Journal Pre-proof

Micropillar compression using discrete dislocation dynamics and machine learning

Jin Tao , Dean Wei , Junshi Yu , Qianhua Kan , Guozheng Kang , Xu Zhang

PII: S2095-0349(23)00055-7
DOI: https://doi.org/10.1016/j.taml.2023.100484
Reference: TAML 100484

To appear in: *Theoretical and Applied Mechanics Letters*

Received date: 28 September 2023
Revised date: 20 November 2023
Accepted date: 21 November 2023

Please cite this article as: Jin Tao , Dean Wei , Junshi Yu , Qianhua Kan , Guozheng Kang , Xu Zhang , Micropillar compression using discrete dislocation dynamics and machine learning, *Theoretical and Applied Mechanics Letters* (2023), doi: https://doi.org/10.1016/j.taml.2023.100484

Highlights：

- Enhanced machine learning is performed by incorporating easily activated dislocation features of single crystal copper micropillar.
- The stress-strain curve of single crystal copper micropillar is predicted by the enhanced machine learning method.
- Effective learning of pillar size and crystal orientation effects has identified the dominant contribution of the size effect through feature analysis.

Micropillar compression using discrete dislocation dynamics and machine learning

Jin Tao[a], Dean Wei[a,b], Junshi Yu[a], Qianhua Kan[a], Guozheng Kang[a], Xu Zhang[a, *]

[a] (*School of Mechanics and Aerospace Engineering, Southwest Jiaotong University, Chengdu 610031, China*)

[b] (*State Key Laboratory of Nonlinear Mechanics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China*)

Abstract Discrete dislocation dynamics (DDD) simulations reveal the evolution of dislocation structures and the interaction of dislocations. This study investigated the compression behavior of single-crystal copper micropillars using few-shot machine learning with data provided by DDD simulations. Two types of features are considered: external features comprising specimen size and loading orientation and internal features involving dislocation source length, Schmid factor, the orientation of the most easily activated dislocations and their distance from the free boundary. The yielding stress and stress-strain curves of single-crystal copper micropillar are predicted well by incorporating both external and internal features of the sample as separate or combined inputs. It is found that the Machine learning accuracy predictions for single-crystal micropillar compression can be improved by incorporating easily activated dislocation features with external features. However, the effect of easily activated dislocation on yielding is less important compared to the effects of specimen size and Schmid factor which includes information of orientation but becomes more evident in small-sized micropillars. Overall, incorporating internal features, especially the information of most easily activated dislocations, improves predictive capabilities across diverse sample sizes and orientations.

Keywords discrete dislocation dynamics simulations, machine learning, size effects, orientation effects, microstructural features

1. Introduction

Micro-nano materials have become increasingly demanded in various fields, such as

precision manufacturing, semiconductor devices, and aerospace. Dislocations are local irregularities in atomic arrangements within crystal materials and act as internal defects. At the micro-nano scale, materials show a unique phenomenon: the strength increases with decreasing size. This phenomenon relates to the formation and movement of dislocations.

Simulation methods such as Molecular Dynamics (MD) [1], Discrete Dislocation Dynamics (DDD) [2], and Crystal Plasticity Finite Element Method (CPFEM) [3] are widely used to study the dislocation behavior of crystals. Among these methods, the DDD based on dislocation elasticity theory enables direct simulation on the dislocation microstructure evolution during deformation. The three-dimensional (3D) DDD [2] discretizes the dislocation system within the crystal into a network of line segments connected by nodes. By integrating the interaction mechanisms between dislocations and other defects within the DDD framework, it can study the deformation mechanisms of grain boundary strengthening [4], twin boundary strengthening [5], and particle/precipitate strengthening [6].

The DDD captures the dislocation structure evolution process, but its calculation is relatively time-consuming, while the computational complexity of calculating dislocation interactions is O (n^2). Furthermore, the dislocation microstructure evolution at the micro-nano scale shows strong randomness, requiring many DDD simulations to explore the relationship between microstructure evolution and mechanical behavior. Therefore, with limited computing resources, it is crucial to use the DDD simulation results effectively to examine the correlation between microstructure parameters and mechanical behavior.

Machine learning (ML) [7] is a data analysis method that can map from input features to output labels, which can uncover hidden relationships between inputs and outputs[8, 9]. Materials informatics [10] integrates data analysis methods such as machine learning into traditional material research involving multiple scales. Large-scale simulation tools offer valuable guidance in practice, but they are more empirical-driven. Small-scale simulations capture the essence of phenomena better, but they face challenges of computational efficiency constraints when simulating complex problems. Machine learning serves as a bridge, connecting different scales. Through machine learning, Evinan et al. [11] developed an MD learning potential that combines the accuracy of quantum mechanics with the efficiency of classical molecular potentials. The ML can reveal the relationships between microscopic structural and material properties, facilitating

theoretical and simulation. For example, using artificial neural networks (ANNs), Yassar et al. [12] predicted the flow stress of micropillars with dislocation density and model size as inputs. Salmenjoki et al. [13] utilized a two-dimensional (2D) DDD and inputted dislocation field information to explore the mechanical behavior of materials with or without preloading. Steinberger et al. [14] classified specimen size based on dislocation microstructure features derived from 3D DDD simulations. Zhang et al. [15] identified and extracted dislocation microstructures from real electron backscatter diffraction images using convolutional neural networks (CNNs) for further post-processing. Zhou et al. [16] established connections between macroscopic models and microscopic DDD simulations using neural networks (NNs). Hiemer et al. [17] linked plasticity with dislocation characteristics using data analysis and studied the influence mechanism of strain rate, dislocation density, and the strengthening mechanism in FCC metals using ML. Bertin et al. [18] learned changes in the graph structure during the dislocation motion process using graph NNs, aiming to accelerate DDD simulation.

Materials show a localized deformation behavior and plastic flow mechanism at the microscale due to the low density of dislocations. Thus, the elastoplastic behavior of materials at the microscale relies on the activation and interactions of locally active dislocations. However, the collective feature of the dislocation field was considered only in previous ML studies. Without enough data, models fail to accurately identify and learn from the dislocation information. Moreover, using existing ML methods such as the CNNs to reduce the dimensionality of dislocation information and find significant dislocation patterns lacks a physical explanation and consumes considerable computational resources.

This work is based on the 3D-DDD simulation platform ParaDiS [19]. ML methods are applied to predict the yielding stress and stress-strain curve using the external feature information and the information of the most easily activated dislocations as features. This approach considers both size or orientation effect and dislocation characteristics to enhance the accuracy of predicting the yielding stress and stress-strain curve of materials.

The paper is organized as follows. The theoretical framework is explained in Section 2, including the DDD framework and ML model. The results and discussion are given in Section 3, including predictive performance, feature analysis, size and orientation effects. Finally, this work is summarized in Section 4.

2. Theoretical Framework

2.1 Discrete Dislocation Dynamics (DDD) Framework

In the framework of DDD, the force on the discretized dislocation node i, denoted as $\boldsymbol{F}_i$, is the negative gradient of the stored energy E, and can be expressed as [19]:

$$\boldsymbol{F}_i = -\frac{\partial E\left[\left(\boldsymbol{x}_i, \boldsymbol{b}_{jk}, \boldsymbol{T}^{\mathrm{S}}\right)\right]}{\partial \boldsymbol{x}_i} \tag{1}$$

where $\boldsymbol{x}_i$ is the spatial coordinates of node i, $\boldsymbol{b}_{jk}$ is the Burgers vector of the dislocation segment jk, and $\boldsymbol{T}^{\mathrm{S}}$ is the surface traction. The energy E depends on the positions of all nodes, the Burgers vectors of dislocation segments, and the externally applied surface traction $\boldsymbol{T}^{\mathrm{S}}$. It consists of two parts: stored energy associated with the long-range elastic distortion (E^{el}) and the local atomic configuration of the dislocation cores (E^{c}), represented as [19]:

$$E = E^{\mathrm{c}} + E^{\mathrm{el}} \tag{2}$$

Thus, the calculation of the force on the node $\boldsymbol{F}_i$ involves both the spatial force derivatives of elastic energies $\boldsymbol{F}_i^{\mathrm{el}}$ and the dislocation core F_i^{c}, given by [19]:

$$\boldsymbol{F}_i = \boldsymbol{F}_i^{\mathrm{c}} + \boldsymbol{F}_i^{\mathrm{el}} \tag{3}$$

The elastic force on the dislocation node i is contributed by all adjacent dislocation segments ij and is the sum $\boldsymbol{F}_i^{\mathrm{el}}=\sum_j \boldsymbol{F}_{ij}^{\mathrm{el}}$. In an elastic continuum, the elastic force $\boldsymbol{F}_{ij}^{\mathrm{el}}$ on the dislocation segment ij is composed of the externally applied force $\boldsymbol{F}_{ij}^{\mathrm{ext}}$, the self-force of the dislocation segment $\boldsymbol{F}_{ij}^{\mathrm{s}}$, and the interaction forces from all other dislocation segments. It can be written as [19]:

$$\boldsymbol{F}_{ij}^{\mathrm{el}} = \boldsymbol{F}_{ij}^{\mathrm{ext}} + \boldsymbol{F}_{ij}^{\mathrm{s}} + \sum_{k=1}^{N-1}\sum_{l=k+1}^{N} \boldsymbol{F}_{ij}^{kl} \tag{4}$$

where k and l are the dislocation nodes of other dislocation segments, N is the total number of dislocation nodes in the system, and $\boldsymbol{F}_{ij}^{kl}$ is the elastic interaction force between the dislocation segment kl and dislocation segment ij, which is the main source of computational cost in DDD simulation.

After calculating the forces on the dislocations, it is necessary to establish the relationship between the dislocation velocity v and the applied force, known as the mobility equation. Based on the MD simulation results [20], a nonlinear mobility equation for dislocations is established as

[21]:

$$v = v_{\mathrm{m}}\left(1-\exp^{-k\tau}\right) \tag{5}$$

where τ is the applied shear stress on the dislocation, v_{m} is the saturation velocity of the dislocation at subsonic speeds, and k is the slope of the exponent. Different types of dislocations have different saturation velocities and slopes. For screw dislocations, v_{m}=1455 m·s^{-1} and k = 9.24×10^{-9} Pa^{-1}, respectively; for edge dislocations, v_{m}=1529 m·s^{-1} and k = 8.58×10^{-9} Pa^{-1}, respectively. If the velocity of the dislocation smoothly transitions with the direction of the Burgers vector under the same applied shear stress, the mobility equation for mixed dislocations can be expressed as:

$$v = v_{\mathrm{s}}\cos^2\theta + v_{\mathrm{e}}\sin^2\theta \tag{6}$$

where θ is the angle between the Burgers vector and the dislocation line direction, v_{s} and v_{e} are the velocities of screw and edge dislocations.

Subsequently, the node positions are updated using a time integration algorithm, and topological changes of the dislocations are performed to simulate the process of dislocation interaction and slip under external loading [19].

The edge size d of the micropillar in this study ranges from 250 to 4000 nm, with a height ratio equal to 2. The initial structure of dislocations is simplified as Frank-Read (FR) dislocation sources and randomly distributed within the model. To simulate the experimentally observed stress-strain response of micropillar compression [22], the length of FR dislocation sources in DDD simulation, denoted as L_{FR}, is also linearly scaled with the specimen size, following the distribution L_{FR} = 90 (nm) + 0.05 d .

At the microscale, the stress-strain response exhibits dislocation size and orientation effects. For the same dislocation source length, the closer to the screw dislocation, the higher the shear stress; while the closer to the edge dislocation, the lower the shear stress. The shear stress can be expressed as [23]:

$$\tau = \frac{T}{br} = \frac{\mu b^2}{4\pi br(1-\nu)}\left\{\left[1-\frac{\nu}{2}(3-4\cos^2\theta)\right]\log\frac{L_{\mathrm{FR}}}{\rho}-1+\frac{\nu}{2}\right\} \tag{7}$$

where T represents the self-energy (line tension) of the dislocation line, r is dislocation line curvature, L_{FR} is the length of the FR dislocation source, ρ represents the truncation radius of the

dislocation core (usually as $0.5b$), μ represents the shear modulus, and ν is Poisson's ratio.

Rao et al. [24] simplified it as:

$$\tau = \frac{\alpha\mu b}{L_{\mathrm{FR}}} \log \frac{L_{\mathrm{FR}}}{b} \tag{8}$$

where α is related to the dislocation line orientation angle. When it is a screw dislocation source, α = 0.19, and when it is an edge dislocation source, α = 0.13.

Considering the influence of dislocations with different orientations on the tensile stress, the activation shear stress of the dislocation source related to the orientation and length is:

$$\tau(\cos\theta, L_{\mathrm{FR}}) = \frac{\mu b^2}{L_{\mathrm{FR}}} \left[\alpha_{\mathrm{s}} \cos^2\theta + \frac{\alpha_{\mathrm{e}}}{1-\nu}(1-\cos^2\theta) \right] \log \frac{L_{\mathrm{FR}}}{b} \tag{9}$$

where α_{s} and α_{e} are fitting parameters, with values of 0.26 and 0.13, respectively.

2.2 Machine learning (ML) method

The study on the collective behavior of massive dislocations was conducted using the 3D-DDD simulation to consider the plastic response and predict the yield stress. By adjusting the micropillar size (d), loading orientation (α), and dislocation density (ρ), the compression behaviors of kinds of samples can be obtained. It should be pointed out that due to the time-consuming computations of DDD simulations, a small dataset of the DDD simulations was conducted, as given in Table 1. For the convenience of subsequent machine learning analysis and discussion, sample size (α) and loading orientation (α) are defined as external features, while dislocation density (ρ) is an internal feature. In addition, considering the significant effect of the most easily activated dislocation sources on the initial yielding of materials at the micro and nano scales, the relevant features of the most easily activated dislocation sources have also been selected as additional internal features, including the length (L_{FR}), orientation (β^{FS}), and distance to the free surface (d^{FS}) of the most easily activated dislocation source, as well as the Schmid factor (m) of the slip system where it is located. The geometric descriptions of internal and external features are shown in Fig. 1a. Four groups of features were used as inputs for the dataset, as shown in Table 2.

The strain rate was uniformly set at 5000 s^{-1}. The yield stress was defined as the offset stress at a plastic strain of 0.2%. The simulation consisted of 140 sets, forming a small dataset. Figure 1b illustrates the data distribution characteristics, showing a clear size effect where the yield stress increases with decreasing the size. An apparent orientation effect was also observed in the specimens. An ML model that can capture both effects needs to be established to predict the yield

stress better.

Table 1. Dataset parameters and value ranges

Parameter	Symbol	Range
Micropillar size (nm)	d	250, 500, 1000, 2000, 4000
Loading axis orientation (°)	α	0, 15, 30, 45, 60, 75, 90
Initial Dislocation Density (10^{13} m^{-2})	ρ	1, 5

Table 2. Features included in four types of feature groups

Group number	Feature Group Type	Features
1	External features	d, α
2	Internal features (excluding L_{FR})	$d^{\mathrm{FS}}, \beta^{\mathrm{FS}}, m, \rho$
3	Internal features (including L_{FR})	$d^{\mathrm{FS}}, \beta^{\mathrm{FS}}, m, \rho, L_{\mathrm{FR}}$
4	External features+ Internal features	$d, \alpha, d^{\mathrm{FS}}, \beta^{\mathrm{FS}}, m, \rho, L_{\mathrm{FR}}$

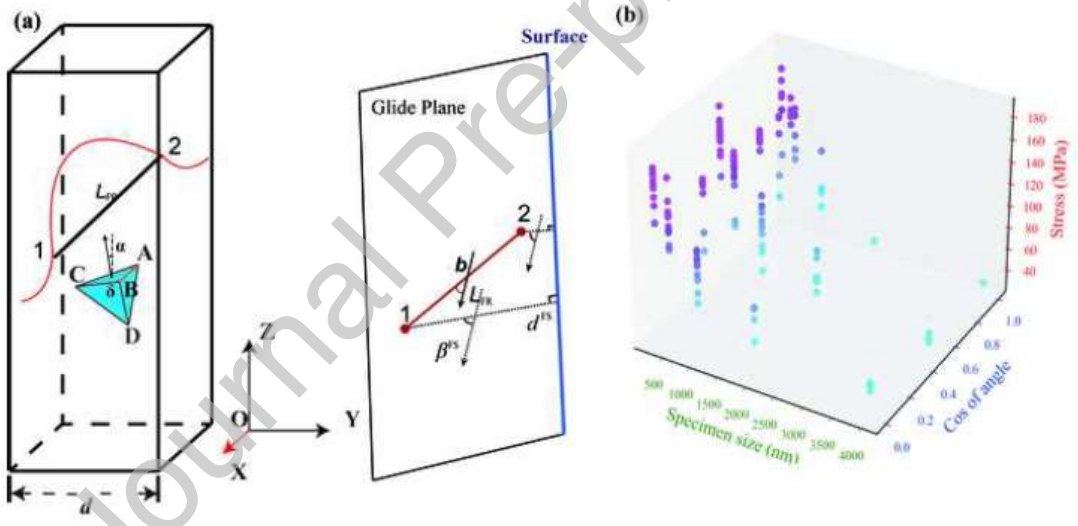


Fig. 1. (a) The external features (micropillars) and internal features (easily activatable dislocations) from DDD simulations (b) The orientation and size distributions of micropillars.

The data also show a kind of non-uniformity. Due to computational limitations, the number of data points for large specimens is significantly lower than that for small specimens. Thus, the adaptability of the model to the imbalanced dataset needs to be considered to reduce the impact of data imbalance on the prediction results. Several statistical ML methods, including linear regression [25], decision trees [26], random forests [27], and gradient boosting trees [28], were selected and compared for this imbalanced small dataset, and these algorithms were implemented

using Sklearn [29]. Moreover, feature selection was performed to remove features that have little or negative impact on the prediction results. The data was divided into training and testing sets in a 4:1 ratio. Data cleaning, normalization, and comparison of the predictive performance of different ML methods were conducted to select the model with the most stable predictive capability. Figure 2 illustrates the ML workflow. The ML process for predicting the yield stress includes a feature selection, data cleaning, data normalization, model selection, and error evaluation.

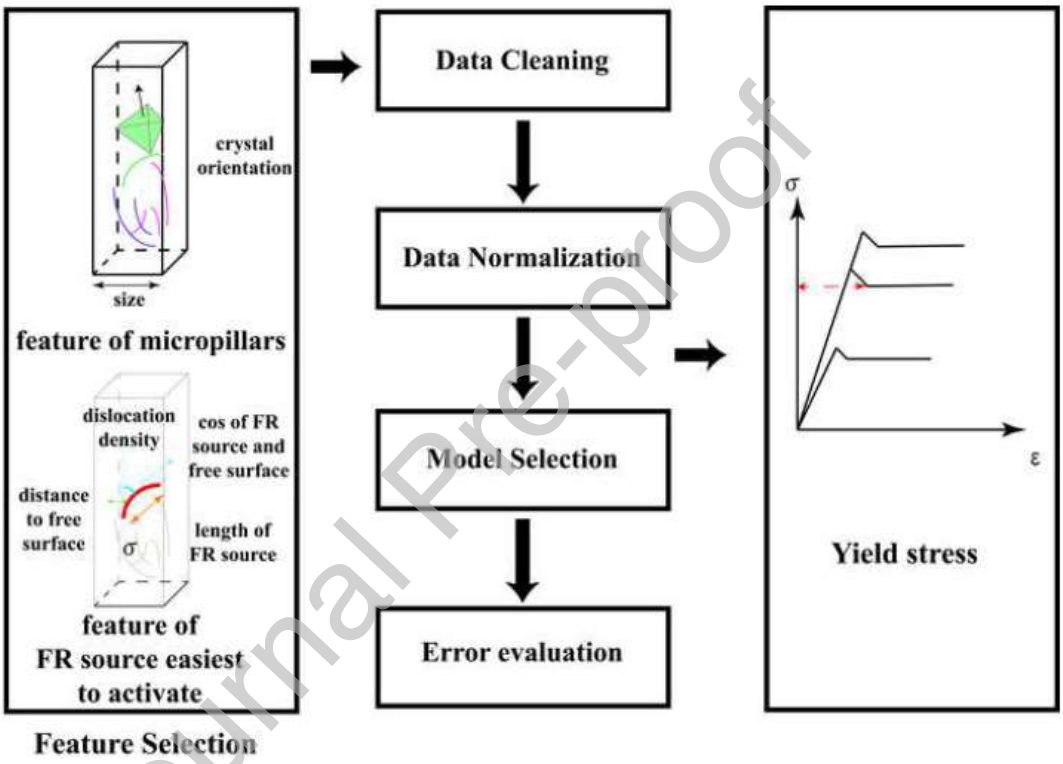


Fig. 2. ML workflow framework

The prediction results of ML were evaluated using the coefficient of determination (r^2), root mean square error (*RMSE*) and the standard deviation (*STD*) which are defined as:

$$r^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{10}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{11}$$

$$STD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{y}_i\right)^2} \tag{12}$$

where $\hat{y}_i$ and $\bar{y}_i$ represent the predicted values generated by the ML model and the mean stress obtained from the simulation, respectively. y_i is the actual observation of point *i*. *n* is the data point count. The learning performance is better when the r^2 is higher and *RMSE* and *STD* are smaller.

SHAP (Shapley Additive Explanations) [30] is a game-theoretic method that measures the marginal effects and interactions of individual features on the output of an ML model. It enhances the interpretability of ML models by calculating the contribution of each input variable to the output. SHAP assigns a score to each variable to indicate its importance to the output. The expression of the SHAP value is:

$$f\left(x\right) = \phi_0 + \sum_{i=1}^{M}\phi_i \tag{13}$$

where ϕ_0 is the output of the model when no factors are influencing it, whose value is $E(f(x))$. *M* is the number of features, and ϕ_i represents the Shapley value, which gives the impact of each feature on the output.

3. Results and discussion

3.1 Predictive performance

For the feature of internal + external (group 4), a random forest to the training set and obtained an ML model is applied. Figure 3a shows the learning performance on the training set, where the data points mainly near the $y = x$ axis. The coefficient of determination (r^2) and root mean square error (*RMSE*) for the training set are 0.96 and 24.71 MPa. For the test set shown in Fig. 3b, the r^2 and *RMSE* are 0.94 and 32.65 MPa, slightly lower than the training set but still capable of effectively predicting the yield stress. In both the training and test sets, the data points in different colors represent the samples with different sizes. It is shown that the model can capture the influence of sample size. The shaded region between the two dashed lines in the figure represents the confidence interval, which spans approximately 50 MPa (three times the standard deviation), and about 95% of the results are distributed in this range.

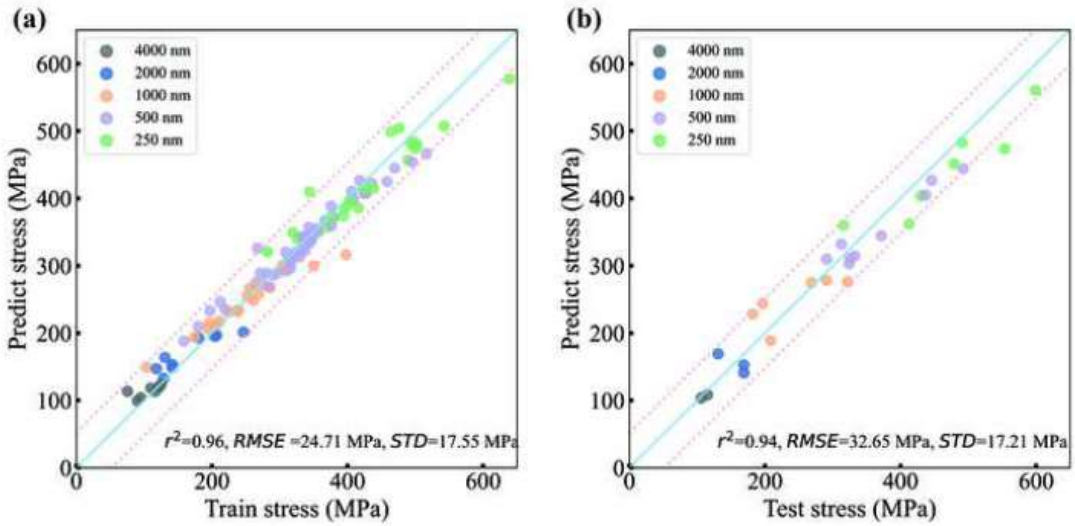


Fig. 3. Yield stress prediction performance of combined internal and external feature groups: (a) Training set; (b) Testing set. The *X*-axis represents the simulated stress, and the *Y*-axis represents the predicted stress.

The same ML model (random forest) is used to predict the yield stress without considering external features, but only internal features including the length of the FR dislocation source and the Schmid factor of easily activated dislocation source (group 3). These internal features contain size and orientation information. Figure 4 shows that the predicted distribution in group 3 for the test set is similar to the results from using both internal and external features in group 4, indicating a correlation between the internal and external features.

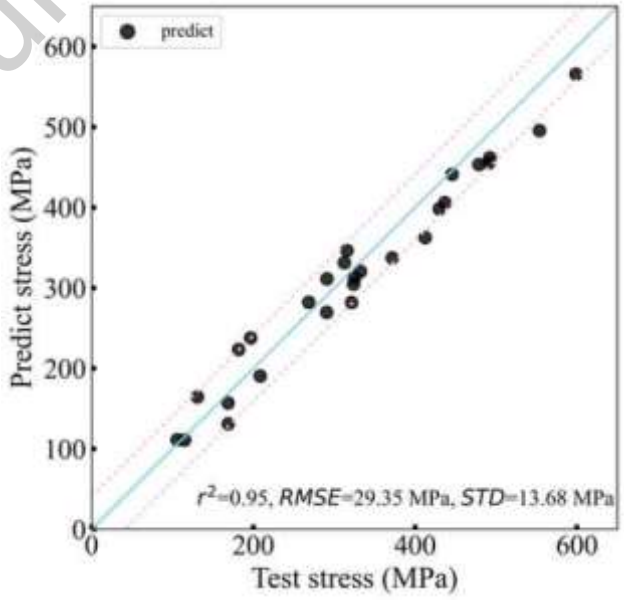


Fig. 4: Yield stress prediction of the testing set using internal feature group with dislocation source length

By calculating the activation stress of the most susceptible dislocation source in each micropillar, the critical shear stress is obtained, which is used to predict the yield stress of the micropillars. The prediction performance is shown in Fig. 5, which is compared to the ML prediction on the test set. Circular points represent the ML predictions, while star-shaped points represent the theoretical calculations. Compared to the theoretical calculations, the ML prediction distribution is closer to the centerline. The coefficient of determination (r^2) and root mean square error (*RMSE*) for the test set of theoretical calculations are 0.79 and 47.39 MPa. Predicting the loading stress at which the weakest dislocation source reaches the critical shear stress can provide a good estimate of yield stress. The ML model learns the relationship between microscale features and yield stress, predicting with greater accuracy based on relevant features of the most activated dislocation source.

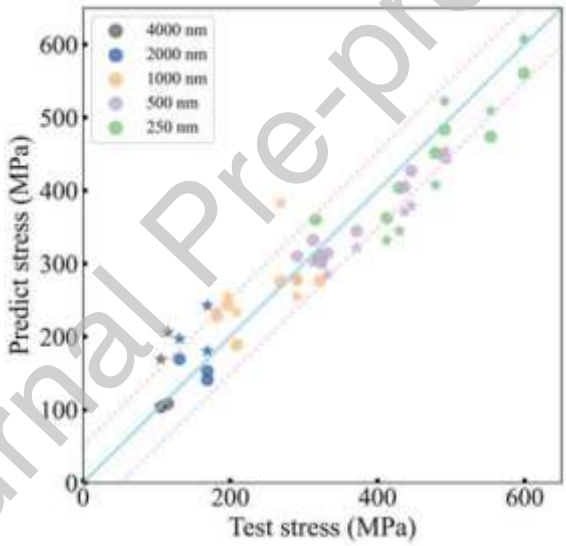


Fig. 5. Yield stress prediction of ML (circular points) and theoretical calculations (star-shaped points)

Moreover, the predictive performance is investigated on the stress-strain curves during the loading. Microstructure and stress are set as the features and labeled at every 0.01% plastic strain. Due to computational resource constraints, the simulation stops at the yield point (a plastic strain of 0.2%). Different-sized specimens yield at different strain levels, so the dataset with a unit length of 0.01% strain has unequal dataset lengths. Previous experimental and simulation studies [31-33] have shown that there is a plateau after yielding, with no significant additional strengthening, in the stress-strain curve of micropillars under uniaxial compression. To obtain an

aligned dataset, stress is assumed to remain unchanged after yielding. Then, 100 sampled points were performed by ML from 0% to 1% strain. Using the initial internal and external features, the model can predict the stress-strain curve well, as Fig. 6a shows. Figure 6b shows the trend of r^2 over time, where values below 0 are treated as 0. The r^2 initially decreases and then increases. It drops sharply when the strain value approaches 0.2%. The r^2 increases to around 0.9 and stabilizes when most samples enter the yielding state. The model can accurately predict the stress of the samples at the corresponding strain.

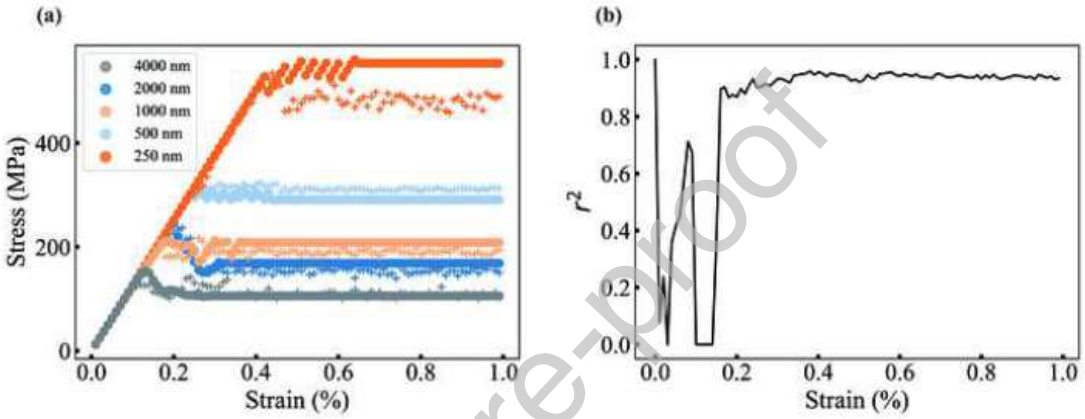


Fig. 6. The prediction to the elastoplastic behavior of the loading process: (a) Stress-strain curve, where circles represent actual values and crosses represent predicted values; (b) Variation of r^2 with strain

3.2 Feature Analysis

Using ML model, two types of features are included: external features and internal features. External features consist of specimen size and crystal orientation, which have the greatest impact on the yield stress and are the main features capturing the size effect and orientation effect in micropillars. Good prediction performance can be achieved solely using the external features (group 1), as shown in Table 3. This phenomenon indicates that the ML model can effectively capture the variations in yield stress with respect to size and orientation. However, during the dislocation activation, the initial structure of dislocations also influences the yield stress. Even under the same size and orientation, differences in dislocation distribution can lead to different yielding behaviors. To assess the influence of internal features on prediction performance, it is necessary to include as much relevant information as possible, such as easily activatable dislocations, as inputs. Since the FR dislocation source length is highly correlated with the

specimen size, the effect of the FR dislocation source length on the prediction performance is evaluated. The results show that without incorporating the FR dislocation source information (group 2), the prediction performance solely based on internal features is significantly lower than other feature combinations. This indicates that relying solely on the dislocation information for prediction, without considering the size-related features, makes it challenging to predict the yielding accurately. However, once the FR dislocation source length is included (group 3), the yielding can be well predicted solely based on the dislocation source information. Employing combined features of internal and external (group 4) and these complex models yields better prediction results than using only the external or internal features. It showes that the ML accuracy predictions for single-crystal micropillar compression can be improved by incorporating easily activated dislocation features with external features.

Table 3. r^2 of various ML methods on the testing set for each feature group

Machine learning methods	Group 1 External features	Group 2 Internal features (excluding L_{FR})	Group 3 Internal Features (including L_{FR})	Group 4 External Features+ Internal Features
Linear regression	0.755	0.588	0.747	0.779
Decision trees	0.893	0.628	0.558	0.516
Random forests	0.896	0.742	0.950	0.939
Gradient boosting trees	0.886	0.767	0.926	0.930

The number of dislocations grows exponentially as the sample size of the specimen increases. This can make it difficult to align the dislocation features at different sizes. In Section 2.2, the dislocation sources can be either weak or strong (weaker ones being closer to the edge dislocations and thus more prone to activation). Additionally, the influence of single-arm source effects is considered on activation, thus ranking the susceptibility of dislocation sources to activation. The influence of number of sampled dislocations is considered to study the effects of other dislocation source. Table 4 shows that a random forest regression model is used to study the effect of the number of sampled dislocations on the ML prediction. The sampling numbers are in the range from 1 to 4, with the features including the orientation and distance of dislocation from the free surface. The prediction accuracy initially rises and then declines as the number of sampled dislocations increases. The highest prediction accuracy occurs when the sampling number is 2.

However, more dislocation information can be collected to learn the overall dislocation characteristics of the sample if the dataset is large enough. In conclusion, selecting an appropriate number of dislocation features can enhance the accuracy of the ML model. In the limited data situations, it is necessary to balance the dislocation information quantity with the predictive capability of the model to achieve the best prediction performance.

Table 4. Variation of r^2 and *RMSE* value with the number of sampled dislocations

Number of dislocations	r^2	*RMSE* (MPa)
1	0.903	41.23
2	0.939	32.65
3	0.888	44.30
4	0.879	46.10

By analyzing the features using SHAP, Fig. 7a and 7c display the SHAP feature importance plots for the samples with internal + external features and the samples with internal features including FR dislocation source length, respectively. These plots show the average impact of features on the model's output value. The horizontal bars represent the average absolute SHAP value for each feature. The bars are ordered by their average absolute SHAP values from top to bottom. The features, which are introduced in Table 2, are ranked by importance. The FR dislocation source length and specimen size contain the size information, while the Schmid factor contains the orientation information. These two types of information affect the prediction most, followed by the most easily activated dislocation information, which has a much lower influence than the size and orientation. The dislocation density feature reflects the Taylor hardening effect, and its SHAP value is similar to the most easily activated dislocation information. This suggests that the Taylor hardening effect is much smaller than the size and orientation effects from dislocation activation in the investigated specimens. Moreover, Fig. 7c shows that when only FR dislocation source length is used, its average absolute SHAP value is about equal to the sum of that using only specimen size and FR dislocation source length. Both the feature sets show similar patterns of size effect. Simplified features may yield better learning results. Figures 7b and 7d show the SHAP summary plots, which are similar to the feature importance plots. The features are ordered by importance from top to bottom. Each sample has points in each feature row, where the horizontal position of a point shows its influence on the prediction result, and the color shows the

feature value (from green to red, from small to large). The plots show that the yield stress is negatively correlated with the size, Schmid factor, and cosine of the FR dislocation source orientation. Also, in Fig. 7b, the distribution of SHAP values for the FR dislocation source length and size is similar, indicating that the model gets similar information from both, which is the specimen size effect.

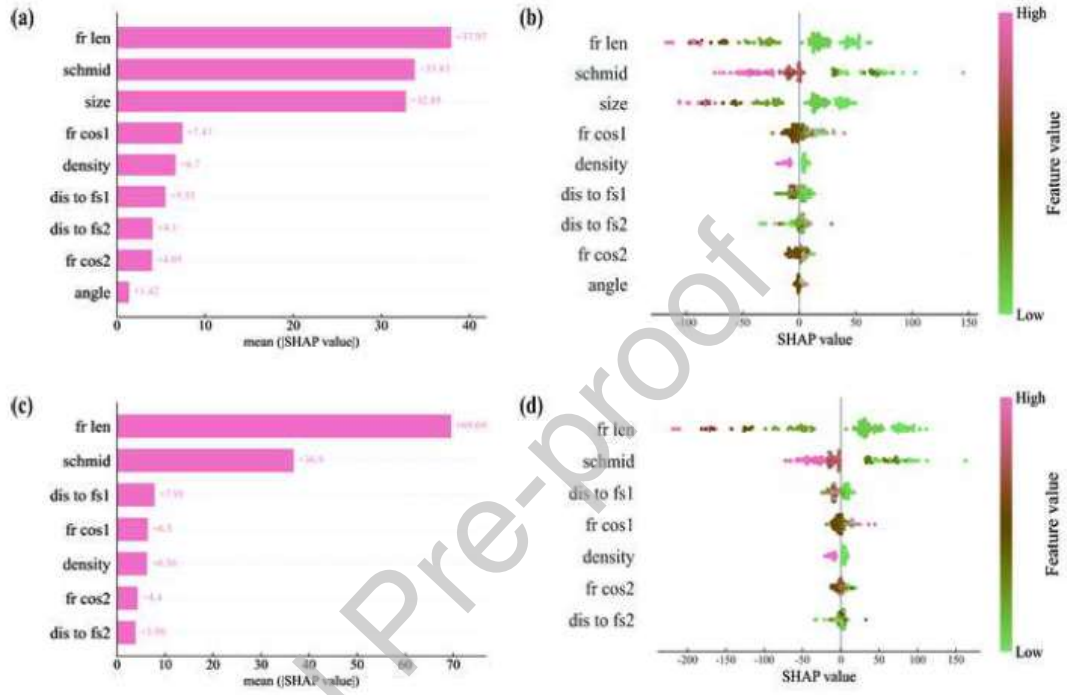


Fig. 7: Analyzing features using the SHAP tool (a) SHAP feature importance plot and (b) SHAP summary plot of internal + external feature group, the parameters of the Y-axis in the figure are L_{FR}, m, d, β^{FS} of dislocation 1(β_1^{FS}), ρ, d_1^{FS}, d_2^{FS}, β_2^{FS}, a (c) SHAP feature importance plot and (d) SHAP summary pot of internal feature group (including FR dislocation source length), the parameters of the Y-axis in the figure are L_{FR}, m, d_1^{FS}, β_1^{FS}, ρ, β_2^{FS}, d_2^{FS}.

3.3 Size effect and orientation effect

Figure 8 shows the ML results for the size effect and orientation effect. The average is taken of the yield stress in the training set of micropillars for each size or orientation, where circles represent the simulation points and triangles indicate the predicted points. Face-centered cubic (FCC) single-crystal micro/nanopillars exhibited a universal power-law relationship between yield/flow stresses and specimen size [34, 35]. Figure 8a depicts the yield stress of simulation and ML variation with size. The yield stress is fitted using the formula $\sigma/\mu = A(d/b)^m$ where μ is

shear modulus, d is the diameter and b is Burgers vector. The fitting parameters indicate A = 0.42 and m = -0.54 for the mean yield stress of micropillars at each size. Figure 8b shows the yield stress variation with the cosine of the loading orientation angle. In micropillars oriented at 0°, the primary slip system's Schmid factor is 0.272, resulting in a higher yield stress. The ML model can effectively capture the orientation effect.

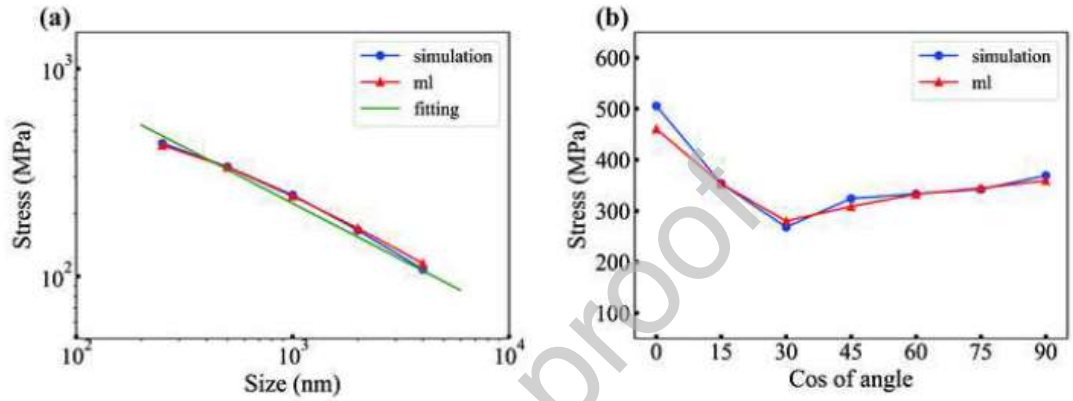


Fig. 8: Prediction performance for size effects (a), as well as orientation effects (b). Take the average yield stress of all micropillars in the training set for each size/orientation. Circles represent simulated values, and triangles represent predicted values by ML. The straight line represents the fitting of a power-law relationship to the size effect of graph (a).

Table 5 shows the *RMSE* values for the stress predictions at different scales, using either only the external features or both the internal and external features. The table indicates that the 4000 nm sample size group has a much lower prediction error than other specimens with smaller sizes. Moreover, using only the external features, namely the sample size and orientation features, yields better prediction results for 4000 nm specimens. This is because many dislocations reduce the influence of individual dislocations on the overall mechanical behavior, and the yielding process involves the activation of multiple dislocations. Thus, the feature related to the most easily activated dislocation may be redundant and decrease the prediction accuracy. Conversely, adding the internal features improves the prediction accuracy of specimens with smaller sizes. At these scales, information about more easily activated dislocations affects the yielding behavior of the specimen, and including this information enables a better prediction of the mechanical behavior. Moreover, Table 6 demonstrates that adding the dislocation information generally improves the

prediction accuracy across different orientations.

Table 5: *RMSE* values for specimens of different sizes under external feature and internal + external feature groups

RMSE values (MPa) for different specimen size features (nm)	External feature	Internal + External feature
250	43.43	42.08
500	48.50	25.96
1000	37.93	34.27
2000	45.20	28.97
4000	1.58	4.90

Table 6: *RMSE* values for specimens with different loading orientations under external feature and internal + external feature groups

RMSE values (MPa) for different orientation (°) feature	External feature	Internal + External feature
0	56.38	45.31
15	19.94	27.09
30	94.02	20.46
45	21.40	18.47
60	29.06	28.96
75	69.65	14.20
90	40.81	35.32

In summary, the predictive performance of the model has been enhanced by incorporating the internal features into external characteristics. By employing ML techniques, the stress-strain curves of single-crystal micropillars have been predicted, effectively accommodating the considerations of size and orientation effects.

In the size range investigated (250 nm to 4000 nm), the primary consideration in this study revolves around the activation mechanisms of F-R dislocation sources and the single-arm source as the dominant mechanisms. Feature selection in machine learning is closely aligned with these two mechanisms. However, when the sample size diminishes to the nanoscale (less than 100 nm), the activation mechanism of dislocation sources at free surfaces becomes the predominant factor [36]. Additionally, besides size and orientation, the strain rate [17, 21] also significantly influences the yielding of the sample. This study offers insights into the effects of various dislocation mechanisms and loading parameters on the mechanical properties of micro and nanoscale pillars. It should be noted that, in FCC or BCC micro/nanopillars with specific orientations, deformation

twinning is another dominant plastic deformation mechanism [8, 9, 37]. For example, the deformation mechanism of <001>-oriented Cu nanopillars transfers from dislocation slip to deformation twinning with decreasing diameter [8]. Moreover, second-phase particles [6] can result in the entanglement of dislocations. Therefore, in the future, the ML method should be further extended to include the features pertinent to twinning and secondary phases. Importantly, the current dataset is informative but constrained in size. Expanding the dataset has a potential to enhance the predictive accuracy significantly and help to learn more about mechanical mechanisms.

4. Conclusions

This study uses ML methods to predict the yield stress of single-crystal copper micropillar by combining the features related to the most easily activated dislocations and the external features such as size and orientation:

(1) Features related to the most easily activated dislocations in the specimen, including the distance and orientation of FR dislocation source with free surfaces, are extracted and combined with the external features, including the size and loading orientation of micropillars. Using hybrid features, various ML methods are applied to predict the yield stress and stress-strain curves of single-crystal micropillar compression.

(2) The results show that the prediction performance improves when adding the features related to the most easily activated dislocations besides the external features. With the current small dataset, the prediction performance firstly improves and then declines as the number of sampled easily activated dislocations increases, reaching the best performance when sampling two easily activated dislocations.

(3) The SHAP analysis (feature importance analysis) of the model indicates that the size has the greatest impact on the prediction results, followed by the schmid factor of main active slip system, which includes orientation information and has a significant impact on the yield stress. Moreover, the influence by easily activated dislocation is much lower than that of size and orientation.

(4) Regarding the prediction, the error for the 4000 nm sample size groups is much lower than that for other groups with smaller specimen sizes. Adding internal features notably improves

the prediction performance of specimens with sizes of 500, 1000, and 2000 nm, and enhances the prediction performance across different orientations.

Declaration of Competing Interest

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

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used chat-gpt to improve English language fluency. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (12192214, 12222209).

References

[1] B. Isralewitz, M. Gao, K. Schulten, Steered molecular dynamics and mechanical functions of proteins, Current opinion in structural biology 11(2) (2001) 224-230.

[2] L.P. Kubin, G. Canova, M. Condat, B. Devincre, V. Pontikis, Y. Bréchet, Dislocation microstructures and plastic flow: a 3D simulation, Solid State Phenomena, Trans Tech Publications, 1992, pp. 455-472.

[3] F. Roters, P. Eisenlohr, L. Hantcherli, D.D. Tjahjanto, T.R. Bieler, D. Raabe, Overview of constitutive laws, kinematics, homogenization and multiscale methods in crystal plasticity finite-element modeling: Theory, experiments, applications, Acta materialia 58(4) (2010) 1152-1211.

[4] X. Zhang, S. Lu, B. Zhang, X. Tian, Q. Kan, G. Kang, Dislocation–grain boundary interaction-based discrete dislocation dynamics modeling and its application to bicrystals with different misorientations, Acta Materialia 202 (2021) 88-98.

[5] D. Wei, M. Zaiser, Z. Feng, G. Kang, H. Fan, X. Zhang, Effects of twin boundary orientation on plasticity of bicrystalline copper micropillars: A discrete dislocation dynamics simulation study, Acta Materialia 176 (2019) 289-296.

[6] Y. Chen, Q. Fang, S. Luo, F. Liu, B. Liu, Y. Liu, Z. Huang, P.K. Liaw, J. Li, Unraveling a novel precipitate enrichment dependent strengthening behaviour in nickel-based superalloy, International Journal of Plasticity 155 (2022) 103333.

[7] I. Pan, L.R. Mason, O.K. Matar, Data-centric Engineering: integrating simulation, machine learning and statistics. Challenges and opportunities, Chemical Engineering Science 249 (2022) 117271.

[8] R. Niu, X. An, L. Li, Z. Zhang, Y.-W. Mai, X. Liao, Mechanical properties and deformation behaviours of submicron-sized Cu–Al single crystals, Acta Materialia 223 (2022) 117460.

[9] Q. Zhang, R. Huang, J. Jiang, T. Cao, Y. Zeng, J. Li, Y. Xue, X. Li, Size effects and plastic deformation mechanisms in single-crystalline CoCrFeNi micro/nanopillars, Journal of the Mechanics and Physics of Solids 162 (2022) 104853.

[10] K. Frydrych, K. Karimi, M. Pecelerowicz, R. Alvarez, F.J. Dominguez-Gutiérrez, F. Rovaris, S. Papanikolaou, Materials informatics for mechanical deformation: A review of applications and challenges, Materials 14(19) (2021) 5764.

[11] J. Han, A. Jentzen, W. E, Solving high-dimensional partial differential equations using deep learning, Proceedings of the National Academy of Sciences 115(34) (2018) 8505-8510.

[12] R.S. Yassar, O. AbuOmar, E. Hansen, M.F. Horstemeyer, On dislocation-based artificial neural network modeling of flow stress, Materials & Design 31(8) (2010) 3683-3689.

[13] H. Salmenjoki, M.J. Alava, L. Laurson, Machine learning plastic deformation of crystals, Nature Communications 9(1) (2018) 5307.

[14] D. Steinberger, H. Song, S. Sandfeld, Machine learning-based classification of dislocation microstructures, Frontiers in Materials 6 (2019) 141.

[15] Y. Zhang, A.H. Ngan, Extracting dislocation microstructures by deep learning, International Journal of Plasticity 115 (2019) 18-28.

[16] Z. Zhou, Y. Zhu, J. Luo, X. Yang, X. Guo, Characterisation of dislocation patterning behaviour with a continuum dislocation dynamics model on two parallel slip planes equipped with a deep neural network resolving local microstructures, International Journal of Solids and Structures 198 (2020) 57-71.

[17] S. Hiemer, H. Fan, M. Zaiser, Relating plasticity to dislocation properties by data analysis: scaling vs. machine learning approaches, Materials Theory 7(1) (2023) 1.

[18] N. Bertin, F. Zhou, Accelerating discrete dislocation dynamics simulations with graph neural networks, Journal of Computational Physics (2023) 112180.

[19] A. Arsenlis, W. Cai, M. Tang, M. Rhee, T. Oppelstrup, G. Hommes, T.G. Pierce, V.V. Bulatov, Enabling strain hardening simulations with dislocation dynamics, Modelling and Simulation in Materials Science and Engineering 15(6) (2007) 553.

[20] E. Oren, E. Yahel, G. Makov, Dislocation kinematics: a molecular dynamics study in Cu, Modelling and Simulation in Materials Science and Engineering 25(2) (2016) 025002.

[21] H. Fan, Q. Wang, J.A. El-Awady, D. Raabe, M. Zaiser, Strain rate dependency of dislocation plasticity, Nature communications 12(1) (2021) 1845.

[22] O. Kraft, C. Volkert, Size effects on deformation and fatigue of thin films and small structures, CAMTEC, Cambridge University (2006).

[23] P.M. Anderson, J.P. Hirth, J. Lothe, Theory of dislocations, Cambridge University Press2017.

[24] S.I. Rao, D. Dimiduk, M. Tang, M. Uchic, T. Parthasarathy, C. Woodward, Estimating the strength of single-ended dislocation sources in micron-sized single crystals, Philosophical Magazine 87(30) (2007) 4777-4794.

[25] G.K. Uyanık, N. Güler, A study on multiple linear regression analysis, Procedia-Social and Behavioral Sciences 106 (2013) 234-240.

[26] J.R. Quinlan, Simplifying decision trees, International journal of man-machine studies 27(3) (1987) 221-234.

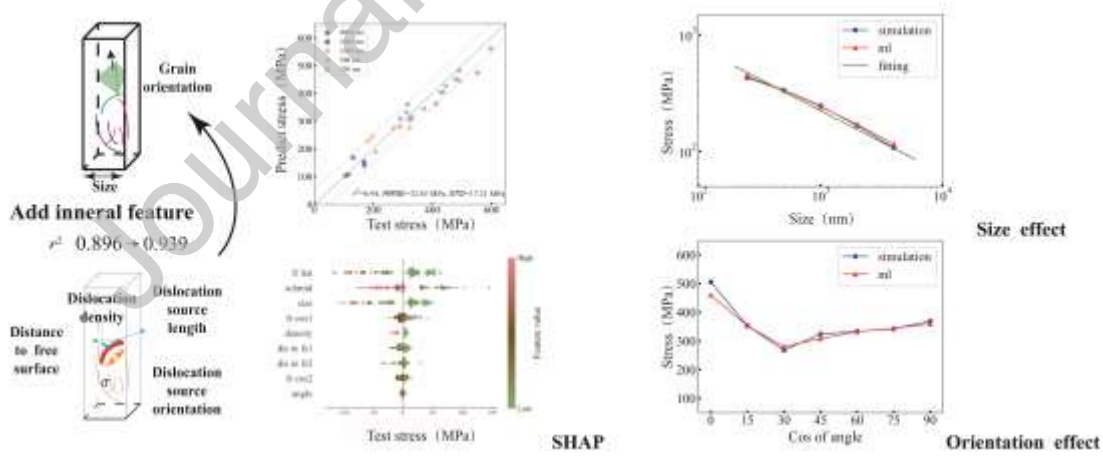
[27] L. Breiman, Random forests, Machine learning 45 (2001) 5-32.

[28] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, Lightgbm: A highly efficient gradient boosting decision tree, Advances in Neural Information Processing Systems 30 (2017).

[29] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, Scikit-learn: Machine learning in Python, the Journal of Machine

Learning Research (2011) 2825-2830.
[30] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, Advances in Neural Information Processing Systems 30 (2017).
[31] J.R. Greer, W.C. Oliver, W.D. Nix, Size dependence of mechanical properties of gold at the micron scale in the absence of strain gradients, Acta Materialia 53(6) (2005) 1821-1830.
[32] D. Dimiduk, M. Uchic, T. Parthasarathy, Size-affected single-slip behavior of pure nickel microcrystals, Acta Materialia 53(15) (2005) 4065-4077.
[33] J.A. El-Awady, S.B. Biner, N.M. Ghoniem, A self-consistent boundary element, parametric dislocation dynamics formulation of plastic flow in finite volumes, Journal of the Mechanics and Physics of Solids 56(5) (2008) 2019-2035.
[34] R. Dou, B. Derby, A universal scaling law for the strength of metal micropillars and nanowires, Scripta Materialia 61(5) (2009) 524-527.
[35] J.R. Greer, J.T.M. De Hosson, Plasticity in small-sized metallic systems: Intrinsic versus extrinsic size effect, Progress in Materials Science 56(6) (2011) 654-724.
[36] A.T. Jennings, C.R. Weinberger, S.-W. Lee, Z.H. Aitken, L. Meza, J.R. Greer, Modeling dislocation nucleation strengths in pristine metallic nanowires under experimental conditions, Acta materialia 61(6) (2013) 2244-2259.
[37] Q. Zhang, R. Huang, X. Zhang, T. Cao, Y. Xue, X. Li, Deformation mechanisms and remarkable strain hardening in single-crystalline high-entropy-alloy micropillars/nanopillars, Nano Letters 21(8) (2021) 3671-3679.

Graphical Abstract



Declaration of interests

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

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: