

ARTICLE OPEN



Application of machine learning to assess the influence of microstructure on twin nucleation in Mg alloys

Biaobiao Yang^{1,2}, Valentin Vassilev-Galindo¹ and Javier Llorca^{1,2}✉

Twin nucleation in textured Mg alloys was studied by means of electron back-scattered diffraction in samples deformed in tension along different orientations in more than 3000 grains. In addition, 28 relevant parameters, categorized in four different groups (loading condition, grain shape, apparent Schmid factors, and grain boundary features) were also recorded for each grain. This information was used to train supervised machine learning classification models to analyze the influence of the microstructural features on the nucleation of extension twins in Mg alloys. It was found twin nucleation is favored in larger grains and in grains with high twinning Schmid factors, but also that twins may form in the grains with very low or even negative Schmid factors for twinning if they have at least one smaller neighboring grain and another one (or the same) that is more rigid. Moreover, twinning of small grains with high twinning Schmid factors is favored if they have low basal slip Schmid factors and have at least one neighboring grain with a high basal slip Schmid factor that will deform easily. These results reveal the role of many-body relationships, such as differences in stiffness and size between a given grain and its neighbors, to assess extension twin nucleation in grains unfavorably oriented for twinning.

npj Computational Materials (2024)10:26; <https://doi.org/10.1038/s41524-024-01212-x>

INTRODUCTION

Magnesium (Mg) and its alloys have emerged as promising candidates for structural components in transport, owing to their light weight, high strength-to-weight ratio, and good recyclability^{1,2}, as well as in biomedical applications due to their biodegradable and biocompatible behavior and have been recently used in scaffolds for bone tissue engineering^{3,4}. Mg has a hexagonal close-packed (hcp) crystal lattice and its plastic deformation is dominated by $\langle a \rangle$ basal slip, which presents a very low critical resolved shear stress (CRSS) (<1 MPa in pure Mg⁵). However, $\langle a \rangle$ basal slip can only accommodate deformation in the basal plane, and plastic strains along the $\langle c \rangle$ axis have to be accommodated through different mechanisms. The CRSS for $\langle c + a \rangle$ pyramidal slip on Mg is very high (98 MPa in pure Mg⁶) and, thus, plastic deformation along the $\langle c \rangle$ in Mg is often accommodated by twinning, a mechanism that involves the shearing of the crystal lattice at one side of the twin plane to mirror the atomic positions with respect to the other side of the twin plane. The twinning systems are defined by the twin plane and the twin direction, as well as by the shear deformation that is accommodated in the twin direction⁷. In the case of Mg hcp lattice, the $\{01\bar{1}2\} \langle 0\bar{1}11 \rangle$ extension twins are often nucleated during plastic deformation because of their low CRSS, as compared with $\{01\bar{1}1\} \langle 0\bar{1}12 \rangle$ compression twins^{8–10}. It should be noted, however, that twinning is a polar mechanism that only occurs when the shear deformation is applied in the appropriate direction, so the twinned region has a mirror symmetry with the parent region across the twin plane. Thus, extension twinning in Mg only occurs under stress states that lead to an extension of the $\langle c \rangle$ axis of the hcp Mg lattice. As a result, twinning deformation leads to a large difference between the tensile and the compressive yield strengths and work hardening of textured Mg alloys, and this marked plastic anisotropy has negative effects on the ductility and formability of wrought Mg alloys^{11,12}.

The accommodation of plastic deformation by twinning involves two successive steps: twin nucleation and twin growth (thickening). Twin nucleation is a heterogeneous process that takes place in regions with large stress concentrations in the microstructure, such as grain boundaries (GBs). Twin nucleation has been widely studied^{13–17} but a definitive theory is still lacking. Several models for twin nucleation were proposed in the past, including the pole mechanism of Thompson and Millard¹⁸, the slip dislocation dissociation mechanism of Mendelson¹⁹, and the disconnection mechanism proposed by Serra, Bacon and Pond^{20–23}. Later, a zonal-twinning mechanism based on atomistic simulations was proposed by Wang et al.^{24,25}, in which a stable twin nucleus was created by simultaneous nucleation of a partial dislocation with a Burgers vector of $-50/107 [10\bar{1}1]$ and multiple twinning dislocations with a Burgers vector of $1/15 [10\bar{1}1]$. In addition, Wang et al.²⁶ presented a pure-shuffle mechanism for twin nucleation in Mg at GB due to the stress concentration and the presence of GB dislocations. Besides, He et al.¹³ experimentally reported a dual-step mechanism for extension twin nucleation through in situ high-resolution transmission electron microscopy (HRTEM). The nucleation of extension twins was initiated by disconnections on the prismatic $|$ basal interfaces which establish the lattice correspondence of the twin with a minor deviation from the ideal orientation. Subsequently, the formation of coherent twin boundaries was achieved through the rearrangement of the disconnections at the prismatic $|$ basal interface¹³. Once the twin has been formed, it is generally accepted that twin thickening is mediated by the glide of twinning dislocations along the twin planes^{13,21,27–29}, and this process is controlled by the resolved shear stress on the twin plane and direction.

From the polycrystal viewpoint, the most common criterion used to explain twin nucleation in one grain is the apparent Schmid factor (SF), based on the hypothesis that the stress state in one grain is identical to the macroscopic applied stress^{30–32}. This

¹IMDEA Materials Institute, C/Eric Kandel 2, Getafe, 28906 Madrid, Spain. ²Department of Materials Science, Polytechnic University of Madrid/Universidad Politécnica de Madrid, E.T.S. de Ingenieros de Caminos, 28040 Madrid, Spain. ✉email: javier.llorca@upm.es

criterion is supported, for instance, by the micro-tensile tests in pure Mg single crystals by Ventura et al.¹⁵ who reported that the appearance of extension twins followed the SF criteria³³. Similar results were found in an AZ31 alloy deformed in compression along the extrusion direction³⁴ and in a hot rolled AZ31 alloy during successive in-plane compression tests along two different directions³⁵. However, many recent works have revealed that other microstructural features, besides the SF, also have remarkable influence on twin nucleation^{36–39}. Beyerlein et al. reported that the GB misorientation angle can affect twin nucleation in polycrystalline pure Mg, when it comes to twin pairs nucleated at a GB⁴⁰. Guan et al. revealed, for a WE43 Mg alloy, that the Luster-Morris geometric compatibility factor (m') between $\langle a \rangle$ basal slip of neighboring grains and the active extension twin variant plays a more critical role in twin nucleation than the apparent SF of the extension twin⁴¹. A similar conclusion was also drawn by Zhou et al. in their analysis of the extension twin variant selection in Mg-5Y (wt%) alloy by means of in situ electron back-scattered diffraction (EBSD)⁴². Koike et al.²⁵ found that the intragranular localized $\langle a \rangle$ basal slip was responsible for the formation of anomalous extension twins with low or even negative SFs in rolled AZ31 Mg alloy sheets deformed in tension along the rolling direction⁴³. Furthermore, extension twinning was also found to be very sensitive to the grain size of Mg alloys^{44,45}. Ghaderi and Barnett studied the effect of grain size on extension twinning in an extruded AZ31 Mg alloy and found that the macroscopic stress required for the activation of extension twins decreased as the grain size increased⁴⁶ and a similar behavior was observed by Dobroń et al.⁴⁷. Thus, there is still no consensus on the underlying factors leading to twin nucleation^{13–16} because extension twins not necessarily occur in all large grains, at all GBs, or in all grains with favorable orientations^{48,49} and it is important to ascertain the main microstructural features that lead to the nucleation of extension twins because of the relevance of this mechanism in the deformation and fracture of Mg alloys.

In this investigation, machine learning (ML) strategies (and, in particular, Bayesian inference) are used to establish the relationship between microstructural features and twin nucleation in two different Mg alloys. To this end, the microstructural features and the nucleation of twins was ascertained by means of 2D EBSD in more than 3000 grains, including 28 relevant parameters for each grain, categorized in four different groups (loading condition, grain shape, apparent SFs, and GB features). The information provided by 2D EBSD does not take into account the nucleation of twins nor the microstructure beneath the surface layer and this may induce some errors. However, the construction of a large dataset of 3D EBSD (as the one necessary for ML)⁵⁰ is an extremely challenging task and, thus, only 2D EBSD was used in this investigation. The information was used to train supervised ML classification models to analyze twinning. The combination of a large experimental dataset and the potential of ML tools allowed us to determine the most important microstructural features promoting twin nucleation. This work, therefore, provides results on the influence of the microstructural features on the nucleation of extension twins in Mg alloys. This information can help to design polycrystal microstructures with controlled twinning during deformation.

RESULTS

Microstructures

The development of deformation twins in Mg alloys is very sensitive to the microstructure^{50–52}. The microstructures of extruded AZ31 Mg and rolled Mg-1Al (at%) alloys before deformation are depicted in Fig. 1. Most grains in the inverse pole figure (IPF) of AZ31 alloy were colored in red because their $\langle c \rangle$ axis is parallel to the normal direction (ND, Fig. 1a). This agrees

well with the strong intensity of the (0002) pole figure around the ND in Fig. 1b, which is a 2D graphical representation of orientation showing the orientation of (0002) plane normal with respect to the sample reference frame. In contrast, most grains in the Mg-1Al alloy (Fig. 1e) were colored in green and blue, indicating that their $\langle c \rangle$ axis is perpendicular to the extrusion direction (ED). Accordingly, the (0002) pole diagram of the Mg-1Al alloy is near 90° away from the extrusion direction (Fig. 1f). This means that hot extrusion results in a strong prismatic texture where most of the grains have the $\langle c \rangle$ axis perpendicular to the ED, whereas hot rolling generates a strong basal texture where the $\langle c \rangle$ axis is oriented parallel to the ND of the rolled sheet.

The average grain size of AZ31 Mg and Mg-1Al alloys are $12.4 \pm 8.8 \mu\text{m}$ and $20.4 \pm 10.8 \mu\text{m}$, respectively (Fig. 1c, g) but both alloys present a wide size distribution that also includes a few large grains ($>40 \mu\text{m}$). Besides, the Mg-1Al alloy with a prismatic texture (mean: $40.6 \pm 19.8^\circ$) exhibits a higher GB misorientation angle than the AZ31 alloy with a basal texture (mean: $32.5 \pm 15.5^\circ$) (Fig. 1d, h). This is because the strong basal texture (with the $\langle c \rangle$ axis of many grains parallel to the ND) leads to misorientation angles in the range 0° to 30°. However, the GB misorientation angle varies from 0° to 90° for strong prismatic texture (with $\langle c \rangle$ axis of many grains perpendicular to the extrusion direction). The schematic illustration of the difference between two textures is shown in Supplementary Fig. 1 in the Supplementary Information (SI). Such a difference of GB misorientation angle can influence the localized twinning behavior⁴⁰. Overall, the different textures as well as the wide grain size distributions allowed us to collect a comprehensive dataset for twin nucleation.

Mechanical behavior

The tensile stress-strain curves of the AZ31 alloy in three different orientations (S0, parallel to the ND; S90, parallel to the transverse direction (TD); and S45 at 45° between the TD and ND) and of the Mg-1Al along the ED are plotted in Fig. 2. The sigmoidal shape and the parabolic shape of the curves are generally associated with twin-dominated and slip-dominated deformation, respectively^{53,54}. Thus, the nucleation and growth of extension twins control the deformation of sample S0, while slip should be the dominant deformation mode for samples S90 and Mg-1Al. The stress-strain curve of sample S45 is neither sigmoidal nor parabolic, suggesting that both slip and twinning may simultaneously contribute to the deformation. These hypotheses were corroborated by the experimental EBSD maps of AZ31 Mg and Mg-1Al alloys before and after deformation (shown in Supplementary Figs. 2 and 3 in the SI). In fact, after tensile deformation up to ~6%, the percentage of grains containing extension twins is 57%, 25%, 2.4%, and 15%, for samples S0, S45, S90, and Mg-1Al, respectively (Table 1). Note that most active extension twins exhibit low, or even negative SFs, in samples S90 and Mg-1Al (Table 1). This behavior is conventionally (and ambiguously) related to stress concentrations at GBs^{55,56} but its linkage with the microstructure has not been analyzed.

Database of microstructural features

The EBSD data of all our samples before deformation were exported using in-house codes based on MTEX (version 5.7.0), an open-source MATLAB toolbox^{57,58}, to collect the information of all microstructural features. Given that the microstructural features of the grains at the edge of the EBSD map were not fully captured, these grains were removed from the dataset. A total of 28 features were selected for each grain, which can be categorized into loading condition, grain shape parameters, apparent SFs, and GB parameters.

The loading condition is given by the *Orientation* feature, with values of 0, 45, and 90 for S0, S45, and S90, respectively, while all grains have a loading condition of 90 for the Mg-1Al alloy, as

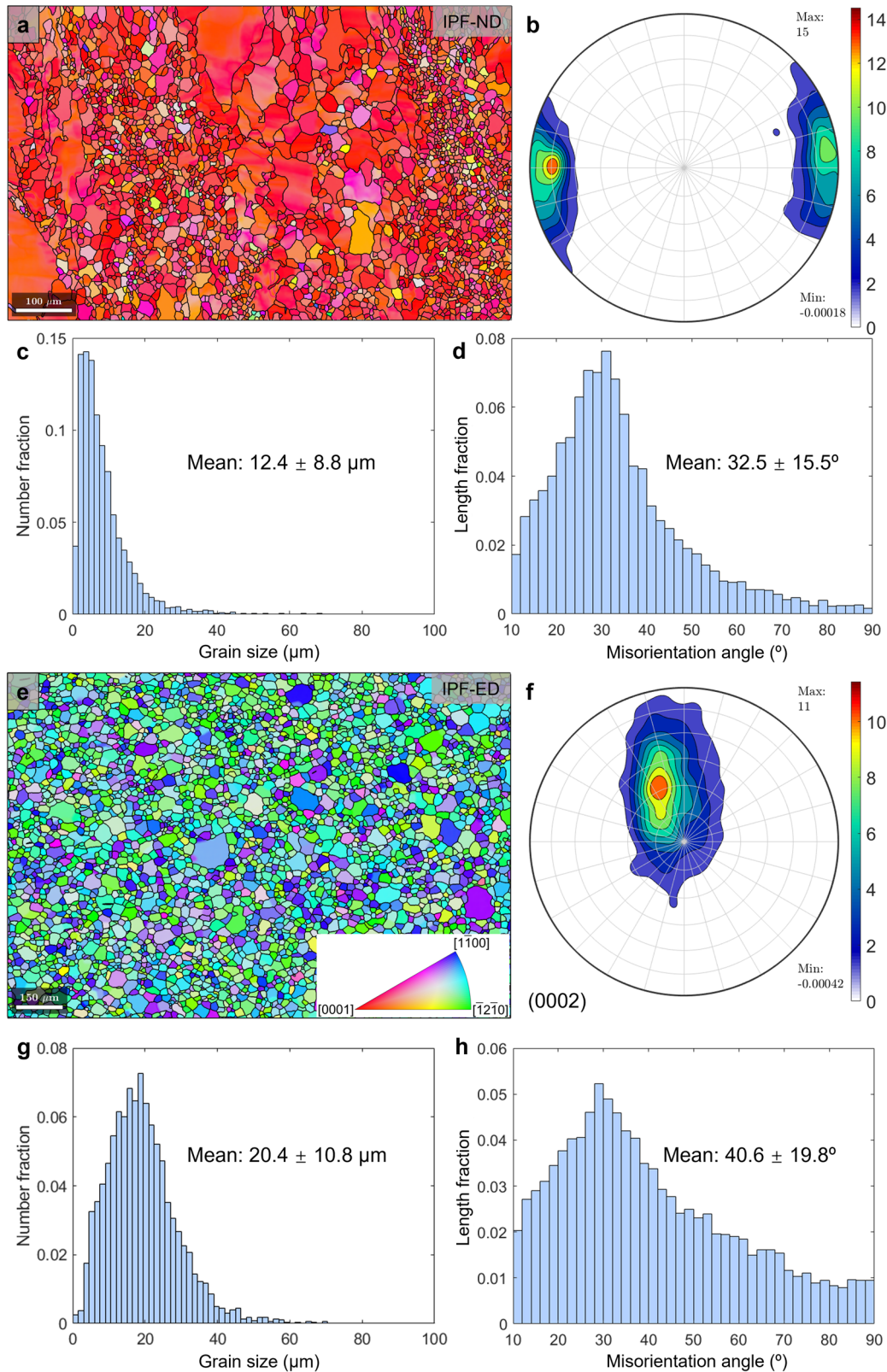


Fig. 1 Microstructures of AZ31 Mg and Mg-1Al alloys before deformation. EBSD inverse pole figure (IPF) maps in the (a) normal direction (ND) of AZ31 Mg alloy and (e) extrusion direction (ED) of Mg-1Al alloy. (0002) pole figures of (b) AZ31 Mg alloy and (f) Mg-1Al alloy. Distributions of (c, g) Grain size and (d, h) grain boundary misorientation angle of (c, d) AZ31 Mg alloy and (g, h) Mg-1Al alloy.

shown in Fig. 3a. The shape parameters of the grains include the diameter of a circle with the same area of the grain (*Grain_size*; in μm), the number of triple points (*Triple_points*), and the number of neighboring grains (*Neighbor_grain_n*) (Fig. 3b). The apparent SFs for the 6 possible extension twin variants and for the 3 $\langle a \rangle$ basal slip systems were considered for each grain, which are the dominant plastic deformation mechanisms during tensile deformation, as shown in grain reference orientation deviation (GROD) maps of Supplementary Fig. 4. The values of the Schmid factors were ordered from the highest to the lowest and named T_SF_n ($n = [1,6]$) and S_SF_n ($n = [1,3]$) for the extension twin variants and the $\langle a \rangle$ basal slip systems, respectively (Fig. 3c). It should be noted that the local SFs (that account for the local stress state) may play more decisive roles than the apparent SFs on the activation of slip and twinning in polycrystalline Mg alloys^{50,59}. However, the determination of local SFs for a large dataset (as the one necessary for ML) is extremely challenging and would require either costly diffraction experiments or 3D computational polycrystalline simulations based on the actual 3D grain structure^{50,60–62}. Considering that the main aim of this investigation is to assess the microstructural features that lead to the formation of extension twins, the apparent SF, which is a geometrical factor that takes into account interaction between the macroscopic stress and the grain orientation, is a good descriptor for ML.

Finally, the GB parameters were subdivided into (1) the Luster-Morris geometric compatibility factor (m'), (2) the GB misorientation (i.e., disorientation) angle (GB_misang ; in $^\circ$), (3) the difference of grain size (ΔGs ; in μm), and (4) the difference between the $\langle a \rangle$ basal slip SF of a given grain and its neighbors (ΔBSF). The

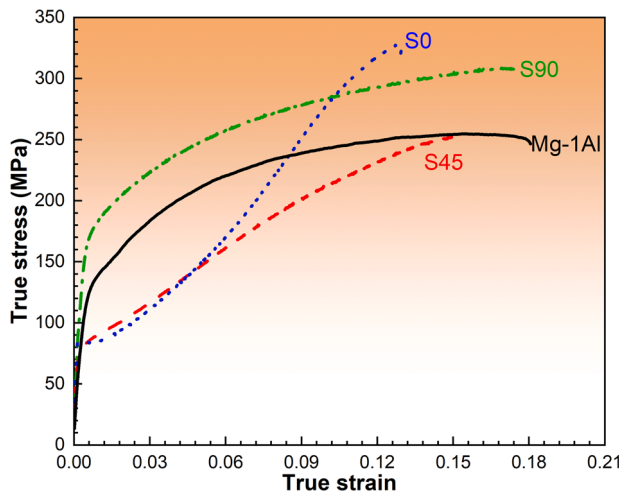


Fig. 2 Mechanical behavior of Mg alloys. True stress-true strain curves of samples S0, S45, S90 and Mg-1Al under tension.

maximum (*max*), the mean (*mean*), and the minimum (*min*) values of all GB features for each grain were included in the dataset.

The Luster-Morris geometric compatibility factor (m') is one of the most relevant criteria to assess slip/twin transfer as well as slip-induced twinning events at GB^{41,42,63}. It is based upon the angles between the active slip/twin plane normal directions ψ and the Burgers vector/twin shear directions κ according to⁵⁹:

$$m' = (\vec{n}_{in} \vec{n}_{out}) (\vec{d}_{in} \vec{d}_{out}) = \cos(\psi) \cdot \cos(\kappa) \quad (1)$$

in which \vec{n}_{in} , \vec{n}_{out} , \vec{d}_{in} , and \vec{d}_{out} stand for the vectors normal to the slip/twin plane and parallel to incoming and outgoing slip/twin directions, respectively (Supplementary Fig. 5). The m' between $\langle a \rangle$ basal slip systems of a given grain and its neighbors (B_b_m') and the m' between the 6 extension twin variants of a given grain and the $\langle a \rangle$ basal slip systems of its neighbors (B_t_m') were chosen as features, as schematically depicted in Fig. 3d. Hence, transmission of $\langle a \rangle$ basal slip across the GB and nucleation of extension twins at the GB induced by $\langle a \rangle$ basal slip in the neighbor grain are considered. It is worth noting that, although there are 3 possible $\langle a \rangle$ basal slip systems, only the one with the highest SF was considered to compute m' . However, all 6 extension twin variants were considered since the nucleation of extension twins induced by $\langle a \rangle$ basal slip in the neighbor grain is triggered by the stress concentration at the GB and all extension twin variants are possible. Even though extension twins in neighboring grains have also been reported to induce nucleation of extension twins^{30,60,61}, this feature was not considered because it is strongly correlated with the GB misorientation angle (GB_misang)^{40,62}.

The differences of grain size (ΔGs ; in μm) and $\langle a \rangle$ basal slip SF (ΔBSF) were calculated by subtracting the value of the feature of a given grain to the value of the feature for each of its neighbors. For ΔBSF , only the highest SF for the $\langle a \rangle$ basal slip systems (S_SF1) in the given grain and its neighbors was considered.

The mean grain orientations were used to calculate the geometric compatibility factors and the theoretical Schmid factors for twinning and slip. The slight deviations in the grain orientation, indicated by the intragranular grain orientation (mostly $<5^\circ$), may lead to small errors in the calculation of the GB characters.

The dataset of AZ31 Mg alloy includes 2301 entries, corresponding to 338 twinned grains and 1963 not twinned grains. The dataset of Mg-1Al alloy includes 811 entries, corresponding to 115 twinned grains and 696 not twinned grains. The variable *Twinned* in the dataset indicates whether a grain is twinned (1) or not twinned (0).

Presentation of some microstructural features

The distributions of some of the microstructural features used for descriptors in the dataset are plotted in Fig. 4. The distributions of

Table 1. Summary of the twinning features for various samples.

	Tensile strain (true)	All grains	Twinned grains	Twinned fraction	Theoretical twinning SF	Active twinning SF
S0	~6%	369	210	57%	0.40	0.40
S45		693	171	25%	0.22	0.23
S90		1578	38	2.4%	0.036	0.083
Mg-1Al		941	138	15%	-0.0096	-0.062
	9.5%	555	78	14%	-0.014	-0.16

Note that the number of grains in the table is slightly larger than the number of grains in database used for ML because the grains on the edge of the EBSD region were not included in the ML database. The theoretical twinning SF was determined as the average for all grains of the highest SFs of all twin variants in each grain. For each grain, the average grain orientation obtained from the initial EBSD map was used to calculate the twinning SF.

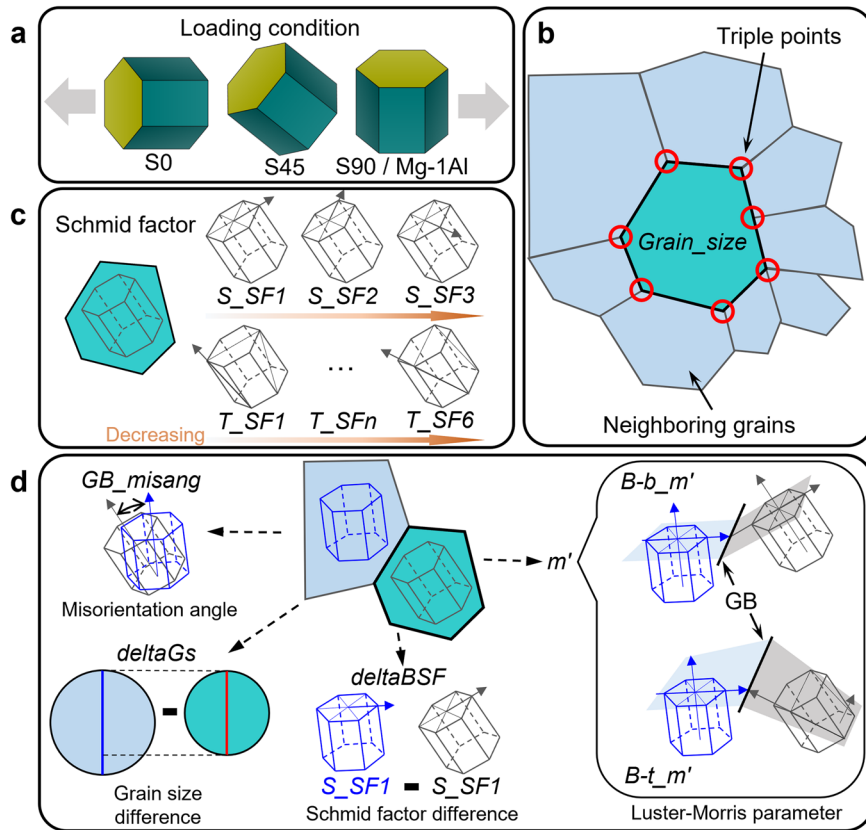


Fig. 3 Schematic of the microstructural features considered in the ML models. a Loading conditions for various samples, **b** grain shape parameters, **c** apparent SFs, and **d** grain boundary (GB) parameters.

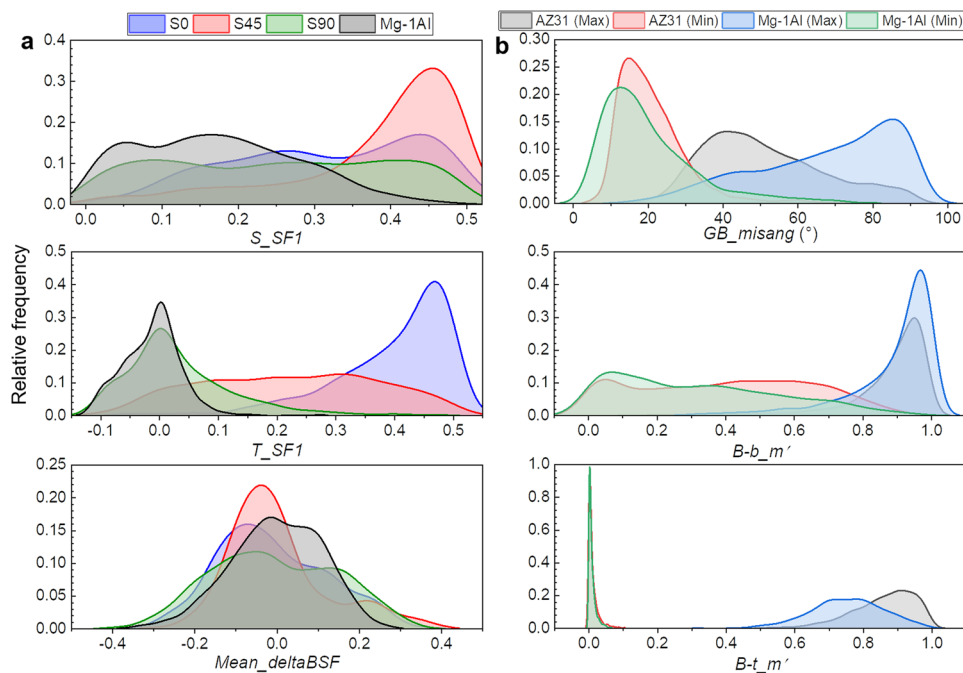


Fig. 4 Distributions of some microstructural features. a S_SF1 (upper), T_SF1 (middle), and $Mean_deltaBSF$ (bottom) values for different loading conditions of the AZ31 alloy and for the Mg-1Al alloy, and **b** Min/Max values of GB_misang (upper), $B-b_{m'}$ (middle) and $B-t_{m'}$ (bottom) for both alloys.

the maximum SF for $\langle a \rangle$ basal slip (S_{SF1}) and extension twinning (T_{SF1}) in Fig. 4a confirm that the three loading orientations in AZ31 lead to the activation of different dominant deformation mechanisms. For instance, the S45 sample exhibits the highest SF for $\langle a \rangle$ basal slip, while the highest SF is found for extension twinning in the S0 sample. Hard-oriented samples (S90 and Mg-1Al) show the lowest SFs for both $\langle a \rangle$ basal slip (although there are grains with high SFs) and extension twinning (Fig. 4a). These data are in good agreement with the experimental results in Fig. 2 and Supplementary Figs. 1 and 2.

Other features, such as *Mean_deltaBSF* (Fig. 4a Bottom), are, however, independent of the orientation, and are distributed symmetrically around 0 within a wide range from -0.3 to 0.2 . Other GB parameters, such as *GB_misang*, *B-b_m'*, and *B-t_m'* are also independent of the loading orientations and their distributions for the full AZ31 dataset (including the three orientations) are provided in Fig. 4b. The *GB_misang* feature for the Mg-1Al alloy is considerably higher than that of the AZ31 alloy (Fig. 4b upper), which could be ascribed to the different textures of both alloys (Fig. 1). Despite that, the distributions of *B-b_m'*, and *B-t_m'* are independent of the alloy. These distributions demonstrate the comprehensive sampling of different features in the database.

Machine learning for twin nucleation

With the goal of finding the apparent causal relations of twinning, the datasets were used to train ML models that can predict if a given grain will twin or not. This can be achieved through ML classification methods, a type of supervised ML approaches whose objective is the prediction of labels (in our case, “twinned” or “not twinned”). Several ML models (e.g., support vector machines, decision trees, random forests, AdaBoost, Gradient boosting, Bayesian networks or BNs, etc.) were initially trained on the AZ31 dataset to select the most suitable method for predicting twinning. As a preprocessing step before training, “MinMax” scaling procedure was used to scale all features in the dataset to values in a range between 0 and 1, and removed all highly correlated features (absolute values of Pearson coefficient ≥ 0.95) that led us to a final dataset containing 24 features. A stratified 10-fold cross-validation procedure was used for training to avoid possible biases (and overfitting). The area under the receiver operating characteristic curve (ROC AUC) score^{64–66} was used to evaluate the performance of our ML models since it is a good metric to select optimal models independently from the cost context or the class distribution.

The accuracy in terms of the ROC AUC score for the 6 best performing ML models as well as the individual accuracy on predicting if the grain twins or not is shown in Table 2. All models present a rather good overall prediction accuracy (over 0.8) that indicates that the dataset contains enough information to differentiate the grains that will twin from those that will not. However, there are differences in the accuracy when predicting twinned (or not twinned) grains between the ensemble methods (gradient boosting [GDB]⁶⁷, AdaBoost⁶⁸, and random forests⁶⁹) and the “Bayes-based” methods (naïve Bayes⁷⁰ and the BNs⁷¹). While the former show a contrast in accuracy between twinned (0.663 in average) and not twinned (0.964 in average) grains, the latter have a more balanced prediction accuracy of 0.850 and 0.866 in average for twinned and not twinned grains, respectively. This means that ensemble methods are prone to bias toward more populated classes in the datasets, compromising their potential to learn from the less populated classes. In our case, both datasets are significantly unbalanced, with twinned grains accounting only for $\sim 15\%$ of the total grains. Hence, the ensemble methods are not considered suitable to achieve our goal of predicting twinning and ascertain the microstructural factors responsible for it.

Table 2. Overall accuracy in terms of the ROC AUC score and individual (twinned and not twinned) accuracy for the five best performing ML methods for the AZ31 dataset.

Method	ROC AUC	Twinned accuracy	Not twinned accuracy
Bayesian networks	0.871	0.879	0.863
Gaussian naïve Bayes	0.851	0.834	0.868
GDB	0.825	0.683	0.966
Random forests	0.812	0.660	0.954
AdaBoost	0.807	0.645	0.972

The scores presented here are the mean over the 10 cross-validation tasks. The “individual” scores were obtained by calculating the ratio between correct predictions and total samples (i.e., a value of 0 would mean that all predictions were incorrect and a value of 1 that all predictions were correct). The default settings were used as implemented in the pyAgrum⁸⁴ and scikit-learn⁸⁵ Python packages for the Bayesian network and for all other methods, respectively. The highest accuracies are highlighted with bold font.

Regarding the Bayes-based models, BNs outperform naïve Bayes (Table 2). Compared to naïve Bayes, BNs provide both a higher overall accuracy (ROC AUC of 0.871) and a more reliable prediction of twinned grains (0.879), while keeping a similar accuracy when predicting not twinned grains (0.863). Moreover, BN models, as their name indicates, are constructed by building a network (a directed acyclic graph) from data, where nodes represent all features available in the dataset (including the target variable) and the edges connecting the nodes indicate dependencies between features (see Fig. 5a, for an example of a BN for the AZ31 dataset). For instance, the BN in Fig. 5a shows that the model is able to learn the connection between features belonging to the same category (size, SF, or angle features) directly from data, without any prior bias. In addition to a remarkable prediction accuracy, BNs can “explain” what the model is learning. Hence, they offer the appropriate framework to obtain insights into the most relevant features defining twin nucleation. Henceforward, we will only focus on training BN models.

The next step to find the best possible model to describe twinning is to optimize the hyperparameters (i.e., those parameters that need to be fixed before training) of the BN model. A grid search cross-validation procedure was used to this end. The description of the optimized hyperparameters and their optimal values for each model are provided in the “Methods” section and in the Supplementary Methods, respectively. The prediction accuracy in terms of the ROC AUC score and the individual accuracy on predicting twinned and not twinned grains for different optimized BN models are shown in Table 3. The BN model for the AZ31 dataset (BN1 in Table 3) shows a slight improvement in accuracy with respect to the non-optimized one (ROC AUC score increased from 0.871 to 0.877). Figure 5a shows the BN obtained from model BN1. Focusing on the target variable (*Twinned*), there are only three nodes directly connected to it: *Grain_size*, T_{SF1} , and S_{SF1} . Such a group of directly connected nodes is the Markov blanket (MB) of the variable *Twinned*. A MB is a subset of all the available features in the dataset that alone contains all the useful information to infer the random variable to which the MB belongs (*Twinned*, in this case). This is confirmed after building a model for the AZ31 dataset using for training only the MB (BN2 in Table 3). The overall accuracy is kept (ROC AUC score of 0.878), and there are only small differences in the accuracy on predicting (not) twinned grains (from 0.885 to 0.893 and from 0.869 to 0.862 for twin and not twins, respectively).

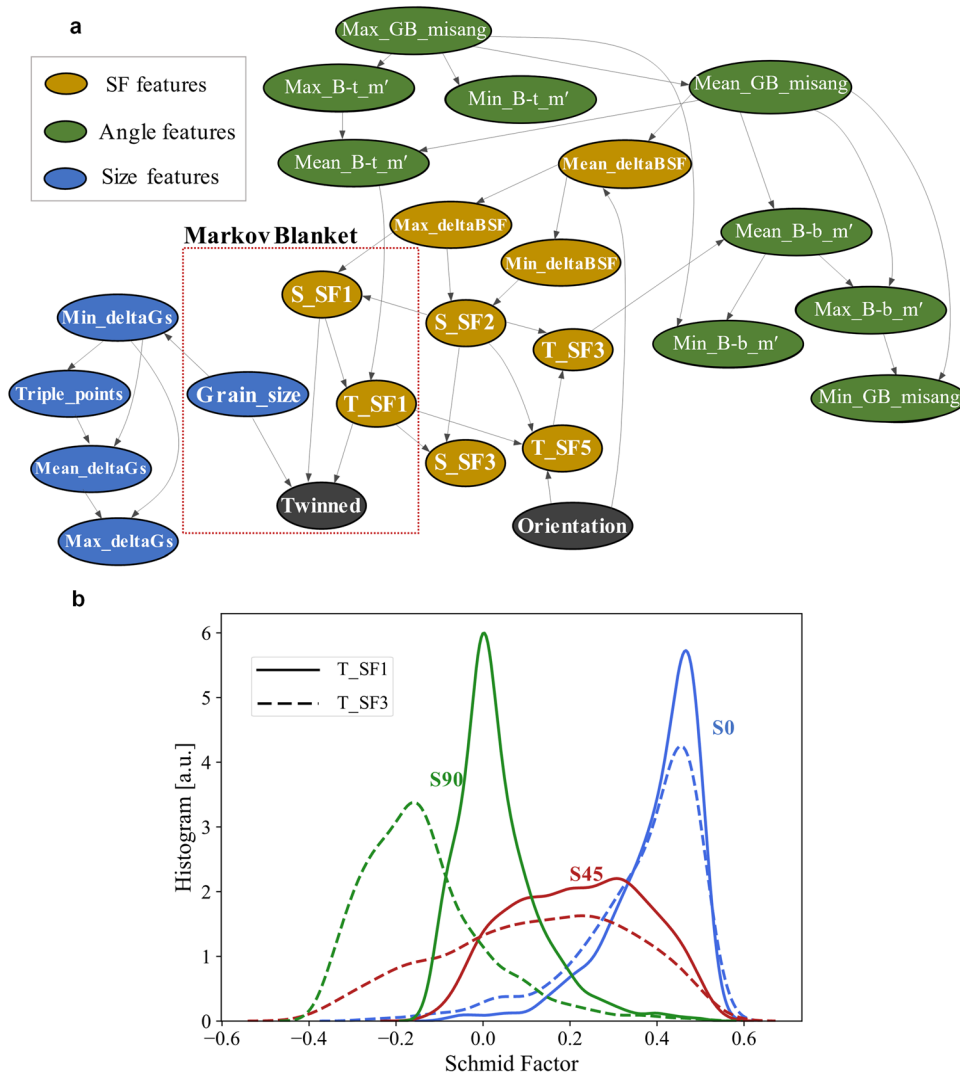


Fig. 5 Selection of the members of the Markov blanket (MB) for twinning. **a** Example of a Bayesian network for the AZ31 dataset with the Markov blanket for the target variable “Twinned” delimited with a dashed red square. Different colors for the nodes are used to indicate different types of features. **b** Distribution of values of the T_{SF1} and T_{SF3} features for the different loading conditions in the AZ31 dataset.

For the AZ31 dataset, the members of the MB of twinning can change depending on the training set. For example, T_{SF3} is sometimes selected instead of T_{SF1} . However, both features provide almost the same information since there is not much differences between the distribution of the former with respect to the one of the latter for the samples corresponding to S0 and S45 samples (see Fig. 5b). Moreover, the highest twinning SF (represented by the T_{SF1} feature) endows a more relevant physical meaning because it corresponds to the twin variant with the highest resolved shear stress, under the macroscopic loading. Therefore, only T_{SF1} was used to train MB-based models for the AZ31 dataset. Another possible variation in the MB is the choice of $Max_deltaBSF$ over S_{SF1} . For this case, training a BN including the former in the MB produces a model (BN3 in Table 3) with an accuracy equal to the BN2 model (ROC AUC score of 0.878 in the case of S_{SF1}).

So far, the accuracy of the models has been discussed. The next step is to analyze the insights they provide in describing twinning. Two features, $Grain_size$ and T_{SF1} , can be highlighted from the MB of the AZ31 dataset (Fig. 5a). They are also generally considered as the most important factors to predict twin nucleation from experimental observations^{33–35,44–47}. The decision

surface of a BN trained on the AZ31 dataset considering only the $Grain_size$ and T_{SF1} features is shown in Fig. 6a. Remarkably, the BN is learning from our dataset a well-known causal relation: a grain has a high probability of twinning if its size is rather large ($Grain_size > 7\ \mu\text{m}$) and its highest twinning SF has a high value ($T_{SF1} > 0.16$)^{33–35,44–47}. This conclusion is true for all samples, regardless of the orientation of the sample.

However, the interplay between the size of the grains and their twinning SFs accounts only for 97% of the correct “twinned” predictions of the BN2 (or BN3) model. In total, the BN2 (or BN3) model predicts correctly as “twinned” 302 samples out of 338, meaning that the third member of the MB (either S_{SF1} or $Max_deltaBSF$) helps in correctly classifying 9 twinned grains more than when using only $Grain_size$ and T_{SF1} . These additional correct predictions correspond to grains that have a grain size lower than $7\ \mu\text{m}$ and a very large twinning SF ($T_{SF1} > 0.46$) identified in Fig. 6a with an orange square. The scatter plots of all twinned samples of the AZ31 dataset considering all variables in the MBs are presented in Fig. 6b, one with S_{SF1} and another one with $Max_deltaBSF$. The analysis of the new correctly predicted samples (surrounded by an orange square like in Fig. 6a) shows that these samples possess both a very high value of

Table 3. Overall accuracy in terms of the ROC AUC score and individual (Twinned and Not twinned) accuracy for different BN models with optimized hyperparameters.

Model	Dataset	ROC AUC	Twinned accuracy	Not twinned accuracy
BN1	AZ31	0.877	0.885	0.869
BN2	AZ31 MB (S_{SF1})	0.878	0.893	0.862
BN3	AZ31 MB ($Max_deltaBSF$)	0.878	0.894	0.862
BN4	AZ31 [$T_{SF1} < 0.16$]	0.639	0.417	0.862
BN5	Mg-1Al	0.626	0.629	0.622
BN6	AZ31 [$T_{SF1} < 0.16$] MB	0.735	0.767	0.703
BN7	Mg-1Al MB	0.685	0.664	0.705
BN8	AZ31 [$T_{SF1} < 0.16$] MB ($Min_deltaGs$)	0.686	0.667	0.704
BN9	Mg-1Al MB ($Grain_size$)	0.676	0.652	0.701

The scores presented here are the mean over the 10 cross-validation tasks. The “individual” scores were obtained by calculating the ratio between correct predictions and total samples (i.e., a value of 0 would mean that all predictions were incorrect and a value of 1 that all predictions were correct). The models were trained with the BN implementation available in the pyAgrum Python package⁸⁴.

$Max_deltaBSF$ (i.e., neighboring grains have higher basal slip SFs) and a very small value of S_{SF1} . Specifically, $Max_deltaBSF > 0.22$ and $S_{SF1} < 0.22$. These results indicate that small grains with a high twinning SF still have a high probability to twin if they do not have favorable conditions to deform (e.g., low basal slip SFs), but also have at least one neighboring grain with high basal slip SFs that will deform easily because the CRSS for basal slip is very low.

Besides small-sized grains that twin, the decision surface in Fig. 6a also shows that there are some grains in the S45 and S90 samples that twin despite having low, even negative, twinning SFs ($T_{SF1} < 0.16$). Given that the BN trained on the full AZ31 dataset does not provide any clue on why these grains twin, a BN model was trained using only the samples in the AZ31 dataset with values of T_{SF1} lower than 0.16 (the value at the boundary in the decision surface) and another model was trained on the Mg-1Al dataset (for which almost all twinned grains have very low, and negative twinning SFs). The performance of the “reduced” AZ31 model (BN4 in Table 3) shows that, despite the number of twinned samples is very small (22) and the dataset is extremely unbalanced (1465 not twinned samples), the model is able to correctly predict almost half of the twinned grains (0.417) while achieving a good accuracy on the not twinned samples (0.862). As per the model trained on the Mg-1Al dataset (BN5 in Table 3), the overall accuracy is comparable to that of the BN4 model (0.626) but shows a more balanced prediction accuracy between twinned (0.629) and not twinned (0.622) samples than

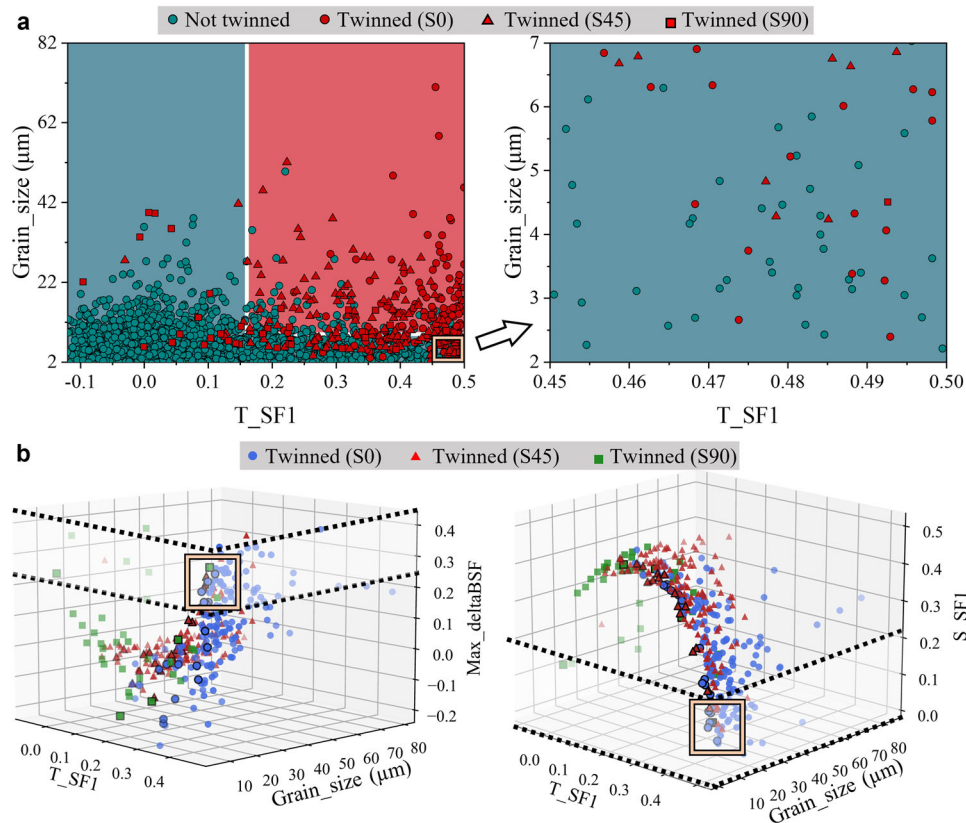


Fig. 6 Feature analysis of small-sized grains that twin. **a** Decision surface of a BN model trained on the AZ31 dataset using only the $Grain_size$ and T_{SF1} features. Blue and red colors identify the zones of the xy plane where the model will predict “Not twinned” or “Twinned”, respectively. **b** Scatter plots of all twinned grains in the AZ31 dataset considering all variables in the MBs. The dashed lines delimit the range of values of the third member of the MB ($Max_deltaBSF$ or S_{SF1}) within which the probability of twinning is high for small grains ($Grain_size < 7 \mu m$) with very high twinning SF ($T_{SF1} > 0.46$). The orange squares surround the small-sized grains that are correctly predicted as “twinned” by the BN model when considering for training the three features in the Markov Blanket.

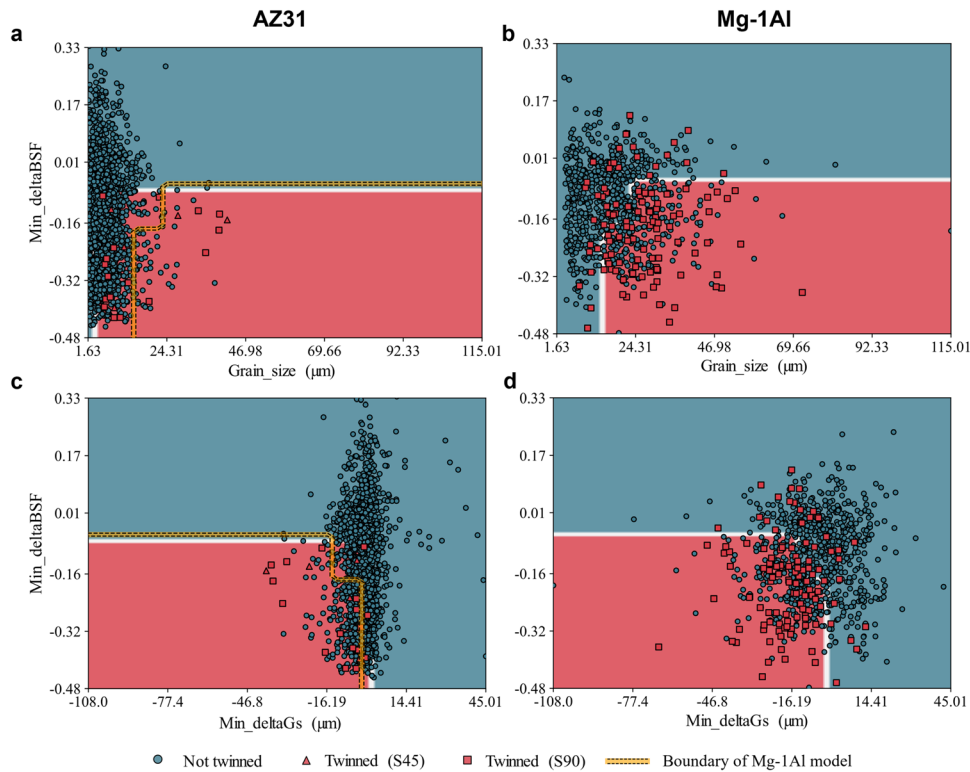


Fig. 7 Decision surfaces of Bayesian Networks models. Decision surfaces of different BN models trained on the (a, c) “reduced” AZ31 and (b, d) the Mg-1Al datasets. Models were trained using *Min_deltaBSF* and either (a, b) *Grain_size* or (c, d) *Min_deltaGs*. Blue and red colors identify the zones of the *xy* plane where the model will predict “Not twinned” or “Twinned”, respectively. For comparison, the boundary of the Mg-1Al models is drawn with a yellow line on top the decision surfaces of the AZ31 models.

the BN4 model. Apart from the comparable performances of both models, their MBs share similar information. Their MBs indicate that *Min_deltaBSF* and a size-related feature (*Grain_size* for the AZ31 alloy and *Min_deltaGs* for the Mg-1Al alloy) are the most important features defining twinning in grains with low twinning SFs. Indeed, training BN models using only the features in the MBs leads to an increment of the prediction accuracy. For the AZ31 dataset (BN6 in Table 3), the ROC AUC score increases from 0.639 to 0.735, whereas the model for the Mg-1Al (BN7 in Table 3) presents an increment from 0.626 to 0.685.

The similarities in the MBs suggest that the information both models are learning is, if not the same, very similar. In view of this, it was decided to exchange between the datasets the non-shared member of the MB and retrain the models (i.e., *Grain_size* was used for the Mg-1Al dataset and *Min_deltaGs* for the AZ31 dataset). As expected, the accuracy of the models dropped (see BN8 and BN9 models in Table 3), but it was possible to compare the BN models of both datasets. The decision surfaces are plotted in Fig. 7a, c for the “reduced” AZ31 models and in Fig. 7b, d for the Mg-1Al models. Focusing on the shared feature in the MB, all models set almost the same upper limit for *Min_deltaBSF* (around -0.06) for large grains (larger than $24\ \mu\text{m}$ in Fig. 7a, b, or at least $16\ \mu\text{m}$ larger than their smallest neighboring grain in Fig. 7c, d). This means that large grains that have at least one neighboring grain more rigid (i.e., lower basal slip SFs) than them have a high probability to twin. Note that there is a difference in the limits learnt for the *Min_deltaBSF* feature between the AZ31 and the Mg-1Al models for smaller grains. Namely, the latter indicates that grains between $15\ \mu\text{m}$ and $24\ \mu\text{m}$ (or that are $2\ \mu\text{m}$ to $16\ \mu\text{m}$ larger than their smallest neighboring grain) will be prone to twin only if at least one of their neighbors is far more rigid than them (*Min_deltaBSF* values lower than -0.16). Also, the Mg-1Al model suggests that grains

smaller than $15\ \mu\text{m}$ are very unlikely to twin regardless of the stiffness of their neighbors. Conversely, the “reduced” AZ31 model trained with its MB shows that twin nucleation in a grain is favorable for almost all grain sizes (*Grain_size* $> 2.4\ \mu\text{m}$) given that at least one neighbor is more rigid.

The differences between the AZ31 and Mg-1Al models are probably a consequence of the lack of data in the “reduced” AZ31 dataset (there were only 22 twinned samples compared to the 115 available in the Mg-1Al) and not a difference in the causal relations of twinning between the alloys. To prove this, the Mg-1Al models of Fig. 7b, d were used to predict the labels of the samples in the “reduced” AZ31 dataset. The boundary of the decision surface of the Mg-1Al models is also shown in Fig. 7a, c for an easy visual comparison between the models. While the models trained using *Grain_size* differ considerably in the region of very small grains (smaller than $15\ \mu\text{m}$), there is not much difference between the models trained with *Min_deltaGs* (the feature included in the MB of the Mg-1Al dataset). This is reflected in the accuracy of the Mg-1Al models in predicting the “reduced” AZ31 dataset. The Mg-1Al model trained using *Grain_size* achieves a ROC AUC score of 0.639 with a very unbalanced accuracy between twinned (0.318) and not twinned (0.960) grains. On the other hand, the Mg-1Al model trained using *Min_deltaGs* has an overall accuracy of 0.782 with a good accuracy for both twinned (0.818) and not twinned (0.747) grains. Such accuracies are even higher than the ones obtained by constructing any model from the “reduced” AZ31 dataset directly (compared to models BN4, BN6, and BN8 in Table 3). This suggests that *Min_deltaGs* is more crucial than *Grain_size* to define twin nucleation for grains with low twinning SFs. Therefore, it can be concluded from our BN models that twin nucleation in grains with low (even negative) twinning SFs is the consequence of many-body relationships, where one needs to consider not only the grain itself but also its neighbors. Namely, these grains will have a

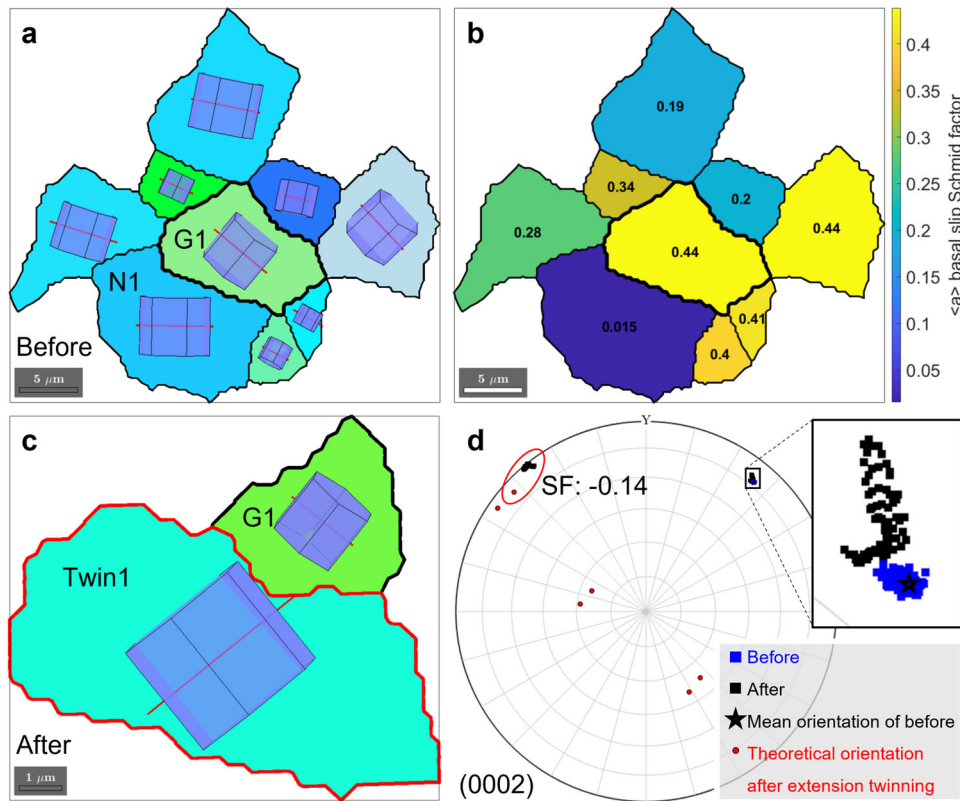


Fig. 8 Experimental evidence of twinning nucleation in grain with a low twinning SF in sample S90. **a** IPF-Z map of grain G1 and its neighboring grains before deformation. Based on the mean orientation for each grain, the crystal lattice as well as the basal plane trace were determined and overlaid in this figure as a red line. **b** $\langle a \rangle$ basal slip Schmid factor map of the same grains under the tension along the horizontal direction. **c** IPF-Z map of grain G1 and formed Twin1 after deformation. The rotation of the crystal lattice and basal plane trace reveals a possible twinning behavior. **d** Projections of the orientation of grain G1 and Twin1 on the (0002) pole figure before (blue) and after (black) deformation. Based on the mean orientation of grain G1 before deformation, the orientations of all six possible extension twin variants were obtained, with their projections added in the pole figure.

high probability of twinning if they have at least one smaller neighboring grain and another one (or the same) that is more rigid because its basal SF is very low.

The BN models suggested that the presence of a more rigid neighboring grain promotes twin nucleation in grains with unfavorable twinning SFs. To validate this conclusion, one of the grains that twinned with a low twinning SF was analyzed. The results of this analysis are presented in Fig. 8. Grain G1 with a size of 11 μm and Euler angles of (140°, 77.4°, 6.5°) has 8 neighbors (Fig. 8a). The SFs for $\langle a \rangle$ basal slip of the neighboring grains span a broad range from 0.015 to 0.44, while the maximum SF for $\langle a \rangle$ basal slip of G1 is equal to 0.44 (Fig. 8b). After deformation, one twin (Twin1) with Euler angles of (38.8°, 91.7°, 11.5°) nucleates inside G1 (Fig. 8c). According to the projection of Twin1 on the (0002) pole shown in Fig. 8d—and taking the mean orientation of G1 as reference—Twin1 is located near the projection of an extension twin variant with a SF of -0.14 . It is interesting to note that the deflection of G1 before and after deformation reveals that a slip deformation mode takes place at the same time that twinning (cf. Fig. 8d)⁷². More remarkable is, however, that Twin1 nucleates near the GB between G1 and the neighboring grain (N1), that has the lowest $\langle a \rangle$ basal slip SF (0.015) of all neighbors. This suggests that the nucleation of Twin1 tends to satisfy the strain compatibility between G1 and N1. Furthermore, the SFs for $\langle a \rangle$ basal slip of G1 and N1 lead to a Δ BSF equal to -0.425 (it is also the $\text{Min_}\Delta$ BSF value of G1), which is in full agreement with the criterion provided by the BN models.

DISCUSSION

The finding that twin nucleation is preferable in large grains with high twinning SF is in accordance with the current knowledge. For instance, Hong et al.⁷³ studied twin nucleation by post-mortem EBSD in hot-rolled AZ31 Mg alloy deformed in compression perpendicular to the $\langle c \rangle$ axis and in tension parallel to the $\langle c \rangle$ axis. They found that most active ET variants followed the SF criteria under both strain paths and similar results can be found in refs.^{30–33}. Moreover, GB are recognized as the preferred sites for twin nucleation as a result of the local stress concentrations to maintain compatibility during the deformation⁷⁴. Besides, Raeesinia and Agnew⁴⁴ carried out uniaxial tension and uniaxial compression experiments on a series of cast polycrystalline pure Mg and binary Mg-Zn alloys with various grain sizes, and the experimental results were subsequently analyzed using an elastic-viscoplastic self-consistent model. The larger the grain size of the sample, the smaller the CRSS to activate extension twinning. Choi et al.⁴⁵ studied the deformation behavior of a series of Mg alloys with different grain sizes ranging from 120 μm to 60 nm. Their results showed that extension twinning is gradually suppressed as the grain size decreases. It was hypothesized that the larger the grain size, the higher the dislocation accumulation caused by slip at the grain boundary, and the higher the stress concentrations near the grain boundary, leading to twin nucleation.

Furthermore, this study also reveals that extension twins may nucleate in grains with negative twinning SFs if the grain has at least one smaller neighboring grain and another one (or the same) that is more rigid. The condition that “at least one smaller

neighboring grain” indicates that the number of neighboring grains may be larger, providing more GBs with the optimum conditions to nucleate twins. As a result, *Min_deltaGs* is more important than *Grain_size* to define twin nucleation for grains with low twinning SFs (Fig. 7). In addition, the condition that “at least one more rigid neighboring grain” points out to the development of a stress concentration at the GB to maintain the compatibility between the grain and its hard neighbor, which should be relieved by the activation of twinning, a hard slip system (prismatic or pyramidal) or even cracking^{75–80}. Koike et al.⁴³ reported that anomalous extension twins with negative twinning Schmid factors are formed to minimize strain incompatibility caused by the greater activity of basal slip in the twinned grain than in the surrounding grains, in agreement with the ML predictions and the experimental results shown in Fig. 8.

In conclusion, twin nucleation of Mg alloys was investigated by the combination of large database (over 3000 grains \times 28 features) obtained from in situ EBSD and state-of-the-art machine learning tools. The Bayesian network (BN) models reveal that twin nucleation is favored in larger grains and in grains with high twinning Schmid factors, but also point out that twins may form in the grains with very low or even negative SFs for twinning (<0.16) if they have at least one smaller neighboring grain and another one (or the same) that is more rigid. Moreover, twinning of small grains with high twinning SFs is favored if they have low basal slip SFs and have at least one neighboring grain with a high basal slip SF that will deform easily. Very likely, twinning will be triggered in these small grains because it is the only way to maintain the deformation compatibility with soft neighbor grain. These results reveal that many-body relationships, such as differences in stiffness and size between a given grain and its neighbors, are crucial to assess extension twin nucleation in grains with characteristics commonly considered unfavorable for twinning (e.g., small grain size, low twinning SF).

Finally, the combination of the strategy presented in this work in combination with 3D microstructural characterization (to get information about the microstructure and of twinning beneath the surface) and of crystal plasticity simulations (to obtain information about the local SFs)^{81,82} is a promising path for future work to understand the physical mechanisms responsible for twin nucleation.

METHODS

Sample preparation

Slabs of $80 \times 65 \times 500$ mm³ of a rolled AZ31B-O Mg alloy were purchased from Magnesium Elektron Ltd. (Manchester, UK). The nominal chemical composition of the alloy is 2.89 wt% Al, 1.05 wt% Zn, and 0.42 wt% Mn. In addition, a Mg-1Al (at.%) alloy was prepared by casting, homogenization treatment (400 °C, 2 h), hot extrusion (temperature: 300 °C, extrusion ratio: 16:1, ram speed: ≈ 2 mm s⁻¹), followed by another heat treatment (400 °C, 2 h).

Samples for microstructural and mechanical characterization were cut into dog-bone shape via electro-discharge machining. The dimensions (length \times width \times thickness) of the central gauge of the specimens were $15 \times 5 \times 2.5$ mm³ (to obtain whole tensile stress-strain curves) and $10 \times 3 \times 1.5$ mm³ (to perform interrupted EBSD measurements), for both Mg-1Al and AZ31 Mg alloys. The longest dimensions of the AZ31 samples were parallel to the normal direction (ND; denominated S0), at 45° between transverse direction (TD) and ND (denominated S45), and to the TD (denominated S90). The sample surface was always perpendicular to the rolling direction (RD). In the case of Mg-1Al alloy, the longest dimension of the sample was parallel to the extrusion direction (ED). For EBSD measurement, the surfaces of all samples were manually ground on an abrasive SiC paper with a grit size of 3000, followed by four polishing steps with 3 μ m, 1 μ m, 0.25 μ m diamond paste and with a suspension containing oxide particles of 40 nm.

Mechanical tests

Tensile tests along the longest direction of each sample were conducted in an Instron model 8501 universal testing machine under displacement control at room temperature. Deformation was measured with an extensometer (Instron model 2620-602, gauge length: 12.5 mm) at an average strain rate of 10^{-3} s⁻¹ while the applied load was monitored with a load cell. The tension for EBSD measurement before and after deformation was carried out in the micromechanical testing machine (Kammrath and Weiss Technologies, Inc., Model MZ.Sb) under the displacement control at 1 μ m/s, which leads to an approximate strain rate of 1×10^{-4} s⁻¹.

EBSD observations

The sample surface was analyzed within the gage length before and after deformation using a scanning electron microscope (SEM, Apreo 25 LoVac, FEI Company, Portland, OR, USA; beam current: 2.7 nA, accelerating voltage: 20 kV) equipped with EBSD (Oxford HKL Channel 5, Oxford Instruments, Abingdon, UK; step size: 0.4 μ m for AZ31 Mg alloy and 0.5 μ m for Mg-1Al alloy, working distance: ~ 10 mm). The area of EBSD observation for the initial microstructure of AZ31 Mg and Mg-1Al alloys was 560×900 μ m² and 1130×1645 μ m², respectively. The step size of EBSD observation for the initial microstructure of both alloys was 1 μ m. The microstructures in terms of inverse pole figure (IPF) of samples S0, S45, S90 and Mg-1Al before and after deformation are shown in Supplementary Figs. 2 and 3 in the SI.

ML methods

The performance of several ML classification methods was tested in order to select the best method to predict twin nucleation, including nearest neighbors, logistic regression, support vector machines, decision trees, Gaussian processes, neural networks, random forests, extremely randomized trees, Adaboost, gradient boosting (GDB), naïve Bayes, local discriminant analysis, and Bayesian networks (BNs). All of them were used as implemented in the scikit-learn Python package⁸³, except from the BN models that were constructed as implemented in the pyAgrum Python package⁸⁴. However, only those methods that performed the best (i.e., the methods whose results are presented later in Table 2) are described here. Such methods are Adaboost⁶⁸, GDB⁶⁷, random forest⁶⁹, naïve Bayes⁷⁰, and BNs⁷¹.

The first three methods (Adaboost, GDB, and random forest) are part of a family of methods known as ensemble methods. Ensemble methods combine the predictions of different base estimators built with a given algorithm (e.g., a decision tree) with the goal of providing a more robust/more general model over the one provided by a single estimator. The difference between different ensemble methods relies on the strategy followed in order to combine the predictions of all base estimators. In general, these methods either take the average of the predictions of all estimators (averaging methods) or build the base estimators sequentially in order to reduce the bias of the combined estimator, for instance, by applying weights depending on how difficult is to correctly predict a given sample (boosting methods). The remaining two methods (Gaussian naïve Bayes and BNs) are based in Bayesian statistics. The main pillar of Bayesian statistics is Bayes' theorem:

$$P(y|x_1, \dots, x_n) = \frac{P(y)P(x_1, \dots, x_n|y)}{P(x_1, \dots, x_n)} \quad (2)$$

where y is the class (target) variable and vector x_1 through x_n correspond to the vector of features. $P(y)$ is known as the prior probability of the target variable y , which expresses the probability of having a given value of y before evidence is considered. $P(x_1, \dots, x_n|y)$ is the likelihood function that contains

the probability of having x_1, \dots, x_n given that y is true. $P(y|x_1, \dots, x_n)$ is the posterior probability, the probability of having y after taking the evidence x_1, \dots, x_n into account. Finally, $P(x_1, \dots, x_n)$ is the probability of the evidence. The objective of Bayes-based classifiers is to find the class that has the maximum posterior probability:

$$\hat{y} = \arg \max_y P(y, |, x_1, \dots, x_n) \quad (3)$$

The difference between different Bayes-based models depends on the adopted assumptions on probability distributions. For instance, naïve Bayes methods assume a “naïve” conditional independence between every pair of features given the value of the class variable:

$$P(x_1, \dots, x_n | y) = \prod_{i=1}^n P(x_i | y) \quad (4)$$

In the case of the Gaussian naïve Bayes method used in this work, the likelihood of the features is assumed to be Gaussian:

$$P(x_i | y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}} \quad (5)$$

where σ_y^2 and μ_y are the variance and mean of the continuous variable x_i computed by a maximum likelihood estimation for a given class y .

On the other hand, BNs assume conditional independence in the joint distribution following the Markov condition (i.e., every node in a BN is conditionally independent of its non-descendants, given its parents):

$$P(y, x_1, \dots, x_n) = P(y, |, \text{Parents}(y)) \prod_i P(x_i, |, \text{Parents}(x_i)) \quad (6)$$

where the class (target) variable y also forms part of the BN.

Preprocessing of ML models

Before constructing any ML model, two preprocessing steps were carried out: (1) feature scaling and (2) removal of highly correlated features. Feature scaling is important to avoid biases toward features having values with big magnitudes. For instance, an ML model could consider angles to be more important than twinning SFs solely because the former can be as large as 90° , while the latter cannot exceed a value of 0.5. The feature scaling procedure that we followed was a MinMax scaling setting all features in a range between 0 and 1. For a given feature vector \vec{X} and a sample i , the scaled value was obtained by:

$$X_i^{\text{scaled}} = \frac{X_i - \min(\vec{X})}{\max(\vec{X}) - \min(\vec{X})} \quad (7)$$

To assess the correlation between features we computed Pearson coefficients (r) for every pair of features in our dataset. Such coefficients were calculated with:

$$r = \frac{\sum(\vec{x} - m_x)(\vec{y} - m_y)}{\sqrt{\sum(\vec{x} - m_x)^2 \sum(\vec{y} - m_y)^2}} \quad (8)$$

in which \vec{x} and \vec{y} are the feature vectors whose correlation is being calculated, and m_x and m_y are the mean of vectors \vec{x} and \vec{y} , respectively. We removed one feature of all pairs whose absolute value of Pearson coefficient was above 0.95. Specifically, the features that we removed with this procedure were *Neighbor_grain_n* (neighboring grain numbers), *T_SF2*, *T_SF4*, and *T_SF6* (correlated with *Triple_points* (triple point numbers), *T_SF1*, *T_SF3*, and *T_SF5*, respectively).

Performance of ML models

The accuracy of our ML models was assessed by the area under the receiver operating characteristic curve (ROC AUC) score^{64–66}. The ROC curve is created by plotting the rate of true positive predictions against the false positive ones at various threshold settings (i.e., the threshold used to decide whether a prediction is positive or negative in a binary classification task). It has advantages over other evaluation measures since it decouples classifier performance from class skew and error costs. However, a ROC curve is a two-dimensional depiction of classifier performance and, as such, it does not provide a single scalar value to compare easily the expected performance of different classifiers. Hence, a common method to solve this issue is to calculate the AUC of the ROC curve. Since the AUC is always a portion within the area of the unit square defined by the rate of true positive and false positive predictions, its value will always be between 0 and 1. However, no realistic classifier will present a ROC AUC below 0.5 because this is a result random guessing would produce. The AUC performs very well and is often used when a general measure of predictiveness is desired⁶⁶.

Model selection

The final step in the ML model pipeline was the model selection. The objective at this step was to select the best performing model out of several models trained using both different hyperparameters as well as different training/test sets. The procedure to achieve this is cross-validation (CV). In general terms, CV is a resampling method that uses different portions of the data to train and test a model on different iterations. Among the different CV approaches, the stratified 10-fold CV was used because it ensures that each portion of the dataset approximately contains the same percentage of samples of each target class in the complete dataset. Here, 10-fold means that the dataset was split into 10 subsets and repeated the training and testing 10 times. At each of these iterations, 9 subsets were used for training and the remaining one for testing. Formally, the stratified 10-fold CV was not used directly to select a model, but for assessing the performance of different models as the average of all CV tasks. This allowed to avoid overfitting.

In addition to stratified CV, the robustness of our BN models was ensured by implementing a grid search CV procedure for selecting the best hyperparameters. The main idea behind grid search techniques is to find the optimal parameters that are not learnt from data by training models with different combinations taken from a grid of parameter values. The best model was the one that achieved the highest CV score (we used again as CV procedure a stratified 10-fold CV). Specifically, we optimized the number of discretization bins (“discretizationNbBins”), the discretization strategy (“discretizationStrategy”), the learning method (“learningMethod”), and whether to use or not the threshold of precision-recall curves to make predictions (“UsePR”). All other hyperparameters were kept with default values since initial tests showed that they remain unchanged after the grid search. The reader is referred to PyAgrum’s documentation for an explanation of all hyperparameters for training a BN model⁸⁴.

DATA AVAILABILITY

The datasets of microstructural features for both the AZ31 and Mg-1Al alloys can be found in Zenodo⁸⁵.

Received: 1 August 2023; Accepted: 18 January 2024;

Published online: 30 January 2024

REFERENCES

1. Dziubińska, A., Gontarz, A., Dziubiński, M. & Barszcz, M. The forming of magnesium alloy forgings for aircraft and automotive applications. *Adv. Sci. Technol. Res. J.* **10**, 158–168 (2016).

2. Masood Chaudry, U., Tekumalla, S., Gupta, M., Jun, T.-S. & Hamad, K. Designing highly ductile magnesium alloys: current status and future challenges. *Crit. Rev. Solid State Mater. Sci.* **47**, 194–281 (2022).
3. Zhao, D. et al. Current status on clinical applications of magnesium-based orthopaedic implants: a review from clinical translational perspective. *Biomaterials* **112**, 287–302 (2017).
4. Li, M. et al. Microstructure, mechanical properties, corrosion resistance and cytocompatibility of WE43 Mg alloy scaffolds fabricated by laser powder bed fusion for biomedical applications. *Mater. Sci. Eng. C* **119**, 111623 (2021).
5. Sánchez-Martín, R. et al. Measuring the critical resolved shear stresses in Mg alloys by instrumented nanoindentation. *Acta Mater.* **71**, 283–292 (2014).
6. Wang, J., Molina-Aldareguía, J. M. & Llorca, J. Effect of Al content on the critical resolved shear stress for twin nucleation and growth in Mg alloys. *Acta Mater.* **188**, 215–227 (2020).
7. Christian, J. W. & Mahajan, S. Deformation twinning. *Prog. Mater. Sci.* **39**, 1–157 (1995).
8. Barnett, M. R. Twinning and the ductility of magnesium alloys: part I: “Tension” twins. *Mater. Sci. Eng. A* **464**, 1–7 (2007).
9. Barnett, M. R. Twinning and the ductility of magnesium alloys: part II. “Contraction” twins. *Mater. Sci. Eng. A* **464**, 8–16 (2007).
10. Yang, B., Wang, J., Li, Y., Barnett, M. & Llorca, J. Suppressed transformation of compression twins to double twins in Mg by activation of <a> non-basal slip. *Scr. Mater.* **235**, 115620 (2023).
11. Ball, E. A. & Prangnell, P. B. Tensile-compressive yield asymmetries in high strength wrought magnesium alloys. *Scr. Metall. Mater.* **31**, 2 (1994).
12. Cepeda-Jiménez, C. M. & Pérez-Prado, M. T. Microplasticity-based rationalization of the room temperature yield asymmetry in conventional polycrystalline Mg alloys. *Acta Mater.* **108**, 304–316 (2016).
13. He, Y., Li, B., Wang, C. & Mao, S. X. Direct observation of dual-step twinning nucleation in hexagonal close-packed crystals. *Nat. Commun.* **11**, 2483 (2020).
14. Li, B. & Ma, E. Atomic shuffling dominated mechanism for deformation twinning in magnesium. *Phys. Rev. Lett.* **103**, 035503 (2009).
15. Pond, R. C., Hirth, J. P., Serra, A. & Bacon, D. J. Atomic displacements accompanying deformation twinning: shears and shuffles. *Mater. Res. Lett.* **4**, 185–190 (2016).
16. Jiang, L. et al. Visualization and validation of twin nucleation and early-stage growth in magnesium. *Nat. Commun.* **13**, 20 (2022).
17. Xin, R. et al. Understanding common grain boundary twins in Mg alloys by a composite Schmid factor. *Int. J. Plast.* **123**, 208–223 (2019).
18. Thompson, N. & Millard, D. J. Twin formation, in cadmium. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **43**, 422–440 (1952).
19. Mendelson, S. Dislocation dissociations in hcp metals. *J. Appl. Phys.* **41**, 1893–1910 (2003).
20. Serra, A., Bacon, D. J. & Pond, R. C. The crystallography and core structure of twinning dislocations in H.C.P. metals. *Acta Met.* **36**, 3183–3203 (1988).
21. Serra, A. & Bacon, D. J. A new model for 1012 twin growth in hcp metals. *Philos. Mag. A* **73**, 333–343 (1996).
22. Serra, A., Bacon, D. J. & Pond, R. C. Dislocations in interfaces in the h.c.p. metals—I. Defects formed by absorption of crystal dislocations. *Acta Mater.* **47**, 1425–1439 (1999).
23. Pond, R. C., Serra, A. & Bacon, D. J. Dislocations in interfaces in the h.c.p. metals—II. Mechanisms of defect mobility under stress. *Acta Mater.* **47**, 1441–1453 (1999).
24. Wang, J., Hirth, J. P. & Tomé, C. N. (1012) Twinning nucleation mechanisms in hexagonal-close-packed crystals. *Acta Mater.* **57**, 5521–5530 (2009).
25. Wang, J. et al. Nucleation of a (1012) twin in hexagonal close-packed crystals. *Scr. Mater.* **61**, 903–906 (2009).
26. Wang, J., Yadav, S. K., Hirth, J. P., Tomé, C. N. & Beyerlein, I. J. Pure-shuffle nucleation of deformation twins in hexagonal-close-packed metals. *Mater. Res. Lett.* **1**, 126–132 (2013).
27. Lloyd, J. T. A dislocation-based model for twin growth within and across grains. *Proc. R. Soc. Math. Phys. Eng. Sci.* **474**, 20170709 (2018).
28. Hu, Y. et al. Disconnection-mediated twin embryo growth in Mg. *Acta Mater.* **194**, 437–451 (2020).
29. Wang, F. et al. Dislocation induced twin growth and formation of basal stacking faults in {1012} twins in pure Mg. *Acta Mater.* **165**, 471–485 (2019).
30. Yang, B. et al. Quasi-in-situ study on {10-12} twinning-detwinning behavior of rolled Mg-Li alloy in two-step compression (RD)-compression (ND) process. *J. Magnes. Alloy* **10**, 2775–2787 (2022).
31. Godet, S., Jiang, L., Luo, A. A. & Jonas, J. J. Use of Schmid factors to select extension twin variants in extruded magnesium alloy tubes. *Scr. Mater.* **55**, 1055–1058 (2006).
32. Yang, B., Wang, J., Li, Y., Barnett, M. & Llorca, J. Deformation mechanisms of dual-textured Mg-6.5Zn alloy with limited tension-compression yield asymmetry. *Acta Mater.* **248**, 118766 (2023).
33. Della Ventura, N. M. et al. {10-12} twinning mechanism during in situ micro-tensile loading of pure Mg: role of basal slip and twin-twin interactions. *Mater. Des.* **197**, 109206 (2021).
34. Barnett, M. R., Ghaderi, A., Quinta da Fonseca, J. & Robson, J. D. Influence of orientation on twin nucleation and growth at low strains in a magnesium alloy. *Acta Mater.* **80**, 380–391 (2014).
35. Chapuis, A., Xin, Y., Zhou, X. & Liu, Q. {10-12} Twin variants selection mechanisms during twinning, re-twinning and detwinning. *Mater. Sci. Eng. A* **612**, 431–439 (2014).
36. Kumar, M. A. et al. Deformation twinning and grain partitioning in a hexagonal close-packed magnesium alloy. *Nat. Commun.* **9**, 4761 (2018).
37. Liu, C., Roters, F. & Raabe, D. Finite strain crystal plasticity-phase field modeling of twin, dislocation, and grain boundary interaction in hexagonal materials. *Acta Mater.* **242**, 118444 (2023).
38. Beyerlein, I. J. & Arul Kumar, M. The stochastic nature of deformation twinning: application to HCP materials. in *Handbook of Materials Modeling: Methods: Theory and Modeling* (eds Andreoni, W. & Yip, S.) 1–39 (Springer International Publishing, 2018).
39. Abdolvand, H. et al. On the deformation twinning of Mg AZ31B: a three-dimensional synchrotron X-ray diffraction experiment and crystal plasticity finite element model. *Int. J. Plast.* **70**, 77–97 (2015).
40. Beyerlein, I. J., Capolungo, L., Marshall, P. E., McCabe, R. J. & Tomé, C. N. Statistical analyses of deformation twinning in magnesium. *Philos. Mag.* **90**, 2161–2190 (2010).
41. Guan, D., Wynne, B., Gao, J., Huang, Y. & Rainforth, W. M. Basal slip mediated tension twin variant selection in magnesium WE43 alloy. *Acta Mater.* **170**, 1–14 (2019).
42. Zhou, B. et al. Revealing slip-induced extension twinning behaviors dominated by micro deformation in a magnesium alloy. *Int. J. Plast.* **128**, 102669 (2020).
43. Koike, J., Sato, Y. & Ando, D. Origin of the anomalous {10-12} twinning during tensile deformation of Mg alloy sheet. *Mater. Trans.* **49**, 2792–2800 (2008).
44. Raeesinia, B. & Agnew, S. R. Using polycrystal plasticity modeling to determine the effects of grain size and solid solution additions on individual deformation mechanisms in cast Mg alloys. *Scr. Mater.* **63**, 731–736 (2010).
45. Choi, H. J., Kim, Y., Shin, J. H. & Bae, D. H. Deformation behavior of magnesium in the grain size spectrum from nano- to micrometer. *Mater. Sci. Eng. A* **527**, 1565–1570 (2010).
46. Ghaderi, A. & Barnett, M. R. Sensitivity of deformation twinning to grain size in titanium and magnesium. *Acta Mater.* **59**, 7824–7839 (2011).
47. Dobroň, P. et al. Grain size effects on deformation twinning in an extruded magnesium alloy tested in compression. *Scr. Mater.* **65**, 424–427 (2011).
48. Arul Kumar, M., Capolungo, L., McCabe, R. J. & Tomé, C. N. Characterizing the role of adjoining twins at grain boundaries in hexagonal close packed materials. *Sci. Rep.* **9**, 3846 (2019).
49. Shi, Z.-Z. et al. On the selection of extension twin variants with low Schmid factors in a deformed Mg alloy. *Acta Mater.* **83**, 17–28 (2015).
50. Wu, M. et al. On the deformation behavior of heterogeneous microstructure and its effect on the mechanical properties of die cast AZ91D magnesium alloy. *J. Magnes. Alloy* **10**, 1981–1993 (2022).
51. Yang, B. et al. Underlying slip/twinning activities of Mg-xGd alloys investigated by modified lattice rotation analysis. *J. Magnes. Alloy* **11**, 998–1015 (2023).
52. Cho, J.-H. & Kang, S.-B. Deformation and recrystallization behaviors in magnesium alloys. in *Recent Developments in the Study of Recrystallization* (ed. Wilson, P.) (InTech, 2013).
53. Kelley, E. W. & Hosford, W. F. Plane-strain compression of magnesium and magnesium alloy crystals. *Trans. Metall. Soc. AIME* **242**, 5–13 (1968).
54. Shi, D. F., Ma, A., Pérez-Prado, M. T. & Cepeda-Jiménez, C. M. Activation of second-order pyramidal slip and other secondary mechanisms in solid solution Mg-Zn alloys and their effect on tensile ductility. *Acta Mater.* **244**, 118555 (2022).
55. Lou, C., Zhang, X. & Ren, Y. Non-Schmid-based {10-12} twinning behavior in polycrystalline magnesium alloy. *Mater. Charact.* **107**, 249–254 (2015).
56. Shi, D. F., Pérez-Prado, M. T. & Cepeda-Jiménez, C. M. Effect of solutes on strength and ductility of Mg alloys. *Acta Mater.* **180**, 218–230 (2019).
57. Hielscher, R. & Schaeben, H. A novel pole figure inversion method: specification of the MTEX algorithm. *J. Appl. Crystallogr.* **41**, 1024–1037 (2008).
58. Bachmann, F., Hielscher, R. & Schaeben, H. Texture analysis with MTEX—free and open source software toolbox. *Solid State Phenom.* **160**, 63–68 (2010).
59. Luster, J. & Morris, M. A. Compatibility of deformation in two-phase Ti-Al alloys: dependence on microstructure and orientation relationships. *Metall. Mater. Trans. A* **26**, 1745–1756 (1995).
60. Arul Kumar, M., Beyerlein, I. J., McCabe, R. J. & Tomé, C. N. Grain neighbour effects on twin transmission in hexagonal close-packed materials. *Nat. Commun.* **7**, 13826 (2016).

61. Kumar, M. A., McCabe, R. J., Taupin, V., Tomé, C. N. & Capolungo, L. Statistical characterization of twin transmission across grain boundaries in magnesium. *Mater. Charact.* **194**, 112457 (2022).
62. Song, X. et al. Quantitative prediction of grain boundary misorientation effect on twin transmission in hexagonal metals. *Mater. Des.* **192**, 108745 (2020).
63. Bieler, T. R., Alizadeh, R., Peña-Ortega, M. & Llorca, J. An analysis of (the lack of) slip transfer between near-cube oriented grains in pure Al. *Int. J. Plast.* **118**, 269–290 (2019).
64. Hanley, J. A. & McNeil, B. J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**, 29–36 (1982).
65. Bradley, A. P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit.* **30**, 1145–1159 (1997).
66. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **27**, 861–874 (2006).
67. Friedman, J. H. Stochastic gradient boosting. *Comput. Stat. Data Anal.* **38**, 367–378 (2002).
68. Freund, Y. & Schapire, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **55**, 119–139 (1997).
69. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
70. Zhang, H. The optimality of naive Bayes. In *Proceedings of 17th International Florida Artificial Intelligence Research Society Conference* 562–567 (2004).
71. Pearl, J. *Bayesian Networks: A Model of Self-activated Memory for Evidential Reasoning* (UCLA Computer Science Department, 1985).
72. Yang, B. et al. Identification of active slip systems in polycrystals by Slip Trace-Modified Lattice Rotation Analysis (ST-MLRA). *Scr. Mater.* **214**, 114648 (2022).
73. Hong, S.-G., Park, S. H. & Lee, C. S. Role of {10-12} twinning characteristics in the deformation behavior of a polycrystalline magnesium alloy. *Acta Mater.* **58**, 5873–5885 (2010).
74. Yamasaki, S., Matsuo, H., Morikawa, T. & Tanaka, M. Acquisition of microscopic and local stress-strain curves by combination of HR-EBSD and DIC methods. *Scr. Mater.* **235**, 115603 (2023).
75. Haouala, S., Alizadeh, R., Bieler, T. R., Segurado, J. & Llorca, J. Effect of slip transmission at grain boundaries in Al bicrystals. *Int. J. Plast.* **126**, 102600 (2020).
76. Benjamin Britton, T. & Wilkinson, A. J. Stress fields and geometrically necessary dislocation density distributions near the head of a blocked slip band. *Acta Mater.* **60**, 5773–5782 (2012).
77. Sangid, M. D., Maier, H. J. & Sehitoglu, H. The role of grain boundaries on fatigue crack initiation—an energy approach. *Int. J. Plast.* **27**, 801–821 (2011).
78. Andani, M. T., Lakshmanan, A., Sundararaghavan, V., Allison, J. & Misra, A. Quantitative study of the effect of grain boundary parameters on the slip system level Hall-Petch slope for basal slip system in Mg-4Al. *Acta Mater.* **200**, 148–161 (2020).
79. Guo, Y. et al. Dislocation density distribution at slip band-grain boundary intersections. *Acta Mater.* **182**, 172–183 (2020).
80. Jamali, A., Ma, A. & Llorca, J. Quantitative assessment of the microstructural factors controlling the fatigue crack initiation mechanisms in AZ31 Mg alloy. *Acta Mater.* **239**, 118263 (2022).
81. Zeng, X. et al. Three-dimensional study of grain scale tensile twinning activity in magnesium: a combination of microstructure characterization and mechanical modeling. *Acta Mater.* **255**, 119043 (2023).
82. Venkatraman, A., Mohan, S., Joseph, V. R., McDowell, D. L. & Kalidindi, S. R. A new framework for the assessment of model probabilities of the different crystal plasticity models for lamellar grains in α - β titanium alloys. *Model. Simul. Mater. Sci. Eng.* **31**, 044001 (2023).
83. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
84. Ducamp, G., Gonzales, C. & Wuillemin, P.-H. aGrUM/pyAgrum: a toolbox to build models and algorithms for Probabilistic Graphical Models in Python. *Proceedings of the 10th International Conference on Probabilistic Graphical Models* **138**, 609 (2020).
85. Yang, B., Vassilev-Galindo, V. & Llorca, J. Dataset used in the publication entitled ‘Application of machine learning to assess the influence of microstructure on twin nucleation in Mg alloys’. <https://doi.org/10.5281/zenodo.10225600> (2023).

ACKNOWLEDGEMENTS

This investigation was supported by the project (MAD2D-CM)-IMDEA Materials funded by Comunidad de Madrid and by the Recovery, Transformation and Resilience Plan and by NextGenerationEU from the European Union, and by the María de Maeztu seal of excellence from the Spanish Research Agency (CEX2018-000800-M). B.Y. wishes to express his gratitude for the support of the China Scholarship Council (202106370122).

AUTHOR CONTRIBUTIONS

B. Y. and J. LL. conceived the idea of this work, who was supervised by J. LL. B. Y. carried out the experimental work to acquire the dataset, while V. V. G. performed the data analysis by machine learning. The results were analyzed and discussed by all authors who also wrote the original draft of the manuscript and reviewed the final manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41524-024-01212-x>.

Correspondence and requests for materials should be addressed to Javier Llorca.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



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 license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024