Archives of Computational Methods in Engineering
https://doi.org/10.1007/s11831-026-10641-4

REVIEW

AI-Driven Image Processing for Microstructure and Surface Characterization: A Systematic Review of Methods, Materials, and Applications

Ali Ercetin[1,2] · Ibrahim Arpaci[3,4] · Samet Memiş[5] · Amila Akagic[6] · Özgür Özgün[7]

Received: 1 March 2026 / Revised: 5 May 2026 / Accepted: 9 May 2026
© The Author(s) 2026

Abstract

Artificial intelligence (AI) has rapidly become a key enabler in materials science, facilitating automated, high-precision analysis of microstructure and surface characteristics. This review presents a structured overview of both deep learning (DL) and classical image processing approaches used in tasks such as segmentation, classification, super-resolution, and surface defect detection across different imaging techniques, including optical microscopy, scanning electron microscopy (SEM), atomic force microscopy (AFM), and electron backscatter diffraction (EBSD). The reviewed studies show that modern DL architectures such as convolutional neural networks (CNN), U-Net, Vision Transformers, and YOLO achieve high performance, often reporting accuracy levels above 90% in segmentation and defect detection tasks. In addition, generative models improve robustness by enabling data augmentation and the generation of synthetic data, particularly in data-limited scenarios. Despite these advances, important challenges remain, including limited data availability, domain shifts across imaging conditions, and limited interpretability of deep models. This review therefore examines explainable AI (XAI) approaches and highlights that visual explanations alone are insufficient unless supported by physical microstructural knowledge. At the same time, classical feature-based methods such as GLCM, LBP, and Gabor filters continue to offer advantages in interpretability and computational efficiency, especially in low-data scenarios. Overall, the review emphasizes the complementary use of classical and deep learning approaches and outlines future directions, including standardized datasets, domain adaptation, and physics-informed modeling to improve reliability and generalization in AI-driven materials characterization.

✉ Ali Ercetin
aercetin03@gmail.com

✉ Ibrahim Arpaci
ibrahimarpaci@uludag.edu.tr

Samet Memiş
smemis@bandirma.edu.tr

Amila Akagic
aakagic@etf.unsa.ba

Özgür Özgün
oozgun@bingol.edu.tr

[1] Department of Naval Architecture and Marine Engineering, Maritime Faculty, Bandırma Onyedi Eylül University, Bandırma 10200, Türkiye

[2] University College, Korea University, 145 Anam-ro, Seoul 02481, Republic of Korea

[3] Department of Computer Engineering, Faculty of Engineering, Bursa Uludağ University, Bursa 16059, Türkiye

[4] Department of Computer Science and Engineering, College of Informatics, Korea University, Seoul 02841, Republic of Korea

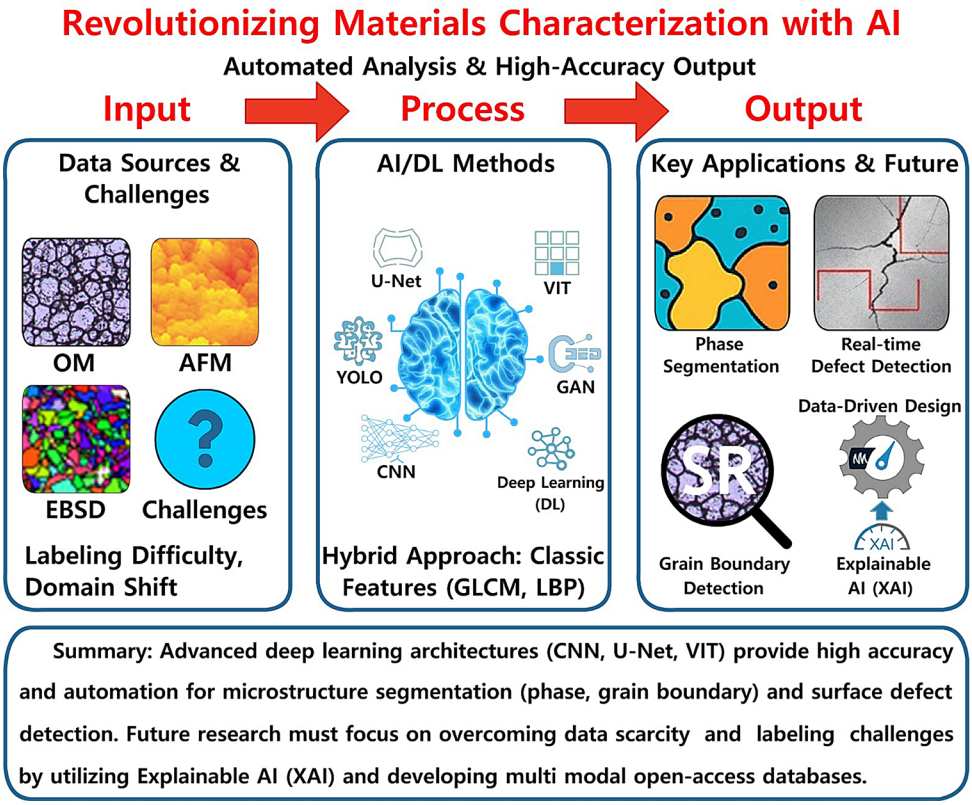
[5] Department of Marine Engineering, Faculty of Maritime, Bandırma Onyedi Eylül University, Bandırma 10200, Türkiye

[6] Department of Computing and Informatics, Faculty of Electrical Engineering, University of Sarajevo (UNSA), Bosnia and Herzegovina, Sarajevo, Bosnia and Herzegovina

[7] Occupational Health and Safety, Faculty of Health Sciences, Bingol University, Bingol, Türkiye

Published online: 30 May 2026

Graphical Abstract



1 Introduction

The mechanical, thermal, and chemical properties of materials are closely related to their microstructures. Microstructure encompasses elements such as phase distribution, grain size, orientation, porosity, and interface properties, and these components are fundamental factors that determine a material's performance [1]. Detailed characterization of microstructure is a critical step in understanding the "microstructure-property relationship," a fundamental goal of materials science. Unraveling this relationship plays a key role in the design of new materials and the improvement of existing ones [2]. Traditional approaches to microstructure analysis often rely on qualitative evaluation of microscopic images (Figs. 1 and 2). However, visual inspection alone is insufficient for accurately defining complex phase morphologies and multi-scale structural features [3], which has increased the need for quantitative and automated analysis methods.

In materials engineering applications, surface topography is as decisive factor as microstructure. Parameters such as surface roughness, waviness, texture orientation, and surface energy directly affect surface-related behaviors, including wear, friction, corrosion, and coating adhesion [6]. In addition, advances in high-resolution imaging techniques such as optical microscopy, SEM, AFM, and EBSD have enabled the acquisition of large, complex datasets, further increasing the need for advanced analytical methods. Understanding the process–microstructure–property relationship in modern engineering materials is not limited to experimental characterization. It also encompasses the microstructural effects of thermomechanical, diffusion, and phase-transformative processes [7]. In this context, microstructure characterization has become an interdisciplinary field integrating experimental observations and modeling approaches.

Microstructure characterization has long relied heavily on human labor and expert interpretation. However, today, classical image processing techniques alone are insufficient for understanding the complex morphologies of microstructures [8]. As a result, AI and computer vision (CV) methods have emerged as effective tools for enabling automated, rapid, and reproducible analysis of large-scale microscopy data [9]. Specifically, the creation of broader, open-access databases for AI-supported microstructure analysis (Fig. 3), the development of interpretable DL models, and the integration of multi-scale characterization approaches are considered fundamental requirements for progress in this field [10].

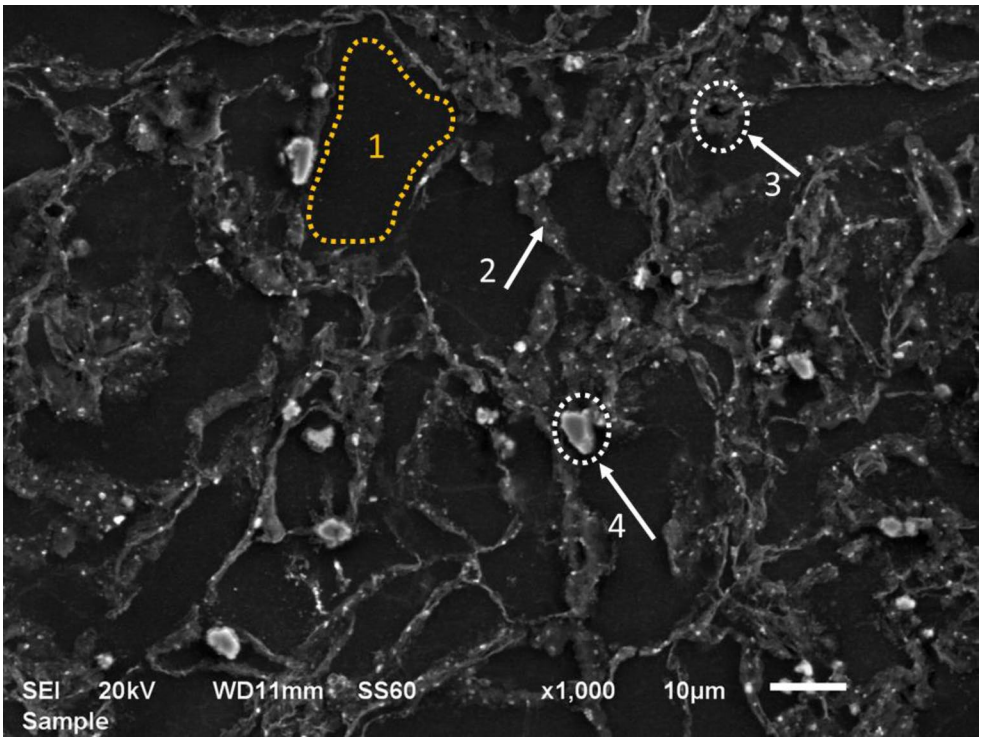


Fig. 1 SEM image (1000× magnification) of 2 wt% Nb-reinforced TAZ532 matrix composite [4]

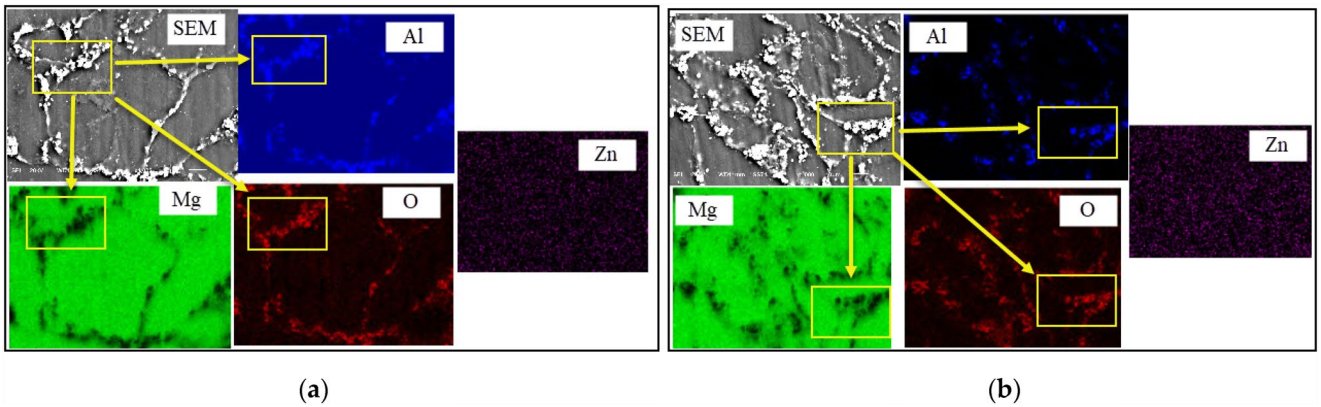


Fig. 2 SEM/EDS mapping analysis before immersion (**a**) Mg2Zn-2 wt% Al2O3 magnesium composite, (**b**) Mg2Zn-8 wt% Al2O3 magnesium composite [5]

Image processing and DL-based models can identify microstructural patterns with significantly higher accuracy and speed than human observers, particularly in material systems such as steel microstructures, and detect complex phase distributions [11]. For example, the classification of steel microstructures using DL methods has enabled groundbreaking advances in the automatic determination of phase separation and grain morphology [11]. Such approaches not only shorten the analysis time but also significantly reduce human error [8]. One of the most important contributions of DL models is their ability to jointly interpret differences in parameterizations of the same material, particularly under varying production conditions such as heat treatment, cooling rates, or deformation processes [9]. Furthermore, applying pre-trained convolutional neural network-based encoders to microscope data has enabled high-accuracy segmentation even with limited data [12].

Another strength of AI-based image processing techniques is their ability to reconstruct microstructure data at different resolutions. DL-based super-resolution models enable more precise, detailed examination of microstructures by converting low-resolution microscope images into high-resolution ones [13]. Such approaches provide high sensitivity, particularly in detecting surface or phase distinctions at the nanoscale [14]. Furthermore, "few-shot learning" approaches offer significant advantages in situations where data labeling is complex or expert interpretation is limited. Using these methods, fast and flexible segmentation

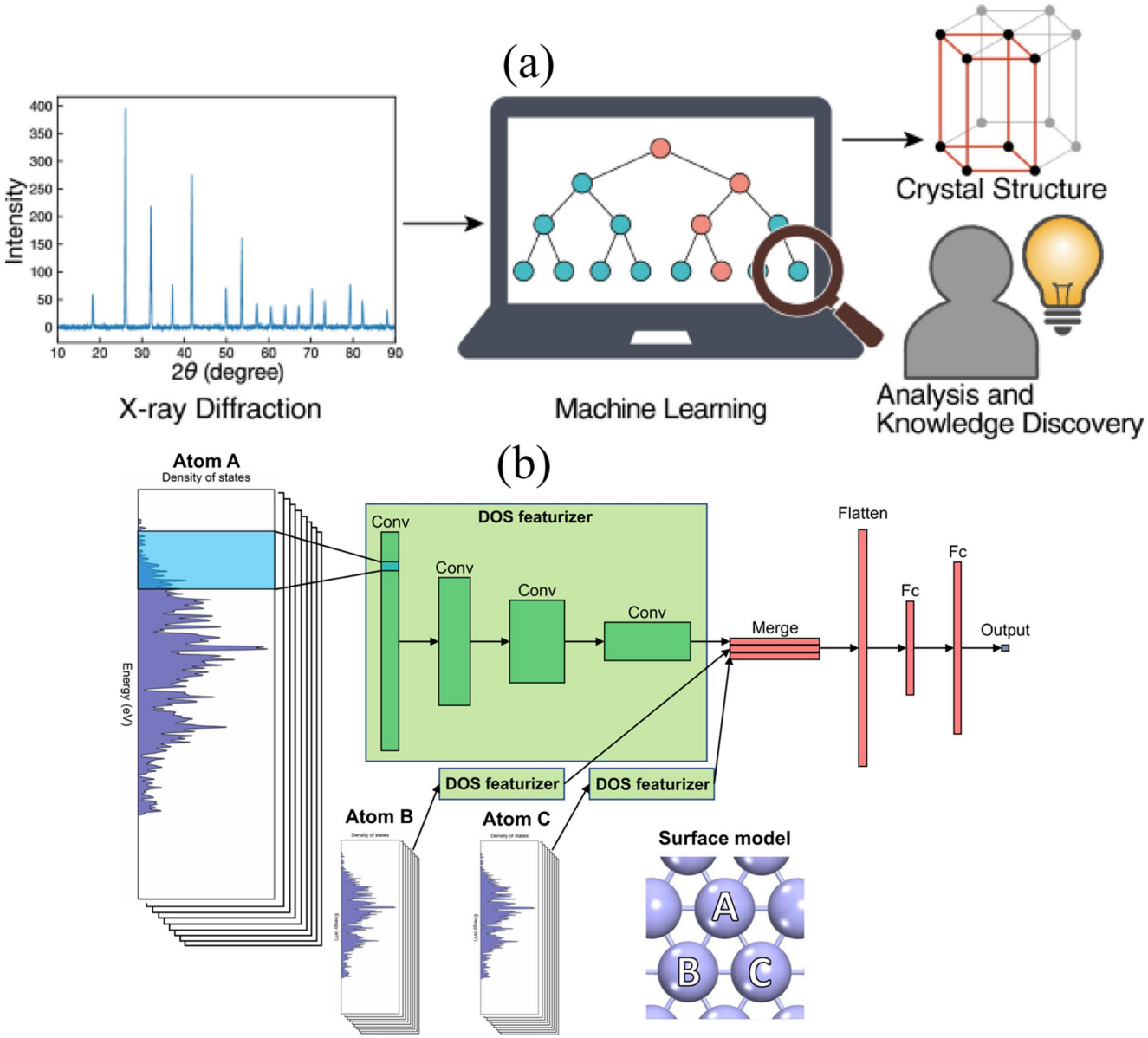


Fig. 3 Practical demonstrations of deep learning methods in the processing and analysis of spectral information [10]

of electron microscope data can be achieved using only a small number of labeled images [15]. Similarly, DL-based 3D segmentation techniques enable volumetric analysis of grain boundaries and phase distributions in polycrystalline materials [16]. One of the biggest challenges in DL-based microstructure analysis is the lack of high-quality and accurately labeled data. Classical interpretation of microscope images is time-consuming and requires expert knowledge [17]. This makes it difficult to obtain sufficient data for model training and limits the generalization ability. In recent years, data augmentation and synthetic data generation strategies have been developed to mitigate this problem. However, it has also been reported that these techniques may increase the risk of overfitting and may not fully capture the model's actual microstructural variations. In learning environments with limited data, optimization strategies are used to enable the model to generalize effectively; these approaches are referred to as "training strategies" [18]. On the other hand, understanding the decision mechanisms of deep networks in high-resolution microscopy data poses a separate challenge. Understanding which features or textural patterns deep networks learn and decide upon is critical for increasing the model's reliability and interpretability [19].

Recently, vision foundation models have emerged as a viable option for image analysis in natural images, which could be considered as a direct result of developments in natural language processing [20, 21]. The most prominent approach is Segment Anything (SAM) [22]. These

foundation models are able to generalize to tasks and data distributions beyond those seen during training, which is a capability enabled through prompt engineering. Subsequently, Segment Anything for Microscopy (μSAM) was introduced in [23] as a tool for segmentation and tracking in multidimensional microscopy data. Based on SAM, it was extended by fine-tuning generalist model for light and electron microscopy with a goal of greatly improving segmentation quality across different modalities. Tung and Harrison [24] propose a machine learning (ML) workflow that combines SAM with a few-shot learning framework, automating segmentation and minimizing user bias. In this work, SAM is used to generate precise annotations from a limited subset of 2D images with prompt engineering, and then used for few-shot learning model. The method was benchmarked with a 3D FIB-SEM tomographic dataset. The method achieved intersection over union (IoU) score of 80.62%, which reportedly achieves better IoU score then previous methods. The material used in this study was metamorphosed CM-like carbonaceous chondrite meteorite (WIS91600). It should be noted that SAM was trained on very large dataset of natural images (i.e., scenes containing people, animals, objects), and it has learned to recognize objects with clear boundaries, familiar shapes and textures, and semantic objects (e.g., cars, trees, persons, etc.). Microstructure images, such as SEM or FIB-SEM in materials science, have very different structures. They might contain abstract patterns that are not typically found in natural images. The contrast between specific objects can be very subtle or noisy. Structures might be defined by material properties, and not visual semantics. SAM is biased toward the statistics and patterns of natural images, and as such it will try to interpret microstructures using the rules learned from natural images. It might miss important features such as grains or phases, or it might incorrectly segment noise or background. Pan et al. [25], authors found that the performance of SAM is slightly inferior to ECA-U-Net (modified version of the U-Net architecture with Efficient Channel Attention) when it comes to details. However, the benefit of using SAM is that the labeling of microstructures is not required, as compared to ECA-U-Net. Whitman and Latypov [26] the authors presented results of experiments with CLIP, DINOv2 and SAM on microstructure images (SEM, EBSD, and metals and alloys dataset). They showed that these models outperformed traditional, task-specific models for learning microstructure–property relationships. Vision foundation models represent a promising tool for segmentation in multidimensional microscopy data, but their effectiveness should be further validated through studies that address specific scenarios within the materials science domain.

Another important concept has gained increasing importance in the materials science community, and this explainable AI. For example, Grad-CAM-based approaches have been used to highlight phase boundaries and relevant microstructural features in steel microstructures [19, 27], thereby providing improved insight into model decision-making processes. Explainable models aim not only to enable AI systems to make predictions but also to reveal which microstructural features contribute most to the prediction process [27]. This approach facilitates the correlation of model outputs with physical fundamentals, thereby increasing transparency and reliability in materials science research [27]. Furthermore, standardizing datasets and code developed by different research groups is critical for reproducibility. Pre-trained deep networks and shared benchmark datasets for microstructure segmentation provide an objective basis for model comparisons [12]. However, most studies in the literature still focus on a limited number of material types, microscopy techniques, or imaging conditions. This limits the models' generalizability across different data types and their effectiveness in multi-material systems [10].

The main objective of this review is to systematically present the current state of AI-supported image processing approaches in microstructure and surface characterization, their methodological diversity, and their application areas. Although DL models currently achieve high success rates in extracting meaningful structural information from microscope data, many obstacles remain, including data diversity, labeling challenges, and limited interpretability [10]. Therefore, this review covers not only technical advances but also the limitations of these technologies and future research needs. In this context, the review aims to provide a structured overview of existing methods, highlight key challenges, and outline future research directions.

2 Methods

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines. The literature search was performed in Scopus using three keyword combinations: (i) "image processing" AND "microstructure," (ii) "microstructure analysis" AND "machine learning," and (iii) "microstructure analysis" AND "deep learning," yielding an initial pool of 4,013 records. As the search was confined to a single database, no duplicate removal was necessary at the identification stage. Figure 4 presents PRISMA 2020 flow diagram illustrating the systematic literature search and selection process.

A publication filter was applied to restrict the pool to studies published from 2015 onward, a threshold that

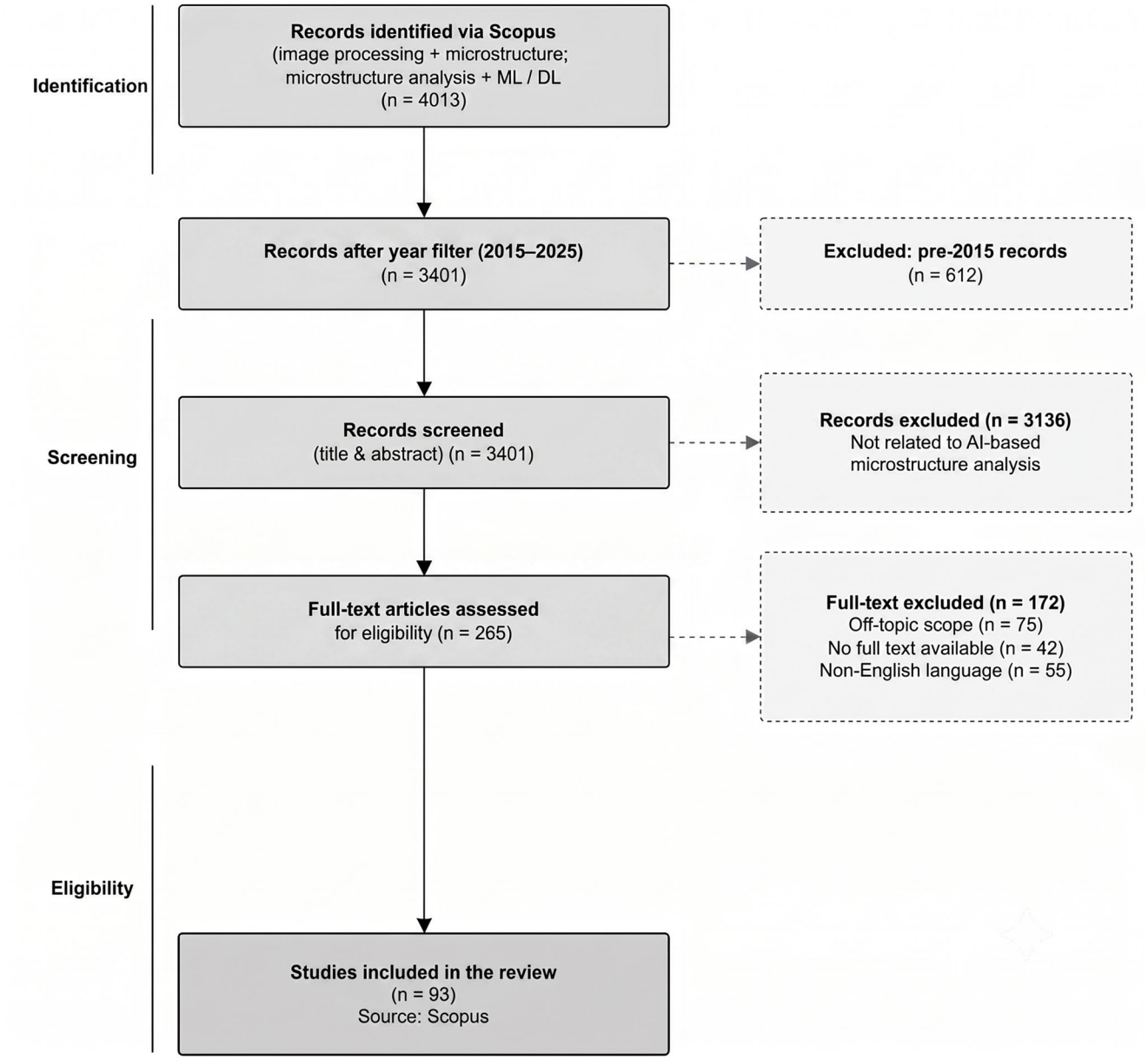


Fig. 4 PRISMA 2020 flow diagram illustrating the systematic literature search and selection process. The search was conducted in Scopus using three keyword combinations targeting AI-based microstructure analysis. The year distribution and leading journals of the 93 included studies are presented alongside the flow diagram

corresponds to the period of accelerating deep learning adoption in materials science. This step reduced the pool to 3,401 records; the 612 pre-2015 entries were excluded from systematic consideration but selectively retained where they provided foundational methodological grounding. Subsequent title and abstract screening of all 3,401 records resulted in the exclusion of 3,136 entries because they did not address AI-based or image-processing-driven microstructure analysis. The remaining 265 records proceeded to full-text assessment. At this stage, 172 were excluded: 75 fell outside the defined topical scope, 42 could not be retrieved in full text, and 55 were published in languages other than English. The 93 studies that met all inclusion criteria were included in the final synthesis.

3 Data Ecosystem and Open-Access Data Repositories

3.1 Image Acquisition Modalities (Optical Microscopy, SEM, AFM, EBSD)

In materials science research, image data is the most fundamental source of information for analyzing microstructures. Data acquired using different microscopy techniques, such as optical microscopy (OM), scanning electron microscopy (SEM), atomic force microscopy (AFM), and electron backscatter diffraction (EBSD), are critical for training and validating AI-based image-processing approaches [12]. These techniques provide multidimensional information about material surface morphology, grain structure, phase distribution, and texture, forming the basis for data-driven materials engineering [28]. Figure 5 summarizes the dataset: (a) and (b) contain longitudinal and transverse views of

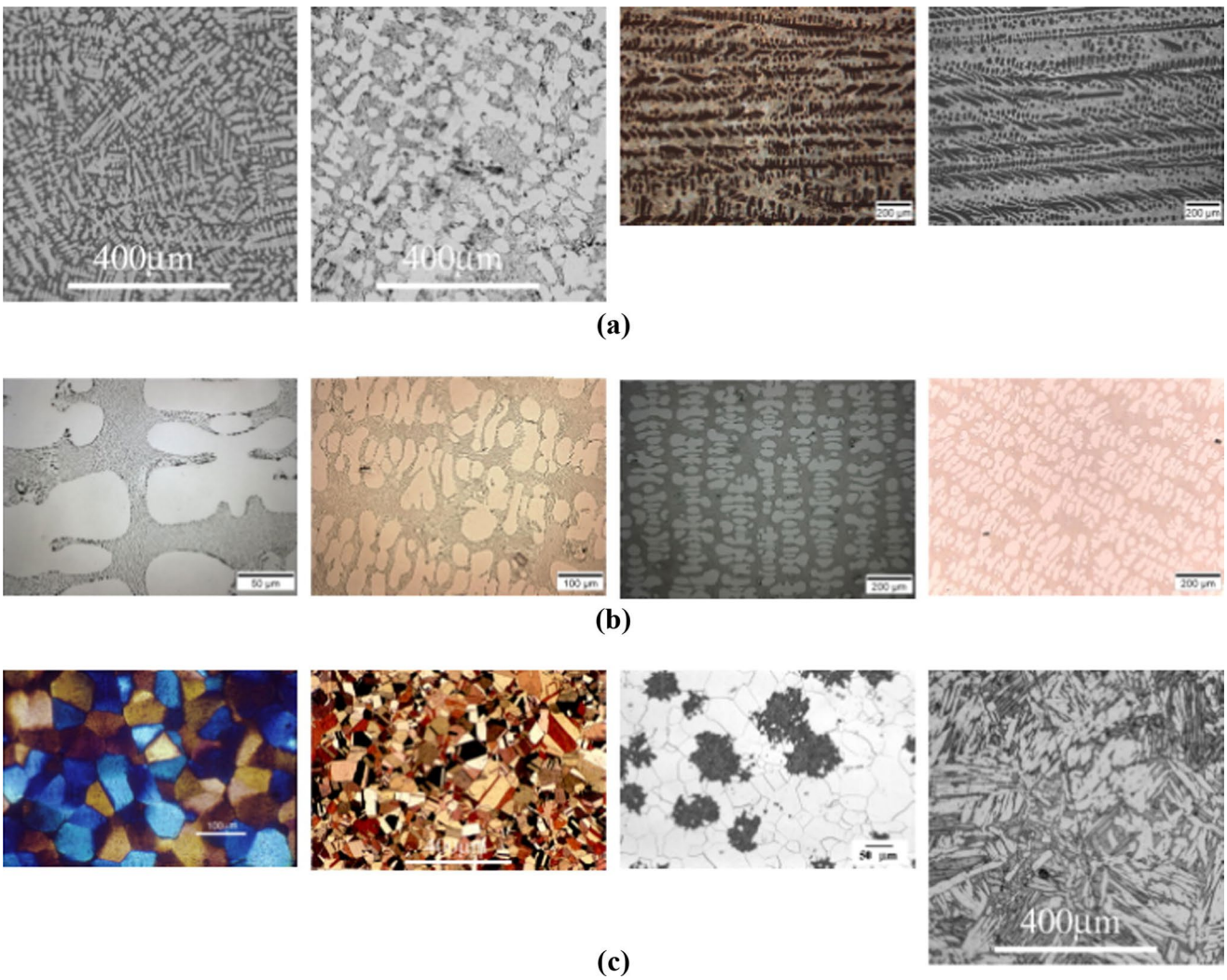


Fig. 5 Representative micrographs for the classification study. Panels (**a**) and (**b**) show longitudinal and transverse cross-sections of dendrites, respectively, while (**c**) contains no dendrites. Images (**a**)–(**c**) were used for Task 1; only (**a**) and (**b**) were used for Task 2 [29]

dendrites, in contrast to (c), which lacks dendritic features. Task 1 refers to the classification of micrographs based on the presence or absence of dendritic structures, where panels (a), (b), and (c) are used as input data. Task 2 focuses on distinguishing between different dendritic orientations (longitudinal and transverse), for which only panels (a) and (b) are used [29].

Today, researchers use experimental microscopy data not only for qualitative analysis but also to create datasets for training DL models. For example, the dataset developed by Pei et al. for 9% Cr steels enabled the prediction of material properties through ML by correlating high-resolution microscope images with phase morphology and grain structure [28]. Such datasets serve as essential references for understanding the connection between microstructure and mechanical behavior, thereby strengthening data-driven modeling. Similarly, Zhao and colleagues presented a comprehensive dataset on the microstructure and mechanical properties of a $(CrCoNi)_{97}Al_{1.5}Ti_{1.5}$ medium-density alloy. In this study, microstructural changes resulting from torsion tests conducted at different temperatures were supported by high-resolution optical and electron microscope images [30]. The resulting dataset is a valuable resource for both understanding material behavior and using artificial intelligence models in multi-scale structural analysis.

Stuckner and colleagues' work demonstrates how images obtained from different microscopy techniques can be used for DL-based segmentation models. In the study, automatic segmentation of microstructural components was successfully achieved using pre-trained encoders applied to microscope data. This approach demonstrated that data from different sources, such as optical, SEM, or EBSD, can be combined into a common representation and that AI can learn generalizable features across different microscopy modalities [12].

3.2 Labeling, Annotation, and Ground-Truth Reliability

One of the biggest challenges in AI-based analysis of microstructure images is creating reliable "ground-truth" (accurate reference) data. DL models require large amounts of labeled data to produce highly accurate results. However, manual annotation of microstructure images is often quite limited due to its high expertise requirements and time-consuming nature [17]. This situation directly affects the accuracy of model training, especially in materials with complex phase distributions or multi-component microstructures [18].

Wang and colleagues developed a DL-based labeling method to mitigate this issue [17]. In this method, automatic segmentation results were progressively improved through expert user review, creating a "semi-automatic annotation" process based on human-machine collaboration. This approach reduces labeling time while enhancing the model's ability to adapt to diverse microstructural textures [17].

Insufficient labeling in data-scarce settings can limit the model's generalization ability. Therefore, Ma and colleagues proposed several "training strategies" strategies for segmenting steel microstructures with limited data [18]. These strategies include data augmentation, dynamic learning-rate adjustment, and loss-function balancing. This enables more effective feature extraction from a small number of labeled data and increases the model's accuracy [18].

Labeling and annotation processes also contribute to a better understanding of the microstructure manifold. A study by DeCost and Holm showed that different microstructure images can be represented in a high-dimensional feature space, allowing their similarities to be measured. This approach revealed that labeled data can be used not only for model training but also for discovering structural relationships between microstructures [9].

3.3 Open-Access Repositories, Shareability, and Domain Shift

The success of AI-based methods in microstructure and surface characterization depends on the availability of well-structured, accessible datasets. However, data sharing in this field remains limited; most research relies on internal or project-based closed data sources [12]. Several open-access datasets have become particularly useful benchmarks for developing and evaluating AI-based microstructure and microscopy image analysis methods. For example, the NFFA-Europe 100% SEM Dataset provides a large collection of SEM images categorized for microscopy image classification tasks, while the Ultrahigh Carbon Steel Micrographs dataset offers microstructure images acquired under systematically varied heat-treatment conditions together with processing metadata and analysis tools. Such datasets are valuable not only for training and testing segmentation or classification models but also for evaluating model robustness, transferability, and domain-shift behavior across different imaging conditions.

Another critical step in promoting scientific data sharing is the development of multimodal datasets. In particular, the study titled "Multi-modal Dataset of a Polycrystalline Metallic Material" published in the Scientific Data journal has created a valuable resource for researchers by providing an open dataset containing optical, SEM, and EBSD data of polycrystalline metallic materials [31]. Such datasets facilitate the integration of data from different microscopy modalities, simplifying both AI training and experimental analysis. However, data sharing is not merely an access issue; it is also crucial for data consistency and for

addressing distribution differences known as "domain shift." Domain shift arises from changes in the statistical distributions of images obtained from different devices, operators, or preparation conditions, and can negatively affect model performance [32]. Ma and colleagues have shown that data augmentation and diversification strategies in microscopic images are effective in mitigating this effect [32]. In particular, synthetic manipulations such as contrast shifting, rotation, noise addition, or scaling enable the model to generalize better to different data sources.

3.4 Data Augmentation and Synthetic Data Generation

The success of DL models depends on the diversity and quantity of training data. However, in materials science, datasets consisting of microscopic images are often limited in size, and labeling processes require high expertise, reducing the model's generalization ability [17]. To overcome this problem, data augmentation and synthetic data generation methods are increasingly being used in AI-assisted microstructure analysis [32].

Data augmentation techniques involve diversifying existing microscopic images through operations such as rotation, reflection, scaling, adding noise, and adjusting brightness and contrast. These operations make the model more resilient to variations in imaging conditions and microstructure. Ma and colleagues comprehensively investigated the effects of data augmentation in microscopic images on material data mining. They demonstrated that these techniques reduce the risk of overfitting in DL-based models [32].

Recent studies have shown that data augmentation methods can indirectly improve not only model performance but also annotation quality. Wang and colleagues reported that their DL-based labeling method, which utilizes data augmentation strategies, improved segmentation accuracy and made the labeling process more efficient. This approach both reduces human error in ground-truth data production and strengthens the model's ability to learn with limited data [17].

A comprehensive review published by Ma and colleagues in 2023 systematically classified data augmentation and synthetic data generation methods used in DL-based microstructure segmentation [33]. This study emphasized that, in addition to classical image transformations, generative models such as generative adversarial networks (GAN) and variational autoencoders (VAE) are increasingly effective at producing synthetic microscope images. Synthetic data enables more generalizable results by balancing model training, especially in systems with a limited number of real photos [33]. More recently, diffusion-based generative models have emerged as a powerful alternative to GAN- and VAE-based approaches for synthetic microstructure generation and reconstruction. Lee and Yun proposed a denoising diffusion model framework for microstructure reconstruction and demonstrated its applicability to microstructures with different spatial distributions and morphological characteristics [34]. Düreth et al. developed a conditional diffusion-based reconstruction approach and showed that diffusion models can generate realistic microstructure images while preserving descriptor-based statistical similarity and diversity [35]. Similarly, Azqadan et al. investigated denoising diffusion probabilistic models for predictive SEM microstructure image generation and demonstrated their potential for generating microstructures under both known and unseen processing conditions [36]. Compared with conventional GAN-based methods, diffusion models offer more stable training and can provide flexible conditional generation, although their computational cost and validation against physical microstructural descriptors remain important challenges.

Section Conclusions Microstructure analysis and AI-based modeling in materials science rely on a data ecosystem comprising multi-source, high-resolution image data. These data, obtained from different microscopy techniques such as optical microscopy, SEM, AFM, and EBSD, are critical not only for structural characterization but also for training and validating AI models. However, issues such as the need for reliable labeling ("ground-truth"), limited data quantity, and human-induced inconsistencies constrain model performance. Methods such as semi-automatic annotation, data augmentation, and synthetic data generation can improve data quality and enhance generalization. Taken together, these data-related considerations shape how microstructural information is represented and analyzed. The following section, therefore, focuses on classical feature extraction and representation methods.

4 Feature Engineering and Classical Image Representations

4.1 Morphological, Statistical, and Texture Descriptors (GLCM, LBP, Filter Banks)

The analysis of morphological and textural features of images in microstructural and surface characterization is essential for microstructural classification and quality assessment in materials science. This section focuses on classical texture representation methods, including the Gray Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), and filter-based methods (Laws and Gabor filters).

The Gray Level Co-occurrence Matrix (GLCM) method, developed by Haralick et al. (1973), is based on statistical measures derived from the spatial relationship of gray levels in an image [37]. This approach defines metrics such as energy, contrast, homogeneity, and entropy by calculating the joint occurrence probabilities of pixels at specific directions and distances [37]. These features are particularly useful for distinguishing the texture pattern and orientation of microstructures. GLCM representation is frequently preferred in microstructural images to assess the continuity of grain or phase textures across different scales.

Local Binary Patterns (LBP), developed by Ojala and colleagues (2002), capture local texture patterns by encoding the neighborhood structure around each pixel in a gray-scale image as a binary pattern [38]. This method is widely used in multi-resolution microstructure analysis due to its robustness against scale and rotation changes [38]. The LBP method performs exceptionally well in distinguishing surface roughness and grain morphology in SEM or optical microscope images.

The filter bank approach proposed by Laws (1980) relies on applying different kernel filters (e.g., edge, point, wavelet) to capture energy distribution at specific spatial frequencies [39]. These filters enhance texture discrimination by capturing roughness and directional differences on microstructure surfaces. Gabor filters are direction-sensitive filters that carry information in both the spatial and frequency domains. Clausi and Jernigan (2000) demonstrated that tissue discrimination power can be optimized by correctly selecting the parameters of Gabor filters [40]. Bianconi and Fernández (2007) systematically investigated the effects of changes in Gabor parameters (frequency, direction, scale) on classification success and determined the most suitable parameter ranges [41]. These filter-based approaches yield effective results in capturing the directional anisotropy of microstructures and inter-grain boundary contrasts. To demonstrate the influence of preprocessing, selected filters were applied to the raw image to enhance its clarity and definition before segmentation. As illustrated in Fig. 6, the upper-right section presents the hydrate content estimated using GeoDict's PoroDict module after arithmetic correction, while empty circles indicate regions where poor image quality prevented successful segmentation [42]. Zhu et al. (2022) integrated classical feature extraction methods, such as GLCM and LBP, with ML algorithms to classify microstructure in hot-formed ultra-high-strength steels. In their work, they demonstrated that these statistical texture features can serve as powerful explanatory variables for the training of AI-supported models [43].

4.2 Segmentation, Object Detection, and Quantification Algorithms

In microstructure and surface characterization, segmentation is a fundamental step for determining the boundaries of different phases, grains, or defects and for quantitatively analyzing their geometric properties. This stage is critical for automating microstructural analysis and ensuring the accuracy of object-based statistical evaluation. Figure 7 presents segmentation results for Q&P steel, with the performance metrics—pixel accuracy (PA) and mean intersection over union (MIoU)—illustrated alongside representative micrographs. In the figure, secondary electron (SE) images of two selected test regions are compared with their corresponding EBSD reference images and segmentation results, demonstrating how the model accurately captures grain boundaries and phase distinctions across complex microstructural regions. This comparison highlights the efficiency and reliability of segmentation techniques in achieving consistent quantitative characterization across heterogeneous materials [44].

Otsu (1979) developed a thresholding method that minimizes intra-class variance based on the gray-level histogram. This method provides basic binary segmentation by calculating an optimal threshold value to separate the image into two distinct regions—for example, grains and matrix [45]. In microstructure images, particularly those obtained by optical microscopy or SEM, the Otsu method is a practical starting point for detecting low-contrast features such as phase separation or porosity.

Vincent and Soille (1991) adapted the watershed method to the digital domain with an "immersion" algorithm that simulates the coalescence of water droplets on elevation maps. The watershed approach treats an image as a topological surface and enables region-based boundary determination through gradient information [46]. This method allows accurate boundary separation in microstructures where grains are in contact. Campbell and colleagues (2018) adapted the watershed algorithm to material microstructures, developing a new digitization approach for automatic grain identification and measurement of grain-size distribution (Fig. 8). This method reduced the need for manual intervention and provided high accuracy, particularly in the quantitative analysis of polycrystalline structures [47].

The "edge-free active contours" approach developed by Chan and Vese (2001) separates regions within an image based on region statistics rather than gradients. This model enables the smooth detection of phase boundaries even in microstructure regions where edge contrast is weak [48]. Active contour-based methods exhibit high stability, particularly in metallographic or composite material images where phase transitions are ambiguous.

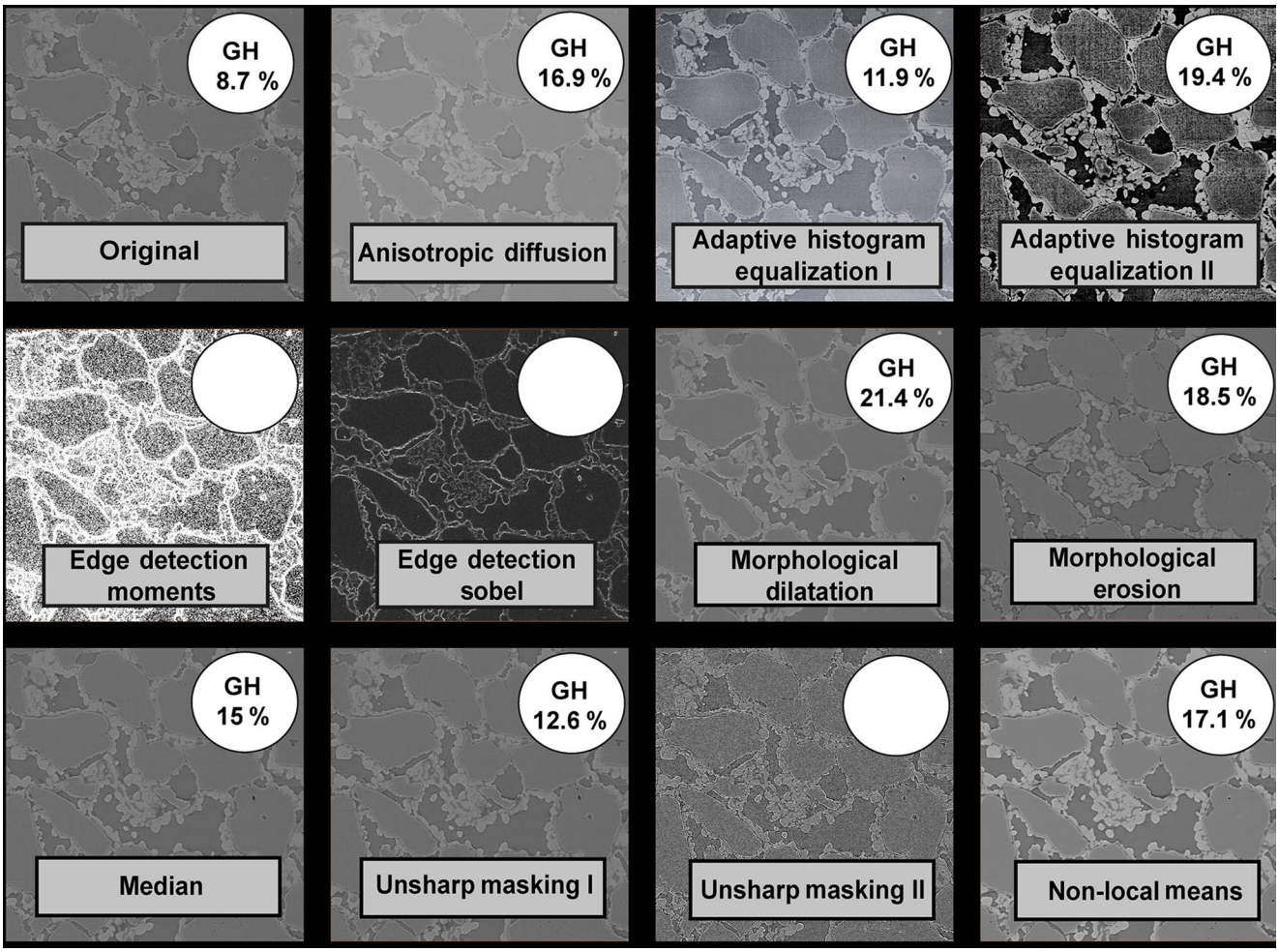


Fig. 6 Selected preprocessing filters were applied to the raw image to improve quality for subsequent segmentation. The upper-right panel shows the numerically estimated hydrate content from GeoDict's PoroDict (after arithmetic correction). Empty circles mark regions of insufficient image quality where segmentation could not be performed [42]

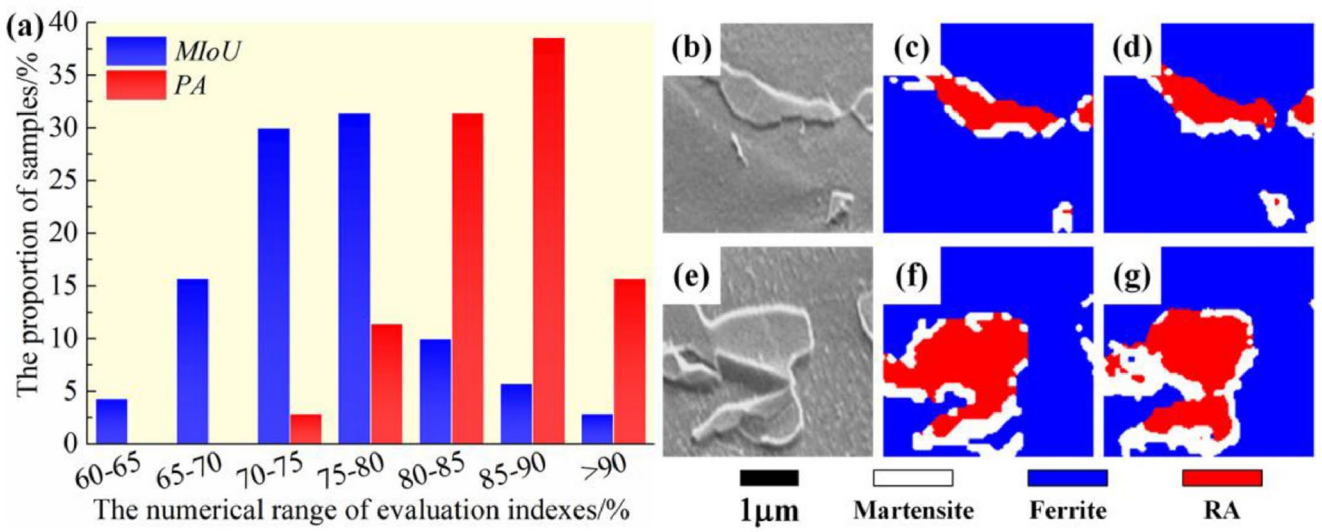


Fig. 7 Segmentation outcomes for Q&P steel: (**a**) distribution of evaluation metrics PA and MIoU; (**b**, **e**) secondary electron (SE) micrographs from two selected test regions; (**c**, **f**) corresponding EBSD reference images; and (**d**, **g**) their respective segmentation maps [44]

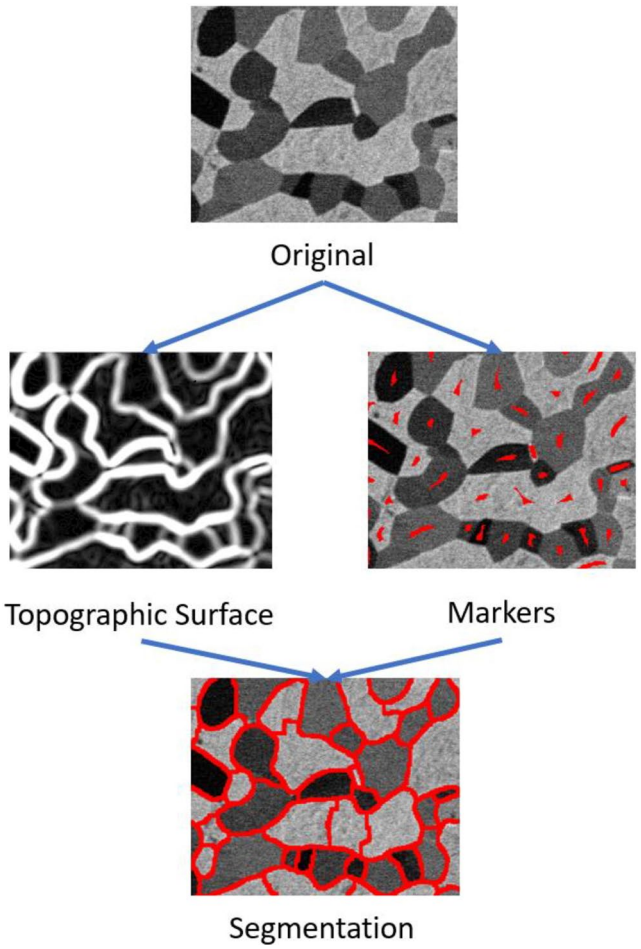


Fig. 8 Visualization of the Watershed Transform [47]

Peregrina-Barreto et al. (Measurement, 2013) propose an automated image-processing method for grain-size determination. The approach uses the gray-level histogram to set a single threshold and divide the image into two classes, thereby separating the grains from the grain-boundary/background areas. As an example, Fig. 9 (a, c) shows the original micrograph and its corresponding histogram; (b) shows the binary output obtained by applying the selected global threshold value (T = 126). This visual representation highlights how histogram-based thresholding enables fast and reproducible grain/background separation [49].

Fotos et al. (2023) improved microstructure quantification by combining the classical watershed algorithm with a DL-supported framework. This approach enables more accurate separation of grains and phases by utilizing probability maps generated by a pre-trained deep network [50]. It reduces the over-segmentation issue in the traditional watershed algorithm, enabling automatic counting of both grains and pores.

Kang et al. (2023) developed a U-Net-based segmentation model for the automatic identification of ferrite–pearlite microstructures and grain size grading. This work achieved high-accuracy grain counting by combining texture maps generated with classical image processing methods with deep neural networks [51]. Such hybrid approaches minimize human intervention while ensuring the reliable acquisition of microstructure statistics from large databases.

4.3 Traditional ML Methods for Feature Extraction and Microstructure Classification

In image analysis studies of microstructure and surface characterization, feature selection is one of the most critical steps to achieving accurate classification. Feature selection aims to improve the model's accuracy and generalization by identifying the most informative variables from the high-dimensional data set extracted from the image [52].

Peng and colleagues (2005) developed the mRMR (minimum Redundancy Maximum Relevance) method, which combines the principles of maximum dependency, maximum relevance, and minimum redundancy within the framework of reciprocal information theory. This approach eliminates unnecessary variables and reduces model complexity by selecting the most meaningful features from among the numerous textural and morphological features obtained from microstructure images [52]. The RELIEF algorithm, developed by Kononenko (1994), and its extended version, ReliefF, evaluate each feature's contribution to class separation based on distances between instances [53]. This method is effective in identifying discriminating features, which is particularly useful for distinguishing similar phases in microstructure classification [53]. Principal Component Analysis (PCA), systematized by Jolliffe (2002), transforms data into a lower-dimensional space by eliminating correlations between features. PCA is frequently used in microstructure analysis to combine multiple parameters based on texture, color, or form to obtain a more compact and explanatory feature set [54].

Support Vector Machines (SVMs), introduced by Cortes and Vapnik (1995), optimize linear or non-linear decision boundaries using kernel functions to maximize the margin between classes. This method is widely preferred in microstructure classification because it exhibits high generalization performance even in high-dimensional feature spaces [55]. The Random Forests algorithm, developed by Breiman (2001), is an ensemble-based approach that improves prediction performance by combining multiple decision trees. This method achieves high discrimination accuracy across various classes, such as phase, grain, or defect type, in complex microstructure data [56].

Gola et al. (2019) used an SVM model to objectively classify the microstructures of low-carbon steels based on morphological and textural features. In their work, the

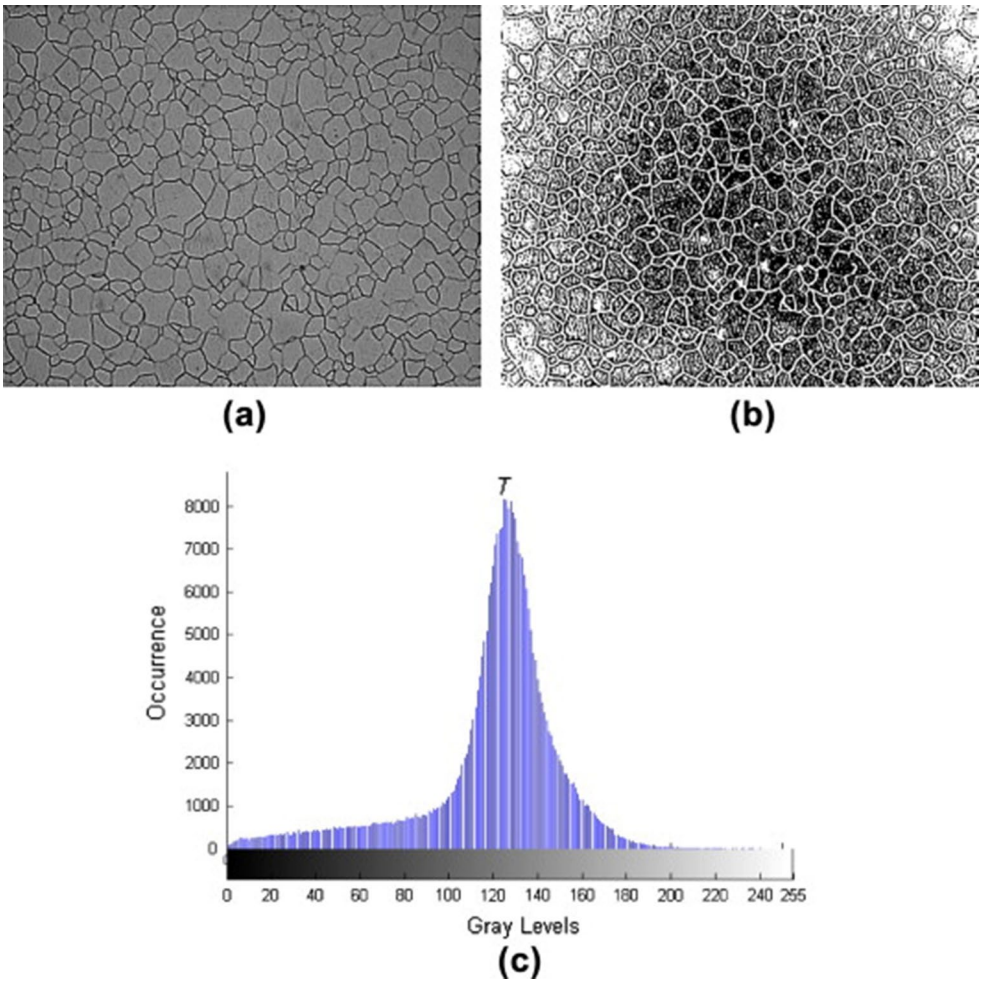


Fig. 9 (**a**, **c**) Original image with its histogram; (**b**) binary result using threshold T = 126 [49]

combination of GLCM-based texture parameters and geometric morphology features resulted in high accuracy in microstructural phase separation [57]. Similarly, Müller et al. (2021) applied traditional ML techniques to classify bainitic subclasses in low-carbon multiphase steels and emphasized the effect of feature selection on model accuracy [58]. Zhu et al. (2022) noted that classical textural features, such as GLCM, offer a more interpretable feature space than convolutional neural networks (CNN)-based approaches when used with SVM classifiers. This demonstrates that traditional methods still have a significant advantage in terms of interpretability [43].

Section Conclusions Classical image processing and feature engineering methods continue to offer explainable, reliable, and data-efficient solutions for microstructure and surface characterization. GLCM, LBP, and filter-based texture analyses provide strong discrimination capabilities by capturing the morphological and orientational features of microstructures. Segmentation techniques such as Otsu, watershed, and active contour algorithms enable the accurate determination of phase boundaries, grains, and defects. Feature selection methods (mRMR, RELIEF, PCA) and traditional ML models (SVM, Random Forests) help optimize high-dimensional image data, improving classification accuracy and enabling generalizable results even on small datasets. While these classical approaches provide interpretability and efficiency, recent developments have increasingly shifted focus toward data-driven methods. The following section, therefore, presents deep learning frameworks that further enhance microstructural analysis.

5 Deep and Hybrid Learning Frameworks for Microstructural Analysis

This section outlines the core AI architectures employed in material science. To provide a conceptual framework, these methodologies are categorized into a functional taxonomy: (i) Classification for phase and group identification, (ii) Segmentation for pixel-level feature extraction, (iii) Detection for surface defect monitoring, and (iv) Reconstruction for 3D microstructural synthesis.

5.1 Core Deep Vision Architectures (CNN, Vision Transformers, Autoencoders) and XAI

Building on classical approaches, this section presents deep learning-based frameworks that have recently become central to microstructural analysis. DL-based image processing approaches have undergone a significant transformation, with core architectures developed to deliver high accuracy and automation for microstructure and surface characterization. Although these architectures are often grouped under the broader term of deep learning, they differ considerably in their internal mechanisms, data requirements, and suitability for specific microstructural characterization tasks. CNNs extract hierarchical local features through convolutional filters and are therefore effective for recognizing texture, edges, phase morphology, pores, and defect patterns in microscopy images. Encoder–decoder architectures such as U-Net and DeepLab extend CNN-based feature extraction to pixel-level segmentation by combining low-level spatial detail with high-level semantic information, making them particularly suitable for grain-boundary, phase, and pore segmentation. In contrast, Vision Transformers divide images into patches and use self-attention mechanisms to model long-range dependencies, which can be advantageous when global spatial relationships or distributed defect patterns are important. Object-detection models such as Faster R-CNN and YOLO are designed to localize and classify discrete objects or defects, with YOLO-type architectures being especially useful for real-time industrial inspection due to their high inference speed. Autoencoders and generative models, including GAN-based frameworks, are mainly used for representation learning, reconstruction, super-resolution, data augmentation, and synthetic microstructure generation. Therefore, the selection of an architecture should depend on the target task: classification, segmentation, detection, reconstruction, or real-time monitoring.

A representative example of this workflow is illustrated in Fig. 10, which conceptually outlines how CNN encoders are initially trained on the MicroNet dataset and subsequently adapted for segmentation through transfer learning (TL). This pre-training and fine-tuning strategy enhances feature-extraction efficiency, enabling the network to generalize effectively across diverse microstructural patterns while reducing the need for large task-specific datasets. Consequently, TL has become a cornerstone in developing advanced segmentation models for materials science applications, bridging the gap between limited experimental data and high-performance deep learning systems [12].

AlexNet, developed by Krizhevsky and colleagues, marked a turning point in the proliferation of deep CNN and demonstrated near-human performance in large-scale image classification tasks [59]. The ResNet architecture introduced the concept of "residual learning" to overcome learning difficulties in deep networks, thereby enabling the stable training of multi-layer networks [60]. In subsequent years, the Vision Transformer (ViT) model offered an alternative paradigm to CNN by using self-attention to partition images into small patches and processing them sequentially.

In image segmentation tasks, the U-Net architecture (Fig. 11) has stood out with its encoder–decoder structure and has become the standard, particularly for pixel-level classification of microstructure maps [61, 62]. The DeepLab family has increased boundary sensitivity through "atrous convolution" and conditional random fields (CRF), enabling more

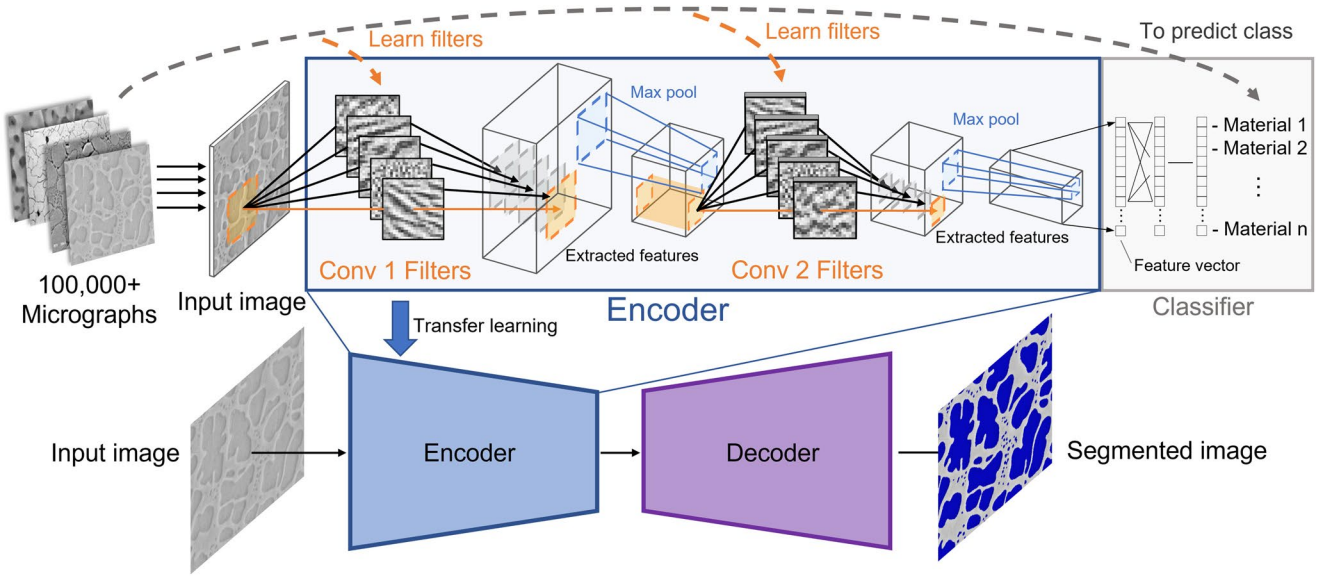


Fig. 10 Conceptual illustration depicting the pre-training of convolutional neural network (CNN) encoders on the MicroNet dataset, followed by their incorporation into a segmentation architecture using transfer learning principles [12]

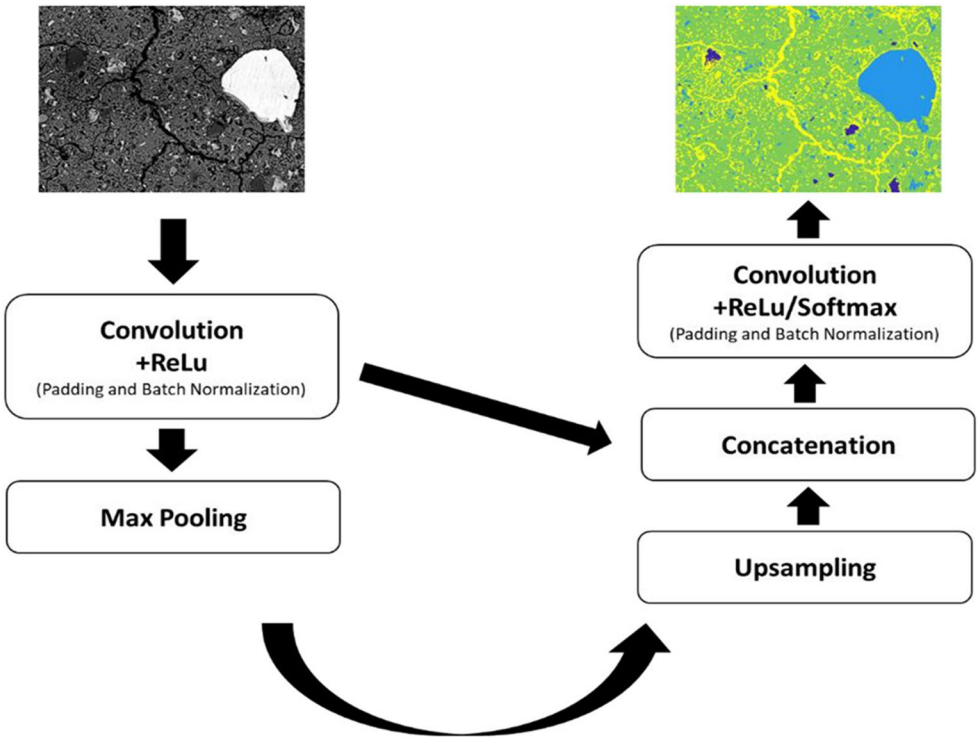


Fig. 11 Basic U-Net architecture with key layers and operations in the encoder (contracting) and decoder (expansive) branches [62]

effective separation of complex phase boundaries [63]. In object detection, Faster R-CNN and YOLOv3 have been widely used to quickly and accurately identify particles, pores, or phase regions in microstructures.

Generative models have come to the fore in applications such as data augmentation and microstructure synthesis. The GAN architecture has enabled the generation of realistic microstructure images [64]. SRGAN, a derivative of GAN, has enabled more precise analysis of surface details by converting low-resolution microscope data into high-resolution data [65]. Among image translation-based approaches, Pix2Pix has enabled conversion from one microstructure modality (e.g., SEM) to another (e.g., optical microscope) [66]. XAI methods, developed to increase the interpretability of deep models, have played an important role in visualizing their decision-making processes. Grad-CAM visually explains which regions of an image a network attends to by using activation maps from convolutional layers [67]. The combined use of these core DL architectures and XAI methods has not only produced highly accurate results in microstructure and surface characterization but also made the reasons behind these results explainable, paving the way for reliable AI-based analysis in materials engineering research. In addition to GANs, VAEs, SRGAN, and Pix2Pix, diffusion-based generative models have recently attracted attention for microstructure reconstruction and predictive image generation. Denoising diffusion probabilistic models and conditional diffusion frameworks can progressively generate realistic microstructure images from noise while preserving statistical and morphological descriptors, making them promising tools for synthetic data generation, microstructure reconstruction, and inverse materials design.

5.2 Deep Learning Applications in Material Microstructure and Surface Characterization

DL-based image processing approaches have gained increasing importance in materials science for microstructure and surface analysis. These methods provide high accuracy, particularly in microstructure segmentation, phase classification, super-resolution (SR), and defect detection. One of the most significant contributions of DL in microstructure characterization is the super-resolution reconstruction of electron backscatter diffraction (EBSD) data. Jung and colleagues demonstrated the enhancement of EBSD microstructure maps with super-resolution (SR) and the contribution of this improvement to the analysis of material mechanical properties [13]. Similarly, Jangid and his team increased the interpretability of EBSD data by reconstructing orientation maps at higher resolution using a physics-based single-image super-resolution (SISR) approach [68]. These studies demonstrate that DL effectively improves the quality of microstructural data.

Durmaz and his team achieved high-accuracy phase identification in complex-phase steels using a U-Net-based approach for microstructure extraction [69]. Furthermore, Fotos and colleagues successfully separated grain boundaries and phases by combining boundary-class semantic

segmentation and Watershed algorithms in Ti-6Al-4 V alloy [50]. Warren and his team also effectively detected grains and grain boundaries using a CNN trained on multiple datasets, demonstrating the model's generalization [70].

Some studies have focused on material synthesis or on generating virtual samples from microstructure data. Huang and colleagues developed a GAN-based model, 3DPm-mGAN, that enables the three-dimensional reproduction of porous medium microstructures [71]. This approach offers significant potential for both data generation in microstructure engineering and virtual experiment design.

CNN- and ViT-based approaches have also emerged for microstructure classification. Kaufmann and colleagues classified EBSD patterns according to space groups using few-shot TL [72]. Azimi and his team achieved high performance in phase separation by classifying steel microstructures at the pixel level using FCNN [11]. In recent years, the work of Bárcena and Liu has also demonstrated the effectiveness of CNN and ViT models for microstructure identification in cast iron and ultra-high-carbon steels [73]. Beyond purely data-driven approaches, Physics-Informed Deep Learning (PIDL) has emerged as a promising direction for integrating domain-specific constraints into neural network architectures. In microstructure analysis, physics-based priors—such as crystallographic symmetry, thermodynamic stability, and phase-field governing principles—can be incorporated into loss functions or model formulations. For example, PIDL frameworks may penalize physically inconsistent predictions, such as violations of volume-fraction conservation or interfacial energy constraints. By constraining the solution space with physically meaningful priors, these approaches have the potential to improve model generalization, particularly in scenarios with limited experimental data, while promoting physically plausible microstructure representations.

For surface characterization, DL models are used for both defect detection and surface quality estimation. Vasan and colleagues detected surface defects in hot-rolled steel strips with high accuracy using a ViT-based system [74]. Gao and his team developed the YOLOv5 architecture with BiFPN and attention mechanisms to identify steel surface defects in real time [75]. Zhang and his team also developed online defect monitoring applications for production lines using a derivative model based on YOLOv8 [76]. A key challenge in surface defect detection is the inherent class imbalance, where minority defect types (e.g., micro-cracks) are significantly underrepresented. To address this issue, many studies employ strategies such as cost-sensitive learning, data augmentation, and loss function modifications (e.g., Focal Loss) to improve the detection of hard-to-classify samples. These approaches aim to mitigate bias toward dominant classes and enhance model robustness in industrial inspection scenarios. Furthermore, the trade-off between precision and recall is a critical consideration in industrial defect detection systems. While high recall reduces the risk of missing critical defects, it may increase false positives. Therefore, selecting an appropriate operating point is application-dependent. In recent studies, performance evaluation increasingly incorporates Precision–Recall (PR) curve analysis and F1-score as complementary metrics to accuracy, particularly in scenarios where the cost of false negatives (missed defects) is significantly higher than that of false positives.

Section Conclusions DL-based methods have established a new paradigm in microstructure and surface characterization. Advanced architectures such as CNN, U-Net, ViTs, and GAN achieves high accuracy in tasks including segmentation, classification, super-resolution, defect detection, and data synthesis. These approaches improve data processing efficiency and also enhance transparency in decision-making through explainable AI methods. As a result, DL models enable reliable, generalizable analysis of both microstructure characterization and surface quality assessment. These methodological advances have led to a wide range of practical applications. The following section presents how these approaches are applied in microstructure and surface characterization.

6 Microstructure and Surface Characterization

6.1 Methods and Case Studies

Based on the discussed methodologies, this section focuses on their practical applications in microstructure and surface characterization. Microstructure characterization is a fundamental research area in materials science, aimed at quantitatively analyzing structural properties, phase ratios, grain boundaries, and surface morphology. In recent years, AI- and DL-supported image processing methods have revolutionized this field by offering faster, more accurate, and more repeatable results than classical micrograph analysis.

Azimi and colleagues accelerated microstructure identification processes by automating phase separation in steel microstructures using a fully connected convolutional neural network (FCNN)-based pixel-level classification method [11]. Similarly, Durmaz and his team extracted microstructures across multiple imaging regimes in complex-phase steels using the U-Net architecture, achieving high accuracy in phase segmentation [69]. Fotos and colleagues presented a hybrid approach combining the watershed algorithm with boundary-class semantic segmentation in Ti-6Al-4

V alloy, thereby enabling more accurate determination of phase boundaries [50]. Furthermore, Warren and colleagues demonstrated high generalization performance in grain and grain boundary segmentation using CNN trained on real and synthetic datasets [70]. Ostormujof and his team developed a CNN-based method for EBSD phase segmentation in dual-phase steels, enabling automated microstructure classification [77]. Kaufmann and colleagues used a few-shot TL method to classify EBSD patterns by space group, thereby establishing a model capable of distinguishing between crystal structures [72]. To reduce reliance on large-scale labeled datasets, recent research in microstructural characterization has increasingly explored self-supervised learning (SSL) and contrastive learning frameworks. Approaches such as SimCLR and MoCo aim to learn meaningful feature representations by maximizing agreement across different augmented views of the same sample.

In materials science, SSL-based approaches offer a promising way to extract textural and morphological features from unlabeled microscopy data, which can then be fine-tuned with a limited set of annotated samples. This is particularly relevant in microstructure analysis, where pixel-level annotation is labor-intensive and subject to inter-observer variability.

Although the application of SSL in this domain is still emerging, these approaches have the potential to improve generalization and reduce annotation costs, especially in scenarios involving rare phases and multi-scale structural features. Patrick and his team optimized the data processing pipeline and post-processing steps by detecting grain boundaries in bright-field TEM images using the U-Net architecture [78]. In another study by Bals and Epple, the size and shape distribution of nanoparticles in SEM images were automatically measured using a CNN-based approach [79].

GAN-based approaches are attracting attention in microstructure reconstruction. Huang and colleagues produced porous medium microstructures in three dimensions using the 3DPmmGAN model, synthesizing new microstructures while preserving statistical properties [71]. Supporting this work, Mosser and his team also demonstrated that 3D microstructure reconstruction using GAN can produce statistically accurate samples [80]. Amiri and colleagues developed a GAN model capable of performing 3D reconstruction from 2D microstructure images, creating a fast and reliable workflow [81]. Furthermore, Nguyen and his team synthesized microstructures with target physical properties using an active learning-supported GAN model [82]. Wanni and colleagues created representative texture datasets (TACS) from EBSD data, improving the texture-based characterization performance of DL models [83]. Sharma and his team demonstrated that twin activity in MgAZ31 alloy can be predicted using ML [84]. These two studies show that microstructural deformation mechanisms can be understood using data-driven methods.

Additionally, classical approaches based on image processing have also made significant contributions. Ercetin et al. determined the phase ratio and grain size in Mg-Al-Sn alloy microstructures using image processing and verified these results using manual methods (Fig. 12) [85]. Furthermore, the automatic grain counting algorithm developed by Akkoyun has been an essential step in determining microstructural parameters in Mg alloys [86]. Ercetin et al. emphasized the importance of image processing methods for evaluating machined surface and tool condition, and also highlighted their potential for microstructure characterization [87]. This approach shows that the relationship between microstructure and surface quality in production processes can be evaluated more comprehensively.

In general, these studies show that DL, GAN, CNN, and image-processing-based systems offer advantages in accuracy, automation, and speed for microstructure characterization. The methods are no longer limited to laboratory measurements but have advanced to the point where they can be used in industrial applications.

6.2 Surface Defect Detection and Classification

Surface characterization plays a critical role in determining material performance and production quality. AI-supported image processing techniques are rapidly replacing classical methods, especially in defect detection and similar areas. DL-based models enable the automatic, accurate, and repeatable evaluation of surface properties.

In recent years, object detection algorithms powered by DL models have been effectively used to identify defects on steel and other metal surfaces. Vasan and colleagues developed a model based on the ViT architecture and accurately detected and classified defects on hot-rolled steel surfaces [74]. Similarly, Wang and his team successfully detected defects on steel strip surfaces in real time using a lightweight model based on Swin-Transformer–YOLOv5 [88]. These studies demonstrate that transformer-based architectures generalize better than classical CNN in industrial quality control processes.

Various YOLO-based variants have also achieved significant success in identifying surface defects. Gao and colleagues classified steel surface defects using the YOLOv5-KBS algorithm and achieved more accurate feature extraction through attention mechanisms in the BiFPN [75]. Zhang and colleagues detected surface defects on hot-rolled steel at high speed using an improved YOLOv5 model, thereby enhancing the model's overall performance [89]. Meanwhile, Song and his team successfully detected

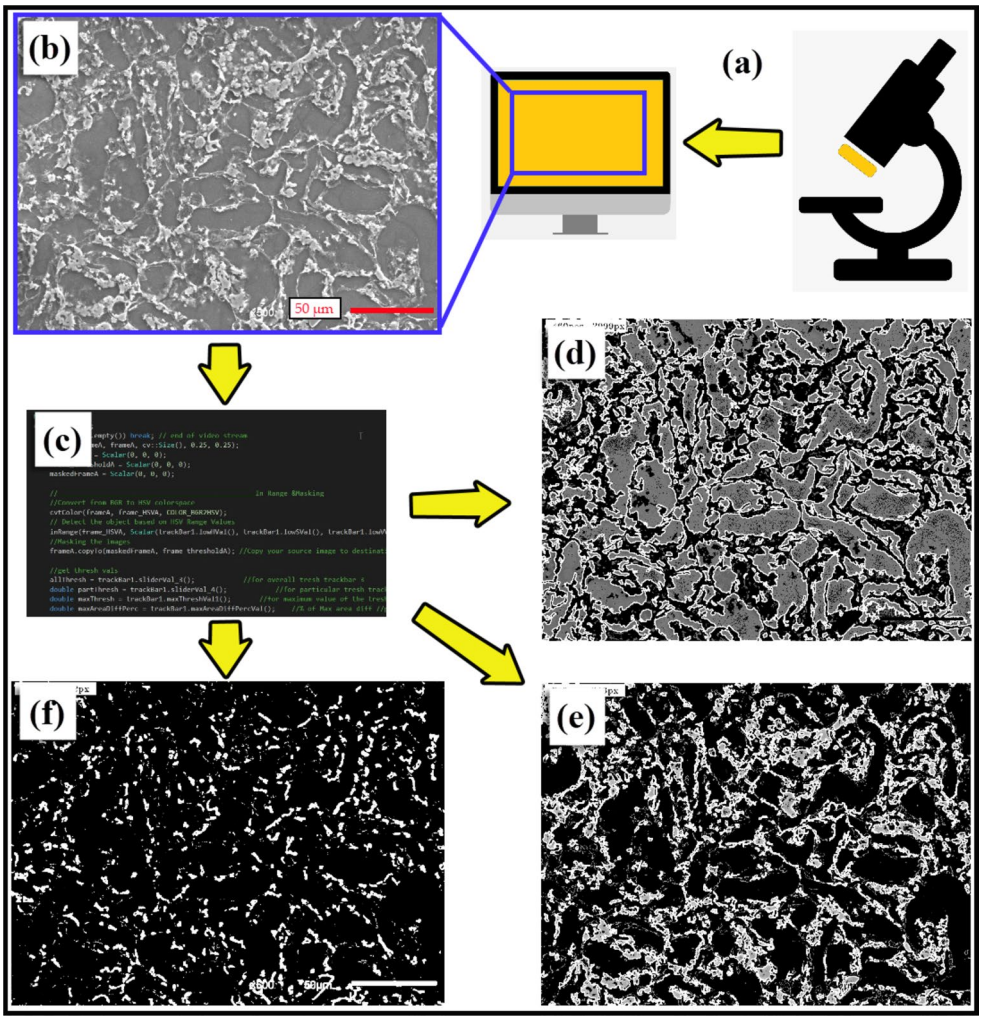


Fig. 12 (**a**) SEM imaging step; (**b**) raw SEM micrograph and its processed outputs; (**c**) software code developed for analysis; (**d**) detected grain boundaries; (**e**) $Mg_{17}Al_{12}$ phase zones; (**f**) Mg_2Sn phase zones extracted through segmentation [85]

both minor and complex defects by proposing an improved version of the YOLOv8 architecture [90]. Similarly, Zhang and colleagues were able to monitor surface defects in steel products online using an algorithm named DDI-YOLOv8n [76].

Defect detection is not limited to steel products. Chen and his team developed a system that detects surface defects in industrial components using DL-supported image processing [91]. Chen and colleagues also demonstrated that real-time quality control is possible in production processes, using a model developed to identify surface defects in steel strips [92]. Furthermore, Das and colleagues used DL to detect corrosion in steel structures and found that this method significantly reduced errors compared to human observation [93]. Arı and colleagues developed a system that detects gear defects using a machine vision approach in gear systems and demonstrated the applicability of AI-based imaging techniques in fault diagnosis on production lines [94]. These examples demonstrate that surface defects can be effectively identified not only in metallurgical components but also in mechanical system components (Fig. 13).

Section Conclusions AI- and DL-based methods have significantly advanced microstructure characterization and surface defect detection. Architectures such as CNNs, U-Net, GANs, and Transformer-based models have enabled automated analysis for tasks including phase identification, grain boundary segmentation, microstructure reconstruction, and defect detection. These developments have reduced reliance on manual interpretation and facilitated data-driven workflows increasingly applicable in both academic research and industrial quality control. However, reported model performance must be interpreted in the context of the specific task and evaluation protocol. As summarized in Table 1, different problem categories rely on distinct performance metrics, including accuracy (classification), IoU and F1-score (segmentation), and mAP (object detection). This distinction is essential to avoid misleading cross-task comparisons. In terms of validation, the reviewed studies present a heterogeneous landscape. While some works employ rigorous strategies such as k-fold cross-validation and stratified sampling [69, 73, 84], a substantial portion of recent deep learning studies, particularly in segmentation and detection, relies

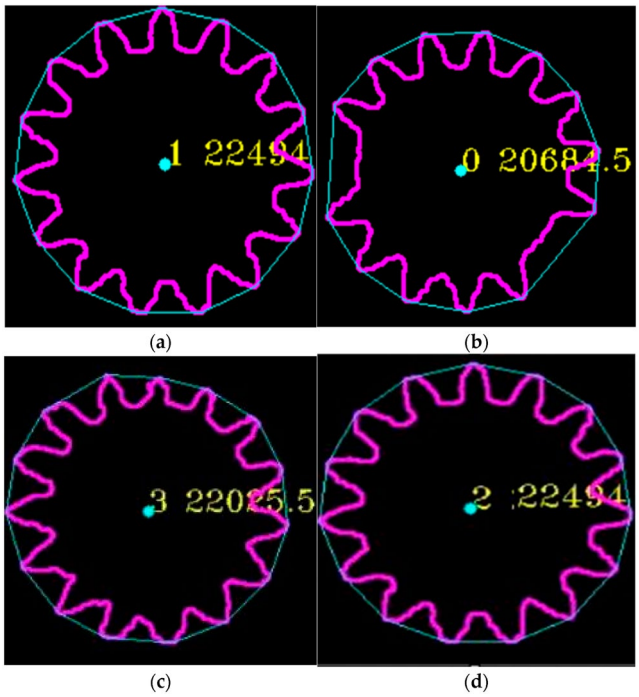


Fig. 13 Processing output illustrates various gear conditions: (**a**) a perfectly manufactured gear; (**b**) a defective gear missing one tooth; (**c**) a gear showing an error in the central hole diameter; and (**d**) a gear with visible surface defects resulting from improper layer deposition during the printing process [94]

on hold-out validation with explicit train/validation/test splits, often supported by data augmentation and early stopping to control overfitting. Moreover, although metrics such as accuracy, IoU, and F1-score are widely reported, only a limited number of studies provide measures of statistical variability (e.g., standard deviation across folds), and formal significance testing remains relatively uncommon in the field. The following Table 1 provides a comprehensive and structured benchmark of the reviewed studies, categorizing them by task, validation strategy, and reported performance metrics.

Despite the rise of deep learning, a systematic comparison reveals that classical ML approaches still hold significant advantages in specific industrial contexts. While deep learning models typically demand large-scale labeled datasets, such as the 4000 images utilized in [93], classical methods like LBP and GLCM are far more efficient regarding data requirements, functioning effectively with minimal samples. In terms of interpretability, these traditional algorithms offer 'glass-box' transparency, allowing for physical verification of spatial texture correlations, whereas DL models often remain 'black-boxes.' Furthermore, the computational cost of classical methods is remarkably lower; for instance, the rule-based approach in [94] achieves a processing time of only 42 ms on standard hardware. This contrast highlights a trade-off: while deep learning provides superior accuracy for complex patterns, classical ML offers a more interpretable and resource-efficient solution for straightforward industrial tasks. A comprehensive summary of these diverse methodologies is presented in Table 1.

The choice of AI architecture in materials science is fundamentally driven by the specific characterization task, as each framework presents distinct strengths and trade-offs. CNN-based models, such as U-Net, remain the gold standard for high-precision segmentation [12, 77] but are constrained by intensive pixel-wise labeling requirements. In contrast, Transformer-based architectures [74, 88] offer superior global feature extraction for defect classification, albeit at a higher computational complexity. For 3D reconstruction from 2D data, GAN-based frameworks [80–82] provide unparalleled structural fidelity (e.g., $< 5\%$ error in [82]), though they face training stability challenges compared to supervised methods. Lastly, real-time industrial detection necessitates lightweight models like YOLOv5s [70], which balance high inference speeds (70 FPS) with robust accuracy (79.9% mAP). A systematic comparison of these architectures, regarding their data requirements, computational costs, and suitability for specific microstructural tasks, is synthesized in Table 1.

The systematic categorization in Table 1 shows that the adoption of deep learning methods is accompanied by diverse validation practices rather than a single standardized protocol. Although high performance is frequently reported across tasks, the reliability and generalizability of these results depend strongly on the chosen validation strategy and dataset characteristics. In particular, studies using cross-validation tend to yield more robust estimates of model performance, whereas those relying on hold-out validation may be more sensitive to dataset size and distribution. This variation underscores the need for more consistent evaluation frameworks and improved statistical reporting in future research. Table 1 shows that recent advances in AI-driven microstructure analysis are driven not only by architectural improvements but also by evolving validation methodologies and performance reporting practices.

Although these advanced methods have shown remarkable success in handling data with high noise, uncertainty, and complex class distributions, they have not yet been systematically applied to the challenging field of microstructure characterization and classification in materials science. Microstructural images inherently contain high uncertainty due to unclear grain boundaries, phase inhomogeneities, and various artifacts. For example, accurately identifying the start and end of a grain boundary or classifying a phase can often be unclear. In such cases, a classifier based on FPFS-kNN, AIFPIFSC1, or AIFPIFSC2 can treat transition regions between different microstructural features (like

Table 1 Comparative analysis of AI-driven methodologies in microstructure and surface characterization: Materials, methods, validation strategies, and performance metrics

Category	Ref	Material / Application	AI Method	Validation Strategy	Key Findings	Performance Metrics
Classification	[11]	Low Carbon Steel (Martensite, Pearlite, Bainite, Ferrite)	FCNN + Max-Voting	Hold-out (Separate Training, Validation, and Test sets; validated on unseen samples)	Demonstrated high-accuracy automated classification of complex steel phases without manual feature extraction.	93.94% Overall Accuracy
	[72]	EBSD Patterns (Space Group Classification)	Few-shot TL (InceptionResNetV2 based)	Benchmark against Train-from-scratch (Using varying training set sizes: 5 to 500 images)	TL significantly outperforms training from scratch, especially when labeled data is extremely scarce (few-shot).	Precision: 0.97, Recall: 0.96 (TL) vs. 0.93 / 0.92 (from scratch)
	[73]	Cast Iron Alloys (Optical Microscopy Images)	MobileNet (Classification) + U-Net (Segmentation)	5-Fold CV (for Segmentation) & Hold-out (for Classification)	Synthetic dataset generation + CNN enables highly accurate and objective graphite morphology classification, outperforming manual analysis	98.9% ± 0.4 Accuracy (Pure images); ~84–93% (Complex mixed morphology)
	[74]	Hot-rolled steel surface defects (NEU-DET dataset, 6 classes)	ViT	Hold-out Validation, Train/test split (80/20) with validation-based model selection	Hyperparameter-optimized ViT achieves robust defect classification and outperforms conventional ML and CNN approaches	Accuracy: 96.39% (optimized); 75–93.89% across hyperparameter variations
	[84]	Mg AZ31 Alloy	Decision Tree (REPTree + Bagging, ML)	10-fold CV	Local microstructural features (GB misorientation, Schmid factors) successfully predict twin activity; enables "intelligent microscopy"	Accuracy: ~85% (5 × 5 neighborhood), AUC ≈ 0.79
Segmentation	[12]	General Microscopy Images (MicroNet; Nisuperalloy & EBC segmentation)	CNN (TL with MicroNet pretrained encoders; UNet / UNet++ / DeepLabV3+)	Train/Validation/Test split with IoU-based evaluation (no cross-validation; multiple dataset splits: Super-1/2/3/4, EBC-1/2/3)	Domain-specific pretraining (MicroNet) significantly improves segmentation accuracy, especially in few-shot and one-shot cases, and enhances generalization to out-of-distribution data	IoU: up to 93.0% (one-shot), but typically ~ 55–65% across architectures; significant improvement over ImageNet baseline
	[50]	Ti6Al4V alloy microstructures (SEM images, grain size & phase analysis)	Deep Learning (BCSS semantic segmentation, ResNet-50) + Watershed (Hybrid HADMA)	Validation against ASTM standards (E112, E562) and comparison with ground-truth measurements	Hybrid approach significantly improves segmentation accuracy and measurement reliability, outperforming conventional watershed methods	BF score: ~97–98%, Precision/Recall: ~97–98%; ~93.9% agreement (grain size), ~ 91.2% (volume fraction)
	[69]	Complex-phase steel microstructures (LOM & SEM, lath-bainite segmentation)	Deep Learning (U-Net, VGG16 U-Net CNN)	5-fold CV (IoU-based evaluation)	DL enables segmentation of complex microstructures with limited data, with performance strongly dependent on imaging modality and resolution	IoU: ~70–72% (LOM), ~ 77–80% (SEM) for foreground; background IoU: ~68–87%
	[77]	Dual-phase steel microstructures (EBSD maps, ferrite–martensite segmentation)	Deep Learning (U-Net CNN; inputs: BS, BC, KAM, quaternions)	Train/validation/test split with pixel-wise evaluation and confusion matrix analysis	DL enables highly accurate automated phase segmentation from EBSD data, outperforming traditional methods and maintaining robustness under noise and complex microstructures	CA: 98.5%; Precision: ~98.5%; Recall: ~98–99%
	[79]	SEM nanoparticle images (SE & STEM)	UNet++ (segmentation), ResNet34 & AlexNet (classification)	Hold-out Validation, Train/validation split with independent test evaluation	Automated particle size and shape extraction from SEM images with enhanced separation of overlapping particles	Accuracy: ~96–99.7%, IoU: ~90–94%, F1: ~94–97% (performance varies by imaging mode and scenario)

Table 1 (continued)

Category	Ref	Material / Application	AI Method	Validation Strategy	Key Findings	Performance Metrics
Detection	[75]	Steel surface defect detection (NEU-DET, 6 classes)	YOLOv5-KBS (YOLOv5 + SE + BiFPN)	Hold-out Validation, Train/test split (90/10)	Combined feature enhancement and fusion improve detection of complex and multi-scale defects	mAP: 0.799 (vs. 0.757 baseline, + 4.2%); FPS: ~70; class-dependent improvements
	[76]	Steel surface defect detection (NEU-DET, 6 classes)	DDI-YOLO (Improved YOLOv8n with C2f_DWR, C2f_DRB, Inner-IoU)	Hold-out Validation Train/validation/test split (8:1:1)	Improved multiscale feature extraction significantly enhances detection of rolled-scale defects while maintaining lightweight structure	mAP: 78.3% (vs. 75.9%, + 2.4%); FPS: 158 (+ 59); class-dependent performance (up to + 10.6%)
	[88]	Hot-rolled steel surface defect detection (NEU dataset, 6 classes)	Swin-Transformer–YOLOv5 (GhostNet + C3STR + CoordAttention + BiFPN)	Hold-out Validation (7:2:1 split; manually annotated via LabelImg)	Lightweight transformer-enhanced YOLO architecture improves multi-scale defect detection and reduces computational complexity	mAP: 74.9% (+ 8.39% vs. YOLOv5s), FPS: 43.2, Params ↓36.6%, GFLOPs ↓40.0%
	[90]	Steel Surface (Defect Detection)	Improved YOLOv8 (DCNv2, BiFPN, BiFormer & WIoUv3)	Hold-out Validation (8:1:1 split; comparative benchmarking)	Optimized YOLOv8 by replacing Bottleneck convolutions with DCNv2 for larger receptive fields and integrating BiFormer attention. Switched PAFPN to BiFPN and CIoU to WIoUv3 to focus on ordinary quality anchors via a dynamic non-monotonic mechanism.	84.8% mAP@0.5; Demonstrated superior accuracy over baseline YOLOv8 (81.4%) and YOLOv7 (80.9%) through advanced feature fusion and attention-based enhancement.
	[93]	Steel Surfaces (Corrosion Segmentation)	Hybrid U-Net (with EfficientNetB7 & DenseNet121 backbones)	Hold-out Validation (8:1:1 split; validated via TL vs. Fully Trained modes)	Proved that EfficientNetB7 as a backbone for U-Net significantly outperforms other architectures in segmenting corroded areas, especially when using pre-trained weights.	95.6% IoU, 97.8% F1-score, and 98.5% Accuracy (with EfficientNetB7 in TL mode)
	[94]	3D-Printed Gears and Sprockets	Image Processing + Masking + HSV (Rule-based)	Repetitive Statistical Evaluation (30 trials under varying lighting/orientation)	Developed a robust, non-ML detection system using HSV masking; demonstrated that rule-based systems can achieve higher speeds and precision for structured industrial objects compared to complex DL models.	98.5% Accuracy; 42 ms processing time; Standard deviation below 1% across repeated trials.

Table 1 (continued)

Category	Ref	Material / Application	AI Method	Validation Strategy	Key Findings	Performance Metrics
Reconstruction	[71]	Porous Media (Sands/Rocks)	Slice-to-Volume GAN (S2VGAN)	Statistical Consistency Validation	Successfully reconstructed 3D volumes from 2D slices by capturing long-range connectivity in porous structures.	Relative Error < 5% (in porosity/surface area)
	[80]	Porous Media (Bead Pack, Sandstone, Limestone)	Deep Convolutional GAN (DCGAN)	Morphological & Statistical Validation (Two-point probability, Euler characteristic, Minkowski functionals)	Demonstrated that a volumetric DCGAN can reconstruct 3D porous microstructures while maintaining high fidelity in spatial correlation and physical characteristics.	High Porosity (ϕ) Accuracy: Bead Pack (0.363 vs. 0.361), Berea (0.196 vs. 0.210); Strong agreement in Minkowski functionals (V_0, V_1, V_2, V_3).
	[81]	Porous Media (Berea Sandstone)	Modified Slice-GAN (3D Generator with three 2D Discriminators)	Cross-modality Validation (Validated against μ CT ground truth; statistical analysis via $S_2$and FID)	Proved that realistic 3D microstructures can be synthesized using only 2D SEM/BSE or optical images, capturing essential pore network characteristics without requiring expensive 3D training data.	High Porosity Correlation: 0.232 (μ CT) vs. 0.226 (BSE-gen); Low FID values (0.60, 0.69, and 0.61 for the x-, y-, and z-directions, respectively,) for BSE images indicating superior structural similarity.
	[82]	Porous Media (Berea Sandstone, Bead Pack)	GAN-AC (GAN + Actor-Critic Reinforcement Learning)	Inverse Design Validation (Comparison of 500 test cases against targeted physical Quantities of Interest)	Developed a reinforcement learning-guided GAN framework that synthesizes 3D microstructures tailored to specific user-defined physical properties, bridging the gap between stochastic generation and controlled design.	Target Matching Accuracy: Mean error of 2.54% for Porosity, 3.51% for Surface Area, and 4.06% for Permeability; High correlation ($R^2 \geq 0.86$) for all properties.
Analytical / Hybrid	[70]	W-Mo-Nb Alloys	Angular Dependent Potential (ADP) + comparison with ML Interatomic Potentials (SNAP, MTP, tabGAP)	Benchmarking against DFT calculations and experimental data	Developed a computationally efficient interatomic potential capable of accurately reproducing thermomechanical properties and defect behavior in CCAs	RMSE ≈ 48 meV/atom (energy), 368 meV/Å (forces) (vs. DFT)
	[84]	EBSD datasets	Unsupervised ML (TACS: K-means clustering + density-based sampling)	Statistical validation via Kolmogorov–Smirnov (K-S) test	Proposed TACS method generates reduced-order datasets that accurately preserve crystallographic texture, outperforming traditional EBSD reduction techniques and enabling 1–2 orders of magnitude data reduction	K–S score ≈ 0.94 (TACS) vs. 0.001–0.23 (conventional methods)
	[85]	Mg–Al–Sn alloy microstructures (grain size & phase ratio analysis from SEM images)	Computer Vision (OpenCV-based image processing) + ML-supported feature detection	Validation via comparison with manual measurements (no cross-validation)	Automated method accurately detects grain size and intermetallic phases, significantly reducing manual effort while maintaining high consistency	> 92% accuracy (grain counting), ~ 99% accuracy (ASTM grain size), > 95–97% phase detection accuracy
	[86]	Mg alloys (SEM-based microstructure grain counting & ASTM grain size)	Computer Vision (OpenCV-based image processing, contour detection)	Validation via comparison with manual measurements (no cross-validation)	Proposed automated method accurately counts grains and determines ASTM grain size, reducing user-dependent errors and achieving consistent results across different alloy compositions	~ 94% accuracy (grain counting), ~ 6% standard deviation vs. manual

ferrite, pearlite, bainite) as "partial members" and incorporate this uncertainty into the decision process, resulting in more realistic and robust classification. Combining these methods with DL models can produce even more reliable results; for example, features extracted by CNN can be processed by such classifiers, enabling the model to learn complex patterns and quantify decision uncertainty. Such a hybrid system could serve as a decision support tool for materials scientists, offering both high accuracy and greater interpretability.

6.3 Uncertainty Modeling in Microstructure Analysis via Fuzzy, Intuitionistic, and Soft Sets-Based Hybrid Approaches

In recent years, the limitations in accuracy, generalization, and interpretability of traditional ML algorithms have driven the development of fuzzy, soft, and intuitionistic extensions. In this context, new-generation algorithms based on these mathematical tools provide an alternative to traditional methods by their ability to model uncertainties.

Integrating traditional methods with fuzzy logic provides a powerful way to overcome the limitations of classical ML algorithms. For example, the Fuzzy kNN algorithm [95, 96] softens the strictness of classical kNN by assigning equal voting rights to each neighbor and introducing a voting mechanism based on each neighbor's class membership. This reduces the influence of uncertain and noisy data points at class boundaries, leading to more robust classification decisions. Similarly, Fuzzy SVM [97, 98] softens the rigid decision boundaries of support vector machines by assigning each data point a distinct membership value based on its significance in the model.

These basic fuzzy approaches were later improved with more sophisticated mathematical structures, such as fuzzy parameterized fuzzy soft matrices (*fpfs*-matrices). For instance, FPFS-kNN achieved higher classification performance than classical Fuzzy kNN by modeling not only the data but also the uncertainty of the parameters through distance-based similarity measures [99]. Going further, classifiers built with intuitionistic fuzzy parameterized intuitionistic fuzzy soft matrices (*ifpifs*-matrices) [100, 101] incorporated hesitation information by considering the degree of membership and non-membership, leading to significant improvements in accuracy and stability. picture fuzzy soft k-nearest neighbor (PFS-kNN) [102], which introduces the concept of a neutrality degree, has been especially effective in fields such as medical diagnosis, where hesitation is common among experts. One of the latest studies, adaptive ML approaches, i.e., AIFPIFSC1 and AIFPIFSC2 [103], have used the concept of *ifpifs*-matrices to enable models to adapt spontaneously to different datasets.

7 Future Research Trends and Challenges

Drawing on the reviewed methods and applications, this section outlines current challenges and potential future research directions. The rapid advancement of AI-supported image-processing methods for microstructure and surface characterization presents new opportunities and significant challenges. This section summarizes future research directions and the main problems that may be encountered.

7.1 Data Diversity and Open Access Databases

One of the most important future research trends is the creation of more comprehensive, multi-modal (optical, SEM, AFM, EBSD, etc.) and well-labeled open databases. Because current datasets are often limited to specific material types, DL models exhibit limited generalization. Standardizing data across laboratories, devices, and sample-preparation conditions will increase the model's reliability and reproducibility. Furthermore, the development of domain-adaptation and transfer-learning-based approaches will enable merging images from different data sources into a common representation.

7.2 Data Deficiency and Labeling Challenges

The labeling process remains one of the biggest bottlenecks in AI-based microstructure analysis. Manual annotation based on expert knowledge is expected to be replaced by semi-automatic labeling, active learning, and self-supervised learning techniques. These approaches will reduce human effort while enabling the model to learn with high accuracy even with limited data. Furthermore, synthetic data generation strategies (e.g., GAN- or VAE-based) are expected to be used more extensively, particularly for representing rare microstructural types.

7.3 Multi-Scale and Time Series Analyses

Current studies are mostly limited to two-dimensional micrographs. However, in the future, the analysis of multi-scale, time-resolved characterization data will enable dynamic monitoring of microstructural evolution processes (e.g., phase transformations, grain growth, deformation). In this regard, the development of DL and image fusion algorithms for three-dimensional (3D) and four-dimensional (4D, i.e., time-dimensional) microstructure reconstruction will be an important research direction.

7.4 Industrial Applications and Real-Time Monitoring

AI-based image systems for real-time quality control in industrial production lines will continue to be developed. In particular, YOLO, ViT, and edge AI applications offer suitable solutions for low-latency and high-accuracy analysis in production environments. These systems are expected to contribute to "smart manufacturing" and Industry 4.0 approaches by working in conjunction with material processing parameters.

Regarding future perspectives, a significant portion of current research remains focused on metallic systems, particularly steels and aluminum alloys. However, the methodological frameworks discussed, such as generative models and advanced attention mechanisms, demonstrate substantial potential for cross-material generalization. Future studies should address the challenges of varying contrast levels and morphological scales inherent in non-metallic material classes, such as ceramics, polymers, and complex composites. By leveraging domain adaptation and transfer learning techniques, the high-performance metrics reported in recent literature, such as the 79.9% mAP achieved in [70], could be replicated across a broader spectrum of material systems, leading to a more universal AI-driven material characterization framework.

7.5 Ethics, Transparency, and Standardization

The reliability of AI-based decision systems depends not only on technical success but also on ethical values and standards. In the future, creating international standards for data privacy, model transparency, and reproducibility will be of great importance. Furthermore, initiatives that encourage the sharing of open-source code among education and research communities will accelerate progress in this field.

8 Conclusions

This review systematically examined the current state of AI-supported image processing methods for microstructure and surface characterization, the range of algorithms used, and their potential applications. The results reveal that DL and computer vision (CV) based approaches go beyond classical micrograph analysis, enabling a revolutionary transformation in materials science research through automation, speed, and accuracy. Modern architectures such as CNN, U-Net, ViT, YOLO, and GAN have achieved high success rates in microstructure classification, phase segmentation, and surface defect detection.

DL-based models have become an effective tool not only for defining structural features but also for understanding physical fundamentals. Approaches such as super-resolution, TL, few-shot learning, and physics-informed DL (PIDL) have significantly reduced limitations, including data scarcity, labeling costs, and generalization deficiencies. However, the scarcity of high-quality, open-access datasets, domain shift issues, and the lack of explainability remain significant obstacles to the field's development.

Image representations and classical feature engineering methods (GLCM, LBP, Gabor filters, etc.) remain essential due to their high reliability, especially when data is scarce or labeling is labor-intensive. Combining these methods with DL models in a hybrid approach offers advantages in both accuracy and explainability. Furthermore, integrating XAI techniques enhances the interpretability of deep models' decision-making processes, enabling results consistent with physical principles.

In surface characterization, DL-supported defect-detection analyses have enabled real-time quality control in production processes. Transformer-based ViT and YOLO derivatives enable high-speed, low-error defect identification. At the same time, PIDL and hybrid neural network models have demonstrated strong performance in correlating surface topography with process parameters. These developments pave the way for new-generation measurement and monitoring approaches that can analyze the microstructure–surface–performance relationship in industrial systems in a multidimensional manner.

Future research should focus on creating more comprehensive, multimodal open databases, developing models suitable for domain adaptation, and designing hybrid artificial intelligence systems with high interpretability. Furthermore, microstructure and surface data are expected to be evaluated not only as analysis tools but also as input sources for data-driven material design and process optimization. In this regard, AI-supported image processing methods are expected to become integral to standard characterization protocols in both academic and industrial applications.

Author contributions The authors contributed equally in each section.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK). This research received no external funding.

Data Availability No datasets were generated or analysed during the current study.

Declarations

AI Declaration In this study, LLMs (DeepL and Grammarly Pro) were only used cautiously for routine tasks such as enhancing language, readability, and translation.

Copyrights permissions for Figures Copyright permissions have been obtained for the figures reproduced from non-open access sources (Figs. 5, 7, 9 and 11).

Competing interests The authors declare no competing interests.

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

References

1. Holm EA, Cohn R, Gao N, Kitahara AR, Matson TP, Lei B et al (2020) Overview: Computer Vision and Machine Learning for Microstructural Characterization and Analysis. Metall Mater Trans A 51:5985–5999. https://doi.org/10.1007/s11661-020-06008-4
2. Chen Y, Fu Z, Wu Y, Xu Y, Xiao Y, Wang J et al (2021) Giant heterogeneous magnetostriction induced by charge accumulation-mediated nanoinclusion formation in dual-phase nanostructured systems. Acta Mater 213:116975. https://doi.org/10.1016/j.actamat.2021.116975
3. Mu S, Al-Samman T, Mohles V, Gottstein G (2011) Cluster type grain interaction model including twinning for texture prediction: Application to magnesium alloys. Acta Mater 59:6938–6948. https://doi.org/10.1016/j.actamat.2011.07.045
4. Ercetin A, Özgün Ö, Aslantaş K, Der O, Yalçın B, Şimşir E Aamir M (2023) Microstructural and Mechanical Behavior Investigations of Nb-Reinforced Mg–Sn–Al–Zn–Mn Matrix Magnesium Composites. Metals 13:1097. https://doi.org/10.3390/met13061097
5. Ercetin A, Pimenov DY (2021) Microstructure, Mechanical, and Corrosion Behavior of Al2O3 Reinforced Mg2Zn Matrix Magnesium Composites. Materials 14:4819. https://doi.org/10.3390/ma14174819
6. Mauro GM, Iasiello M, Bianco N, Chiu WKS, Naso V (2022) Mono- and Multi-Objective CFD Optimization of Graded Foam-Filled Channels. Materials 15:968. https://doi.org/10.3390/ma15030968
7. Ren Q-Q (2024) Understanding Processing–Microstructure–Property Relationships of Structural Alloys. Cryst (Basel) 14:666. https://doi.org/10.3390/cryst14070666
8. DeCost BL, Holm EA (2015) A computer vision approach for automated analysis and classification of microstructural image data. Comput Mater Sci 110:126–133. https://doi.org/10.1016/j.commatsci.2015.08.011
9. DeCost BL, Francis T, Holm EA (2017) Exploring the microstructure manifold: Image texture representations applied to ultrahigh carbon steel microstructures. Acta Mater 133:30–40. https://doi.org/10.1016/j.actamat.2017.05.014
10. Choudhary K, DeCost B, Chen C, Jain A, Tavazza F, Cohn R et al (2022) Recent advances and applications of deep learning methods in materials science. NPJ Comput Mater 8:59. https://doi.org/10.1038/s41524-022-00734-6
11. Azimi SM, Britz D, Engstler M, Fritz M, Mücklich F (2018) Advanced Steel Microstructural Classification by Deep Learning Methods. Sci Rep 8:2128. https://doi.org/10.1038/s41598-018-20037-5
12. Stuckner J, Harder B, Smith TM (2022) Microstructure segmentation with deep learning encoders pre-trained on a large microscopy dataset. NPJ Comput Mater 8:200. https://doi.org/10.1038/s41524-022-00878-5
13. Jung J, Na J, Park HK, Park JM, Kim G, Lee S et al (2021) Super-resolving material microstructure image via deep learning for microstructure characterization and mechanical behavior analysis. NPJ Comput Mater 7:96. https://doi.org/10.1038/s41524-021-00568-8
14. Sharma R, Chelladurai I, Orme AD, Miles MP, Giraud-CARRIER C, Fullwood DT. (2018) A step towards intelligent EBSD microscopy: machine-learning prediction of twin activity in MgAZ31. J Microsc 272(1):67–78. https://doi.org/10.1016/j.isci.2020.101220
15. Akers S, Kautz E, Trevino-Gavito A, Olszta M, Matthews BE, Wang L et al (2021) Rapid and flexible segmentation of electron microscopy data using few-shot machine learning. NPJ Comput Mater 7:187. https://doi.org/10.1038/s41524-021-00652-z
16. Hirabayashi Y, Iga H, Ogawa H, Tokuta S, Shimada Y, Yamamoto A (2024) Deep learning for three-dimensional segmentation of electron microscopy images of complex ceramic materials. NPJ Comput Mater 10:46. https://doi.org/10.1038/s41524-024-01226-5
17. Wang X, Su H, Li N, Chen Y, Yang Y, Meng H (2023) A Deep Learning Labeling Method for Material Microstructure Image Segmentation. Processes 11:3272. https://doi.org/10.3390/pr11123272
18. Ma X, Yu Y (2023) Training Tricks for Steel Microstructure Segmentation with Deep Learning. Processes 11:3298. https://doi.org/10.3390/pr11123298
19. Horwath JP, Zakharov DN, Mégret R, Stach EA (2020) Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images. NPJ Comput Mater 6:108. https://doi.org/10.1038/s41524-020-00363-x
20. Awais M, Naseer M, Khan S, Anwer RM, Cholakkal H, Shah M et al (2025) Foundation Models Defining a New Era in Vision: A Survey and Outlook. IEEE Trans Pattern Anal Mach Intell 47:2245–2264. https://doi.org/10.1109/TPAMI.2024.3506283
21. Khan S, Naseer M, Hayat M, Zamir SW, Khan FS, Shah M (2022) Transformers in Vision: A Survey. ACM Comput Surv 54:1–41. https://doi.org/10.1145/3505244
22. Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, et al (2023) Segment anything. In: Proc IEEE/CVF Int Conf Comput Vis (ICCV), IEEE, pp 3992–4003. https://doi.org/10.1109/ICCV51070.2023.00371
23. Archit A, Freckmann L, Nair S, Khalid N, Hilt P, Rajashekar V et al (2025) Segment Anything for Microscopy. Nat Methods 22:579–591. https://doi.org/10.1038/s41592-024-02580-4
24. Tung P-Y, Harrison RJ (2025) Efficient microstructure segmentation in three-dimensional imaging: Combining few-shot learning with the segment anything model. Next Mater 8:100663. https://doi.org/10.1016/j.nxmate.2025.100663
25. Pan A, Xiang Q, Tang H, Wu K, Cheng L (2025) Microstructure segmentation and quantitative study of hot-rolled strip steel based on deep learning and SAM big model. Mater Charact 229:115486. https://doi.org/10.1016/j.matchar.2025.115486
26. Whitman SE, Latypov MI (2025) Machine learning of microstructure–property relationships in materials leveraging microstructure

representation from foundational vision transformers. Acta Mater 296:121217. https://doi.org/10.1016/j.actamat.2025.121217
27. Zhong X, Gallagher B, Liu S, Kailkhura B, Hiszpanski A, Han TY-J (2022) Explainable machine learning in materials science. NPJ Comput Mater 8:204. https://doi.org/10.1038/s41524-022-00884-7
28. McHugh JG, Mouratidis P, Jolley K (2022) Dataset for stacking-mediated diffusion of ruthenium nanoclusters in bilayer graphene and graphite. Data Brief 45:108702. https://doi.org/10.1016/j.dib.2022.108702
29. Chowdhury A, Kautz E, Yener B, Lewis D (2016) Image driven machine learning methods for microstructure recognition. Comput Mater Sci 123:176–187. https://doi.org/10.1016/j.commatsci.2016.05.034
30. Zhao Y, Guo N, Long S, Huang X, Song B, Hu J et al (2020) Dataset for microstructure and mechanical properties of (CrCoNi)97Al1.5Ti1.5 medium entropy alloy twisted by free-end-torsion at room and cryogenic temperatures. Data Brief 33:106333. https://doi.org/10.1016/j.dib.2020.106333
31. Stinville JC, Hestroffer JM, Charpagne MA, Polonsky AT, Echlin MP, Torbet CJ et al (2022) Multi-modal Dataset of a Polycrystalline Metallic Material: 3D Microstructure and Deformation Fields. Sci Data 9:460. https://doi.org/10.1038/s41597-022-01525-w
32. Ma B, Wei X, Liu C, Ban X, Huang H, Wang H et al (2020) Data augmentation in microscopic images for material data mining. NPJ Comput Mater 6:125. https://doi.org/10.1038/s41524-020-00392-6
33. Ma J, Hu C, Zhou P, Jin F, Wang X, Huang H (2023) Review of Image Augmentation Used in Deep Learning-Based Material Microscopic Image Segmentation. Appl Sci 13:6478. https://doi.org/10.3390/app13116478
34. Lee K-H, Yun GJ (2024) Microstructure reconstruction using diffusion-based generative models. Mech Adv Mater Struct 31:4443–4461. https://doi.org/10.1080/15376494.2023.2198528
35. Düreth C, Seibert P, Rücker D, Handford S, Kästner M, Gude M (2023) Conditional diffusion-based microstructure reconstruction. Mater Today Commun 35:105608. https://doi.org/10.1016/j.mtcomm.2023.105608
36. Azqadan E, Jahed H, Arami A (2023) Predictive microstructure image generation using denoising diffusion probabilistic models. Acta Mater 261:119406. https://doi.org/10.1016/j.actamat.2023.119406
37. Haralick RM, Shanmugam K, Dinstein I (1973) Textural Features for Image Classification. IEEE Trans Syst Man Cybern SMC–3:610–621. https://doi.org/10.1109/TSMC.1973.4309314
38. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24:971–987. https://doi.org/10.1109/TPAMI.2002.1017623
39. Laws KI (1980) Rapid texture identification. In: Wiener TF (ed), pp 376–381. https://doi.org/10.1117/12.959169
40. Clausi DA, Ed Jernigan M (2000) Designing Gabor filters for optimal texture separability. Pattern Recognit 33:1835–1849. https://doi.org/10.1016/S0031-3203(99)00181-8
41. Bianconi F, Fernández A (2007) Evaluation of the effects of Gabor filter parameters on texture classification. Pattern Recognit 40:3325–3335. https://doi.org/10.1016/j.patcog.2007.04.023
42. Sell K, Saenger EH, Falenty A, Chaouachi M, Haberthür D, Enzmann F et al (2016) On the path to the digital rock physics of gas hydrate-bearing sediments – processing of in situ synchrotron-tomography data. Solid Earth 7:1243–1258. https://doi.org/10.5194/se-7-1243-2016
43. Zhu B, Chen Z, Hu F, Dai X, Wang L, Zhang Y (2022) Feature Extraction and Microstructural Classification of Hot Stamping Ultra-High Strength Steel by Machine Learning. JOM 74:3466–3477. https://doi.org/10.1007/s11837-022-05265-5
44. Shen C, Wang C, Huang M, van der Xu N, Xu W (2021) A generic high-throughput microstructure classification and quantification method for regular SEM images of complex steel microstructures combining EBSD labeling and deep learning. J Mater Sci Technol 93:191–204. https://doi.org/10.1016/j.jmst.2021.04.009
45. Otsu N (1979) A Threshold Selection Method from Gray-Level Histograms. IEEE Trans Syst Man Cybern 9:62–66. https://doi.org/10.1109/TSMC.1979.4310076
46. Vincent L, Soille P (1991) Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans Pattern Anal Mach Intell 13:583–598. https://doi.org/10.1109/34.87344
47. Campbell A, Murray P, Yakushina E, Marshall S, Ion W (2018) New methods for automatic quantification of microstructural features using digital image processing. Mater Des 141:395–406. https://doi.org/10.1016/j.matdes.2017.12.049
48. Chan TF, Vese LA (2001) Active contours without edges. IEEE Trans Image Process 10:266–277. https://doi.org/10.1109/83.902291
49. Peregrina-Barreto H, Terol-Villalobos IR, Rangel-Magdaleno JJ, Herrera-Navarro AM, Morales-Hernández LA, Manríquez-Guerrero F (2013) Automatic grain size determination in microstructures using image processing. Measurement 46:249–258. https://doi.org/10.1016/j.measurement.2012.06.012
50. Fotos G, Campbell A, Murray P, Yakushina E (2023) Deep learning enhanced Watershed for microstructural analysis using a boundary class semantic segmentation. J Mater Sci 58:14390–14410. https://doi.org/10.1007/s10853-023-08901-w
51. Kang W, Chen Q, Huang L, Zhang J, Hou Z, Wang X et al (2023) Investigation of the Mesoscale Damage Evolution Process of AA5754O Aluminum Alloy CMT Welded Joints. Met (Basel) 13:555. https://doi.org/10.3390/met13030555
52. Hanchuan Peng F, Ding L (2005) Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell 27:1226–1238. https://doi.org/10.1109/TPAMI.2005.159
53. Kononenko I (1994) Estimating attributes: analysis and extensions of RELIEF. pp 171–182. https://doi.org/10.1007/3-540-57868-4_57
54. Principal component analysis. Springer-Verlag, New York (2002). https://doi.org/10.1007/b98835
55. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297. https://doi.org/10.1007/BF00994018
56. Breiman L, Random Forests (2001) Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324
57. Gola J, Webel J, Britz D, Guitar A, Staudt T, Winter M et al (2019) Objective microstructure classification by support vector machine (SVM) using a combination of morphological parameters and textural features for low carbon steels. Comput Mater Sci 160:186–196. https://doi.org/10.1016/j.commatsci.2019.01.006
58. Müller M, Britz D, Staudt T, Mücklich F (2021) Microstructural Classification of Bainitic Subclasses in Low-Carbon Multi-Phase Steels Using Machine Learning Techniques. Met (Basel) 11:1836. https://doi.org/10.3390/met11111836
59. Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. Commun ACM 60:84–90. https://doi.org/10.1145/3065386
60. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proc IEEE Conf Comput Vis Pattern Recognit (CVPR), IEEE, pp 770–778. https://doi.org/10.1109/CVPR.2016.90
61. Ronneberger O, Fischer P, Brox T (2015) U-Net: convolutional networks for biomedical image segmentation. pp 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
62. Bangaru SS, Wang C, Zhou X, Hassan M (2022) Scanning electron microscopy (SEM) image segmentation for microstructure

analysis of concrete using U-net convolutional neural network. Autom Constr 144:104602. https://doi.org/10.1016/j.autcon.2022.104602
63. Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL (2018) DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Trans Pattern Anal Mach Intell 40:834–848. https://doi.org/10.1109/TPAMI.2017.2699184
64. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S et al (2020) Generative adversarial networks. Commun ACM 63:139–144. https://doi.org/10.1145/3422622
65. Hong JH, Zach C, Fitzgibbon A (2017) Revisiting the variable projection method for separable nonlinear least squares problems. In: Proc IEEE Conf Comput Vis Pattern Recognit (CVPR), IEEE, pp 5939–5947. https://doi.org/10.1109/CVPR.2017.629
66. Isola P, Zhu J-Y, Zhou T, Efros AA (2017) Image-to-image translation with conditional adversarial networks. In: Proc IEEE Conf Comput Vis Pattern Recognit (CVPR), IEEE, pp 5967–5976. https://doi.org/10.1109/CVPR.2017.632
67. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2017) Grad-CAM: visual explanations from deep networks via gradient-based localization. In: Proc IEEE Int Conf Comput Vis (ICCV), IEEE, pp 618–626. https://doi.org/10.1109/ICCV.2017.74
68. Jangid DK, Brodnik NR, Goebel MG, Khan A, Majeti S, Echlin MP et al (2022) Adaptable physics-based super-resolution for electron backscatter diffraction maps. NPJ Comput Mater 8:255. https://doi.org/10.1038/s41524-022-00924-2
69. Durmaz AR, Müller M, Lei B, Thomas A, Britz D, Holm EA et al (2021) A deep learning approach for complex microstructure inference. Nat Commun 12:6272. https://doi.org/10.1038/s41467-021-26565-5
70. Starikov S, Grigorev P, Olsson PAT (2024) Angular-dependent interatomic potential for large-scale atomistic simulation of W-Mo-Nb ternary alloys. Comput Mater Sci 233:112734. https://doi.org/10.1016/j.commatsci.2023.112734
71. Huang Y, Xiang Z, Qian M (2022) Deep-learning-based porous media microstructure quantitative characterization and reconstruction method. Phys Rev E 105:015308. https://doi.org/10.1103/PhysRevE.105.015308
72. Kaufmann K, Lane H, Liu X, Vecchio KS (2021) Efficient few-shot machine learning for classification of EBSD patterns. Sci Rep 11:8172. https://doi.org/10.1038/s41598-021-87557-5
73. Bárcena M, Lloret Iglesias L, Ferreño D, Carrascal I (2024) Classification of cast iron alloys through convolutional neural networks applied on optical microscopy images. Steel Res Int 95. https://doi.org/10.1002/srin.202400120
74. Vasan V, Sridharan NV, Vaithiyanathan S, Aghaei M (2024) Detection and classification of surface defects on hot-rolled steel using vision transformers. Heliyon 10:e38498. https://doi.org/10.1016/j.heliyon.2024.e38498
75. Gao Y, Lv G, Xiao D, Han X, Sun T, Li Z (2024) Research on steel surface defect classification method based on deep learning. Sci Rep 14:8254. https://doi.org/10.1038/s41598-024-58643-1
76. Zhang T, Pan P, Zhang J, Zhang X (2024) Steel Surface Defect Detection Algorithm Based on Improved YOLOv8n. Appl Sci 14:5325. https://doi.org/10.3390/app14125325
77. Martinez Ostormujof T, Purushottam Raj Purohit RRP, Breumier S, Gey N, Salib M, Germain L (2022) Deep Learning for automated phase segmentation in EBSD maps. A case study in Dual Phase steel microstructures. Mater Charact 184:111638. https://doi.org/10.1016/j.matchar.2021.111638
78. Zhang AC, Maguire SM, Ford JT, Composto RJ (2023) Using Focused Ion Beam Time-of-Flight Secondary Ion Mass Spectrometry to Depth Profile Nanoparticles in Polymer Nanocomposites. Microsc Microanal 29:1557–1565. https://doi.org/10.1093/micmic/ozad085
79. Bals J, Epple M (2023) Deep learning for automated size and shape analysis of nanoparticles in scanning electron microscopy. RSC Adv 13:2795–2802. https://doi.org/10.1039/D2RA07812K
80. Mosser L, Dubrule O, Blunt MJ (2017) Reconstruction of three-dimensional porous media using generative adversarial neural networks. Phys Rev E 96:043309. https://doi.org/10.1103/PhysRevE.96.043309
81. Amiri H, Vogel H, Plümper O (2024) New 2D to 3D reconstruction of heterogeneous porous media via deep generative adversarial networks (GANs). J Geophys Res: Mach Learn Comput 1. https://doi.org/10.1029/2024JH000178
82. Nguyen PCH, Vlassis NN, Bahmani B, Sun W, Udaykumar HS, Baek SS (2022) Synthesizing controlled microstructures of porous media using generative adversarial networks and reinforcement learning. Sci Rep 12:9034. https://doi.org/10.1038/s41598-022-12845-7
83. Wanni J, Bronkhorst CA, Thoma DJ (2024) Machine learning enhanced analysis of EBSD data for texture representation. NPJ Comput Mater 10:133. https://doi.org/10.1038/s41524-024-01324-4
84. Sharma R, Chelladurai I, Orme AD, Miles MP, Giraud-CARRIER C, Fullwood DT (2018) A step towards intelligent EBSD microscopy: machine-learning prediction of twin activity in MgAZ31. J Microsc 272(1):67–78. https://doi.org/10.1111/jmi.12748
85. Ercetin A, Akkoyun F, Şimşir E, Pimenov DY, Giasin K, Gowdru Chandrashekarappa MP et al (2021) Image Processing of Mg-Al-Sn Alloy Microstructures for Determining Phase Ratios and Grain Size and Correction with Manual Measurement. Materials 14:5095. https://doi.org/10.3390/ma14175095
86. Akkoyun F, Ercetin A (2022) Automated Grain Counting for the Microstructure of Mg Alloys Using an Image Processing Method. J Mater Eng Perform 31:2870–2877. https://doi.org/10.1007/s11665-021-06436-2
87. Ercetin A, Der O, Akkoyun F, Gowdru Chandrashekarappa M, Şener R, Çalışan M et al (2024) Review of Image Processing Methods for Surface and Tool Condition Assessments in Machining. J Manuf Mater Process 8:244. https://doi.org/10.3390/jmmp8060244
88. Wang Q, Dong H, Huang H (2024) Swin-Transformer -YOLOv5 for lightweight hot-rolled steel strips surface defect detection algorithm. PLoS ONE 19:e0292082. https://doi.org/10.1371/journal.pone.0292082
89. Zhang H, Li S, Miao Q, Fang R, Xue S, Hu Q et al (2024) Surface defect detection of hot rolled steel based on multi-scale feature fusion and attention mechanism residual block. Sci Rep 14:7671. https://doi.org/10.1038/s41598-024-57990-3
90. Song X, Cao S, Zhang J, Hou Z (2024) Steel Surface Defect Detection Algorithm Based on YOLOv8. Electron (Basel) 13:988. https://doi.org/10.3390/electronics13050988
91. Chen Z, Feng X, Liu L, Jia Z (2023) Surface defect detection of industrial components based on vision. Sci Rep 13:22136. https://doi.org/10.1038/s41598-023-49359-9
92. Chen H, Qiu J, Gao D, Qian L, Li X (2024) Research on surface defect detection model of steel strip based on MFFA-YOLOv5. IET Image Process 18:2105–2113. https://doi.org/10.1049/ipr2.13086
93. Das A, Dorafshan S, Kaabouch N (2024) Autonomous Image-Based Corrosion Detection in Steel Structures Using Deep Learning. Sensors 24:3630. https://doi.org/10.3390/s24113630
94. Arı PD, Akkoyun F, Ercetin A (2025) A Machine Vision System for Gear Defect Detection. Processes 13:1727. https://doi.org/10.3390/pr13061727

95. Keller JM, Gray MR, Givens JA (1985) A fuzzy K-nearest neighbor algorithm. IEEE Trans Syst Man Cybern SMC–15:580–585. https://doi.org/10.1109/TSMC.1985.6313426
96. Bian Z, Vong CM, Wong PK, Wang S (2022) Fuzzy KNN Method With Adaptive Nearest Neighbors. IEEE Trans Cybern 52:5380–5393. https://doi.org/10.1109/TCYB.2020.3031610
97. Lin C-F, Wang S-D (2002) Fuzzy support vector machines. IEEE Trans Neural Netw 13:464–471. https://doi.org/10.1109/72.991432
98. Gao Y, Jiang J, Pan J, Yuan B, Zhang H, Zhu Q (2025) Contrastive learning-based fuzzy support vector machine. Neurocomputing 637:130101. https://doi.org/10.1016/j.neucom.2025.130101
99. Memiş S, Enginoğlu S, Erkan U (2022) Fuzzy parameterized fuzzy soft k-nearest neighbor classifier. Neurocomputing 500:351–378. https://doi.org/10.1016/j.neucom.2022.05.041
100. Enginoğlu S, Arslan B (2020) Intuitionistic fuzzy parameterized intuitionistic fuzzy soft matrices and their application in decision-making. Comput Appl Math 39:325. https://doi.org/10.1007/s40314-020-01325-1
101. Memiş S, Arslan B, Aydın T, Enginoğlu S, Camcı Ç (2023) Distance and Similarity Measures of Intuitionistic Fuzzy Parameterized Intuitionistic Fuzzy Soft Matrices and Their Applications to Data Classification in Supervised Learning. Axioms 12:463. https://doi.org/10.3390/axioms12050463
102. Memiş S (2023) Picture Fuzzy Soft Matrices and Application of Their Distance Measures to Supervised Learning: Picture Fuzzy Soft k-Nearest Neighbor (PFS-kNN). Electron (Basel) 12:4129. https://doi.org/10.3390/electronics12194129
103. Memiş S, Şola Erduran F, Aydoğan H (2025) Adaptive machine learning approaches utilizing soft decision-making via intuitionistic fuzzy parameterized intuitionistic fuzzy soft matrices. PeerJ Comput Sci 11:e2703. https://doi.org/10.7717/peerj-cs.2703

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