International Journal of Mechanical Sciences 315 (2026) 111424

Contents lists available at ScienceDirect

International Journal of Mechanical Sciences

journal homepage: www.elsevier.com/locate/ijmecsci

Deep learning characterization of stress-state-dependent anisotropic ductile damage

Zhichao Wei [a,*,1], Yijia Mo [a,1], Shuwei Zhou [a], Steffen Gerke [b], Sebastian Münstermann [a], Michael Brünig [b]

[a] Institute of Metal Forming, RWTH Aachen University, Intzestr. 10, D-52072 Aachen, Germany
[b] Institute for Engineering Mechanics and Structural Analysis, University of the Bundeswehr Munich, Werner-Heisenberg-Weg 39, D-85577 Neubiberg, Germany

ARTICLE INFO

Keywords:
Anisotropic damage mechanisms
Deep learning
Stress state
SEM images-driven
Biaxial experiments
Quantitative micro-defect characterization

ABSTRACT

Ductile damage results from the growth and coalescence of micro-voids and micro-shear-cracks, with a transition between these mechanisms. This paper investigates anisotropic ductile damage and its stress-state-dependent transition using scanning electron microscopy (SEM) images at the microscopic level combined with deep learning. Conventional optical observation, thresholding, and feature-based machine learning show apparent limitations. Deep learning with DeepLabv3 was evaluated using different backbone architectures on annotated SEM datasets of varying sizes. Comparisons with traditional recognition methods and different dataset sizes demonstrate the advantages of deep learning, particularly in resolving fine boundaries and discriminating defects under heterogeneous contrast. SEM data were obtained from fracture surfaces of novel biaxial experiments covering a wide range of stress states and used for training and validation. Numerical simulations were conducted to determine the stress states for the selected biaxial loading conditions. The transition of stress-state-dependent microscopic damage behavior was then investigated by combining numerically predicted stress states with deep learning-based recognition of micro-voids. Overall, deep learning provides an effective and efficient solution for quantitatively characterizing fracture-surface microstructures. The proposed workflow reduces annotation costs, operates on standard computing resources, and enables automated analysis of stress-state-dependent damage and transitions in ductile metals. The transition from micro-shear-crack-dominated to micro-void-dominated damage behavior is independently identified based on stress triaxiality, with a critical value of 0.38, and based on the Lode parameter, with a corresponding critical value of −0.44. The calibrated critical transition stress states provide a solid foundation for further development of the advanced continuum damage model.

1. Introduction

Lightweight ductile metals are widely used in engineering applications, such as automobiles, high-speed trains, and aerospace structures, owing to their high strength and reduced weight, which enhances the performance of transport vehicles. Engineering structures subjected to different complex loading conditions exhibit elastic–plastic–damage behavior, where irreversible plastic deformation initiates damage evolution and leads to damage-induced softening. Moreover, ductile damage and fracture behavior are known to be stress-state-dependent [1–13]. For example, under a tensile-dominated stress state (see Fig. 1(a)), damage is caused by the nucleation, growth, and coalescence of micro-voids, whereas under shear or compressive stress states, it results from the growth and coalescence of micro-shear-cracks (see Fig. 1(b)). For mixed tensile-shear stress states, both mechanisms contribute to the damage process. The stress-state-dependent evolution of damage mechanisms during loading is referred to as anisotropic damage, as noted by Murakami [14], Brünig [15], and Brünig et al. [16]. It should be noted that the anisotropic damage behavior may also be caused by anisotropic material properties resulting from different crystallographic structures [17–24]. As these defects evolve, the effective load-bearing area decreases, leading to a reduction in mechanical properties. Because the current stress state governs this evolution, it is inherently anisotropic.

Scanning electron microscope (SEM) analysis of the fracture surface is an effective method for identifying anisotropic damage mechanisms. However, despite its widespread use, conventional optical observation

* Corresponding author.
E-mail address: zhichao.wei@ibf.rwth-aachen.de (Z. Wei).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ijmecsci.2026.111424
Received 2 December 2025; Received in revised form 17 February 2026; Accepted 18 February 2026
Available online 19 February 2026
0020-7403/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

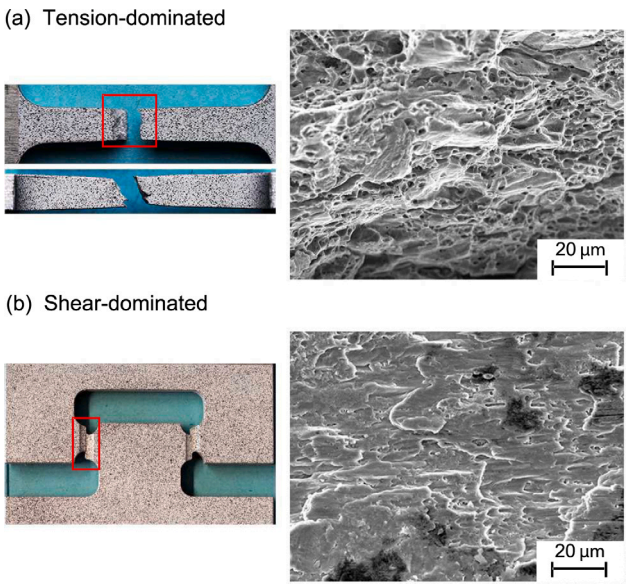

Fig. 1. Fracture morphologies under tension-dominated and shear-dominated loading conditions: (a) macroscopic fracture photograph with the corresponding SEM image taken from the region highlighted by the red box, showing a micro-void-dominated fracture surface under tension-dominated loading, and (b) macroscopic fracture photograph and SEM image taken from the red-boxed region, exhibiting micro-shear-crack-dominated fracture features under shear loading. All SEM images are shown at the same magnification with a scale bar of 20 μm.

makes it difficult to determine whether damage evolution under mixed tensile-shear stress states is induced by micro-voids or micro-shear-cracks. This difficulty becomes more pronounced under complex loading conditions, where variations in cyclic loading sequences, patterns, or amplitudes significantly alter micro-defect evolution, leading to changes in both micro-crack length [25] and micro-void porosity [26]. Although advanced techniques such as 3D in-situ laminography provide detailed insights into void growth and shear-localized damage [27], their experimental complexity limits large-scale application.

In contrast, quantitative fractography coupled with machine learning has also been shown to predict macroscopic properties, such as fracture toughness, or to distinguish brittle and ductile failure regimes directly from fracture surfaces [28,29]. This motivates the development of efficient methods to quantitatively identify stress-state-dependent damage mechanisms at the microscopic level using readily available SEM data. Moreover, accurate quantitative ductile damage analysis enables the validation of numerical predictions of damage and fracture using advanced material models under different engineering loading conditions, contributing to a more reliable characterization of material behavior and more efficient use of structural materials.

To quantitatively evaluate damage mechanisms under different loading conditions, image-based analysis can be used to segment micro-void regions and measure crack lengths from fracture-surface SEM images. Early image analysis relied on rule-based tools such as ImageJ [30], which extract handcrafted features from grayscale images, as exemplified by threshold-based porosity quantification methods [31]. For SEM-based datasets, Joos et al. [32] applied an Otsu-based segmentation workflow [33] to FIB-SEM images, enabling automated pore-solid separation and porosity quantification. However, such global thresholding approaches depend primarily on grayscale distributions and often assume bimodality, which makes them sensitive to heterogeneous contrast and fuzzy boundaries commonly observed in fracture-surface SEM images [34].

Supervised learning has been increasingly used for SEM segmentation, ranging from feature-engineered classifiers such as random forest [35] to end-to-end CNN-based models. By leveraging spatial and textural information, these approaches typically outperform global thresholding and provide probabilistic outputs that facilitate operating-point selection and class-imbalance handling through sampling or class weighting [36]. Related workflows have been reported for FIB/SEM phase segmentation and fracture-morphology grouping [37–39].

Feature-engineered machine learning (ML) improves flexibility over rule-based segmentation, but it typically classifies pixels or patches using precomputed descriptors. As a result, performance can be sensitive to feature design, acquisition-related artifacts, and domain shift, and it may struggle to capture long-range spatial context that is important in high-resolution SEM imagery. In addition, computational cost and memory demand increase rapidly with image resolution and dataset size, which can become a practical bottleneck for large SEM datasets.

In contrast, deep learning performs dense, end-to-end semantic segmentation using convolutional neural networks (CNNs) [40,41], learning task-specific representations directly from data. Modern architectures such as FCNs, U-Net, and the DeepLab family [42–45] combine multi-scale context aggregation with fine-scale boundary recovery and have demonstrated strong performance on heterogeneous SEM microstructure datasets. Recent studies have applied DeepLabv3+ to mudrock SEM segmentation [46] and U-Net/FCN models to FIB-SEM datasets of complex ceramics for pore segmentation [47]. Similar deep networks have also been successfully deployed across SEM and X-ray microstructure segmentation tasks in concrete, metals, and rocks [48–53]. Two-stage deep neural networks have been used to identify and quantify precipitate areas in chromium-based alloys, further illustrating the capability of deep models for high-throughput microstructural quantification [54]. Beyond segmentation, data-driven deep learning has been explored for predicting void growth in aluminum under monotonic tension [55]. Furthermore, CNNs have also been widely used in engineering applications to address different analysis problems [56–64]. These capabilities are particularly relevant to fracture-surface SEM images, where micro-voids and micro-shear-cracks can overlap in gray levels and exhibit heterogeneous local contrast.

To study stress-state-dependent anisotropic damage behavior, biaxial or multiaxial experiments are an effective way to generate a wide range of stress states. Multiaxial experiments are typically performed using tension-torsion cylindrical specimens [65–67], whereas biaxial experiments [25,68–70] are conducted with cruciform thin metal sheets. Since sheet metals are widely used in automobiles and high-speed trains, biaxial experiments can more accurately capture material behavior, as specimen geometry can influence the observed response. In recent decades, the lightweight aluminum alloy has been extensively studied by Brünig's research group [25,26,71–80], resulting in a large number of experimental datasets, including force–displacement curves and strain fields measured with digital image techniques and SEM images. To systematically analyze the stress-state-dependent microstructural behavior induced by different loading histories, quantitative analysis of SEM images is required. Traditional image processing, machine learning, and deep learning approaches have been applied to such images, yet their performance differs significantly. Therefore, a detailed comparison of these methods is essential to identify their respective advantages and limitations. In this work, the DeepLabv3 [44] model, incorporating atrous convolutions and an Atrous Spatial Pyramid Pooling (ASPP) module, is employed as the representative deep learning framework. The network is evaluated with different backbone architectures on annotated SEM datasets of varying sizes, enabling a systematic assessment of its performance in resolving fine boundaries and capturing microstructural features under complex stress states. In terms of numerical methods, the stress states obtained from numerical calculations are combined with micro-defect properties extracted from SEM images, enabling a comprehensive investigation of stress-state-dependent damage behavior.

This paper is structured as follows. The investigated materials and the corresponding testing methods are briefly introduced in Section 2.

Table 1
The chemical composition of EN-AW 6082-T6 (AlSiMgMn) in weight percent.

Al	Si	Mg	Mn	Fe
to balance	0.90%	0.70%	0.47%	0.37%
	Cu	**Zn**	**Ti**	**Others**
	0.09%	0.09%	0.03%	0.06%

Section 3 summarizes the constitutive model used in the numerical simulations to calculate the stress states for different loading conditions. The methodology of the different image recognition methods is described in Section 4. The image segmentation implementation and neural network parameter setups are discussed in Section 5. The evaluation of the deep learning network is provided in Section 6. Based on the trained deep learning network, a detailed analysis of the SEM images for different biaxial loading experiments is shown in Section 7. Finally, the conclusion is provided in Section 8.

2. Materials and experiments

In this section, the investigated ductile aluminum alloy EN AW-6082-T6 is described, followed by the definition of the relevant stress states. Finally, the experimental setups, including the biaxial loading experiments and the SEM characterization, are introduced.

2.1. Material composition and microstructure

The ductile aluminum alloy EN-AW 6082-T6 offers high specific strength and good weldability, making it widely used in structural, automotive, and high-speed train applications. The chemical composition is summarized in Table 1.

Furthermore, the microstructure of the alloy was statistically characterized by Electron Backscatter Diffraction (EBSD). The EBSD results showed no significant texture in the investigated material, as shown in Fig. 2. The corresponding pole figures reveal only a weak preferential alignment along the ⟨100⟩ orientation, with a maximum texture intensity of approximately twice that of a random distribution. This low intensity indicates that the overall texture is weak. Consistent with this observation, uniaxial tensile tests performed along different loading directions show negligible anisotropy in the mechanical response, confirming that the weak crystallographic texture has minimal influence on the macroscopic mechanical behavior.

2.2. Experimental setups

To generate different stress states, biaxially loaded specimens, as illustrated in Fig. A.21 and proposed by Gerke et al. [69] and Wei et al. [79], can cover a wide range of conditions. As shown in Fig. A.21(a), the X0-specimen has bi-symmetric notches around its notch section, which produce identical stress states for loads applied to the horizontal and vertical axes. The H-specimen and HC-specimen have different geometries compared to the X0-specimen, where tensile/compressive and shear stress states are obtained by applying loads along the horizontal and vertical axes, respectively. Moreover, the only difference between the H- and HC-specimens is the central opening: in the HC-specimen, the larger distance between the vertical and horizontal axes allows cyclic loading conditions to be performed. Therefore, it is evident that different load combinations along the vertical and horizontal axes result in distinct stress states. Based on the experimental and numerical analysis of the different loading cases [69,79], the proposed biaxial experiments cover stress triaxiality values from −0.6 to 0.8, as shown in Fig. 3.

The stress triaxiality is defined as the ratio of the mean stress to the von Mises equivalent stress

$$\eta = \frac{\sigma_m}{\sigma_{eq}} = \frac{I_1}{3\sqrt{3J_2}} \tag{1}$$

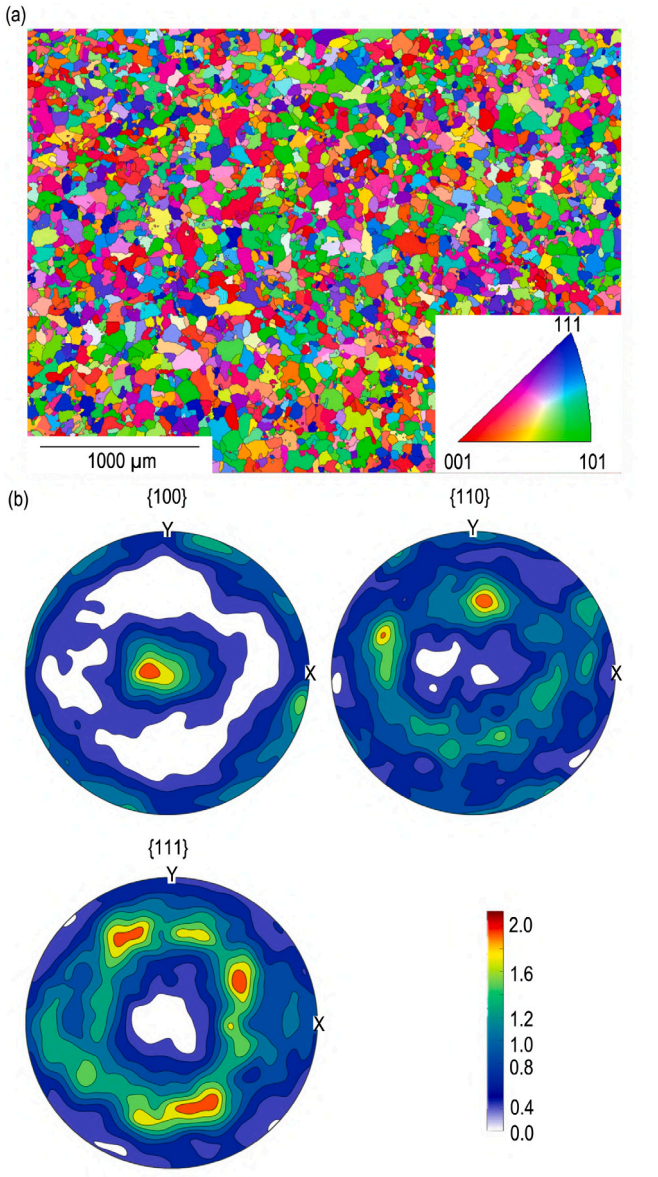

Fig. 2. EBSD analysis results of the investigated aluminum alloy EN AW-6082-T6: (a) inverse pole figure (IPF) map showing grain morphology and crystallographic orientations, and (b) {100}, {110}, and {111} pole figures, illustrating the crystallographic texture of the material calculated from the grain orientations. The color scale represents the orientation density relative to a random distribution.

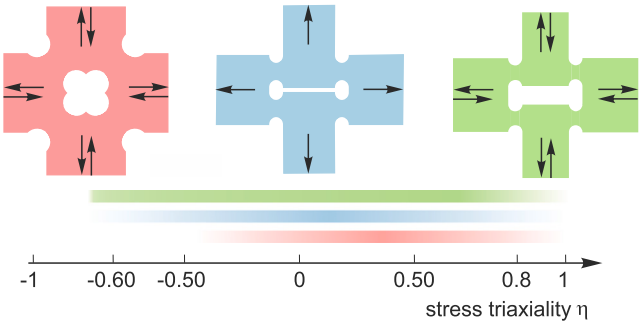

Fig. 3. Attainable stress triaxiality η for the biaxially loaded X0-, H-, and HC-specimens (red, blue, and green, respectively) under different combinations of biaxial loading.

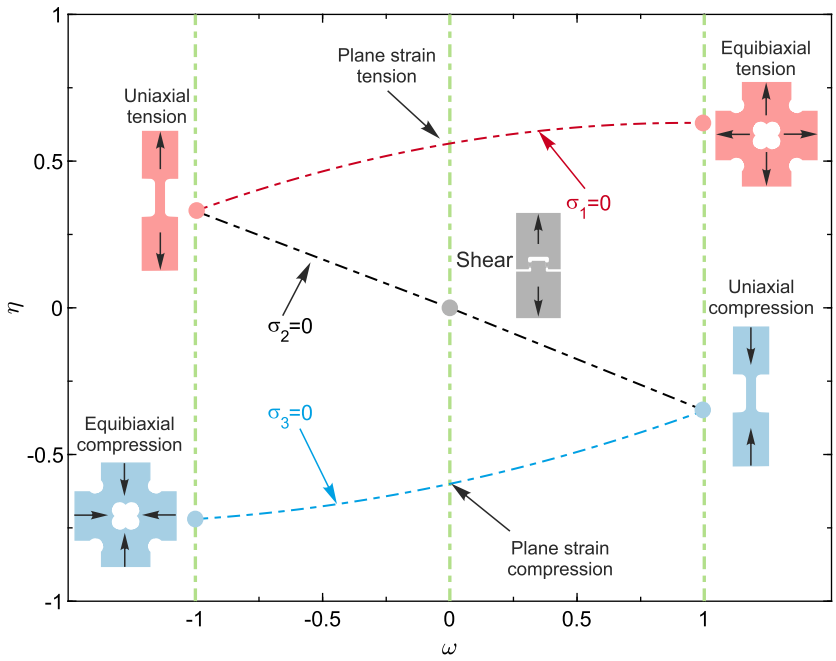

Fig. 4. Stress states in the space of stress triaxiality η and the Lode parameter ω. Representative loading conditions, including uniaxial and equibiaxial tension and compression as well as shear, are illustrated in the η–ω space. The red, black, and blue dashed lines represent loading paths corresponding to $\sigma_1 = 0$, $\sigma_2 = 0$, and $\sigma_3 = 0$, respectively.

with

$$I_1 = \mathrm{tr}(\mathbf{T}) = \sigma_1 + \sigma_2 + \sigma_3 \tag{2}$$

and

$$J_2 = \frac{1}{2}\mathrm{dev}\mathbf{T} \cdot \mathrm{dev}\mathbf{T} = \sqrt{\frac{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2}{6}}, \tag{3}$$

where $\mathbf{T}$ denotes the Kirchhoff stress tensor, and σ_1, σ_2, and σ_3 are the principal stresses. It is evident that stress triaxiality can be used to identify the loading case. For example, under compressive and tensile loading, $\eta = \pm 1/3$, whereas for shear loading, $\eta = 0$. In addition, the Lode parameter ω is an important parameter, defined as a function of the third stress invariant of the stress deviator. It is effectively used to distinguish stress states in three dimensions, particularly under low positive and shear stress states. It can be calculated from the principal stresses $\sigma_1 \geq \sigma_2 \geq \sigma_3$ in the form of

$$\omega = \frac{2\sigma_2 - \sigma_1 - \sigma_3}{\sigma_1 - \sigma_3}. \tag{4}$$

The stress states in the space of stress triaxiality η and the Lode parameter ω are illustrated in Fig. 4. Thus, stress triaxiality and the Lode parameter can effectively distinguish different loading conditions, and both are key parameters in ductile damage and fracture behavior.

In this work, a series of proportional and non-proportional loading experiments were conducted to study the dependence of the stress state on damage and fracture behavior. All experiments were carried out using a horizontally arranged electro-mechanical testing machine, LFM-BIAX 20 kN, manufactured by Walter + Bai AG, Switzerland, see Fig. 5. The loading speed under static loading was 0.004 mm/s. Further details of the experimental setups for biaxial material testing can be found in [79–81]. After the experiments, the fracture surfaces were examined using a SEM equipped with a ZEISS EVO 15, featuring a lanthanum hexaboride (LaB6) electron emitter and an Everhart–Thornley (ET) detector, to investigate damage and fracture behavior at the microscopic level. Imaging was performed in a high-vacuum environment, maintaining a working distance of about 11 mm, an electron beam diameter

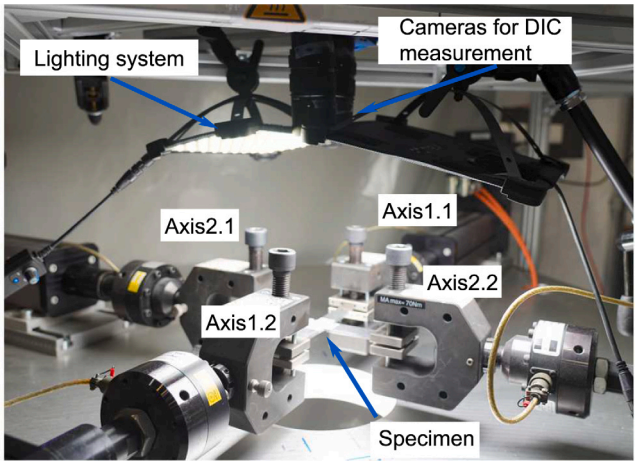

Fig. 5. Experimental setup for the designed biaxial experiments: DIC cameras and the lighting system are used to measure specimen deformations, and the LFM-BIAX testing machine, with a capacity of 20 kN, is equipped with four loading axes.

(spot size) of 300 nm, and an accelerating voltage of 20 kV. As illustrated in the overview of the fracture surface in Fig. 6, three positions (left, middle, and right) were selected as representative locations for the SEM features in each loading case. It can be observed that the selected positions differ only slightly from the visual observations, and in the respective cases the damage is caused by the growth and coalescence of micro-voids.

3. Constitutive modeling

Since stress states cannot be directly measured during experiments, numerical simulations are required to calculate the corresponding stress triaxiality and Lode parameter. To accurately capture the material response, the advanced plastic–damage model proposed in [75,76] is

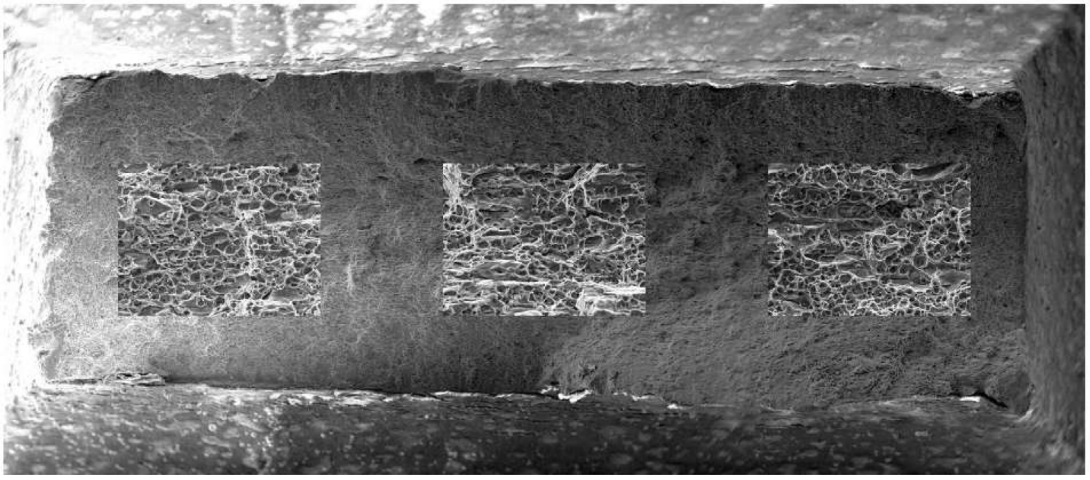

Fig. 6. Overview of the representative fracture surface at the notch of the tested specimen. The zoomed SEM images indicate representative locations on the fracture surface; their positions and sizes are shown schematically and do not correspond to the exact locations or a 1:1 scale.

employed. In this section, only the most important plastic and damage conditions are summarized; for further details, please refer to [75,76].

The Drucker–Prager yield condition [75,82] is used to characterize the material behavior

$$f^{\mathrm{pl}} = \sqrt{\frac{1}{2}\mathrm{dev}(\bar{\mathbf{T}} - \bar{\boldsymbol{\alpha}})\cdot\mathrm{dev}(\bar{\mathbf{T}} - \bar{\boldsymbol{\alpha}})} - \bar{c}\left(1 - \frac{a}{\bar{c}}\,\mathrm{tr}(\bar{\mathbf{T}} - \bar{\boldsymbol{\alpha}})\right) = 0\,, \tag{5}$$

where $\bar{\mathbf{T}}$ denotes the effective Kirchhoff stress tensor, $\bar{\boldsymbol{\alpha}}$ describes the effective back stress tensor, $a/\bar{c}$ is the hydrostatic coefficient.

The non-associated plastic flow is used to characterize the evolution of effective plastic strain

$$\dot{\bar{\mathbf{H}}}^{\mathrm{pl}} = \dot{\lambda}\frac{\partial g^{\mathrm{pl}}}{\partial(\bar{\mathbf{T}} - \bar{\boldsymbol{\alpha}})} = \dot{\lambda}\frac{1}{2\sqrt{\bar{J}_2}}\,\mathrm{dev}(\bar{\mathbf{T}} - \bar{\boldsymbol{\alpha}}) \tag{6}$$

with

$$g^{\mathrm{pl}} = \sqrt{\frac{1}{2}\mathrm{dev}(\bar{\mathbf{T}} - \bar{\boldsymbol{\alpha}})\cdot\mathrm{dev}(\bar{\mathbf{T}} - \bar{\boldsymbol{\alpha}})}\,, \tag{7}$$

where $\dot{\lambda}$ denotes the non-negative plastic multiplier.

Moreover, the damage condition is given by

$$f^{\mathrm{da}} = \hat{\alpha}(\eta,\omega)\,\mathrm{tr}(\mathbf{T}) + \hat{\beta}(\eta,\omega)\sqrt{\frac{1}{2}\mathrm{dev}\mathbf{T}\cdot\mathrm{dev}\mathbf{T}} - \tilde{\sigma} = 0\,, \tag{8}$$

where $\hat{\alpha}(\eta,\omega)$ and $\hat{\beta}(\eta,\omega)$ are stress-state-dependent variables, $\mathbf{T}$ denotes the Kirchhoff stress tensor, and $\tilde{\sigma}$ represents the initial equivalent damage stress. For tension-dominated stress states, damage is governed primarily by micro-void growth and is therefore controlled mainly by the hydrostatic stress component, i.e., $\mathrm{tr}(\mathbf{T})$. In contrast, for shear-dominated stress states, damage is driven by deviatoric deformation of micro-defects, as the hydrostatic stress vanishes, i.e., $\hat{\alpha}\,\mathrm{tr}(\mathbf{T}) = 0$. In pure shear loading, the transition is clearly associated with zero stress triaxiality and Lode parameter. However, for mixed tension–shear stress states, the corresponding transition value is not obvious and requires further identification, which motivates the present study. Although Brünig et al. [16], based on macroscopic experiments and numerical simulations of uniaxial tensile tests with different notch radii, suggested a transition from shear-dominated to tensile-dominated behavior at a stress triaxiality of $1/\sqrt{3}$, this transition value is primarily inferred from macroscopic experimental and numerical analyses. It lacks direct experimental validation at the microscopic scale. The details of the formulation proposed by Brünig et al. [16] are shown in Appendix B.

The damage strain rate tensor captures the evolution of micro-defect deformation under different damage mechanisms and is defined as

$$\dot{\mathbf{H}}^{\mathrm{da}} = \dot{\mu}\left(\tilde{\alpha}(\eta,\omega)\frac{1}{\sqrt{3}}\mathbf{1} + \tilde{\beta}(\eta,\omega)\tilde{\mathbf{N}}\right), \tag{9}$$

where $\dot{\mu}$ is the rate of the equivalent damage strain, $\tilde{\alpha}(\eta,\omega)$ and $\tilde{\beta}(\eta,\omega)$ are stress-state-dependent variables, $\mathbf{1}$ denotes the identity tensor, and $\tilde{\mathbf{N}}$ represents the transformed deviatoric stress tensor. Similar to the damage condition in Eq. (8), a critical value is required to describe the transition from shear-dominated to tension-dominated damage mechanisms in the formulation of different damage evolution equations.

It should be noted that the parameters $\hat{\alpha}$ and $\hat{\beta}$ in Eq. (8), as well as $\tilde{\alpha}$ and $\tilde{\beta}$ in Eq. (9), are identified through numerical simulations, which provide reasonable results and enable an accurate characterization of the stress states discussed in the following analysis. This work aims to identify stress-state-dependent damage transitions, based on microscopic experimental evidence, that can be used to modify further and refine these parameters.

4. Methodology

This section reviews the segmentation methodologies considered in this study, from traditional thresholding to feature-based machine learning and deep learning. Each approach is outlined with its underlying principle, advantages, and limitations, followed by the rationale for method selection.

4.1. Traditional image processing

Thresholding is the simplest and most widely used method of segmentation in materials characterization. An example of the thresholding technique is shown in Fig. 7. In this study, ImageJ was used as a baseline implementation. After grayscale conversion, Otsu's method [33] automatically determined an intensity threshold that separates foreground objects like micro-voids from the background. This technique is computationally efficient and requires no prior training data. However, because it relies on global intensity statistics, it often fails in fracture-surface SEM images where micro-voids overlap with micro-shear-cracks and local contrast varies significantly.

4.2. Feature-based machine learning

To address the limitations of global thresholding, machine learning approaches typically integrate preprocessing, handcrafted descriptors, and supervised classifiers. Images are first standardized by normalization and smoothing, after which a multi-cue descriptor bank is applied. These descriptors include local intensity statistics, edge and gradient measures, morphological filters, and texture operators such as Gabor

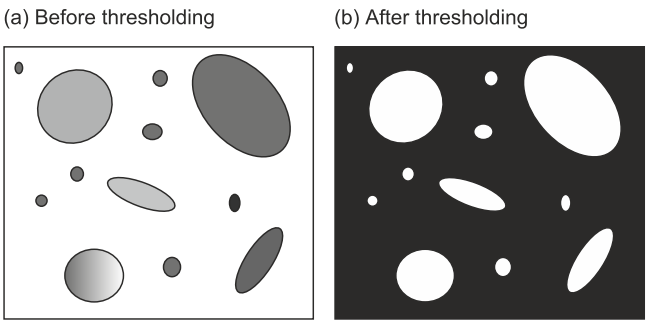

Fig. 7. Example of threshold-based image segmentation using Otsu's method: (a) original grayscale SEM image containing microstructural features, and (b) binary segmentation mask obtained after thresholding, where white regions correspond to segmented features and black regions represent the background.

filters and local binary patterns. For each pixel, the descriptors are concatenated into a feature vector, producing a design matrix

$$\mathbf{X} \in \mathbb{R}^{HW \times C}, \tag{10}$$

where $H \times W$ is the image size and C the number of feature channels. The corresponding ground truth $\mathbf{y} \in \{0,1\}^{HW}$ is obtained from manually annotated binary masks, indicating for each pixel whether it belongs to a micro-void or to the background.

A random forest classifier is then trained on these data to predict the per-pixel probability of belonging to the micro-void class. These probabilities are thresholded to produce binary masks, and simple morphological post-processing may be applied to refine the boundaries.

4.3. Deep learning for image segmentation

Deep learning eliminates the need for handcrafted descriptors by learning hierarchical features directly from raw images. CNNs extract low-level patterns (edges, textures) in early layers and progressively build higher-level representations (shapes, regions) in deeper layers. Encoder–decoder architectures such as U-Net [43] combine fine spatial details with semantic context, producing sharper segmentation boundaries. Variants like DeepLab [44,45] further extend the receptive field using dilated (atrous) convolutions and pyramid pooling, enabling robust multi-scale analysis. Transfer learning from large-scale datasets such as ImageNet [83] accelerates training and improves generalization, even for grayscale SEM data.

In this study, DeepLabv3 is adopted for micro-void segmentation. Fig. 8 shows its main components. First, the backbone network acts as a feature extractor, converting SEM images into multi-level feature maps that capture both local edges and broader patterns. This is analogous to automatically detecting descriptors of micro-voids and micro-shear-crack structures, without manually designing filters. Next, the ASPP module applies convolutions with multiple dilation rates (6, 12, 18) and global pooling in parallel. This step enables the network to capture features at various spatial scales, ranging from the fine details of small voids to larger surrounding crack patterns. Finally, the prediction head combines these outputs to produce dense, per-pixel probability maps. After thresholding, the probability maps yield binary masks delineating micro-voids, which can then be used for quantitative morphometric analysis.

4.4. Summary and method selection

In summary, thresholding offers simplicity but lacks robustness under heterogeneous contrast. Feature-based machine learning improves adaptability but depends on handcrafted descriptors. Deep learning, represented by DeepLabv3, learns multi-scale features directly and achieves the most reliable performance, albeit at a higher cost.

For this reason, the present study adopts Otsu's thresholding as a baseline, a random forest classifier as a representative of feature-based machine learning, and DeepLabv3 with different backbones and ASPP module as the deep learning framework for systematic comparison. This selection covers the range from rule-based to modern data-driven approaches.

5. Image segmentation implementation and neural network parameters setup

Based on the methodological overview in Section 4, this section focuses on the practical implementation of the selected segmentation strategies. The aim is to provide a reproducible description of dataset preparation, software environment, algorithmic pipelines, and parameter settings. The implementations are organized in increasing order of algorithmic complexity: (i) baseline thresholding in ImageJ, (ii) random forest classifier with handcrafted descriptors, and (iii) DeepLabv3 as a representative deep neural network.

5.1. Baseline segmentation with ImageJ

For baseline comparison, Otsu thresholding was applied using the built-in functions in ImageJ. SEM images were first converted into 8-bit grayscale and the global threshold was computed automatically. The resulting binary masks were used to calculate micro-void area fractions within user-defined regions of interest, serving as a simple, unsupervised reference for subsequent approaches.

5.2. Dataset preparation and annotation

To enable supervised learning for the machine learning and deep learning approaches, the original SEM images (3072 × 2304 pixels) were partitioned into 480 × 480 patches, which means 24 small cropped images for each SEM image. This patching strategy reduces memory usage, accelerates annotation, and suppresses peripheral artifacts such as scale bars. The LabelMe tool [84] was used to create polygon annotations, assisted by EfficientSAM [85] for semi-automatic boundary generation. Annotations were exported in JSON format and converted into binary masks. The resulting paired image-mask dataset was used consistently across all supervised segmentation methods. A schematic of the data preparation workflow is shown in Fig. 9.

5.3. Feature-based random forest pipeline

Following the dataset preparation, handcrafted descriptors were extracted from each grayscale patch, including intensity statistics, gradients, morphological filters, and texture operators. The specific filter bank and parameter ranges are summarized in Table 2. The stacked feature maps were reshaped into a design matrix and aligned with pixel-level labels from the masks.

Training used stratified pixel sampling to handle class imbalance under memory limits. A random forest with 200 trees was fitted using class-weight balancing and parallelized execution. A 900 annotated images dataset was randomly assigned by hand to a single 76/24 train/validation split, this split was made once before training, froze the file lists, and reused them for all experiments. To avoid data leakage, each image and its mask appear in only one split. The validation set was strictly held out, which was not used to fit the model, or to tune the hyperparameters of the network. During inference, the same preprocessing and feature extraction were applied, and the trained model produced per-pixel probability maps. A validated operating threshold was applied to generate binary masks, with optional morphological post-processing for boundary refinement. The full workflow is summarized in Fig. 10.

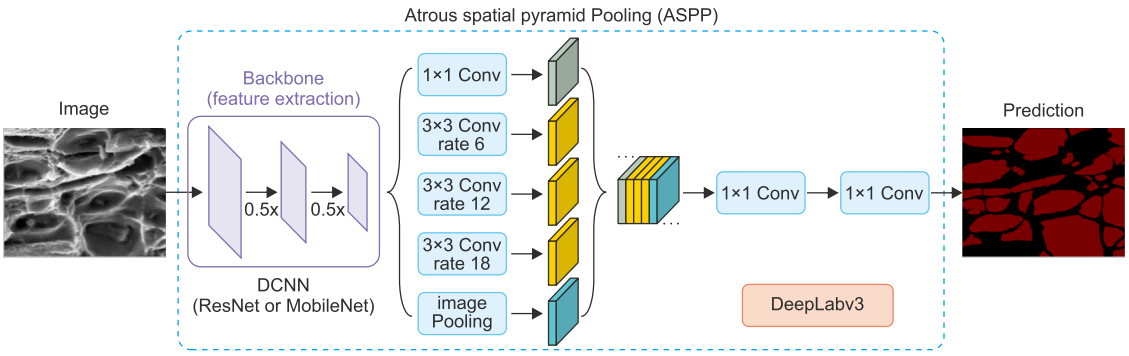

Fig. 8. Schematic overview of the DeepLabv3 architecture used in this study. The input SEM image is first processed by a ResNet or MobileNet backbone to extract features, which are subsequently refined by an atrous spatial pyramid pooling (ASPP) module to capture multi-scale information. The resulting features are used to generate dense per-pixel predictions for robust segmentation under heterogeneous contrast conditions.

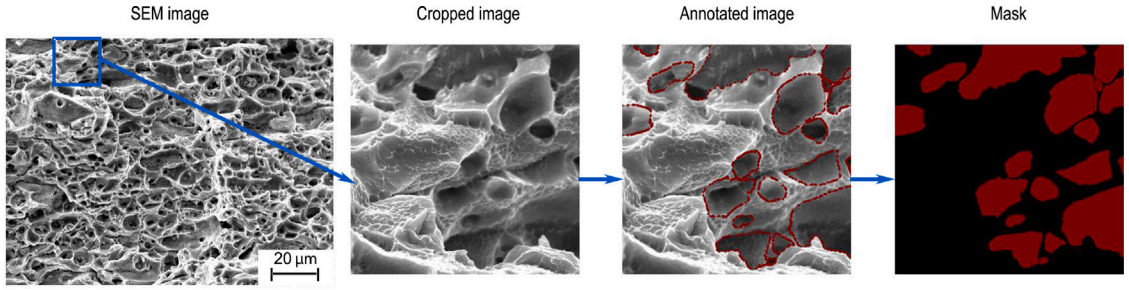

Fig. 9. A schematic example of the image preparation process used to construct the training dataset. An original SEM image is cropped into several small images. These small images are manually annotated using red polygons to identify micro-voids. The annotations are then converted to binary masks, with red regions indicating the identified micro-void area.

Table 2
Summary of feature extraction parameters used in the descriptor bank.

Feature type	Description	Parameters
Intensity	Raw grayscale intensity	–
Smoothing	Gaussian smoothing	$G_\sigma \in \{1, 2, 4\}$
Local statistics	Mean, standard deviation	5×5 window
First-order derivatives	Sobel/Scharr gradients	x, y directions
Gradient features	Magnitude, orientation	Normalized orientation components
Second-order derivatives	Laplacian	Standard kernel
Multi-scale edges	Laplacian-of-Gaussian	$G_\sigma \in \{1, 2, 4\}$
DoG	Difference-of-Gaussians	$(1-2), (2-4)$
Hessian	Approx. Hessian eigenvalues	Ridge/valley detection
Morphology	Top-hat, black-hat	Elliptical kernel
Texture (Gabor)	Multi-orientation, multi-scale	Orientations: $0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$; Wavelengths: $\{4, 8, 16, 24\}$; $G_\sigma \approx 0.56\ell$, ρ=0.5, kernel size 9
Texture (LBP)	Local binary pattern	Radius = 1, 8-neighborhood

5.4. Deep learning-based segmentation

Deep learning experiments were carried out using the PyTorch implementation of DeepLabv3. Paired images and masks were loaded through a custom dataset class, with joint data augmentation (flips, normalization) applied during training. Several backbones were tested (ResNet-50/101 [83], MobileNetV3_Large [86]), all configured for two-class segmentation of background and micro-voids class.

Training employed cross-entropy loss optimized with Adam optimizer at a learning rate of 10^{-4} for 200 epochs. Mixed-precision training was used to improve memory efficiency. Table 3 lists the key hyperparameters, which were chosen following established conventions and constrained by available GPU memory. These values were subsequently held fixed across all experiments, and the validation set was used exclusively for performance evaluation rather than hyperparameter tuning.

Table 3
Training hyperparameters for the DeepLabv3 model.

Batch size	4
Learning rate	10^{-4}
Training epochs	200
Optimizer	Adam
Loss function	Cross-entropy
Backbone	ResNet-50
Initialization	Random
Precision	Mixed precision
Train/validation split	76/24

Performance was monitored via pixel accuracy, Dice score, precision, and recall on a held-out validation set. Precision quantifies

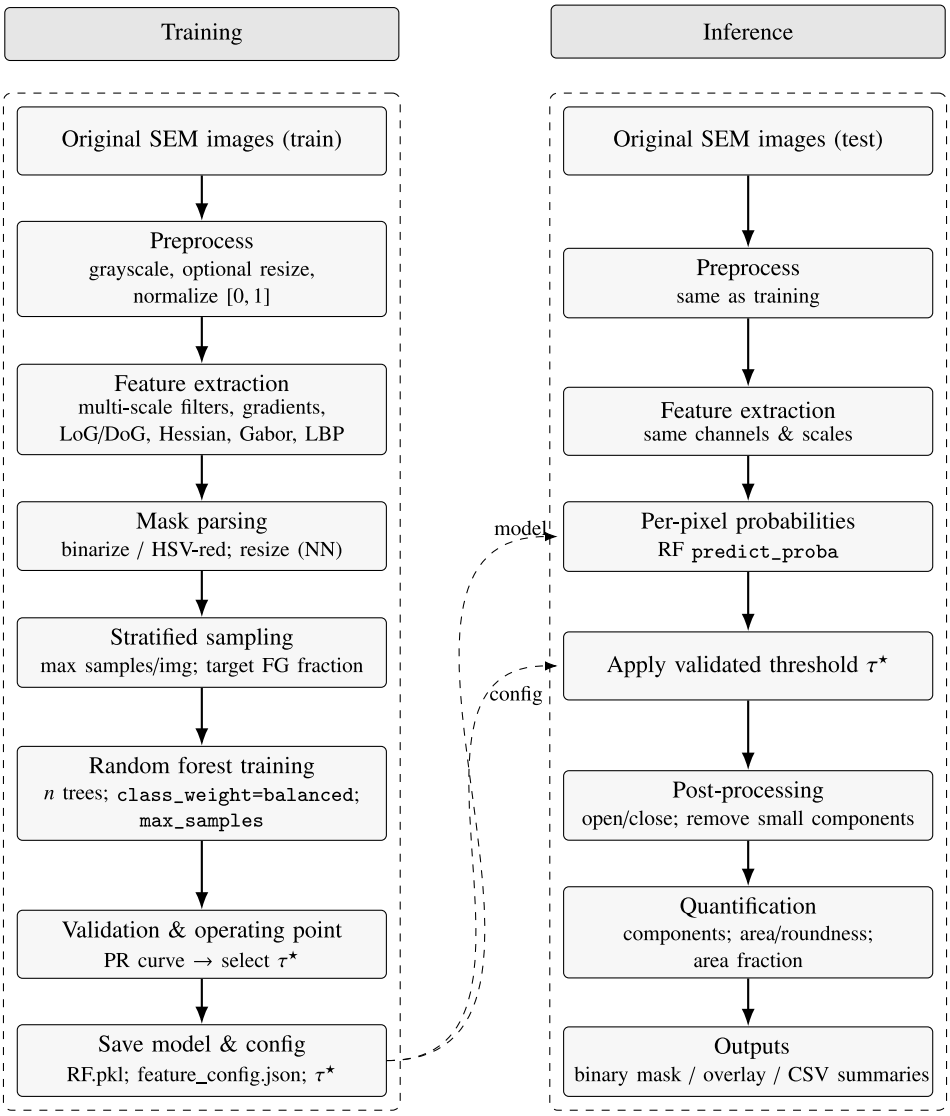

Fig. 10. Flowchart of the traditional machine learning pipeline for SEM image segmentation. The training stage (left) shows image preprocessing, hand-crafted feature extraction, random forest (RF) training, and selection of an operating threshold $\tau^\star$. The trained hyperparameter configuration is saved for future use. The inference stage (right) applies the same preprocessing and feature extraction process to unseen SEM images. The RF classifier then uses the saved configuration to generate binary prediction masks, after which quantitative micro-void statistics are performed.

the proportion of correctly identified target pixels among all pixels predicted as belonging to the target class, reflecting the model's ability to avoid false detections. Recall, or sensitivity, measures the proportion of target pixels in the ground truth that are successfully detected by the model, indicating how completely the relevant structures are captured. The Dice coefficient combines both precision and recall into a single measure of overlap between the predicted and reference masks, with higher values corresponding to greater segmentation accuracy and fewer false positives or missed regions. Training and validation curves, as well as confusion matrices, were recorded for reproducibility.

To assess model robustness and mitigate variance arising from dataset partitioning, a five-fold cross-validation (5-fold CV) protocol was implemented. In each fold, the dataset was split into training and validation subsets while reserving a fixed 15% portion as an independent test set. For inference, SEM images were normalized consistently with the training phase. Predicted masks were overlaid on the original images for visualization, and quantitative morphometric measures (e.g., void area fraction, particle roundness, connected-component statistics) were extracted in physical units.

To characterize micro-shear-crack features complementary to micro-void segmentation, Canny edge detection was applied to the background region (i.e., excluding predicted void pixels). The procedure included gap bridging, skeletonization, and length estimation with $\sqrt{2}$-weighted diagonals, with optional pruning of spurious branches. For each connected component, both individual lengths and aggregated statistics (mean, median, total, maximum) were computed. Results were reported in physical units (μm) when pixel size information was available, and visual overlays were generated for qualitative assessment.

6. Deep learning network evaluation

In this section, the study of different backbones and cross-validation performance is described. Then, a comprehensive comparison of different image segmentation methods is discussed, and the influence of data size is also shown.

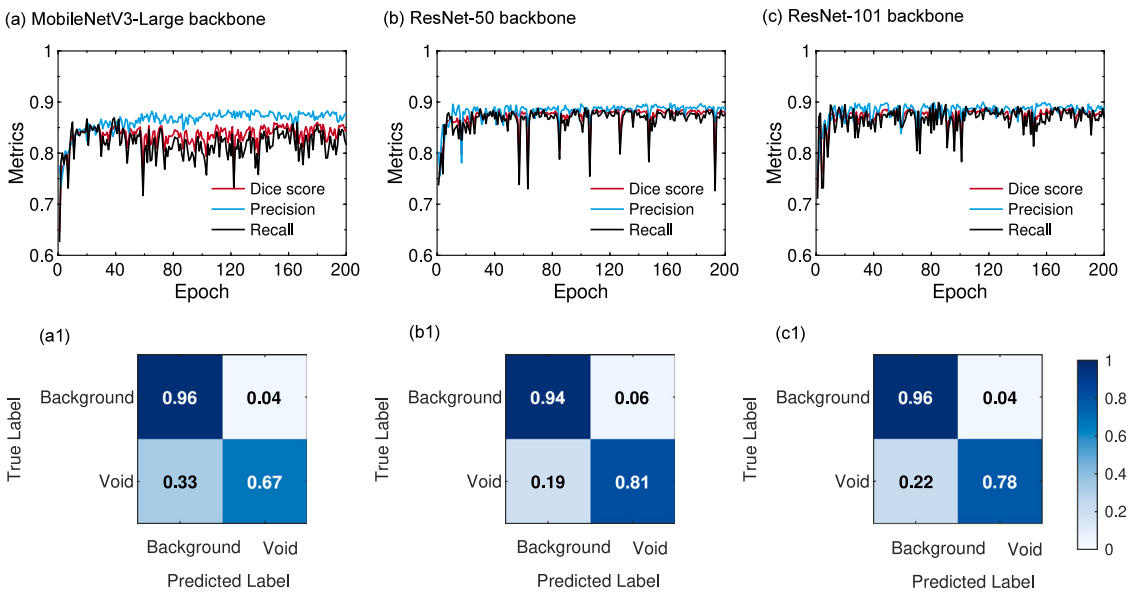

Fig. 11. Segmentation performance metrics and confusion matrices for models trained with different backbones: (a)–(c) the relationship between the Dice score, precision, and recall versus training epoch for the MobileNetV3-Large, ResNet-50, and ResNet-101 backbones, respectively, and (a1)–(c1) the corresponding confusion matrices for the background and void classes.

6.1. Backbone comparison

To investigate the influence of backbone depth and architecture on segmentation performance, three DeepLabv3 variants were trained using MobileNetV3-Large, ResNet-50, and ResNet-101 backbones. All trainings were performed under identical hyperparameter settings, differing only in backbone complexity. The corresponding training times were approximately 2.5 h, 5 h, and 6.5 h, respectively. All trainings were conducted on a laptop equipped with an AMD Ryzen 9 5900HS CPU and an NVIDIA GeForce RTX 3050 Ti Laptop GPU with 4 GB of VRAM. The software environment was based on Python 3.12.7 and PyTorch 2.7.1.

The training and validation accuracy curves (Fig. C.23) reveal apparent differences in convergence dynamics across the three backbones. MobileNetV3-Large converges rapidly within the first 20 epochs but plateaus early, suggesting a limited capacity for fine-grained features learning. ResNet-50 exhibits smoother convergence and achieves higher validation accuracy, whereas ResNet-101 converges more slowly but maintains stable generalization, ultimately reaching the highest final validation accuracy. Notably, no significant overfitting was observed for any backbone within 200 epochs.

The corresponding loss curves (Fig. C.23) provide further insights. For MobileNetV3-Large, the training loss decreases sharply during the first epochs but quickly stabilizes, while the validation loss remains relatively high and fluctuates after approximately 40 epochs, indicating limited representational capacity. ResNet-50 exhibits steady and consistent decreases in both training and validation loss, with a small gap between them, indicating stable convergence and strong generalization. In contrast, ResNet-101 exhibits a slower decline in training loss. It maintains a higher validation loss relative to training, suggesting a mild tendency toward overfitting despite ultimately achieving the lowest final loss. Taken together, the loss dynamics confirm that ResNet-50 offers the most favorable balance between convergence speed, stability, and generalization for the present dataset.

The evolution of Dice, precision, and recall over 200 training epochs is summarized in Fig. 11. MobileNetV3-Large exhibited stable convergence but plateaued at a lower recall, consistent with its tendency to under-segment smaller or faint voids. ResNet-50 achieved a higher and more stable void recall while maintaining precision, leading to the highest mean Dice score of 0.85. ResNet-101 displayed comparable recall and Dice, but its learning curve fluctuated more strongly, suggesting increased sensitivity to initialization and stochastic effects.

Table 4
Comparison of different backbone architectures in terms of training efficiency and segmentation performance. Note that BG denotes background and acc. denotes accuracy.

Backbone	Time [h]	BG acc.	Void acc.	Mean Dice
MobileNetV3-Large	2.5	0.96	0.67	0.80
ResNet-50	5.0	0.94	0.81	0.85
ResNet-101	6.5	0.96	0.78	0.84

Confusion matrices in Fig. 11 provide further insights. MobileNetV3 achieved intense background discrimination of 96% but frequently misclassified void pixels, resulting in fragmented segmentation masks. ResNet-50 balanced background accuracy of 94% with higher void recall, producing the best overall trade-off. ResNet-101 slightly improved the background accuracy of 96% but did not substantially increase void detection compared to ResNet-50.

A summary of training efficiency and segmentation accuracy is given in Table 4. MobileNetV3 offered the fastest training but sacrificed recall, while ResNet-101 demanded the longest training time without delivering proportional performance gains. ResNet-50 emerged as the optimal compromise between efficiency and accuracy for this dataset.

Representative segmentation masks are shown in Fig. 12. MobileNetV3 tended to capture only the largest voids, often missing fine boundaries and producing fragmented predictions. ResNet-50 yielded sharper boundaries and more complete detection of small micro-voids, aligning more closely with ground truth. ResNet-101 produced slightly smoother masks but showed no significant improvements over ResNet-50, despite doubling the training cost. This observation suggests diminishing returns when increasing backbone depth under the current dataset size and image characteristics.

Taken together, these results demonstrate that ResNet-50 provides the most favorable balance between segmentation accuracy and computational efficiency. Lightweight backbones such as MobileNetV3 are

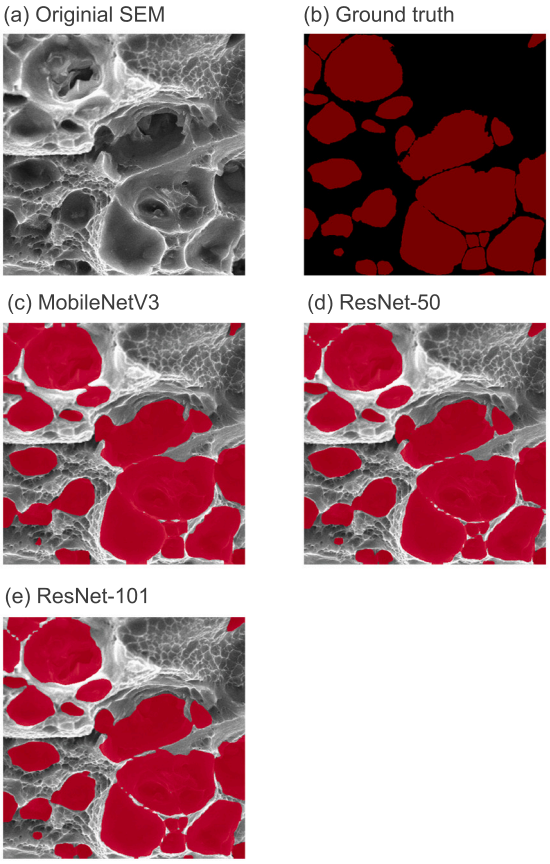

Fig. 12. Qualitative comparison of segmentation outputs with different backbones: (a) original SEM image, (b) ground truth, and segmentation results predicted by (c) MobileNetV3, (d) ResNet-50, and (e) ResNet-101. Red regions indicate the segmented micro-voids overlaid on the SEM images.

suitable for rapid prototyping but sacrifice recall on subtle structures, whereas deeper backbones such as ResNet-101 incur higher training costs without yielding proportional accuracy gains.

6.2. Cross-validation performance evaluation

To evaluate the robustness and generalization of the segmentation model, a five-fold cross-validation was conducted. A 400 annotated images dataset was used. In each fold, 80% of the data were used for training and 20% for validation, with the test partition held constant across all runs. The averaged performance across folds is summarized in Table 5.

The model exhibited consistently high pixel-level accuracy and Dice scores across all folds, indicating stable learning behavior with minimal dependence on data partitioning. The Dice coefficient showed limited variation (standard deviation < 0.01), confirming the robustness of the model across different training splits. Precision and recall remained well-balanced, suggesting that the network effectively captured micro-void regions while avoiding over-segmentation. Overall, these results demonstrate that the DeepLabv3-based framework achieves reliable and repeatable performance even with moderate dataset sizes and under standard computational constraints.

6.3. Comparison of image recognition approaches

A comparison of image recognition using ImageJ, traditional machine learning (ML), and convolutional neural networks (CNN) with

Table 5
Summary of the results of 5-fold cross-validation, averaged over 100 epochs until the end of training.

Fold	Accuracy (%)	Dice	Precision	Recall
1	94.44	0.8121	0.8299	0.8008
2	91.49	0.8002	0.8146	0.7907
3	92.43	0.7898	0.8254	0.7612
4	91.94	0.8059	0.8130	0.8028
5	90.77	0.8103	0.8202	0.8043
Mean	92.21	0.8036	0.8206	0.7920
Std.	1.24	0.0081	0.0064	0.0161

two examples, one tension-dominated and one shear-dominated SEM image, is shown in Fig. 13. For the tensile loading condition H-0/1 [71], the micro-void ratio identified by ImageJ with the default threshold is 59.4%, while those obtained by ML and CNN are 21.52% and 33.09%, respectively. In the SEM image shown in Fig. 13 for the shear loading condition H-1/0, the micro-void ratios identified by ImageJ with the default threshold, ML, and CNN are 79.64%, 8.59%, and 1.26%, respectively. From optical observation, it is clear that ImageJ with default setting leads to an overestimation of micro-voids, particularly in the shear loading case, with a value of 79.64%, where no voids can be observed in the SEM image. In contrast, the micro-voids identified in the red mask using CNN appear more realistic compared with the other methods. Additionally, by manually adjusting the threshold, as shown in Fig. 13, the results are comparable to those predicted by the CNN. However, this manual adjustment is time-consuming, highly subjective, and provides inconsistent results. By comparing the SEM images with the adjusted threshold to the ML-identified SEM images, it can be observed that the traditional ML method fails to distinguish micro-voids accurately and yields incorrect predictions. Based on the above analysis, it can be concluded that the advanced deep learning method (CNN) demonstrates advantages in SEM image recognition compared with traditional ML and threshold-based methods.

6.4. Effect of dataset size

To assess how annotation volume affects performance, we trained DeepLabv3 (ResNet-50 backbone) with three dataset sizes: 500, 700, and 900 images. All runs used identical hyperparameters and train/validation splits to ensure comparability. The corresponding training/validation curves and confusion matrices are shown in Fig. 14.

With 500 images, the model achieved rapid convergence, with Dice stabilizing near 0.86 after 100 epochs (Fig. 14(a)). Despite the smaller dataset, both precision and recall were competitive, demonstrating that the model can learn discriminative features from relatively modest annotation volumes. Increasing the dataset to 700 images resulted in smoother training curves and reduced oscillations in Dice and recall, although the average performance gains were modest (Fig. 14(b)). At 900 images, the curves were the most stable, with precision consistently above 0.9 and recall improving slightly compared to the 700-image setting (Fig. 14(c)). These results suggest diminishing returns in performance improvements once the dataset exceeds 700 annotated samples under the present configuration. However, when dealing with entirely new SEM data, having a larger dataset can help to ensure more consistent results when evaluating the model.

The confusion matrices highlight similar trends. With 500 images, background classification accuracy was already strong at 95%, and void recall reached 81%. Expanding to 700 images slightly increased background accuracy to 96% but reduced void recall (76%), reflecting a more conservative classifier. At 900 images, void recall improved again to 79% while maintaining high background accuracy of 96%, confirming that additional annotations primarily benefit the minority class of micro-voids area.

A summary of segmentation performance versus dataset size is presented in Table 6. The results confirm that even the smallest dataset

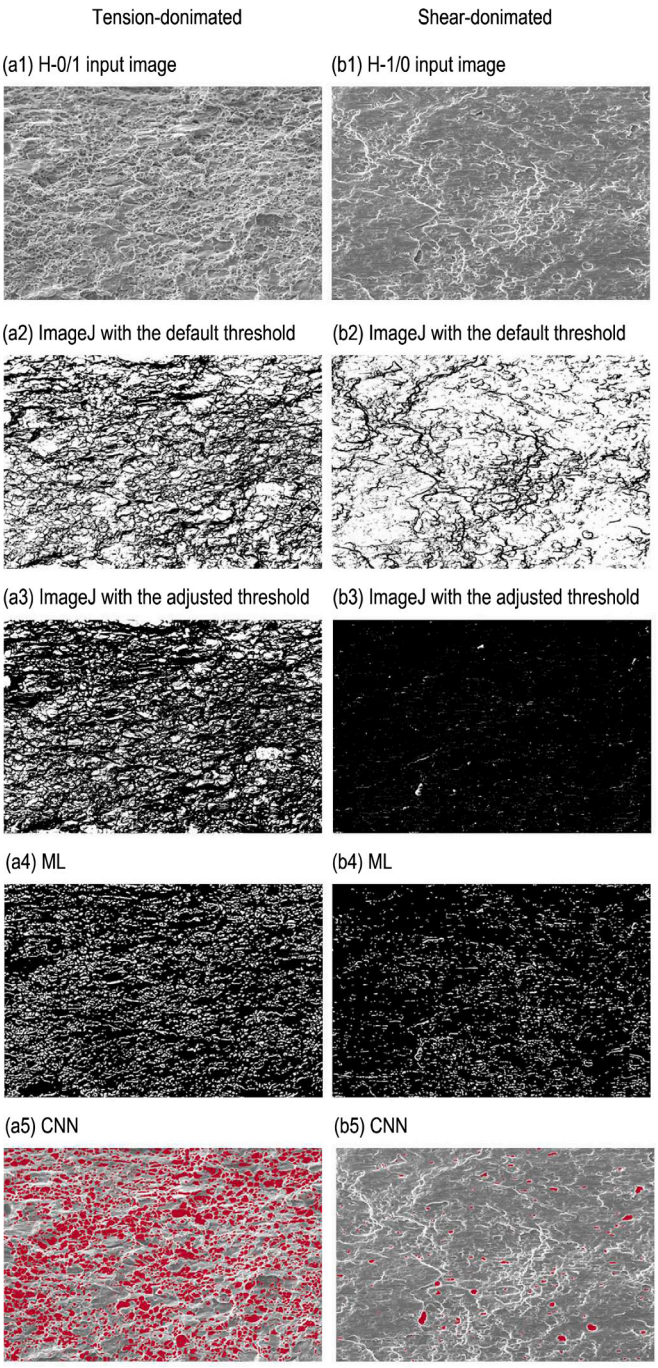

Fig. 13. Comparison of void segmentation results obtained with different methods under tension-dominated (left) and shear-dominated (right) loading conditions: (a1) and (b1) input SEM images for tension-dominated and shear-dominated loading conditions, respectively; (a2) and (b2) segmentation results obtained using ImageJ with the default threshold; (a3) and (b3) segmentation results obtained using ImageJ with an adjusted threshold; (a4) and (b4) segmentation results from a traditional machine learning (ML) approach; (a5) and (b5) segmentation results predicted by the deep learning CNN model, where red regions indicate the identified voids. For the ImageJ- and ML-identified images, white regions indicate micro-voids.

provides competitive performance, while larger datasets mainly improve training stability and void-class detection. However, beyond 700 samples, the incremental benefits diminish, with only marginal improvements at 900 images compared to 700 using this validation dataset.

Table 6
Performance versus dataset size (evaluated on validation dataset).

Data size	BG acc.	Void recall	Mean Dice	Notes
500	0.95	0.81	0.86	Fast convergence
700	0.96	0.76	0.86	Stable, small gain overall
900	0.96	0.79	0.87	Smoothest curves

Overall, these experiments demonstrate that DeepLabv3 can achieve reliable segmentation performance with as few as 500 annotated images, reducing labeling costs significantly. While enlarging the dataset improves robustness and stabilizes training, the marginal utility of additional annotations diminishes beyond 700 samples with current validation dataset. For the remainder of this study, we adopt the 900-image configuration to ensure maximal robustness, while noting that the 500-image setting already provides competitive accuracy.

7. SEM-based analysis of anisotropic damage behavior

Since ductile damage mechanisms are stress-state-dependent, micro-void growth and coalescence dominate under high stress triaxiality, whereas the growth and coalescence of micro-shear-cracks become dominant under low positive and negative stress states [16,76]. Accordingly, as introduced in Section 6, both micro-voids and micro-shear-cracks can be analyzed using deep learning, enabling the identification of characteristic micro-defect properties, such as size, length, and porosity. In addition, stress states obtained from numerical simulations are used, together with the identified defect characteristics, to interpret stress-state-dependent damage mechanisms.

In this section, the above stress-state-dependent damage mechanisms are examined at the microstructural level using SEM images from the fracture surfaces of the designed biaxial experiments under monotonic and cyclic loading, and the corresponding analyses are discussed in detail.

7.1. Transition stress states in ductile damage mechanisms

In this work, 18 biaxial loading cases and one uniaxial loading condition are used to study the relation between micro-defects and stress states, including the mean micro-void ratio $\bar{f}$ and the mean area of the micro-voids $\bar{s}$, as well as to verify the proposed machine learning approach. The fracture surfaces and selected representative SEM images for the biaxial experiments are shown in Fig. 15, and those for the uniaxial tension test are depicted in Fig. 1(a). It should be noted that these images were not used for training in either traditional machine learning or CNN models, as different dataset sizes were employed for training. Since stress states cannot be directly measured in experiments, numerical simulations with the elastic–plastic–damage constitutive model summarized in Section 3 and proposed by Wei et al. [76] were performed to obtain the corresponding results. Comparison of the numerical results with experimentally measured load–displacement curves shows that the proposed model can accurately capture the material behavior. Thus, the numerically predicted stress states can be assumed to represent the actual experimental conditions. In addition, the corresponding biaxial results with HC-specimens, taken from previous works, are shown in Fig. A.22 in Appendix A.

Fig. 16 shows the contour plots of the stress triaxiality η and Lode parameter ω over the notch cross-section for all experiments, obtained from numerical calculations performed just before fracture. The stress states (stress triaxialities and Lode parameters) exhibit an almost homogeneous distribution with edge gradients. Wei et al. [76] proposed an area-weighted mean value method to better represent the stress states instead of using contour plots for a given loading condition, and the

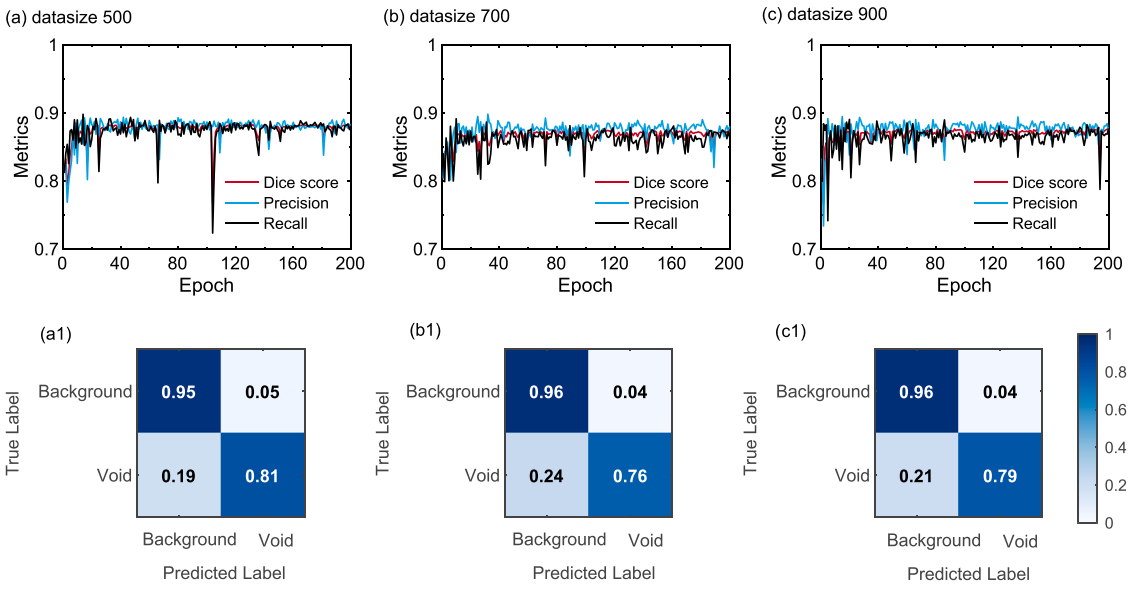

Fig. 14. Segmentation performance metrics and confusion matrices for CNN training with different training dataset sizes of SEM images: (a)–(c) relationship between the Dice score, precision, and recall versus training epoch for dataset sizes of 500, 700, and 900 SEM images, respectively, and (a1)–(c1) the corresponding confusion matrices for the background and void classes.

Table 7
Micro-void properties analyzed using different SEM image recognition approaches and numerically predicted mean stress states.

Test name	Stress state		Traditional ML			CNN 500			CNN 700			CNN 900		
	$\bar{\eta}$ [−]	$\bar{\omega}$ [−]	N [−]	$\bar{f}$ [%]	$\bar{s}$ [μm^2]	N [−]	$\bar{f}$ [%]	$\bar{s}$ [μm^2]	N [−]	$\bar{f}$ [%]	$\bar{s}$ [μm^2]	N [−]	$\bar{f}$ [%]	$\bar{s}$ [μm^2]
HC-TN5	−0.13	0.15	368	0.94	0.78	298	3.69	3.14	270	2.85	2.94	272	3.69	3.62
H-1/-1	−0.12	0.13	3	0.02	2.42	21	0.09	10.55	46	0.12	6.99	22	0.07	8.27
HC-TN3	−0.06	0.04	230	0.63	0.89	79	0.38	1.86	82	0.26	1.18	124	0.61	1.83
X0-1/-1	0.03	−0.04	178	5.06	3.57	34	0.83	3.16	41	0.93	2.89	71	2.23	4.64
HC-T0	0.08	−0.12	1022	4.42	1.18	64	0.49	3.24	62	0.41	2.51	83	0.61	2.73
X0-1/−0.5	0.09	−0.19	497	23.63	6.08	39	0.77	2.99	58	1.10	2.64	62	1.53	3.41
H-1/0	0.10	−0.17	1516	4.54	5.94	107	0.54	11.40	92	0.38	8.15	217	1.26	12.66
X0-1/0	0.19	−0.41	4756	22.98	9.94	301	1.11	8.08	777	2.91	8.31	879	3.90	9.84
HC-T3	0.22	−0.26	676	2.37	1.18	91	0.69	3.00	81	0.44	2.11	107	0.83	3.06
HC-S5	0.28	−0.34	2225	12.46	1.94	273	2.49	3.13	262	1.88	2.67	218	1.67	2.90
HC-T5	0.32	−0.44	1916	31.46	6.30	931	21.39	8.90	893	16.71	7.31	861	21.40	9.76
HC-S4.5	0.33	−0.43	2262	12.26	1.87	595	10.81	6.42	533	6.64	4.35	558	7.44	5.04
HC-T6	0.39	−0.56	2807	10.92	7.97	954	12.70	27.78	988	8.71	18.61	1148	12.79	24.37
T	0.41	−0.90	2695	8.85	3.94	655	47.78	43.29	1852	29.7	37.25	1623	24.08	34.88
HC-T7	0.44	−0.69	2686	20.74	2.76	803	37.84	18.61	895	29.48	13.08	808	30.15	14.91
X0-1/1	0.47	−0.49	403	15.25	4.87	88	20.21	33.36	150	19.97	19.45	149	23.84	22.95
HC-T9	0.58	−0.75	1799	11.93	2.33	707	47.46	26.47	847	40.33	18.71	753	42.71	22.30
H-0/1	0.70	−0.48	3554	16.54	9.62	1450	35.39	55.02	1977	25.64	28.85	1951	33.09	37.67
HC-S0	0.71	−0.46	2525	13.83	1.90	487	55.16	45.41	826	35.08	16.98	686	30.55	17.65

corresponding numerical simulations confirmed its effectiveness in representing and distinguishing different loading conditions. In addition, the mean values change only slightly during loading under monotonic proportional loading conditions. In this work, since the SEM images are taken from different positions of the notch cross-section to represent the microstructure of the whole cross-section, the mean value over this region is suggested to better represent the corresponding loading case. Thus, the mean stress triaxiality $\bar{\eta}$ and mean Lode parameter $\bar{\omega}$ are defined as follows

$$\bar{\eta} = \frac{1}{S}\iint_S \eta \, ds \quad \text{and} \quad \bar{\omega} = \frac{1}{S}\iint_S \omega \, ds, \tag{11}$$

where S denotes the total area of the notch cross-section, and the stress triaxiality and Lode parameter are calculated as defined in Eqs. (1) and (4), respectively. The numerically predicted mean stress triaxiality $\bar{\eta}$ and mean Lode parameter $\bar{\omega}$ for different loading cases are summarized in Table 7. A wide range of stress triaxialities is obtained from the designed biaxial experiments with different loading ratios, covering values from negative $\bar{\eta} = -0.13$ to high positive $\bar{\eta} = 0.71$. The obtained Lode parameters range from −0.9 to around 0.2.

Moreover, as illustrated in Fig. 15, at high stress states the fracture lines appear straighter with a zigzag form, as shown in Figs. 15(i) and (m). From optical observations, it is evident that the growth of micro-voids mainly causes ductile damage. In contrast, for stress states close to zero or negative values, the fracture lines exhibit shear band formation, and at negative stress states, distortions appear at the notch edges, as shown in Fig. 15(h). Compared to high stress states, SEM images at low stress states reveal almost no voids, but instead only micro-shear-cracks. A detailed analysis of micro-defect properties (number, micro-void ratio, and mean area) predicted by traditional machine learning and CNN models with different dataset sizes is summarized in Table 7.

In addition, the training data and corresponding trend fitting are illustrated in Fig. 17. It is evident that the relation between the mean micro-void ratio $\bar{f}$ and the mean stress triaxiality $\bar{\eta}$, as well as that

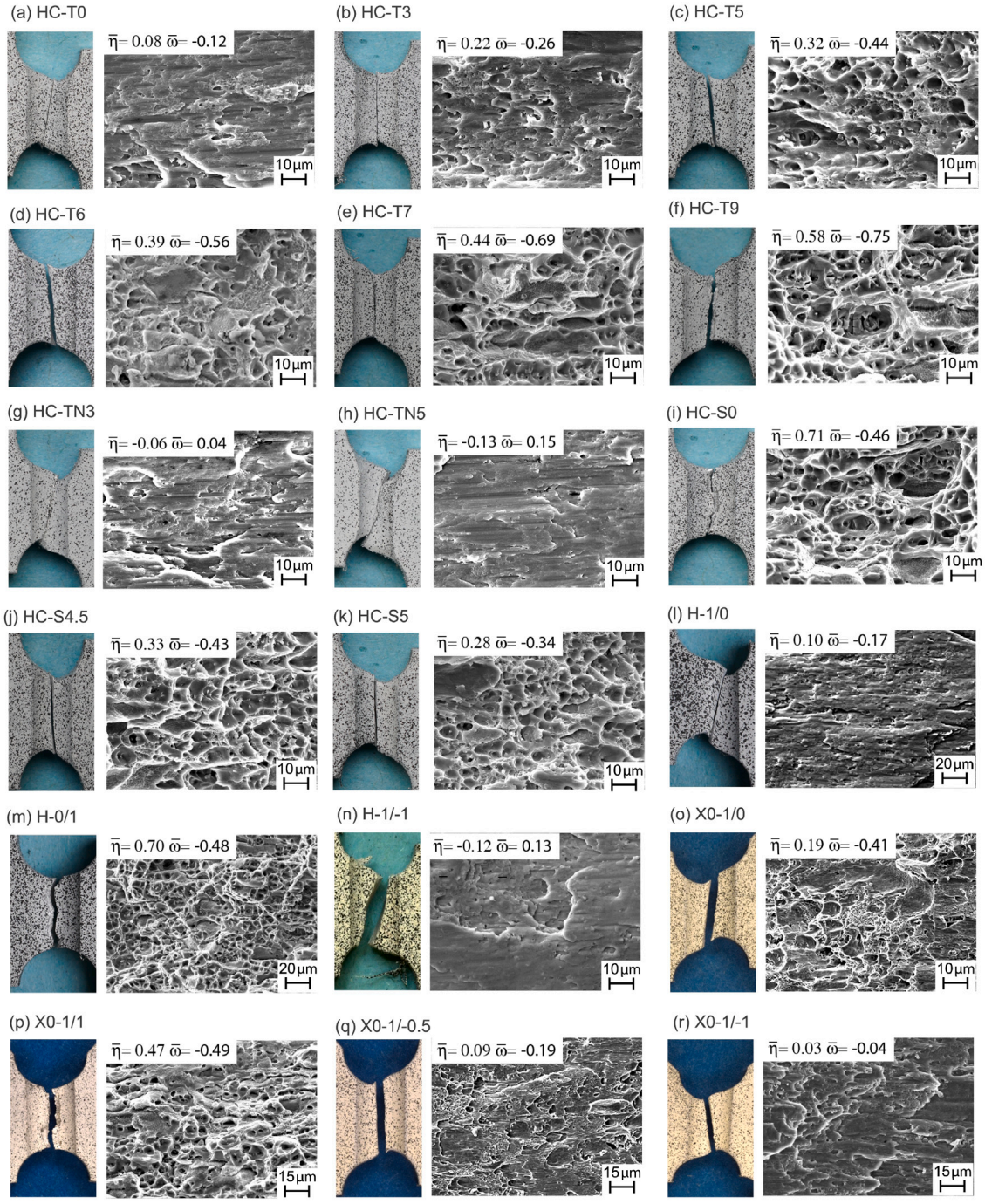

Fig. 15. Fracture photos and representative SEM images for various biaxial loading conditions, covering a wide range of stress states. For each loading condition, the left side shows the fracture photograph of the notched region, and the right side shows the corresponding SEM image taken from the notch surface. The mean stress triaxiality $\bar{\eta}$ and Lode parameter $\bar{\omega}$ are calculated based on numerical simulations, and the calculation equation is given in Eq. (11).

between the mean micro-void ratio $\bar{f}$ and the mean Lode parameter $\bar{\omega}$, obtained using traditional machine learning differs from that obtained with the CNN approach. As shown in Figs. 17(a) and (b), no evident transition trend is observed. The predicted transition stress triaxiality point at $\bar{\eta} = 0.06$ is significantly non-physical, since the SEM images under low positive stress states still show micro-shear-cracks dominating, as illustrated in Figs. 15(b) and (o). Similarly, a non-physical prediction of the transition in the micro-void ratio is obtained based on the Lode parameter criterion, with an identified value of $\bar{\omega} = -0.14$. Moreover, by comparing the predicted micro-void ratios $\bar{f}$ under loading condition HC-T0 with mean stress triaxiality $\bar{\eta} = 0.08$ and mean Lode parameter $\bar{\omega} = -0.12$, and experiment X0-(1/–0.5) with $\bar{\eta} = 0.09$ and $\bar{\omega} = -0.19$, the values predicted by traditional machine learning are significantly higher than those from the deep learning method. It is evident from the corresponding SEM images in Fig. 15 that the traditional machine learning predictions of $\bar{f}$ are incorrect, clearly highlighting the advantage of the deep learning method over the traditional approach.

For the relation between the mean micro-void ratio $\bar{f}$ and the mean stress triaxiality $\bar{\eta}$, a tanh-like distribution trend is observed. Although the upper shelf differs slightly between CNN models trained with different dataset sizes, the transition from shear-dominated to tensile-dominated behavior consistently appears around $\bar{\eta} = 0.38$. This value is close to that obtained in uniaxial tension tests ($\bar{\eta} = 0.41$),

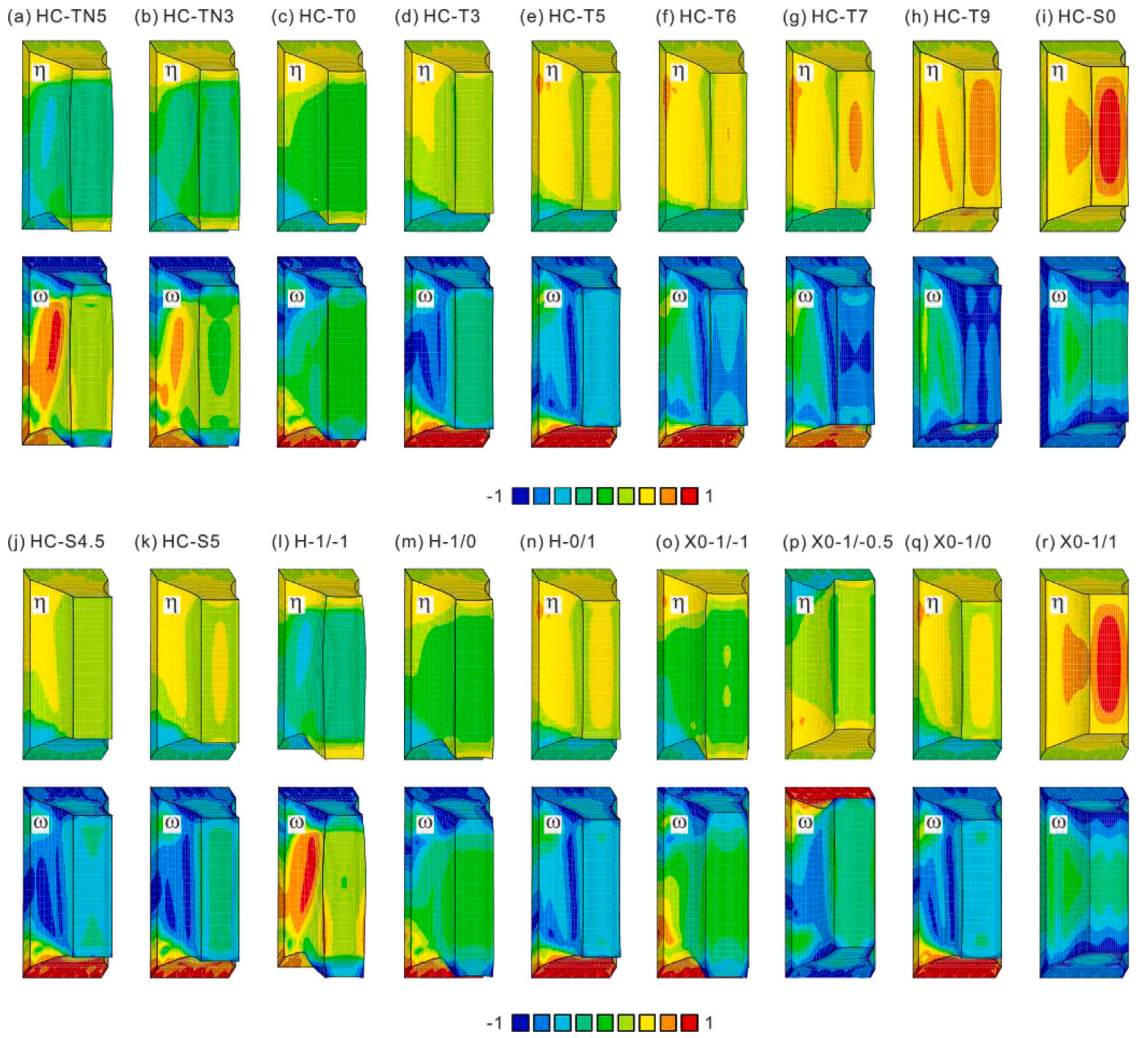

Fig. 16. Distributions of the numerically predicted stress triaxiality η (upper row) and Lode parameter ω (lower row) over the notched cross-section of the biaxially loaded HC-, H-, and X0-specimens under different complex loading conditions. The details of the geometries can be found in Fig. A.21.

where the stress triaxiality increases due to necking and therefore deviates from the analytical value of 1/3. The predicted transition also confirms the proposed damage mechanism: under high stress states, damage is governed by micro-void growth, whereas under low positive and negative stress triaxialities, it is governed by micro-shear-crack deformation. These findings quantitatively identify the transition point ($\bar{\eta} = 0.38$) from shear-dominated to tensile-dominated ductile fracture behavior.

Furthermore, the transition of the tanh function is steeper for the relation between the mean micro-void ratio $\bar{f}$ and the mean Lode parameter $\bar{\omega}$, whereas a smoother transition is observed in the relation between the mean micro-void ratio $\bar{f}$ and the mean stress triaxiality $\bar{\eta}$, as shown in Fig. 17. In addition, it can be observed that the dependence on the Lode parameter is less pronounced. For values above $\bar{\omega} > -0.44$, the porosity remains relatively small and nearly constant. In contrast, for $\bar{\omega} \leq -0.44$, although a constant fitting is suggested for the micro-void ratio, deviations around the predicted line are observed as the Lode parameter decreases, highlighting its secondary but non-negligible role in controlling void evolution.

It is worth noting that the calibrated transition stress triaxiality of $\bar{\eta} = 0.38$ is not coupled with the corresponding Lode parameter value of $\bar{\omega} = -0.44$, since the transition points based on stress triaxiality and the Lode parameter are identified separately. As shown in Fig. 15(d), the stress states for the HC-T6 experiment lie near the transition point defined by stress triaxiality ($\bar{\eta} = 0.38$), whereas the corresponding Lode parameter is −0.56, which differs significantly from the transition Lode parameter. In addition, due to the definition of the Lode parameter, which reaches its maximum ($\omega = 1$) at uniaxial compression and equibiaxial tension and its minimum ($\omega = -1$) at uniaxial tension and equibiaxial compression, as illustrated in Fig. 4, there is no monotonic trend with stress triaxiality. For example, the Lode parameter decreases from −0.26 to −0.9 for tests HC-T3 to T and then increases again, as listed in Table 7. This non-monotonic trend makes it unsuitable to use the Lode parameter alone as an independent criterion. While the Lode parameter may contribute as a secondary indicator of stress states and influence damage transition behavior, identifying its coupling with stress triaxiality requires denser experimental coverage around the transition states. Owing to the limited number of available data points and experimental techniques, this effect cannot be quantified in the present work. Consequently, although the Lode parameter plays a secondary but non-negligible role, the critical transition position is determined primarily by the stress triaxiality criterion.

Concerning the relation between the stress states ($\bar{\eta}$ and $\bar{\omega}$) and the mean area of the micro-voids $\bar{s}$, distinct trends are observed. The mean area of the micro-voids increases significantly with increasing stress triaxiality, as shown by the data in Table 7 and the fitted curve in Fig. 18(b). Unlike the distribution of the mean void ratio $\bar{f}$, no upper shelf appears at higher stress states, suggesting that void coalescence becomes more dominant than void growth under such conditions. This interpretation is physically reasonable, as specimens subjected to high

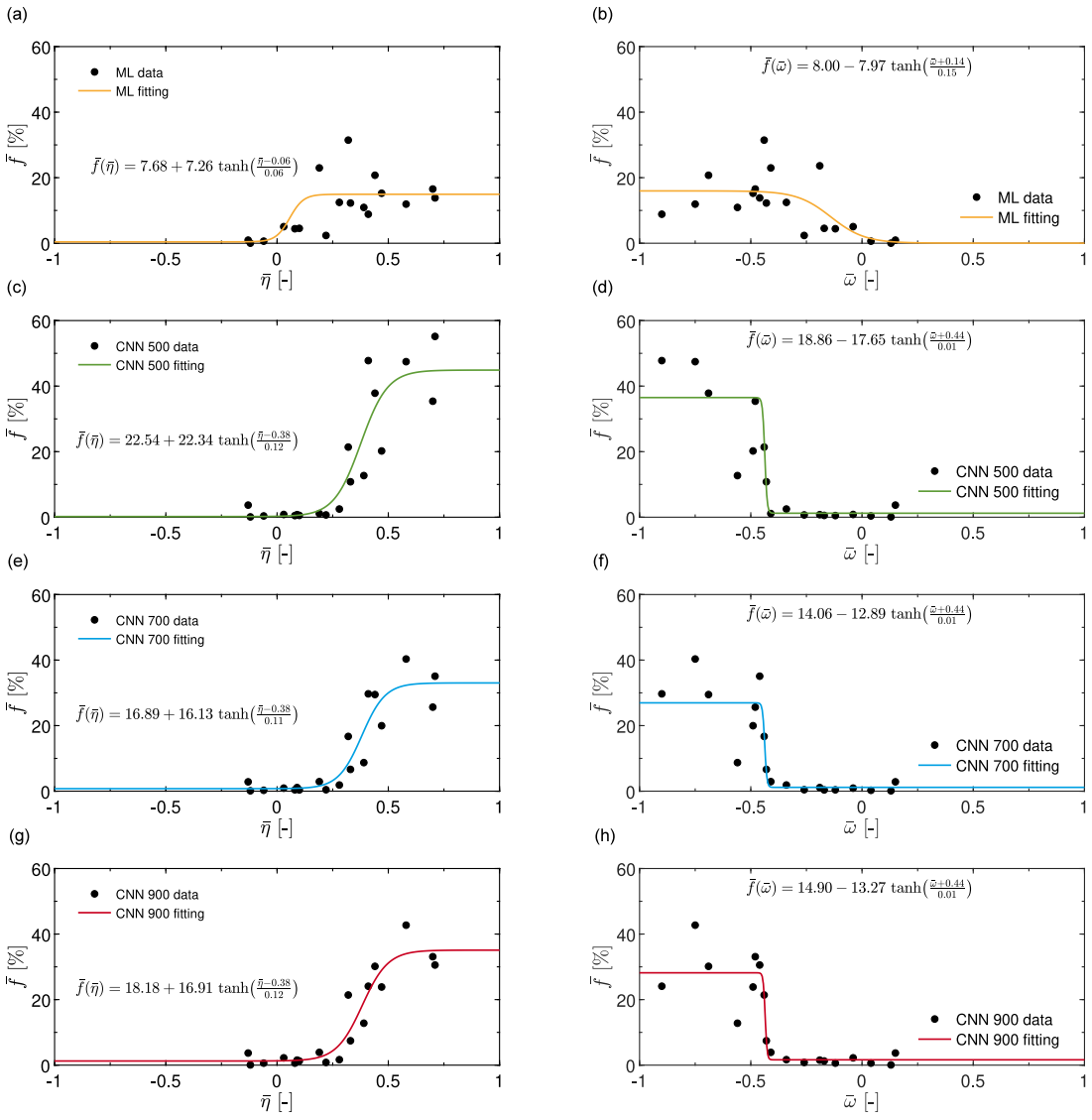

Fig. 17. Statistically analyzed relationships between the mean micro-void ratio $\bar{f}$ and the corresponding mean stress triaxiality $\bar{\eta}$ (left column) and mean Lode parameter $\bar{\omega}$ (right column): (a) and (b) results obtained using the traditional machine learning method, and (c)–(h) results obtained using deep learning methods with different training dataset sizes of 500, 700, and 900 SEM images. Scatter points represent the predicted data, and the solid lines correspond to curve fitting.

stress states fail much earlier than under other loading conditions. For the Lode parameter, the behavior is different: above the cut-off value $\bar{\omega} \geq -0.44$, the mean void areas remain relatively small with only minor differences, whereas the mean area increases rapidly with decreasing Lode parameter, as illustrated in Fig. 18(d).

By comparing the prediction results of the CNN models with dataset sizes of 500, 700, and 900, along with the corresponding fitted transition curves, it can be observed that as the dataset size increases from 700 to 900, the CNN predictions become more stable and better aligned with the data, as shown in Fig. 18. Furthermore, the trend curves predicted by CNN-700 and CNN-900 differ only slightly, indicating that while larger datasets improve CNN accuracy, training with 700 data points already yields sufficiently reliable results. Most importantly, unlike the ML model, which exhibits an unphysical transition point around $\bar{\eta} = 0.06$ or $\bar{\omega} = -0.14$, the CNN provides more realistic transitions that are consistent with expected physical behavior.

The above findings confirm the effectiveness of the proposed deep learning approach in distinguishing between micro-voids and micro-shear-cracks, thereby providing a quantitative understanding of how stress states influence microstructural evolution. The identified micro-defect characteristics enable a more distinct separation of the transition values associated with different damage mechanisms, as well as the transition parameters related to combined stress states. In future work, the identified ratio of micro-voids and the corresponding transition values can be incorporated into the damage evolution law in Eq. (9), thereby improving the numerical accuracy of damage and fracture predictions. This approach further allows the analysis of SEM images across different temperatures and strain rates, providing insights into their influence on microstructural evolution at the micro-level.

7.2. Evolution of micro-defect properties under cyclic loading

Cyclic loading is another common condition encountered in engineering applications. Based on experimental and numerical analyses of shear cyclic tests under biaxial loading, Wei et al. [25] indicated that changes in loading direction can cause shear-cracks to close. As a result, the mean crack length may decrease under cyclic loading and exhibits a non-monotonic trend between the number of loading cycles and the mean crack length.

In this work, four cyclic loading tests with varying amplitudes were conducted using HC-specimens (Fig. A.21(c)). A tensile preload of 3 kN was applied and maintained in the horizontal loading direction, while

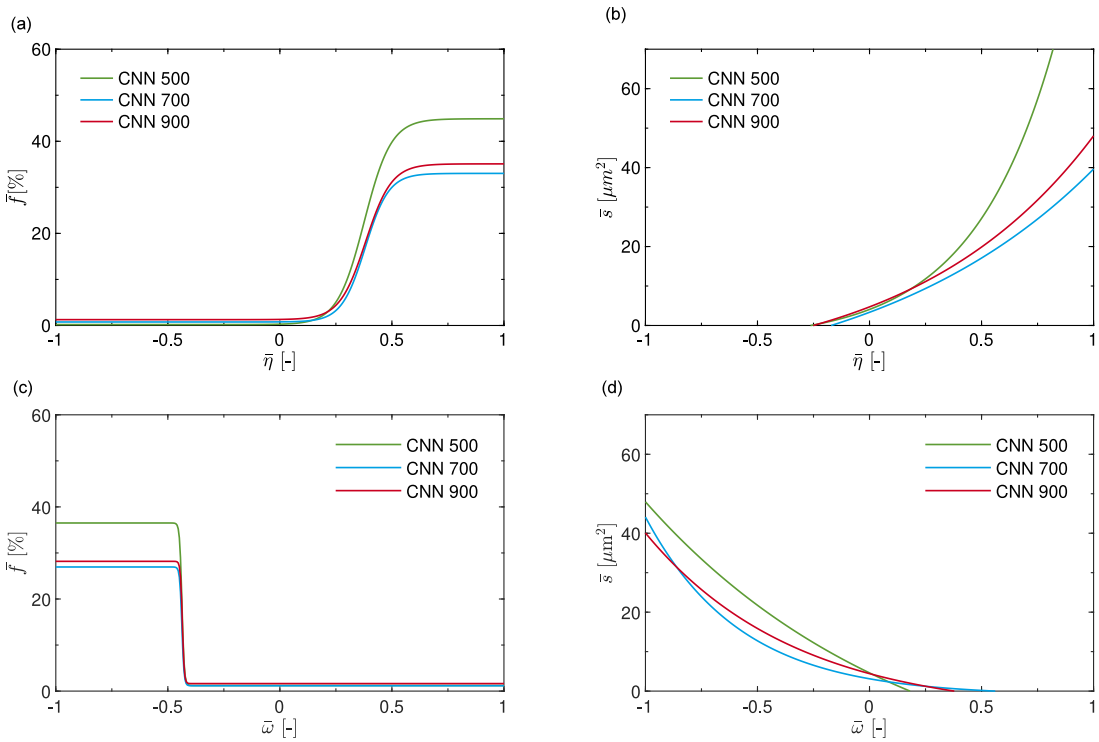

Fig. 18. Comparison of statistically analyzed relationships between the mean micro-void ratio $\bar{f}$ and the mean micro-void area $\bar{s}$ with the corresponding mean stress triaxiality $\bar{\eta}$ and mean Lode parameter $\bar{\omega}$, obtained using deep learning methods with training dataset sizes of 500, 700, and 900 SEM images, respectively.

cyclic shear loading was superimposed in the vertical direction. The numerical and experimental results are presented in Fig. 19. It is clearly shown that the numerical simulations agree well with the experimental results, and the corresponding stress triaxialities are quite similar, with values of 0.27, 0.29, 0.26, and 0.25 for the experiments with one, two, six, and twenty loading cycles, respectively. In contrast, the Lode parameter shows more pronounced differences, with values of −0.17, −0.17, −0.30, and −0.29 for the experiments with one, two, six, and twenty loading cycles, respectively. The respective SEM images of the fracture surfaces are illustrated in Fig. 20. It is evident that damage and fracture mechanisms are not governed solely by micro-voids but also involve micro-shear-cracks. Compared with the CNN-identified critical stress triaxiality of $\bar{\eta} = 0.38$ or the Lode parameter value of $\bar{\omega} = -0.44$, the above-mentioned stress states all satisfy this criterion, as experimentally confirmed by the SEM images.

Since the deep learning prediction results with CNN 900 show the best performance, as discussed in Section 7.1, the analyzed micro-defects (micro-voids and micro-shear-cracks) are evaluated using CNN 900. The corresponding results are summarized in Table 8. One can observe that the stress triaxialities under monotonic and cyclic loading conditions differ only slightly. The mean micro-void ratio $\bar{f}$ increases with the number of loading cycles, while the mean area of the micro-voids shows no distinct trend. In contrast, the mean crack length under cyclic loading is significantly smaller than under monotonic loading, and the mean area of the micro-voids is larger compared to monotonic loading. These findings indicate that micro-shear-cracks can alter deformation directions during cyclic loading, leading to either an increase or a decrease in crack length $\bar{l}$, depending on the loading history. The same phenomenon has also been confirmed by Wei et al. [25] through an analysis of SEM images of fracture surfaces under different shear cyclic loading conditions, which indicate that changes in the loading direction alter the evolution direction of micro-shear-cracks. Consequently, the crack length may increase or decrease depending on the loading direction and cyclic loading amplitude. The corresponding

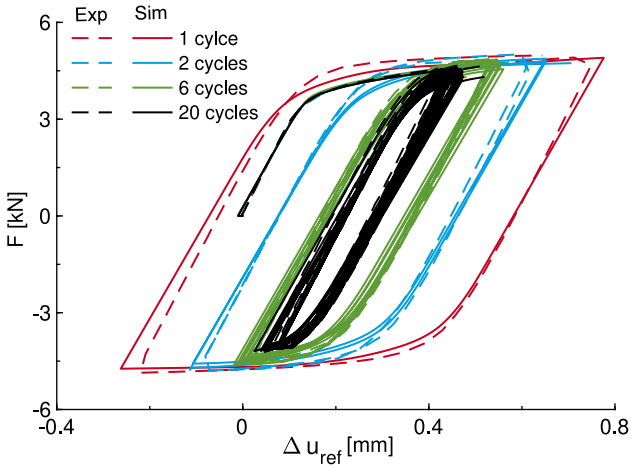

Fig. 19. Experimental (dashed lines) and numerical (solid lines) force–displacement curves for biaxial cyclic loading tests using HC-specimens (Fig. A.21(c)) with loading cycles of 1, 2, 6, and 20.

numerical simulations are consistent with this observation, showing that damage evolution under different cyclic loading patterns results in distinct damage accumulation and, thus, different fracture behavior.

Furthermore, variations in cyclic loading may also alter the shape of pre-existing micro-voids, making them more prone to coalescence with neighboring voids and thereby increasing the mean area of the micro-voids, as can be observed from the SEM images in Fig. 20. Similar observations have been reported after compressive loading, where penny-shaped micro-voids tend to coalesce with adjacent voids [75, 87], leading to an increased mean micro-void area $\bar{s}$. In addition,

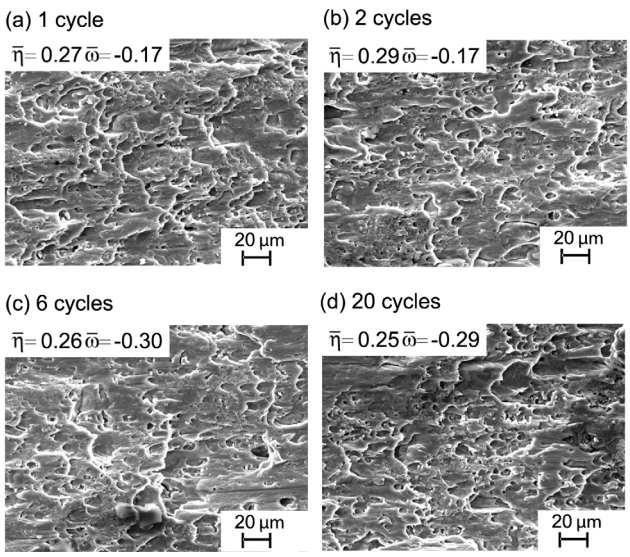

Fig. 20. SEM images taken from the fracture surface of biaxially loaded HC-specimens subjected to 1, 2, 6, and 20 loading cycles. The mean stress triaxiality $\bar{\eta}$ and Lode parameter $\bar{\omega}$, defined in Eq. (11), are calculated based on numerical simulations.

Table 8
Properties of micro-defects (micro-voids and micro-shear-cracks) identified by CNN analysis for biaxial shear cyclic tests.

Loading cycle(s)	Stress states		Micro-voids		Micro-shear-cracks
	$\bar{\eta}$ [−]	$\bar{\omega}$ [−]	$\bar{f}$ [%]	$\bar{s}$ [μm^2]	$\bar{l}$ [μm]
mon.	0.22	−0.26	0.83	3.06	33.00
1	0.27	−0.17	0.68	2.82	16.39
2	0.29	−0.17	1.17	3.39	12.62
6	0.26	−0.30	1.27	3.20	21.66
20	0.25	−0.29	1.34	3.30	20.43

Wei et al. [80] performed a statistical analysis of area changes in micro-defects based on metallographic images obtained during cyclic loading experiments interrupted at different loading stages, revealing the influence of loading patterns and loading histories on damage behavior.

The above-mentioned analysis of micro-defect properties using the deep learning CNN approach clearly demonstrates that the proposed method can also be applied to capture microstructural changes (micro-shear-crack length and micro-void area and porosity) in cyclic loading tests with different loading histories. However, a more detailed investigation of the influence of loading history on the evolution of micro-defect properties requires a larger number of cyclic loading tests with varying conditions in future studies.

7.3. Physics-informed deep learning extensions and future directions

Although this study achieved good results using a mature semantic segmentation neural network, introducing more physically meaningful information during training or recognition remains an important direction for future work to improve robustness and interpretability further. One approach is to introduce physical constraints during model training and testing. For example, under certain stress triaxiality and Lode parameter conditions, the identified porosity should not deviate significantly from a prescribed range. Another way to improve robustness is to adopt a hybrid local–global strategy, in which models are trained for specific acquisition settings or sub-domains and then combined through shared physical constraints [88]. Such enhancements not only strengthen generalization across diverse experimental setups but also bridge the gap between SEM-based morphometrics and continuum damage modeling, ultimately reducing reliance on manual annotation.

Building on these physical foundations and the observed critical stress-state transition behavior, the next step is to characterize damage evolution through micromechanical numerical simulations of a representative volume element (RVE) containing one or more voids to predict micro-void deformation under complex loading conditions. To determine stress-state-dependent parameters within the evolution equations, physics-informed recurrent models can be used, with the evolution laws embedded in the network architecture or the loss function [89]. The outputs of these predictions can then be used as dynamic constraints for the CNN model, creating a closed-loop refinement process that aligns image recognition with the underlying mechanics of material failure. Experimental validation is necessary to confirm deep learning results, and additional biaxial loading paths need to be designed to cover a wider range of loading conditions.

8. Conclusion

This study presents a comparative framework for segmenting micro-voids in SEM fracture-surface images, evaluating both traditional machine learning approaches and modern deep learning models. Random forest combined with handcrafted multi-scale descriptors provides a competitive baseline, particularly for capturing intensity and texture features; however, its performance is limited by reliance on manual feature design. In contrast, convolutional neural networks (DeepLabv3) automatically learn hierarchical representations that integrate local details with global context, yielding more accurate and robust segmentation.

Moreover, systematic experiments revealed that both backbone architecture and dataset size influence segmentation performance. Among the tested variants, ResNet-50 provided the best balance between computational efficiency and accuracy, outperforming lightweight alternatives such as MobileNetV3 while avoiding the diminishing returns observed with deeper backbones like ResNet-101. Dataset size experiments further showed that reliable performance can be achieved with as few as 500 annotated images, although larger datasets improve training stability and modestly enhance void accuracy. These results highlight the potential to reduce annotation costs without a substantial loss of accuracy.

Different biaxial experiments covering a wide range of stress states are performed, and the corresponding SEM images are used to validate the proposed deep learning approach. The analysis reveals a clear transition in the micro-void ratio with respect to stress triaxiality, with a critical value at $\bar{\eta} = 0.38$. A transition region around $\bar{\omega} = -0.44$ is separately identified in the Lode parameter analysis; however, under tensile-dominated conditions, damage and fracture are primarily governed by stress triaxiality. Therefore, the critical stress triaxiality ($\bar{\eta} = 0.38$) is adopted as the primary criterion. Moreover, the mean area of the micro-voids increases with rising stress triaxialities, while the void ratio remains nearly constant, indicating that void coalescence contributes more significantly to ductile damage and fracture behavior.

The above findings clearly indicate that deep learning substantially outperforms thresholding and feature-based machine learning, particularly in capturing heterogeneous contrast and subtle microstructural features. Moreover, the proposed workflow achieves a practical balance between accuracy, computational efficiency, and annotation effort.

CRediT authorship contribution statement

Zhichao Wei: Writing – original draft, Visualization, Methodology, Investigation, Data curation, Conceptualization. **Yijia Mo:** Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Shuwei Zhou:** Writing – review & editing, Validation, Investigation, Data curation.

Steffen Gerke: Writing – review & editing, Validation, Methodology, Investigation, Conceptualization. **Sebastian Münstermann:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Michael Brünig:** Writing – review & editing, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The project has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 322157 331, this financial support is gratefully acknowledged. The SEM images of the fracture surfaces presented in this paper were performed at the Institut für Werkstoffe im Bauwesen, University of the Bundeswehr Munich and the support of Wolfgang Saur is gratefully acknowledged. The authors gratefully acknowledge the computing resources provided by RWTH Aachen University, Germany, under projects rwth1791 and rwth1654.

Appendix A. Geometries of biaxially loaded specimens and representative experimental and numerical results

See Figs. A.21 and A.22.

Appendix B. Stress-state-dependent damage criterion based on Brünig et al.'s theory [16]

Based on macroscopic experimental and numerical results from simple shear tests and uniaxial tensile tests with different notch radii, the transition between shear-dominated, mixed-mode, and tensile-dominated damage is defined using the third deviatoric stress invariant and the Lode parameter, which take zero values for the corresponding reference stress states. Accordingly, the damage condition is expressed as follows.

For tension-dominated stress states, with the critical value $\eta \geq \frac{1}{\sqrt{3}}$, the damage condition is given by

$$f^{\mathrm{da}} = \frac{1}{3}\,\mathrm{tr}(\mathbf{T}) - \tilde{\sigma} = 0\,. \tag{B.1}$$

For shear-dominated stress states, with $\eta \leq 0$, the damage condition is defined as

$$f^{\mathrm{da}} = \sqrt{\frac{1}{2}\,\mathrm{dev}\mathbf{T}\cdot\mathrm{dev}\mathbf{T}} - \tilde{\sigma} = 0\,. \tag{B.2}$$

For mixed stress states ($0 < \eta < \frac{1}{\sqrt{3}}$), the damage condition is formulated as

$$f^{\mathrm{da}} = \frac{1}{3}\,\mathrm{tr}(\mathbf{T}) + \hat{\beta}\sqrt{\frac{1}{2}\,\mathrm{dev}\mathbf{T}\cdot\mathrm{dev}\mathbf{T}} - \tilde{\sigma} = 0 \tag{B.3}$$

with

$$\hat{\beta} = 1 - (d\eta)^m\,, \tag{B.4}$$

where d and m are material constants.

Appendix C. Effect of backbone and dataset size on training and validation accuracy

See Fig. C.23.

Data availability

Data will be made available on request.

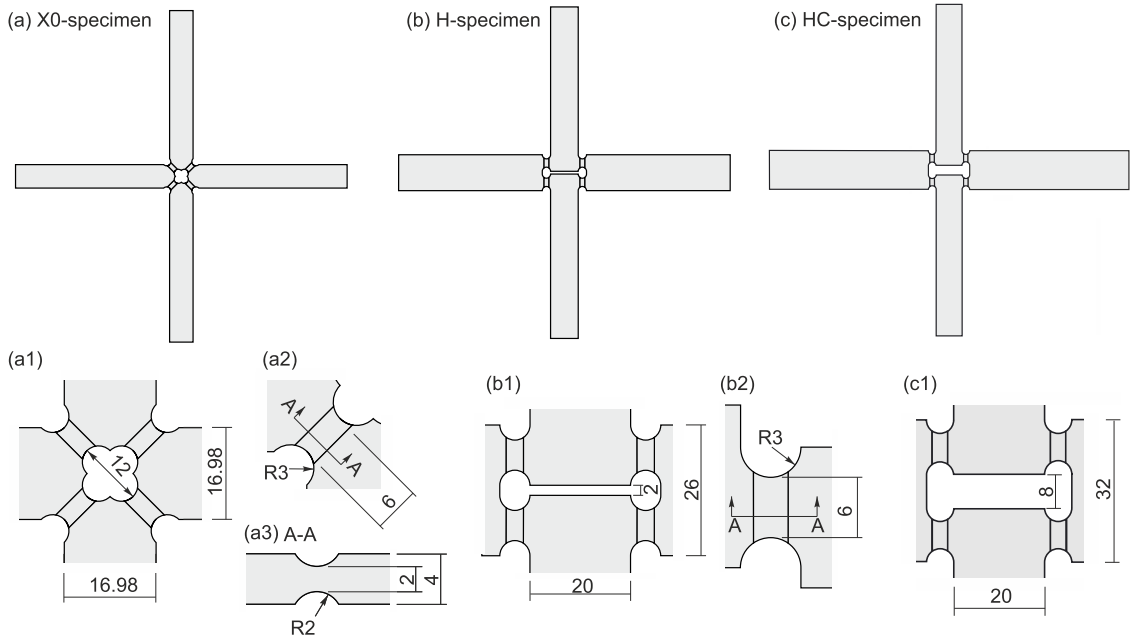

Fig. A.21. Cruciform specimens used for the designed biaxial experiments: (a) X0-specimen [69], (b) H-specimen [69], and (c) HC-specimen [79]. (a1)–(a3), (b1)–(b2), and (c1) show the corresponding detailed dimensions of the central opening, notched regions, and cross-sectional views. All dimensions are given in mm.

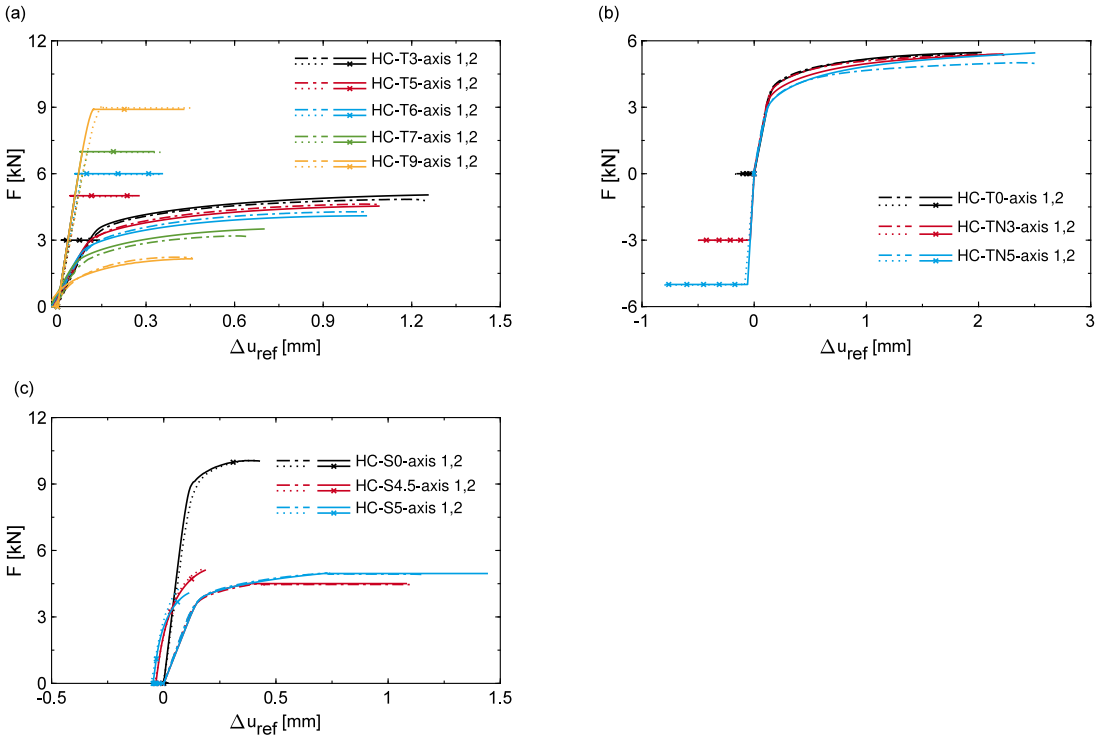

Fig. A.22. Force–displacement curves from HC experiments (dashed lines) and numerical simulations (solid lines) for biaxial loading cases: (a) and (b) shear loading combined with tensile or compressive preloads [76]; (c) tensile loading combined with shear preloads [78]. In the subfigures, T0–T9 indicate tensile preloads, TN3–TN5 indicate compressive preloads, and S0–S5 indicate shear preloads, where the numbers specify the applied magnitude. The figure is reproduced from [26] under the terms of the Creative Commons CC BY 4.0 license.

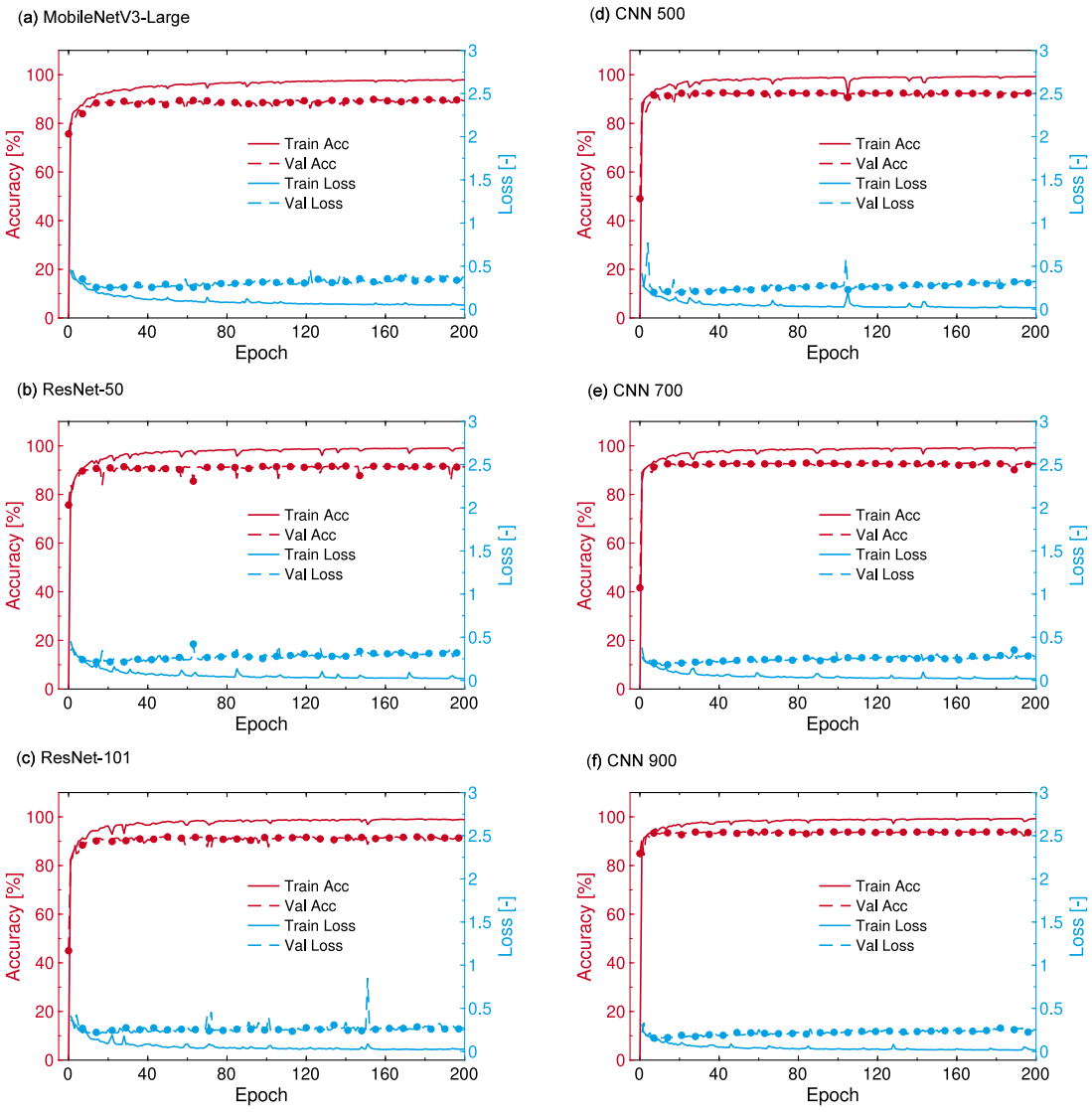

Fig. C.23. Comparison of training and validation accuracies and losses across different backbones and dataset sizes: (a)–(c) relation between training and validation accuracy and loss and training epoch for deep learning models using the MobileNetV3-Large, ResNet-50, and ResNet-101 backbones, respectively, and (d)–(f) the corresponding results for CNN models trained with dataset sizes of 500, 700, and 900 SEM images. Solid lines denote training accuracy and loss, whereas dashed lines indicate validation accuracy and loss.

References

[1] Zeng C, Fang X. A new yield surface model considering non-uniform strain rate hardening behavior. Int J Mech Sci 2025;301:110504.

[2] Zeng C, Fang X. Specimen geometry design for plasticity and fracture characterization of sheet metal under high testing speed and various stress states. Thin-Walled Struct 2023;186:110688.

[3] Yu F, Jar P-YB, Hendry MT. Constitutive analysis of pressure-insensitive metals under axisymmetric tensile loading: A stress triaxiality-dependent plasticity damage model. Int J Mech Sci 2018;142–143:21–32.

[4] Shang X, Zhang H, Wang L, Zhu G, Cui Z, Fu MW, Zeng X. The effect of stress state and strain partition mode on the damage behavior of a Mg-Ca alloy. Int J Plast 2021;144:103040.

[5] Wu H, Zhuang X, Zhao Z. Extended GTN model for predicting ductile fracture under a broad range of stress states. Int J Solids Struct 2022;239–240:111452.

[6] Tomstad AJ, Boåsen M, Faleskog J, Børvik T, Hopperstad OS. On the influence of stress state on ductile fracture of two 6000-series aluminium alloys with different particle content. Int J Solids Struct 2023;269:112149.

[7] Li Z, Chang Y, Liu W, Lian J. Predicting edge fracture in dual-phase steels: Significance of anisotropy-induced localization. Int J Mech Sci 2024;274:109255.

[8] Dou W, Xu Z, Han Y, Huang F. A ductile fracture model incorporating stress state effect. Int J Mech Sci 2023;241:107965.

[9] Chandran S, Verleysen P. Non-monotonic plasticity and fracture in dp1000: Stress-state, strain-rate and temperature influence. Int J Mech Sci 2024;267:109011.

[10] Zhang P, Chen A, Xing J, Wang Y. New unified model for predicting ductile fracture in structural steels. Int J Mech Sci 2025;111085.

[11] Zhao W, Zhao C, Zhang G, Wang T, Huang Q. Damage evolution under negative stress triaxiality: Experiments and modeling. Int J Mech Sci 2025;305:110777.

[12] Pan B, Zhu L, Wei Z, Tekkaya B, Stebner S, Münstermann S. Analyses of stress-state-dependent ductile damage and fracture behavior of zirconium. Materials 2025;19(1).

[13] Wei Z, Mao G, Li D, Tekkaya B, Stebner S, Münstermann S. Ductile-to-brittle transition in high-strength steel: Experiments and damage-coupled constitutive modeling. Int J Solids Struct 2026;331:113896.

[14] Murakami S. Mechanical modeling of material damage. J Appl Mech 1988;55(2):280–6.

[15] Brünig M. An anisotropic ductile damage model based on irreversible thermodynamics. Int J Plast 2003;19(10):1679–713.

[16] Brünig M, Chyra O, Albrecht D, Driemeier L, Alves M. A ductile damage criterion at various stress triaxialities. Int J Plast 2008;24(10):1731–55.

[17] Lou Y, Yoon JW. Alternative approach to model ductile fracture by incorporating anisotropic yield function. Int J Solids Struct 2019;164:12–24.

[18] Brünig M, Koirala S, Gerke S. A stress-state-dependent damage criterion for metals with plastic anisotropy. Int J Damage Mech 2023;32(6). 1811–832.

[19] Rong H, Hu P, Ying L, Hou W, Dai M. Modeling the anisotropic plasticity and damage of AA7075 alloy in hot forming. Int J Mech Sci 2022;215:106951.

[20] Wu H, Zhuang X, Zhang W, Zhao Z. Anisotropic Gurson–Tvergaard–Needleman model considering the anisotropic void behaviors. Int J Mech Sci 2023;248:108229.

[21] Zhang Y, Hu Z, Yu Y, Xu X, Cai X, Peng Q, Zeng W, Zhong W, Blunt L, Jiang X. Anisotropic plastic deformation and damage evolution of sapphire under nanoindentation. Int J Mech Sci 2025;307:110769.

[22] Li C, Piao Y, Zhang F, Zhang Y, Hu Y, Wang Y. Understand anisotropy dependence of damage evolution and material removal during nanoscratch of mgf 2 single crystals. Int J Extrem Manuf 2023;5(1):015101.

[23] Piao Y, Li C, Hu Y, Cui H, Luo X, Geng Y, Zhang F. Nanoindentation induced anisotropy of deformation and damage behaviors of mgf2 crystals. J Mater Res Technol 2024;28:4615–25.

[24] Wei Z, Koirala S, Gerke S, Brünig M. Numerical analysis of anisotropic plasticity and damage based on the inelastic predictor-elastic corrector method. Int J Solids Struct 2026;326:113770.

[25] Wei Z, Harting M, Gerke S, Brünig M. Ductile damage analysis under extreme low-cycle biaxial shear loadings: Experiments and simulations. Int J Solids Struct 2025;313:113292.

[26] Wei Z, Mao G, Gerke S, Münstermann S, Brünig M. Experimental analysis and modeling of anisotropic ductile damage in non-proportional extreme low-cycle biaxial loading with shear-tension histories. Int J Plast 2025;194:104474.

[27] Roth CC, Morgeneyer TF, Helfen L, Mohr D, Tancogne-Dejean T. Quantifying damage mechanisms through fe-based void tracking: Application to shear and tension in-situ laminography experiments on aa2198-t851. Acta Mater 2025;288:120783.

[28] Mototake Y, Ito K, Demura M. Quantitative prediction of fracture toughness (k i c) of polymer by fractography using deep neural networks. Sci Technol Adv Mater 2022;2(1):310–21.

[29] Pal S, Sarkar T, Saha S, Mahapatra PRS. An autoencoder-based deep learning model forfractographic characterization of tungstenheavy alloys. Int J Image Graph Signal Process 2025;17(1):45–58.

[30] Schneider CA, Rasband WS, Eliceiri KW. Nih image to imagej: 25 years of image analysis. Nature Methods 2012;9(7):671–5. http://dx.doi.org/10.1038/nmeth.2089.

[31] Grove C, Jerram DA. jpor: An imagej macro to quantify total optical porosity from blue-stained thin sections. Comput Geosci 2011;37(11):1850–9. http://dx.doi.org/10.1016/j.cageo.2011.03.002.

[32] Jie M, Chow CL, Wu X. Damage-coupled FLD of Sheet Metals for Warm Forming and Nonproportional Loading. Int J Damage Mech 2011;20(8):1243–62.

[33] Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 1979;9:62–6.

[34] Wang X, Xi Z, Zhang S, Tang S, Yan Z. Quantitative characterization of different pore types in marine shale based on machine learning and sem images. Nat Resour Res 2025;34(5):2559–78. http://dx.doi.org/10.1007/s11053-025-10506-w.

[35] Breiman L. Random forests. Mach Learn 2001;45(1):5–32. http://dx.doi.org/10.1023/A:1010933404324.

[36] He H, Garcia EA. Learning from imbalanced data. IEEE Trans Knowl Data Eng 2009;21(9):1263–84.

[37] Osenberg M, Hilger A, Neumann M, Wagner A, Bohn N, Binder JR, Schmidt V, Banhart J, Manke I. Classification of fib/sem-tomography images for highly porous multiphase materials using random forest classifiers, http://dx.doi.org/10.48550/arXiv.2207.14114.

[38] Röding M, Fager C, Olsson A, von Corswant C, Olsson E, LorÉn N. Three-dimensional reconstruction of porous polymer films from fib-sem nanotomography data using random forests. J Microsc 2021;281(1):76–86.

[39] Tsopanidis S, osovski S. Unsupervised machine learning in fractography: Evaluation and interpretation. Mater Charact 2021;182:111551.

[40] Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998;86(11):2278–324.

[41] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Commun ACM 2017;60(6):84–90.

[42] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. 2015 IEEE conference on computer vision and pattern recognition. CVPR, IEEE; 2015, p. 3431–40. http://dx.doi.org/10.1109/CVPR.2015.7298965.

[43] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation, http://dx.doi.org/10.48550/arXiv.1505.04597.

[44] Chen L-C, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for semantic image segmentation, http://dx.doi.org/10.48550/arXiv.1706.05587.

[45] Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Trans Pattern Anal Mach Intell 2018;40(4):834–48.

[46] Bihani A, Daigle H, Santos JE, Landry C, Prodanović M, Milliken K. Mudrocknet: Semantic segmentation of mudrock sem images through deep learning. Comput Geosci 2022;158:104952. http://dx.doi.org/10.1016/j.cageo.2021.104952.

[47] Hirabayashi Y, Iga H, Ogawa H, Tokuta S, Shimada Y, Yamamoto A. Deep learning for three-dimensional segmentation of electron microscopy images of complex ceramic materials. Npj Comput Mater 2024;10(1).

[48] Bangaru SS, Wang C, Zhou X, Hassan M. Scanning electron microscopy (sem) image segmentation for microstructure analysis of concrete using u-net convolutional neural network. Autom Constr 2022;144:104602.

[49] Rühle B, Krumrey JF, Hodoroaba V-D. Workflow towards automated segmentation of agglomerated, non-spherical particles from electron microscopy images using artificial neural networks. Sci Rep 2021;11(1):4942.

[50] Yang S, Zhu J, Cao Z, Yang M, Wang Y. Sem image segmentation method for copper microstructures based on enhanced u-net modeling. Coatings 2025;15(8):969.

[51] Li C, Han X, Yao C, Guo Y, Li Z, Jiang L, Liu W, Huang H, Fu H, Ban X. A novel training-free approach to efficiently extracting material microstructures via visual large model. Acta Mater 2025;290:120962.

[52] Tung P-Y, Harrison RJ. Efficient microstructure segmentation in three-dimensional imaging: Combining few-shot learning with the segment anything model. Next Mater 2025;8:100663.

[53] Zhang Y, Wu X, You J. A benchmark dataset and baseline methods for rock microstructure interpretation in sem images. Sci Data 2025;12(1):1671.

[54] Xia Z, Ma K, Cheng S, Blackburn T, Peng Z, Zhu K, Zhang W, Xiao D, Knowles AJ, Arcucci R. Accurate identification and measurement of the precipitate area by two-stage deep neural networks in novel chromium-based alloys. Phys Chem Chem Phys 2023;25(23):15970–87.

[55] Wang X-J, Li Y-F, Gu T, Xiang P, Cheng S, Jia L-J. Data-driven void growth prediction of aluminum under monotonic tension using deep learning. J Constr Steel Res 2024;222:109002.

[56] Wang Y, Luo Q, Xie H, Li Q, Sun G. Digital image correlation (DIC) based damage detection for CFRP laminates by using machine learning based image semantic segmentation. Int J Mech Sci 2022;230:107529.

[57] Xie J, Hu P, Chen J, Han W, Wang R. Deep learning-based instantaneous cutting force modeling of three-axis CNC milling. Int J Mech Sci 2023;246:108153.

[58] Tian X, Li S, Zhang H, Li Q, Ai S. A digital twin framework for turbine blade crack propagation prediction. Int J Mech Sci 2025;297–298:110396.

[59] Liu Y, Zou M, Qi Y, Guo Z, Yu H, Shi L, Song J, Xu S, Fan W, Yu Q. Mechanical property prediction of complex-configuration mechanical metamaterials based on multimodal CNN. Int J Mech Sci 2025;306:110872.

[60] Rosenberger J, Tlatlik J, Münstermann S. Deep learning based initial crack size measurements utilizing macroscale fracture surface segmentation. Eng Fract Mech 2023;293:109686.

[61] Rosenberger J, Tlatlik J, Beckmann C, B. Rohrmüller S Münstermann. Extracting ductile cast iron microstructure parameters from fracture surfaces: A deep learning based instance segmentation approach. Eng Fract Mech 2024;311:110586.

[62] Zhang J, Cheng J, Li E, Li B. Data-driven inverse design framework for broadband mode-selective elastic metasurfaces. Int J Mech Sci 2026;309:111008.

[63] Zhang Y, Wang Y, Tao Q, Liu Y, Wang C. Deep learning of buckling instability in geometrically symmetry-breaking kirigami. Int J Mech Sci 2024;280:109331.

[64] Liu X, Duan L, Li X, Zhu P, Yan Z, Zhong Z. A deep learning framework for predicting ceramic viscosity during sintering. Int J Mech Sci 2025;111067.

[65] Papasidero J, Doquet V, Mohr D. Ductile fracture of aluminum 2024-T351 under proportional and non-proportional multi-axial loading: Bao–Wierzbicki results revisited. Int J Solids Struct 2015;69–70:459–74.

[66] Cortese L, Nalli F, Rossi M. A nonlinear model for ductile damage accumulation under multiaxial non-proportional loading conditions. Int J Plast 2016;85:77–92.

[67] Chen K, Scales M, Kyriakides S. Material response, localization and failure of an aluminum alloy under combined shear and tension: Part II analysis. Int J Plast 2019;120:361–79.

[68] Hou Y, Lee M-G, Lin J, Min J. Experimental characterization and modeling of complex anisotropic hardening in quenching and partitioning (Q&P) steel subject to biaxial non-proportional loadings. Int J Plast 2022;156:103347.

[69] Gerke S, Adulyasak P, Brünig M. New biaxially loaded specimens for the analysis of damage and fracture in sheet metals. Int J Solids Struct 2017;110–111:209–18.

[70] Raj A, Verma RK, Singh PK, Shamshoddin S, Biswas P, Narasimhan K. Experimental and numerical investigation of differential hardening of cold rolled steel sheet under non-proportional loading using biaxial tensile test. Int J Plast 2022;154:103297.

[71] Brünig M, Zistl M, Gerke S. Numerical analysis of experiments on damage and fracture behavior of differently preloaded aluminum alloy specimens. Metals 2021;11(3):381.

[72] Zistl M, Brünig M, Gerke S. Analysis of damage and fracture behavior in ductile metal sheets undergoing compression and shear preloading. Int J Mater Form 2022;15(4):59.

[73] Gerke S, Zistl M, Bhardwaj A, Brünig M. Experiments with the X0-specimen on the effect of non-proportional loading paths on damage and fracture mechanisms in aluminum alloys. Int J Solids Struct 2019;163:157–69.

[74] Gerke S, Zistl M, Brünig M. Experiments and numerical simulation of damage and fracture of the X0-specimen under non-proportional loading paths. Eng Fract Mech 2020;224:106795.

[75] Wei Z, Zistl M, Gerke S, Brünig M. Analysis of ductile damage and fracture under reverse loading. Int J Mech Sci 2022;228:107476.

[76] Wei Z, Gerke S, Brünig M. Damage and fracture behavior under non-proportional biaxial reverse loading in ductile metals: Experiments and material modeling. Int J Plast 2023;171:103774.

[77] Wei Z, Gerke S, Brünig M. Mechanical responses of ductile aluminum alloy under biaxial non-proportional tensile reverse loading patterns. Metals 2023;13(12):1922.

[78] Wei Z, Gerke S, Brünig M. Numerical analysis of non-proportional biaxial reverse experiments with a two-surface anisotropic cyclic plasticity-damage approach. Comput Methods Appl Mech Engrg 2024;419:116630.

[79] Wei Z, Gerke S, Brünig M. Novel uniaxial and biaxial reverse experiments for material parameter identification in an advanced anisotropic cyclic plastic-damage model. Mech Mater 2025;205:105294.
[80] Wei Z, Gerke S, Brünig M. Ductile damage and fracture characterizations in bi-cyclic biaxial experiments. Int J Mech Sci 2024;276:109380.
[81] Gerke S, Wei Z, Brünig M. Experiments on low–cycle ductile damage and failure under biaxial loading conditions. Exp Mech 2024;64:1021–36.
[82] Drucker DC, Prager W. Soil mechanics and plastic analysis or limit design. Quart Appl Math 1952;10(2):157–65.
[83] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2016 IEEE conference on computer vision and pattern recognition. CVPR, IEEE; 2016, p. 770–8. http://dx.doi.org/10.1109/CVPR.2016.90.
[84] Russell BC, Torralba A, Murphy KP, Freeman WT. Labelme: A database and web-based tool for image annotation. Int J Comput Vis 2008;77(1–3):157–73.
[85] Xiong Y, Varadarajan B, Wu L, Xiang X, Xiao F, Zhu C, Dai X, Wang D, Sun F, Iandola F, Krishnamoorthi R, Chandra V. Efficientsam: Leveraged masked image pretraining for efficient segment anything, http://dx.doi.org/10.48550/arXiv.2312.00863.
[86] Howard A, Sandler M, Chu G, Chen L-C, Chen B, Tan M, Wang W, Zhu Y, Pang R, Vasudevan V, Le V Q, Adam H. Searching for MobileNetV3. In: iCCV 2019. 2019, http://dx.doi.org/10.48550/arXiv.1905.02244.
[87] Kanvinde AM, Deierlein GG. Cyclic void growth model to assess ductile fracture initiation in structural steels due to ultra low cycle fatigue. J Eng Mech 2007;133(6):701–12.
[88] Zhang J, Ruan Z, Li Q, Zhang Z-Q. Toward robust and efficient musculoskeletal modeling using distributed physics-informed deep learning. IEEE Trans Instrum Meas 2023;72:1–11.
[89] Lahariya M, Karami F, Develder C, Crevecoeur G. Physics-informed lstm network for flexibility identification in evaporative cooling system. IEEE Trans Ind Inf 2023;19(2):1484–94.