

ORIGINAL PAPER

Open Access



Toward automated microstructure characterization of stainless steels through machine learning-based analysis of replication micrographs

Hamza Ghauri¹, Reza Tafreshi^{1,2} and Bilal Mansoor^{1,2,3*}

Abstract

Machine learning-driven automated replication micrographs analysis makes possible rapid and unbiased damage assessment of in-service steel components. Although micrographs captured by scanning electron microscopy (SEM) have been analyzed at depth using machine learning, there is no literature available on the technique being attempted on optical replication micrographs. This paper presents a machine-learning approach to segment and quantify carbide precipitates in thermally exposed HP40-Nb stainless-steel microstructures from batches of low-resolution optical images obtained by replication metallography. A dataset of nine micrographs was used to develop a random forest classification model to segment precipitates within the matrix (intragranular) and at grain boundaries (intergranular). The micrographs were preprocessed using background subtraction, denoising, and sharpening to improve quality. The method achieves high segmentation accuracy (91% intergranular, 97% intragranular) compared to human expert classification. Furthermore, segmented micrographs were quantified to obtain carbide size, shape, and density distribution. The correlations in the quantified data aligned with expected carbide evolution mechanisms. Results from this study are promising but necessitate validation of the method on a larger dataset representative of evolution of thermal degradation in steel, given that characterization of the evolution of microstructure components, such as precipitates, applies to broad applications across diverse alloy systems, particularly in extreme service.

Keywords Microstructure analysis, Replication metallography, Machine learning, Automated segmentation

Introduction

Microscopy is commonly used to record high-magnification images and a planar view of features of a material's microstructure. Typically, alloy samples are polished,

etched, and imaged using optical, scanning electron (SEM) or transmission electron microscopes (TEM) in a laboratory setting to obtain microstructural images, called micrographs. In the case of metals, metallography techniques are used to obtain and analyze micrographs to develop an overall sense of the microstructural features, including precipitate shape and size distribution, grain sizes, and defects, such as voids and cracks. These features represent important material properties and behavior, such as strength and ductility.

Replication metallography is a specialized nondestructive technique for obtaining in situ micrographs in industrial environments (Marder 1989). The ability to

*Correspondence:

Bilal Mansoor
bilal.mansoor@tamu.edu

¹ Mechanical Engineering Program, Texas A&M University at Qatar, Doha, Qatar

² Department of Mechanical Engineering, Texas A&M University, College Station, TX, USA

³ Department of Materials Science and Engineering, Texas A&M University, College Station, TX, USA



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

generate micrographs without part removal or modification is crucial to the process industry, where the length of maintenance downtime poses significant financial consequences. Replication micrographs have been conventionally used to assess corrosion, creep damage, and crack formation in metals subjected to high temperature–pressure conditions (Jana 1995). Replicas are created on-site and imaged with an optical microscope for subsequent evaluation. To prepare a replica, the material surface is cleared of any corrosion and oxidation products, polished, and etched before the application of a polymeric film that captures a microstructural stamp. The polymeric film is then removed and secured onto a glass slide before being imaged by an optical microscope for subsequent evaluation (E1351-01 2020). Features seen in optical micrographs are a product of depth contrast created by etching, which are captured by polymer casts as height contours. The lower reflectivity of polymer surfaces along with inversion of depth features leads to a loss of contrast in optical images taken from replicas. This results in the addition of noise/artifacts and an overall reduction in the contrast and resolution of feature boundaries, as noted in the ASTM E1351-01 standard (E1351-01 2020). Therefore, it is harder to distinguish between precipitates, grain boundaries, and defects such as micro voids and cracks in optical images of replicas.

Material science experts usually evaluate the micrographs collected from the field to assess microstructural evolution, damage, and defects incurred in operation. However, such manual evaluations are laborious, prone to subjectivity, and require significant expertise, factors that can introduce uncertainty and bias into the analysis (Azimi et al. 2018). Martin et al. (2022) reported the industry standard methods for replica microstructure quantification to be (i) interpolation from previous analysis of similar microstructures and (ii) point counting estimation where an arbitrary number of equidistant points is plotted on a micrograph and the phase area fractions are calculated by the fraction of points falling in each phase.

The development of computer vision techniques has made microstructure analysis quantitative, repeatable, and consistent. Microstructures are seen as matrices of gray value pixel data that can be subsequently manipulated and quantified in a traceable manner. Once labeled, the grayscale values are used to train supervised classification models and segment microstructures according to the features labeled in the training data (Holm et al. 2020). Numerous studies have demonstrated the capability of machine learning (ML) models to correctly identify key microstructural features in micrographs, such as grain sizes, precipitates, texture, and defects (Holm et al. 2020; Perera et al. 2021; Baskaran et al. 2020). While deep

learning methods such as convolutional neural networks (CNNs) are more robust for computer vision applications (Kordijazi et al. 2021), they are also relatively slower to train and execute and carry significant computational costs. DeCost et al. (2019) applied a pixel-wise CNN to segment a labeled open-source steel microstructure dataset obtained using scanning electron microscopy (SEM). They achieved an average precision of 0.96 and a sensitivity of 0.92 for classifying intergranular networked carbides, albeit their model had a higher rate of misclassification for damaged networked carbides. On the other hand, machine learning (ML) methods such as random forests and support vector machines (Shmilovici 2005) are less complex yet provide comparable results when the dataset is small, and the microstructures have distinguishable edges between the features being segmented (Lai et al. 2019).

Although SEM images are preferred for automated segmentation due to their clarity, high resolution, and contrast, Bulgarevich et al. (2018) demonstrated the ability of a supervised random forest model to segment the different phases found within a typical steel microstructure from optical micrographs. Papa et al. (2013) compared the performance of the support vector machine, optimum-pat forest (OPF), and Bayesian models to identify carbon-rich precipitates in optical images obtained from service-exposed industrial cast iron samples and got the highest universal image quality index of 0.85 for the OPF model. However, replication micrographs are lower in quality than those captured directly on an optical microscope due to the increased presence of artifacts, background hue that affects contrast, blur/low resolution, and excessive noise (Jana 1995). To the best of our knowledge, a ML-based segmentation and quantification approach to analyze microstructure obtained by replication has not been done.

In materials science, the distribution, morphology, orientation, and composition of microstructural features such as grains, precipitates, and defects are all relevant to material properties and are considered during microstructural analysis. Due to the limitations of replication metallography, the matrix and precipitates are the only features that can be discerned with confidence (Jana 1995). Precipitates are one of the most important microstructural features of interest, as their morphology and distribution can impact material properties such as tensile strength (Bonaccorsi et al. 2014). In the case of HP40-Nb steel, a common material used in high-temperature applications in the oil and gas industry, the carbon-rich precipitates, known as carbides, and their transformation with exposure to extreme conditions such as high temperature play a defining role in the service life of components. Carbide transformations are used as markers

of material creep and help material scientists gauge the extent of the remaining design life of a component. Other alloy systems exhibit similar precipitate transformations when subjected to extreme environments. In nickel alloy 263 steam turbines, thermal exposure leads to the formation of the η precipitate, and an increasing precipitate density correlates to a shorter creep life (Detrois et al. 2021). By efficiently quantifying and tracking the microstructural features such as precipitates, the adoption of a predictive inspection–maintenance–replacement philosophy can be facilitated.

This study develops and implements a ML segmentation model to identify distinct carbide phases in low-quality replication metallographs of service-exposed HP40-Nb steel and subsequently quantify the precipitate shape and size features. Image preprocessing methods are utilized to improve the image quality of the microstructures. Then, a random forest ML model is used to identify and segment the intragranular and boundary precipitates. Segmentation performance is benchmarked against manual labeling to gauge performance. The segmented output images are then analyzed to obtain the carbide counts, density, shape, and size distribution for every carbide instance identified via segmentation. The entire process is automated and consolidated for batch image processing, simplifying the workflow, and increasing the repeatability of results, compared to manual replication microstructure analysis. By gaining rapid insights into the microstructural evolution process under different service conditions, we can optimize maintenance strategies, anticipate potential failures, and proactively plan for component replacements across a wide range of scenarios and diverse industrial sectors.

Methods

This section details the methodology employed to obtain and preprocess the microstructural images for developing a ML model and to identify and quantify the presence of the carbides within micrographs.

Materials and data description

Microstructures of nine HP40-Nb stainless steel samples with distinct thermal histories were replicated and imaged using optical microscopy. One parent microstructure image was obtained for each sample at $50\times$ magnification resulting in 1306×623 pixels. The resultant measurement ratio for all images was $0.64\ \mu\text{m}/\text{pixel}$.

The microstructure of an HP40-Nb steel contains two morphological distributions of carbide precipitates: (i) large intergranular (boundary) precipitates along the grain boundaries (interface between two or more crystallites that a metal is composed of) and (ii) relatively

smaller intragranular precipitates within the matrix (continuous, base phase present in the material) (Vaché et al. 2020). Figure 1a shows the two carbide types labeled on an optical image of a HP40-Nb replica microstructure. The distribution of precipitates depends on carbide transformations that occur as the material incurs damage under service conditions. After a short duration of exposure to high-temperature service conditions, the HP40-Nb microstructure shows an extensive, thinly distributed carbide network along the grain boundaries, while the precipitates within the matrix are expected to be few and small (Shi and Lippold 2008). After extensive thermal exposure, the carbide network breaks down, resulting in disconnected, thicker precipitates on the grain boundaries and the appearance of more intragranular precipitates that eventually agglomerate and grow (Shi and Lippold 2008). The precipitates are easier to observe in SEM micrographs since a better magnification and a clearer image can be produced. To visualize the complexity of identifying precipitates in a replication micrograph as compared to an SEM micrograph, an SEM micrograph obtained from an ex-service HP40-Nb steel exhibiting both the intergranular (grain boundary) and intragranular (within the matrix) carbides is shown in Fig. 1b.

Image preprocessing

ImageJ (Schindelin et al. 2012), an open-source image processing software, was used to preprocess the raw micrographs. The original images were in an RGB format and had a blue/green hue due to the choice of replica material and the lighting conditions during imaging. Since the hue is inconsequential to the microstructure, all the images were converted to an 8-bit grayscale format to save storage and processing time.

The images contained a significant pixel intensity gradient in the background, which was an artifact of replication imaging and unrelated to the microstructure. Therefore, it was necessary to remove the background gradient and ensure that any variation seen in pixel intensities was due to microstructural features. Background subtraction was performed using the rolling ball method (Sternberg 1983), which creates a topographical map out of the pixel intensities and rolls a ball of a defined pixel radius over it, picking any spatial variations in the grayscale values and subtracting them from the original images. Figure 2 shows (a) an image before the background removal, (b) the background gradient removed from the image, and (c) the image after the background removal.

Due to the nature of replication and optical imaging, significant noise was observed in the images that could negatively affect the performance of an ML segmentation model. The images were denoised (Fig. 2d) using the

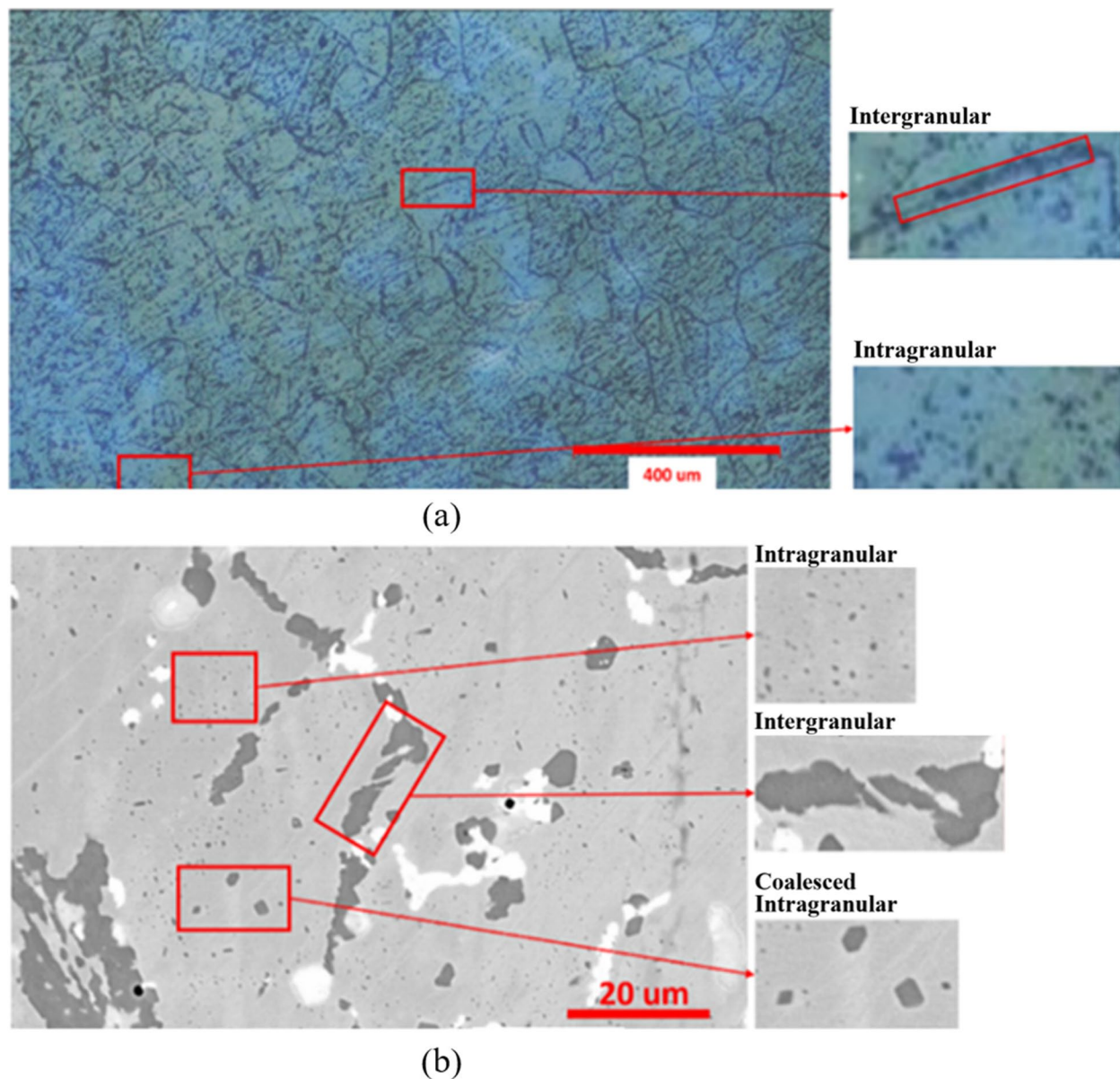


Fig. 1 Distribution of intergranular (boundary) and intragranular carbide phases in **a** an optical image of an aged HP40-Nb replication microstructure and **b** aged HP40-Nb SEM micrograph

nonlocal means method (Buades et al. 2011), which registers the pixel intensities across the image and compares neighbors of similar pixels to identify and subtract noise. The denoising also blurs object boundaries which then need to be restored. The image was sharpened (Fig. 2e) using the unsharp mask method that applies a low-pass filter on the pixel intensities, creating a secondary image that captures the blur from the parent image (Polesel et al. 2000). This blur is subtracted from the original image, giving a sharpened output. Lastly, the contrast was enhanced (Fig. 2f) using normalized contrast stretching,

which normalizes the pixel values in the image between 0 and 255, utilizing the full range of pixel intensities available (Jain 1988).

Data preparation, ML modeling, and implementation

While numerous ML models such as support vector machine (Shmilovici 2005), k-nearest neighbors (Guo et al. 2003), and voted perceptron have been used for pixel classification problems, the random forest (RF) method was more suited to the carbide identification study due to its ability to handle an input comprised of

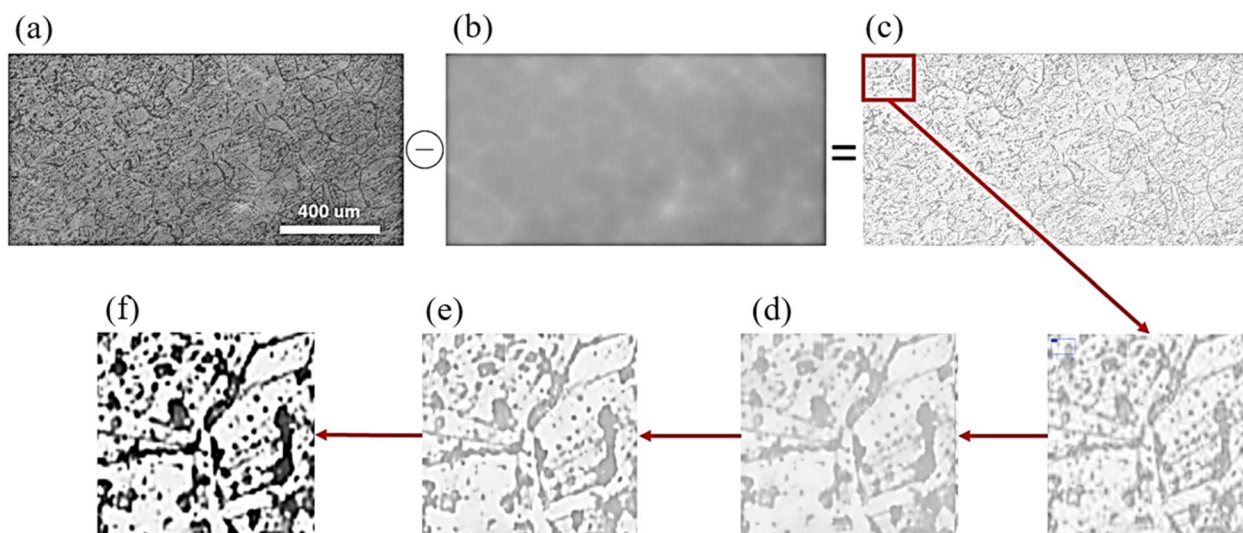


Fig. 2 Image preprocessing schematic diagram: **a** original gray scale, **b** rolling ball background, **c** post background removal, **d** nonlocal means denoising, **e** unsharp mask sharpening, and **f** contrast stretching

a large number of features for small datasets. A low carbide-to-matrix ratio in the micrographs results in class-imbalanced data. This is partially handled by random forests, an ensemble method in which each tree is trained independently, minimizing the effect of trees that may be biased toward the majority class. Figure 3a shows a schematic layout of an RF ensemble with n decision trees, while Fig. 3b shows the overall process workflow used in this study, including data split and ML. RF consists

of multiple decision trees where the data used to train each decision tree is randomly sampled from the training labels using the random bootstrapping process, which accommodates smaller training datasets by using statistical resampling.

Data split

Nine parent micrographs were each split into 4, resulting in a set of 36 images with 653×313 ($1306/2 \times 623/2$)

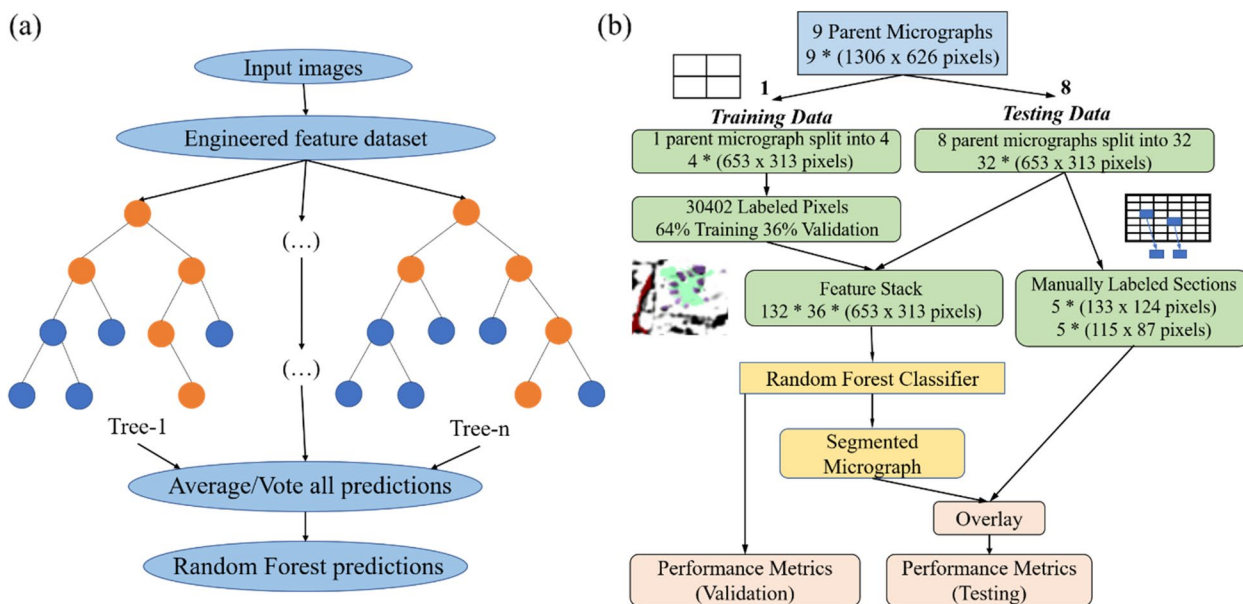


Fig. 3 Process flowchart detailing **a** a typical random forest with n decision trees and **b** the workflow from parent images to machine learning outputs

pixels. One of the 9 parent images (4 out of $9 \times 4 = 36$ split images) was specified for training and validation. The remaining eight parent images (32 out of 36 split images) were specified as testing data.

- Data labeling for training and validation: The microstructural features present in the micrographs were divided into three classes: (i) intergranular (grain boundary) carbides, (ii) intragranular (peppered within the matrix) carbides, and (iii) the material matrix. The image specified for training and validation data was partially labeled by hand, in which 6017 pixels were labeled as boundary carbides, 3261 pixels as intragranular carbides, and 21,124 pixels as matrix. A sample of labeled input prepared for training and validation is shown in Fig. 4a.
- Data labeling for testing: Additional data were labeled to test the RF model's segmentation performance for both intergranular and intragranular carbides. The sizes of the labeled image sections were selected between the range of 50×50 and 150×150 pixels randomly (133×124 pixels for intergranular carbides and 115×87 pixels for intragranular carbides). Five sections for each carbide were then selected at random locations from micrographs in the testing dataset and labeled manually by material science experts.

- Feature generation and model training: Mathematical transformations, such as Laplacian, Gaussian, Hessian, Sobel membrane projection, Lipschitz, Kuwahara, and anisotropic diffusion, were applied to the dataset at multiple intensities to engineer a set of 132 (default software settings) features. RF classifier randomly chooses between the input feature stack and generates decision trees based on them. The default value of 100 decision trees was used. The output from all the decision trees for each pixel is then voted to classify the pixel as one of three defined classes, and a segmented version of the micrographs is produced as the output. The Trainable WEKA Segmentation Toolbox in FIJI (Arganda-Carreras et al. 2016), an open-source image analysis software, was used for generating and training the RF model.

Validation

The model's performance was validated in two ways: (i) by segmenting the training image and visually inspecting the segmentation accuracy versus training labels, as shown in Fig. 4b, and (ii) by calculating the out-of-bag error (OOBE) where the leftover training data (i.e., 36% of training and validation labels) from the RF bootstrapping

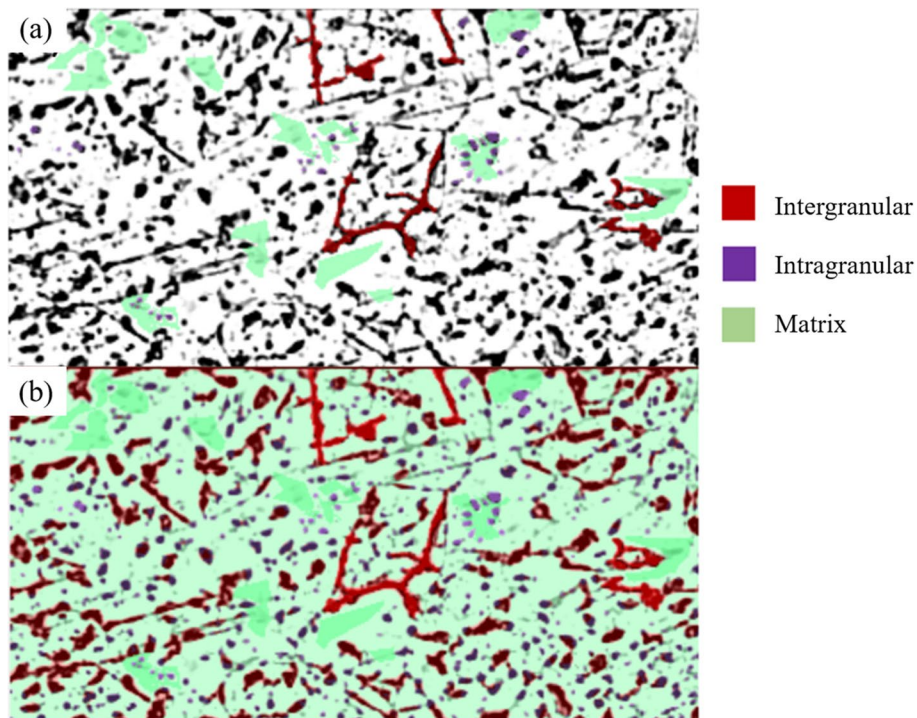


Fig. 4 **a** A sample image manually labeled for training (red, intergranular/boundary carbides; purple, intragranular/pepper carbides; green, matrix). **b** RF segmentation of training image used for visual validation

process are compared to the ML segmentation to compute validation error (Bulgarevich et al. 2018).

Testing

The model segmented all micrographs in the testing dataset. The segmented images were then threshold using RGB values for each class to obtain two separate slices per image, isolating the intergranular and intragranular carbides. To compute performance metrics for RF segmentation specific to each carbide class, the problem was reduced from a three-class classification to two binary-class classification problems. For intergranular carbides, the intragranular carbides and matrix together were considered as the negative class. For intragranular carbides, the intergranular carbides and matrix together were considered as the negative class.

Model performance (accuracy, precision, sensitivity, specificity, fallout, and F1 score) was evaluated by overlaying the image sections labeled for testing onto the corresponding regions of the intergranular and intragranular slices produced from the segmented testing images. The overlay method is widely used to determine the performance of an automated microstructure segmentation, as used by Stan et al. (2020) to test neural network segmentation of (i) X-ray tomography microstructures of an Al-Zn alloy obtained during solidification with an aim to distinguish the solid and liquid phases and (ii) optical images of an Sn-Pb alloy exhibiting dendritic structures with an aim to identify the dendrites and track their coarsening. Hwang et al. (2020) used the ground truth overlay to evaluate ML segmentation accuracy for identifying individual material phases in multiphase composite microstructures.

Carbide quantification

The quantification was performed using the extended particle analyzer tool from the FIJI BioVoxel plug-in (Brocher 2022). For both intergranular and intragranular carbides, the area fraction (%), average precipitate area (μm^2), circularity, solidity, aspect ratio, and Feret diameters (diameters along the principle X- and Y-axes, μm) for each identified carbide instance were calculated and saved as a spreadsheet. The mean morphological properties for both carbide classes were calculated for each micrograph and visualized. Correlations between quantified parameters were also plotted. The methodology applied to calculate the parameters from applied threshold image slices is available in FIJI/ImageJ documentation (Brocher 2022). The RF segmentation, thresholding, and quantification process were consolidated as a Java macro script in ImageJ, allowing for automated batch execution.

Results and discussion

The RF segmentation performance was evaluated, and the distribution of carbide properties within the dataset was explored and compared to the expected carbide evolution of thermally exposed HP40-Nb steel.

ML segmentation

The preprocessed images were the input to the trained RF model. The model classified the micrographs as intergranular carbides, intragranular boundary carbides, and matrix. The segmentation output was then converted into separate slices for each carbide class against a null background (negative class). Examples of segmented micrograph slices showing labeled intragranular and intergranular carbides are shown in Fig. 5a and b, respectively. The performance of the RF classification model trained on a labeled micrograph was gauged using validation and testing data. The out-of-bag validation error yielded a high accuracy of 98.7%.

Intergranular carbide segmentation

The ML segmentation performance analysis compared the carbide identification by the RF model with that of a group of domain experts. Figure 6 demonstrates one of the RF analyses conducted with respect to boundary carbides: (a) the original section, (b) the manually labeled image, (c) the original RF output, and (d) the intergranular carbide slice of the RF-labeled image against a null background (the negative class, which consists of intragranular carbides plus matrix pixels).

The overlay of the expert and ML-labeled images in Fig. 6e shows that the RF model correctly identifies the larger boundary carbides but misidentifies some of the smaller broken-up boundary carbides as intragranular carbides, leading to false negatives (pixels belonging to the intergranular carbide class but incorrectly identified as matrix/intragranular carbide by the model) in regions with smaller boundary carbides. False positives (pixels belonging to the negative class but incorrectly identified as intergranular carbide) are seen at the boundaries of the carbides, indicating that the model identified a slightly larger region as a carbide than the actual carbide while maintaining geometry.

The model performance metrics for intergranular carbide segmentation are listed in Table 1. An overall accuracy of 91% is achieved. Since the area occupied by the negative class (matrix plus intragranular carbides) is large, accuracy alone becomes less conclusive metric. Therefore, the positive prediction value (precision) of 80% and the true-positive rate (sensitivity) of 82%

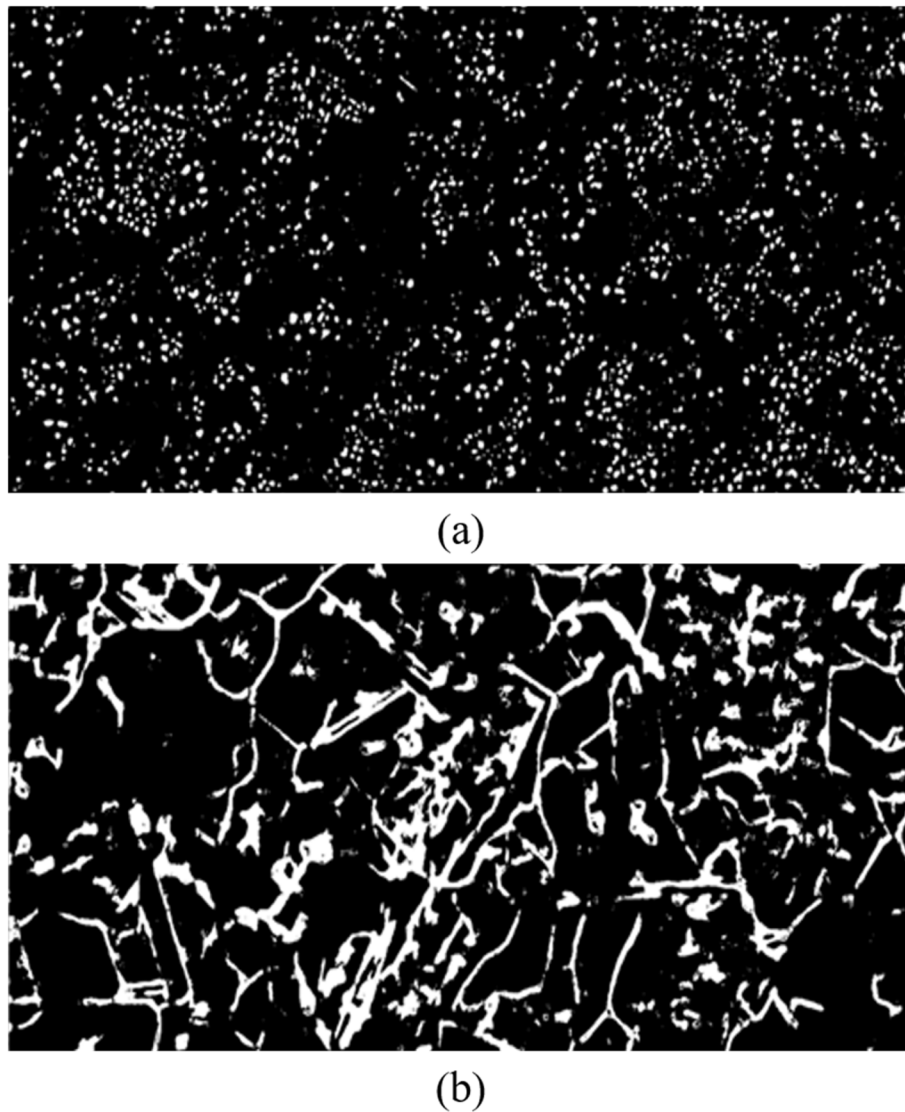


Fig. 5 Carbide slices obtained from a segmented micrograph showing **a** intragranular and **b** intergranular precipitates against a null background

are considered as better indicators of the segmentation performance.

Intragranular carbide segmentation

A separate ground truth analysis was conducted for intragranular carbides. Figure 7 compares the manually labeled and RF segmented sections: (a) the original section, (b) the expert labeled image, (c) the original RF output, and (d) the intragranular carbide slice of the RF-labeled image against a null background (negative class/intergranular carbides plus matrix pixels). The overlay of manual and RF labels in Fig. 7e shows that the ML model was able to correctly identify the majority of the precipitates in the segmented intragranular section.

Misclassified false-positive pixels are seen entirely on the edges of intergranular carbides. The intragranular carbides misclassified by the model are smaller than those correctly identified and are similarly sized as the residual replication artifacts in the image. On the other hand, majority of false-negative misclassifications happen within a small pixel depth along the perimeters of correctly identified intragranular carbides. While the base geometries of the carbides are correctly identified, the misclassified pixels around the outer perimeter introduce uncertainty in the overall carbide area by a small percentage.

The model performance metrics for intragranular carbide segmentation are also listed in Table 1. An

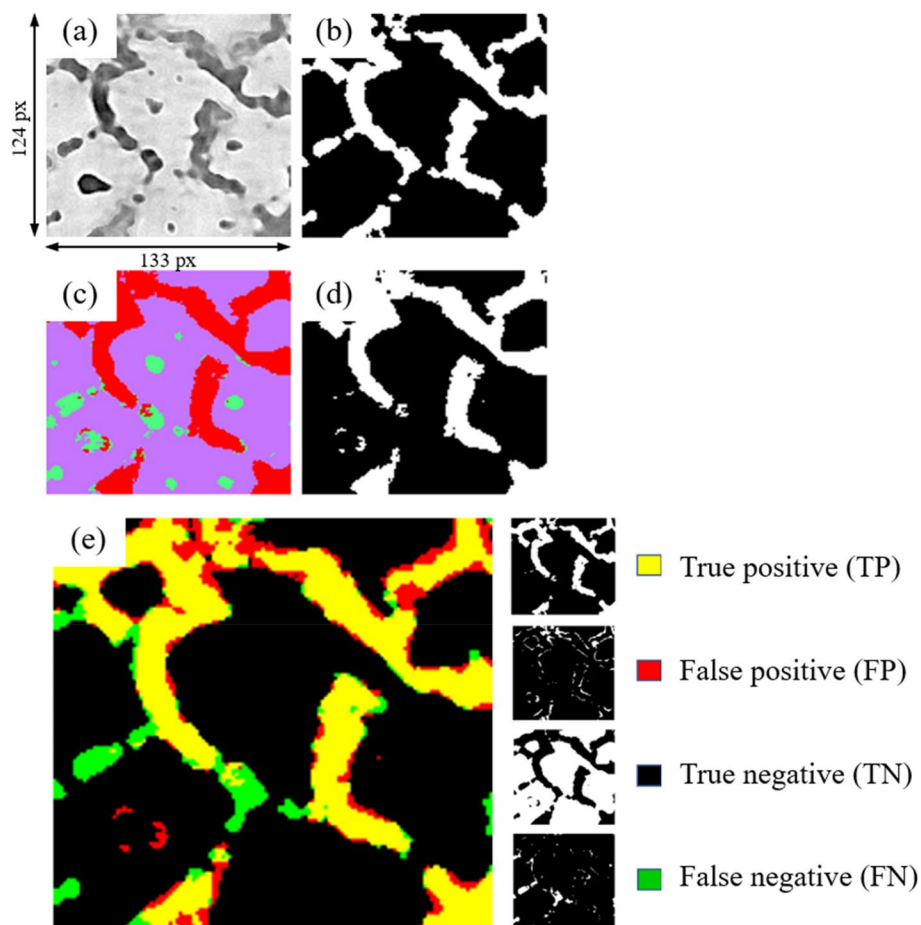


Fig. 6 Sample preparation for intergranular carbide segmentation performance analysis. **a** Preprocessed image section, **b** manual labels, **c** original RF segmentation results, **d** isolated boundary carbide slice, and **e** overlay of segmented intergranular carbides and manual labels

Table 1 Carbide identification performance metrics for intergranular and intragranular classes

Positive class — carbides	Accuracy	Precision	Specificity	Fallout	Sensitivity	F1 score
Intergranular	0.91	0.80	0.92	0.08	0.82	0.85
Intragranular	0.97	0.92	0.99	0.01	0.67	0.77

overall accuracy of 97% is achieved. Similar to the case of intergranular carbides, the area occupied by the negative class (matrix plus intergranular carbides) is comparatively large; therefore, accuracy alone becomes a misleading metric. Hence, the positive prediction value (precision) of 92% and the true-positive rate (sensitivity) of 67% are considered as better indicators of the segmentation performance. The low sensitivity can be attributed to the smaller intragranular carbides below the minimum detection threshold of the model that were identified in the manual labels but entirely filtered

in preprocessing and thus not identified by the RF model.

Carbide quantification

Conventional microstructural analysis focuses on precipitates that typically include size, morphology, area fraction, spatial distribution, and orientation. This work particularly focuses on the morphology of intergranular and intragranular carbides within the micrographs, as morphological evolution of carbide phases is of particular importance to thermally exposed microstructures.

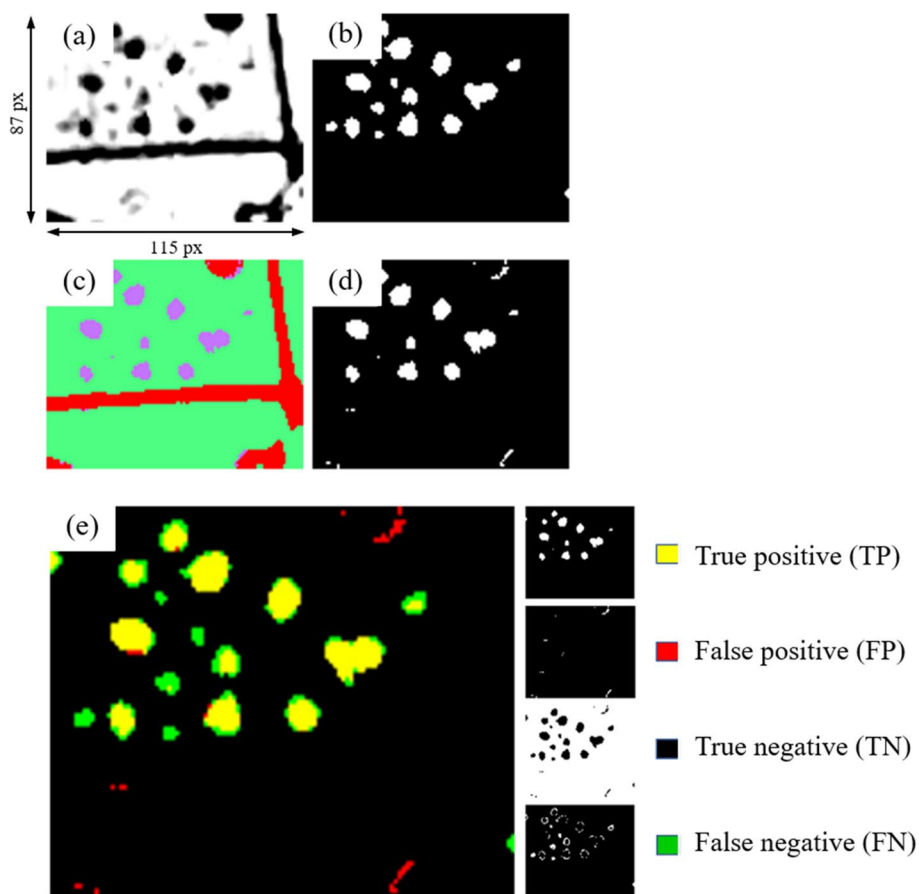


Fig. 7 Sample preparation for intragranular carbide segmentation performance analysis. **a** Preprocessed image section, **b** manual labels, **c** segmentation results, **d** intragranular carbide slice from segmentation results, and **e** overlay of segmented intergranular carbides and manual labels

The nine parent micrographs used in this study were each carefully chosen with distinct thermal histories to test the random forest model's suitability for segmenting a range of microstructures with different precipitate morphologies.

Intergranular carbides

The intergranular precipitates are predominantly comprised of primary carbides. They were quantified from the intergranular carbide slices produced by segmentation. The analysis demonstrated the possibility of capturing the variation in the shape and sizes of carbide deposition along grain boundaries in HP40-Nb replication microstructures using ML. The count/mm² of intergranular carbides ranged between 490 and 1730 across the micrographs. The average precipitate area ranged from 33 to 221 μm^2 . The area fraction of intergranular carbides ranged from 2.5 to 12.5%. There was a significant variation in circularity and aspect ratio in the micrographs, with circularity ranging from 0.55 to 0.84 and aspect ratio ranging from 0.3 to 0.5. The minimum

solidity observed was 0.75, while the maximum was 0.9. Figure 8 shows the distributions of intergranular carbide (a) average precipitate area, (b) count/mm², (c) area fraction, (d) circularity, (e) aspect ratio, and (f) solidity. The range of intergranular carbide morphologies captured via quantification demonstrates that due to thermal exposure during service, the network of thin carbides first breaks down with time into smaller precipitates before agglomerating as isolated, thick, and relatively rounded deposits.

Figure 9a shows correlations observed within the quantified data for intergranular carbides. The carbide count and area fraction have a positive correlation of 0.39 with a p -value of 0.03. Breakdown of the carbide network and decreasing carbon solubility in the matrix result in splintered, thicker precipitates and an increase in both the count and total area occupied by carbides simultaneously, which is indicative of the accumulation of thermal damage (Vaché et al. 2020; Shi and Lippold 2008). The carbide count and average area have a negative correlation of 0.27 with a p -value of 0.06,

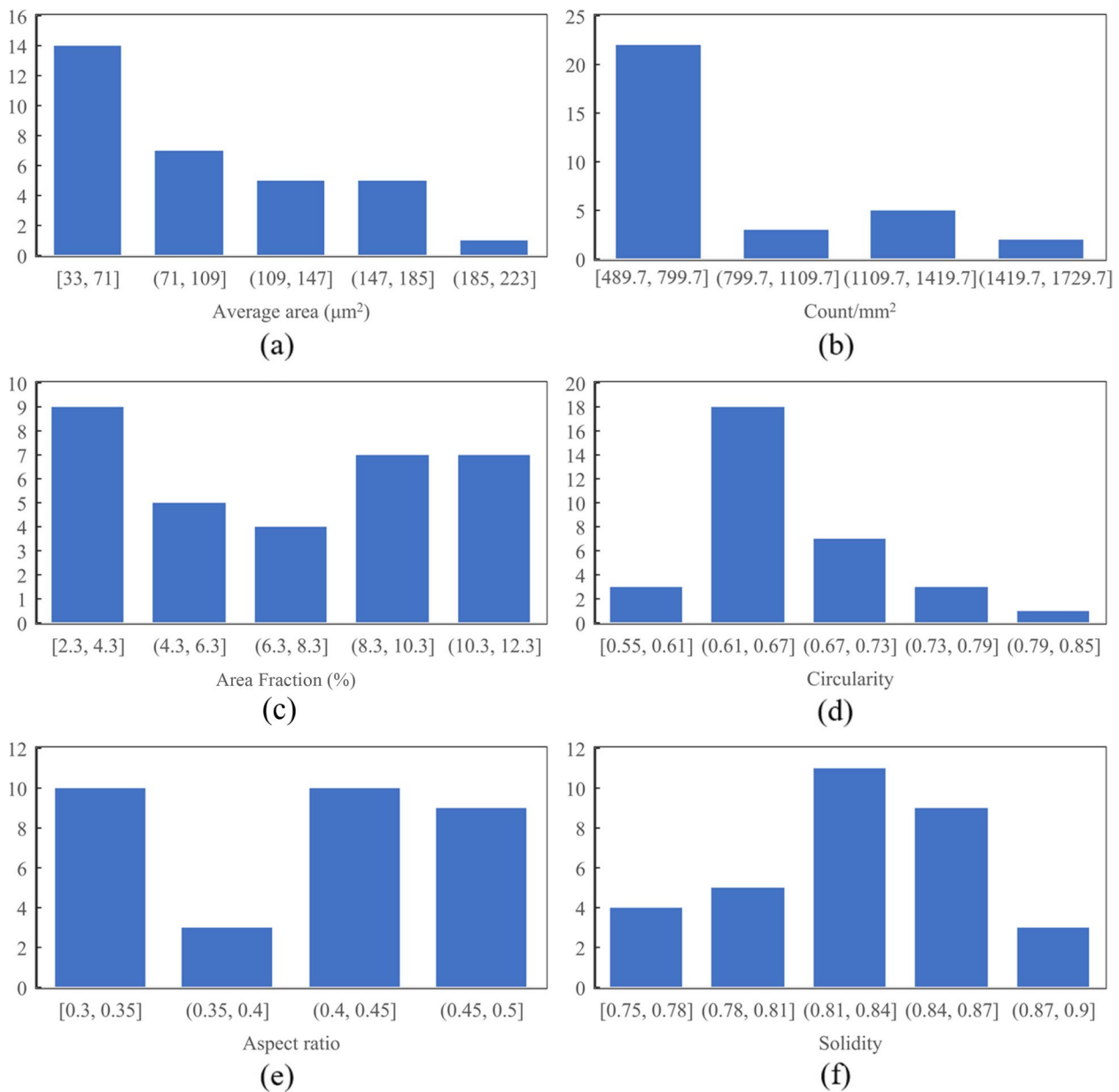


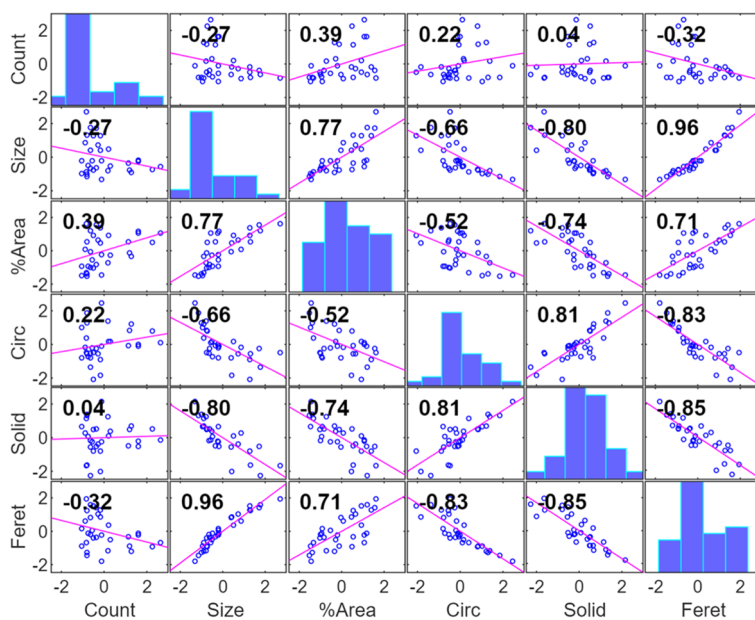
Fig. 8 Distribution of intergranular carbide. **a** average area (μm^2), **b** count/mm², **c** area fraction (%), **d** circularity, **e** aspect ratio, and **f** solidity across parent images

corresponding to the breakdown of bodies of larger, networked carbides into numerous smaller precipitates. The average area has a negative correlation with both circularity and solidity by 0.66 and 0.80, respectively, with a p -value of ~ 0 for each. Networked carbides with low thermal damage have high aspect ratios and have relatively complex perimeters due to their deposition along grain boundaries; therefore, an increase in

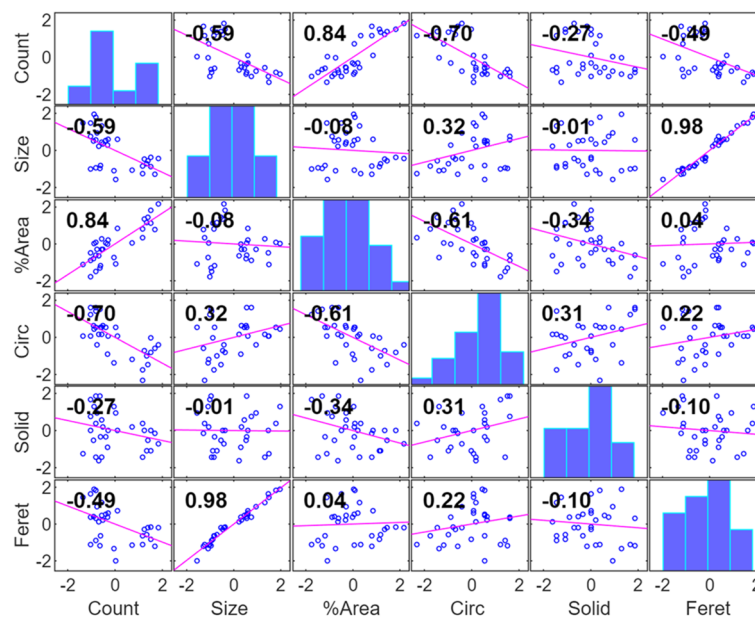
average carbide area correlates to decreasing circularity and solidity.

Intragranular carbides

The intragranular precipitates, predominantly comprised of secondary carbides, were quantified from the intergranular carbide slices produced by segmentation. The analysis revealed the extent of variation in the shape and sizes of intragranular carbides within the matrix in



(a)



(b)

Fig. 9 Correlation plots for quantified **a** intergranular carbide and **b** intragranular carbide data

the thermally exposed HP40-Nb microstructures. The distributions of intragranular carbide (a) count/mm², (b) area fraction, (c) circularity, and (d) average precipitate area are shown in Fig. 10. The count was seen to range between 140 and 510 across the micrographs. The area fraction of intragranular carbides with respect to the general microstructure ranges from 1 to 3.5%. As

intragranular carbides coalesce, they grow and become increasingly circular, with the overall circularity for the analyzed dataset varying from 0.87 to 0.92. The smallest recorded average precipitate area was 4.2 μm², and the largest was 9.6 μm². The range of intragranular carbide morphologies captured via quantification demonstrates that with the onset of thermal damage, specks

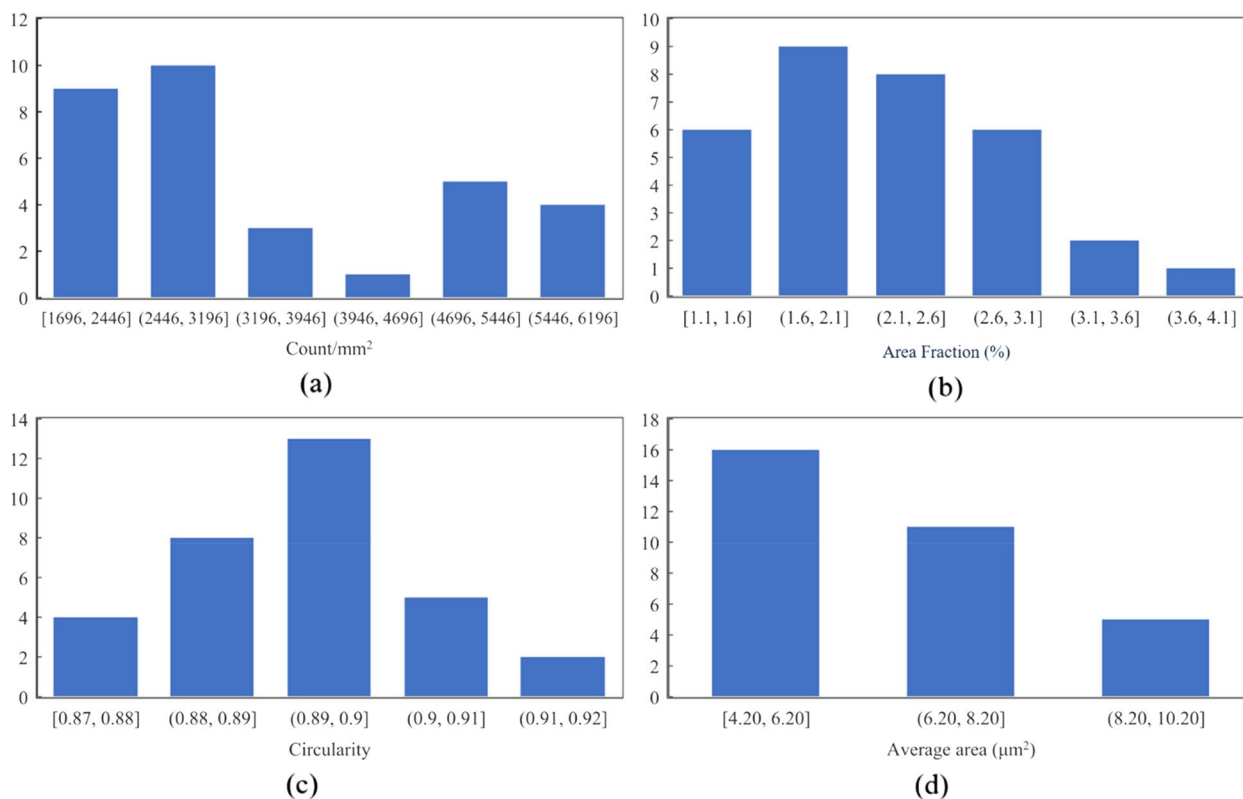


Fig. 10 Distribution of intragranular carbide. **a** count/mm², **b** area fraction (%), **c** circularity, and **d** average area (µm²) across parent images

of fine carbide precipitates appear within the material matrix, which then grow and coalesce as further damage is incurred. The large variation seen in the count, area fraction, and the average area within the microstructural dataset reaffirm that quantification of micrographs can provide an objective view toward microstructure evaluation.

Figure 9b shows correlations observed within the quantified data for intragranular carbides. There is a negative correlation of 0.59 between carbide count and average area, while count and circularity have a negative correlation of 0.70, with p -values of 0.0004 and 0, respectively. The correlations confirm that as the average area of intragranular precipitates increases, the count decreases because they merge and grow. The increase in average area is also seen to make the precipitates rounder, confirmed by the increase in circularity correlated to decreasing count. There is also a positive correlation of 0.84 between the area fraction and carbide count with a p -value of ~ 0 . When intragranular carbides first appear in the matrix, they are small and numerous. At this stage, the fresh nucleation results in higher carbide counts and area fractions. They then agglomerate when subjected to further thermal exposure, reducing the carbide count, while their migration to the grain boundaries and

coalescence with the intergranular precipitates leads to a concurrent decrease in area fraction. Hence, micrographs with higher counts of freshly nucleated intragranular carbides also demonstrate a higher area fraction. Stainless steel that is fresh or has low levels of thermal damage also exhibits low intragranular carbide counts and area fractions; however, this dataset did not contain any representative samples.

The distribution of carbide properties and correlations observed in both intergranular (grain boundary) and intragranular (within the matrix) carbide precipitate quantification demonstrate that the batch ML segmentation and subsequent quantification produced results in line with the expected evolution of a thermally aging HP40 stainless steel microstructure. These research findings highlight the significant potential of ML-driven microstructure characterization by analyzing replication micrographs, with a particular emphasis on studying precipitates in HP40 steels. Furthermore, this innovative approach has the capacity to significantly advance our understanding of the behavior and intricate evolution of various microstructural features, such as precipitates and grain size, within alloy components subjected to severe environments. In the future, we will leverage a more comprehensive and curated dataset of replicated micrographs

to track the evolution of relevant microstructural features during creep of HP40 steels to achieve automated, efficient, and precise microstructure characterization. This advancement will not only provide a deeper understanding of the creep behavior and properties of an important engineering alloy system, but its applications extend beyond steel to address the challenges posed by a wide range of demanding environments. Notably, high-strength steels, nickel-based superalloys, titanium alloys, and aluminum alloys are examples of such alloy systems extensively employed in severe applications where they are subjected to creep and fatigue damage in high-temperature environments, hydrogen embrittlement, and corrosion degradation.

The continued exploration of ML-driven characterization of important microstructural components is crucial for the future development of safer and more efficient alloys capable of withstanding the challenges posed by demanding environments. By employing ML in the analysis of low-resolution replication micrographs, researchers can design materials with enhanced performance, durability, and reliability, specifically tailored to meet the rigorous requirements of such applications.

Conclusion

This paper presented an automated computational method to segment and quantify multiple distinct carbides within batches of field replication micrographs captured by optical microscopy. It provides motivation to the scientific community interested in applying automated ML-based processes to classify and quantify microstructural features in low-resolution replication micrographs. The micrographs inherited substantial noise and artifacts due to the replication process and required preprocessing to improve image quality. It was noted that denoising and background subtraction techniques introduced a lower bound for the smallest observable microstructural feature, as precipitates similar in size to artifacts and noise were filtered out. While this negatively affected the intragranular carbide prediction precision, the overall clarity of the image and ML segmentation performance was seen to improve the RF classifier used to segment the microstructures achieving an overall accuracy of 91% for intergranular and 97% for intragranular carbides against manually labeled microstructures, which is comparable to previous carbide identification studies conducted on SEM and optical micrographs.

It was observed from the variation and correlations within the quantified shape and distribution properties of both carbide classes that the numbers obtained from segmented micrographs aligned with the expected breakdown and subsequent agglomeration of intergranular carbides and the nucleation, coalescence and

migration of intragranular carbides as HP40-Nb steel thermally ages. This approach has the capability to correctly identify and compare precipitate instances across a set of micrographs and allows for robust quantitative evaluations as opposed to manual visual inspections, which can be laborious, time-consuming, and prone to random error. The ML model's performance can be further improved by using high-resolution, noise-free replication micrographs, having a larger training dataset, and utilizing further mathematical transformations as features for the ML model. If a larger micrograph dataset can be acquired, deep learning methods such as convoluted neural networks (CNNs) could potentially produce a better segmentation performance than RFs. Comparing the carbide morphological features quantified via machine learning using optical replication micrographs to those obtained from SEM micrographs would help further ascertain the usefulness of ML models applied to replication micrographs for microstructure analysis.

This work serves to prove that ML can be implemented in industrial cases where tracking thermal damage to metals is of importance. Building on this initial study, the viability of ML methods must be explored for quantifying replicas obtained from other alloy systems where precipitate transformations across the service life due to extreme operating conditions such as exposure to corrosive media, thermal loads, and pressurized environments impact the material's functional properties. The dataset used herein did not contain any representative micrographs exhibiting defects such as voids and cracks. There were also no micrographs from unexposed or low thermal exposure samples, limiting the range of quantified morphological properties and the extent to which they could be linked to theoretical postulates related to precipitate evolution. The scope of this paper is restricted to demonstrating the ability to use ML methods to automate and quantify microstructural features from replication micrographs. Future work would include the identification and quantification of both defects and grains and establishing a criterion for determining the creep stage, fatigue life, or other relevant outcome parameters as per use case that a quantified micrograph exhibit.

Acknowledgements

This publication was made possible by the grant NPRP No.: 13S-0114-200074 from the Qatar National Research Fund (a member of the Qatar Foundation). The statements made herein are solely the responsibility of the authors.

Authors' contributions

HG, methodology, investigation, writing — original draft, analysis, and review and editing. RT, methodology, investigation, supervision, writing — analysis, and review and editing. BM, conceptualization, methodology, writing — analysis, review and editing, supervision, project administration, and funding acquisition.

Availability of data and materials

The data presented in this study are available upon reasonable request from the corresponding author.

Declarations**Ethics approval and consent to participate**

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 5 April 2024 Accepted: 13 June 2024

Published online: 26 June 2024

References

- Arganda-Carreras I, Verena K, Curtis R, Johannes S, Albert C, Sebastian SH (2016) "Trainable_Segmentation: release v3.1.2." <https://doi.org/10.5281/ZENODO.59290>
- Azimi SM, Britz D, Engstler M, Fritz M, Mücklich F (2018) Advanced steel microstructural classification by deep learning methods. *Sci Rep* 8(1):2128. <https://doi.org/10.1038/s41598-018-20037-5>
- Baskaran A, Kane G, Biggs K, Hull R, Lewis D (2020) Adaptive characterization of microstructure dataset using a two stage machine learning approach. *Comput Mater Sci* 177:109593. <https://doi.org/10.1016/J.COMMATSCI.2020.109593>
- Bonaccorsi L, Guglielmino E, Pino R, Servetto C, Sili A (2014) "Damage analysis in Fe–Cr–Ni centrifugally cast alloy tubes for reforming furnaces. *Eng Failure Anal.* 36:65–74. <https://doi.org/10.1016/j.engfailanal.2013.09.020>
- Brocher (2022) BioVoxel/BioVoxel-Toolbox: BioVoxel Toolbox. *BioVoxel Toolbox*. <https://doi.org/10.5281/ZENODO.5986130>
- Buades A, Coll B, Morel J-M (2011) Non-local means denoising. *Image Processing on Line* 1:208–212
- Bulgarevich, Dmitry S, Susumu T, Tadashi K, Masahiko D, Makoto W (2018) Pattern recognition with machine learning on optical microscopy images of typical metallurgical microstructures. *Sci Rep.* 8(1):1–8. <https://doi.org/10.1038/s41598-018-20438-6>
- DeCost BL, Lei Bo, Francis T, Holm EA (2019) High throughput quantitative metallography for complex microstructures using deep learning: a case study in ultrahigh carbon steel. *Microsc Microanal* 25(1):21–29. <https://doi.org/10.1017/S1431927618015635>
- Detroit M, Jablonski PD, Hawk JA (2021) The effect of η phase precipitates on the creep behavior of alloy 263 and variants. *Mater Sci Eng, A* 799:140337. <https://doi.org/10.1016/J.MSEA.2020.140337>
- E1351–01, ASTM (2020) Standard practice for production and evaluation of field metallographic replicas. ASTM International, West Conshohocken
- Guo G, Wang H, Bell D, Bi Y, Greer K (2003) KNN model-based approach in classification. *Lecture Notes Comput Sci* 2888:986–996. https://doi.org/10.1007/978-3-540-39964-3_62/COVER
- Holm EA, Cohn R, Gao N, Kitahara AR, Matson TP, Lei Bo, Yarasi SR (2020) Overview: computer vision and machine learning for microstructural characterization and analysis. *Metall Mater Trans A* 51(12):5985–5999. <https://doi.org/10.1007/S11661-020-06008-4/FIGURES/9>
- Hwang H, Choi SM, Jiwon Oh, Bae SM, Lee JH, Ahn JP, Lee JO, An KS, Yoon Y, Hwang JH (2020) Integrated application of semantic segmentation-assisted deep learning to quantitative multi-phased microstructural analysis in composite materials: case study of cathode composite materials of solid oxide fuel cells. *J Power Sources* 471:228458. <https://doi.org/10.1016/J.JPOWSOUR.2020.228458>
- Jain, Anil K (1988) *Fundamentals of Digital Image Processing: United States Edition*. Edited by Thomas Kailath. Pearson. <https://doi.org/10.1016/B978-012077790-7/50005-9>
- Jana S (1995) Non-destructive in-situ replication metallography. *J Mater Process Technol* 49(1–2):85–114. [https://doi.org/10.1016/0924-0136\(94\)01314-Q](https://doi.org/10.1016/0924-0136(94)01314-Q)
- Kordijazi A, Zhao T, Zhang J, Alrfou K, Rohatgi P (2021) A review of application of machine learning in design, synthesis, and characterization of metal matrix composites: current status and emerging applications. *JOM* 73(7):2060–2074. <https://doi.org/10.1007/S11837-021-04701-2/FIGURES/4>
- Lai C, Song L, Han Y, Li Q, Hui Gu, Wang B, Qian Q, Chen W (2019) Material image segmentation with the machine learning method and complex network method. *MRS Advances* 4(19):1119–1124. <https://doi.org/10.1557/ADV.2019.7>
- Marder AR (1989) Replication microscopy techniques for NDE. *ASM Handbook: Nondestructive Evaluation and Quality Control*
- Martin LP, Switzner NT, Oneal O, Curiel S, Anderson J, Veloo P (2022) Quantitative evaluation of microstructure to support verification of material properties in line-pipe steels. *Proceedings of the Biennial International Pipeline Conference, IPC 3*. <https://doi.org/10.1115/IPC2022-87063>
- Papa JP, Nakamura RYM, De Victor Hugo C, Albuquerque AX, Falcão, and João Manuel R S Tavares. (2013) Computer techniques towards the automatic characterization of graphite particles in metallographic images of industrial materials. *Expert Syst Appl* 40(2):590–597. <https://doi.org/10.1016/J.ESWA.2012.07.062>
- Perera R, Guzzetti D, Agrawal V (2021) Optimized and Autonomous machine learning framework for characterizing pores, particles, grains and grain boundaries in microstructural images. *Comput Mater Sci* 196:110524. <https://doi.org/10.1016/J.COMMATSCI.2021.110524>
- Poesele A, Ramponi G, Mathews VJ (2000) Image enhancement via adaptive unsharp masking. *IEEE Trans Image Process* 9(3):505–510. <https://doi.org/10.1109/83.826787>
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S et al (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods.* 9(7):676–82. <https://doi.org/10.1038/nmeth.2019>
- Shi S, Lippold JC (2008) Microstructure evolution during service exposure of two cast, heat-resisting stainless steels—HP-Nb modified and 20–32Nb. *Mater Charact* 59(8):1029–1040. <https://doi.org/10.1016/j.matchar.2007.08.029>
- Shmilovici A (2005) Support vector machines. Edited by Oded Maimon and Lior Rokach. *Data Mining and Knowledge Discovery Handbook*. Springer US. https://doi.org/10.1007/0-387-25465-X_12
- Stan T, Thompson ZT, Voorhees PW (2020) Optimizing convolutional neural networks to perform semantic segmentation on large materials imaging datasets: X-ray tomography and serial sectioning. *Mater Charact* 160:110119. <https://doi.org/10.1016/J.MATCHAR.2020.110119>
- Sternberg S (1983) Biomedical image processing. *Computer* 16(01):22–34. <https://doi.org/10.1109/MC.1983.1654163>
- Vaché N, Steyer P, Duret-Thual C, Perez M, Douillard T, Rauch E, Véron M et al (2020) Microstructural study of the NbC to G-phase transformation in HP-Nb alloys. *Materialia* 9:100593. <https://doi.org/10.1016/J.MTLA.2020.100593>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.