Nuclear Science and Techniques (2026) 37:133
https://doi.org/10.1007/s41365-026-01961-y

User-guided interactive machine learning for high-throughput, multiscale helium-bubble segmentation and quantification

Zhi-Wei Zheng[1] · **Xue-Zheng Yue**[1,4] · **Jin-Cheng Wang**[2,3] · **Juan Hou**[1]

Received: 21 October 2024 / Revised: 26 January 2025 / Accepted: 23 February 2025 / Published online: 12 May 2026

Abstract
Microstructural characterization of materials is fundamental for predicting macroscopic properties; however, it is typically influenced by characterization techniques and user subjectivity. In terms of irradiated materials, electron microscopy reveals significant variations in helium-bubble characteristics owing to diverse imaging conditions and task-specific requirements, thus necessitating robust statistical analysis. This study presents a novel approach that combines deep-learning and machine-learning methodologies to address these challenges. We introduce an interactive image-segmentation technique that utilizes minimal user annotations to guide a model in identifying and segmenting regions of interest. To overcome the limitations of existing deep-learning methods in detecting helium nanobubbles, we implement advanced computer vision-based machine-learning algorithms. This user-guided interactive machine-learning framework enables the extraction of relevant helium bubbles customized to specific user requirements, thereby mitigating the potential biases inherent in previous model annotations. Our methodology demonstrates enhanced accuracy and adaptability in helium-bubble analysis within irradiated materials, thereby contributing to more precise microstructural characterizations. The proposed approach is applicable to a wide range of material-science applications, thus offering a more objective and user-centric method for analyzing complex microstructural features.

Keywords Nuclear irradiation · TEM analysis · Helium bubble · Microstructural analysis · Artificial intelligence

1 Introduction

Helium bubbles contribute significantly to the degradation of nuclear-reactor components. These microscopic structures are ubiquitous in irradiated structural materials and serve as primary nucleation sites for voids and grain-boundary creep cavities [1, 2]. In nuclear environments, radiation-induced point defects can evolve into larger defect clusters, including helium bubbles, dislocation loops, and voids. This evolutionary process can ultimately result in macroscopic material failure [3]. This issue is particularly pronounced in future fusion reactors, where helium atoms are generated through (n,α) particle reactions. Because of their low solubility in metallic alloys, helium atoms accumulate within the metal matrix and form nanoscale bubbles [4–6]. Under sustained irradiation, helium bubbles contribute substantially to material swelling and performance degradation [6], thus rendering them a critical concern with respect to the longevity of core materials in nuclear power plants [7].

This work was supported by the Natural Science Foundation of China (No. U22B2067) and Shanghai Engineering Research Center of High-Performance Medical Device Materials (No. 20DZ2255500).

✉ Xue-Zheng Yue
usst-yzyz@usst.edu.cn

1 School of Materials Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200082, China

2 Department of Mechanical Engineering, The University of Melbourne, Parkville, VIC 3010, Australia

3 School of Engineering, M050, The University of Western Australia, 35 Stirling Highway, Crawley, Perth, WA 6009, Australia

4 Graduate School of Aerospace Engineering, Gyeongsang National University, Jinju 52828, Korea

Advanced characterization techniques, particularly electron microscopy, have revolutionized material science by providing detailed insights into microscopic structures [8, 9]. In situ transmission electron microscopy (TEM) is particularly valuable for analyzing complex defect structures in radiation material science [10–12]. TEM allows for the direct observation of bubbles and voids, primarily through bright-field imaging in thin samples, where changes in contrast indicate the bubble size. Complementary techniques, such as small-angle neutron scattering and small-angle X-ray scattering, offer additional information on nanoscale features, although they require meticulous data interpretation [13].

Recent studies have extensively used helium-ion irradiation to simulate neutron irradiation and employed TEM to analyze the size and distribution of helium bubbles in various metals [14–18]. The characteristics of these bubbles serve as indicators of material swelling rates [19] and the extent of radiation-induced effects [20], thus providing metrics for evaluating radiation resistance [21–25]. Studies showed that helium bubbles exhibit distinct shapes in different materials: squares in martensitic steels, predominantly circles or hexagons (cubic structures) in 304 L stainless steel, and primarily circles or polygons in tungsten alloys [26–28]. The characteristics of these shapes, which vary significantly with temperature, not only facilitate more accurate measurements of helium-induced swelling but may also enable the identification of base materials through imaging alone [29]. Furthermore, studies have demonstrated the potential of differentiating between various inclusions by observing bubble characteristics, such as distinguishing helium bubbles from argon (Ar) bubbles based on contrast differences [29, 30]. This capability underscores the importance of obtaining a comprehensive understanding into the dynamic evolution of helium bubbles, particularly their size and distribution. Such knowledge is crucial for advancing material science in nuclear applications because it provides valuable insights into material behavior under irradiation and facilitates the development of radiation-resistant materials [31].

Machine learning in nuclear material science has focused primarily on predicting material properties and extracting information from the literature [32]; however, the segmentation of helium bubbles based on TEM images remains a relatively underexplored area. Despite the efforts expended in this direction, research on helium-bubble segmentation remains limited. Initially, Anderson et al. successfully employed convolutional neural network (CNN)-based neural networks [33], whereas subsequent studies utilized Mask-RCNN [34–36]. These approaches require substantial annotated training data and are limited to the recognition of large, well-defined helium bubbles in neutron-irradiated components. By contrast, methods based on clustering [37] or computer vision models [38, 39] demonstrate potential in identifying densely packed small helium bubbles during the early stages of nanobubble coalescence; however, they typically exhibit instability during training. The persistent challenge of accurately recognizing helium bubbles, particularly nanoscale bubbles formed under low-temperature irradiation, hinders their quantitative analysis and creates a bottleneck in research. This underscores the urgent necessity for high-throughput, repeatable methods that can detect helium bubbles at various developmental stages and in diverse conditions. Addressing these challenges is crucial for advancing our understanding of nuclear materials and their behavior under irradiation, which may potentially advance the field significantly.

This study introduces a user-guided interactive machine-learning approach to improve microscopic segmentation in nuclear material analysis. The proposed method addressed the following two key challenges:

1. Insufficient training data to include all possible scenarios in nuclear material imaging.
2. The instability of computer vision models when autonomously identifying helium bubbles.

By incorporating user expertise, we aim to enhance the accuracy and adaptability of helium-bubble analysis under various irradiation conditions. Additionally, we investigate the swelling rate induced by these bubbles. This approach not only improves segmentation accuracy but also facilitates in understanding the damage mechanisms in nuclear materials. Ultimately, this study seeks to advance the design of radiation-resistant materials, which are crucial for the development of next-generation nuclear energy technologies.

2 Materials and methods

2.1 Materials and dataset preparation

Our dataset was obtained primarily from two sources. One is a widely used open-source dataset published by Anderson [33], which is derived from nickel-based superalloys used in CANDU (Canada Deuterium Uranium reactors) [7], specifically the neutron-irradiated ex-service Inconel X-750—a precipitation-hardened variant of Alloy-600—irradiated to approximately 55 dpa and 18000 appm helium. This dataset has been used as a benchmark in numerous experiments.

The other source was obtained from our previous dataset on the ion irradiation of additively manufactured and conventionally rolled 304 L stainless steel [38]. This section describes the simulation of various stages of neutron irradiation using ion irradiation. The samples were irradiated with a 350 keV He^+ ion beam at a density of 1×10^{17} ions/cm^2, thus resulting in the formation of helium bubbles within the material with a damage level of 2.35 dpa and a helium concentration of 5.5 at.%. The irradiated samples were cut to 50 nm and examined using TEM (Tecnai G2 F20 S-TWIN; FEI Company, USA). The Mask R-CNN training dataset is publicly available [33]. It contains 224 images as the training data, and the validation set comprises 19 images. For the test set, we selected eight bright-field TEM images from this dataset, as well as four bright-field TEM images from our in-house dataset for testing.

We selected underfocused images from both datasets to be included in our study.

2.2 User-guided interactive machine learning

In this section, we present a novel methodology that integrates human expert knowledge with multimodel fusion for the analysis of helium bubbles of various sizes and characteristics. Our approach leverages user-guided interactive machine-learning techniques to combine the advantages of human expertise and advanced computational models. This integration enables a comprehensive and accurate assessment of helium-bubble formation and behavior across different scales.

To facilitate the adoption of our method, we provide an example code that includes the necessary libraries and demonstrate the training of the generative convergence model (GCM) using a few helium-bubble annotations. Users can easily follow this example to install the required libraries in a Python environment and replace the provided images and annotations with their own data to apply the model to similar images. This integration enables a comprehensive and accurate assessment of helium-bubble formation and behavior across different scales.

2.2.1 Morphological features of helium bubbles

TEM is an effective tool for examining the microstructure of materials and can reveal the distinct characteristics of helium bubbles based on their size and shape. As illustrated in Fig. 1, which presents both simulated and actual TEM images of helium bubbles, the diameter of Fresnel rings serves as a crucial indicator of the bubble size. Bubbles exceeding 2 nm in diameter exhibited clearly discernible first bright and dark Fresnel rings. This phenomenon is due to the interaction between the electron beam and helium-filled cavity, which results in a phase contrast that manifests as characteristic rings. As the bubble size decreased, an interesting optical effect occurred: The relative size of these diffraction rings increased in proportion to the bubble diameter. This inverse relationship provides valuable information regarding bubble dimensions. However, when the bubble size decreased to below 1 nm, the imaging characteristics changed significantly. For these nanoscale bubbles, the first bright ring overlapped with the bubble, thus effectively obscuring its shape. Simultaneously, the dark ring became indistinct and blended into the background, thus rendering its differentiation from the surrounding matrix challenging. This transition in imaging features presents a unique challenge in the detection and analysis of miniscule helium bubbles.

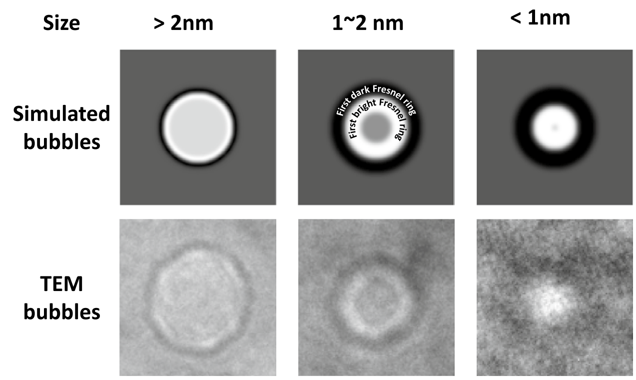

Fig. 1 (Color online) Fresnel-diffraction characteristics of helium bubbles in simulation and actual imaging. Columns from left to right show bubbles > 2 nm, 1-2 nm, and < 1 nm in diameter. Bubbles > 2 nm exhibit a faint first bright Fresnel ring. The 1-2 nm bubbles show the most distinct boundaries between dark and bright Fresnel rings. For bubbles < 1 nm, boundary definition is challenging as the first bright Fresnel ring exceeds the bubble's imaging range

To address these challenges and achieve a comprehensive detection and quantitative analysis of helium bubbles across all size ranges, this study proposes a novel model-fusion approach. This method allows each model to focus on the size range in which it excels. Mask-RCNN—a deep-learning model—was employed to detect helium bubbles with clear Fresnel rings, typically those larger than 1 nm. Mask-RCNN is particularly effective in identifying and segmenting objects with well-defined boundaries, thus rendering it ideal for bubbles with distinct ring structures. To complement this, an improved GCM was utilized to detect bubbles smaller than 1 nm with blurred boundaries. The GCM was adapted to address the challenges posed by the overlapping bright rings and indistinct dark rings, which are characteristic of these miniscule bubbles. By combining the results from these two models, we aim to comprehensively detect helium bubbles across the entire size spectrum. This fusion approach leverages the advantages of each model to overcome the limitations inherent in analyzing bubbles of varying sizes.

To enhance the accuracy and reliability of our analysis, we incorporated user-guided interactive machine learning. This novel approach eliminates biases that may be introduced during model training, whether from manual annotation or automated processes. The interactive component allows the model to learn from expert inputs, thus continually improving its performance and adapting to the nuances in different TEM images. This comprehensive methodology not only enables the detection of helium bubbles across all size ranges but also ensures high accuracy and reliability in quantitative analysis. By combining sophisticated deep-learning techniques with expert human guidance, we aim to advance material characterization using TEM, thereby providing invaluable insights into the behavior of helium bubbles in various materials and conditions.

2.2.2 Machine-learning models

Among the various methods for segmenting helium bubbles, the Mask-RCNN model demonstrated effectiveness in helium-bubble detection, as reported in previous studies [34]. However, this approach relies on pre-annotated training data, which can introduce inherent limitations. The model may inherit biases from human annotators, thus potentially restricting its ability to recognize helium bubbles with characteristics different from those in the training set. This limitation is particularly evident when examining helium bubbles smaller than 1 nm, which do not present clear Fresnel-diffraction rings and exhibit highly variable shapes. Consequently, the Mask-RCNN model may struggle to identify and segment these smaller structures accurately.

Hence, the GCM is introduced to computer vision-based machine learning, particularly for detecting minute nanosized helium bubbles. The innovation of the GCM is its unsupervised-learning framework, which differs from conventional methods using predefined shape features. Specifically, the GCM utilizes a series of trainable filters (or convolution kernels) that are thoughtfully designed and stacked together.

Unlike conventional CNNs, which typically process various color and grayscale images using randomly initialized convolution kernels, the GCM is specifically optimized for grayscale electron microscopy images. This is achieved by significantly simplifying the model structure and replacing random convolution kernels with filters of predefined distributions and functions. These filters include Gaussian, adaptive histogram equalization, and adaptive binarization filters. The key stages of the workflow are as follows:

1. *Preprocessing with Gaussian and Uniform Filters*: In the initial stage, the model preprocesses grayscale electron microscopy images using Gaussian and adaptive histogram equalization filters. These filters suppress noise and enhance the clarity of helium-bubble structures by emphasizing relevant features. This preprocessing step generates a substantial amount of training data, thus establishing a foundation for the subsequent stages.
2. *Feature Extraction via Adaptive Binarization*: Following preprocessing, the GCM applies adaptive binarization to extract helium bubbles by analyzing the contrast between the bubbles and their surrounding environment. This step isolates the key features of the helium bubbles while maintaining structural integrity, thus enabling effective feature extraction. Additionally, the model outputs the helium-bubble count.
3. *Parameter Refinement via Iterative Training*: Based on a specific loss function's evaluation of the reasonableness of the helium-bubble count, the model iteratively refines the parameters of the filters, including their size, intensity, and distribution. These refinements are guided by results from previous iterations, thus allowing the GCM to progressively optimize the filter parameters and improve detection accuracy.

The filters utilized in the GCM can be efficiently implemented using libraries such as Python-OpenCV, thus rendering the approach computationally efficient and accessible for practical applications.

Owing to this unique design, fully random kernels need not be trained; thus, learning parameters such as filter size and intensity as well as other characteristics can be prioritized. Although this approach limits the applicability of the GCM to grayscale images, it provides significant advantages in terms of interpretability and reduces the number of annotations required for training.

The initial processing stages of the GCM involve Gaussian filtering as a preprocessing step, which can be represented by the following Eq. (1):

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \tag{1}$$

where $G(x, y)$ is the weight of the filter and σ is the standard deviation that determines the degree of blurriness, and x and y are the distances from the center point.

The application of Gaussian filtering as a preprocessing step in the GCM serves a dual purpose in enhancing the detection of helium bubbles. Primarily, it serves as a noise-reduction mechanism that smooths out random fluctuations in an image that may potentially interfere with accurate bubble detection. Simultaneously, this filtering process enhances the distinguishing features of the helium bubbles. Because helium bubbles feature indistinct or fuzzy boundaries, the blurring effect of a Gaussian filter is advantageous for the detection process. By applying a controlled amount of blur, the filter homogenizes the appearance of these diffuse boundaries, thus rendering them more uniform and thus more easily recognizable in subsequent processing steps. This homogenization effect is particularly beneficial for creating a more consistent representation of helium bubbles across the image, thus allowing the GCM's trainable filters to learn and detect these structures more effectively, even when they vary slightly in size or shape. Essentially, the Gaussian filter prepares the image data such that they align well with the GCM's unsupervised-learning approach, thus providing a basis for more robust and accurate helium-bubble detections.

However, the GCM may not exhibit instability when it is trained without manual annotations. Hence, we have enhanced the model by integrating user-guided interactive machine learning. This approach eliminates the necessity to focus on the specific shapes of the nanosized helium bubbles, which is impractical owing to their variable morphology. Instead, users provide point annotations at the central positions of the bubbles. Subsequently, the GCM learns to infer the potential bubble shapes by analyzing the features surrounding these labeled points. This interactive method enables an interpretable training of the GCM, thus resulting in more reliable helium-bubble segmentations.

Figure 2 illustrates our ensemble model approach for segmenting helium bubbles of various sizes by incorporating user annotations. When the user-guided point annotations are provided, the ensemble model trains the GCM to segment the corresponding helium bubbles based on these annotations. Otherwise, the model employs the GCM with default parameters alongside the Mask R-CNN for helium-bubble segmentation and analysis. This dual approach enables the merging of results, thus ensuring a comprehensive segmentation of helium bubbles across all size ranges. This figure shows the manner by which this adaptive method addresses different input scenarios to maximize the accuracy and completeness of bubble detection and characterization.

Examples of our workflow demonstrate the flexibility and efficiency of the ensemble model. In scenarios with user-guided point annotations (rows 1 and 2), the model precisely captured the requirements of the annotator by employing a user-guided approach to segment helium bubbles corresponding to the annotated sizes. This underscores the significance of user-guided point annotations in enhancing the model performance, particularly in the detection of microscopic bubbles. Conversely, the scenario without user annotations (row 3) demonstrates the autonomous capabilities of the model, although it primarily focused on the detection of larger bubbles. This comparison highlights the adaptability of the model to various input conditions. Notably, the output encompasses not only the location and morphology of the bubbles but also the size distribution analysis, which is crucial for material-science research. The model integrates the unsupervised-learning advantages of the GCM with the object-detection capabilities of the Mask R-CNN. In the absence of user annotations, the results of these two methodologies are intelligently merged to achieve the most comprehensive bubble detection outcomes. This adaptive approach not only enhances detection accuracy and completeness but also provides researchers with the flexibility to select optimal analysis methods under diverse experimental conditions.

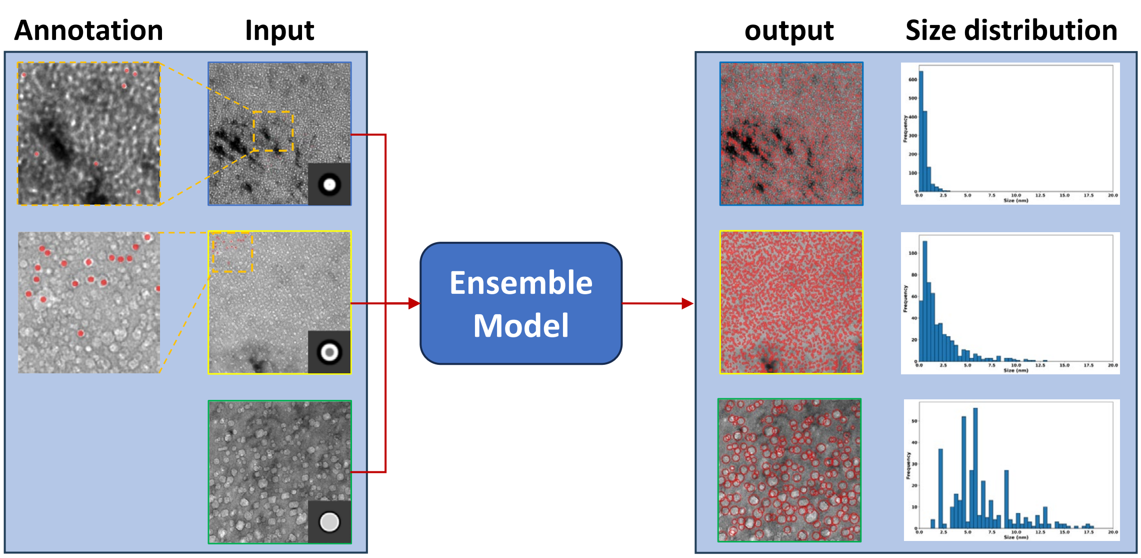

Fig. 2 (Color online) Example workflow of helium-bubble detection model. Input comprises the original image and a user-annotated image. Output includes identified helium bubbles (indicated by red regions or red circles) and their size distribution. Rows 1 and 2 show the model's operation with user-guided point annotations, which results in the precise segmentation of all helium bubbles with accurate shape and size information. Row 3 illustrates the model's performance without user annotations, where it effectively detects larger helium bubbles, which are denoted by circular indicators. This comparison highlights the effect of user input on the model's segmentation capabilities across various bubble sizes

2.2.3 User-guided point annotations

The proposed method incorporates an innovative technique that enables domain experts or users to optimize a model with minimal point annotations customized to their specific requirements. Specifically, although precisely defining the boundaries of nanoscale helium bubbles is challenging, users are only required to provide center-point annotations as inputs. Meanwhile, domain experts may further refine these annotations by specifying additional details, such as bubble size or shape. Although point annotations may appear simplistic, Gaussian filtering transforms these annotations into a normal probability distribution with the center point representing the highest likelihood of being part of a helium bubble, while decreasing probabilities are represented by the distribution radiating outward. This interactive segmentation approach significantly enhances the reliability and interpretability of the model by learning critical features from sparse annotations.

As shown in the first and second rows of Fig. 2, sparse point annotations can significantly improve the output quality of the model. This outstanding efficiency is primarily due to the adoption of the GCM as a core component in our integrated framework. Unlike conventional neural networks, which typically struggle to effectively utilize sparse point labels and often require full shape annotations for accurate training, the unique architecture of the GCM allows one to leverage the diffuse information surrounding point annotations. This allows the model to be trained effectively, even when the precise shape of the helium bubbles cannot be determined, using point annotations as the primary input. The GCM achieves this by analyzing the image features around the annotated points and inferring the approximate shape and characteristics of the helium bubbles. Notably, the GCM effectively incorporates this probabilistic information into its training process and synthesizes it with the surrounding image features to infer more complete target contours.

Through its iterative convergence mechanism, the GCM progressively refines its understanding of the target's shape. By considering both the point annotations and the relationships among surrounding pixels, it generates high-quality segmentation results and accurate quantitative analyses. This ability to process sparse input is a key factor contributing to the success of our approach, as it enables the model to extract meaningful and comprehensive information from minimal user input.

Compared with conventional deep-learning methods, such as the region-based convolutional neural network (RCNN) and Mask-RCNN, our approach achieves comparable performance with significantly fewer point annotations, thereby reducing both data-preparation effort and cost. Furthermore, as the model is optimized based on user-guided point annotations, its decision-making process becomes more transparent and interpretable, thus allowing users to intuitively understand the manner by which the model utilizes these annotations to improve its output. This interpretability is particularly important in domains such as medical diagnosis and scientific research. Additionally, our method is highly adaptable and flexible, as it enables users to select the number and location of point annotations based on domain-specific knowledge and requirements. This allows the model to adapt promptly to different applications and datasets without requiring extensive retraining. Whereas the GCM serves as the core of our approach, its integration with other models—such as conventional CNNs—for capturing global features further enhances its robustness and generalizability.

2.3 Evaluation metrics

Standard metrics such as precision and the F1-score are widely used to quantify the overlap between a segmentation map and its ground truth.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

We evaluated our model's predictions using point annotations as the ground truth, as shown in Fig. 3. True positive (TP) instances were defined as cases in which the model accurately identified a positive sample, with a predicted bubble region encompassing a ground-truth point annotation. False positive (FP) cases represent misclassifications in which the model predicts a bubble in an area without a corresponding ground-truth annotation. False negative (FN) instances indicate situations in which the model fails

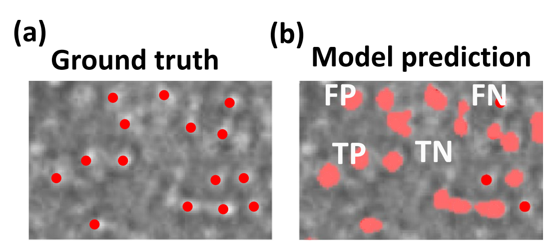

Fig. 3 (Color online) Evaluation of model predictions using point annotations as ground truth. **a** Original image with point annotations serving as ground truth. **b** Model prediction overlaid with labeled examples showing the manner by which true positive (TP), false positive (FP), true negative (TN), and false negative (FN) are defined

to detect a bubble at a location denoted by a ground-truth point annotation. Because of the challenge in precisely defining helium-bubble edges in TEM images, we utilized point annotations placed at the core of each bubble as our ground-truth reference. This approach allowed us to calculate the TP, FP, and FN while accommodating the inherent difficulties of exact boundary delineation at the nanoscale. Figure 3a shows an original image with point annotations serving as the ground truth, whereas Fig. 3b presents the model's predictions overlaid on the image, which shows the manner by which the TP, FP, and FN are defined in the context of our evaluation framework.

3 Results and discussion

In this study, we evaluated the performance of our model in analyzing helium-bubble sizes and investigated the effect of varying user inputs on its effectiveness. Our assessment involved testing the ensemble model on selected underfocused images and examining the manner by which different numbers of annotation labels affect the model's performance using a smaller dataset. This approach allowed us to gauge the accuracy of the model in bubble-size analysis and its sensitivity to the extent of the user guidance provided.

3.1 Model performance

In this study, we compared the performance of our ensemble model with that of a previously used Mask R-CNN for analyzing larger helium bubbles. This comparison is particularly relevant because helium-bubble sizes typically vary with irradiation temperature. The benchmark micrographs used in this analysis showed helium-bubble formation in the helium-dense regions of 304 L stainless steel subjected to high-temperature irradiation at 750 °C. This approach allowed us to evaluate the effectiveness of our ensemble model in accurately detecting and characterizing larger helium bubbles under specific irradiation conditions.

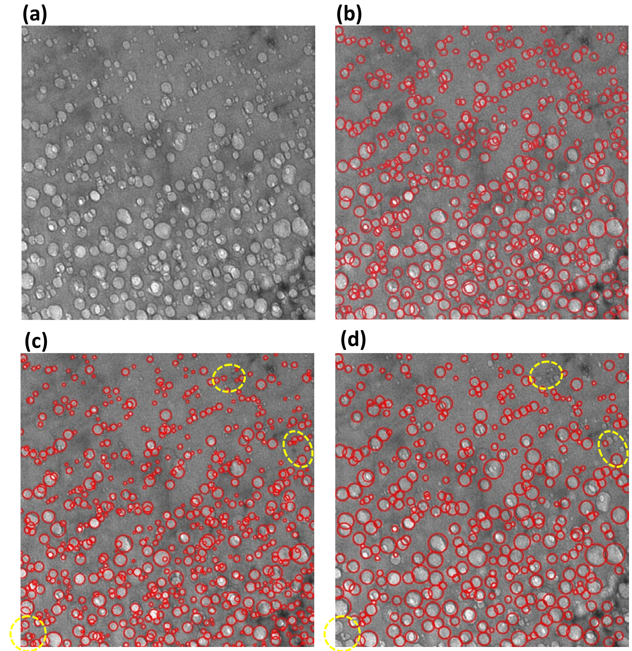

Fig. 4 (Color online) Comparison of helium-bubble segmentation methods. **a** Original transmission electron microscopy image. **b** Expert-annotated ground truth. **c** Results from our proposed ensemble model. **d** Results from previous state-of-the-art method (Mask R-CNN)

Figure 4 presents a comparative analysis of helium-bubble segmentation in 304 L stainless-steel samples after 100 h of irradiation annealing. Individual helium bubbles are indicated by red circles in all the images. Our model, which integrates the Mask R-CNN and GCM, demonstrated significantly improved segmentation accuracy compared with previous state-of-the-art methods, particularly for smaller bubbles and those located near image edges, as highlighted by the dashed yellow circles. These areas were particularly challenging to detect and segment accurately using previous methods.

The ground truth shown in Fig. 4b was annotated by domain experts. Compared with the results from the previous state-of-the-art method shown in Fig. 4d, our ensemble model presented in Fig. 4c demonstrated superior performance in comprehensively detecting and delineating helium bubbles of various sizes and locations within the image. This advancement is attributable to the synergistic combination of the Mask R-CNN and GCM, where their respective strengths were leveraged. However, owing to its tendency to detect all bubbles, our model may split overlapping helium bubbles into multiple segments. This behavior contrasts with the primary error observed in the previous Mask R-CNN method, which merged overlapping bubbles into one and struggled to detect smaller bubbles.

Statistically, this resulted in an error of approximately 5% in the detected bubble count and a slight reduction in the average bubble size (tested on the 304 L stainless-steel dataset). This effect may be more pronounced in images with specific characteristics. Therefore, we recommend that users compute the helium-bubble count and average size across multiple images to obtain more reliable results.

Notably, our model achieved improved and more comprehensive segmentations without relying on handcrafted annotations, which represents a significant achievement in the automated helium-bubble detection and analysis of irradiated steel samples. The ability to accurately segment smaller bubbles and those near edges, which were previously overlooked or inaccurately segmented, enables more precise quantifications and characterizations of the microstructural features. Consequently, this provides deeper insights into the microstructural evolution of irradiated materials and ultimately contributes to the development of radiation-resistant materials for nuclear energy, space exploration, and other critical applications.

Figure 5 shows two examples in which our ensemble model excels in accurately detecting and segmenting smaller and densely packed helium bubbles—a challenge that previous methods struggled to address effectively. Conventional approaches are typically inadequate in managing complex scenarios with numerous small and closely arranged bubbles. By contrast, our model demonstrated exceptional ability in precisely delineating individual helium bubbles, even in intricate situations. This capability is particularly significant because the accurate detection and analysis of smaller bubbles are vital for the comprehensive characterization of irradiated materials.

As shown in Table 1, our ensemble model achieved significant improvements in terms of the overall segmentation performance. In particular, it obtained an F1-score of 0.85, which surpassed those achieved by previous state-of-the-art methods, including those of Wu et al. [37] (0.799), Anderson et al. [33] (0.78), Chen et al. [35] (0.77), and Jacobs et al. [34] (0.82). Notably, our model achieved a well-balanced performance between precision (0.81) and recall (0.89), thereby outperforming competing methods in both metrics.

Chen et al. [35] and Jacobs et al. [34] employed Mask R-CNN models. Specifically, Chen et al. [35] achieved a higher recall (0.88) through scaling techniques but at the cost of reduced precision (0.75). Meanwhile, Jacobs et al. [34] reported a high F1-score of 0.82 but did not provide detailed precision and recall metrics. In comparison, our model not only achieved the highest F1-score but also maintained an optimal balance between precision and recall, thus demonstrating superior segmentation performance in general.

Table 1 Helium-bubble segmentation performance

Method	F1-score	Precision	Recall
Wu et al. [37]	0.799	0.694	0.940
Anderson et al. [33]	0.78	0.84	0.72
Chen et al. [35]	0.77	0.75	0.88
Jacobs et al. [34]	0.82	N/A	N/A
Ours	0.85	0.81	0.89

Notably, the improvement in the evaluation metrics shown in Table 1—while significant—is not particularly pronounced in terms of numerical increase. This is partly because our test dataset excludes results from low-temperature irradiation, which contains fewer nanometer-scale helium bubbles. Meanwhile, previous methods, such as the Mask R-CNN, almost failed completely (e.g., the results shown in Fig. 5c). Additionally, most of our test data were derived from a publicly available dataset, which exhibited a much higher similarity to the training data used by the competing methods. Consequently, the performance gap between our method and the other methods may appear to be numerically smaller owing to these factors. Nevertheless, our proposed approach significantly reduced the annotation burden and enabled the analysis of nanoscale

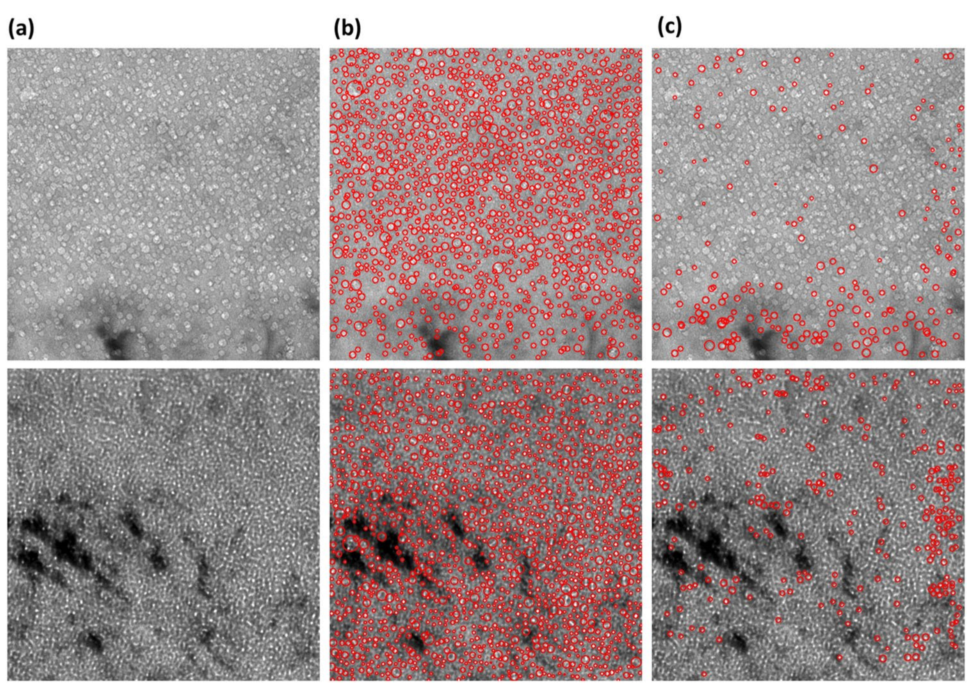

Fig. 5 (Color online) Comparison of segmentation results for smaller helium bubbles. **a** Original image; **b** output from our ensemble model; **c** output from Mask-RCNN

helium bubbles, which was previously challenging or infeasible.

3.2 User-guided point annotations and effect on segmentation

We conducted a comprehensive evaluation to investigate the effect of the number of user-provided point annotations on the segmentation performance of the model and then compared the model with existing neural-network-based interactive learning methods [40]. The assessment began with a single-point annotation, which was then incrementally increased across multiple iterations. At each step, the model was retrained using the point annotations provided as input data, and the segmentation outputs were analyzed in detail.

Specifically, we selected a diverse set of images and manually annotated one to five points on each image to simulate user-guided point annotations of different densities. Subsequently, these annotated images were used to independently train separate instances of the segmentation model. Systematically varying the number of user-guided point annotations—from sparsely annotated cases with a single point to more densely annotated examples with five points—enabled us to analyze the correlation between user-input quantity and model performance, thus ultimately enabling us to determine the optimal number of annotations required for effective segmentation.

We benchmarked our approach against two established methods: the pretrained Mask R-CNN model and current neural network-based interactive segmentation techniques [40]. To provide a more comprehensive evaluation, we expanded our analysis by incorporating helium-bubble contours as the background input, thus offering additional context for segmentation-performance assessment.

Figure 6 illustrates the effect of increasing the number of user-provided point annotations on segmentation performance. The figure shows the model input—comprising the original image with point annotations, the resulting segmentation map, and an overlay of the segmentation prediction on the original image—progressing from a single annotation to five annotations.

As the number of point annotations increased, the model progressively learned the approximate shapes and distinctive features of helium bubbles, thus yielding more accurate and stable segmentation outcomes. Because of only one or two annotations (Fig. 6a and b), the model struggled with inconsistency and yielded erratic segmentations that failed to capture the actual boundaries of the bubbles. However, by introducing three or four annotations (Fig. 6c and d), the

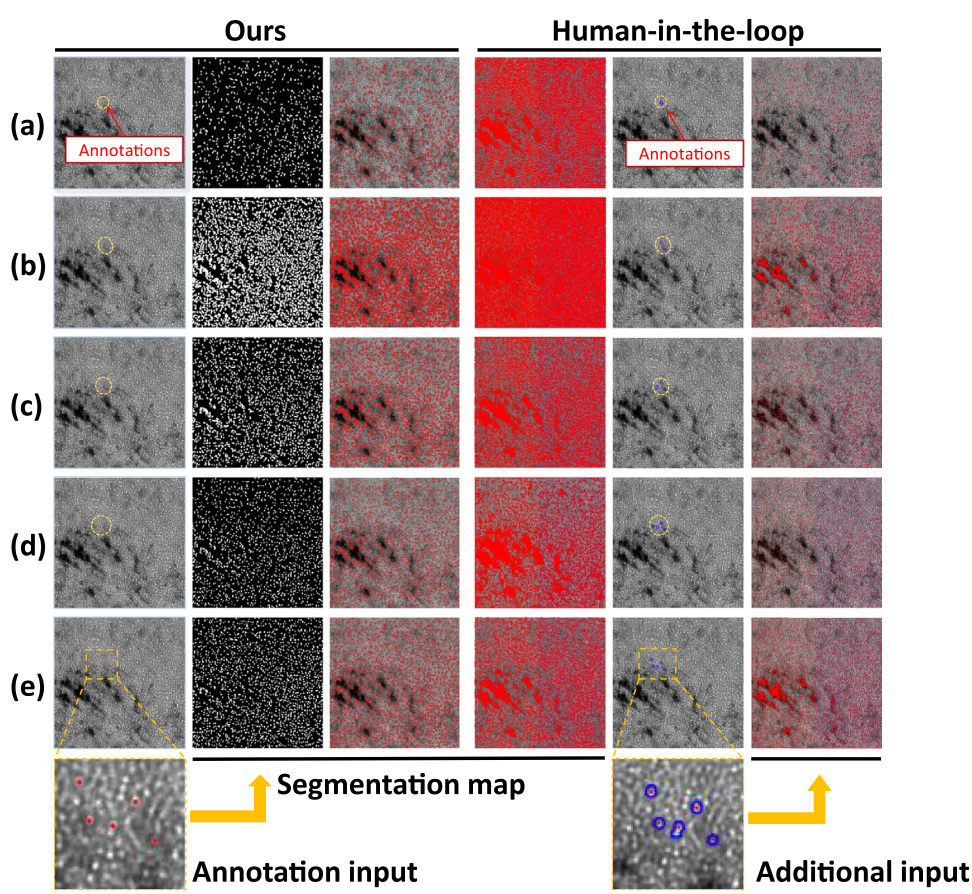

Fig. 6 (Color online) Performance comparison between our approach and neural network-based interactive methods under increasing point annotations. **a**–**e** illustrate the model predictions corresponding to one to five point annotations, respectively. Each column includes the original image with point annotations, the segmentation map, and the overlay of model predictions. For comparisons with neural network-based interactive methods (human-in-the-loop), each column presents the output from the original input, the modified input with additional background annotations, and the final segmentation results

model began to effectively recognize the essential characteristics of helium bubbles and successfully segmented most of them, albeit with occasional omissions. Further increasing the number of annotations to five (Fig. 6e) significantly enhanced the performance, thus enabling the model to segment the helium bubbles almost completely. At this stage, the model not only captured most of the bubbles but also began to approximate their shapes more precisely by replacing irregular boundaries with smoother contours. Moreover, it became less sensitive to variations in contrast and noise, thereby achieving robust segmentation across all helium bubbles. This progression highlights the ability of the model to generalize better as more annotations are provided; specifically, its segmentation capabilities are refined by leveraging the increased information regarding the bubble morphology.

This observation suggests that the model effectively leverages the limited user input, learns the visual features of helium bubbles from a few point annotations, and extrapolates this information to achieve high-quality segmentations. A quantitative analysis of this trend is presented in Table 2, which presents the F1-scores for various numbers of annotations across multiple images. In general, the F1-score exhibited an increasing or stabilizing trend with as the number of user-provided points increased. For instance, in Image 1, the F1-score began at 0.623 with a single annotation and then improved marginally to 0.634 with two annotations. A notable increase was observed with three annotations, where an F1-score of 0.712 was achieved. Four annotations further enhanced the score to 0.723, and five annotations yielded an F1-score of 0.801, thus demonstrating the exceptional performance of the model. The results for the other images corroborated this trend, thereby highlighting the ability of the model to efficiently learn the features of the target object from a limited number of interactive annotations, which consequently yielded accurate and comprehensive segmentations. This capability is advantageous for the rapid analysis of cells, particles, defects, and other entities in various applications.

The training time of our method was affected by the image size and the number of helium bubbles. In the test experiment, under an image size of 768×768 pixels, the model converged in approximately 300 iterations, which required approximately 10 min. By contrast, the human-in-the-loop [40] method converged within 2000 iterations per training cycle on average, with a runtime of approximately 5 min. The human-in-the-loop interactive segmentation approach demonstrated significant limitations when trained solely on helium-bubble labels. It incorrectly classified substantial portions of the ambiguous background and other defects, such as helium bubbles (column 4 in Fig. 6). Even with additional background annotations (column 5 in Fig. 6), the training remained unstable, as evidenced by noisy outputs and persistent defect misclassifications. This limitation stems from the inability of the method to effectively extract helium-bubble features from the annotations provided, which requires comprehensive background labeling of all defects for accurate performance.

Table 2 F1-score by increasing annotations

Points	Image 1	Image 2	Image 3
1	0.623	0.697	0.567
2	0.634	0.715	0.635
3	0.712	0.763	0.659
4	0.723	0.755	0.733
5	0.801	0.759	0.812

This challenge is particularly significant as helium bubbles are smaller than 1 nm, which renders precise contour delineation extremely difficult and time intensive. Our method addresses this limitation by requiring only central point annotations for helium bubbles, thereby reducing the labeling time by approximately 90%. This efficiency is achieved by implementing Gaussian filtering in the GCM, which transformed point annotations into probability distributions, i.e., high at the center and gradually decreasing outward, thus creating more informative spatial labels compared with simple binary annotations.

By contrast, the conventional Mask R-CNN approach required extensive training data: 389 images containing 33,517 annotations, with an average of 86 helium-bubble labels per image. Although such a comprehensive annotation is valuable for diverse, large-scale datasets, it is impractical for typical applications or smaller datasets. Our method achieved comparable results with only five point annotations per image, i.e., approximately 6% of the original labeling effort, thus rendering it significantly more efficient for practical applications.

4 Summary

This paper presents a novel approach for segmenting and quantitatively analyzing helium bubbles of various sizes using TEM micrographs. Our method employs a user-guided interactive machine-learning approach and trains a GCM to segment helium bubbles in accordance with user requirements. By integrating the GCM with a Mask R-CNN, we created an ensemble model that can manage helium bubbles across a wide range of TEM imaging conditions.

Our approach demonstrated excellent semantic segmentation ability, as evidence by its accurate identification of various helium-bubble types in underfocused bright-field TEM micrographs. It provided comprehensive data on the bubble size, shape, and other characteristics. Compared with previous machine-learning methods, our training process is more stable and accurate. Notably, our method required minimal input in the form of weak point labels but achieved results comparable or superior to those of methods requiring numerous precise labels.

Despite its effectiveness, our method struggled to accurately separate individual bubbles into overlapping regions under undesirable imaging conditions. The observed limitations were attributable to inherent challenges in semantic segmentation models. To further evaluate our method, we conducted extensive in-house tests on helium bubbles within Au, Cu, Fe-10-Cr, ceramics, and non-fullerene acceptors (NFAs) thin films, as well as on Ar bubbles and bubbles induced by H/He irradiation in various metals. Our model performed well in metals and alloys, with slightly reduced accuracy in ceramics, whereas its performance in NFA thin films was significantly affected by background interference. In H/He irradiation scenarios, the results were promising. However, for Ar bubbles and bubbles induced by other gases, additional data are required to validate the model's performance comprehensively.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Zhi-Wei Zheng, Xue-Zheng Yue, and Juan Hou. The first draft of the manuscript was written by Zhi-Wei Zheng, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability The data that support the findings of this study are openly available in Science Data Bank at https://cstr.cn/31253.11.sciencedb.j00186.00893 and https://www.doi.org/10.57760/sciencedb.j00186.00893.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. T. Cheng, G. Wei, S. Jiang et al., Enhanced resistance to helium irradiations through unusual interaction between high-entropy-alloy and helium. Acta Mater. **248**, 118765 (2023). https://doi.org/10.1016/j.actamat.2023.118765
2. M. Gilbert, S. Dudarev, D. Nguyen-Manh et al., Neutron-induced dpa, transmutations, gas production, and helium embrittlement of fusion materials. J. Nucl. Mater. **442**(1–3), S755–S760 (2013). https://doi.org/10.1016/j.jnucmat.2013.03.085
3. S.J. Zinkle, L.L. Snead, Designing radiation resistance in materials for fusion energy. Annu. Rev. Mater. Res. **44**(1), 241–267 (2014). https://doi.org/10.1146/annurev-matsci-070813-113627
4. R.J. Kurtz, H.L. Heinisch, The effects of grain boundary structure on binding of He in Fe. J. Nucl. Mater. **329**, 1199–1203 (2004). https://doi.org/10.1016/j.jnucmat.2004.04.262
5. R. Kemp, G. Cottrell, K. Bhadeshia, Classical thermodynamic approach to void nucleation in irradiated materials. Energy Mater. **1**(2), 103–105 (2006). https://doi.org/10.1179/174892406X142714
6. H. Trinkaus, B. Singh, Helium accumulation in metals during irradiation–where do we stand? J. Nucl. Mater. **323**(2–3), 229–242 (2003). https://doi.org/10.1016/j.jnucmat.2003.09.001
7. C.D. Judge, N. Gauquelin, L. Walters et al., Intergranular fracture in irradiated Inconel X-750 containing very high concentrations of helium and hydrogen. J. Nucl. Mater. **457**, 165–172 (2015). https://doi.org/10.1016/j.jnucmat.2014.10.008
8. J. Nuñez, R. Renslow, J.B. Cliff et al., NanoSIMS for biological applications: current practices and analyses. Biointerphases **13**(3), 03B301 (2018). https://doi.org/10.1116/1.4993628
9. M. Petrushevska, K. Pavlovska, J. Laskova et al., Transmission electron microscopy: Novel application of established technique in characterization of nanoparticles as drug delivery systems. Prilozi **40**(2), 67–72 (2019). https://doi.org/10.2478/prilozi-2019-0016
10. K. Burns, P.C. Reuel, F. Guerrero et al., Thermal stability and radiation tolerance of lanthanide-doped cerium oxide nanocubes. crystals **11**(11), 1369 (2021). https://doi.org/10.3390/cryst11111369
11. J. Dumbre, S.K. Kairy, E. Anbere et al., Understanding the formation of (Al, Si)$_3$Sc and V-phase (AlSc$_2$Si$_2$) in Al-Si-Sc alloys via ex situ heat treatments and in situ transmission electron microscopy studies. J. Alloys Compd. **861**, 158511 (2021). https://doi.org/10.1016/j.jallcom.2020.158511
12. M. Liu, Y.-F. Yan, Z.-B. Zhu et al., Investigated the influence of He ion irradiation on the microstructure and hardness of Ni–TiC NP composites. Nucl. Sci. Tech. **32**, 121 (2021). https://doi.org/10.1007/s41365-021-00961-4
13. Y. Dai, G. Odette, T. Yamamoto, 1.06-The effects of helium in irradiated structural alloys. Comprehensive Nuclear Materials **1**, 141–193 (2012). https://doi.org/10.1016/B978-0-08-056033-5.00006-9
14. S.-H. Li, J.-T. Li, W.-Z. Han, Radiation-induced helium bubbles in metals. Materials **12**(7), 1036 (2019). https://doi.org/10.3390/ma12071036
15. J.-T. Li, I.J. Beyerlein, W.-Z. Han, Helium irradiation-induced ultrahigh hardening in niobium. Acta Mater. **226**, 117656 (2022). https://doi.org/10.1016/j.actamat.2022.117656
16. W. Yang, J. Pang, S. Zheng et al., Interface effects on He ion irradiation in nanostructured materials. Materials **12**(16), 2639 (2019). https://doi.org/10.3390/ma12162639
17. W. Qi, Z.-T. He, B.-L. Zhang et al., Behaviors of fine (IG-110) and ultra-fine (HPG-510) grain graphite irradiated by 7 MeV Xe^{26+} ions. Nucl. Sci. Tech. **28**(10), 144 (2017). https://doi.org/10.1007/s41365-017-0292-x
18. X.-H. Yan, L. Sun, D. Zhou et al., Effects of interstitial cluster mobility on dislocation loops evolution under irradiation of austenitic steel. Nucl. Sci. Tech. **35**(8), 132 (2024). https://doi.org/10.1007/s41365-024-01493-3
19. M. Wurmshuber, M. Balooch, X. Huang et al., Helium-induced swelling and mechanical property degradation in ultrafine-grained W and W-Cu nanocomposites for fusion applications. Scr. Mater. **213**, 114641 (2022). https://doi.org/10.1016/j.scriptamat.2022.114641
20. N. Li, M.J. Demkowicz, N.A. Mara, Microstructure evolution and mechanical response of nanolaminate composites irradiated with helium at elevated temperatures. JOM **69**, 2206–2213 (2017). https://doi.org/10.1007/s11837-017-2580-7
21. X. Liang, Y. Wang, J. Zhao et al., Size-and ion-dose-dependent microstructural evolution and hardening in He-irradiated miscible Cu/Zr crystalline/crystalline nanolaminates. Surf. Coat. Technol. **366**, 255–265 (2019). https://doi.org/10.1016/j.surfcoat.2019.03.037

22. Z. Shang, J. Ding, C. Fan et al., He ion irradiation response of a gradient T91 steel. Acta Mater. **196**, 175–190 (2020). https://doi.org/10.1016/j.actamat.2020.06.019
23. T. Niu, J. Li, Y. Zhang et al., In-situ studies on the mechanical properties of He ion irradiated nanotwinned Ag. J. Nucl. Mater. **540**, 152392 (2020). https://doi.org/10.1016/j.jnucmat.2020.152392
24. Z. Yan, T. Yang, Y. Lin et al., Effects of temperature on helium bubble behaviour in Fe–9Cr alloy. J. Nucl. Mater. **532**, 152045 (2020). https://doi.org/10.1016/j.jnucmat.2020.152045
25. I. Beyerlein, M. Demkowicz, A. Misra et al., Defect-interface interactions. Prog. Mater Sci. **74**, 125–210 (2015). https://doi.org/10.1016/J.PMATSCI.2015.02.001
26. M.-S. Ding, J.-P. Du, L. Wan et al., Radiation-induced helium nanobubbles enhance ductility in submicron-sized single-crystalline copper. Nano Lett. **16**(7), 4118–4124 (2016). https://doi.org/10.1021/acs.nanolett.6b00864
27. O. El-Atwani, W. Cunningham, D. Perez et al., Temperature threshold for preferential bubble formation on grain boundaries in tungsten under in-situ helium irradiation. Scr. Mater. **180**, 6–10 (2020). https://doi.org/10.1016/j.scriptamat.2020.01.013
28. Q. Wei, N. Li, K. Sun et al., The shape of bubbles in He-implanted Cu and Au. Scr. Mater. **63**(4), 430–433 (2010). https://doi.org/10.1016/j.scriptamat.2010.04.043
29. R.E. Zartman, G. Wasserburg, J. Reynolds, Helium, argon, and carbon in some natural gases. J. Geophys. Res. **66**(1), 277–306 (1961). https://doi.org/10.1029/JZ066i001p00277
30. S. Xie, P.J. Tackley, Evolution of helium and argon isotopes in a convecting mantle. Phys. Earth Planet. Inter. **146**(3–4), 417–439 (2004). https://doi.org/10.1016/j.pepi.2004.04.003
31. Z.J. Bergstrom, D. Perez, O. El-Atwani, Helium bubble facetation in tungsten thin films. Scr. Mater. **220**, 114918 (2022). https://doi.org/10.1016/j.scriptamat.2022.114918
32. D. Morgan, G. Pilania, A. Couet et al., Machine learning in nuclear materials research. Curr. Opin. Solid State Mater. Sci. **26**(2), 100975 (2022). https://doi.org/10.1016/j.cossms.2021.100975
33. C.M. Anderson, J. Klein, H. Rajakumar et al., Automated detection of helium bubbles in irradiated X-750. Ultramicroscopy **217**, 113068 (2020). https://doi.org/10.1016/j.ultramic.2020.113068
34. R. Jacobs, P. Patki, M.J. Lynch et al., Materials swelling revealed through automated semantic segmentation of cavities in electron microscopy images. Sci. Rep. **13**(1), 5178 (2023). https://doi.org/10.1038/s41598-023-32454-2
35. Q. Chen, C. Zheng, Y. Cui et al., A deep learning model for automatic analysis of cavities in irradiated materials. Comput. Mater. Sci. **221**, 112073 (2023). https://doi.org/10.1016/j.commatsci.2023.112073
36. S. Liu, Q. Zhao, S. Han et al., Dynamics of catalyst nanoparticles quantified from in situ TEM video. Nano Today **59**, 102505 (2024). https://doi.org/10.1016/j.nantod.2024.102505
37. Z.-H. Wu, J.-J. Bai, D.-D. Zhang et al., Statistical analysis of helium bubbles in transmission electron microscopy images based on machine learning method. Nucl. Sci. Tech. **32**(5), 54 (2021). https://doi.org/10.1007/s41365-021-00886-y
38. Z. Zheng, S. Qiu, X. Yue et al., Detecting irradiation defects in materials: A machine learning approach to analyze helium bubble images. J. Nucl. Mater. **596**, 155117 (2024). https://doi.org/10.1016/j.jnucmat.2024.155117
39. S. Liu, H. Xu, D. Liu et al., Identify the activity origin of Pt single-atom catalyst via atom-by-atom counting. J. Am. Chem. Soc. **143**(37), 15243–15249 (2021). https://doi.org/10.1021/jacs.1c06381
40. J. Na, S.-J. Kim, H. Kim et al., A unified microstructure segmentation approach via human-in-the-loop machine learning. Acta Mater. **255**, 119086 (2023). https://doi.org/10.1016/j.actamat.2023.1190861