Materials Characterization 174 (2021) 110978

Contents lists available at ScienceDirect

Materials Characterization

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

An automated and unbiased grain segmentation method based on directional reflectance microscopy

Mallory Wittwer [a], Bernard Gaskey [a], Matteo Seita [a,b,*]

[a] School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore
[b] School of Materials Science and Engineering, Nanyang Technological University, Singapore

ARTICLE INFO

Keywords:
Grain segmentation
Directional reflectance microscopy
Polycrystalline metals
Automated image processing
Machine-learning
Dimensionality reduction

ABSTRACT

Identifying individual grains from sectioned polycrystalline metals is a foundational task of microstructure analysis. However, traditional grain segmentation methods applied to optical micrographs may suffer from the lack of optical contrast between grains and require the manual selection of adjustable parameters to achieve acceptable segmentation results. We propose an alternative method which takes advantage of a multi-angle optical microscopy technique termed directional reflectance microscopy. By combining dimensionality reduction, similar-dissimilar classification, and multi-region merging of surface directional reflectance, our method enables fully automated and reliable grain segmentation of polycrystalline surfaces. We apply our method to metal samples with different crystal structures and grain orientation distributions. Our results suggest applicability of the method to a wide range of microstructures, enabling a more objective, robust, and universal characterization of polycrystalline metals.

1. Introduction

Many relevant properties of polycrystalline metals—including mechanical strength [1], toughness [2], and fracture resistance [3]—are governed by the morphology and size distribution of their constituent grains. Characterizing these quantities is essential to establish structure-property relationships and to predict performance of engineering materials. Several well-established techniques, including scanning electron microscopy [4–8], transmission electron microscopy [9], and three-dimensional X-ray diffraction [10] already enable the accurate analysis of grain structures in polycrystals. However, these techniques suffer from a high equipment cost and are both time- and resource-intensive. A fast, simple and affordable alternative method to analyze polycrystalline microstructures is to characterize the cross-section of chemically etched samples by optical microscopy [11,12]. Some etchants reveal the microstructure of a polycrystal by dissolving grain boundaries, leading to recognizable dark traces at the interfaces between grains. On the basis of the resulting optical contrast, micrographs can be segmented into grains semi-automatically, for instance using image processing methodologies that combine watershed segmentation [13,14] with region merging operations [15–17]. The result is a collection of separate domains corresponding to the constituent grains in the microstructure, which forms a simplified partitioning of the original image. Pixels belonging to the same grain are grouped under a common integer identifier.

Despite its simplicity and convenience, optical grain segmentation suffers from two main limitations. First, generic algorithms such as watershed usually fail to segment the image when grain boundaries do not etch deeply enough to create a clear optical contrast between adjacent grains, or when etching is not uniform across the sample surface [14]. Some low energy grain boundaries, such as coherent twins, are known to be less reactive to chemical etching and thus leave a comparatively light optical contrast which is difficult to detect [18]. There have been attempts at overcoming this limitation using artificial neural networks [19,20]. However, these approaches require the manual annotation of a training set for each category of samples and have so far only produced moderate improvements. The second limitation stems from the variability in both sample microstructure and surface preparation, which often leads to differences in optical contrast from sample to sample. To compensate for this recurrent variability, the standard practice is to let a human operator adjust parameters of the image processing methodology to meet the correct level of segmentation for each individual microstructure. The watershed algorithm, for example, relies on the definition of appropriate markers to be used as

---

* Corresponding author at: School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore.
E-mail address: mseita@ntu.edu.sg (M. Seita).

https://doi.org/10.1016/j.matchar.2021.110978
Received 3 November 2020; Received in revised form 10 February 2021; Accepted 12 February 2021

water sources for the segmentation. Similarly, in region-merging algorithms, a stopping criterion for the merging operation must be predetermined before running the algorithm. This workflow is reliant on prior assumptions about the grain structure, which makes grain segmentation semi-subjective. It also requires that users develop in-depth knowledge about the functioning of the image processing algorithms in order to correctly process the data and interpret the results. Overall, the systematic requirement for parameter tuning makes the analysis more cumbersome and may lead to uncertain or irreproducible results.

To avoid these limitations, we propose to employ directional reflectance microscopy (DRM) [21] as an optical technique for grain segmentation. DRM differs from conventional optical microscopy in several critical aspects. First, it requires the use of a crystallographic etchant rather than one which selectively attacks grain boundary traces. Crystallographic etching yields the formation of microscopic surface facets whose geometry and direction are dictated by the underlying grain orientation [22–24]. The micro-faceting leads to grain-specific anisotropic reflectance: the reflection intensity across the sample surface varies from grain to grain and as a function of the illumination direction relative to the viewer. DRM captures this anisotropic surface reflectance by collecting a series of optical micrographs of the polycrystalline sample while incrementally varying the position of a light source to control the incoming light direction [25]. Therefore, a key difference with traditional methods is that the optical contrast resulting from crystallographic etching takes place between differently oriented grains, as opposed to the contrast obtained between grain boundaries and grains interior when using a grain boundary etchant. Moreover, while standard optical microscopy techniques are based on a single micrograph of the sample surface, DRM produces a multi-dimensional dataset consisting of many micrographs; each one characterized by the optical contrast observed under a different illumination direction.

In this paper, we present a computational technique for segmenting the grain structure of different polycrystalline metals based on DRM. The principle of our technique is illustrated in Fig. 1 and detailed in the following sections. First, we use a dimensionality reduction algorithm to extract a reduced set of features from the DRM dataset. This first step enables the unsupervised selection of salient features in the reflectance signal that preserve distinguishability between grains and, at the same time, reduces computational complexity. We then train a logistic regression classifier (LRC) to distinguish locations belonging to different grains based on the corresponding reduced directional reflectance signal. Finally, we combine the trained LRC with a multi-region merging (MRM) operation into a unified algorithm—which we term LRC-MRM—to produce a segmentation of the polycrystalline microstructure. In contrast to standard optical techniques, our method is only weakly parametrized by tunable variables that are implemented primarily to enable faster computation, as opposed to influencing segmentation results. This distinctive feature improves consistency of the results and avoids human input, thereby providing a fully automated, unbiased grain structure analysis of polycrystals.

2. Materials and methods

We apply our workflow to three annealed samples of pure metals (nickel, aluminium, and titanium) and one sample of nickel-based alloy Inconel 718. Prior to the DRM measurement, we ground the samples using a sequence of silicon carbide paper from 320 grit to 4000 grit, and then etched them by immersion in a chemical bath at room temperature. The first sample, **Ni**, is a sample of randomly textured polycrystalline nickel (>99.99%) with equiaxed grains, which we annealed at 1300 °C for 30 min. We etched **Ni** in Marble's reagent (10 g copper sulfate, 50 mL 36% hydrochloric acid and 50 mL deionized water) for 180 min. The second sample, **Al**, is a sample of polycrystalline aluminium (>99.99%) with a pronounced {100} crystallographic texture in the out-of-plane direction. We annealed **Al** at 400 °C for 6 h and etched it in a solution of 36% hydrochloric acid for 6 min. **Ti** is a sample of commercially pure titanium (>99%) with equiaxed grains, which we annealed at 800 °C for 12 h. We etched **Ti** in Kroll's reagent (2 mL hydrofluoric acid, 6 mL nitric acid, 92 mL deionized water) for 2 min. Finally, **I718** is a sample of Inconel 718 (54% Ni, 18% Cr, 20% Fe, 5% Mo, 2%Nb, 1%Ti), which we produced by Directed Energy Deposition (DED) using a BeAM Magic 800 machine and laser power and tool speed of 1600 W and 610 mm/s, respectively. The sample contains a mixture of equiaxed and columnar grains that result from the intrinsic variations in solidification

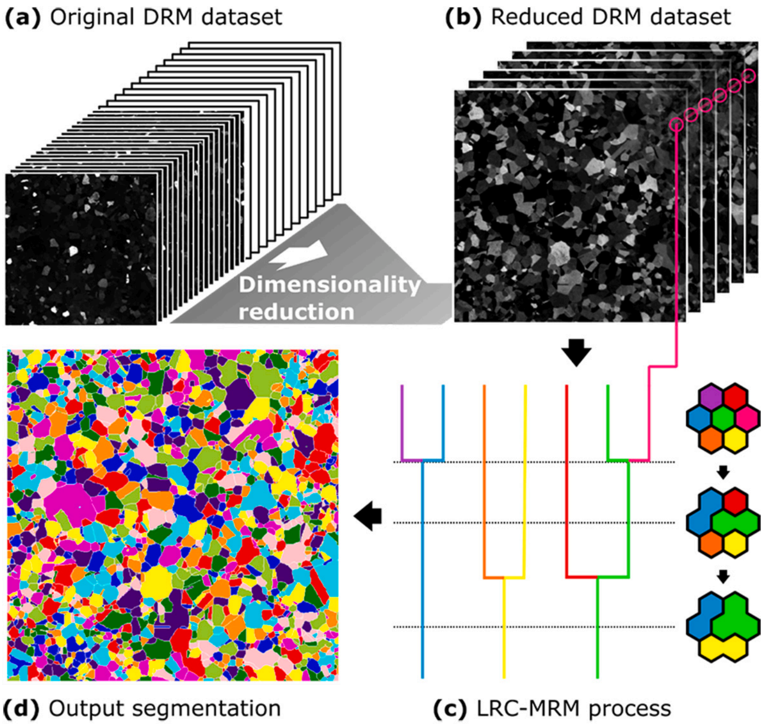


Fig. 1. Segmentation of a polycrystal surface based on directional reflectance microscopy (DRM). (a) A DRM dataset consists of a collection of optical micrographs captured under varying illumination directions. (b) The original dataset is reduced to a condensed set of features that preserve contrast between grains while minimizing computational complexity. (c) A logistic regression classifier (LRC) combined with a multi-region merging (MRM) operation is used to segment the microstructure automatically. (d) A segmentation of individual grains is generated by the LRC-MRM algorithm.

conditions during the DED process [26]. We heat-treated **I718** following the AMS5663 standard (60 min solution treatment at 954 °C followed by age hardening at 718 °C for 8 h, followed by age hardening at 621 °C for 8 h), ground and polished it to a mirror-like finish using a silica suspension solution, and then etched in a bath of Kalling's 2 reagent (5 g cupric chloride, 100 mL hydrochloric acid, 100 mL ethanol) for 10 min.

For each sample, we acquired a separate DRM dataset according to the procedure described in reference [25]. Our DRM apparatus includes an Olympus SZ6145 stereomicroscope equipped with a 110ALK0.3× objective lens, an industrial monochromatic CMOS camera with a resolution of 2448 × 2048 pixels, and a collimated white LED as light source. We sequentially capture micrographs of the sample surface while controlling the incident illumination angle by incrementally varying the light source azimuthal (φ) and elevation (θ) coordinates through a motorized stage. We vary θ between 15° and 65° by steps of 2.5°, φ all around the sample by steps of 5°, and collect one micrograph for each (φ, θ) combination. This wide range of motion—especially in the elevation—is enabled by the large working distance of our optical microscope (up to 35 cm), which provides enough clearance for the light source to move around the sample without masking the field of view. The complete DRM dataset contains 1′512 different micrographs, captured for each 72 × 21 combination of azimuth and elevation pair, respectively. The DRM dataset for **I718** consists of a reduced set of 72 × 8 = 432 micrographs taken by varying θ from 15° to 65° by steps of 10°.

We characterized crystal orientation in each sample by means of electron backscatter diffraction (EBSD, Oxford instruments Nordlys 2S detector) using a JEOL 7600F field emission scanning electron microscope. We employed a 15 mA emission current, 20 kV accelerating voltage, and a 15 μm step size. EBSD is a well-established electron microscopy technique that measures local crystal orientation on polycrystalline surfaces and is routinely used for accurate grain segmentation [27,28]. We provide the inverse pole figure (IPF) grain map of each sample in the Supplementary Materials. We used these measurements to compile a grain segmentation map for each sample through the built-in grain segmentation algorithm implemented in the Matlab MTEX software [5], which relies on a pre-defined misorientation angle threshold. We chose a threshold of 5° following ASTM standards [29]. We employ the EBSD-based grain segmentation as ground truth against which to compare our DRM-based segmentations results. For this purpose, we registered the two segmentation maps (from EBSD and DRM measurements) manually using a vector graphics editor software (Inkscape). During registration, we applied minor transformations—including rescaling, stretching, rotations, and shearing—to the EBSD map until we obtained a satisfactory alignment with the DRM map. We then converted the aligned EBSD map back to a labeled array to be used for further analysis through an in-house script written in the Python programming language.

3. Results and discussion

3.1. Sample selection and preparation

We select a database of polycrystals to be segmented that consists of three pure metals and one metal alloy. The rationale for using pure metals is to produce samples with large grains through annealing. While a coarse microstructure is not a requirement for DRM, it facilitates the analysis of the directional reflectance signal across individual grains as well as the evaluation of our segmentation method performance when compared against a reference measurement (e.g., EBSD). We include the results from **I718** to showcase the applicability of our DRM-based method to engineering alloys with complex microstructure and higher relevance to industrial applications.

Similarly to any other optical and electron microscopy technique, DRM is sensitive to the quality of the sample surface. Thus, sample preparation is a critical step to ensure correct grain segmentation. Previous studies have shown that surface imperfections, such as scratches, may yield to the identification of spurious grains [25], and that the reflectance signal varies with etching time and chemical etchant used [30]. However, DRM does not require a mirror-like surface finish (in contrast to EBSD). Instead, the desired surface state is usually obtained by fine mechanical grinding followed by chemical etching. This process yields reproducible results when using the same etching time, temperature, and chemical reagent on the same material. For more details on surface preparation for DRM, we refer to our previous works [21,25].

3.2. Reflectance datasets

Chemical etching induces different surface morphologies depending upon the material being etched and the reagent used. Using a crystallographic etchant on a pure metal often results in the formation of micro-facets across the sample surface, which correspond to specific crystallographic planes. Previous studies have shown that, in **Ni**, the etched surface is dominated by {111}-type facets, which exhibit the lowest surface energy in face centered cubic (FCC) materials (Fig. 2a) and thereby have higher resistance to chemical dissolution [22]. In **Al**, which is also FCC, the crystallography of the etched-induced facets evolves over time upon etching towards {100} planes, which are less thermodynamically stable than {111}, but kinetically favored in this material [31,32]. In metal alloys, such as **I718**, the etch-induced surface topography is more complex and dominated by the different corrosion resistance of the constituent phases (e.g., the matrix and the precipitates) or the elemental segregation that may be induced during processing of the material (e.g., the rapid solidification during DED) [12].

Regardless of their nature and crystallography, all these etch-induced structures impart directional reflectance to the sample surface. This directional reflectance signal measured across all illumination directions at a specific location on the sample surface—corresponding to a pixel in the optical micrograph—can be visualized versus the angles φ and θ in plots such as those shown in Fig. 2b. A comparison of such plots from pixels belonging to differently oriented grains shows a variety of configurations of reflection intensity maxima appearing at different (φ, θ) coordinates. Previous studies have shown that the angular coordinates of these reflection intensity maxima coincide with specular reflections of the incoming light beam at the etch-induced surface micro-facets [22,30]. Since the orientation of the micro-facets is determined by the underlying crystal orientation, each individual grain carries a distinctive reflectance pattern. Minimal differences exist between the reflectance patters of pixels belonging to the same grain, which may be attributed to surface contaminants or defects [25], minor intragranular misorientation [30], or measurement noise.

In the specific cases of **Ni** and **Al**, the sharp and well-defined peaks making up the reflectance patterns can be indexed and utilized to retrieve information about the underlying crystal orientation of the grains using analytical methods [22,30]. This feature provides the unique capability of using DRM to map grain orientation over the surface of crystalline solids in a similar fashion to EBSD. However, developing robust analytical models to correlate the reflectance patterns to crystal orientation is not a trivial task that can be carried out in all materials. In **Ti** and **I718**, for example, the reflectance patterns are difficult to interpret because of the complex etch-induced surface structure. In **Ti**, we observed that different grains exhibit different surface micro-faceting. However, we could not link the morphology of these facets to any unique family of crystallographic planes with certainty. In **I718**, the directional reflectance results from the reflection of secondary phase precipitates, which are more resistant to corrosion than the matrix and cause the formation of a corrugated surface (see Supplementary Fig. S3). However, the small size of the etched precipitates produces a more diffused directional reflectance signal, whose features are difficult to distinguish and to analyze quantitatively. In both these cases, developing a robust analytical method for indexing grain orientation by DRM is challenging. However, resolving the grain structure

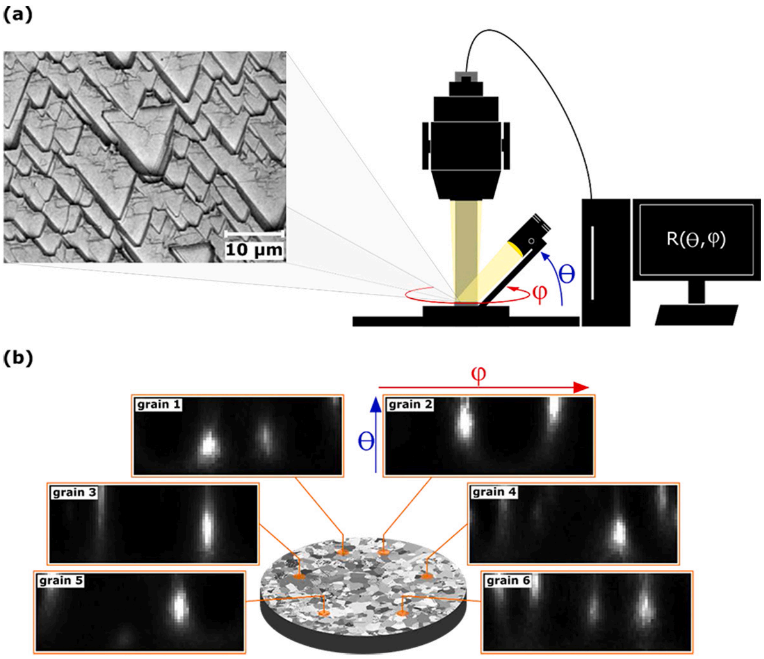
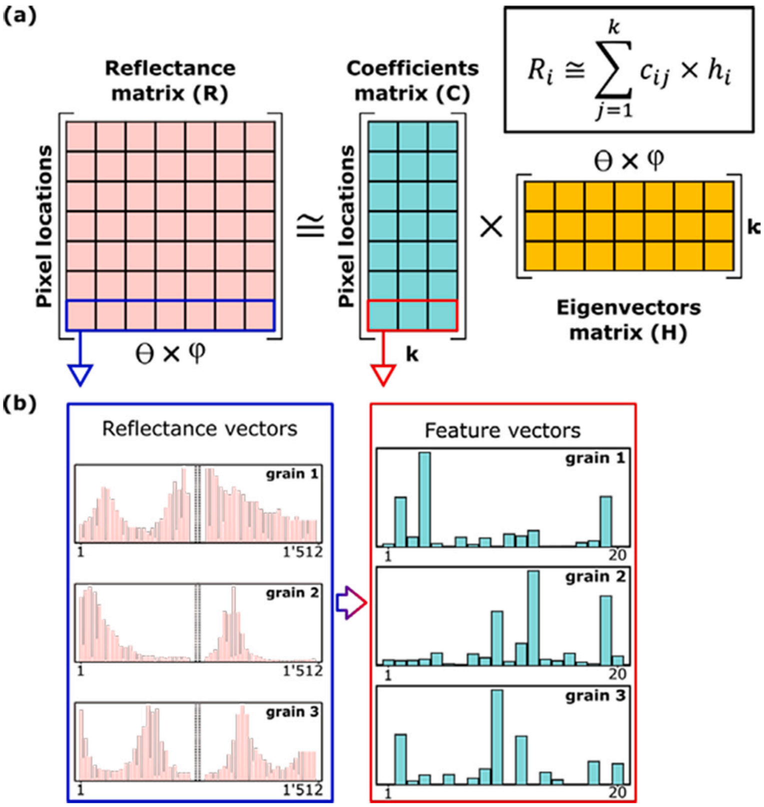


Fig. 2. Surface characterization by DRM. (a) Example of etch-induced micro-faceting within a grain (here, in the **Ni** sample). This micro-faceting is responsible for the generation of grain-specific, directional surface reflectance. The DRM setup is designed to capture optical micrographs of the sample surface under multiple incident illumination directions parametrized by the azimuth (**φ**) and elevation (**θ**) coordinates of the light source. (b) Directional reflectance patterns of pixels belonging to differently oriented grains show reflection intensity maxima that coincide with specular reflections of the incoming light beam at the etch-induced surface facets.

and performing automated and unbiased grain segmentation based on directional reflectance is still feasible, as we demonstrate in the following sections.

3.3. Dimensionality reduction

The large number of variables captured by DRM—which describe illumination direction and local surface reflectance—produces a high-dimensional signal, which is conducive to enhancing separability between grains and thereby to improving segmentation as compared to conventional, single-image optical techniques. However, analyzing high-dimensional data can be computationally complex. To facilitate the analysis, we first apply dimensionality reduction to the DRM dataset. The goal of dimensionality reduction is to transform the original, high-dimensional dataset to a lower-dimensional representation that retains some of the meaningful properties of the original dataset, such as separability between grains. To achieve this goal, we select Non-Negative Matrix Factorization (NMF) [33] as a dimensionality



$$R_i \cong \sum_{j=1}^{k} c_{ij} \times h_i$$

Fig. 3. Decomposition of a directional reflectance dataset by non-negative matrix factorization (NMF). (a) The original reflectance matrix ***R*** is factorized into a product of two smaller matrices: the matrix of coefficients ***C*** and the matrix of eigenvectors ***H***. The matrix ***C*** is a low-dimensional representation of ***R***. (b) In this example, the feature vectors forming the rows of the matrix ***C*** are determined by only 20 component variables, which is significantly less than the original 1′512 reflection intensities constituting the DRM signals.

reduction algorithm. NMF is suitable to model additive mixtures of signals and applicable to datasets where only non-negative signals exist, which is the case for reflection intensity measurements.

To apply NMF, we consider the DRM dataset in the form of a 2D matrix, R. The columns of R represent the measured illumination directions, as defined by (φ, θ) coordinates, while the rows represent reflection intensities measured at individual pixel locations. The NMF algorithm decomposes R into a product of two smaller positive matrices: the matrix of coefficients, C, and that of eigenvectors, H (Fig. 3a). The factorization is designed so that the product of C and H closely approximates R. In other terms, the matrix C is a reduced (or compressed) representation of R. Referring to the rows of R as reflectance vectors and to the rows of C as feature vectors, each reflectance vector is associated to a corresponding feature vector, which serves as its compressed, low-dimensional representation (Fig. 3b). The length of the feature vectors, k, corresponds to the size of the NMF basis. It can be significantly reduced as compared to the length of the reflectance vectors, $\varphi \times \theta$, while preserving distinguishability between grains. Pixel locations with similar reflectance vectors share similar feature vectors and, conversely, pixel locations with different reflectance vectors map to different feature vectors.

The NMF model may be fitted on the entire dataset of reflectance vectors. To reduce computation time, however, it is more practical to fit the model on a randomly selected subset of examples. In this study, we use a subset of 10′000 examples of reflectance vectors in the dataset to fit the NMF model. The number 10′000 is arbitrarily chosen to account for a statistically representative number of examples that well-represent the original dataset. The size of this sampling can therefore be viewed as an *optional* tunable parameter of our method. Similarly, the size of the NMF basis remains a hyperparameter of the method. In all our case-study samples, we set the size of the dimensionality reduction basis to 20 components. The tradeoff when determining the size of the NMF basis happens primarily between running time and the risk of over-compressing the dataset, which could negatively impact performance. As a guideline, we recommend that the number of components exceeds a reasonable minimum of 15.

3.4. Similar-dissimilar classification

We perform grain segmentation on the reduced DRM dataset following a pairwise similar-dissimilar classification approach [34,35]. The goal of this classification is to model the level of dissimilarity between two arbitrary feature vectors, A and B, which indicates whether A and B are part of different grains or the same grain. To this end, we first compute the distance between A and B as $d = (A - B)^2$ elementwise: the output of the distance operation is a vector that contains the squared differences between the elements of A and B. We refer to this vector as the “distance vector” and term its components “distance components”. We call S and D the multivariate statistical distributions of distance components belonging to the same grain and to different grains, respectively. The objective of the similar-dissimilar classification is to draw a boundary between S and D which provides optimal separation between the two distributions. To infer this boundary, we rely on a multivariate logistic regression classifier (LRC), which is a type of supervised linear classifier [36]. The LRC is used to predict the probability, p, of a distance vector, $(x_1, x_2, \ldots x_n)$, to be part of D (or S), which is expressed as

$$p = \frac{1}{1 + exp(-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n))} \tag{1}$$

The LRC derives this probability by applying a logistic function to a linear combination of the distance components. The weights $(\beta_1, \beta_2, \ldots \beta_n)$ of the linear combination as well as the bias term, β_0, are trainable parameters of the model, which are to be inferred for each dataset. While it is common to regularize logistic regression, for example by L1 or L2 penalty, in this study we do not use a regularization term to avoid introducing extra hyperparameters. In the code provided in the Supplementary Materials, however, we provide the option to use L2 regularization and to cross-validate regularization strength to identify the highest performing value. The trainable parameters of the logistic regression are determined by fitting the model on a training set, which is a subset of examples marked with the correct class labels. Labeling the training set based on grain membership (class S or D) could be achieved by manually marking grains on a representative portion of each sample. However, this approach is highly time-intensive and introduces operator bias, which is incompatible with the purpose of an automated and unbiased grain segmentation workflow.

As an alternative, we use pixel adjacency in the optical micrographs as a proxy for grain membership (Fig. 4a). We consider *adjacent* the four pixels located directly north, south, west, and east of a reference pixel. We term $\widehat{S}$ the distribution of distance components corresponding to A-B pairs of feature vectors for which A and B are adjacent on the surface. We assume that $\widehat{S}$ approximates S because neighboring pixels have a high probability of belonging to the same grain. Likewise, we approximate D by the distribution of distance components for which A and B are non-adjacent, denoted as $\widehat{D}$, based on the assumption that most instances within the ensemble of non-adjacent pixels correspond to pairs of pixels that belong to different grains. Therefore, we construct a training set made of extracted samples of $\widehat{S}$ and $\widehat{D}$ by randomly selecting examples of adjacent and non-adjacent pairs of feature vectors in the dataset, respectively. This training set is balanced: it contains the same number of examples from both classes. In this study, we select 10′000 examples of each class for each dataset to account for a statistically representative number of examples that well-represent the class distributions. The size of the subset of examples chosen to form the training set of the logistic regression model can be considered as another *optional* parameter of our method, as in the case of the NMF model. It has strong impact on running time and weak impact on the results, so long as the number of samples selected is statistically representative of the entire dataset (we recommend a minimum of 1000).

We fit the LRC on the extracted training set (Fig. 4c) to derive a linear classification boundary between $\widehat{S}$ and $\widehat{D}$. The algorithm used to train the model was a Broyden–Fletcher–Goldfarb–Shannon (BFGS) solver [37]. If our assumption that pixel adjacency accurately approximates grain membership is correct, the predictions of the LRC model are equally applicable to determining grain membership, under the terms of the original similar-dissimilar classification.

3.5. Multi-region merging

To produce a grain map, we integrate the trained LRC to a multi-region merging (MRM) algorithm. We call the combined method LRC-MRM. A “region” is a cluster of pixels which are grouped under a common integer identifier—and therefore considered part of the same grain—at any given iteration. Moreover, we assign to each region the feature vector of one representative pixel of that region called the “master pixel”. The algorithm starts with a segmentation in which every pixel is a master pixel forming an independent region. Then, pairs of adjacent regions are recursively merged to form progressively larger and larger regions. The merging of two regions is governed by the output of the LRC applied to the distance vector associated with the master pixels of these regions. Considering two adjacent regions X and Y and their associated master pixels x and y, and calling the LRC output probability p and the distance function d, the merging of X and Y is conducted when $p(d(x,y)) \leq 0.5$; namely when the computed probability that X and Y are part of different grains is less than 50%. Moreover, we impose that each newly formed region, Z, resulting from the merging of X and Y inherits the master pixel of the largest of the two constituent regions (in pixel count). The rationale for this choice is that the largest constituent region is more likely to be representative of Z and, ultimately, of the underlying grain. Since the LRC can discriminate between pairs of master pixels

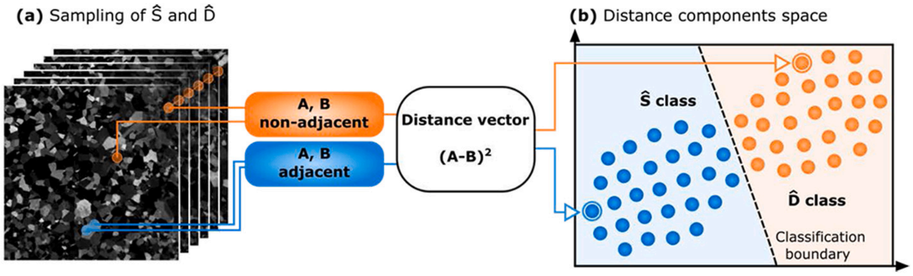


Fig. 4. Similar-dissimilar classification of feature vectors by a logistic regression model. (a) We randomly extract pairs of feature vectors at adjacent and non-adjacent pixel locations. We subtract and square the feature vectors to produce distance vectors belonging to the distributions $\widehat{S}$ and $\widehat{D}$, which approximate the distributions of distance vectors from the same grain (S) and from different grains (D), respectively. (b) Distance vectors are defined in a distance components space, in which we use a logistic regression classifier to draw a boundary between the ensembles $\widehat{S}$ and $\widehat{D}$ (the figure represents a simplified distance components space in two dimensions).

which belong to the same grain and to different grains, LRC-MRM gradually leads to a segmented grain map. The process terminates when all remaining regions and their neighbors are considered different grains by the LRC model.

3.6. Segmented grain maps

The segmentations produced by LRC-MRM applied to the studied samples **Ni**, **Al**, **Ti**, and **I718** are shown in Fig. 5. For each sample, we select a square portion of the original field of view with well-defined

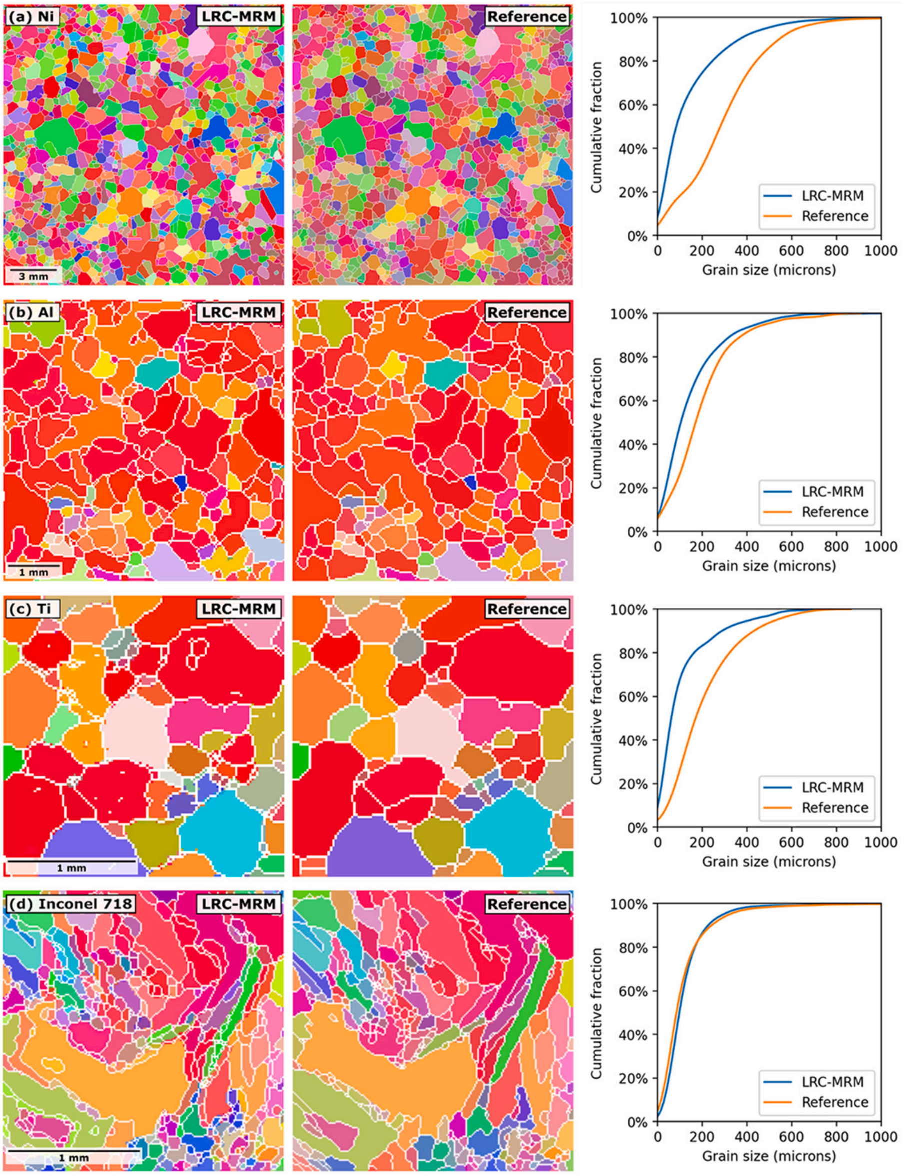


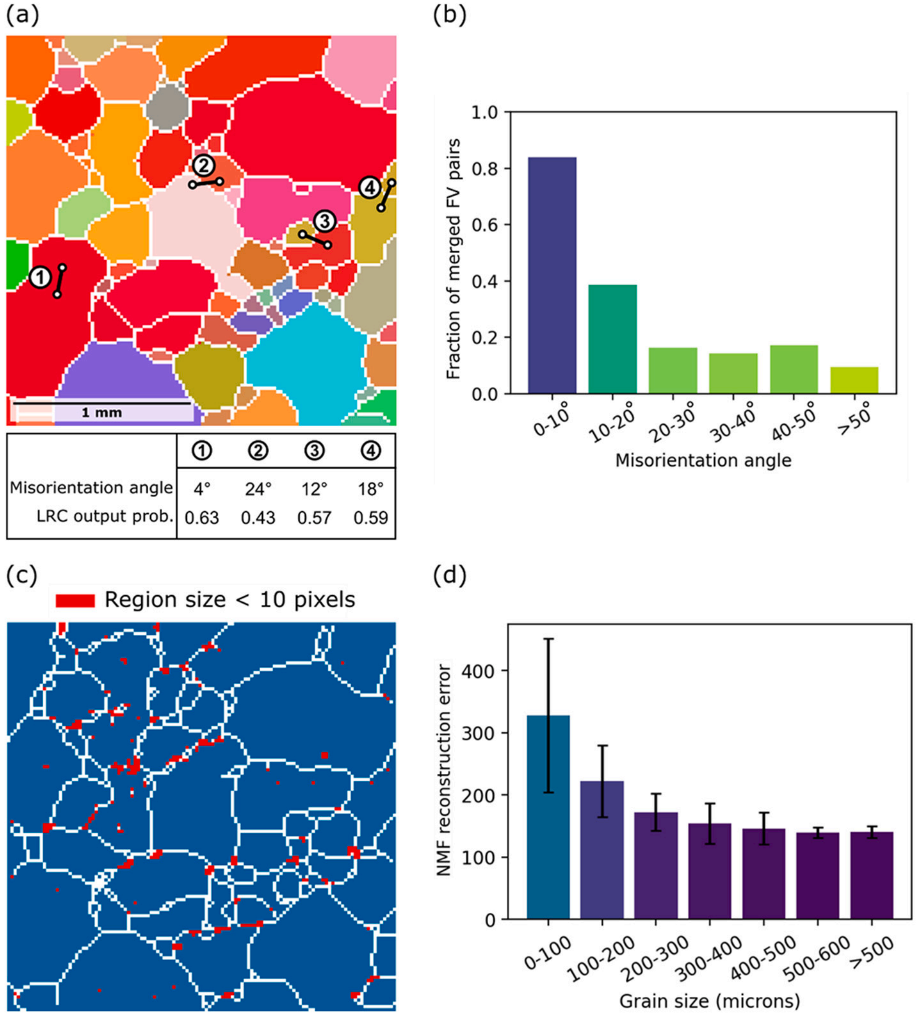
Fig. 5. Segmentation of grains by the LRC-MRM method. Comparison between LRC-MRM and EBSD segmentation and grain size distribution in (a) **Ni**, (b) **Al**, (c) **Ti**, and (d) I718. The colour coding in the grain maps is based on the inverse pole figure (IPF) plotted along the out-of-plane direction (i.e., the sample surface normal) as measured by EBSD.

grain structure as region of interest, which we use for testing the LRC-MRM algorithm. In **Ni**, the region of interest is an array of 694 × 694 pixels, in **Al,** 179 × 179 pixels, in **Ti,** 120 × 120 pixels, and in **I718**, 190 × 190 pixels. We process the data and implement the LRC-MRM algorithm using an in-house script written in the Python programming language. We use available models from the Scikit-learn package [38] for the NMF and logistic regression. When running on a personal computer, the segmentation of **Ni** is completed in four minutes and 18 s, 40 s for **Al,** 33 s for **Ti**, and 27 s for **I718**. To quantify the accuracy of our DRM-based method, we compare the LRC-MRM segmentations against the registered EBSD references (also shown in Fig. 5).

The results in Fig. 5 demonstrate that grains are generally correctly identified in all four samples, despite the differences in composition, crystal structure, etch-induced surface structures, grain morphology, and grain orientation distribution. The exact same algorithm was applied in all four cases without adjusting any parameter. These results suggest that DRM combined with LRC-MRM is a suitable technique for performing automated and unbiased grain segmentation of polycrystalline surfaces.

However, the maps in Fig. 5 also reveal some limitations of our method. Failures at larger grain scales in the form of (i) the erroneous merging of two or more distinct grains and (ii) the fractioning of single grains into multiple regions can be noticed. In many cases, these errors are not systematically reproducible when running the algorithm multiple times. We notice that keeping the training set to fit the LRC model constant across multiple runs always leads to the same segmentation output. Thus, we deduce that the observed differences in-between runs stem from the random sampling of the training set used to fit the LRC model, and do not depend on crystal orientation or grain geometry. We hypothesize that different training sets extracted from different random samplings can induce minor variations in the trained parameters of the LRC model, leading to a different order in the merging of regions. The exact order in which regions are merged can moderately affect the output segmentation by influencing the selection of master pixels. We study the occurrence of these events in **Ti**, by following four specific grain boundaries for which the adjacent grains are not always merged or kept separated over the course of multiple runs of the LRC-MRM segmentation algorithm (Fig. 6a). We notice that these grain boundaries are all characterized by a misorientation angle between 3° and 25°. Moreover, for each of these sensitive grain boundaries the output probability of the LRC model—computed in-between feature vectors of pixels from the corresponding adjacent grains—falls within an intermediate range between 0.40 and 0.65. This result indicates a high uncertainty in the decision to merge the grains or keep them separated. These selected grain boundaries are both in the vicinity of the classification boundary of the LRC model and near the misorientation threshold used to define separate grains. Grain boundaries with intermediate misorientation angle are probably the most difficult to identify and may cause discrepancies when comparing the segmentation from the LRC-MRM model against one that is based on a fixed misorientation threshold applied to a reference measurement—such as EBSD in our case.

Fig. 6b shows an estimate of the fraction of pairs of feature vectors that are merged by the LRC model in different ranges of misorientation angles in our **Ti** dataset. This estimate was obtained by computing the



	①	②	③	④
Misorientation angle	4°	24°	12°	18°
LRC output prob.	0.63	0.43	0.57	0.59

Fig. 6. Limitations of the LRC-MRM method. (a) In **Ti**, we select four grain boundaries that have been handled inconsistently over multiple runs of our LRC-MRM algorithm. These grain boundaries are associated with intermediate misorientation angles and classification probability in the range of 0.40–0.65, indicating a high uncertainty regarding the decision to merge the corresponding adjacent grains or not. (b) The fraction of feature vector pairs that are merged by the LRC model decreases sharply with misorientation angle. Intermediate misorientations in the range of 10°-20° are frequently merged and susceptible to produce segmentation errors when compared to references based on a 5° misorientation threshold. (c) In the output segmentation by our method (here, in the **Ti** sample), small regions concentrated at the interfaces between grains are mistakenly marked as separate grains, leading to an overrepresentation of small grains in the size distributions. (d) The reconstruction error of the NMF model increases in smaller segmented regions, which might explain the frequent mishandling of grain boundary regions. The displayed bar chart represents reconstruction error averaged in selected ranges of region sizes. Error bars represent one standard deviation from the mean.

LRC output on 10′000 randomly selected pairs of feature vectors in different misorientation ranges and counting the fraction of LRC outputs probabilities above 0.5. As expected, we observe that the merging fraction decreases rapidly with misorientation angle. However, a significant proportion of feature vector pairs corresponding to intermediate misorientations (between 10° and 20°), as well as a minor fraction of the higher misorientation pairs, are merged, which is erroneous in the framework of a misorientation threshold of 5° for grain boundary tracing. These results outline that the criterion used to determine the reference segmentations—namely the misorientation threshold—is critical, albeit arbitrarily defined. Consequently, failure to identify grain boundaries close to the misorientation threshold might be rationalized and—to some extent—deemed more acceptable than failure to recognize lowly (<3°) or highly (>25°) misoriented grains.

Furthermore, the cumulative grain size distribution from the LRC-MRM segmentation shows a significant overrepresentation of small grains when compared to the reference. We find that this effect is mainly caused by small clusters of pixels remaining in the vicinity of grain boundaries, which are erroneously interpreted as separate regions (Fig. 6c). A possible explanation for these artefacts is the limited ability of the NMF model to reconstruct grain boundary pixels systematically. The reconstruction error of the NMF model—computed as the mean squared error between the original reflectance signals contained in the reflectance matrix (R) and the product of the coefficients matrix (C) and the eigenvectors matrix (H)—increases in regions of smaller sizes (Fig. 6d). The bar chart in Fig. 6d was drawn by computing the reconstruction error of every master pixel in the LRC-MRM segmentation of **Ti** and grouping the values based on grain size. In this example, the ensemble of reconstruction errors above the median (middle quantile) contains 78% of grains below 50 pixels in size, while this fraction is only 52% in the total dataset. We observe similar trends in the other datasets. The difficulty to represent grain boundary signals systematically in the NMF model may stem from the fact that grain boundary pixels contain a mixture of reflectance signals coming from both adjoining grains, which are combined in proportions that depend on the position of the pixel relative to the physical grain boundary. As a result, substantial dissimilarities may exist between grain boundary signals and the signals of neighboring grain interiors. This inherent variability of the reflectance signals could be reflected in the corresponding feature vectors, leading to the erroneous handling of pixels nearby grain boundaries in the segmentation process.

3.7. Performance evaluation

We hypothesize that any segmentation failure can ultimately be explained by a series of one or more misclassification events in the record of all merging decisions taken by the algorithm over the course of the region-merging process. There are fundamentally two types of classification errors:

1. Instances of grains that were mistakenly merged in the segmentation. This category represents pairs of feature vectors that were predicted to be part of the ensemble $\widehat{S}$ (and therefore of S), but were in fact part of D. These erroneous merging decisions directly reflect the false negative (FN) instances of the LRC model.
2. Regions belonging to the same grain that were marked as different grains in the segmentation. These artefacts are the result of false positive (FP) instances of the LRC model, which correspond to pairs of feature vectors predicted as being part of the ensemble $\widehat{D}$ while they were in fact part of S.

As a metric to account for both types of errors, we define the accuracy of the LRC model as the ratio of all mislabeled instances (FP and FN) to the total number of instances including true positives (TP) and true negatives (TN):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

Moreover, to quantify the accuracy of the segmentation produced by the LRC-MRM method, we compute the mutual information score (MIS) between a grain map produced by the algorithm, U, and the reference, V:

$$MIS(U,V) = \sum_{i=1}^{|U|}\sum_{j=1}^{|V|}\frac{|U_i \cap V_j|}{N} log \frac{N|U_i \cap V_j|}{|U_i||V_j|} \tag{3}$$

MIS is bound between 0 and 1, with high values indicating high similarity between the two segmentations. We evaluate MIS and accuracy on a test set constructed by randomly sampling 10′000 distance vectors from the S and D classes in each of the three studied samples. The test set is labeled based on grain membership as inferred from the reference maps. Moreover, we benchmark performance of the LRC model against a baseline model. We design the baseline model as a simple, naive algorithm with intentionally limited ability to produce nuanced outputs. The aim of the baseline model is to define trivially attainable performance in order to put performance of the LRC model into context. The baseline model performs a classification based on a comparison between the mean intensity, I_m, of the distance components of a distance vector, $\boldsymbol{v}$, and the global mean intensity, I_g, of all distance components in the training set. Both I_m and I_g are scalar variables. The model assigns $\boldsymbol{v}$ to class D when I_m is greater than I_g, and to class S otherwise. We integrate the baseline model to the MRM algorithm (in place of the LRC), segment the datasets, compute the MIS and classification accuracy, and compare them against that obtained using the LRC model.

We report baseline performance metrics alongside performance values of the LRC-MRM algorithm in Table 1. To account for the variability in-between runs caused by the randomly selected training sets, we repeat the segmentation process 50 times for each sample and report the mean and standard deviation of the performance metrics. We find that both MIS and accuracy of the LRC-MRM method are consistently superior to baseline values. This result shows that a simple, naive classifier is not sufficient to successfully achieve segmentation, and confirms the necessity of using a multivariate model instead, such as logistic regression. Moreover, the reported values of classification accuracy of the LRC-MRM algorithm are consistently high, demonstrating the ability of the LRC model to correctly discriminate between pixels belonging to the same and to different grains. This result confirms our hypothesis that labeling distance vectors based on pixel adjacency is a suitable approximation for grain membership. Similarly, the high values of MIS indicate that the classification performed by the LRC model can effectively provide a criterion to stop the region merging process, leading to accurate segmentations. The values of accuracy and MIS in Table 1 may be used as benchmark to compare performance against future iterations of the LRC-MRM method or against other segmentation algorithms applied to the same datasets. These results should be interpreted considering that the segmentation errors discussed above—including failures in the intermediate misorientation range and artefacts at grain

Table 1
Classification accuracy and MIS (mean, based on 50 repetitions) of the LRC-MRM method and baseline. Standard deviations from the mean are indicated in parenthesis.

		Accuracy	MIS
Ni	LRC-MRM	0.969 (±0.001)	0.871 (±0.001)
	Baseline	0.920 (±0.003)	0.733 (±0.046)
Al	LRC-MRM	0.940 (±0.002)	0.883 (±0.005)
	Baseline	0.883 (±0.007)	0.764 (±0.024)
Ti	LRC-MRM	0.956 (±0.002)	0.840 (±0.006)
	Baseline	0.851 (±0.019)	0.641 (±0.086)
I718	LRC-MRM	0.989 (±0.001)	0.791 (±0.003)
	Baseline	0.932 (±0.010)	0.717 (±0.018)

boundaries—are what separate the reported MIS and accuracies from perfect scores. Adding up to these sources of errors, any misregistration between the EBSD and DRM datasets will further contribute to lowering the performance values. Because we register the datasets manually, we expect the registration to be slightly imperfect and to inevitably introduce a minor fraction of mislabeled data in the datasets.

Finally, we attempt to benchmark our method against a conventional optical segmentation method. Focusing on the **Ni** dataset, we compare the results of LRC-MRM against the segmentation of a single micrograph of the sample illuminated under diffuse (non-directional) white light using the watershed method. We use our DRM stereomicroscope to capture the image to avoid optical hardware bias, and the exact same etched surface to avoid surface preparation bias. We compute grain boundary contrast using a Sobel operator and 1′500 selected seed pixels to be used as markers for the water sources. We implement the watershed method provided by the Python Scikit-learn package. The resulting segmentation is available in the Supplementary Materials. The MIS compared to the EBSD reference is 0.841, which is lower than the LRC-MRM result (0.871). Moreover, a visual comparison between two segmentations suggests a clear advantage in favor of our method. However, this conclusion should be nuanced by the subjective nature of the watershed result. The chosen number of markers (1′500) is influential on the segmentation and was determined based on the visual judgement of an operator after trying different values. This result underlines the difficulty of defining a single, objective benchmark against which to test our method. Since most optical segmentation methods tend to rely strongly on surface and imaging conditions—as well as on tunable parameters—the produced benchmarks can be subjective and therefore potentially unreliable.

3.8. Grain size specification

Three of the studied samples are pure metals with coarse grain size in the range of several hundreds of micrometers. We note that these characteristics are not commonly found in engineering applications. By contrast, the **I718** sample exhibits a microstructure that is representative of DED processes [39], which are relevant for many industrial applications [40]. We find that the quality of grain segmentation in this sample is difficult to assess, especially in regions characterized by a small grain size (i.e., of the order of a few tens of microns). The lack of resolving power of our DRM apparatus makes it difficult to distinguish the grains in these regions and to register the DRM dataset with the EBSD reference for a quantitative analysis. If it were not for this constraint, we speculate that the minimum grain size that could be accurately segmented by our method would depend on the pixel resolution—which is set by the sensor resolution—as well as the size of the etch-pits covering the grains. Optimal results may be obtained when the pixel size is small compared to the grain size—such that each grain contains many pixels and the fraction of grain boundary pixels is minimized—but at the same time large compared to the etch-pit size—such that each pixel carries out the same directional reflectance signal within individual grains [22].

4. Conclusions

As an alternative to traditional methods based on a single image of the sample surface (i.e., conventional optical metallography), we present a new workflow designed for the segmentation of grains in sectioned polycrystalline metals based on directional reflectance microscopy (DRM). This workflow involves preparing the sample surface using a crystallographic etchant (as opposed to a grain boundary etchant) and capturing reflection intensity under varying light incident angles using the DRM apparatus. The measurement at a given pixel location is a reflectance signal which contains the directional reflectance pattern of the underlying grain. Reflectance patterns are specific to each individual grain, which makes them ideal inputs for grain segmentation.

While the original signals are high-dimensional, we show that their reduced representation can be extracted as feature vectors through non-negative matrix factorization (NMF). Feature vectors containing 20 components effectively preserve distinguishability between grains, while significantly alleviating the computational cost of the image analysis. To compile a grain map, we propose to combine a logistic regression classifier (LRC) with a multi-region-merging (MRM) operation. The LRC is trained to differentiate pairs of feature vectors corresponding to pixels which are adjacent and non-adjacent in the image and subsequently used to provide a stopping criterion for the MRM operation. The correct level of segmentation is determined based on the classification boundary of the logistic regression classifier, which is inferred automatically as part of training. As a result, the LRC-MRM method is robust, unbiased, and essentially material-agnostic. The size of the NMF basis and that of the randomly sampled training sets can be considered as *optional* tunable parameters of the method to enable faster computation. However, their impact on the segmentation result is minimal.

We test our algorithm on pure metals (nickel, aluminium, and titanium) and include an additively manufactured nickel-base superalloy sample (Inconel 718) with heterogeneous grain size distribution to showcase the applicability of our method to engineering alloys with complex microstructure. We demonstrate that our method produces accurate grain segmentation in all these cases despite the markedly different surface structures induced by chemical etching and the different types of directional reflectance signals measured by DRM. To quantitatively evaluate performance of our method, we compute classification accuracy and the mutual information score (MIS) between the segmentations produced by our method and references derived from EBSD. Both performance metrics are consistently high across all three studied samples, confirming the capability of the method to produce accurate segmentation results regardless of the material and microstructure considered. Benchmarking our LRC-MRM algorithm and DRM-based segmentation method against a baseline model and a conventional watershed segmentation method also confirms a clear advantage in favor of our approach. However, our method incorrectly handles pixels in the vicinity of grain boundaries, which are often mistakenly interpreted as distinct regions in the segmentations owing to their unique reflectance patterns. Although this is undeniably a downside, it only affects the lower end of the size distribution.

We believe that the presented results underline the potential of a DRM-based approach for the segmentation of a variety of different polycrystalline materials. While in this work we employed a specific optical stereomicroscope with large working distance (to enable a wide range of motion of the light source), we think that DRM measurements could be extended to generic optical microscopes too, although maybe with limited angular coverage. In this regard, it would be interesting to perform a systematic study of grain segmentation accuracy as a function of the angular range and resolution of the DRM dataset. Based on the results shown in here, however, we speculate that different materials prepared with different etching methods would have different requirements and that a universal DRM acquisition methodology does not exist. The simplicity, low cost, and enhanced sample throughput compared to alternative optical microscopy techniques, however, make DRM a promising tool for microstructure analysis; with high potential to be implemented broadly in the academic community. Our DRM-based LRC-MRM method, for instance, generalizes well to different material systems and holds the promise to yield more objective grain segmentation of polycrystalline microstructures as compared to conventional methods based on optical microscopy.

Data availability

The raw data required to reproduce these findings are available to download from doi:10.17632/t4wvpy29fz.2.

Code availability

The code required to reproduce these findings are available to download from doi:10.17632/t4wvpy29fz.2.

Declaration of Competing Interest

None.

Acknowledgements

The authors would like to thank Le Tan Phuc and Wang Yulai for designing and manufacturing the motorized stage used for DRM measurements; Jude Emil Fronda for his support with Matlab, Ekta Jain for providing the DRM and EBSD datasets of the aluminium sample, Yong Chen Yeoh for providing the Inconel 718 sample, and Gan Li Meng for helping to prepare the titanium sample. This research was funded by the Ministry of Education of Singapore, Official Number: MOE2017-T2-2-119. Access to shared experimental facilities used for the experiments described in this work was provided by the School of Mechanical and Aerospace Engineering at NTU.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.matchar.2021.110978.

References

[1] K. Lu, L. Lu, S. Suresh, Strengthening materials by engineering coherent internal boundaries at the nanoscale, Science 324 (5925) (2009) 349–352, https://doi.org/10.1126/science.1159610.

[2] X. Wu, et al., Heterogeneous lamella structure unites ultrafine-grain strength with coarse-grain ductility, Proc. Natl. Acad. Sci. U. S. A. 112 (47) (2015) 14501–14505, https://doi.org/10.1073/pnas.1517193112.

[3] S. Suresh, R.O. Ritchie, Propagation of short fatigue cracks, Int. Met. Rev. 29 (1) (Jan. 1984) 445–475, https://doi.org/10.1179/imtr.1984.29.1.445.

[4] Y. Zhou, et al., Edge-weighted centroid voronoi tessellation with propagation of consistency constraint for 3D grain segmentation in microscopic superalloy images, in: IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work., 2014, pp. 258–265, https://doi.org/10.1109/CVPRW.2014.47.

[5] F. Bachmann, R. Hielscher, H. Schaeben, Grain detection from 2d and 3d EBSD data-specification of the MTEX algorithm, Ultramicroscopy 111 (12) (2011) 1720–1733, https://doi.org/10.1016/j.ultramic.2011.08.002.

[6] Y. Cao, L. Ju, Q. Zou, C. Qu, S. Wang, A multichannel edge-weighted centroidal voronoi tessellation algorithm for 3D super-alloy image segmentation, in: Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., no. May, 2011, pp. 17–24, https://doi.org/10.1109/CVPR.2011.5995590.

[7] Y. Cao, L. Ju, S. Wang, Grain segmentation of 3D superalloy images using multichannel EWCVT under human annotation constraints, in: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 7574 LNCS, no. PART 3, 2012, pp. 244–257, https://doi.org/10.1007/978-3-642-33712-3_18.

[8] H. Chuang, et al., An automated segmentation for nickel-based superalloy, in: Computer (Long. Beach. Calif), School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana, 2008, pp. 2280–2283.

[9] H.M. Guo, L.C. Zhan, The TEM image segmentation based on ICM-MRF algorithm, J. Phys. Conf. Ser. 1087 (2) (2018), https://doi.org/10.1088/1742-6596/1087/2/022017.

[10] C. Tan, M.D.R. Kok, S.R. Daemi, D.J.L. Brett, P.R. Shearing, Three-dimensional image based modelling of transport parameters in lithium-sulfur batteries, Phys. Chem. Chem. Phys. 21 (8) (2019) 4145–4154, https://doi.org/10.1039/c8cp04763d.

[11] G. Petzow, Metallographic Etching, 2nd Edition: Techniques for Metallography, Ceramography, Plastography, ASM International, 1999.

[12] M. Seita, M. Wittwer, X. Wang, Optical Metallography of Fusion-Based Additively Manufactured Metals, Elsevier Ltd., 2020.

[13] S. Beucher, F. Meyer, Chapter 12 The Morphological Approach to Segmentation: The Watershed Transformation. Mathematical Morphology in Image Processing, 1st Edition, CRC Press, 1993.

[14] ASTM, Standard Test Methods for Determining Average Grain Size Using Semiautomatic and, Methods, Vol. 03, No. Reapproved 2015, 1997, pp. 1–24, https://doi.org/10.1520/E1382-97R15.Section.

[15] Y. Chen, J. Chen, A watershed segmentation algorithm based on ridge detection and rapid region merging, in: 2014 IEEE Int. Conf. Signal Process. Commun. Comput., ICSPCC, 2014, pp. 420–424, https://doi.org/10.1109/ICSPCC.2014.6986227.

[16] J. Liu, J. Chen, An improved iterative watershed according to ridge detection for segmentation of metallographic image, Am. J. Sci. Eng. 1 (1) (2012) 8–13.

[17] A. Campbell, P. Murray, E. Yakushina, S. Marshall, W. Ion, New methods for automatic quantification of microstructural features using digital image processing, Mater. Des. 141 (2018) 395–406, https://doi.org/10.1016/j.matdes.2017.12.049.

[18] G. Jakob, A. Rinnhofer, H. Bischof, W. Benesova, Twin and scratch detection and removal in micrograph images of Inconel 718, Mach. Vis. Appl. Ind. Insp. XIV 6070 (2006) 60700L, https://doi.org/10.1117/12.642635.

[19] O. Dengiz, A.E. Smith, I. Nettleship, Grain boundary detection in microstructure images using computational intelligence, Comput. Ind. 56 (8–9) (Dec. 2005) 854–866, https://doi.org/10.1016/j.compind.2005.05.012.

[20] F. Ajioka, Z.-L. Wang, T. Ogawa, Y. Adachi, Development of high accuracy segmentation model for microstructure of steel by deep learning, ISIJ Int. 60 (5) (2020) 954–959, https://doi.org/10.2355/isijinternational.isijint-2019-568.

[21] M. Seita, et al., A high-throughput technique for determining grain boundary character non-destructively in microstructures with through-thickness grains, npj Comput. Mater. 2 (Jun. 2016), https://doi.org/10.1038/npjcompumats.2016.16.

[22] B. Gaskey, L. Hendl, X. Wang, M. Seita, Optical characterization of grain orientation in crystalline materials, Acta Mater. 194 (2020) 558–564, https://doi.org/10.1016/j.actamat.2020.05.027.

[23] J. Achard, et al., Identification of etch-pit crystallographic faces induced on diamond surface by H_2/O_2 etching plasma treatment, Phys. Status Solidi Appl. Mater. Sci. 206 (9) (2009) 1949–1954, https://doi.org/10.1002/pssa.200982210.

[24] G.F. Vander Voort, Macrostructure, in: Metallography: Principles and Practice, 1984.

[25] M. Seita, M.M. Nimerfroh, M.J. Demkowicz, Acquisition of partial grain orientation information using optical microscopy, Acta Mater. 123 (2017) 70–81, https://doi.org/10.1016/j.actamat.2016.10.021.

[26] H. Li, et al., Columnar to equiaxed transition in additively manufactured CoCrFeMnNi high entropy alloy, Mater. Des. 197 (2021) 109262, https://doi.org/10.1016/j.matdes.2020.109262.

[27] F.J. Humphreys, Quantitative metallography by electron backscattered diffraction, J. Microsc. 195 (3) (1999) 170–185, https://doi.org/10.1046/j.1365-2818.1999.00578.x.

[28] T. Manninen, et al., EBSD, Mater. Sci. 86 (1) (2011).

[29] ASTM E2627–-13, Standard Practice for Determining Average Grain Size Using Electron Backscatter Diffraction (EBSD) in Fully Recrystallized Polycrystalline Materials, Astm Int., 2019, pp. 1–5, https://doi.org/10.1520/E2627, vol. 03.01, no. Reapproved 2019.

[30] X. Wang, S. Gao, E. Jain, B. Gaskey, M. Seita, Measuring crystal orientation from etched surfaces via directional reflectance microscopy, J. Mater. Sci. 55 (25) (2020) 11669–11678, https://doi.org/10.1007/s10853-020-04734-z.

[31] P. Centre, Crystallographic Dissolution of High Purity Aluminium, No. May, 2007, pp. 1729–1748, https://doi.org/10.1098/rspa.2007.1846.

[32] E. Steinsland, T. Finstad, A. Hanneborg, Etch rates of (100), (111) and (110) single-crystal silicon in TMAH measured in situ by laser reflectance interferometry, Sensors Actuators A Phys. 86 (1–2) (2000) 73–80, https://doi.org/10.1016/S0924-4247(00)00309-5.

[33] Z.-Y. Zhang, Nonnegative matrix factorization: models, algorithms and applications, in: D.E. Holmes, L.C. Jain (Eds.), Data Mining: Foundations and Intelligent Paradigms: Volume 2: Statistical, Bayesian, Time Series and other Theoretical Aspects, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 99–134.

[34] H. Bao, G. Niu, M. Sugiyama, Classification from pairwise similarity and unlabeled data, in: 35th Int. Conf. Mach. Learn. ICML 2018, Vol. 2, 2018, pp. 763–781.

[35] S. Dan, H. Bao, M. Sugiyama, Learning from Noisy Similar and Dissimilar Data, Available: http://arxiv.org/abs/2002.00995, 2020.

[36] V. Bewick, L. Cheek, J. Ball, Statistics review 14: Logistic regression, No. March 2005, 2014, https://doi.org/10.1186/cc3045.

[37] http://www.umiacs.umd.edu/~hal/docs/daume04cg-bfgs.pdf, 2008. (Accessed 18 February 2021).

[38] G. Varoquaux, L. Buitinck, G. Louppe, O. Grisel, F. Pedregosa, A. Mueller, Scikit-learn, GetMobile Mob. Comput. Commun. 19 (1) (2015) 29–33, https://doi.org/10.1145/2786984.2786995.

[39] Z. Wang, T.A. Palmer, A.M. Beese, Effect of processing parameters on microstructure and tensile properties of austenitic stainless steel 304L made by directed energy deposition additive manufacturing, Acta Mater. 110 (2016) 226–235, https://doi.org/10.1016/j.actamat.2016.03.019.

[40] R. Liu, Z. Wang, T. Sparks, F. Liou, J. Newkirk, Aerospace Applications of Laser Additive Manufacturing, Elsevier Ltd, 2017.