Available online at www.sciencedirect.com

jmr&t Journal of Materials Research and Technology

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

Original Article

Molecular dynamics simulation and machine learning of mechanical response in non-equiatomic FeCrNiCoMn high-entropy alloy

Liang Zhang [a,b,*,1], *Kun Qian* [c,1], *Jun Huang* [d], *Mao Liu* [e], *Yasushi Shibuta* [f]

[a] *International Joint Laboratory for Light Alloys (MOE), College of Materials Science and Engineering, Chongqing University, Chongqing, 400044, China*

[b] *Shenyang National Laboratory for Materials Science, Chongqing University, Chongqing, 400044, China*

[c] *Educational Physiology Laboratory, The University of Tokyo, Tokyo, 113-0033, Japan*

[d] *School of Computer Science and Technology, Anhui University of Technology, 243002, China*

[e] *School of Metallurgy, Northeastern University, Shenyang, 110819, China*

[f] *Department of Materials Engineering, The University of Tokyo, Bunkyo-ku, Tokyo, 113-8656, Japan*

ARTICLE INFO

Article history:
Received 27 February 2021
Accepted 7 June 2021
Available online 12 June 2021

Keywords:
Molecular dynamics simulation
Machine learning
High entropy alloy
Yield stress

ABSTRACT

High-entropy alloys (HEAs) have attracted a wide range of academic interest for their promising properties as structural materials, among which the equiatomic FeCrNiCoMn HEAs have been reported to possess a series of superior properties. However, one may have to change the alloy composition from the equiatomic composition to improve a specific material property. In this study, molecular dynamics simulation combined with machine learning methods was used to study the mechanical properties of non-equiatomic FeCrNiCoMn HEAs. A database describing the relationship between materials composition and mechanical properties was established based on a tensile test of 300 HEA single-crystal samples by MD simulation. We investigated and compared three ML models for the learning task of yield stress, including support vector machine (SVM), kernel-based extreme learning machine (KELM), and deep neural network (DNN). It was found that the DNN model outperformed others for the binary classification of yield stress. The elemental composition strategy was used to guide the design of polycrystal FeCrNiCoMn samples, and the accuracy of the DNN model was further verified by the polycrystal samples. We show in this contribution that computational study combined with machine learning method can provide instructive guidance for the design of high-strength HEA and accelerate the development of new alloy materials.

* Corresponding author.
E-mail address: liangz@cqu.edu.cn (L. Zhang).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jmrt.2021.06.021

1. Introduction

High-entropy alloys (HEAs) have attracted much attention in the community of materials and have been extensively studied by researchers in the past two decades due to their unique element composition, special microstructure, and highly controllable physical and chemical properties [1–4]. For conventional alloys which contain one and rarely two principal elements, the main mechanical behavior is dictated by the dominant elements. The other minor alloying elements are used to enhance some special properties, such as interstitial solute element-assisted solid-solution strengthening and phase-transformation strengthening. In contrast, HEAs generally refer to the solid-solution alloys containing five or more principal elements in equal or near equiatomic percentages (at.%) [5,6], and the properties of HEAs can be different from any of the constituent elements. Compared with conventional alloys, HEAs exhibits many superior mechanical properties, including excellent strength, exceptional ductility and fracture toughness at cryogenic temperatures, superior mechanical performance at high temperatures, etc. [7–9]. Exploring the combination of high strength and good ductility has always been a difficulty and challenge in the development of advanced metallic structural materials, but some recent studies of HEAs have shown that the well-designed HEAs can provide an effective solution to this problem. For example, a new class of transformation-induced plasticity-assisted, dual-phase HEA was designed, which exhibits excellent strength-ductility combinations [10]. An ordered interstitial complex strengthening mechanism was proposed to successfully overcome the strength-ductility trade-off conundrum for a class of HEAs [11]. Tuning the element distribution and composition of HEAs can lead to higher yield strength without compromising strain hardening and tensile ductility [12]. The excellent mechanical properties of HEA make it a promising candidate for advanced structural material for industrial applications in the future.

It has been proved that not all alloys consisting of five principal elements can form solid solutions. The equiatomic face-centered cubic (fcc) FeCrNiCoMn HEA, or namely Cantor alloy, is the most attractive prototype among all the various HEAs due to its stable phase structure and its excellent mechanical properties [9]. The operating mechanisms of plastic deformation of the Cantor alloy have been widely studied in the previous experimental work. Otto et al. [13] researched the influences of temperature on the tensile properties of an equiatomic Cantor alloy, and it was found that the yield strength, ultimate tensile strength, and elongation to fracture of the samples all increased with decreasing temperature. In particular, the nanoscale deformation twins were observed at a low temperature ($T = 77$ K), which produces a high degree of work hardening and enhanced ductility of the tested samples. Similarly, Laplanche et al. [14] investigated the microstructure evolution and critical stress for twinning of the Cantor alloy. They found an enhanced strength-ductility combination of the samples at liquid nitrogen temperature compared to room temperature, which is primarily due to twinning starting earlier in the deformation process and providing additional work hardening. Zhu et al. [15] conducted nanoindentation to characterize the nature of incipient plasticity of a single fcc structure Cantor alloy. As with conventional alloys, the maximum shear stress required to initiate plasticity was found to relatively insensitive to grain orientation but was strongly dependent upon the temperature, indicating a thermally activated process. Schuh et al. [16] studied the mechanical properties and thermal stability of an equiatomic nanocrystalline Cantor alloy with a grain size of approximately 50 nm, and they found that the hardness of the samples increases for a long annealing time due to the nanoscale phases embedded in the HEA matrix.

In addition to the equiatomic HEA, the non-equiatomic HEA has also received attention in recent years. With carefully designed composition and tailored microstructures, the non-equiatomic HEA may exhibit more interesting properties [1], as suggested by Yeh, the pioneer of HEAs. For example, a non-equiatomic sample $Fe_{40}Mn_{27}Ni_{26}Co_5Cr_2$ (at.%) was reported to achieve an excellent tensile ductility of 60% [17], suggesting that the limited tensile ductility is not an intrinsic limitation of HEAs. In addition, the bulk nanocrystalline $Co_{25}Ni_{25}Fe_{25}Al_{7.5}Cu_{17.5}$ (at.%) HEA was reported to exhibit a compressive yield strength of about 1.8 GPa [18], which is dramatically higher than the yield strength of most previously reported fcc structured HEAs. The above experimental results indicate that the strength of the equiatomic HEA can be further improved through the composition regulating of the constituent elements. However, the number of new alloy systems obtained by changing elemental composition is extremely large, it is impossible to evaluate such huge numbers of the HEA system by the empirical trial-and-error method. In recent years, the combination of high-throughput computing and advanced machine learning (ML) algorithms provide a new method for structural design and performance development of materials, making a significant contribution to the discovery of new materials [19–23].

In this work, taking Cantor alloy as the research object, we investigated the non-equiatomic FeCrNiCoMn HEA by combining high-throughput molecular dynamics (MD) simulation and machine learning methods, so as to explore the possible composition for high-strength HEA. The simulation method and ML algorithms are introduced in section 2. In section 3, MD simulations were carried out to investigate the mechanical properties of 300 single-crystal FeCrNiCoMn samples with different compositions (at.%). A database was established based on the MD results for describing the relationship between element composition and yield stress. Three ML algorithms were investigated and compared for the learning target of yield stress, and the optimized ML model was further verified by polycrystal FeCrNiCoMn samples. The main results and conclusions are summarized in section 4.

2. Methodology

2.1. *MD simulation*

MD simulations were carried out using the large-scale atomic/molecular massively parallel simulator (LAMMPS) [24]. MD simulation on a multicomponent HEA system requires interatomic potentials for all the constituent binary alloy systems. Recently, Lee's group developed the interatomic potentials for the constituent binary systems of the FeCrNiCoMn HEA system

[25]. They examined the equilibrium atomic configuration, stability of lattice structure , and mechanical properties of the equiatomic FeCrNiCoMn HEA. Also, they designed non-equiatomic FeCrNiCoMn HEAs with improved strength using the developed potentials, and the results were confirmed experimentally. Therefore, the selected interatomic potential is appropriate to study the mechanical properties of the non-equiatomic FeCrNiCoMn HEA system in the present work. The single-crystal HEA samples without initial structural defects were constructed, and the length of each side of the cube sample is approximately 10 nm. The current model size is a benefit of the computational efficiency and ensures the observation of primary nucleation of dislocations in the grain. The large polycrystal samples were constructed using the Voronoi construction method containing 16 grains with random crystallographic orientations, and the mean grain size is 14.2 nm. Atoms of the five elements are uniformly distributed in both single-crystal and polycrystal HEA samples [26], as shown in Fig. 1. All simulation samples were constructed with a single fcc phase. Many experimental studies have shown that the FeCrNiCoMn HEA maintains a stable fcc single-phase structure from the low-temperature range to the melting point without any polycrystalline phase transition [9]. This high structural stability is generally attributed to the extremely sluggish atomic diffusion in the FeCrNiCoMn HEA due to its chemical complexity and local accumulation disorder.

The equilibrium structure of the HEA samples was obtained after an initial energy minimization and the followed system relaxation procedure at 300 K for 100 ps? The fcc lattice structure of the samples was maintained after system relaxation in an isobaric-isothermal (NPT) environment. Periodic boundary conditions were applied to all directions. A uniaxial tension was applied to test the mechanical properties of HEA samples with different elemental compositions. During the dynamic tension in one direction, the pressure along the other two directions was kept at zero in the NPT ensemble. MD simulations were performed at the temperature of 300 K, the deformation strain was set at a constant rate of 5×10^8 s^{-1}, and the timestep was set as 1 fs throughout the work. The system strain was derived from the positions of the periodic boundaries along the loading direction, and the system stress was attained by calculating the pressure of the entire system of atoms. All the atomic figures in this study are illustrated by the visualization tool OVITO [27], and the dislocation extraction algorithm (DXA) [28] is used to extract the dislocation lines in the sample during deformation.

2.2. Database

A database was established to describe the quantitative relationship between elements composition and the mechanical response of HEA samples based on MD simulations of the single-crystal samples. In the ML task, the composition of the five elements FeCrNiCoMn was used as an input feature to predict the mechanical response. The principle concept of HEAs is based on designing the alloys with multiple principal elements ranging from 5 to 35 atomic percent with a target to form single-phase solid solutions arising from the high

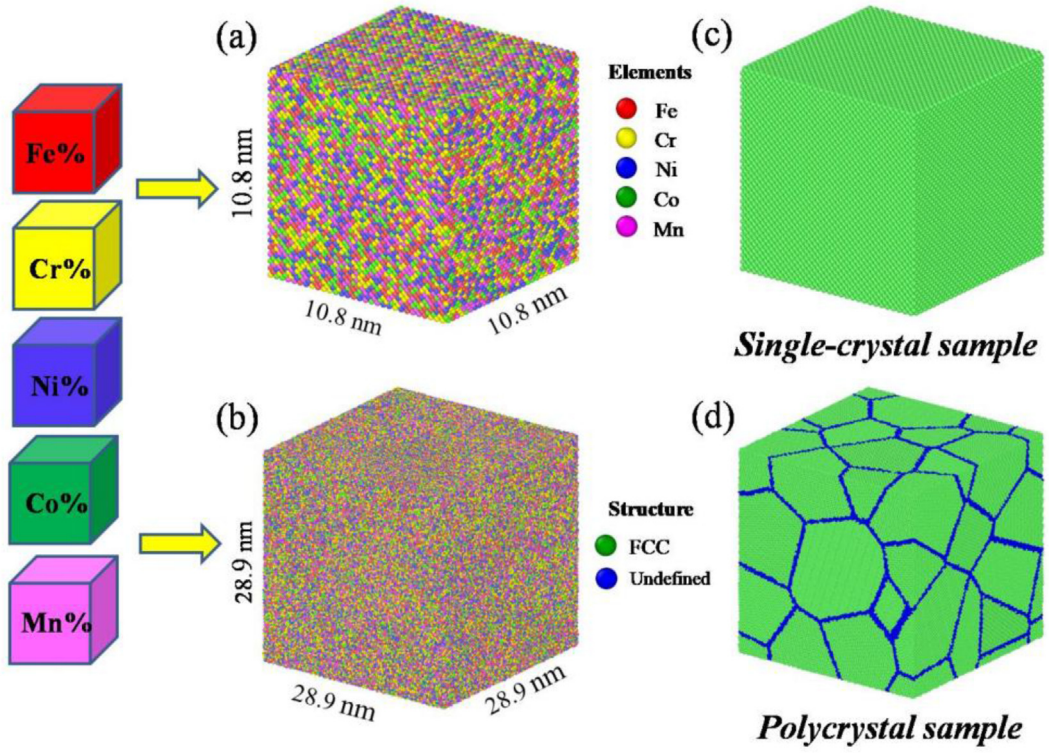


Fig. 1 – Atomic configurations of FeCrNiCoMn single-crystal sample and polycrystal sample. (a)–(b) atoms are colored according to the element types. (c)–(d) atoms are colored by the common neighbor analysis (CNA) method. The green atoms are involved in fcc structure, and the blue atoms are located at the grain boundary region with disordered structure.

entropy of the system [5]. Therefore, when randomly assigning the proportion of the five elements, the upper limit of each element was controlled to be 35%, and the lower limit to be 5%. For metallic materials, the nucleation of the first set of dislocations corresponds to the theoretical yield strength and represents the beginning of the plastic deformation [29]. Also, it plays a significant role in the contribution of materials hardening mechanisms and hence can be used to guide the design and optimization of alloy composition. Therefore, yield stress (Y_S) was set as output features. The classification of the values of yield stress was chosen as the learning target for ML in this study. When establishing the database by high throughput MD simulations on single-crystal samples, we not only changed the composition of the elements in the alloy system but also considered the anisotropy of materials. During the dynamic tension, the tensile strain was applied in three crystallographic orientations, [100], [110], and [111]. For each orientation of the data set, the MD simulation contains a result of the equiatomic HEA sample and 100 results of the non-equiatomic HEA samples based on different elemental compositions.

2.3. ML models

Our task is to predict the mechanical response (output features) of the HEA samples through element composition (input features), so the supervised learning paradigm is selected. In the present study, three ML models are investigated, namely support vector machine (SVM), kernel-based extreme learning machine (KELM), and deep neural network (DNN). Fig. 2 illustrates the basic principle of the three ML models. SVM finds a set of hyperplanes in a multi-dimensional space that instances belonging to different classes can be separated within. When training an SVM classier, a subset of data points with the widest possible gap between the class boundaries (based on the *support vectors*) from the train set will be selected as pivots to support the hyperplane that can maximize the separation between classes. When testing a given instance, the test data (*feature vectors*) will be first mapped to the multi-dimensional space, then the prediction will be given based on which side of the gap they fall on to [30], as shown in Fig. 2(a). An extreme learning machine (ELM) is a kind of a single hidden layer feed-forward artificial neural network, which has no need for tuning the parameters of the network. For an ELM, the hidden neurons are randomly initialized, by which the output weights can be analytically determined. If the feature mapping is unknown, then the kernel trick used in SVM can also be applied to ELM; this model is known as kernel-based ELM (KELM) [31], as shown in Fig. 2(b). DNN is one of the deep learning principles, which extract higher representations from the data via the help of a series of the nonlinear transformation of the inputs. The procedure of training a DNN is to update the parameters of its layers (weights and biases) iteratively for minimizing a predefined loss function, which measures the difference between the target output vectors and the actual output vectors of the network [32], as shown in Fig. 2(c).

3. Results and discussion

3.1. Single-crystal samples

In this section, the mechanical properties of the FeCrNiCoMn single-crystal samples are examined by MD simulations. Fig. 3 shows the stress–strain responses of the single-crystal samples under different tensile orientations at 300 K. For each orientation, the figure contains 100 results based on the random combination of the constituent elements. The results of the equiatomic samples are also plotted for comparison. From a theoretical point of view, the elastic limit and the incipient plasticity in crystals correspond to the nucleation of

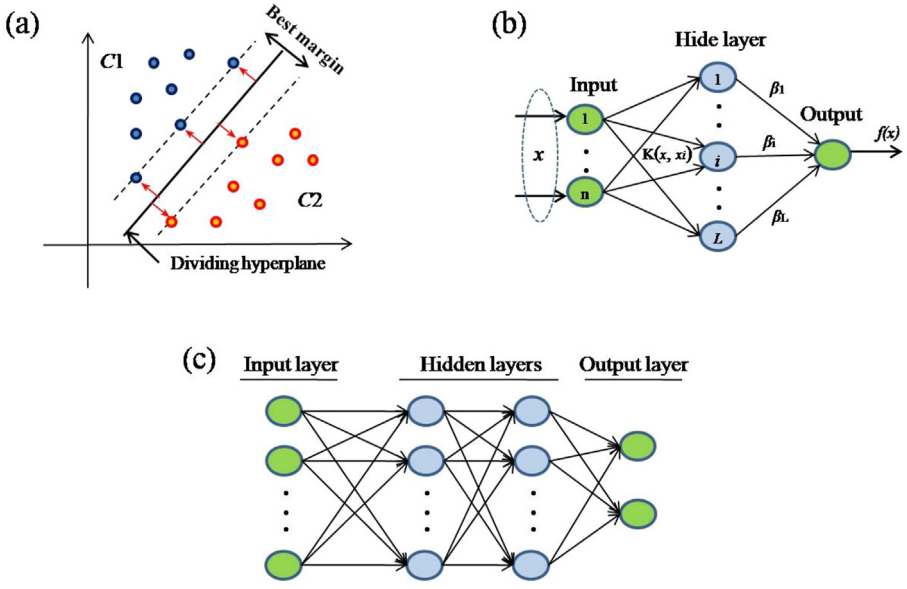
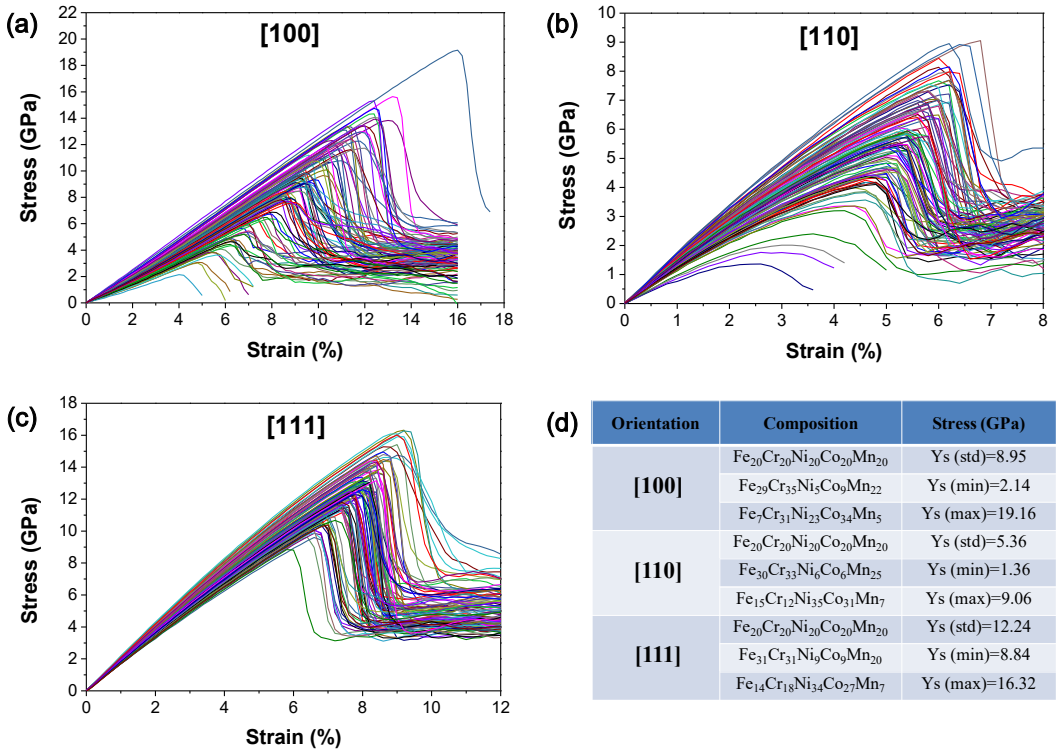


Fig. 2 – Schematic of the machine learning models. (a) SVM learning model, (b) KELM learning model network structure, (c) DNN learning model network structure.



(d)

Orientation	Composition	Stress (GPa)
[100]	$Fe_{20}Cr_{20}Ni_{20}Co_{20}Mn_{20}$	Ys (std)=8.95
	$Fe_{29}Cr_{35}Ni_5Co_9Mn_{22}$	Ys (min)=2.14
	$Fe_7Cr_{31}Ni_{23}Co_{34}Mn_5$	Ys (max)=19.16
[110]	$Fe_{20}Cr_{20}Ni_{20}Co_{20}Mn_{20}$	Ys (std)=5.36
	$Fe_{30}Cr_{33}Ni_6Co_6Mn_{25}$	Ys (min)=1.36
	$Fe_{15}Cr_{12}Ni_{35}Co_{31}Mn_7$	Ys (max)=9.06
[111]	$Fe_{20}Cr_{20}Ni_{20}Co_{20}Mn_{20}$	Ys (std)=12.24
	$Fe_{31}Cr_{31}Ni_9Co_9Mn_{20}$	Ys (min)=8.84
	$Fe_{14}Cr_{18}Ni_{34}Co_{27}Mn_7$	Ys (max)=16.32

Fig. 3 – Stress–strain response of the single-crystal HEA samples with various element compositions along different orientations. (a)–(c) Each branch figure contains the result of one equiatomic sample and 100 non-equiatomic samples. (d) List of the standard yield stress Ys(std) of the equiatomic sample, the maximum yield stress Ys(max), and minimum yield stress Ys(min) of the non-equiatomic samples.

the first set of dislocations associated with a diffusion-controlled process [29,33]. For instance, the quantitative uniaxial tensile strength measurements of pristine Pd nanowires demonstrate that the first measurable inelastic event is attributed to the nucleation of a mobile dislocation from the surface of the nanowires [34]. Visual inspection of the MD result in this study, it is found that the dislocation nucleation occurs near the peak of the stress–strain curve. Therefore, the peak stress (elastic limit) is defined as the yield stress in the simulation work, which is consistent with previous MD studies on the mechanical properties of materials [35–37]. According to the MD simulations, we have the following main results.

Firstly, the mechanical response shows an elastic anisotropy of the single-crystal samples with different crystal orientations. Young's modulus for each crystal orientation is obtained from the linear portion of its corresponding stress–strain between 0 and 0.5% strain. Taking the equiatomic sample $Fe_{20}Cr_{20}Ni_{20}Co_{20}Mn_{20}$ as an example, the tension along [111] orientation shows the maximum Young's modulus of 184.63 GPa, while the [100] orientation has the minimum value of 89.67 GPa. Also, there are remarkable differences in yield stress or elastic limit for different orientations. The maximum yield stress is 12.24 GPa for [111] orientation, and the minimum value is 5.36 GPa for [110] orientation. Note that, the lowest yield stress value does not occur in the orientation with the lowest elastic modulus , indicating that there is no direct correspondence between the two. This phenomenon can be attributed to the nonlinear elastic behavior in the elastic deformation stage [38]. Therefore, considering the anisotropy of mechanical properties of materials is beneficial to establish a reliable database and improve the applicability of the trained ML model to the subsequent complex polycrystal samples.

Secondly, it is found that the yield stresses of the equiatomic samples are at a medium level of all the three tested orientations. Fig. 3(a) plots the stress–strain responses of the FeCrNiCoMn samples with different element compositions along [100] orientation. The maximum yield stress is 19.16 GPa for the sample of $Fe_7Cr_{31}Ni_{23}Co_{34}Mn_5$, which is much higher than the minimum value of 2.14 GPa for the sample of $Fe_{29}Cr_{35}Ni_5Co_9Mn_{22}$. For the equiatomic sample, the yield stress (8.95 GPa) is at a medium level among all the tested samples along [100] orientation. Similar results are observed for the other two orientations. In Fig. 3(b), the yield stress of the equiatomic sample along [110] orientations is 5.36 GPa, which is between the minimum value (1.36 GPa) for the sample of $Fe_{30}Cr_{33}Ni_6Co_6Mn_{25}$ and the maximum value (9.06 GPa) for the sample of $Fe_{15}Cr_{12}Ni_{35}Co_{31}Mn_7$. In Fig. 3(c), the yield stress of the equiatomic sample is 12.24 GPa along [111] orientation, which is near the mean value of the 100 non-equiatomic samples in the corresponding orientation. The simulation results indicate that changing the proportions of elements can either improve or decrease the mechanical properties of the equiatomic FeCrNiCoMn samples, which provide an opportunity for exploring the higher performance of HEA by optimizing its elemental compositions.

Thirdly, the yield stress value is directly affected by the content of the individual element. The maximum and

minimum yield stresses in different orientations are summarized in Fig. 3(d). It can be found that the higher yield stresses are accompanied by the higher content of elements Ni and Co for the tested samples. On the contrary, the yield stresses are significantly reduced when the proportion of elements Fe and Cr increases. To further clarify the influence of the individual element on the yield stress of the samples. We plot the yield stress as a function of the element content along different tensile orientations, as shown in Fig. 4. Although different orientations are considered here, their sensitivity to the content of each element is generally consistent. That is, the increase of Ni and Co content is conducive to the increase of yield stress,

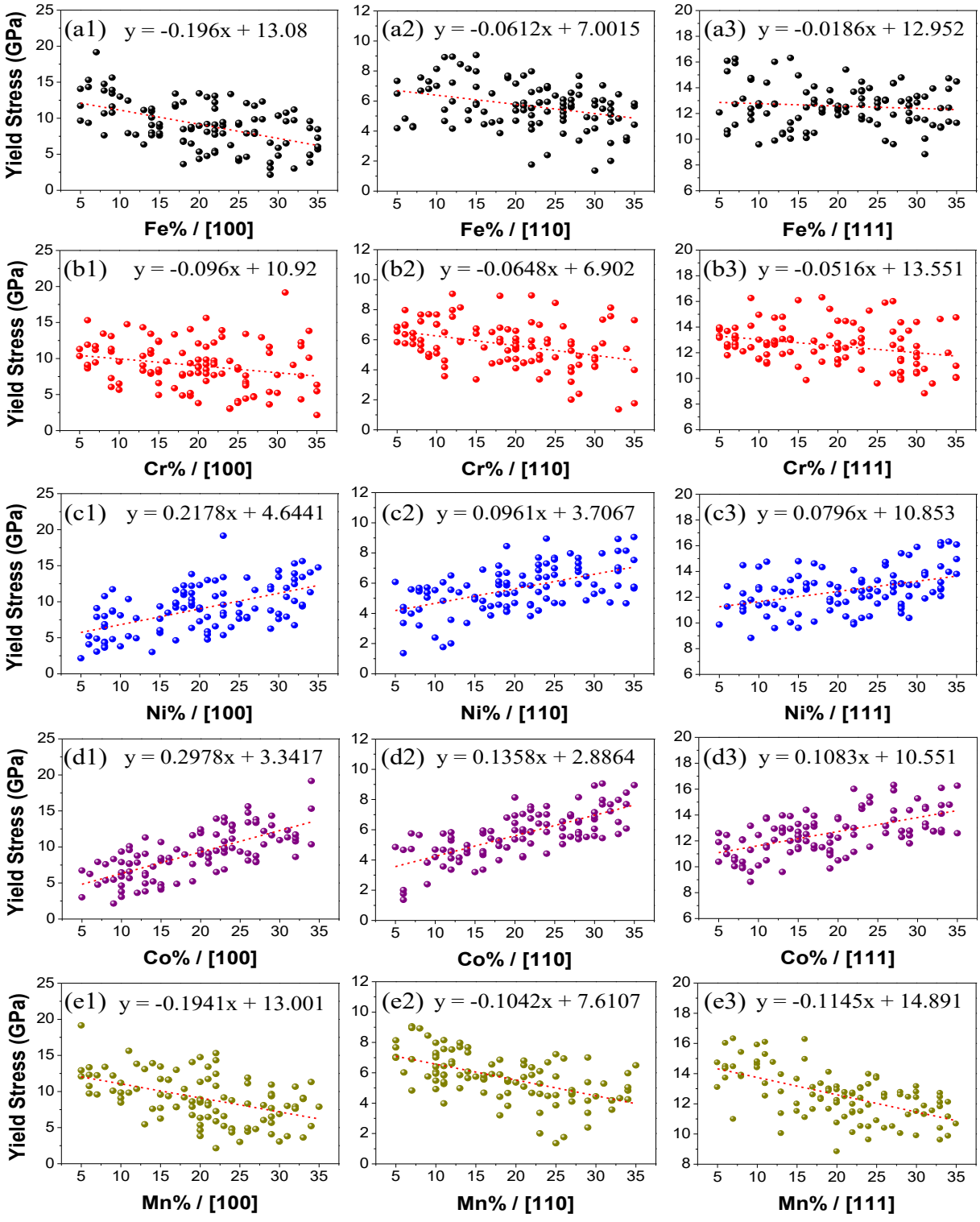
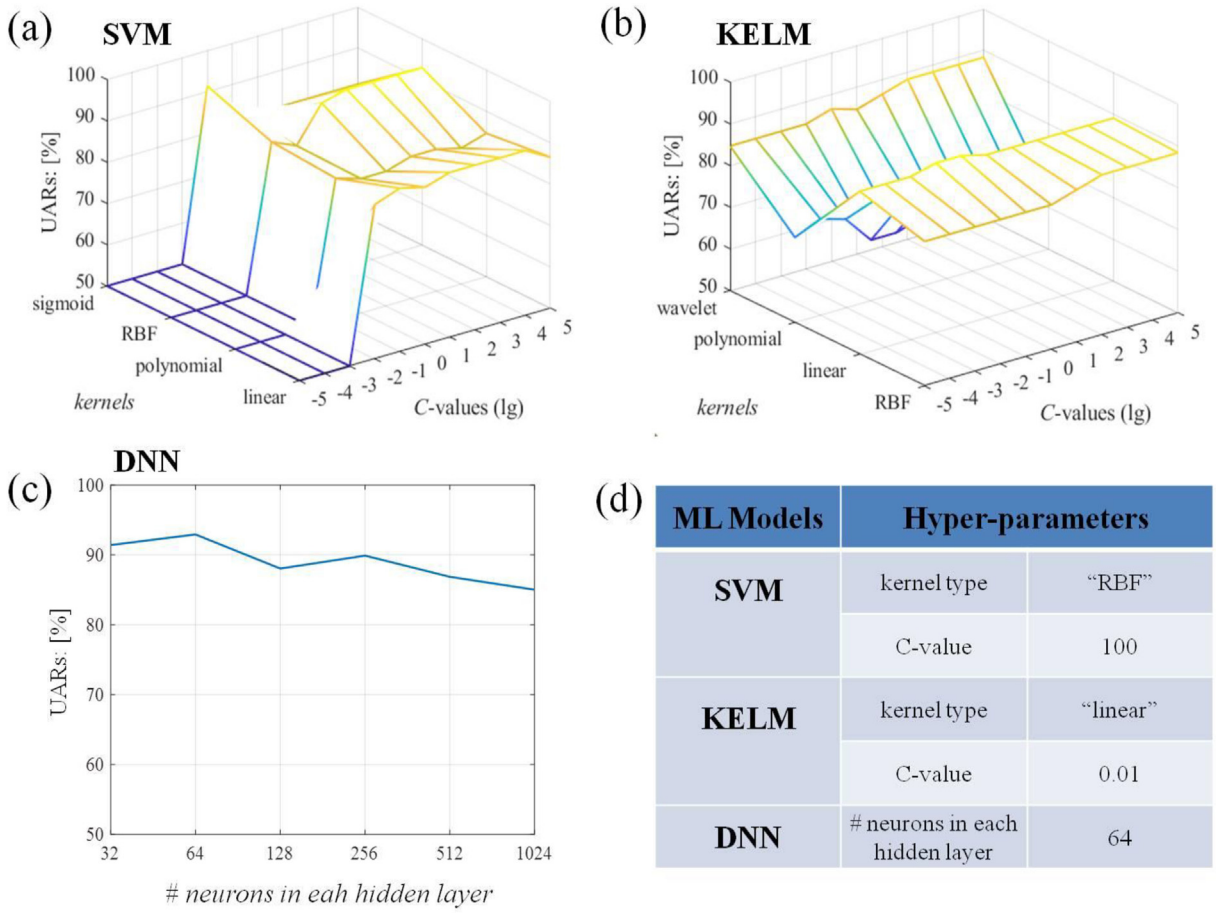


Fig. 4 – The yield stress as a function of individual element content along different tensile orientations. The data are fitted by linear method, and the equations are given.



ML Models	Hyper-parameters	
SVM	kernel type	"RBF"
	C-value	100
KELM	kernel type	"linear"
	C-value	0.01
DNN	# neurons in each hidden layer	64

Fig. 5 – The hyper-parameters tuning process for output feature by different ML models, (a) SVM model, (b) KELM model, and (c) DNN model. (d) The optimal hyper-parameters selected for evaluating the ML models on test set.

while the yield stress shows a gradual decrease with the increase of Fe, Cr, and Mn. Based on the above results, it can be simply considered that a FeCrNiCoMn sample with higher yield stress and improved mechanical properties can be obtained by increasing the content of Ni and Co and reducing the content of Fe, Cr, and Mn. However, the actual preparation and manufacture of material are not as simple as the results suggest, and many factors such as production cost and environmental impact need to be considered. For example, Co is a rare metal that is expensive and toxic, it is one of the goals to reduce the content of Co while obtaining the required mechanical performance of the FeCrNiCoMn. Therefore, it is instructive to develop a tool for fast and accurate prediction of alloy properties based on alloying element composition.

3.2. *Machine learning*

Based on the MD simulations on the 300 FeCrNiCoMn single-crystal samples, a database was constructed to describe the quantitative relationship between element composition (input features) and the yield stress (output feature) of the HEA samples. The ML task is a binary classification problem. The values of yield stress of the non-equiatomic HEA samples are divided into two categories: 'Good' and 'Weak' according to the benchmark value of the equiatomic HEA sample; the values above the benchmark value are classified as 'Good' and other values as 'Weak'. If the prediction by ML model is consistent with the MD simulation result, the predicted result by ML model is regarded as accuracy. Three ML models are investigated for the given tasks, including SVM, KELM, and DNN. The overall database was split into train, development (dev), and test sets, which occupies 60%, 20%, and 20% of the whole instances, respectively. Before being fed into the ML models, all the input features were standardized using a z-score transformation to eliminate the effects of outliers. The feature vectors were scaled to a distribution having an arithmetic mean of zero and a variance of one. Let $x(n)$ be the composition vector of the n-th element in the HEA data, and we standardize the values of $x(n)$ as:

$$\tilde{x}(n) = \frac{x(n) - \mu_x}{\sigma_x}$$

where μ_x and σ_x are the mean and the standard deviation of the vector $x(n)$, respectively. We keep this information (μ_x and σ_x) measured in the train set and apply it to the dev set and test set.

Due to the imbalanced characteristic of the proposed database (the number of instances belonging to each class in the database is not equal), the unweighted average recall (UAR) are used as our main metric to evaluate the model's

performance. Firstly, we define the *recall* (aka. class-wise accuracy) of the i-th class as:

$$Recall_i = \frac{\overline{N_i}}{N_i}$$

where $\overline{N_i}$ and N_i are the correctly recognized instance number and the total instance number of the i-th class, respectively. The weighted average recall (WAR) is defined as:

$$\mathrm{WAR} = \sum_{i=1}^{N_c} \lambda_i \ Recall_i$$

$$\lambda_i = \frac{N_i}{N}$$

where N_c and N are the number of classes in the task and the total number of instances, respectively. We know that, if the data distribution is extremely imbalanced, an ML model will usually be trained much stronger in recognizing the classes which occupy a larger percentage than the other classes in the total dataset. Therefore, using WAR to evaluate the final performance for an imbalanced dataset could be overoptimistic. In contrast, the UAR is defined as:

$$\mathrm{UAR} = \frac{\sum_{i=1}^{N_c} Recall_i}{N_c}$$

We can see that, for a balanced dataset (i.e., λ_i is a constant), WAR is equal to UAR.

The hyper-parameters of all ML models are tuned and optimized on dev set based on the performance (UAR). We use a grid-searching strategy to decide the optimal hyper-parameter for a specific ML model. The KELM and DNN models are implemented by the MATLAB (R2019a) *Statistics and Machine Learning Toolbox* [39], MATLAB (R2019a) *Deep Learning Toolbox* [40], and MATLAB (K) ELM Codes [41]. The SVM model is implemented by the LIBSVM Toolkit [42]. We use the default hyper-parameters designed in these tools except claimed as follows. Fig. 5 shows the grid searching strategy we use for optimizing the hyper-parameters of the ML models. From Fig. 5(a)–(c) we can see that, optimization of the hyper-parameters is necessary for building an efficient ML model. All the hyper-parameters of the models are tuned and optimized based on their performance (UAR) on the dev set. Then, the optimal hyper-parameter will be applied to train a new model based on the combination of train and dev sets. Finally, test set (unseen) will be evaluated by this new model. Fig. 5(d) shows the optimal hyper-parameters selected for evaluating the final ML models on test set. When tuning the hyper-parameters by the grid searching strategy, the earliest value (calculating from the start point of the grid) will be selected if there are more than one peak value (i.e., highest UAR) achieved in the whole grid.

Fig. 6 present the achieved results (UARs in [%]) of the three ML models for the learning task. We can see in Fig. 6(a) that all of the results are above 90% of UAR, indicating that the ML models perform efficiently for the given learning task. The best UAR achieved on the dev set is reached by SVM and DNN models (92.9%). On the test set, the DNN model takes the first place with a UAR of 93.7%. Although the KELM model shows a less accuracy than the other two models, it can also show a good capacity in predicting the results. The confusion matrices of the investigated three ML models are shown in Fig. 6(b) and (c). We can find that, the DNN model has the highest recall for recognizing 'good' HEA samples whereas the KELM model has the best recall performance in detecting the 'weak' HEA samples. The ML results show that the relationship between the element composition (input features) and the yield stress (output feature) of the HEA samples is sufficient to build ML models.

Considering our task is a binary classification problem, we also provide a series of complementary evaluation metrics, including the sensitivity (Sens.), the specificity (Spec.), the precision (Prec.), and the F1-score. They are defined as:

$$\mathrm{sensitivity} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

$$\mathrm{specificity} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}$$

$$\mathrm{precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

$$\mathrm{F1\ score} = 2 \times \frac{\mathrm{precision} \times \mathrm{sensitivity}}{\mathrm{precision} + \mathrm{sensitivity}} = \frac{2\mathrm{TP}}{2\mathrm{TP} + \mathrm{FP} + \mathrm{FN}}$$

where TP, TN, FP, and FN are the number of *true positive* ("Good" correctly identified as "Good"), *true negative* ("Weak" correctly identified as "Weak"), *false positive* ("Weak" incorrectly identified as "Good"), and *false negative* ("Good" incorrectly identified as "Weak"), respectively. The results of the complementary evaluation metrics are listed in Table 1. By comparing the UAR results as well as other complementary evaluation metrics of the above models, we think that the DNN model has a better overall performance for the learning task, and it is selected as the preferred ML model in the subsequent prediction studies on polycrystal samples.

3.3. *Polycrystal samples*

The polycrystal sample is the combination of single crystals with random orientations with a complex grain boundary network, which can better represent the texture of real materials. Both experiments and simulations show that grain boundaries play an important role in the mechanical properties of materials, especially when the grain size drops to the nanoscale region (<100 nm) [43–47]. Therefore, it is desirable to know whether the simulation results based on single-crystal samples can be used to guide the design of polycrystal samples to improve their mechanical performance. Meanwhile, whether the ML model based on the optimization of single-crystal samples can be applied to polycrystalline samples remains to be verified. Here, we constructed one HEA sample with equiatomic composition and four HEA samples with non-equiatomic composition. Considering the influence of elements on the mechanical properties of the HEA sample, we deliberately designed four non-equiatomic samples, including $Fe_{10}Cr_{10}Ni_{35}Co_{35}Mn_{10}$, $Fe_{30}Cr_{20}Ni_{10}Co_{10}Mn_{30}$, $Fe_{15}Cr_{15}Ni_{35}Co_{20}Mn_{15}$, and $Fe_{25}Cr_{25}Ni_{10}Co_{15}Mn_{25}$. Each polycrystal sample contains 16 grains with a mean grain size of 14.2 nm, and the total number of atoms is 2,003,895, as shown in Fig. 7.

Fig. 7(c) shows the mechanical response of the FeCrNiCoMn polycrystal samples with different compositions under

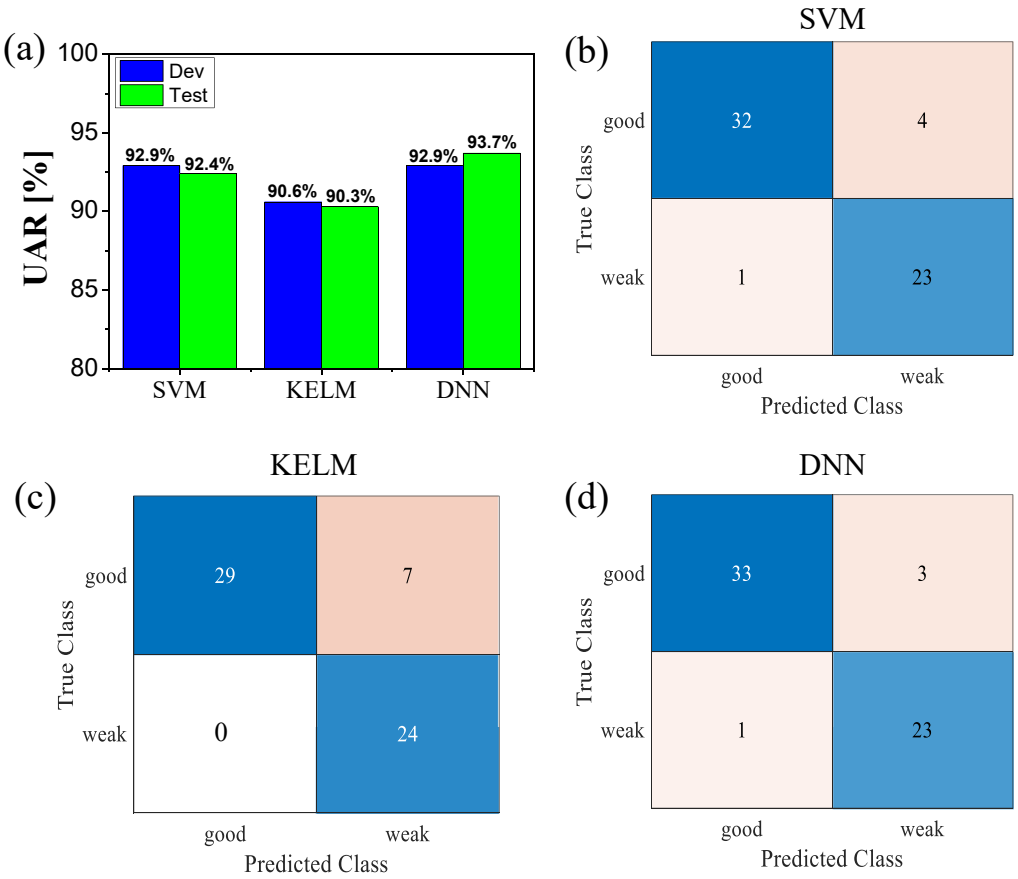


Fig. 6 – (a) The results (UAR in [%]) of three ML models on the task of yield stress. The UAR shown for the dev set are the best UAR achieved by the optimal hyper-parameters tuned for the corresponding model. The UAR shown for the test set are the final performance achieved by the model trained by the train plus dev sets within the optimal hyper-parameters. (b)–(d) The confusion matrices of three ML models.

uniaxial tension at 300 K by MD simulations. The yield stresses of the samples are found between 4–6% strain depending on their element compositions. The sample $Fe_{10}Cr_{10}Ni_{35}Co_{35}Mn_{10}$ has the highest yield stress (5.98 GPa), which is 25.9% higher than the standard value of the equiatomic sample (4.43 GPa). The lowest yield stress (2.74 GPa) is found in the sample $Fe_{30}Cr_{20}Ni_{10}Co_{10}Mn_{30}$, a 38.1% decrease from the standard value. As mentioned above, the yield stress depends primarily on the nucleation of the first set of dislocations [29]. Fig.7(a) shows the configuration of the $Fe_{10}Cr_{10}Ni_{35}Co_{35}Mn_{10}$ sample near the peak stress of the stress–strain curve. It shows that the first group of dislocation nucleation from grain boundaries at 5.1% tensile strain, leaving stacking faults crossing the grains. To clearly show the dislocation activities inside the grain in Fig. 7(d), the dislocation lines are extracted using the DXA method. At the strain of 5.1%, some dislocations are emitted from grain boundaries and they are slipping across the grain, while some embryo dislocations are just nucleated from the grain boundaries, as indicated by the arrows. Since the stress required for dislocation nucleation at structure defects (e.g. grain boundary) is much lower than that required for nucleation of the single-crystal sample with defect-free structure [35,48], it is reasonable to see that the yield stress of the polycrystal samples shows an overall decreasing trend when compared with the single-crystal samples. The stress gradually decreases and tends to be flat, entering the stress flow stage with a higher density of dislocations nucleated from grain boundaries. Note that, the yield stress and the overall stress value of the equiatomic samples are at a medium level comparing with the non-equiatomic counterparts, which is consistent with the result of single-crystal samples. The simulation results again prove that the element composition has an important effect on the mechanical properties of HEA samples.

Although the yield stress of the polycrystal samples is lower than that of single-crystal samples due to the existence of grain boundaries, the dislocation nucleation is still the

Table 1 – The results of the ML models on test set by complementary evaluation metrics (in [%]).

ML Model	WAR	Sens.	Spec.	Prec.	F1-score	Average
SVM	91.7	88.9	95.8	97.0	92.8	**93.24**
KELM	88.3	80.6	100.0	100.0	89.2	**91.62**
DNN	93.3	91.7	95.8	97.1	94.3	**94.44**

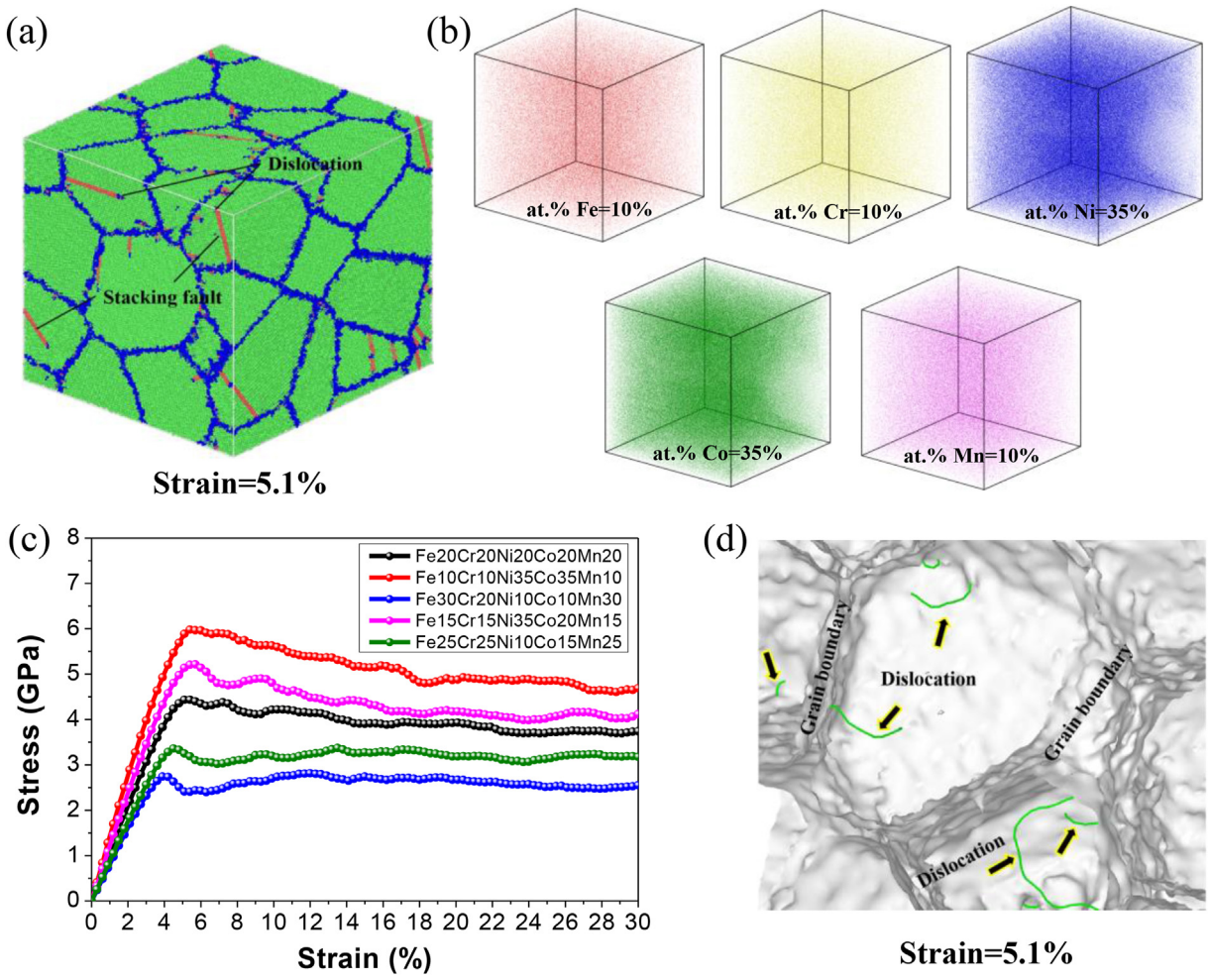


Fig. 7 – (a) Atomic configuration of the polycrystal sample $Fe_{10}Cr_{10}Ni_{35}Co_{35}Mn_{10}$ with 16 randomly orientated grains, the mean grain size is 24.8 nm. Atoms are colored by the CNA method. The green atoms are located in the grain with fcc structure, the blue atoms are located at the grain boundary region with disordered structure, and the red atoms represent the stacking faults created by the slip of the nucleated dislocations from grain boundaries. (b) Elemental composition of the polycrystal sample $Fe_{10}Cr_{10}Ni_{35}Co_{35}Mn_{10}$. (c) Stress–strain response of the polycrystal HEA samples with different element compositions. (d) Dislocations and interfaces extracted by the DXA method.

main reason for the yield of polycrystal samples at the current grain size. Also, it can be seen from Fig. 7(a) that the morphology of the grain boundary network near the yield point has not changed substantially compared with the initial configuration, indicating that the existence of grain boundaries of polycrystal samples has no significant influence on the learning task of yield stress by ML model. The DNN model developed from the single-crystal samples is used to perform the binary classification problem of yield stress for the polycrystal samples. The prediction results ('Good' or 'Weak') of the four non-equiatomic samples given by the DNN model are consistent with the MD simulation result. Considering the consumption of computational resources, we did not test many polycrystal samples in the present study. In the following work, we will further prove the efficiency of the ML model through high-throughput calculation and experimental test results. Nevertheless, the current work has demonstrated that computational simulation combined with the ML method is an efficient way to give a quick prediction on the mechanical properties of HEAs.

4. Conclusions

The high-entropy alloys exhibit many excellent mechanical properties and have a good prospect of engineering application. Most of the previous work has focused on HEAs with equiatomic compositions, while the non-equiatomic HEAs may exhibit surprising properties due to their large adjustable space. In this study, we combined molecular dynamics simulations with machine learning methods to develop a new method for composition design and mechanical property prediction of non-equiatomic FeCrNiCoMn samples.

A database for describing the relationship between elemental composition and yield stress was established based on a tensile test on 300 single-crystal samples by MD simulations. The results

show that the mechanical properties of the HEA samples can change considerably by adjusting their elemental compositions. The random combination of the element composition for the non-equiatomic samples can either improve or decrease the yield stress of the equiatomic samples. Taking the binary classification of yield stress as the learning objective, three ML models (SVM, KELM, and DNN) were investigated and compared. According to the UAR and the complementary evaluation metrics, the DNN model was found to give a prediction with the highest accuracy. The elemental composition strategy was used to guide the design of the polycrystalline HEA samples, the variation of their mechanical properties were in accordance with the expected law. The accuracy of the DNN model was further verified by the polycrystal HEA samples, and the prediction of yield stress is consistent with the MD simulation results. A combination of high-throughput computing and machine learning methods will provide instructive guidance for the sample preparation at the experiment stage and will accelerate the development of new alloys with high performance.

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.

Acknowledgement

This research was supported by the National Key Research and Development Program of China (2020YFA0405900), the National Natural Science Foundation of China (52071034), and the research grants from Japan Society for the Promotion of Science (P19081; 19H02415).

REFERENCES

[1] Tsai M-H, Yeh J-W. High-entropy alloys: a critical review. Mater Res Lett 2014;2:107–23.
[2] Ye YF, Wang Q, Lu J, Liu CT, Yang Y. High-entropy alloy: challenges and prospects. Mater Today 2016;19:349–62.
[3] Miracle DB, Senkov ON. A critical review of high entropy alloys and related concepts. Acta Mater 2017;122:448–511.
[4] George EP, Raabe D, Ritchie RO. High-entropy alloys. Nature Rev Mater 2019;4:515–34.
[5] Yeh J-W, Chen S-K, Lin S-J, Gan J-Y, Chin T-S, Shun T-T, et al. Nanostructured high-entropy alloys with multiple principal elements: novel alloy design concepts and outcomes. Adv Eng Mater 2004;6:299–303.
[6] Cantor B, Chang ITH, Knight P, Vincent AJB. Microstructural development in equiatomic multicomponent alloys. Mater Sci Eng A 2004;375–377:213–8.
[7] Zhang Y, Zuo TT, Tang Z, Gao MC, Dahmen KA, Liaw PK, et al. Microstructures and properties of high-entropy alloys. Prog Mater Sci 2014;61:1–93.
[8] Diao HY, Feng R, Dahmen KA, Liaw PK. Fundamental deformation behavior in high-entropy alloys: an overview. Curr Opin Solid State Mater Sci 2017;21:252–66.
[9] Li Z, Zhao S, Ritchie RO, Meyers MA. Mechanical properties of high-entropy alloys with emphasis on face-centered cubic alloys. Prog Mater Sci 2019;102:296–345.
[10] Li Z, Pradeep KG, Deng Y, Raabe D, Tasan CC. Metastable high-entropy dual-phase alloys overcome the strength–ductility trade-off. Nature 2016;534:227.
[11] Lei Z, Liu X, Wu Y, Wang H, Jiang S, Wang S, et al. Enhanced strength and ductility in a high-entropy alloy via ordered oxygen complexes. Nature 2018;563:546–50.
[12] Ding Q, Zhang Y, Chen X, Fu X, Chen D, Chen S, et al. Tuning element distribution, structure and properties by composition in high-entropy alloys. Nature 2019;574:223–7.
[13] Otto F, Dlouhý A, Somsen C, Bei H, Eggeler G, George EP. The influences of temperature and microstructure on the tensile properties of a CoCrFeMnNi high-entropy alloy. Acta Mater 2013;61:5743–55.
[14] Laplanche G, Kostka A, Horst OM, Eggeler G, George EP. Microstructure evolution and critical stress for twinning in the CrMnFeCoNi high-entropy alloy. Acta Mater 2016;118:152–63.
[15] Zhu C, Lu ZP, Nieh TG. Incipient plasticity and dislocation nucleation of FeCoCrNiMn high-entropy alloy. Acta Mater 2013;61:2993–3001.
[16] Schuh B, Mendez-Martin F, Völker B, George EP, Clemens H, Pippan R, et al. Mechanical properties, microstructure and thermal stability of a nanocrystalline CoCrFeMnNi high-entropy alloy after severe plastic deformation. Acta Mater 2015;96:258–68.
[17] Yao MJ, Pradeep KG, Tasan CC, Raabe D. A novel, single phase, non-equiatomic FeMnNiCoCr high-entropy alloy with exceptional phase stability and tensile ductility. Scripta Mater 2014;72–73:5–8.
[18] Fu Z, Chen W, Wen H, Zhang D, Chen Z, Zheng B, et al. Microstructure and strengthening mechanisms in an FCC structured single-phase nanocrystalline Co25Ni25Fe25Al7.5Cu17.5 high-entropy alloy. Acta Mater 2016;107:59–71.
[19] Jain A, Ong SP, Hautier G, Chen W, Richards WD, Dacek S, et al. Commentary: the Materials Project: a materials genome approach to accelerating materials innovation. Apl Mater 2013;1:011002.
[20] Curtarolo S, Hart GLW, Nardelli MB, Mingo N, Sanvito S, Levy O. The high-throughput highway to computational materials design. Nat Mater 2013;12:191–201.
[21] Xue D, Balachandran PV, Hogden J, Theiler J, Xue D, Lookman T. Accelerated search for materials with targeted properties by adaptive design. Nat Commun 2016;7:11241.
[22] Ramprasad R, Batra R, Pilania G, Mannodi-Kanakkithodi A, Kim C. Machine learning in materials informatics: recent applications and prospects. npj Computational Materials 2017;3:54.
[23] Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science. Nature 2018;559:547–55.
[24] Plimpton S. Fast parallel algorithms for short-range molecular dynamics. J Comput Phys 1995;117:1–19.
[25] Choi W-M, Jo YH, Sohn SS, Lee S, Lee B-J. Understanding the physical metallurgy of the CoCrFeMnNi high-entropy alloy: an atomistic simulation study. npj Computational Materials 2018;4:1.
[26] Zhang L, Shibuta Y. Inverse Hall-Petch relationship of high-entropy alloy by atomistic simulation. Mater Lett 2020;274:128024.
[27] Stukowski A. Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool. Model Simulat Mater Sci Eng 2010;18:015012.
[28] Stukowski A, Albe K. Dislocation detection algorithm for atomistic simulations. Model Simulat Mater Sci Eng 2010;18:025016.

[29] Li J. Dislocation nucleation: diffusive origins. Nat Mater 2015;14:656.
[30] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273–97.
[31] Huang G, Zhou H, Ding X, Zhang R. Extreme learning machine for regression and multiclass classification. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 2012;42:513–29.
[32] LeCun Y, Bengio Y, Hinton G. Deep learning, Nature 2015;521:436–44.
[33] Li J, Van Vliet KJ, Zhu T, Yip S, Suresh S. Atomistic mechanisms governing elastic limit and incipient plasticity in crystals. Nature 2002;418:307–10.
[34] Chen LY, He M-r, Shin J, Richter G, Gianola DS. Measuring surface dislocation nucleation in defect-scarce nanostructures. Nat Mater 2015;14:707–13.
[35] Tschopp MA, McDowell DL. Influence of single crystal orientation on homogeneous dislocation nucleation under uniaxial loading. J Mech Phys Solid 2008;56:1806–30.
[36] Tschopp MA, Spearot DE, McDowell DL. Atomistic simulations of homogeneous dislocation nucleation in single crystal copper. Model Simulat Mater Sci Eng 2007;15:693–709.
[37] Salehinia I, Bahr DF. Crystal orientation effect on dislocation nucleation and multiplication in FCC single crystal under uniaxial loading. Int J Plast 2014;52:133–46.
[38] Zhang L, Lu C, Tieu AK. Nonlinear elastic response of single crystal Cu under uniaxial loading by molecular dynamics study. Mater Lett 2018;227:236–9.
[39] MATLAB. Statistics and machine learning toolbox. Natick, MA, USA: MathWorks, Inc.; 2019a.
[40] MATLAB. Deep learning toolbox. Natick, MA, USA: MathWorks, Inc.; 2019b.
[41] MATLAB. (K)ELM codes: Basic ELM algorithms (MATLAB version). Singapore: Nanyang Technological University; 2006.
[42] Chang CC, Lin CJ. "LIBSVM: a library for support vector machines". ACM Transac Intell Syst Technol 2011;2:1–27.
[43] Dao M, Lu L, Asaro RJ, De Hosson JTM, Ma E. Toward a quantitative understanding of mechanical behavior of nanocrystalline metals. Acta Mater 2007;55:4041–65.
[44] Mishin Y, Asta M, Li J. Atomistic modeling of interfaces and their impact on microstructure and properties. Acta Mater 2010;58:1117–51.
[45] Zhang L, Lu C, Tieu K. A review on atomistic simulation of grain boundary behaviors in face-centered cubic metals. Comput Mater Sci 2016;118:180–91.
[46] Zhang L, Shibuta Y, Huang X, Lu C, Liu M. Grain boundary induced deformation mechanisms in nanocrystalline Al by molecular dynamics simulation: from interatomic potential perspective. Comput Mater Sci 2019;156:421–33.
[47] Greer JR, De Hosson JTM. Plasticity in small-sized metallic systems: intrinsic versus extrinsic size effect. Prog Mater Sci 2011;56:654–724.
[48] Zhang L, Lu C, Tieu K, Su L, Zhao X, Pei L. Stacking fault tetrahedron induced plasticity in copper single crystal. Mater Sci Eng A 2017;680:27–38.