer Science
University of Missouri
Columbia, Missouri, USA

***Abstract*—The role of scientists and engineers is being significantly reshaped by the increased role of artificial intelligence (AI) across various science domains. Materials science is a field that is well-positioned to benefit from the integration of AI in the design of automated experimental systems. This is especially true for systems that are collecting and analyzing massive quantities of data from scanning electron microscopes (SEM). The associated SEM imagery is on the nanoscale, offering unprecedented fidelity to analyze materials in their most basic and simple structures. However, the richness and potential volume of this data necessitate human-machine teaming, specifically AI tools to accelerate scientific discovery. This paper details how the foundation AI vision Segment Anything Model (SAM) is leveraged within an experimental workflow. Specifically, SAM's basic use requires significant human-in-the-loop effort to mark relevant areas. Herein, we detail how SAM creates excessive segmentation masks and describe the necessary post-processing to refine and filter outputs to be suitable for an autonomous materials experimentation pipeline. Our vision system design that leverages SAM coupled with classical computer vision as a post-processing stage achieves 98–100% precision isolating Aluminum nanoparticles within SEM images.**

I. Introduction

The role of scientists and engineers is being significantly reshaped by the increased role of machine learning (ML) and artificial intelligence (AI) across a variety of science domains. AI/ML is drastically impacting and disrupting the scientific community, ranging from materials science and computational chemistry to biotechnology (e.g., protein structure prediction). New AI/ML tools and approaches can enhance the speed of research, expand the scope of inquiry, and generate ever-greater volumes of data. Autonomously operating research systems that plan, conduct, and analyze experiments already exist and will become increasingly pervasive [1]. This trend increasingly frees the researcher from the mundane tasks within scientific experimentation, allowing them to focus on advanced analytical analysis and bringing creative thinking into the research space. All the while, this human-machine teaming for scientific discovery is generating more, higher-quality data.

Energetic material particles are solid-state grains that release significant energy when undergoing an oxidative reaction. The reaction mechanisms of isolated aluminum nanoparticles (AL NPs) and nanoparticle clusters remain poorly defined. These mechanisms are expected to change as a function of their diameter, temperature, and heating rate. One promising methodology to interrogate the reaction mechanisms of isolated AL NP clusters is by using laser-induced photothermal heating. In this study, 80 nm diameter AL NPs were dispersed on an optical grating substrate to isolate and react small clusters of AL NPs. The particles are then irradiated under the microscope using a focused laser. Identifying and characterizing the pre- and post-reaction particle morphologies using conventional optical microscopy is difficult given the limited resolution of optical imagery.

Scanning electron microscopy (SEM) imaging of AL NPs provides image resolution at the nanometer scale compared to the optical diffraction limit of optical microscopy. SEM analysis may also provide energy-dispersive X-ray spectroscopy (EDS) that can provide a spatial map of chemical composition to better inform experiments. However, SEM analysis is typically more expensive and time-consuming than optical microscopy while also reducing the field of view in each image. Image segmentation may be used to rapidly extract the location and size of all particles in an image.

Experimental setup, planning, and characterization may be greatly accelerated if high-quality SEM images could be rapidly and automatically segmented. For example, if the location and size of each AL NP on a substrate could be determined from an SEM image or image sequence before experimentation, a scientist could plan investigational campaigns and target AL NPs with desired attributes without having to

979-8-3503-5409-6/24/$31.00 ©2024 IEEE
DOI 10.1109/CAI59869.2024.00224

slowly search for appropriate AL NPs using inferior imaging capability of an optical microscope. Similarly, rapid analysis of pre- and post-reaction particle image pairs would benefit experimentalists as they try to determine how reaction conditions influence reaction mechanisms. Ultimately, robust image segmentation and analysis, coupled with computer-controlled experimental control, could ultimately lead to autonomous investigations in which the human experimenter is over the loop, instead of being human-in-the-loop.

In [2], the authors show that optical microscopy *before* and *after* images at 100× can be leveraged for automated characterization of AL NP reactions. However, SEM imagery offers much higher magnification which allows for better texture analysis and can image smaller particle clusters or even single particles. This level of detail provides significantly more visual details and potential for the development of automated AL NP reaction analysis. This contrastive resolution of optical microscopy and SEM images is shown in Fig. 1.

Recently, the Segment Anything Model (SAM) [3] has been released as an open-source foundation AI model with robust image segmentation capabilities. However, SAM's basic use requires significant human-in-the-loop effort to mark relevant areas as "foreground". Alternatively, in an automated mode, "segment everything", a grid of seed points is used to generate a robust, full image segmentation. The full image segmentation results, however, include excessive segmentation masks that must be properly filtered if this model is to be leveraged in an automated system and take human intervention out of the loop. Therefore, we describe the necessary post-processing to refine and filter outputs to be suitable for an autonomous materials experimentation pipeline.

This work, to the best of our knowledge, is the first exploration of using foundation AI computer vision models to perform automated analysis of AL NPs captured with an SEM. In fact, the materials science efforts for imaging and analysis of AL NPs are not well studied in the literature. Our integration of foundation AI into an automated analysis pipeline for materials science research is a novel enabler of future advances.

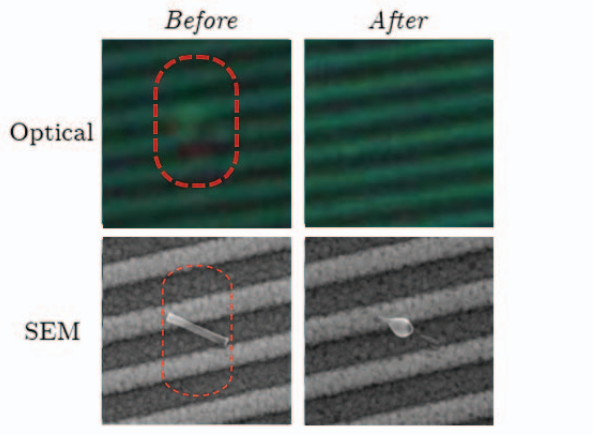


Fig. 1. Single Particle Reaction Image Modality Comparison: Optical - 100x magnification polarization based scattering optical microscopy images, SEM - 100,000x magnification secondary electron microscopy images. It can be observed that significantly more structural detail is visible in the SEM versus the Optical images. Additionally, the true result of the laser irradiation is discernible in the *After* SEM image versus appearing as though the AL NP is gone in the Optical *After* image.

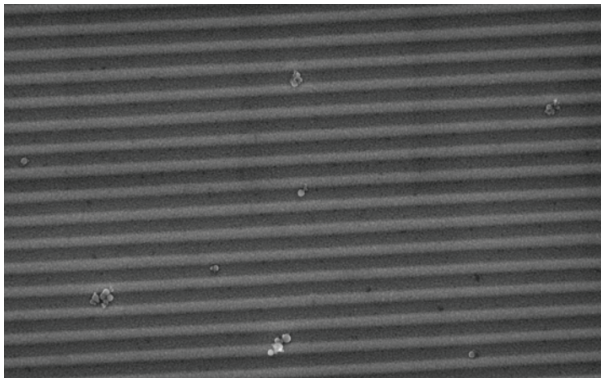


Fig. 2. Example SEM image of AL NPs on an optical grating. We can see a variety of AL NP cluster sizes.

II. SEM Images of Aluminum Nanoparticles

Aluminum particles are dispersed in ethanol to create a stock solution with a concentration of 1 mg/mL and bath sonicated for 1 hour. From this stock solution, 1 mL is diluted with 9mL of ethanol to create a solution with a concentration of 0.1 mg/mL and once again bath sonicated for 1 hour. This solution is further diluted to 0.01 mg/mL before bath sonicating for a final 1 hour. 2 μL of this solution is drop cast onto a substrate and allowed to dry on a hot plate for 30 seconds at 400° C. The substrates are a plasmonic grating or a flat silver substrate capped with 10 nm of alumina. The photothermal heating of aluminum particles follows procedures similar to previous works [4]–[6].

SEM images are acquired using a ThermoFisher Phenom Pharos microscope. Images are acquired at 10 kV acceleration voltage using a secondary electron detector. The image resolution is 1920×1200 pixels. Optical microscope images are acquired using an Olympus BX51W fluorescence microscope and a 32 MP Teledyne Infinity x32 color microscope camera. SEM images are acquired as 16-bit grayscale Tiff images. These images are converted to 3-channel images, with each channel having identical values as the original grayscale image.

A map of the randomly dispersed aluminum particles, and associated details of the particles (diameter, circularity, etc.) are foundational to later reaction experiments in which particles that meet specific physical criteria are selected for laser irradiation. Images acquired from optical microscopy alone are inadequate to measure the morphology of particles or count the number of particles comprising a particle cluster; however, SEM readily resolves the detail of particles with nanoscale

resolution, as demonstrated in Figure 1. The limited field of view of an individual SEM image necessitates the acquisition of many images to traverse a substrate of interest. At a modest magnification, dozens of particles may populate the field of view (Figure 2). By acquiring numerous SEM images dispersed across a substrate, a robust AL NP segmentation pipeline can rapidly segment the particles from the background substrate and provide their spatial coordinates and critical information for later particle testing. Herein lies a critical opportunity to leverage a foundation AI that can accurately segment SEM images within this complex visual modality.

III. Segment Anything

Segment Anything [3] is a transformer-based segmentation model created by Meta AI. Their publicly available pre-trained models were trained on 11 million high-resolution images, in which there are 1.1 billion training segmentation masks. The large volume of highly diverse training samples is why they claim the model has the capability to segment anything – that is, it is a foundation AI model for image segmentation. If this model can segment our SEM imagery, we could then use the masks to extract cluster centers. Due to space constraints and the primary focus of this paper, interested readers are encouraged to review [3] for architectural and related details of Segment Anything.

Segment Anything performs its segmentation in two steps. It first runs the input image through an image encoder to create its embedding. This embedding is then run through a decoder with additional information, an encoded prompt, to create its masks. It should be noted that these masks do not provide any semantic meaning from the model, they are simply layers of a segmentation.

Segment Anything requires prompts through one of two different methods, points or bounding boxes. A user can add points on an image denoting foreground and/or background, and then Segment Anything will segment the image with this added context. Additionally, users can supply bounding boxes signifying foreground areas. Bounding boxes and points may be used in conjunction. The result of the model processing is a collection of segmentation masks. However, in the case of our SEM images of AL NP as detailed above, the need for human-in-the-loop seeding of segmentation is viable for our target of autonomous AL NP identification. Instead of providing the information about where the clusters exist as a prompt, we seek to automatically extract this with the aid of the model.

Segment Anything does have a prompting formula for autonomous use, a grid of foreground points. By laying a grid of foreground points equally over the image, a majority of the image is sampled and one might assume you hit all targets. However, an issue that arises is that the number of sample points to lay out into the grid becomes a new human-controlled parameter. The grid of these seeded foreground points may miss some of the key image contents, particularly the small ones, that are desired from the segmentation. Specifically, if you look at Fig. 2, you can see a variety of AL NP cluster sizes, and to ensure we get small particles like the one about midway down on the left requires a very high number of segmentation seed points in the grid. Otherwise, particles that do not contain a gridded seed point are not marked as a segmentation target and will not be extracted as a segmented cluster.

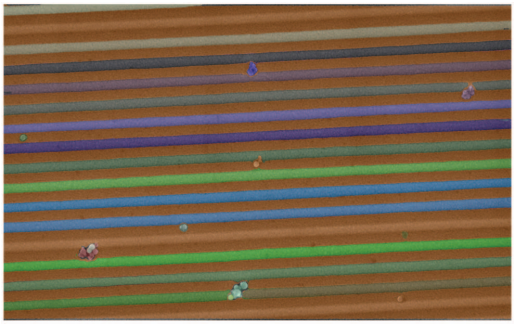


Fig. 3. Segment Anything Prediction of Figure 2

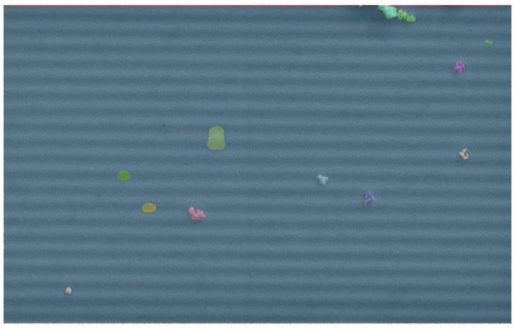


Fig. 4. Segment Anything Prediction with Grating Imperfections

Figure 3 shows how applying a grid of points over an image yields many erroneous mask regions relative to our use case. In the rendering, each mask is a semi-transparent color that is overlayed on the original image. Specifically, we desire to extract each AL NP cluster as a mask layer, and the background substrate as a separate mask layer. We can see that the substrate becomes multiple masks (rainbow striping). Additionally, we see that some AL NP clusters are over-segmented (lower-left quadrant of Fig. 3). These various erroneous mask regions necessitate post-processing combined with appropriate grid seeding in order to incorporate the Segment Anything model into an autonomous experimentation workflow. Fig. 5 shows the processing pipeline, and the subsequent techniques to filter and select only relevant AL NP cluster regions. We can see that the Segment Anything model is a first step that expands the SEM image into a deep band of mask layers, and all the subsequent steps in the pipeline are detailed in Sects. IV and V.

IV. Mask Filtering

For mask predictions from Segment Anything to be useful, the false positive (FP) regions need to be removed. The example in Fig. 3 shows background substrate mask regions that

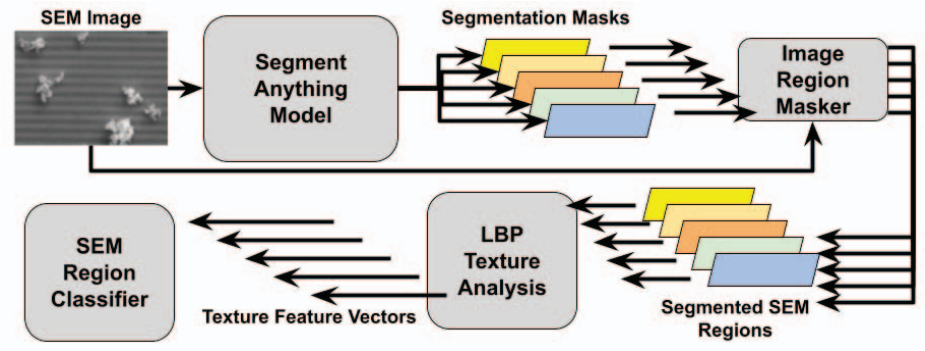


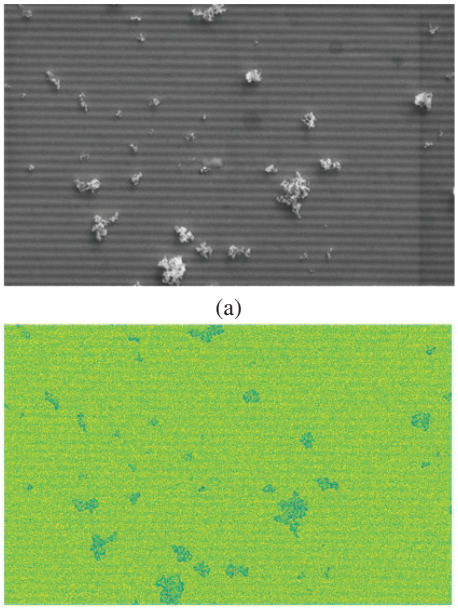
Fig. 5. Autonomous Region Selection Pipeline. SAM produces layers of segmentation masks, which are recombined with the input (*Image Region Masker*) to generate isolated SEM image regions as separate images. These are then fed into *LBP Texture Analysis* and then classified (*SEM Region Classifier*).

are long and easily filterable through eccentricity thresholding. Alternatively, the regions shown in Fig. 4 cannot be filtered in this way. The FP regions shown there are much more circular, versus elongated, and therefore another method for mask filtering must be determined. Because this process is being developed for autonomous laser targeting, we are much more concerned with recognizing the erroneous mask regions, i.e., false positives, rather than any clusters Segment Anything has missed, i.e., false negatives (FN). The laser irradiation process and stage movement are time-consuming and therefore missing one cluster among so many is much less important than avoiding irradiating the background and adding a 'fake' image to our experimental dataset.

A. Texture Analysis for Mask Selection

An advantage of SEM images over optical images is that the minute details in the clusters are visible, that is one can easily see the texture differences between the background and the clusters. Therefore, we propose to use classical texture descriptors to identify the relevant AL NP clusters from the collection of segmentation masks. For our texture descriptor, we chose the local binary pattern [7] (LBP) with rotational invariance. The LBP texture descriptor analyzes each pixel within a mask region and its surrounding pixels to see if they are of greater or less intensity than itself. Each pixel can then be assigned to a pattern depending on which of its neighbors is brighter or not. A histogram of the number of times each pattern was encountered can then be made to describe the texture of the area. We utilized a radius of three pixels at eight equidistant angles. The segmentation masks were eroded to remove three pixels of depth all around the perimeter of each individual region. This ensures that entries into the LBP histogram can only look within the region area.

To begin, we calculate the LBP over an image with grating blemishes. This image is shown in Fig. 6 where the pixel intensity represents which pattern it corresponds to (see [7] for details of the 26 patterns extracted from the 8-neighbor LBP). In the LBP image, the cluster regions are still highly visible, whereas grating blemishes blend in the background



(a)

(b)

Fig. 6. (a) SEM image with large, complex AL NP clusters and (b) the associated LBP image. As noted in the pipeline discussion, each mask is eroded by the radius of the LBP filter, and the resulting masked SEM image regions are processed independently.

completely. This is a great signal that the LBP is highly distinctive when used to compare the two region types.

B. Mask Selection Pipeline

To leverage this texture response for filtering segmentation masks, we create a pipeline to process SEM images through Segment Anything, then process the masked regions with LBP texture analysis, feeding masked regions of SEM into a binary classifier. Figure 5 illustrates the overall flow of SEM images through Segment Anything, LBP texture analysis, and AL NP cluster selection. A key component is the creation of a

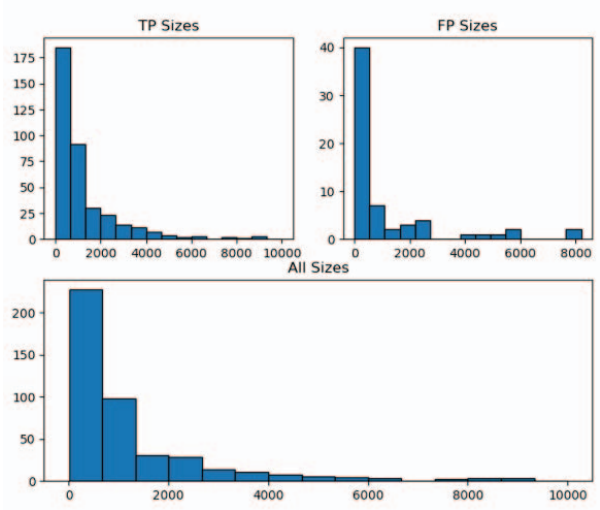


Fig. 7. Size Histogram for All Mask Regions

classifier for the LBP features generated from the masked SEM segmentations.

We create a nearest-prototype classifier by collecting histograms from seven cluster regions, true positives (TPs), and seven regions that should be filtered out, FPs, to use as exemplars. The regions chosen were randomly selected with size brackets. For both the TPs and FPs four regions were selected from pixels sizes $60 \geq r < 1000$, two regions from $1000 \geq r < 4000$, and the last region from $5500 \geq r < 6500$. These size brackets are chosen to reflect the size distribution of our dataset, the distribution is shown in Fig. 7. The histograms of all the exemplars can be seen in Figure 8, the last bin has been removed for visual clarity. The TP and FP histograms are highly characterized by their differences in the bins at edges and centers. The edge bins of an LBP histogram represent the flat unchanging patterns whereas the center bins represent edge and corner patterns. This is highly consistent with what we see in the SEM imagery as real cluster regions have many edges and corners due to the individual particles and the FP regions are mostly uniform in intensity with the exception of the grating lines which have little area per region.

V. Experimental Results

Each segmented region is compared to the exemplars using Euclidean distance. However, there were a few options in how this was done. The first choice is how much of the two histograms to compare. We could simply compare the entire histogram or we could compare only the regions where they typically differ. Thus, we choose three different subsets of the histogram: the entire histogram, the edges (bins 1-5 and 21-25), or the center (bins 6-20). We also had a choice of which exemplars to compare the regions to. We could compare each region to all exemplars or compare them based on size: small regions compared to the small exemplars and small-medium regions(> 750 & < 1500) compared to both small and medium exemplars, medium regions to medium exemplars, and lastly medium-large and large regions (> 3500) to both the medium exemplars and the single large exemplars. After the exemplars for comparison are chosen, the Euclidean distance between the target region and each FP exemplar is computed and recorded. Similarly, the Euclidean distances between the target region and each TP exemplar are calculated separately. The target region is labeled by the label of the minimum distance prototype, i.e., a simple nearest-neighbor classifier.

TABLE I
Texture Based Filtering Results

Exemplar	Bins	Accuracy	Precision	Recall	F1
All	All	0.982	0.982	0.997	0.990
	Center	0.968	0.992	0.971	0.982
	Edges	0.984	0.982	1.0	0.991
Size	All	0.984	0.985	0.997	0.991
	Center	0.961	1.0	0.956	0.977
	Edges	0.984	0.982	1.0	0.991

Table I shows the results from the experimentation over the different exemplar selection techniques and bin subsets. Again, we are highly concerned with minimizing the FPs to ensure we do not add false clusters to our downstream energetics experimentation dataset. Because of this, we happily accept lower performance metrics in overall accuracy, recall, and F1 to search for the model with perfect precision. Interestingly, utilizing only the center bins when taking the Euclidean distance gives the best precision in both exemplar selection methods. While the best model does use size to select exemplars the small difference in performance and our small dataset means we cannot say for certain if this is the overall superior method. However, with this texture-based algorithm to filter out false Segment Anything mask regions, we can confidentially image and target clusters for laser irradiation autonomously.

VI. Conclusion

The reaction of isolated aluminum nanoparticles for energetic materials research requires the acquisition of many SEM images to establish the location and diameter of particles, and the quantity of particles residing in clusters. Because of the dimensions of nanoscale and microscale particles, this information may not be accurately obtained using optical microscopy. Herein, the Segment Anything foundation AI model was used to segment aluminum particles in SEM images and record mask information to elucidate the stochastic distribution of particles on complex substrates.

The proposed workflow demonstrates how foundation AI models can incorporated into energetics research to accelerate research processes and enhance the confidence of results. The application of texture analysis is critical for the identification of substrate defects captured by Segment Anything to mitigate false positive particle identification. This pipeline will be used in the experimental flow of energetics material research to accelerate the data acquisition process and to increase

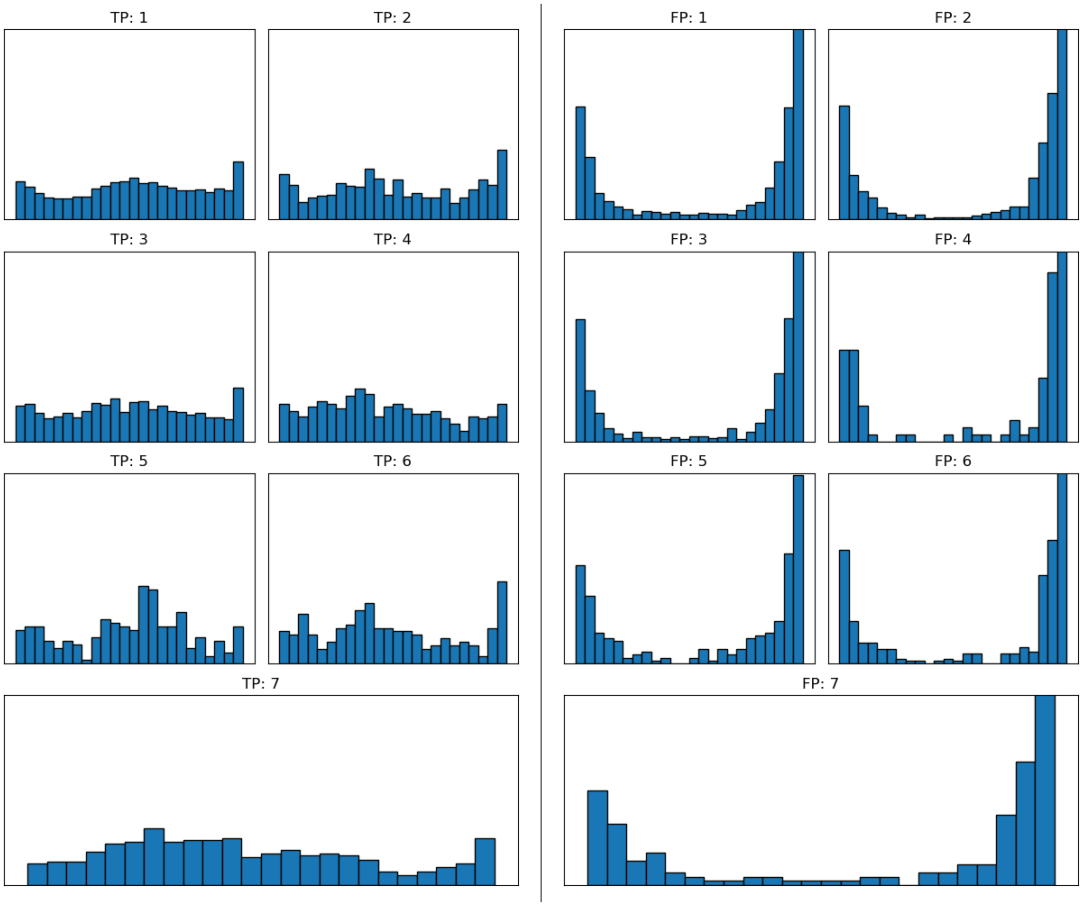


Fig. 8. Histogram of all exemplars for the nearest-neighbor classifier: (left) true positive AL NP clusters (TP) and (right) false positive (FP) segments such as blemishes on the surface grating.

confidence in the attributes of the aluminum particles prior to their testing.

ACKNOWLEDGMENTS

The experiments described and the resulting data presented herein, unless otherwise noted, were funded under M&S for MUM-T, under Contract W81EWF83375468, managed by the US Army Engineer Research and Development Center. The work described in this document was conducted at the University of Missouri, Columbia, Missouri, USA. Permission was granted by the Director of the Geotechnical and Structures Laboratory to publish this information. Computing resources for this research have been supported by the NSF National Research Platform and NSF OAC Award #2322218 (GP-ENGINE).

REFERENCES

[1] P. Nikolaev, D. Hooper, F. Webber, R. Rao, K. Decker, M. Krein, J. Poleski, R. Barto, and B. Maruyama, "Autonomy in materials research: a case study in carbon nanotube growth," *npj Computational Materials*, vol. 2, no. 1, pp. 1–6, 2016.

[2] T. Gaines, C. Boyle, J. A. Hurt, M. R. Maschmann, J. M. Keller, G. J. Scott, and S. R. Price, "Towards automatic characterization of nano-energetic material response to directed energy," in *Energy Harvesting and Storage: Materials, Devices, and Applications XIII*, vol. 12513. SPIE, 2023, pp. 14–22.

[3] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, and R. Girshick, "Segment anything," *arXiv:2304.02643*, 2023.

[4] N. Zakiyyan, C. Boyle, C. J. Mathai, K. Gangopadhyay, J. McFarland, S. Gangopadhyay, and M. R. Maschmann, "Observation of spallation in an isolated aluminum nanoparticles and 2d molybdenum oxide system by rapid plasmonic heating," in *2023 IEEE 23rd International Conference on Nanotechnology (NANO)*, 2023, pp. 243–248.

[5] N. Zakiyyan, C. Mathai, J. McFarland, S. Gangopadhyay, and M. R. Maschmann, "Spallation of isolated aluminum nanoparticles by rapid photothermal heating," *ACS Applied Materials & Interfaces*, vol. 14, no. 49, pp. 55 277–55 284, 2022, pMID: 36445833. [Online]. Available: https://doi.org/10.1021/acsami.2c18678

[6] N. Zakiyyan, C. M. Darr, B. Chen, C. Mathai, K. Gangopadhyay, J. McFarland, S. Gangopadhyay, and M. R. Maschmann, "Surface plasmon enhanced fluorescence temperature mapping of aluminum nanoparticle heated by laser," *Sensors*, vol. 21, no. 5, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/5/1585

[7] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.