Pushing the boundary: Particle edge determination in human segmentation of uranium ore concentrates in scanning electron microscopy images

International Nuclear Materials Management (INMM) and European Safeguards Research and Development Association (ESARDA) 2023 Annual Meeting

Pushing the boundary: Using texture features to understand particle edge determination in human segmentation of uranium ore concentrates in scanning electron microscopy images

Authors: Cole J. Thompson[1], Reid Porter[2], Alexa Hanson[1], and Kari Sentz[1]
Affiliations:
1. Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, NM, 87545
2. Pooka Play, LLC., Los Alamos, NM, 87544

ABSTRACT
Nuclear forensics rely on a mosaic of approaches to support the determination of nuclear material origin and production pathway. Morphology, or the analysis of size, shape, structure and distribution of features, is an emerging approach with the potential of providing signatures of nuclear process history. An expert can perform a morphological analysis using a human-assisted machine learning image analytic toolkit developed at Los Alamos National Laboratory to segment (outline) material particles from scanning electron microscope (SEM) imagery and obtain quantitative statistics. Human-mediated segmentation requires many decisions—like deciding which particles to segment and determining the boundary of a segmented particle—making it a time-consuming effort. This work focuses on assessing ambiguity in the human segmentation of particles from SEM images in nuclear forensics applications. We investigate this through exploring the gray-scale values of pixels surrounding a human segmentation and comparing with variations on fuzzy c-means image segmentation to explore the neighborhood of boundaries at the perimeter of particles.

INTRODUCTION
There are a wide range of analytic tools that nuclear forensic scientists can use to gain insights from a sample of interest, like the sample production pathway [1]. One emergent method for nuclear materials is *morphology,* or the analysis of the analysis of size, shape, structure, and distribution of features of particles, occlusions, pores, agglomerates, and other features in microscopy images, typically with scanning electron microscopy (SEM). Quantitative image analysis for morphology has typically been done through human-mediated segmentation with expert judgment determining distinctive features such as particle type or particle boundary. This process confers a high cost in terms of analyst time. A typical human segmentation can be seen in Figure 1 where the user segmented a complex materials image by identifying the features of interest, delineating them, and labeling them with color labels.

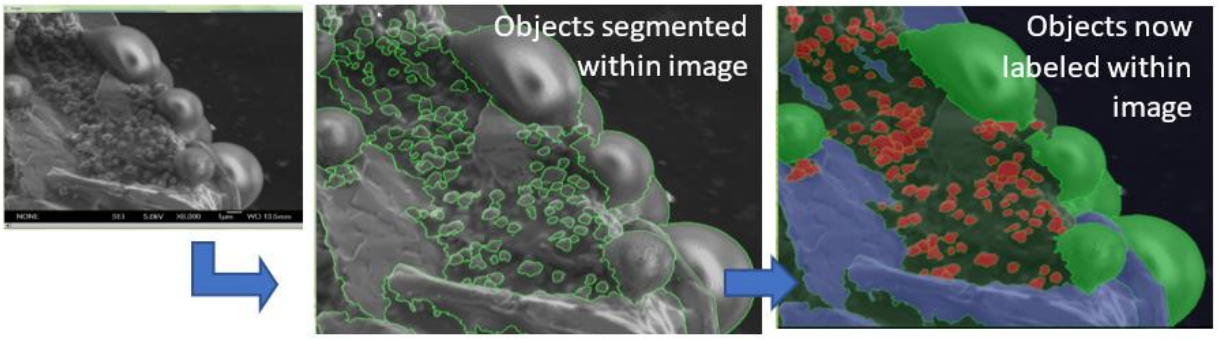

Figure 1. Illustration of the typical segmentation pipeline.

With this segmented image with colored labels, quantitative analysis of each particle type, denoted by color in the image, can be performed. Typical measurement attributes include area, circularity, and ellipse aspect ratio.

To perform morphologic segmentation, Los Alamos National Laboratory has developed a human-assisted machine learning image analytic toolkit which combines the ability to segment (outline) material particles from scanning electron microscope (SEM) imagery and obtain quantitative statistics into a single software called Morphological Analysis of Materials (MAMA) [2, 3]. MAMA has over one hundred registered users at twenty-four US government organizations, academic institutions, and international partners. In nuclear forensic research, MAMA has been used to show that morphological analysis using segmentation and quantitative analysis of uranium ore concentrates has been effective for identifying the process history and conditions of uranium ore concentrates across different synthetic pathways mimicking commercial production pathways [4, 5, 6, 7].These results underscore both the usefulness for morphology signatures and motivate the need for advancement. One outstanding question is what is the uncertainty associated with the segmentation.

In forensics applications, uncertainty is an important measure of confidence in the downstream statistical quantifications. Previous approaches invoke such assumptions of a Gaussian distribution that ignores the features of the image itself that can lend to the ambiguity of segmentation. Such features include pixel intensity, contrast and their relationship to edge location, edge sharpness, and surface detail at different scales, particularly locally, i.e. surrounding segmented particles, and globally, i.e. taking the entire image into account [8]. What is needed is an uncertainty quantification measure is based on the characteristics of the image itself rather than distributional assumptions that are not based on image characteristics.

There are at least two main types of uncertainties that are present in image segmentation: aleatory uncertainty that is the uncertainty associated with randomness and epistemic uncertainty that is the uncertainty associated with a lack of knowledge [9]. Epistemic uncertainty in imagery can be found with low contrast particle edges that can be hard to segment. In the case of the algorithmic segmentation, we as an end user can be epistemically uncertain as to why an algorithm selected a particular segmentation; an uncertainty characterization based on the image characteristics is a technological gap for both human and algorithmic approaches to image segmentation. Figure 2 illustrates poorly defined boundaries in a sample. Focusing particularly on the center of the image, determination of discrete particles is difficult. We can consider this as the type of epistemic uncertainty where we are uncertain as to what class to assign a pixel, i.e. is this pixel a boundary or not?

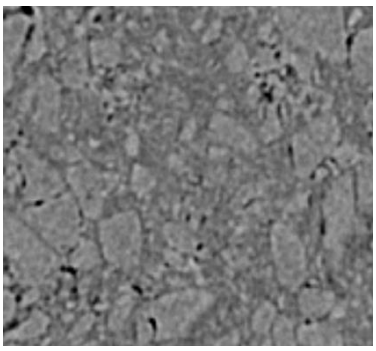

Figure 2. An example SEM image with poorly defined boundaries.

To help answer the question of whether a pixel is *part of a boundary* or *not part of a boundary*, this work explores ambiguity in segmentation by quantifying segmentation uncertainty in an unsupervised manner using the enhanced fuzzy c-means algorithm [10]. This algorithm uses fuzzy logic where pixels can belong to multiple classes simultaneously. This is expressed by a membership function that expresses the degree to which a pixel belongs to different classes. More specifically, we explore modifications to the enhanced fuzzy C-means (EnFCM) algorithm to identify different classes in an unsupervised way and generate values of group membership as a function of gray level value. This grounds the basis of the membership function based on measurable quantities from the image itself.

METHODS

SEM Image Data

We used scanning electron microscope (SEM) imagery from a previous study that examined the effect of diel cycling of temperature, relative humidity, and synthetic production route on the surface morphology of alpha-U308 [5]. This data consisted of 78 images from different samples and replicates of alpha-U308 imaged at many different magnifications. The size of each image is (943, 1024) pixels.

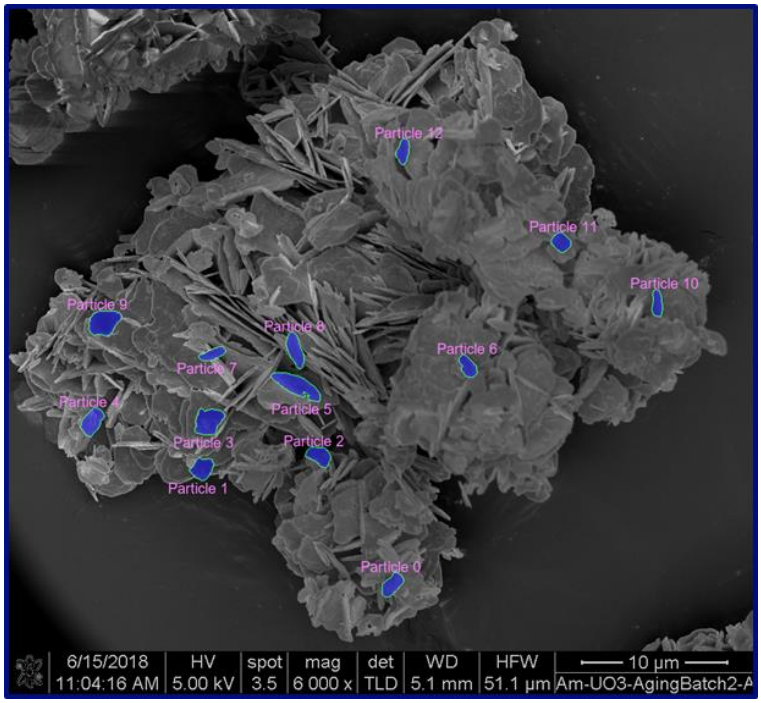

Figure 3. Example SEM image and associated segmented particles.

Enhanced Fuzzy C-Means Algorithm

This work applies the enhanced fuzzy C-means (EnFCM) algorithm developed by Szilagyi et al. [9]. This algorithm builds on the work of the standard fuzzy C-means algorithms (FCM) which groups values x_k, *k*=1, *n,* into *c* clusters by minimizing the objective function given by Eq. (1) [12]:

$$J_B = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^{p} \, (x_k - v_i)^2 \qquad (1)$$

Where v_i is the prototype value of the *i*th cluster, u_{ik}^p is the fuzzy membership of the *k*th pixel with respect to cluster *i*, and *p* is a weighting exponent. This was later modified by Ahmed et al. to allow for local pixel neighborhoods to influence the classification of neighboring pixels [13]. This modified the objective function to be Eq. (2):

$$J_A = \sum_{i=1}^{c}\sum_{k=1}^{N}\left[u_{ik}^p(x_k - v_i)^2 + \frac{\alpha}{N_k}\sum_{r=1}^{N_k}u_{ik}^p\left(x_{k,r} - v_i\right)^2\right] \quad (2)$$

where $x_{k,r}$ represents the neighbor pixels of x_k and N_k is the number of pixels in the neighborhood of the k^{th} pixel. The α value controls the intensity of the neighborhood influence.

The EnFCM algorithm takes these algorithms and performs further modifications, by first normalizing pixel intensity in the neighborhood of interest according to Eq. (3) to use the objective function given by Eq. (4).

$$\xi_k = \frac{1}{1 + \alpha}\left(x_k + \frac{\alpha}{N_k}\sum_{r=1}^{k}x_{k,r}\right) \quad (3)$$

$$J_S = \sum_{i=1}^{c}\sum_{l=1}^{q}\gamma_l u_{il}^p\,(\xi_l - v_i)^2 \quad (4)$$

where γ_l denotes the number of pixels from the image having an intensity equal to l, where l=1,..,q, where q is the number of gray levels in the image.

To minimize this objective function, the algorithm first calculates the fraction of pixels at each gray level and initializes cluster prototypes. Then the membership function of the pixels to each class is updated to Eq. (5) and followed by an update to the new values of the cluster prototypes according to Eq. (6). This is repeated until the convergence threshold or maximum number of iterations is reached.

$$u_{il} = \sum_{j=1}^{c}\left(\frac{\xi_l - v_i}{\xi_l - v_j}\right)^{\frac{2}{p-1}} \quad (5)$$

$$v_i = \left(\sum_{l=1}^{q}\gamma_l u_{il}^p\xi_l\right)\left(\sum_{l=1}^{q}\gamma_l u_{il}^p\right)^{-1} \quad (6)$$

Image Processing

Using the human segmentations from the 78 SEM images, image patches of each of the 2083 segmented particles were extracted. A binary mask version of each segmented particle image patch was also created. For each of the 2083 segmented particle image patches, the EnFCM algorithm was run using a gray-level scale of 255, 4 clusters, fuzziness degree (*m*) of 2, convergence threshold of 0.05, a window size of (3,3), and a neighborhood effect of three.

From the fitting of the EnFCM to each image, the value for each cluster for each gray level are used to represent the membership function for the fuzzy approach to image segmentation. An example of the resulting membership function is shown in Figure 4. Note that in the intersection

of classes, we observe that each of the possible classes contributes a non-zero membership to the ambiguity in class selection (blue, yellow, green, and red distributions). This contrasts with the typical segmentation approach where pixel classes (or clusters in this application) are discrete values and assigned based on the class with the highest probability, without consideration of the probability distribution. These intersections confirm the intuition that the clustering of pixels into classes has uncertainty, which is not acknowledged with typical segmentation approaches.

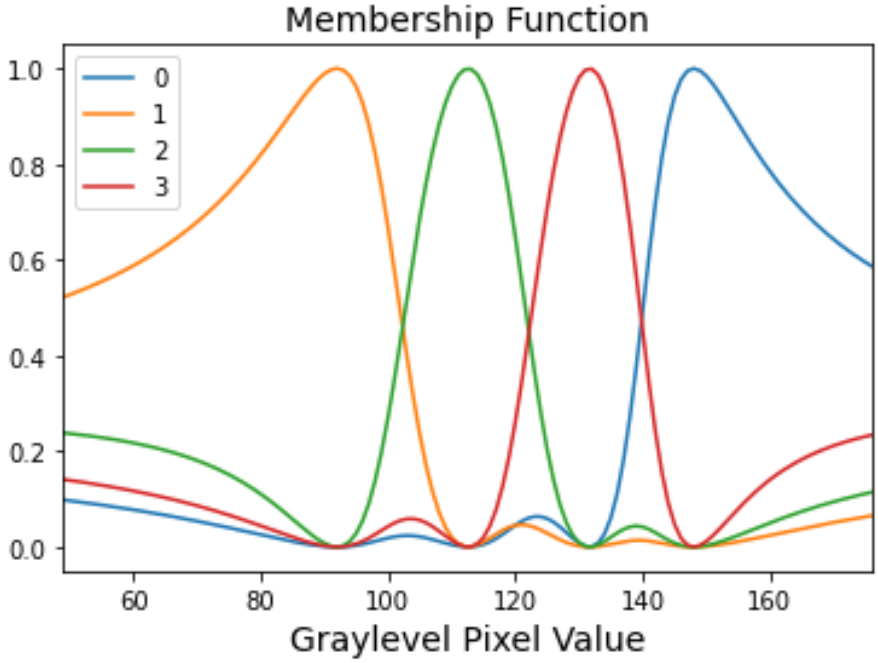

Figure 4. The fuzzy c-means membership function.

The original gray-level image can be converted from gray level pixel space to class space (and vice versa). This is shown in Figure 5. This use of thresholds is found in conventional image segmentation, but unlike normal segmentation, this approach uses the continuum of classes allowing for a more granular thresholding approach.

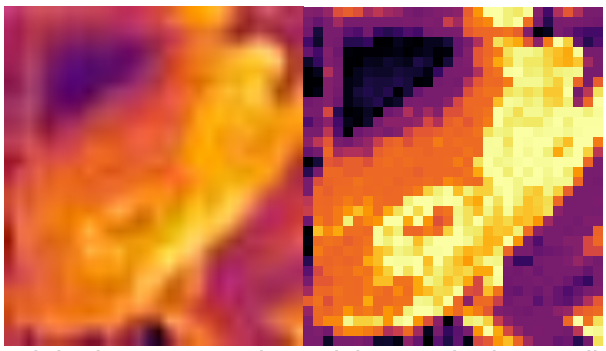

Figure 5. Left: The original segmented particle patch (uses discrete gray levels, 0-255). Right: The Fuzzy C-Means image (uses weighted average class value, continuous between 0 and 3)

The newly created image, called the Fuzzy C-Means image (right in Figure 5), can now be manipulated. Three specific manipulations are performed and an additional image is generated from each, shown in Figure 6:

1) **Outer image**: Threshold the image at the user-specified membership function value to capture the human segmentation and outer (left image in Fig. 6)
2) **Intermediate image**: Human segmentation image with clusters labeled by colors (middle image in Fig. 6)
3) **Inner image**: Threshold the Human segmentation (Intermediate) image (generated above in 2) at the same user specified value from 1 (right image in Fig. 6)

These manipulations create an outer and inner bound on the human segmentation. After the manipulations above, the contour of the largest area in the image is found using the OpenCV library and extracted for quantitative analysis. The largest area will correspond to the particle area of interest given the initial tight crop of the segmented particle image patches. There are many quantitative features that could be calculated for each particle. This analysis focused on particle circularity.

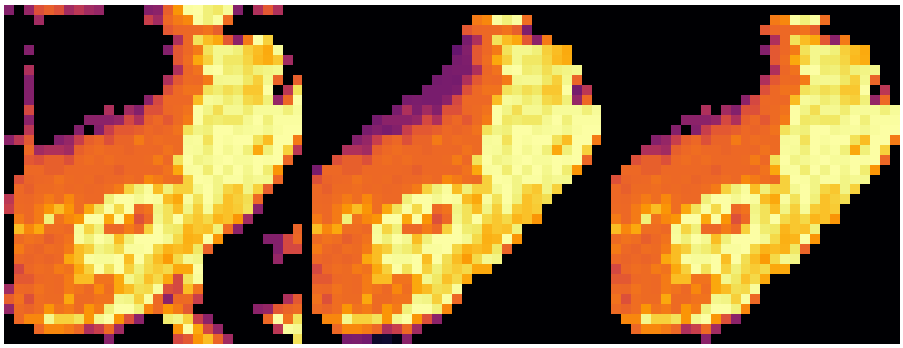

Figure 6. *Left:* **Outer image**: *Fuzzy C-Means image thresholded at input parameter (1.05).*
Middle: **Intermediate image**, *the human segmentation.*
Right: **Inner image**, *threshold of intermediate image at user-input value (1.05)*

RESULTS

Using the quantitative metric of particle circularity to allow for comparisons between the outer segmentation boundary (left image in Figure 6 panel), the intermediate segmentation boundary image from human segmentation (middle image in Figure 6 panels), and the inner segmentation boundary (right image in Figure 6 panels).

Circularity

Circularity is a measure that relates the ratio of the particle area to the particle perimeter. Values closer to 1 indicate a more circular particle. The distributions shown in Figure 7 show that human segmentation is skewed towards more circular than both the inner and outer segmentations. This figure demonstrates that both fuzzy segmentations are vastly different from the human segmentation.

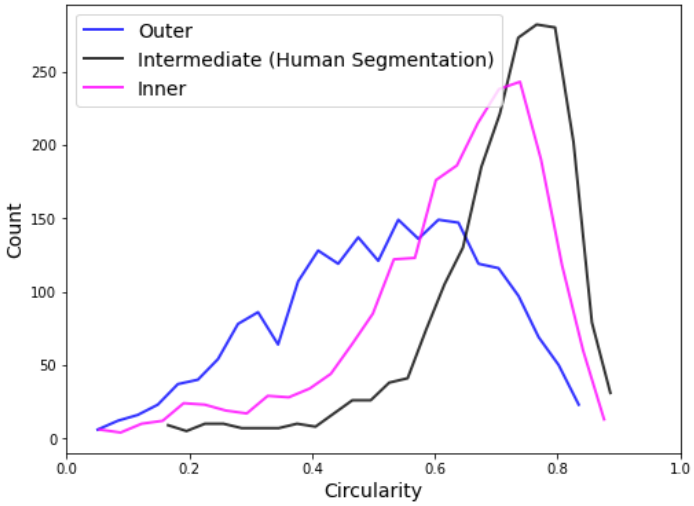

Figure 7. Histogram of circularity for each image type.

This finding is confirmed with the Welch's t-test (also known as the unequal variances test) was performed for two comparisons for each metric using all 2083 data points for circularity:

1) The outer boundary and the intermediate boundary
2) The inner boundary and the intermediate boundary

Table 1 reports the test-statistic and corresponding p-value for each comparison of interest for each metric.

Table 1. Welch's t-test results for comparison of means of distributions for different segmentations.

Metric	Comparison	t-statistic	p-value	Distributions have the same mean?
Circularity	Outer vs. Intermediate	42.23	2.55E-319	No
Circularity	Inner vs. Intermediate	18.79	1.67E-75	No

The test for equivalence of the mean for circularity comparing the Outer Image to the Intermediate as well as the Inner Image to the Intermediate supports the uncertainty in the characterization of circularity. The human segmentation is more circular then either of the fuzzy c-means segmentations. The difference between the Outer, the Inner, and the Intermediate offers a new insight into the uncertainty in the clustering of gray-level pixels into classes.

DISCUSSION

The Welch's t-test comparisons for the three segmentations yield meaningful results with respect to the characterization of uncertainty in gray-scale imagery. The results for circularity show that the intermediate human segmentations skew more circular. This suggests that gray scale values and membership functions within the area of a particle and across the perimeter of the particle is heterogenous and residing more in the intersections between classes. If these are accounted for the shape across the particle is less consistent than the human segmentation implies. Consequently, exploring the area of human segmented boundaries is a meaningful source of gray-scale values. This difference in quantitative features derived from the enhanced fuzzy c-

means membership represent a first step towards quantifying and understanding epistemic uncertainty of particle segmentation.

CONCLUSION

Morphology is an emerging analytical signature for nuclear forensic applications, especially for understanding sample process history. To use morphology signatures, SEM images must be segmented and quantitatively analyzed, either by a human or an algorithm. This work focuses on the exploration and quantification of epistemic uncertainty of the segmentation to understand the impact on quantitative metrics used for data analysis while maintaining an agnostic view towards how segmentations are produced. The enhanced fuzzy c-means algorithm allows us to take advantage of a membership function to represent the uncertain clusters of pixels in a segmented image. Using the membership function allows for the examination of image quantification metrics in comparison to the human segmentation, by looking at neighborhoods of pixels around the segmentation and quantify the ambiguity. This represents an important start towards quantifying uncertainty in SEM morphology signatures in forensic applications.

ACKNOWLEDGEMENTS

The images collected as well as the methodological development for this study was supported by the Department of Energy's National Nuclear Security Administration, Office of Defense Nuclear Nonproliferation Research and Development under: LA21-ML-MorphologySignature-P86-NTNF1b.

REFERENCES

1. The Evolving Missions of Nuclear Forensics, The National Interest, July 2021.
2. Porter, R.B. and C.E. Ruggiero. *LA-UR-14-23625. MAMA Software Features: Quantification Verification Documentation-1*. 2014.
3. Ruggiero, C.E. and R.B. Porter. *LA-UR-14-23579. MAMA Software Features: Quantified Attributes*. 2014.
4. Olsen, A.M., et al., *Quantification of high temperature oxidation of U_3O_8 and UO_2*, in *Journal of Nuclear Materials*. 2018. p. 574-582.
5. Abbott, E.C., et al., *Dependence of UO_2 surface morphology on processing history within a single synthetic route*, in *Radiochimica Acta*. 2019. p. 1121-1131.
6. Hanson, A.B., et al., *Impact of Controlled Storage Conditions on the Hydrolysis and Surface Morphology of Amorphous-UO_3*, in *ACS Omega*. 2021. p. 8605-8615.
7. Schwerdt, I.J., et al., *Uranium oxide synthetic pathway discernment through thermal decomposition and morphological analysis*, in *Radiochimica Acta*. 2018. p. 193-205.
8. OpenCV. (2015). *Open Source Computer Vision Library*.
9. Sentz, K., Ferson, S., *Combination of Evidence in Dempster-Shafer Theory*. 2002.
10. Szilagyi, L., et al., *MR Brain Image Segmentation Using and Enhanced Fuzzy C-Means Algorithm*, in *Proceedings of the 25th Annual Internal Conference of IEEE EMBS.* Cancun, Mexico. 2003.
11. Tamasi, A.L., et al. *A lexicon for consistent description of material images for nuclear forensics* in *J Radioanal Nucl Chem*. 2016. **307,** p 1611–1619.
12. Bezdek, J.C., Pal, S.K., *Fuzzy Models for Pattern Recognition*, 1991.
13. Ahmed, M. N., et al., *A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data* in *IEEE Transactions on Medical Imaging.* 2002. 32, 4, pp. 193-199.