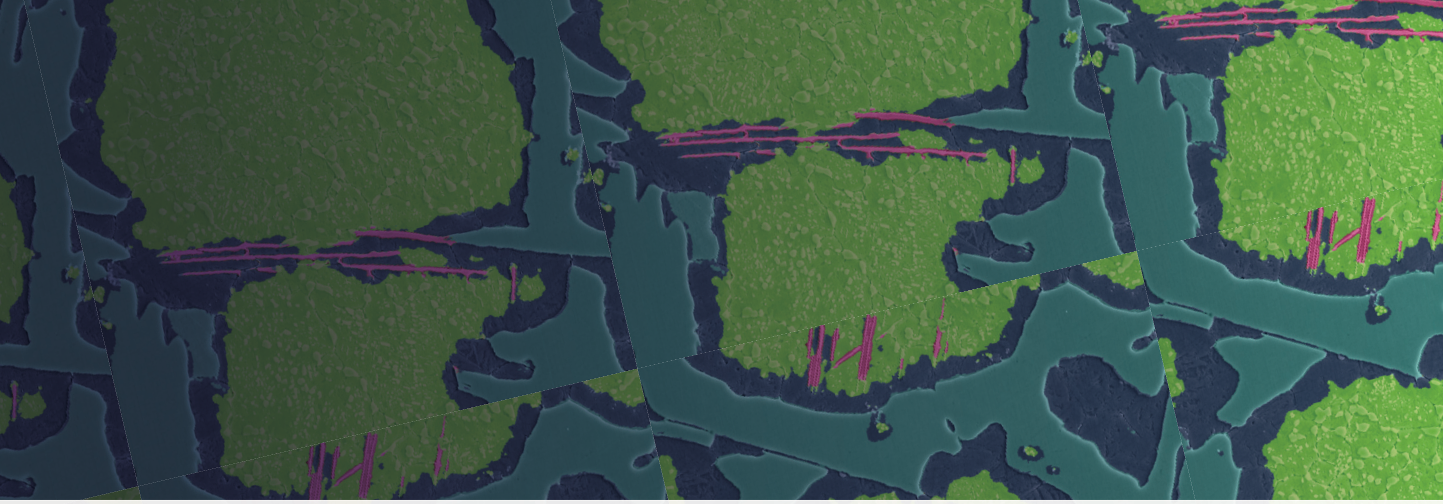


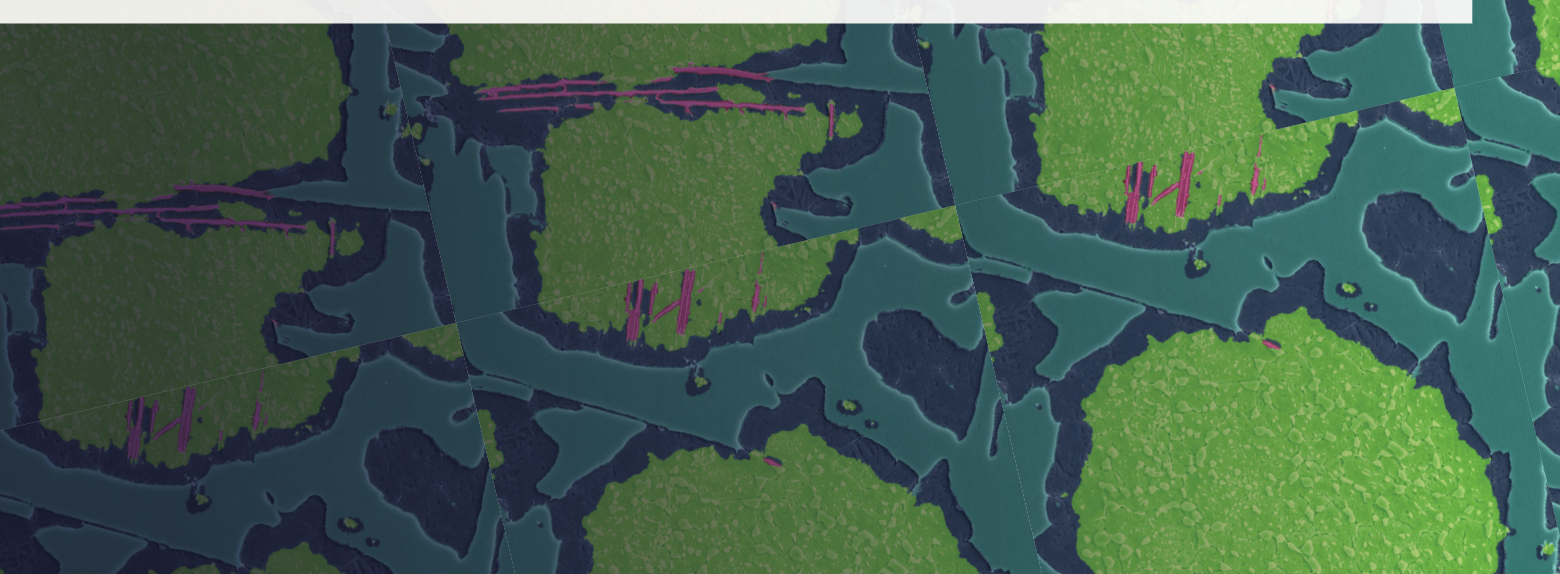
COMPUTER VISION AND MACHINE LEARNING TO QUANTIFY MICROSTRUCTURE

Elizabeth A. Holm, FASM,* Ryan Cohn,* Nan Gao,* Andrew R. Kitahara,*
Bo Lei,* and Srujana Rao Yarasi*
Carnegie Mellon University, Pittsburgh

Thomas P. Matson, Carnegie Mellon University, Pittsburgh
and MIT, Cambridge, Massachusetts

Computer vision and machine learning systems for microstructural characterization and analysis are used for a variety of image analysis tasks, including image classification, semantic segmentation, object detection, and instance segmentation, leading to accurate, autonomous, objective, repeatable results in an indefatigable and permanently available manner.

**Member of ASM International*



Quantitative representation of microstructure is the foundational tool of microstructural science, connecting the material's structure to its composition, process history, and properties. Microstructural quantification traditionally involves a human deciding *a priori* what to measure and then devising a purpose-built method for doing so. However, recent advances in data science, including computer vision (CV) and machine learning (ML), offer new approaches to extracting information from microstructural images[1-7]. The objective of CV is to represent the visual content of an image in numerical form, and ML makes use of these representations to accomplish a given goal. Given a microstructural image, a CV/ML system can perform a variety of analysis objectives, including image classification (e.g., ferritic, austenitic, martensitic), property prediction (e.g., yield strength), feature measurement (e.g., grain size), constituent identification (e.g., phase identification), or a host of other characterization tasks. The CV/ML approach is not a single solution that addresses every microstructural science challenge, but it offers a path toward objective, repeatable, generalizable, and scalable methods that complement the traditional materials characterization workflow.

...THE MOST VALUABLE MICROSTRUCTURAL DATA SETS INCLUDE METADATA THAT ENRICHES THEIR INFORMATION CONTENT.

COMPUTER VISION AND MACHINE LEARNING

Computer vision encompasses an array of methods for creating a numerical representation of a visual image, termed the feature vector[8]. Machine learning methods then extract quantitative visual information from the high-dimensional feature vector[9]. Most high-performance CV/ML systems currently use convolutional neural networks (CNNs), which take an image as input, apply a variety of signal processing operations to it in order to encode it as a vector, and then utilize an artificial neural network or other ML method to draw a conclusion about the visual content of the image[10,11]. The first part of the CNN pipeline—encoding the image as a feature vector—is termed the feature learning stage, and the second part—drawing a conclusion—is the classification stage.

Designing and training a CNN requires deep expertise and a large data set (typically millions of images), making it impractical for most microstructural data sets. However, CNNs that have been optimized and trained on a large set of natural images have been successfully used with other kinds of images, including microstructures. This transfer learning[12] approach enables using pre-trained CNNs (such as the VGG16 network[13] trained on the ImageNet data set[14]) for microstructural representation. However, because the goal is not to classify microstructural images into the ImageNet categories (broccoli, bucket, bassoon), the network is truncated before the classification stage. Instead, the CNN layers themselves are used as the image representations for ML tasks.

Machine learning methods are either supervised (trained using known correct answers, termed ground truth) or unsupervised (finding patterns without knowledge of a ground truth), and there are important roles for each approach. Supervised ML methods make predictions about new data based on information learned from training data with known ground truth answers[10,11,15]. In contrast, unsupervised ML algorithms find relationships between images without ground truth data or human intervention, typically by generating clusters of related images[6,16]. The choice of ML modality and model depends on the nature of the input data and the desired outcome. In this process, it is helpful to include a domain expert in ML algorithms because the best-in-class solutions are ever-evolving.

MICROSTRUCTURAL IMAGE DATA

When assembling an image data set for CV/ML analysis, image quality is less important than exposing the ML system to the full scope of the data space. This is not a recommendation to ignore good microscopy practices, but rather a suggestion that many acceptable images are better than one perfect image. Data collection practices that increase the performance of CV/ML systems include taking redundant images of a sample with non-overlapping fields of view, standardizing imaging conditions (such as instrument, settings, magnification, and orientation), and data augmentation via subsampling or affine transformations such as translation or rotation[17]. Moreover, the most valuable microstructural data sets include metadata that enriches their information content. Metadata may include multiple imaging modalities (e.g., EBSD and backscatter data for the same field of view), as well as information on material system, composition, imaging information, processing history, property measurements, and any other data available related to the image.

Data size is often assumed to be the limiting factor in developing CV/ML methods, and in some cases it is. However, excellent results have been achieved with very small numbers of original micrographs (sometimes fewer than 10). This has to do with the data-richness of microstructural images. The upshot is that a relatively modest investment in data may yield a successful CV/ML system.

IMAGE CLASSIFICATION AND CHARACTERIZATION

Image classification may not seem important, because microstructures are usually known. However, classification of images underlies a host of critical archiving and analysis tasks. Classification relies on the fact that the CV feature vector is a numerical representation of the visual information contained in an image. As such, similarities in the feature vector should correspond to visual similarities. Thus, the distance between two feature vectors can be used to perform visual search, clustering, and classification. For example, for a database of 961 ultrahigh carbon steel (UHCS) microstructures[18] Fig. 1 shows the three

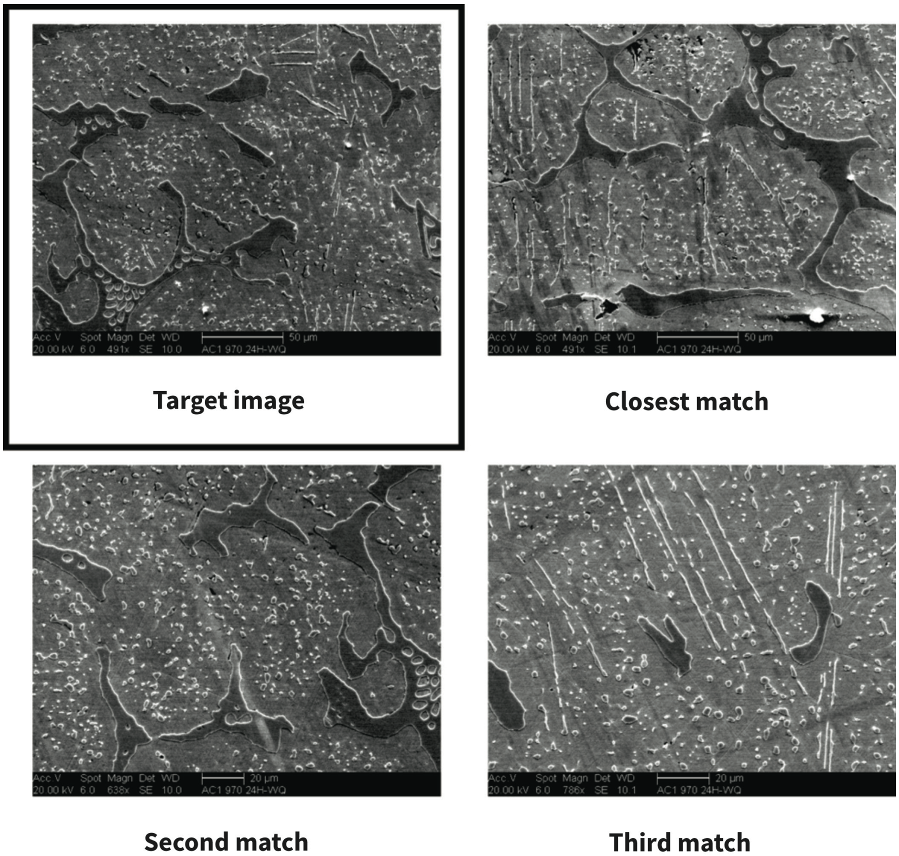


Fig. 1 — Visual search for images with similar feature vectors in a database of 961 ultrahigh carbon steel micrographs[18]. Visually similar micrographs contain similar microstructural constituents, here comprised of a mix of network, Widmanstatten, and spheroidite carbides.

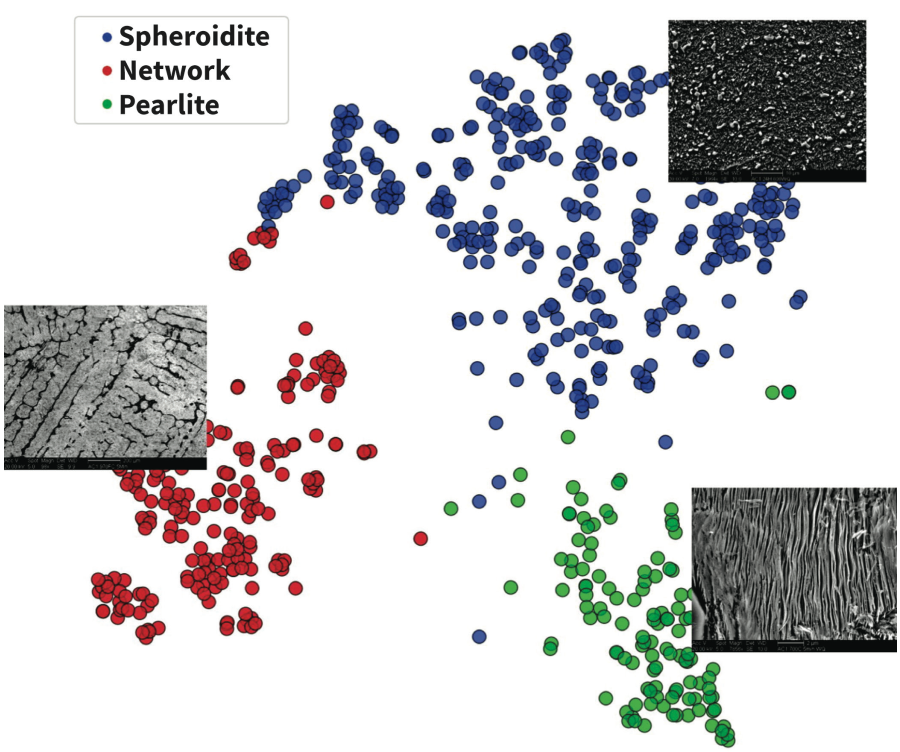


Fig. 2 — Visual clustering plot in which micrographs cluster according to their primary microconstituent: spheroidite (blue), network carbide (red), or pearlite (green)[2]. Example microstructures for each cluster are shown in the insets.

images with feature vectors closest to that of a given target image; obviously, feature vector similarity is reflected in visual similarity. This makes it easy to search an image database for related microstructures. For the same set of UHCS micrographs, Fig. 2 shows a visual clustering map where each point represents an image; point color corresponds to the primary microstructural constituent in each micrograph. Clearly, similar images cluster, which illustrates the visual structure of the data set.

Feature vectors can also be used to quantify microstructural information directly. Figure 3 shows an example that uses the feature vector to measure average grain size in polycrystalline microstructures; the results are within a standard error of 2.3%. Finally, the feature vector can contain visual information that is not perceptible to humans. For instance, chemical composition is not usually measured visually, but rather with specialized tools such as energy dispersive spectroscopy (EDS). Figure 4 shows the results of a ML approach that achieves 76% total accuracy in classifying the composition of inclusions in steel from backscattered SEM images. This demonstrates the ability of the CV/ML system to sense subtle visual details like feature size, shape, contrast, and color distribution with a fidelity that exceeds human perception.

SEMANTIC SEGMENTATION

Quantitative measurement of materials microstructure typically requires the image to be segmented, where each pixel in the image is assigned to a microstructural constituent. Conventional segmentation algorithms, such as those incorporated in ImageJ[19], can work well on suitable microstructures, but become less effective for complex or non-ideal images and often require considerable human intervention. Therefore, we turn to CV/ML methods to address these challenges.

Image segmentation has important applications in robotics and medical imaging among others, so there is considerable research activity in developing segmentation methods. These methods can be adapted to microstructural images via transfer learning. For example, the PixelNet CNN[20] trained on the ImageNet database of natural images[14] has been used to classify pixels according to their microstructural constituent as shown in Fig. 5. In Fig. 5a, the system was trained using 20 hand-annotated images from the UHCS micrograph database[3], and in Fig. 5b, the system was trained on 30 hand-annotated images from a set of tomographic slices of an Al-Zn solidification dendrite[21]. In both cases, the predicted segmentations are arguably equal in quality to the human annotations, and certainly adequate for quantitative analysis. Besides the excellent performance, the CV/ML system is also fast, autonomous, objective, and repeatable, enabling the high throughput necessary for applications such as 3D reconstruction or quality control.

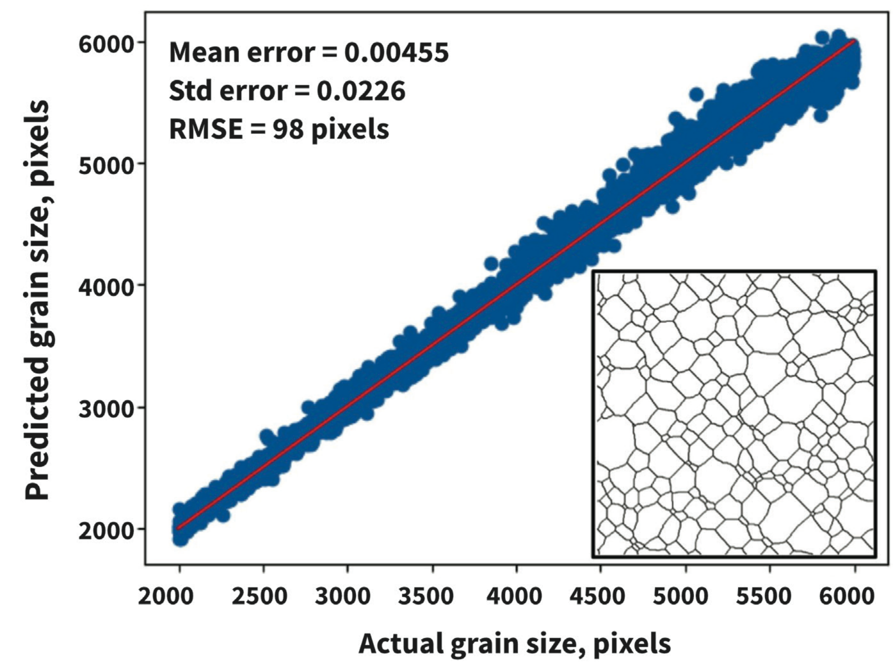


Fig. 3 — Measurement of average grain size from a database of 15,213 synthetic polycrystalline microstructures using deep regression. The red line corresponds to perfect accuracy. Inset shows an example microstructure.

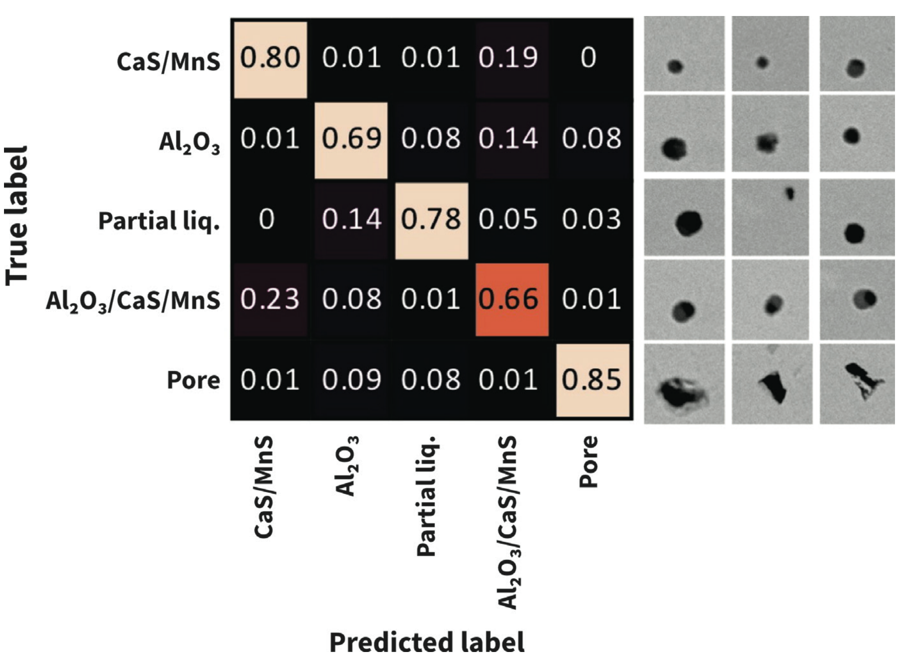


Fig. 4 — Classification results for steel inclusion composition from a database of 2543 backscattered SEM image patches (example images are shown to the right). The prediction accuracy for each inclusion type is shown along the diagonal; overall accuracy is about 76%.

An additional benefit of this approach is the ability to capture human-like judgments about image features. For instance, in Fig. 5a, the spheroidite matrix constituent, comprised of spheroidite particles in a ferrite matrix, is segmented as a single constituent (orange). Likewise, in Fig. 5b, the system learns to ignore sample preparation artifacts such as the sample edge, pores, and the circular beam spot at the center of the image. Conventional segmentation methods would be challenged to handle these complex features. It is this capacity for learning what to look for and what to ignore that distinguishes the CV/ML approach to semantic segmentation.

OBJECT DETECTION AND INSTANCE SEGMENTATION

Object detection entails locating each unique object of its kind in an image, i.e., finding each individual precipitate in a micrograph. Instance segmentation extends this technique to also generate segmentation masks for each individual object. Specialized CNNs have been developed for object detection and instance segmentation[22]. As in the case of semantic segmentation, transfer learning allows models trained on natural images to be adapted to materials science applications.

For example, the presence of small satellite particles is known to affect the flowability of metal powders used in additive manufacturing. A CV/ML approach utilizing object detection and instance segmentation demonstrated the ability to identify individual powder particles and their satellites in dense powder images. Tedious manual annotation yielded five and ten labeled images for powder particles and satellites, respectively. The CV/ML system was trained on these images, and sample predictions are shown in Fig. 6. The powder particle masks showed very good agreement

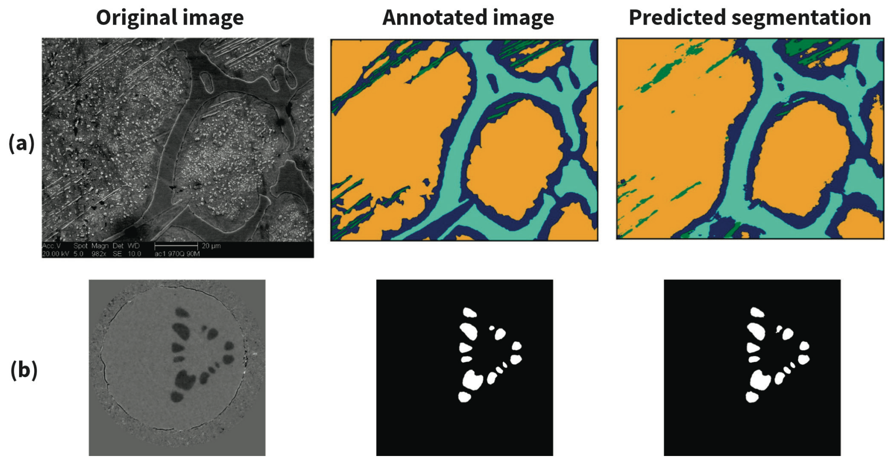


Fig. 5 — Semantic segmentation of microstructural images using a CV/ML system. (a) Segmentation of microstructural constituents in an SEM micrograph of ultrahigh carbon steel. Constituents include network carbide (light blue), ferritic denuded zone (dark blue), Widmanstatten carbide (green), and spheroidite matrix (gold). (b) Segmentation of a tomographic section of an Al-Zn alloy. The solidification dendrite is shown in white on a black background.

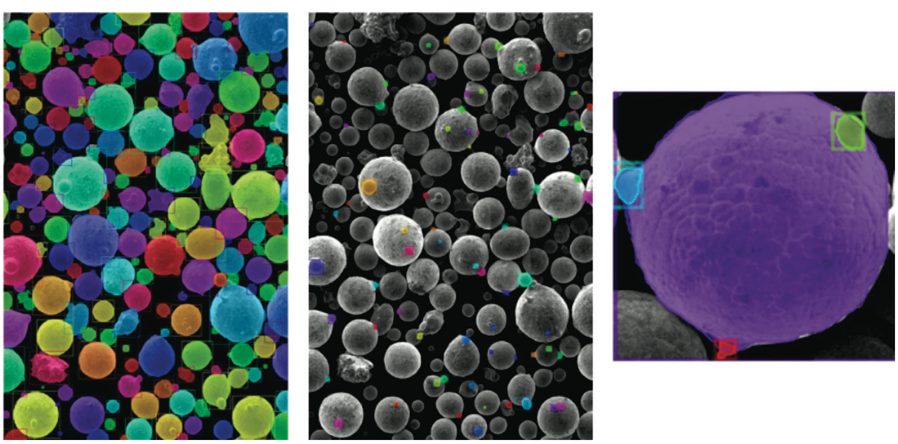


Fig. 6 — Predicted powder particle (left) and satellite (middle) segmentation masks for SEM images of metal powders used in additive manufacturing. Colors are randomly assigned for visual clarity and do not have physical significance. Sample satellited powder particle (right) detected by overlaying the powder particle and satellite masks.

with the manual annotations and indicated that the model approached human-level performance for identifying individual particles. Detecting satellites is a much harder problem, resulting in lower model performance. However, most of the predictions still matched with the annotations, indicating that satellites can consistently be detected in these images. Overlaying the particle and satellite masks to determine the fraction of particles that contain satellites provided a new, objective, and self-consistent method of characterizing the satellite content of powder samples that showed good agreement with the expected trends for images of different powder samples.

CONCLUSIONS

The key function of CV is to numerically encode the visual information contained in a microstructural image for ML algorithms to find associations and trends. CV/ML systems for microstructural characterization and analysis span the gamut of image analysis tasks, including image classification, semantic segmentation, object detection, and instance segmentation. Applications include:

- Visual search, sort, and classification of micrographs via feature vector similarity.
- Extracting information not readily visible to humans, such as chemical composition in SEM micrographs.
- Performing semantic segmentation of microstructural constituents with a high accuracy and human-like judgment about what to look for and what to ignore.
- Finding all instances of individual objects, even when they impinge and overlap.
- Segmenting individual objects to enable new capabilities in microstructural image analysis.

A common characteristic among all of these applications is that they capitalize on the ability of computational systems to produce accurate, autonomous, objective, and repeatable results in an indefatigable and permanently available manner. **~AM&P**

Acknowledgments

This work was supported by the National Science Foundation under grant CMMI-1826218 and the Air Force D3OM2S Center of Excellence under agreement FA8650-19-2-5209.

For more information: Elizabeth Holm, professor, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, eaholm@andrew.cmu.edu, cmu.edu.

References

1. E.A. Holm, et al., Overview: Computer Vision and Machine Learning for Microstructural Characterization and Analysis, *Materials and Metallurgical Transactions A, 51*(12), p 5985-5999, 2020.

2. B.L. DeCost, T. Francis, and E.A. Holm, Exploring the Microstructure Manifold: Image Texture Representations Applied to Ultrahigh Carbon Steel Microstructures, *Acta Materialia, 133,* p 30-40, 2017.

3. B.L. DeCost, T. Francis, and E.A. Holm, High Throughput Quantitative Metallography for Complex Microstructures using Deep Learning: A Case Study in Ultrahigh Carbon Steel, *Microscopy and Microanalysis, 25*, p 21-29, 2019.

4. B.L. DeCost, E.A. Holm, Vision-based Methods in Microstructure Analysis, *Statistical Methods for Materials Science: The Data Science of Micro-*

structure Characterization, CRC Press, p 241-258, 2019.

5. B.L. DeCost, et al., Computer Vision and Machine Learning for Autonomous Characterization of AM Powder Feedstocks, *JOM, 69,* p 456-465, 2017.

6. A.R. Kitahara and E.A. Holm, Microstructure Cluster Analysis with Transfer Learning and Unsupervised Learning, *Integrated Materials and Manufacturing Innovation, 7,* p 148-156, 2018.

7. J. Ling, et al., Building Data-driven Models with Microstructural Images: Generalization and Interpretability, *Materials Discovery, 10,* p 19-28, 2017.

8. R. Szeliski, *Computer Vision: Algorithms and Applications,* Springer, New York, 2010.

9. P. Flach, *Machine Learning: The Art and Science of Algorithms that Make Sense of Data,* Cambridge University Press, 2012.

10. Y. LeCun, Y. Bengio, and G. Hinton, Deep Learning, *Nature, 521,* p 436-444, 2015.

11. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning,* MIT Press, 2016.

12. G. Csurka, et al., Visual Categorization with Bags of Keypoints, Workshop on Statistical Learning in Computer Vision, ECCV 1, p 1-16, 2014.

13. K. Simonyan and A. Zisserman, Very Deep Convolutional Networks for Large-scale Image Recognition, International Conference on Learning Representations, arXiv:1409.1556, p 1-14, 2015.

14. O. Russakovsky, et al., Imagenet Large Scale Visual Recognition Challenge, *International Journal of Computer Vision, 115,* p 211-252, 2015.

15. T.M. Mitchell, *Machine Learning,* McGraw-Hill, 1997.

16. S. Lloyd, Least Squares Quantization in PCM, *IEEE Transactions on Information Theory, 28,* p 129-137, 1982.

17. C. Shorten and T.M. Khoshgoftaar, A survey on Image Data Augmentation for Deep Learning, *Journal of Big Data, 6,* p 60, 2019.

18. B.L. DeCost, et al., UHCSDB: UltraHigh Carbon Steel Micrograph DataBase, *Integrated Materials and Manufacturing Innovation, 6,* p 197-205, 2017.

19. C.A. Schneider, W.S. Rasband, and K.W. Eliceiri, NIH Image to ImageJ: 25 Years of Image Analysis, *Nature Methods, 9,* p 671-675, 2012.

20. A. Bansal, et al., PixelNet: Representation of the Pixels, by the Pixels, and for the Pixels, arXiv:1702.06506 [cs.CV], p 1-17, 2016.

21. T. Stan, Z. Thompson, and P. Voorhees, Optimizing Convolutional Neural Networks to Perform Semantic Segmentation on Large Materials Imaging Datasets: X-ray Tomography and Serial Sectioning, *Materials Characterization, 160,* 110119, 2020.

22. L. Liu, et al., Deep Learning for Generic Object Detection: A Survey, *International Journal of Computer Vision, 128,* p 261-318, 2020.