

Grain and Grain Boundary Segmentation using Machine Learning with Real and Generated Datasets

Peter Warren¹, Nandhini Raju², Abhilash Prasad², Shahjahan Hossain², Ramesh Subramanian³, Jayanta Kapat², Navin Manjooran¹ and Ranajay Ghosh^{2*}

¹*Solve Technology and Research, Inc., 3505 Lake Lynda Drive,
Suite-200, Orlando, 32817, FL, USA.

²*Department of Mechanical and Aerospace Engineering,
University of Central Florida, 4000 Central Florida Blvd,
Orlando, 32816, FL, USA.

³Siemens Energy, Inc., 11842 Corporate Blvd, Orlando, 32817,
FL, USA.

*Corresponding author(s). E-mail(s): Ranajay.Ghosh@ucf.edu;

Abstract

We report significantly improved accuracy of grain boundary segmentation using Convolutional Neural Networks (CNN) trained on a combination of real and generated data. Manual segmentation is accurate but time-consuming, and existing computational methods are faster but often inaccurate. To combat this dilemma, machine learning models can be used to achieve the accuracy of manual segmentation and have the efficiency of a computational method. An extensive dataset of from 316L stainless steel samples is additively manufactured, prepared, polished, etched, and then microstructure grain images were systematically collected. Grain segmentation via existing computational methods and manual (by-hand) were conducted, to create "real" training data. A Voronoi tessellation pattern combined with random synthetic noise and simulated defects, is developed to create a novel artificial grain image fabrication method. This provided training data supplementation for data-intensive machine learning methods. The accuracy of the grain measurements from microstructure images segmented via computational methods and machine learning methods proposed in this

2 Machine Learning Grain Segmentation

work are calculated and compared to provide much benchmarks in grain segmentation. Over 400 images of the microstructure of stainless steel samples were manually segmented for machine learning training applications. This data and the artificial data is available on Kaggle.

Keywords: Grain Boundary, Grains, Neural Network, Machine Learning, Artificial Data Generation, Additive Manufacturing

1 Introduction

Grains are a fundamental hierarchical unit of material organization. For metals, a grain is a region within the material that has a uniform, crystalline structure. Grain size, shape and distribution have a significant effect on the mechanical properties. In general, metals with small, uniform grains are stronger and more ductile than those with large, irregular grains[1–5]. In addition to strength and ductility, the grain size and shape can also affect other properties, such as its electrical and thermal conductivity, corrosion resistance, and wear resistance [6–8]. Considering the singular importance of grain structure, its measurement and quantification remains a key objective of materials analysis and characterization. Recent advances in metal additive manufacturing[9–13], that depend on layer by layer deposition lead to non-traditional grain structures and hence unique material properties that can be distinct from traditional manufacturing methods[14–16]. Such new technological advances further underscore the need for rapid evaluation of grain structures for predictions and certifications.

Traditionally grain structures are visible under either optical or Scanning Electron Microscopes (SEM). The grains are then observed via cross sectional imaging and appear like irregular polygons. Although their geometrical structures can be noted manually in a straightforward way, it is too time consuming and impractical for large sample sizes or high throughput testing. Thus, computational techniques that rely on advance image processing algorithms are often employed [15, 17–20]. However, these techniques suffer from well-known inefficiencies and inaccuracies [14, 21]. These include sensitivity to image quality, parameter tuning requirements, and a lack of standardization [22, 23]. These problems can be addressed using data based Machine Learning (ML) methods. However ML methods are critically dependent on training data sets for their success [24]. Such training data sets are often tedious, expensive and difficult to obtain, severely restricting the use of machine learning.

This problem is addressed in this paper by developing a new approach whereby a hybrid experimental and Voronoi cell generated images are used as training set. Artificial grain images are produced by exploiting the geometries that occur in a Voronoi cell plot, and can therefore be controlled and manipulated through the mathematical properties of the Voronoi method. In addition, simulated defects such as pores and polishing scratches were added,

and guassian noise is added in a myriad of ways to create a finalized and realistic grain image. The artificial grain generation method in this work produces equiaxed grains, which was observed on the experimentally gathered data. This is not a limitation as this method could easily be tuned to create other grain structure patterns as well. The artificial and authentic (naturally occurring) training data was used to train CNN's to perform grain segmentation. Once segmentation is performed the accuracy was calculated from the geometric grain measurements and from the pixel segmentation accuracy. The grain and grain boundary segmentation process can be challenging due to pores, impurities, polishing scratches, and small grain fragments in the images. An illustrative example of these challenges is shown in Fig 1. Artificially generated data methods are often computationally expensive, and this novel method significantly reduces computation by generating the underlying geometric pattern and adding intricacies directly through mathematical pattern manipulation in euclidean space.

Traditional techniques such as the line intercept and planimetric methods have been widely used for this purpose [25, 26]. In this study, a large dataset of metallic grains from cubical samples of stainless steel 316L additively manufactured using an InnoventX ExOne printer (Binder Jet Technology) and an optical microscope is collected and compiled. After printing, sintering, polishing, and etching the samples, a total of 640 images were collected and manually segmented as the baseline for evaluating various segmentation methods. This dataset provides a valuable resource for testing and comparing different machine learning techniques for grain boundary segmentation [27, 28].

The rest of the paper is organized as follows, a description of the traditional image segmentation techniques is given in Section 2. These include manual thresholding, gradient based approach, holistically nested edge detection (HED), and manual segmentation. A description of how methodology accuracy is calculated, and accuracy of segmentation is calculated is provided in Section 3. The design and implementation of the novel artificial grain generation method developed in this work is provided in Section 4. The CNN architecture and the composition of the various training sets is provided in Section 5. Finally the results and discussion is given in Section 6, and then the conclusion in Section 7.

2 Image Segmentation Techniques

The section of the paper will explain the four most common ways to segment an image into grains and grain boundary. The four ways are: (1) Manual Thresholding (2) Gradient Based Methods (3) HED method (4) Manual Segmentation. In practice these methods can be combined in a variety of ways to segment and image. The HED method is an example of a neural network style approach. There are many new developing methods that use neural networks to detect edges of an image, HED is just the most well known and popular at

4 Machine Learning Grain Segmentation

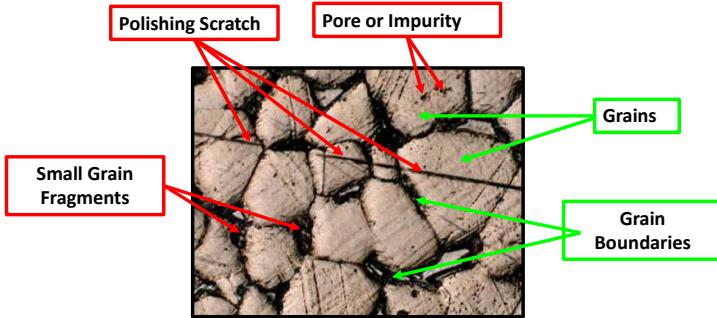


Fig. 1 An example image of a metallic microstructure which illustrates the challenges during classification of pixels into either grains or grain boundaries. These challenges include: polishing scratches, pores, impurities, and small grain fragments.

the time of writing this paper. The HED method was developed specifically for detecting the edges of an image in any image.

2.1 Manual Thresholding

The coding for the manual thresholding was done using OpenCV [29]. All of the code for this work is available on github and kaggle [30]. A step by step approach to manual thresholding can be seen in Fig 2. Step 1 is to collect the microstructure image. The material in this work is 316L stainless steel printed on an EXone printer from an optical microscope at 500X. Step 2: Conversion of image to black and white. Step 3 is a pixel intensity evaluation for pixel cutoff determination. The histogram shown for step 3 is for this purpose, and the X-axis is pixel intensity and the Y-axis shows the number of times that pixel intensity has occurred. The red line on this histogram shows the selected thresholded value, also called the pixel cutoff value. Step 4a and step 4b are both examples of thresholding. Step 4a is an example of standard thresholding where all of the pixels (>130) are set to 255 (white) and all pixels (≤ 130) are set to 0 (black). The value 130 is shown in red on the histogram. Step 4b is an example of an adaptive gaussian threshold. The threshold value is a gaussian-weighted sum of the neighbourhood values minus the constant C. The neighborhood values are a block of pixels, for this case it was (55x55), and the constant subtracted was 2.

In step 5 a median blur is applied 2 times to a 5x5 pixel area, and then 2 more times to a 7x7 pixel area. In step 6 image segmentation is applied, and the grains are separated from the grain boundaries. In step 7 the grains are filtered, usable grains are shown in green and unusable grains are shown in red. Unusable grains are grains that are too small or grains that have too much perimeter per area. Too much perimeter per area indicates the grain did not get properly segmented during the total process. Essentially two or more grains are connected by a small "neck" region.

Image segmentation via threshold is a simple but highly customizable process. There are several steps in this process where the user has to select a value

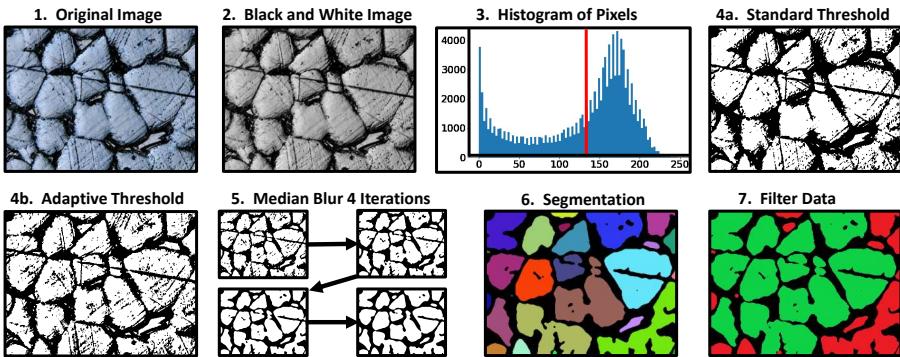


Fig. 2 A step by step example of the manual thresholding technique being applied to a metallic microstructure image with equiaxed grains boundaries.

and check the results. For example, threshold cutoff, block size for blurring, block size for the adaptive gaussian method, etc. All of these values lead to different results, so the user must pick the one that looks the best in order to automate the grain segmentation process. Different materials, imaging equipment, or image settings can effect the images, and could easily require different variables for optimal segmentation.

2.2 Gradient Based Methods

The gradient based approach identifies sharp changes in pixel intensity, and can classify those areas as an edge. The first and second derivatives of pixel intensity as it relates to pixel position are used to determine the gradients. The method used in this work is the "Canny Edge Detection Method," developed by John Canny [31]. First the image is converted to greyscale and then gaussian filters are used on the pixel values to slightly blur the image. This will reduce the amount of predicted edges in the result, which is helpful because the gradient based approach tends predict too many edges. After the gaussian filters, the pixel intensity gradient is calculated (first and second derivatives). Lower changes in pixel intensity are filtered out, unless they directly connected to a high change in pixel intensity.

The application of the Canny Edge Detector can be seen in Fig 3. First the image was converted to black and white. Next the Canny Edge Detector was used. In the third step, two Gaussian blur iterations were applied to a 3x3 pixel block. In the fourth step, the image is thresholded into 0 and 255 (black and white). The threshold value was very high (200), meaning most gray pixels were set to black. Step 5: segmentation occurs. Step 6: Filtering of the data occurs. The same filters were used as the previous section. Too much perimeter per area were filter out, and too little a grain size were also filter out. Image segmentation via gradient detection is a highly customizable process and optimal parameters can be difficult to determine. Small adjustments to pre and post filters will cause significant changes in the final predictions. The

6 Machine Learning Grain Segmentation

engineer or user of this method must tailor the settings and parameters of the algorithm for specific materials or image settings.

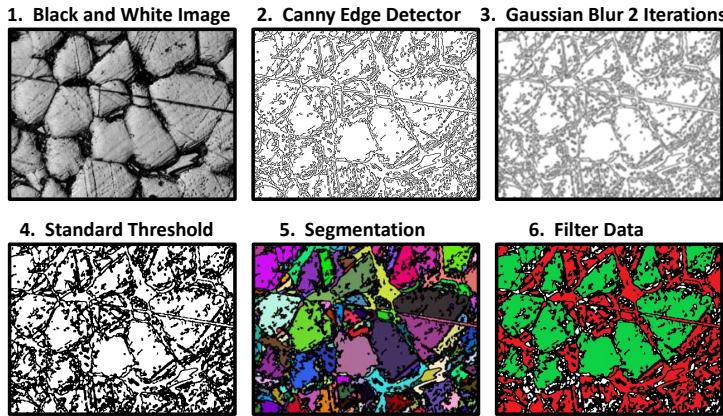


Fig. 3 A step by step example of the gradient based technique being applied to a metallic microstructure image with equiaxed grains boundaries.

2.3 Holistically Nested Edge Detection

The next detection method is deep learning-based method for edge detection in images. The architecture of the HED utilizes a fully convolutional network, which was trained end-to-end on a large dataset of images with ground truth edge maps.[32] CNN's are particularly well-suited for tasks involving image processing and analysis, such as object detection, image classification, and image segmentation [33–35].

In Fig 4, a clear step-by-step guide for utilizing the HED approach is provided. First the image is converted to black and white. Next it is fed into the HED model to determine the edges. The third step is to apply an adaptive gaussian threshold to convert all pixels to either 0 or 255 (white or black). The final two steps are segmentation and data filtration. The HED model outputs a pixel intensity that represents the confidence that the pixel is an edge or not an edge. That is why step 3 (adaptive thresholding) is required. The HED is highly effective at detecting grain boundaries, but its suitability as an end-to-end solution is limited due to its tendency to also detect undesired features such as pores and polishing scratches. This is primarily because HED lacks the ability to distinguish between different types of edges, relying on low-level image features that generalize its edge detection capabilities to various image characteristics.

2.4 Manual Segmentation

The most arduous way to segment the grains of image is the manual segmentation. This is done by the user manually annotating the grains or grain

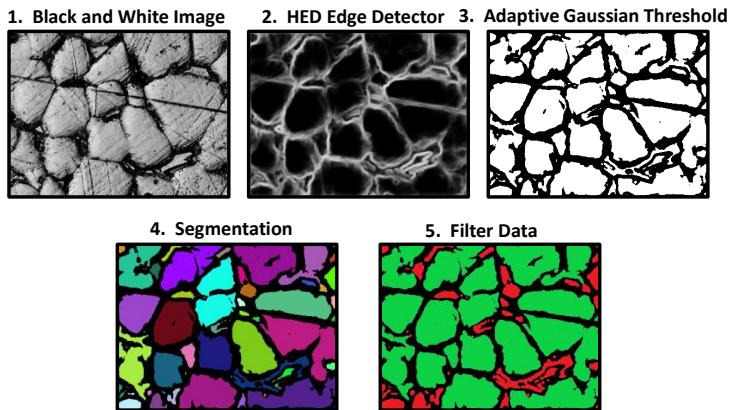


Fig. 4 A step by step example of the (HED) Holistically Nested Edge Detection neural network technique being applied to a metallic microstructure image with equiaxed grains boundaries.

boundaries. For this work it was completed by marking the grain boundaries in a specific color, and then thresholding the image so that all annotated pixels are 0 (black) and all others are 255 (white). After that the same process of segmenting the image computationally and filtering out small particles occurs. This process is shown in Fig 5. All inaccuracies that occur in this method are from human error. Polishing scratches and pores must be avoided during annotation of the grain boundaries. This method cannot be automated, each image requires human attention. For this work all 480 images of grains were painstakingly manually segmented, and this was completed to allow the selected machine learning methods to train and replicate the process. All of the previously mentioned segmentation methods/algorithms were used (HED, Gradient, Threshold) on the dataset to provide benchmarks for each of the methods and the proposed methods in this work. All of this data is available on Github and Kaggle.

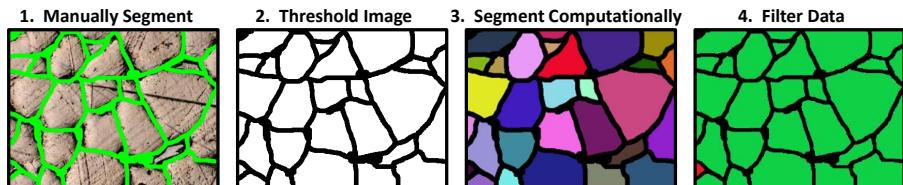


Fig. 5 A step by step example of the manual segmentation being applied to a metallic microstructure image with equiaxed grains boundaries. This was done by hand, by a human.

3 Grain Measurement and Accuracy Quantification

There are multiple ways to quantify how accurate the grain boundary segmentation methods perform. Accuracy quantification can be approached from the perspective of a material scientist or from the perspective of a computer vision scientist. Both approaches will be conducted in this work. The material scientist is interested in the resulting dimensional grain data. Traditional grain measurements techniques are used once the image is segmented. The line intercept (Heyn method) and the planimetric method (Saltykov method) are both used to measure the grain dimensions. These two industry standard methods will be explained in the following sections. The computer vision scientist is interested in the pixel segmentation accuracy. The current standard is the dice score or the Intersection of Union (IoU) score. These two methods are frequently used in computer vision to determine how accurate a neural network is at segmentation, and they will be explained in this section as well.

3.1 Line Intercept Method

The grain boundary image dataset used in this study consists of 480 images, each with dimensions of 400 x 300 pixels. The grain intercept method, initially developed by Heyn and widely employed for rapid grain size measurements [25], involves drawing a line across the image and tallying the number of grains intersected by the line. The average grain size is then determined by dividing the total number of grains by the line length, as shown in Equation 1. To automate this process and conduct measurements at each pixel, code was implemented for all the segmentation techniques outlined in Section 2, resulting in a total of 144,000 X-direction measurements and 192,000 Y-direction measurements. Table 1 presents the average grain size and standard deviation for both the horizontal and vertical directions. Although the grain intercept method provides a substantial number of measurements, it is inherently limited as a 1-dimensional approximation of a complex 3-dimensional structure. Nonetheless, its continued utilization in current research reinforces its relevance in this study.

$$\text{Grain Size} = \frac{\text{Number of Grains in Contact With Line}}{\text{Total Length of the Line}} \quad (1)$$

Table 1 Line Intercept Grain Measurements

| Direction | Manual Threshold | Gradient Method | HED Method | Manual Segmentation |
|-------------|------------------|-------------------|------------------|---------------------|
| X-Direction | 41.45 ± 6.42 | 20.86 ± 10.90 | 33.05 ± 6.53 | 36.87 ± 4.84 |
| Y-Direction | 42.28 ± 7.59 | 22.64 ± 5.99 | 34.02 ± 6.80 | 37.40 ± 5.44 |

3.2 Planimetric Method

The planimetric method, a 2-dimensional approach, enables the extraction of size and shape data regarding the grains. Extensive data from the segmented dataset was collected, which has been categorized into size data (Table 2) and shape data (Table 3). These datasets encompass all the segmentation methods listed in Section 2. The planimetric size data comprises four measurements. The first one is the total grains metric which represents the overall count of grains across all images. The second measurement is the average individual grain area, accompanied by its standard deviation. The third and fourth entries in Table 2 correspond to the total grain area per image and the total grain boundary area per image. These two values can be computed interchangeably using Equation 2.

$$\text{Total Image Area} = \text{Total Grain Area} + \text{Total Grain Boundary Area} \quad (2)$$

$$\text{Circularity} = \frac{4\pi A}{P^2} \quad (3)$$

Table 2 Planimetric Size Measurements

| Measurement | Manual Threshold | Gradient Method | HED Method | Manual Segmentation |
|-------------------------------|-------------------------|------------------------|-----------------------|------------------------|
| Total Grains | 9,575 | 24,740 | 15,297 | 11,002 |
| Individual Grain Area Average | $844 \pm 1,314 \mu m^2$ | $248 \pm 1033 \mu m^2$ | $488 \pm 823 \mu m^2$ | $1068 \pm 929 \mu m^2$ |
| Total Grain Area % | 53.3% | 40.5% | 49.3% | 77.5% |
| Total Grain Boundary Area % | 46.7% | 59.5% | 50.7% | 22.5% |

The planimetric shape data encompasses five distinct measurements. Circularity represents a geometric property of a shape, ranging between 1 and 0, where 1 denotes a perfect circle and 0 signifies the theoretical opposite of a perfect circle. The closer the value is to 0, the rougher the 2-dimensional shape becomes. The equation for circularity can be seen in Equation 3. Another measurement is the average maximum diameter, which represents the maximum distance from one side of the grain to the other, passing through the centroid of the grain. The subsequent three measurements pertain to the X and Y directions. The first is the average width of the grains, denoted as the average X-diameter. The next measurement is the average diameter in the Y-direction, also known as the average height. The final value is the aspect ratio, obtained

by dividing the average X-diameter by the average Y-diameter, providing a measure of the width-to-height relationship.

Table 3 Planimetric Shape Measurements

| Measurement | Manual Threshold | Gradient Method | HED Method | Manual Segmentation |
|----------------------|-------------------|-------------------|-------------------|---------------------|
| Circularity | 0.55 ± 0.18 | 0.28 ± 0.15 | 0.50 ± 0.18 | 0.61 ± 0.12 |
| Average Max Diameter | 39.31 ± 30.69 | 25.14 ± 18.60 | 30.57 ± 22.84 | 47.38 ± 20.52 |
| Average X-Diameter | 32.18 ± 25.39 | 22.64 ± 15.90 | 24.81 ± 19.51 | 39.07 ± 18.49 |
| Average Y-Diameter | 32.94 ± 26.70 | 22.64 ± 15.40 | 25.47 ± 19.76 | 39.06 ± 18.97 |
| Aspect Ratio X/Y | 1.09 ± 0.55 | 1.06 ± 0.53 | 1.07 ± 0.51 | 1.11 ± 0.58 |

3.3 Evaluation of Segmentation

For assessing the segmentation model accuracy, the dice score and the IoU (Intersection over Union) score will be employed. When classifying pixels in a segmented image, there are four possible outcomes: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The dice equation, represented by these variables, can be seen in Equation 4. Additionally, the probabilistic version of the equation is provided in Equation 5. These equations serve as quantitative measures to evaluate the performance of the segmentation models.

$$Dice = \frac{2TP}{2TP + FN + FP} \quad (4)$$

$$Dice(A, B) = \frac{2 * |A \cap B|}{|A| + |B|} \quad (5)$$

To facilitate implementation, it is crucial to establish and indicate a ground truth. In this study, the manually segmented images serve as the ground truth, with the only potential error arising from human inaccuracies during the grain boundary sketching process. Table 4 showcases the comparison of Dice and IoU scores for the manually thresholded method, the gradient method, and the HED method. These scores can serve as benchmarks, alongside other grain properties, to quantify the accuracy of the machine learning methods developed in this research.

Table 4 Planimetric Shape Measurements

| Scoring Method | Manual Threshold | Gradient | HED |
|----------------|-------------------|-------------------|-------------------|
| Dice | 0.55 ± 0.18 | 0.28 ± 0.15 | 0.50 ± 0.18 |
| IoU | 39.31 ± 30.69 | 25.14 ± 18.60 | 30.57 ± 22.84 |

4 Artificially Generated Grains

In this work, a novel approach was developed to generate artificial grains and grain boundaries using a Voronoi tessellation pattern. By leveraging the Voronoi tessellation pattern, specifically tailored to meet the desired conditions, a pattern closely resembling equiaxed grain boundaries was achieved. To ensure realism and account for practical challenges, artificial noise was intentionally introduced to simulate pores, scratches, and various image quality issues. This innovative Voronoi grain method, developed within the scope of this research, enables the generation of representative synthetic grain structures for further analysis and evaluation.

4.1 Voronoi Tessellation

A Voronoi tessellation pattern is generated by dividing a plane into regions based on the distances to specific points in the plane [36]. Each point, also referred to as a seed, corresponds to a region comprising all the areas in the plane that are closer to that point than to any other. This partitioning of the plane into regions defined by proximity to the seeds forms a Voronoi diagram. An example of a Voronoi diagram, utilized in this work, is depicted in Figure 6. The size of the Voronoi cells is determined by the area measurements obtained in the previous section. The average area of the Voronoi cells is approximately equal to the reciprocal of the point density. For instance, if the point density is 10 points per unit area, the average cell area would be approximately 0.1 units. Based on this calculation, the average grain (cell) area for the generated data was set to $1000 \mu\text{m}^2$. Comparatively, the average individual grain area for the manually segmented images was found to be $1068 \mu\text{m}^2$, so $1000 \mu\text{m}^2$ was selected to closely represent the actual data.

4.2 Artificial Noise

To accurately replicate the challenges observed in the images, noise was intentionally introduced into the Voronoi tessellation pattern. The main challenges encountered in the images include polishing scratches and pores, as depicted in Figure 1. To incorporate scratches, random lines were plotted within the Voronoi graph, while pores were simulated by randomly placing dots of varying sizes. An illustrative example showcasing these defects can be seen in Figure 6. While these defects could be mitigated through more meticulous polishing techniques and improved etching methods, achieving perfection in these

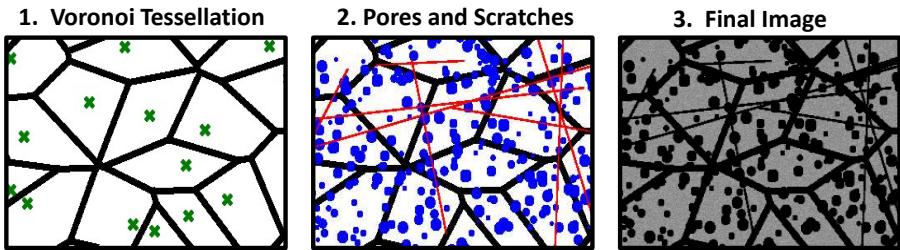


Fig. 6 Artificial grain generation via Voronoi Tessellation pattern: (1) Randomly generated voronoi tessellation pattern with centroids marked in green. (2) Polishing scratches marked in red and large and small pores marked in blue. (3) Gaussian Noise applied to simulate grain color.

complex processes can be challenging, especially for certain metals. Therefore, employing machine learning methods to automatically disregard these defects represents a viable approach to mitigating their undesired effects during the measurement process.

The overall plot area ranges from 0 to 800, but the actual image spans from 200 to 600. Consequently, a 400x400 pixel image, identical to the experimental data, is created. This arrangement allows for the possibility of scratches either traversing the entire sample image or being confined within the image boundaries. During the generation of the artificial data, the quantity of pores and scratches was deliberately varied, and this will be further explained in the subsequent section.

4.3 Artificial Dataset

The quantity of scratches within the entire plot area was determined using a uniform random number generator ranging from 2 to 10. The starting and stopping locations of the scratches were also randomly assigned using uniform random distributions for both the X and Y coordinates. Moreover, the number of randomly sized pores was generated using a uniform random number between 2000 and 4000. Additionally, random black pixels were inserted into the images, with the quantity determined by a uniform random number between 10,000 and 20,000. To impart a gray color to the grains, a subtraction operation was applied to all white pixels (255) using a normally distributed Gaussian distribution centered around 100 ± 10 . Detailed numerical values for all of noise distributions can be found in Table 5. The objective is to train a CNN to accurately predict the mask, representing the initial noiseless Voronoi tessellation, given a noisy Voronoi tessellation. Although a CNN can readily accomplish this task, additional noise will be introduced to enhance the robustness of the model.

To enhance the training process and develop a more robust CNN for image segmentation, the next round of noise was introduced by exploiting the limitations of two segmentation techniques mentioned in Section 2: Manual Thresholding and Gradient-Based Method. This approach aims to leverage the deficiencies of these techniques to further train the model. Specifically, the

Table 5 Applied Noise Distribution Table

| Noise Type | Distribution/Application |
|------------------------------------|-------------------------------------|
| Added Black Pixels | $\mathcal{U} \sim (10,000, 20,000)$ |
| Added Black Dots | $\mathcal{U} \sim (1,000, 2,000)$ |
| Dot Size (Radius) | $\mathcal{U} \sim (5, 25)$ |
| Gaussian Noise Subtracted | $N \sim (100, 10)$ |
| Total Scratches | $\mathcal{U} \sim (2, 10)$ |
| Scratch Start and End Point | $\mathcal{U} \sim (0, 800)$ |
| Total Area Plotted | X and Y : (0, 800) |
| Area in Image (From Original Plot) | X and Y : (200, 600) |

median blur technique described in Section 2.1 and the Gaussian blur technique discussed in Section 2.2 were implemented. Additionally, the inverse of an adaptive threshold, which is similar to an erosion technique, was applied. These techniques were utilized in various iterations and combinations to create a diverse range of noise patterns. The specific application method can be found in Table 6. For a visual representation of the generated noise applied to an original Voronoi tessellation for each set in Table 6, refer to Figure 7.

Table 6 Image Distortion Applications Table

| Set of 100 Images 256x256 | Erosion and Dilation Kernel [3,3] | Gaussian Blur Kernel [15,15] | Median Blur Kernel [5,5] |
|---------------------------|-----------------------------------|------------------------------|--------------------------|
| 1 | YES | NO | YES (1 iteration) |
| 2 | YES | NO | YES (2 iteration) |
| 3 | YES | YES (1 iteration) | NO |
| 4 | YES | YES (2 iteration) | NO |
| 5 | NO | NO | YES (1 iteration) |
| 6 | NO | NO | YES (2 iteration) |
| 7 | NO | YES (1 iteration) | NO |
| 8 | NO | YES (2 iteration) | NO |

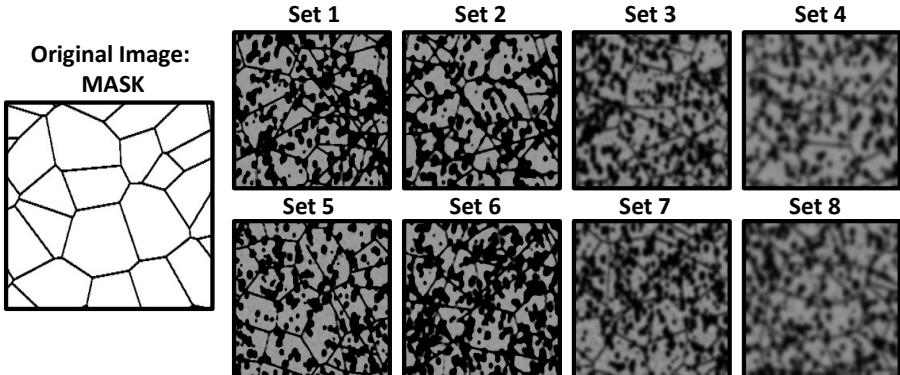


Fig. 7 Artificial grain generation method: The original mask is on the left hand side of the figure and all eight of the different sets of generated grains are labeled Set 1-8. The method of noise generation and application for each set is given in Table 6

5 Machine Learning Methods

The chosen machine learning approach for grain and grain boundary identification is a Convolutional Neural Network (CNN) with a U-Net architecture. The details of the U-Net architecture will be discussed in the section 5.1: U-NET. Multiple training methods were employed and will be elaborated upon in the section covering 5.2: Training Styles.

5.1 U-Net

The U-Net architecture, introduced in 2015 by Ronneberger et al. [37], has emerged as a powerful framework for pixel segmentation tasks. Its design is particularly well-suited for this purpose as it effectively captures both local and global features present in the input image. One key advantage of U-Net is the incorporation of skip connections, enabling the model to leverage both high-level and low-level features during the segmentation process. In many image classification problems, both levels of features contain crucial information necessary for accurate classification.

In the current era of Artificial Intelligence growth, numerous variations of the classic U-Net model have been proposed. However, to maintain consistency and facilitate benchmarking, we have opted to employ the standard U-Net architecture in this study. Comparing the performance of different models can be easily accomplished since all the relevant data is publicly available.

For the activation functions within the layers, the Rectified Linear Unit (ReLU) function was utilized. The final activation function for the classification task was the sigmoid function. The optimizer function employed was ADAM, while the loss function used was binary cross-entropy. During training, a batch size of 16 was utilized, and the training process consisted of 50 epochs. A validation split of 10% was employed to monitor the model's performance during training. The layout of the U-Net architecture employed in this study is depicted in Figure 8.

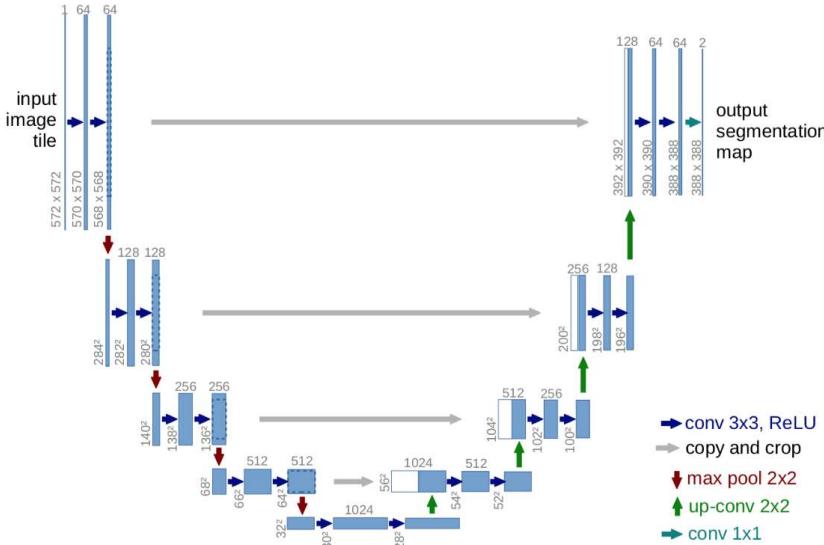


Fig. 8 This is a diagram of a UNET from the First and Original paper on the UNET. caption[37]

5.2 Training Styles

The dataset was split into 80% for training and 20% for testing. Two crucial aspects of the training process are determining the training data and defining the desired outcome, referred to as the mask. The mask represents the specific pixels classified, in this case classified as grain boundaries. In this study, the mask (grain boundary classified pixels) are assigned a value of 0 (black), while all grains are assigned a value of 255 (white). The masks used in this work were primarily manually segmented by human experts, except for the Voronoi-generated data, where the original unaltered Voronoi diagram serves as the mask. An example of such a mask is illustrated in Figure 6.

The total composition of the training data is a critical aspect to consider. The obvious approach is to train the model solely on manually segmented data. However, the question arises: would incorporating the Voronoi tessellation-generated data enhance the model's robustness? This question will be addressed in the results section. The different compositions of the training sets are depicted in Figure 9. The first five compositions involve varying percentages of generated and manually segmented data, allowing for a comparison between the Voronoi method, the Manual method, and a combination of both.

Training sets 6-10 follow a more intricate approach. Traditional segmentation methods from Section 2.1, 2.2, and 2.3 were utilized to preprocess the data. This means that the mask remains the same (manually segmented image), but the original grain image undergoes modifications through the application of

the manual threshold method, the HED method, or the gradient method. Subsequently, the U-Net model is trained to map from the manually segmented image to the mask (manually segmented image). This approach was employed to explore whether it would lead to a more robust model and what effect the incorporation of this data would have on the accuracy of the model.

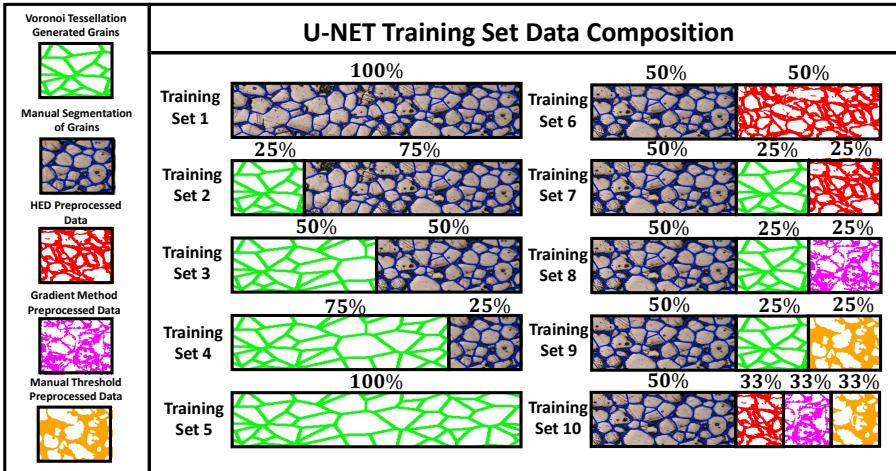


Fig. 9 A visual grouping of the various sets of data used for training neural networks in this work. The training sets are composed of either manually segmented data, artificially generated data, or preprocessed data through the HED method, Gradient Method, or Manual Threshold method

6 Results and Discussions

In Section 3: Grain Measurement and Accuracy Quantification, two methodologies were discussed to assess the accuracy of the results. The first approach utilizes a commonly used metric for segmentation accuracy called the Dice score. This quantification method provides a measure of how well the predicted segmentation masks align with the ground truth masks. The second approach focuses on evaluating the accuracy of dimensional grain data measurements. This involves comparing the measured grain properties obtained from the predicted masks with the ground truth measurements. Both of these evaluation methods were implemented and will be discussed in detail in the subsequent sections. In addition to the quantitative results, the qualitative findings of this work will also be presented and discussed.

6.1 Dice Score

The Dice score is a widely used metric in the machine learning and computer vision communities to assess the overlap between two sets of data, making it particularly suitable for evaluating image segmentation methods. In this work,

the Dice score was employed to compare the performance of different methods on the various training set compositions depicted in Figure 9. The Dice score ranges from 0 to 1, where a score of 1 indicates a perfect overlap between the predicted segmentation and the ground truth. In this study, the focus was on accurately classifying pixels as grain boundaries, and thus the Dice score was calculated based on the ability of the models to accurately identify grain boundaries.

The Dice score results are presented in two graphs in this section. Figure 10 displays the Dice scores for all the traditional methods described in Section 2. Figure 11 depicts the Dice scores for the different U-NET training compositions utilized in this work. It is important to note that the manually segmented image is considered the ground truth in this study, which is why it achieves a perfect score in both figures. The Dice score graphs provide valuable insights into the performance of the segmentation methods and their ability to accurately identify grain boundaries.

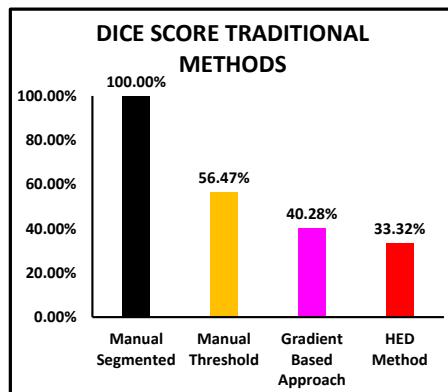


Fig. 10 Bar graph for the Dice scores for the 3 traditional methods of grain segmentation: Manual Thresholding, Gradient Based Thresholding, and HED method. The ground truth for this measurement is the Manually Segmented images.

In the graph in Figure 10, the traditional segmentation methods displayed relatively low Dice scores, with the HED method performing unexpectedly poorly. However, these methods can still be valuable for generating noise to enhance the robustness of machine learning models, as demonstrated in this study. The lower scores emphasize the challenges and limitations of manual and rule-based approaches in accurately capturing complex grain boundary features. The integration of machine learning techniques, like U-NET, offers improved segmentation by capturing both local and global features more effectively.

Figure 11 illustrates that the dice scores for the U-NET models are consistently around 90%. Surprisingly, the inclusion of fully artificially generated data in training sets 2, 3, and 4 did not adversely affect the scores. Even when using solely artificially generated data in training set 5, the dice score only

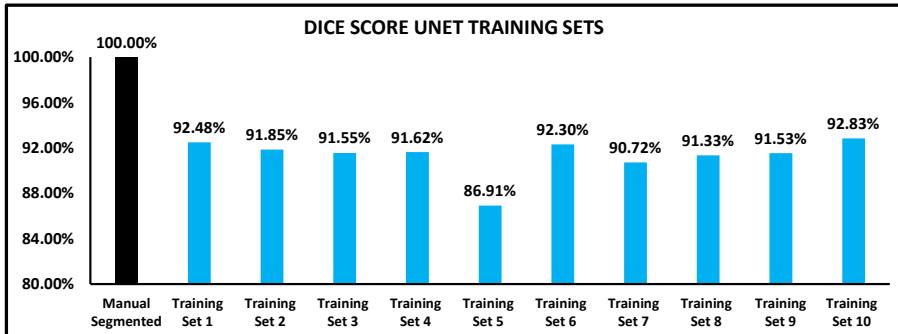


Fig. 11 Bar graph for the Dice scores for the 10 training sets of data used to train the UNET in this work. The ground truth for this measurement is the Manually Segmented images.

dropped to 87%, which is still remarkable considering the absence of real grain images in the neural network's training.

Furthermore, the dice scores for training sets 6 through 10 demonstrate that incorporating noise generated from the traditional methods does not negatively impact the scores. This suggests that such noise can be utilized to train more robust models that are resistant to the effects of traditional segmentation methods. Notably, training set 10, which included noise from all three traditional segmentation methods, achieved the highest dice score.

6.2 Grain Data

The accuracy of grain geometry measurements is of greater interest to the materials science community than the accuracy of pixel classification. To evaluate the accuracy of the methods used in this paper, the standard error equation (Equation 6) is employed. The error values are depicted in Figure 12, where each axis represents a specific grain measurement, such as circularity, average maximum diameter, average X-diameter, average Y-diameter, grain intercept method in the X-direction, grain intercept method in the Y-direction, average grain area, and total grains.

The error values are calculated by comparing the average values obtained from the grain measurements of the different methods to the corresponding average values from the manually segmented images, which are regarded as the ground truth. The colorbar in Figure 12 indicates whether the estimated values overestimate (blue) or underestimate (red) the true values. This analysis provides insights into the accuracy of the grain geometry measurements, highlighting the extent to which the methods over or under predict the values.

$$\text{Error Percent} = \frac{\text{Measured Value} - \text{True Value}}{\text{Measured Value}} \quad (6)$$

Based on the values in Figure 12, the traditional methods (manual threshold, Gradient Approach, and HED Method) exhibit higher error percentages

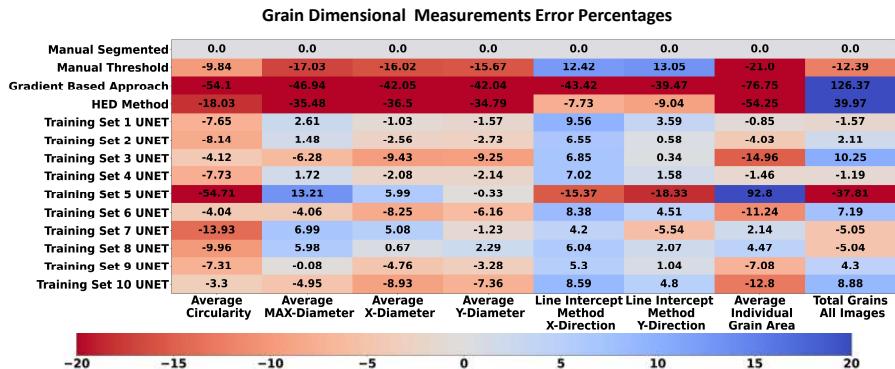


Fig. 12 A color map superimposed on the numerical values for the various grain dimensional measurement errors. The measurement error is provided for the 3 traditional segmentation methods and all ten of the UNETs trained on differing training sets. The training set compositions are given in Figure 9. The manually segmented images are regarded as the true value for this error measurement.

for all grain measurements compared to the other methods. Among the traditional methods, the manual threshold method shows relatively lower errors. However, both the Gradient Approach and HED Method significantly overestimate the total number of grains in the images, leading to skewed values in other measurements. The average maximum diameter, average X-diameter, and average Y-diameter are also affected by overestimation, although the manual threshold method has lower errors in these measurements compared to the other traditional methods.

In contrast, the UNET training set variations generally exhibit lower errors. Training set 5, trained entirely on artificial data, shows the highest error with an underprediction of grain count by approximately 38%. This is attributed to the presence of incomplete boundaries and small neck regions connecting grains, as evidenced by the lower circularity values. Training set 1 performs well across all measurements, and the inclusion of artificial grains in sets 2, 3, and 4 slightly increases the error percentages. However, these sets outperform the traditional methods and some other training sets. Training set 10, which includes noise from the traditional methods, also demonstrates good performance, indicating the beneficial impact of incorporating diverse noise sources in the training process. Overall, the UNET-based methods outperform the traditional methods, with different training set compositions yielding varying levels of accuracy in grain measurements.

6.3 Qualitative Results

The qualitative results provide a closer examination of the approach by presenting individual image results in grain boundary segmentation. This allows for a detailed analysis of the segmentation performance on a case-by-case basis.

UNET TRAINING SET 3: PREDICTION EXAMPLES

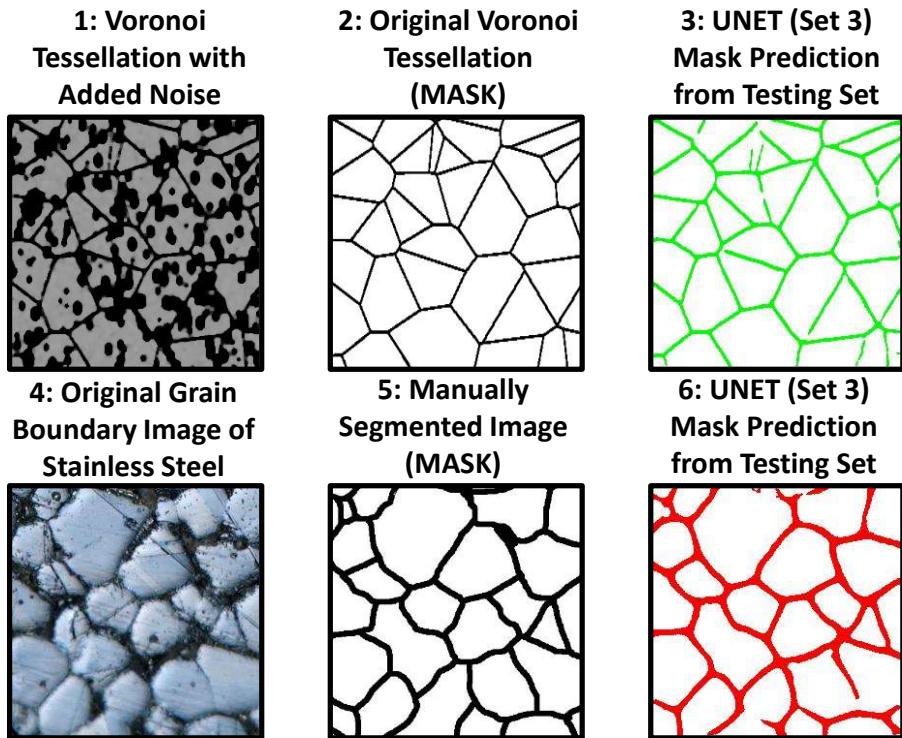


Fig. 13 Qualitative results from the UNET trained on training set 3. This training set is composed entirely of artificially generated data. (1) An artificial grain generated from the Voronoi Tessellation method. (2) The mask from image 1. (3) Prediction of the mask from UNET trained on data set 3. (4) Real grain image. (5) Mask of the manual segmentation of image 4, complete by a human. (6) UNET training set 3 prediction of the mask of image 4.

Figure 13 presents example images from the UNET trained on data set 3, which comprised 50% real grains with manual segmentations and 50% artificially generated grains using the Voronoi tessellation method. Image 3 shows the UNET’s prediction of the original mask (image 2) based on the input image (image 1) from the testing set. The UNET successfully predicted the mask despite not seeing this specific image during training.

Furthermore, the same UNET trained on data set 3 demonstrated its ability to accurately predict grain boundaries in a real grain image. Image 4 represents a real grain image, image 5 shows the corresponding manually drawn segmentations, and image 6 displays the UNET’s prediction of the grain boundaries. The high degree of similarity between images 5 and 6 highlights the exceptional performance of neural networks in grain boundary segmentation tasks. These results showcase the superior capabilities of the UNET model trained

on data set 3 in accurately predicting grain boundaries, even when presented with unseen images during training.

UNET TRAINING SET 5: PREDICTION EXAMPLES

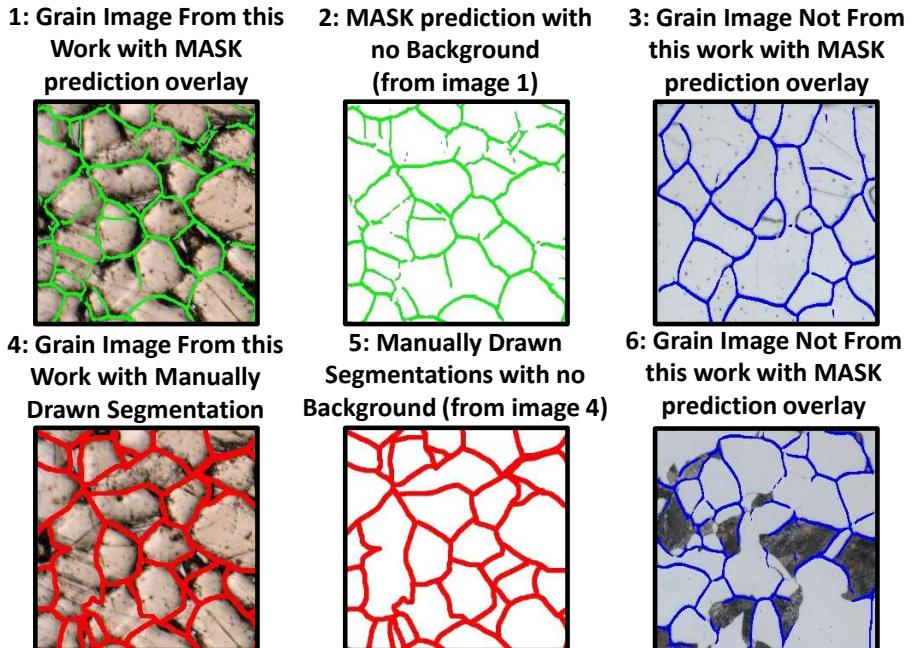


Fig. 14 Qualitative results from the UNET trained on training set 5. This training set is composed of 50% manually segmented real grains and 50% artificially generated grains. (1) Real grain with a prediction of mask from UNET overlaid. (2) The mask from image 1 with background removed. (3) Image of grains this UNET did not train with prediction of mask overlaid (4) Image 1 with manually segmentations overlaid (5) The mask from image 4 with background removed. (6) Image of grains this UNET did not train with prediction of mask overlaid

In Figure 14, the results of a UNET trained exclusively on artificially generated data (training set 5) are presented. This UNET model was never exposed to actual grain boundary images but was solely trained on data generated using the Voronoi tessellation grain image generation method proposed in this work.

Image 1 in Figure 14 displays a real grain image of stainless steel, with the predicted grain boundary mask overlaid in green. Despite not being trained on this specific data, the UNET was able to provide a reasonably accurate estimation of the grain boundaries. For comparison, image 4 shows the manually drawn grain boundaries for the same grain image. Images 2 and 5 show the masks without the grain boundaries, provided for reference. Although the predictions from UNET trained on data set 5 are not perfect, they demonstrate

that the model successfully avoided classifying voids and polishing scratches as grain boundaries.

Images 3 and 6 showcase stainless steel images produced through injection molding, which were not used for training and were not captured specifically for this work. These images serve to demonstrate the potential of using artificial data for training more robust models. The UNET trained on data set 5 achieved reasonably accurate grain boundary predictions for these images as well, even in the cases where some grains appeared darker, like in image 6.

6.4 Summary of Results

The summary of results from this work will be broken down into the different approaches to keep the statements organized and keep the clarity of each statement intact.

The Traditional Methods: The traditional methods (manual thresholding, gradient-based approach, HED method) require parameter tuning and are not universally applicable. However, they can be used to generate noise for training more robust models.

UNET Training Set 1: Training set 1 performed well on both the dice score and the actual grain measurements, as should be expected.

UNET Training Sets 2-4: Training Sets 2-4, which included artificially generated grains, maintained high scores in dice score and grain measurement accuracy. Artificial data can effectively increase training data without compromising accuracy.

UNET Training Set 5: This set is composed entirely of artificially generated data, and performed reasonably well, surpassing the traditional methods. The voronoi grain generation method can be used to supplement training data for grain segmentation.

UNET Training Set 6-10: Training Sets 6-10, incorporating noise from traditional methods, achieved high scores in dice score and grain measurement accuracy. Training Set 10, which included noise from all three traditional methods, performed the best. Incorporating noise from these methods enhances model robustness.

Planimetric and Line Intercept Method: Planimetric analysis yielded more accurate grain measurements compared to the line intercept method. Deep learning-based image segmentation can provide suitable images for planimetric analysis.

In summary, the deep learning approach using UNET models demonstrated superior performance compared to traditional methods for grain boundary segmentation. Incorporating artificially generated data and noise from traditional methods can enhance model performance and robustness. Additionally, planimetric analysis provides more accurate grain measurements than the line intercept method.

7 Conclusion

This study has made significant contributions to the field of grain boundary segmentation by introducing a novel method for generating artificial data and demonstrating its potential for improving segmentation accuracy. The generated artificial data incorporated various realistic defects, such as pores, polishing scratches, and small grain fragments, which proved beneficial in training the UNET algorithm to better identify and avoid these imperfections when predicting grain boundaries. This approach addresses a critical challenge faced in grain measurement science, where the availability of annotated real-world data for training segmentation algorithms is often limited.

However, one notable limitation of the UNET trained on the artificial data was the occasional failure to completely enclose the grains in the predicted boundaries, leading to lower dice scores and less accurate grain measurements. Strategies such as refining the generation algorithm, incorporating additional features to guide grain boundary closure, or exploring alternative network architectures may help overcome this limitation and enhance the accuracy of grain boundary segmentation. Despite this limitation, the combination of real and artificial training data showcased promising results, indicating that a diverse training set comprising both real and artificial data can effectively improve grain boundary segmentation. Moreover, the inclusion of noise generated from traditional methods did not compromise segmentation accuracy, suggesting that these noisy data sources can be leveraged to enhance the robustness of the trained models. This finding opens up avenues for future research, where the integration of various data sources, including artificial, real, and noisy data, can lead to the development of more reliable and adaptable machine learning models for grain boundary analysis.

The artificial data generation method presented in this study offers valuable insights and possibilities for advancing grain measurement science. The developed artificial grain generation method can be adapted for different grain structures and offers a quantitative framework for analyzing grain geometry by utilizing the input properties of the Voronoi tessellation pattern. This adaptability allows researchers to explore diverse grain patterns and investigate their influence on material properties. By systematically varying the input properties, valuable insights can be gained into the relationship between grain morphology and material behavior. This approach has the potential to enhance our understanding of structure-property relationships, improve grain boundary segmentation algorithms, and facilitate the design of materials with tailored microstructures.

Acknowledgements

We thank the Siemens Energy Center and AM center in North Carolina for Providing the Samples Analyzed in this work

References

- [1] Do Lee, C.: Effect of grain size on the tensile properties of magnesium alloy. *Materials Science and Engineering: A* **459**(1-2), 355–360 (2007)
- [2] Schempp, P., Cross, C., Häcker, R., Pittner, A., Rethmeier, M.: Influence of grain size on mechanical properties of aluminium gta weld metal. *Welding in the World* **57**(3), 293–304 (2013)
- [3] Wang, N., Wang, Z., Aust, K., Erb, U.: Effect of grain size on mechanical properties of nanocrystalline materials. *Acta Metallurgica et Materialia* **43**(2), 519–528 (1995)
- [4] Bai, Y., Wagner, G., Williams, C.B.: Effect of particle size distribution on powder packing and sintering in binder jetting additive manufacturing of metals. *Journal of Manufacturing Science and Engineering* **139**(8) (2017)
- [5] Voyiadjis, G., Abed, F.: Transient localizations in metals using microstructure-based yield surfaces. *Modelling and Simulation in Materials Science and Engineering* **15**(1), 83 (2006)
- [6] Heo, S., Yun, J., Oh, K., Han, K.: Influence of particle size and shape on electrical and mechanical properties of graphite reinforced conductive polymer composites for the bipolar plate of pem fuel cells. *Advanced composite materials* **15**(1), 115–126 (2006)
- [7] Uddin, S.M., Mahmud, T., Wolf, C., Glanz, C., Kolaric, I., Volkmer, C., Höller, H., Wienecke, U., Roth, S., Fecht, H.-J.: Effect of size and shape of metal particles to improve hardness and electrical properties of carbon nanotube reinforced copper and copper alloy composites. *Composites Science and Technology* **70**(16), 2253–2257 (2010)
- [8] Ali, H., Stein, Z., Fouliard, Q., Ebrahimi, H., Warren, P., Raghavan, S., Ghosh, R.: Computational model of mechano-electrochemical effect of aluminum alloys corrosion. *Journal of Engineering for Gas Turbines and Power* **144**(4), 041004 (2022)
- [9] Adam, K., Zöllner, D., Field, D.P.: 3d microstructural evolution of primary recrystallization and grain growth in cold rolled single-phase aluminum alloys. *Modelling and Simulation in Materials Science and Engineering* **26**(3), 035011 (2018)
- [10] Chen, F., Cui, Z., Liu, J., Zhang, X., Chen, W.: Modeling and simulation on dynamic recrystallization of 30cr2ni4mov rotor steel using the cellular automaton method. *Modelling and Simulation in Materials Science and Engineering* **17**(7), 075015 (2009)

- [11] Herriott, C., Li, X., Kouraytem, N., Tari, V., Tan, W., Anglin, B., Rollett, A.D., Spear, A.D.: A multi-scale, multi-physics modeling framework to predict spatial variation of properties in additive-manufactured metals. *Modelling and Simulation in Materials Science and Engineering* **27**(2), 025009 (2019)
- [12] Raju, N., Warren, P., Subramanian, R., Ghosh, R., Fernandez, E., Raghavan, S., Kapat, J.: Sintering behaviour of 3d printed 17-4ph stainless steel. In: *Turbo Expo: Power for Land, Sea, and Air*, vol. 86052, pp. 007–17028 (2022). American Society of Mechanical Engineers
- [13] Van Der Giessen, E., Schultz, P.A., Bertin, N., Bulatov, V.V., Cai, W., Csányi, G., Foiles, S.M., Geers, M.G., González, C., Hütter, M., *et al.*: Roadmap on multiscale materials modeling. *Modelling and Simulation in Materials Science and Engineering* **28**(4), 043001 (2020)
- [14] Yan, F., Xiong, W., Faierson, E.J.: Grain structure control of additively manufactured metallic materials. *Materials* **10**(11), 1260 (2017)
- [15] Lin, J., Balint, D., Pietrzyk, M.: *Microstructure Evolution in Metal Forming Processes*. Elsevier, USA (2012)
- [16] Steinbach, I.: Phase-field models in materials science. *Modelling and simulation in materials science and engineering* **17**(7), 073001 (2009)
- [17] Bandyopadhyay, A., Zhang, Y., Bose, S.: Recent developments in metal additive manufacturing. *Current opinion in chemical engineering* **28**, 96–104 (2020)
- [18] Warren, P., Raju, N., Krsmanovic, M., Ebrahimi, H., Kapat, J., Subramanian, R., Ghosh, R.: Shrinkage prediction using machine learning for additively manufactured ceramic and metallic components for gas turbine applications. In: *Turbo Expo: Power for Land, Sea, and Air*, vol. 85987, pp. 002–05023 (2022). American Society of Mechanical Engineers
- [19] Raju, N., Warren, P., Subramanian, R., Ghosh, R., Raghavan, S., Fernandez, E., Kapat, J.: Material properties of 17-4ph stainless steel fabricated by atomic diffusion additive manufacturing (adam). In: *2021 International Solid Freeform Fabrication Symposium* (2021). University of Texas at Austin
- [20] Vayre, B., Vignat, F., Villeneuve, F.: Metallic additive manufacturing: state-of-the-art review and prospects. *Mechanics & Industry* **13**(2), 89–96 (2012)
- [21] Yakout, M., Elbestawi, M., Veldhuis, S.C.: A review of metal additive manufacturing technologies. *Solid State Phenomena* **278**, 1–14 (2018)

26 *Machine Learning Grain Segmentation*

- [22] Frazier, W.E.: Metal additive manufacturing: a review. *Journal of Materials Engineering and performance* **23**, 1917–1928 (2014)
- [23] Tan, J.H.K., Sing, S.L., Yeong, W.Y.: Microstructure modelling for metallic additive manufacturing: A review. *Virtual and Physical Prototyping* **15**(1), 87–105 (2020)
- [24] Perera, R., Guzzetti, D., Agrawal, V.: Optimized and autonomous machine learning framework for characterizing pores, particles, grains and grain boundaries in microstructural images. *Computational Materials Science* **196**, 110524 (2021)
- [25] Heyn, E.: *Physical Metallography*. Wiley, USA (1925)
- [26] Subcommittee, A.: Standard Test Methods for Determining Average Grain Size. ASTM International, USA (1996)
- [27] Warren, P.: Artificial Grains and Real Grains. <https://www.kaggle.com/datasets/peterwarren/voronoi-artificial-grains-gen> (2023)
- [28] Warren, P.: ExONE Stainless Steel 316L Grains 500X. <https://www.kaggle.com/datasets/peterwarren/exone-stainless-steel-316l-grains-500x> (2023)
- [29] Bradski, G.: The OpenCV Library. Dr. Dobb's Journal of Software Tools (2000)
- [30] Warren, P.: SinteringTrajectory Github Repository. <https://github.com/Peterwarren623/GrainBoundaryDetection> (2022 (accessed Dec 16, 2022))
- [31] Canny, J.: A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence* (6), 679–698 (1986)
- [32] Xie, S., Tu, Z.: Holistically-nested edge detection. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1395–1403 (2015)
- [33] O'Shea, K., Nash, R.: An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458 (2015)
- [34] Albawi, S., Mohammed, T.A., Al-Zawi, S.: Understanding of a convolutional neural network. In: *2017 International Conference on Engineering and Technology (ICET)*, pp. 1–6 (2017). Ieee
- [35] Warren, P., Ali, H., Ebrahimi, H., Ghosh, R.: Rapid defect detection and classification in images using convolutional neural networks. In: *Turbo Expo: Power for Land, Sea, and Air*, vol. 84966, pp. 004–05013 (2021). American Society of Mechanical Engineers

- [36] Voronoi, G.: Nouvelles applications des paramètres continus à la théorie des formes quadratiques. deuxième mémoire. recherches sur les parallélopèdres primitifs. *Journal für die reine und angewandte Mathematik* (Crelles Journal) **1908**(134), 198–287 (1908)
- [37] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-assisted Intervention, pp. 234–241 (2015). Springer