

Full length article

A hybrid prediction frame for HEAs based on empirical knowledge and machine learning

Shuai Hou^{a,*}, Mengyue Sun^a, Meijuan Bai^a, Dong Lin^b, Yujiao Li^a, Weiwei Liu^{c,*}^a School of Information and Electrical Engineering, Hebei University of Engineering, Handan, Hebei, 056038, China^b Industrial and Manufacturing Systems Engineering, Kansas State University, Manhattan, KS 66506, United States of America^c School of Mechanical Engineering, Dalian University of Technology, Dalian, Liaoning, 116024, China

ARTICLE INFO

Article history:

Received 24 August 2021

Revised 2 February 2022

Accepted 8 February 2022

Available online 9 February 2022

Keywords:

Hybrid model

High-entropy alloys

Phase prediction

DS evidence theory

ABSTRACT

Phase formation plays key role in the properties of high-entropy alloys (HEAs). If the phases of HEAs can be accurately predicted, the number of experiments can be greatly reduced, and the process of material design can be greatly accelerated. Machine-learning methods have been successfully and widely applied to predict the phases of HEAs. However, the accuracy of a single machine-learning (ML) algorithm is not ideal and different ML algorithms may predict different results. These issues hinder the application of ML in material design. In this paper, a hybrid frame for HEAs phase prediction, which combines machine-learning and empirical knowledge, is proposed. First, for the purpose of solving the problem that a sample may be predicted as inconsistent prediction phases by different algorithms, the Dempster-Shafer (DS) evidence theory is adopted to fuse the inconsistent of the predicted phases among different algorithms, and provide a fusion prediction phase with the highest credibility. Second, a conflict-resolution model with high accuracy based on the improved DS evidence theory is proposed. Last, the empirical knowledge criterion is combined with the conflict-resolution model to improve the efficiency and accuracy of the hybrid prediction frame. The 426 different HEAs samples consisting of 180 quaternaries, 189 senaries, and 57 septenaries were collected to validate against the effectiveness of the proposed methods. The experimental results demonstrate the hybrid prediction frame achieves higher accuracy and better performance than single ML algorithm. Keywords: Hybrid model; High-entropy alloys; Phase prediction; DS evidence theory

© 2022 Acta Materialia Inc. Published by Elsevier Ltd. All rights reserved.

1. Introduction

The HEAs have attracted much attention due to their excellent properties [1, 2], such as high thermal stability, high hardness, superior fracture and oxidation resistance features, etc. [3–7]. Differ from traditional alloys, HEAs contain at least five principal elements which offers a large composition variabilities [8], and have potential applications in aerospace, electronics, and other fields [8, 9]. The phase formation of HEAs is a key factor affecting their comprehensive performances, and different phases show different physical and mechanical properties. The phases of HEAs are mainly single-phase solid solution (SS), amorphous (AM), intermetallic compound (IM), and a mixture of SS and IM (SS+IM) [7, 10, 11].

Many methods have been applied to predict the phase formation in HEAs [12]. Computational methods, such as first-principles

calculations, calculation of phase diagram (CALPHAD), density functional theory (DFT) simulations, and so on, have been developed for this purpose. However, these methods are impractical for the material design process due to large amount of computational effort [10, 13]. The parametric method, which was applied to find the criterions for the phase formation [8], can effectively reduce the amount of computational effort and is less time-consuming [14]. Zhang et al. concluded that the solid-solution phase is easily formed in zone S, which is described by the parameters of δ and ΔH_{mix} [15]. Yang et al. proposed a new parameter Ω and found a phenomenon that SS phase can be formed when $\Omega \geq 1.1$ and $\delta \leq 6.6\%$ [16, 17]. Zhang et al. carried out large number of experiments, and further confirmed the effectiveness of this criteria [18]. However, the parametric method is carried out based on a large quantity of trial-error experiments and it only provides limited representations. The composition variabilities of HEAs are still huge, leading to a long time and significant cost for performing experiments [8], which limits its application scope [8, 19].

Owing to the high efficiency and superior learning ability, the ML method has been accepted as an effective method in predict-

* Corresponding authors.

E-mail addresses: houshuai20072@163.com (S. Hou), liuww@dlut.edu.cn (W. Liu).

ing the phase of HEAs. The ML method is based on principles of mathematical statistics to establish a potential mapping relationship between the input parameters and phases of HEAs, and provide a guidance to the material designers [8]. A machine-learning framework proposed by Islam et al. has been applied to predict the phase of HEAs [20]. Three machine-learning algorithms, including k-nearest neighbors (KNN), support vector machine (SVM), and artificial neural network (ANN), were adopted by Huang et al. to predict the HEAs phases including SS, IM, and SS+IM, respectively [10]. To distinguish stable body-centered cubic (BCC) and face-centered cubic (FCC) HEA phases, Li et al. utilized an SVM model to predict the phase of an alloy system composed of 16 metallic elements [21]. By combining the ML model with data processing method, Dai et al. generated new descriptors by linear transformation method and predicted the phases of HEAs on a small dataset [8]. The previous research works have laid important foundation and provided a solid evidence that a lot of manpower and time can be saved if a phase prediction model using machine learning method is established to reflect the mapping relationship between the input parameters and microstructure of HEAs [22–24]. While the general performance of a single ML algorithm is not ideal both in many theoretically and empirically cases [25, 26]. The accuracy of a single ML is limited since it does not take full advantage of diversity learning ability among different machine learning methods. Additionally, different algorithms may predict a sample with different results due to the different learning abilities among ML algorithms [27]. With improper algorithms selected, low prediction accuracy may lead to wrong decisions [28].

In order to improve the prediction accuracy and inconsistent among results, DS evidence theory was proposed by Harvard University mathematician A. P. Dempster and further developed by G. Shafe [29]. Referring to the conflict-resolution ability of DS evidence theory, Zhang et al. proposed a model based on DS evidence theory to predict an essential site in the plant genome. The model built four sub-classifiers based on different features, and the final results were obtained by fusing different outputs of four sub-classifiers [30]. Zeng et al. proposed a multi-model based on a neural network (NN), support vector machine (SVM), and DS evidence theory to recognize the RGB-D (Red Green Blue-Depth) object. The multi-model first extracted RGB (Red Green Blue) and depth features from the NN, and then applied two SVMs with extracted RGB and depth features to obtain two probable recognition results, and finally integrated two probable results by DS evidence theory [31]. Qin et al. proposed a multi-classifier fusion model to detect video flame based on DS evidence theory. The model first chose three different flame features and applied four sub-classifiers to classify a same-flame feature separately. The model then fused four classification results by DS evidence theory to make each feature's preliminary decision. Finally, the preliminary decision results of three features were fused by DS evidence theory [28]. The conflict-resolution ability of DS evidence theory has been successful validated in many fields [32]. However, to the best of author's knowledge, how to deal with the prediction confliction among different algorithms has not been studied in the phase prediction of HEAs.

The above studies have adopted in many fields owing to the good accuracy [33]. However, these ML methods have poor generalization and require large amounts of training samples [34, 35]. In view of the empirical knowledge is characterized by high generalization and strong interpretability, many scholars carried out research on hybrid models, which combine both empirical knowledge model and ML model. For instance, Kong et al. used a hybrid model to predict the adhesion energies between Cu and other materials. The main information was predicted by the empirical knowledge model, and the error of empirical knowledge was compensated by the ML model [36]. Liu et al. proposed a hybrid model to predict the remaining life of membrane fuel cells. A ML model

was applied to predict the trend of long-term degradation at the first stage, and the remaining useful life was predicted by the empirical model and ML model based on the samples obtained in the first stage [35]. These efforts have proved that the hybrid model can take advantage of the high generalization and strong interpretability of empirical knowledge and the superior learning ability of machine learning. However, to the best of the author's knowledge, the hybrid prediction model has not been applied for phase prediction of HEAs.

For the aforementioned reasons, a hybrid prediction frame for the HEAs phase that combines the advantages of empirical knowledge and the ML model is proposed in this draft. First, the DS evidence theory is adopted to deal with the inconsistent of predicted results from different algorithms and provide a prediction result with the highest reliability. Second, a conflict-resolution model with higher accuracy constituted by an improved DS evidence theory is proposed. Last, the empirical knowledge criterions are integrated with the conflict-resolution model to improve the efficiency and accuracy of a hybrid prediction frame.

The rest of this paper is organized as follows: In chapter 2, background knowledge on HEAs and DS evidence theory is introduced. In chapter 3, the hybrid prediction frame for HEAs proposed in this paper is described in details. In chapter 4, numerous experiments are carried out and the effectiveness of the proposed hybrid prediction frame is verified. In chapter 5, a conclusion is given.

2. Methods

Formulas of HEAs parameters and the DS evidence theory are introduced in this section.

2.1. Background theory of HEAs

Due to their excellent properties, HEAs have been widely applied in automobile, aerospace, transport, and other fields [9, 18, 37, 38]. The phases of HEAs greatly affects their properties and the parametric method has played an important role in the phase formation [39]. Meanwhile, formulas of HEAs' parameters are shown as follows:

The δ is the mean square deviation of the atomic sizes of all elements. The formula of δ is shown in Eq (1):

$$\delta = \sqrt{\sum_{i=1}^n c_i \left(1 - \frac{r_i}{\bar{r}}\right)^2} \quad (1)$$

where c_i is the atomic percentage of the i -th component, r_i is the atomic radius and $\bar{r} = \sum_{i=1}^n c_i r_i$ is the average atomic radius.

The ΔS_{mix} is the entropy of mixing an n-element regular solution. The formula of ΔS_{mix} is calculated as Eq (2):

$$\Delta S_{mix} = -R \sum_{i=1}^n c_i \ln c_i \quad (2)$$

where R is the gas constant.

The ΔH_{mix} is the enthalpy of mixing for a multi-component alloy system with n elements. The formula of ΔH_{mix} is shown as Eq (3):

$$\Delta H_{mix} = \sum_{i=1, i < j}^n 4H_{ij} c_i c_j \quad (3)$$

where H_{ij} is the regular solution interaction parameter between the i -th and j -th element. c^i and c^j are the atomic percentage of the i -th and j -th component.

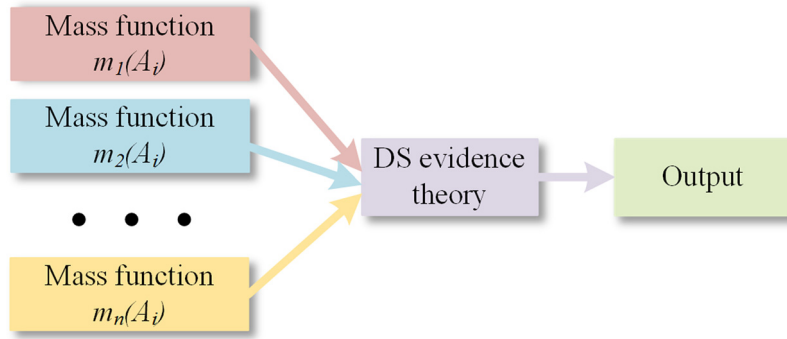


Fig. 1. Schematic diagram of DS evidence theory.

The VEC is the average valence electron concentration, and the formula of VEC is calculated as Eq (4):

$$VEC = \sum_{i=1}^n c_i VEC_i \quad (4)$$

where VEC_i is the average valence electron concentration of i -th element.

The $\Delta\chi$ is the electronegativity, and the formula of $\Delta\chi$ is calculated as Eq (5):

$$\Delta\chi = \sqrt{\sum_{i=1}^n c_i (\chi_i - \bar{\chi})^2} \quad (5)$$

where χ_i is the Pauling electronegativity of i -th element, and $\bar{\chi} = \sum_{i=1}^n c_i \chi_i$ is the average Pauling electronegativity.

The Ω is a parameter which combines the effect between ΔS_{mix} and ΔH_{mix} , and the formula of Ω is calculated as Eq (6):

$$\Omega = \frac{T_m \Delta S_{mix}}{|\Delta H_{mix}|} \quad (6)$$

where T_m^i is the melting point of the i -th component of the alloy, and T_m is calculated by the rule of mixtures $T_m = \sum_{i=1}^n c_i T_m^i$.

2.2. DS evidence theory

DS evidence theory was proposed by mathematician, A. P. Dempster, and it was further improved by his student, G. Shafer. It shows good conflict-resolution ability to provide reliable results [29, 40]. The DS evidence theory is schematically shown in Fig. 1:

According to the DS evidence theory, the frame of discernment Θ is a finite nonempty set of hypotheses [41]. Herein, A_i refers to the hypothesis, and $m(A_i)$ is the mass function of A_i , also called the basic probability assignment function. It represents the probability distribution of hypothesis [28]. The $m(A_i)$ is a belief measurement of the final result, and it satisfies two conditions as Eq (7) and Eq (8):

$$m(\emptyset) = 0 \quad (7)$$

$$\sum_{A \subset \Theta} m(A_i) = 1 \quad (8)$$

where $\emptyset$ is an empty set.

The credibility (Bel) function is defined as Eq (9):

$$Bel(A_i) = \sum_{B \subset A, B \neq \emptyset} m(B_i) \quad (9)$$

The $Bel(A_i)$ is interpreted as the total amount of belief in the hypothesis A_i , which can be considered as the credibility of hypothesis A_i [42].

According to Dempster's combination rule, the max value of the combination rules is taken as the final result with highest credibility [43]. The combination of n masses of hypotheses is defined as Eq (10):

$$(m_1 \oplus m_2 \oplus \dots \oplus m_n)(A) = \frac{1}{K} \sum_{A_1 \cap A_2 \cap \dots \cap A_n = A} m_1(A_1) \cdot m_2(A_2) \cdot \dots \cdot m_n(A_n) \quad (10)$$

where K is the normalization factor. It refers to high conflict when $K = 1$ or infinitely approaching 1 [29]. The formula of K is calculated by Eq (11):

$$K = \sum_{A_1 \cap A_2 \cap \dots \cap A_n \neq \emptyset} m_1(A_1) \cdot m_2(A_2) \cdot \dots \cdot m_n(A_n) = 1 - \sum_{A_1 \cap A_2 \cap \dots \cap A_n = \emptyset} m_1(A_1) \cdot m_2(A_2) \cdot \dots \cdot m_n(A_n) \quad (11)$$

Due to its excellent conflict resolution ability, the DS evidence theory has been adopted to deal with conflict challenges in many fields and provide several reliable results [30]. However, the assumption of traditional DS evidence theory is that each algorithm has the same reliability. Many factors, such as different feature reliability of algorithms and differences of performance among algorithms, would affect the reliability of the conflict-resolution result. Additionally, if the conflation of results among different algorithms is significant, the DS evidence theory combination rule will be invalid and results obtained will be counterintuitive. The high-conflict cases would guide algorithms toward wrong decision. Assigning different weights, depending on the algorithms' performance, can greatly improve the reliability of results and minimize the negative effect of high conflict [44, 45].

3. Hybrid-phase prediction frame for HEAs Alloys

In this section, a hybrid prediction frame (HPF) is proposed based on the empirical knowledge model and conflict-resolution model.

3.1. Empirical knowledge model

Based on many experiments, Zhang et al. proposed a SS phase-formation criterion. The SS phase-formation criterion is that the SS phase will be formed if the sample falls within the scope of $\Omega \geq 1.1$ and $\delta \leq 6.6\%$ [46]. Similar research was also carried out, and the experimental results further confirmed the effectiveness of the SS phase-formation criterion to identify the SS phase [18].

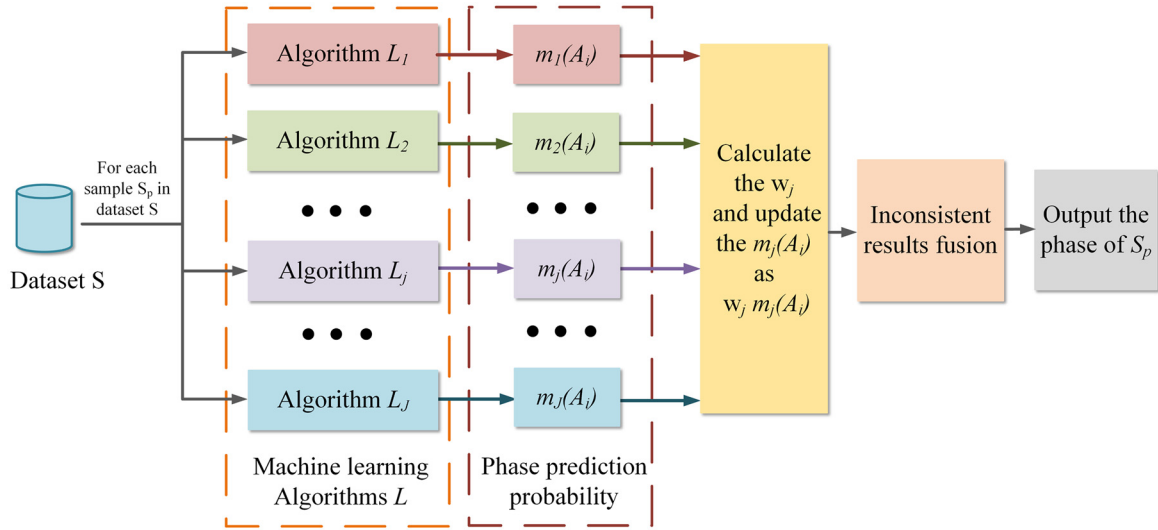


Fig. 2. Schematic diagram of conflict-resolution model.

To improve efficiency, the SS phase-formation criterion is integrated into the hybrid prediction frame. In the prediction process, the dataset $X = \{X_1, X_2, \dots, X_q, \dots, X_Q\}$ is divided into two sub-sets based on the criterion. Samples that fall within the scope form a dataset $E = \{E_1, E_2, \dots, E_i, \dots, E_l\}$, and samples in the rest of the dataset form another dataset $S = \{S_1, S_2, \dots, S_p, \dots, S_P\}$. Samples in dataset $E = \{E_1, E_2, \dots, E_i, \dots, E_l\}$ will be marked as SS phase directly.

3.2. Conflict-resolution model

To resolve inconsistent prediction results by different algorithms, a conflict-resolution model (CRM) based on improved DS evidence theory is proposed. The conflict-resolution model contains five different ML algorithms, including support vector machine (SVM), K-nearest neighbor (KNN), decision tree (DT), logistics regression (LR), and random forest (RF). The improved DS evidence theory based on a standard deviation is adopted to resolve the confliction results among algorithms. The $m_j(A_i)$ is the output from ML algorithm L_j . The weight w_j for each algorithm L_j is adaptively calculated based on standard deviation. The combination results are calculated based on the updated $m_j(A_i)'$, where $m_j(A_i)'$ is updated by $m_j(A_i)$ and w_j .

The model is schematically shown in the Fig. 2:

Steps of the conflict-resolution model are shown as follows:

Input: HEAs test dataset $S = \{S_1, S_2, \dots, S_p, \dots, S_P\}$, ML algorithms $L = \{L_1, L_2, \dots, L_j, \dots, L_J\}$.

Step 1: Train the ML algorithms $L = \{L_1, L_2, \dots, L_j, \dots, L_J\}$.

Step 2: Output the $m_j(A_i)$ by algorithm in L for sample S_p . Calculate the standard deviation D_j of each ML algorithm L_j as Eq (12):

$$D_j = \sqrt{\sum_{i=1}^N \left(m_j(A_i) - \frac{1}{N} \right)^2}, \quad 1 < i < N, 1 < j < J \quad (12)$$

where $m_j(A_i)$ represents the probability that sample S_p belongs to the i -th phase by j -th algorithm; N represents the total number of alloy phases; M represents the total number of ML algorithms.

Step 3: Calculate the weight w_j of each algorithm L_j , the weight w_j is defined as Eq (13):

$$w_j = \frac{D_j}{\sum_{j=1}^J D_j}, \quad 1 < j < J \quad (13)$$

Step 4: Update $m_j(A_i)'$ by the weight w_j as Eq (14):

$$m_j(A_i)' = w_j \cdot m_j(A_i) \quad (14)$$

Step 5: Calculate the combination result of $m(A_i)$ as Eq (15):

$$m(A_i) = \frac{\sum_{i=1}^N A_i=A \prod_{j=1}^J m_j(A_i)'}{1 - \sum_{i=1}^N A_i=A \prod_{j=1}^J m_j(A_i)'} \quad (15)$$

Step 6: Calculate the credibility (Bel) of A_i as Eq (16):

$$Bel(A_i) = \frac{m(A_i)}{\sum_{i=1}^N m(A_i)}, \quad 1 \leq i \leq N \quad (16)$$

Step 7: Assign the phase with the highest credibility to the sample as Eq (17):

$$Bel(A) = \max(Bel(A_1), Bel(A_2), \dots, Bel(A_N)) \quad (17)$$

Beyond this common phenomenon that different algorithms predict a sample as different phases, different algorithms also may predict a sample as the same phase in a few ideal cases. In these ideal cases, the sample is directly recognized as this phase.

3.3. Hybrid prediction frame for HEAs phase

A hybrid-phase prediction frame constituted by an empirical knowledge model and a conflict-resolution model is proposed in this section. The empirical knowledge model is based on the SS phase-formation criterion, while the conflict-resolution model is based on improved DS evidence theory. The schematic diagram of the hybrid prediction frame is shown in Fig. 3:

As shown in Fig. 3, steps of the proposed hybrid prediction frame can be listed as follows:

Input: HEAs dataset $X = \{X_1, X_2, \dots, X_q, \dots, X_Q\}$, ML algorithms $L = \{L_1, L_2, \dots, L_j, \dots, L_J\}$.

Step 1: Train the ML algorithms $L = \{L_1, L_2, \dots, L_j, \dots, L_J\}$.

Step 2: Divide the dataset X into two sub-sets, samples that fall within the scope of SS formation criterion as $E = \{E_1, E_2, \dots, E_i, \dots, E_l\}$, and the rest of the samples as $S = \{S_1, S_2, \dots, S_p, \dots, S_P\}$.

Step 3: Mark samples in dataset $E = \{E_1, E_2, \dots, E_i, \dots, E_l\}$ as SS phase.

Step 4: Predict S_p in dataset S by algorithms $L = \{L_1, L_2, \dots, L_j, \dots, L_J\}$ and output the $m_{jp}(A_i)$ of every sample S_p .

Step 5: For every sample S_p , calculate the weight w_{jp} of sample S_p by algorithms L_j by Eq (13). Calculate the combination result $m_{jp}(A_i)$ by Eq (15). Calculate the $Bel_p(A_i)$ by Eq (16). Assign the phase with highest credibility of sample S_p by Eq (17).

Step 6: Output the prediction results of dataset X .

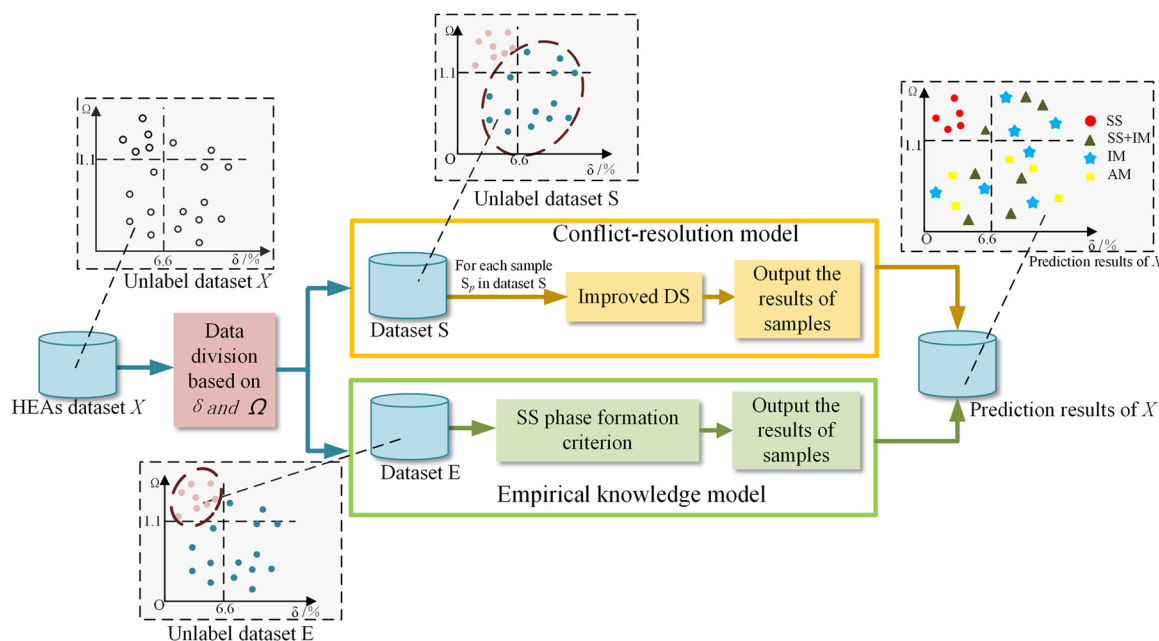


Fig. 3. Schematic diagram of the hybrid prediction frame.

4. Results and discussion

In this paper, the proposed hybrid prediction frame and support vector machine (SVM), K-nearest neighbor (KNN), decision tree (DT), logistics regression (LR), and random forest (RF) are compared in this section. In order to guarantee the performance of ML algorithms, the grid search method is applied to optimize the hyperparameters of algorithms [47]. The optimal hyperparameters of algorithms are shown as follows after the systematic study: the hyperparameter 'criterion' of DT is set as 'gini', and hyperparameter 'max depth' is '4'. The hyperparameter 'solver' of LR is 'newton-cg'. The hyperparameter 'k' of KNN is '4'. The hyperparameter 'criterion' of RF is 'gini'. And the hyperparameter 'gamma' of SVM is '0.1'. Compared with the hyperparameters listed above, other hyperparameters with little effect on the accuracy are set to default values.

The software used is Python 3.7.4.

4.1. Experimental dataset

Dataset X in this paper is collected from Refs [48–51], which contains four collections of HEAs. The data preprocessing is carried out in HEAs dataset X, which includes parameter value calculation, data supplementary, data deduplication, and noise sample elimination. After the data preprocessing, the HEAs dataset including 426 samples is comprised of 180 quinary samples, 189 senary samples, and 57 septenary samples.

The phases are divided into four categories – solid solution phase (SS), intermetallic phase (IM), solid solution, intermetallic phase (SS+IM), and amorphous phase (AM) [10, 11]. The dataset contains 21 elements (Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Y, Zr, Nb, Mo, Pd, Ag, Hf, Ta, Os, Au, Al). As shown in Fig. 4, Fe, Cr, Co, Ni, Cu, and Al appear in more than 200 samples, while Pd, Au, and Sc appear twice, Ag and Os just are observed one time.

Fig. 5 shows a snapshot of the first five rows of the data in the Pandas DataFrame format. The first column identifies an HEA and the remaining six columns describe the corresponding parameters.

More details of parameters including their maximum values, minimum values, and parameter descriptions are listed in Table 1.

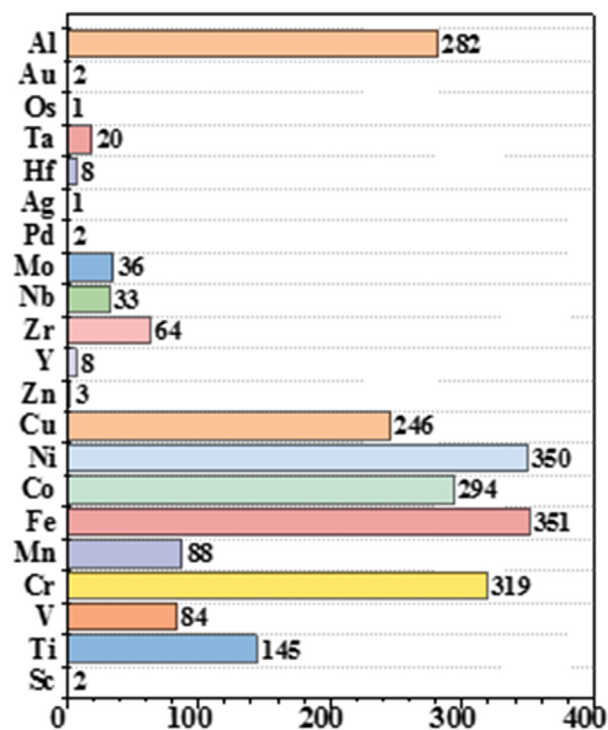


Fig. 4. The summarization of frequency of occurrence of elements.

In order to help understand the data distribution and the relationship between these parameters, a 6×6 scatterplot matrix is plotted in Fig 6. Moreover, in order to quantitatively describe the relationship between parameters, the Pearson correlation coefficient P between parameters x and y is defined as Eq (18):

$$P_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (18)$$

where $\bar{x}$ and $\bar{y}$ are the mean values of parameters x and y , respectively. The $P_{xy} = 1$ and $P_{xy} = -1$ represent a perfect positive cor-

	Alloy	ΔS_{mix}	ΔH_{mix}	δ (%)	$\Delta \chi$	VEC	Ω	Phase
0	AlCoCrCuFe	13.380867	-2.560000	5.108760	0.118423	7.40	8.389176	SS+IM
1	AlCo2CuFeNi	12.975746	-5.222222	5.165383	0.103722	8.33	3.873260	SS+IM
2	AlCo3CuFeNi	12.263784	-4.979592	4.896273	0.097311	8.43	3.911645	SS
3	CoFeMnTiVZr	14.896688	-18.666667	9.464926	0.185712	6.17	1.506960	SS+IM
4	CoFeNiSi0.5	11.238732	-19.428571	2.067322	0.031944	7.75	1.014874	SS+IM

Fig. 5. Pandas snapshot of the first five samples of the data in this paper. Units of ΔS_{mix} and ΔH_{mix} are $\text{kJ}\cdot\text{mol}^{-1}$ and $\text{kJ}^{-1}\cdot\text{mol}^{-1}$, respectively.

Table 1
Snapshot of parameters.

Number	Parameter	Maximum value	Minimum value	Parameter description
1	ΔS_{mix}	16.1782	6.4654	thermodynamic parameter
2	ΔH_{mix}	30.2400	-34.1000	chemical parameter
3	δ	20.3800	0.2100	electronic parameter
4	$\Delta \chi$	0.3324	0.03167	electronic parameter
5	VEC	10.6700	0.3400	electronic parameter
6	Ω	283.5023	0.5655	chemical thermodynamic parameter

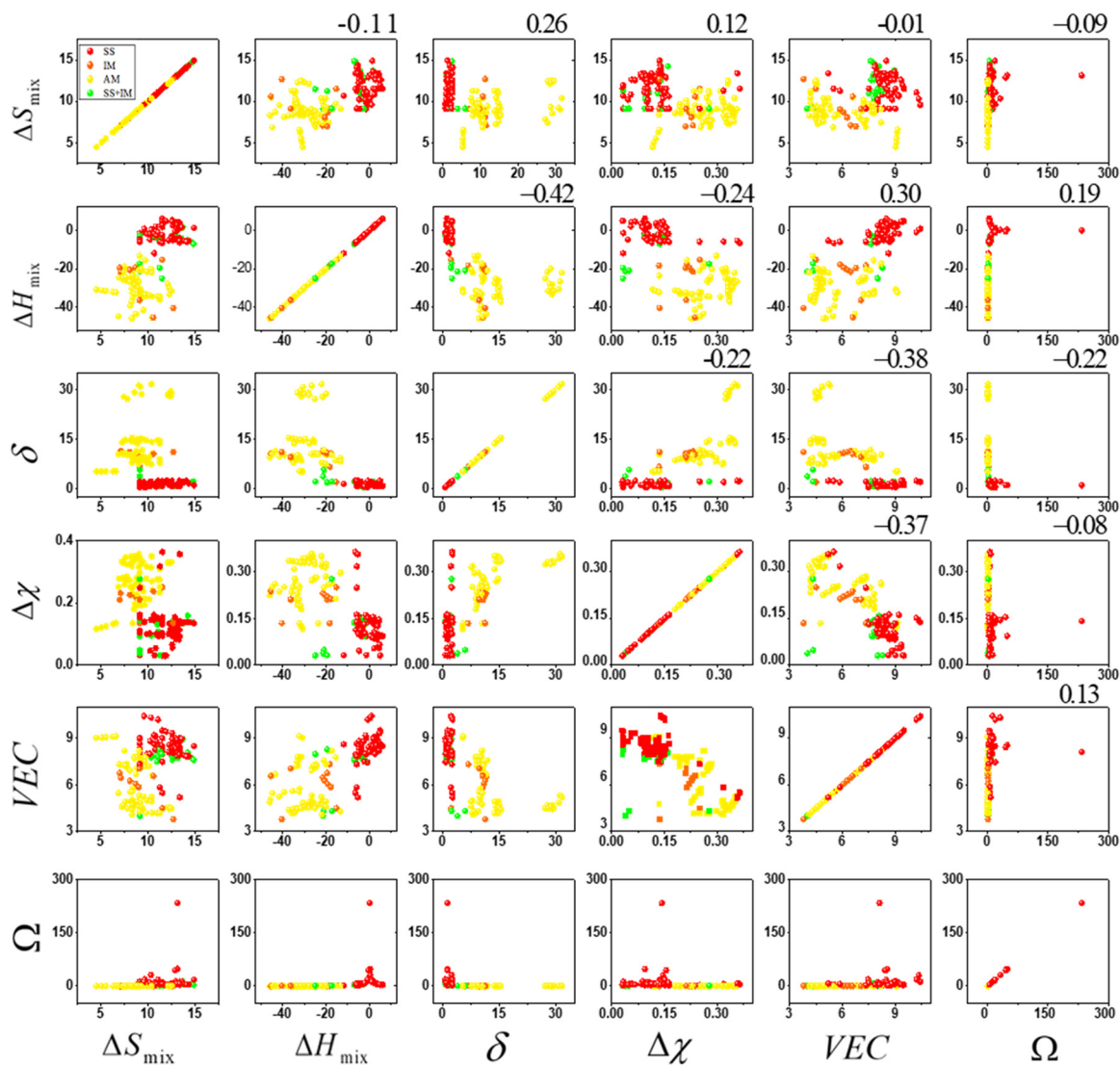


Fig. 6. Scatterplot is plotted according to different parameters. The point in a scatterplot corresponds to a sample, and different colors represent different phases. The blue represents SS phase, red represents IM phase, green represents AM phase, and yellow represents SS+IM phase.

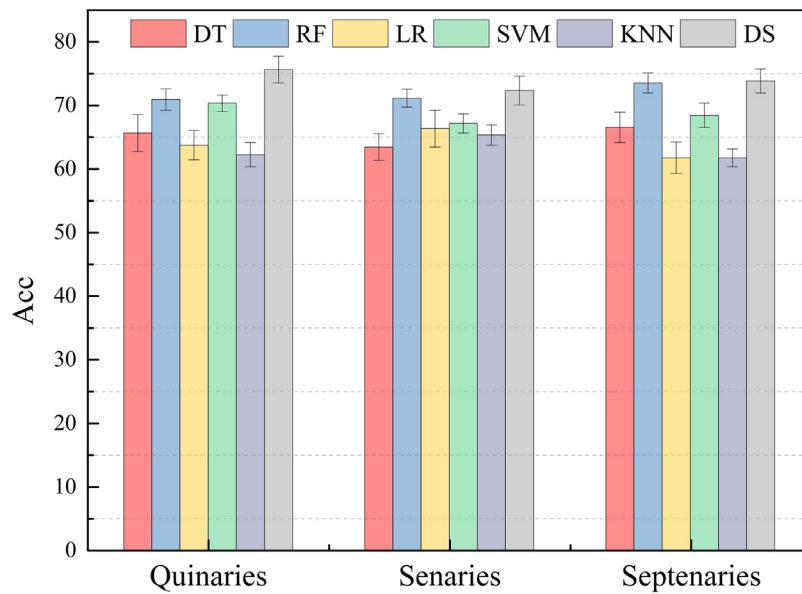


Fig. 7. The accuracy comparison of DS evidence theory and other algorithms.

Table 2
Correlations between the parameters.

	ΔS_{mix}	ΔH_{mix}	δ	$\Delta \chi$	VEC	Ω
ΔS_{mix}	1	-0.11	0.26	0.12	-0.01	-0.09
ΔH_{mix}	-0.11	1	-0.42	-0.24	0.30	0.19
δ	0.26	-0.42	1	-0.22	-0.38	-0.22
$\Delta \chi$	0.12	-0.24	-0.22	1	-0.37	-0.08
VEC	-0.01	0.30	-0.38	-0.37	1	0.13
Ω	-0.09	0.19	-0.22	-0.08	0.13	1

relation and a negative correlation between parameters x and y . Table 2 lists the computed correlation matrix between parameters, which is also marked in the right upper corner of the plot in Fig. 6.

As observed from the scatterplot matrix plot, one parameter or two parameters and a simple linear function cannot draw a clear boundary to distinguish these phases. Moreover, the range of Pearson coefficients correlation P is between -0.42 to 0.30 , and ten of the 15 independent Pearson coefficients correlations are negative while the other five Pearson coefficients correlations are positive. As a result, there is no strong correlation between any two parameters.

In order to avoid the overfitting of ML [10, 52], the k-fold cross-validation method is applied in this experiment. The k value is set as 5 refers to the method proposed by Jung et al. [52]. First the dataset X in this paper is divided into five equal subsets. In each cross-validation process, one subset is selected as the test set and the remaining four subsets are the training sets. Each subset is selected as a test set once. In five cross-validation processes, the five cross-validation methods are repeated six times and the final accuracy is the average value of 30 times accuracies of the cross-validation process.

4.2. Comparison between DS evidence theory and other algorithms

To verify the conflict-resolution capability of DS evidence theory under different material systems, a set of comparison experiments among DS evidence theories DT, RF, LR, SVM, and KNN, are carried out. The selected algorithms refer to the following literature [8, 10, 28]. The comparison is shown in Fig. 7.

The details of experiment results are shown in Table 3 where the first column is the order number of the material system and the remaining columns are different experiment methods.

As shown in Fig. 7 and Table 3, the accuracy of DS evidence theory is higher than DT, RF, LR, SVM, and KNN. The reason may be that DS evidence theory can combine the advantages of these algorithms and merge their outputs to provide a more reliable result [53]. Meanwhile it can also resolve the inconsistent prediction among different algorithms by its conflict-resolution ability, and provide more reliable prediction results. RF has higher accuracy than DT, LR, SVM, and KNN because it belongs to the ensemble algorithm. It can fuse multiple results of different weak algorithms showing a better learning ability. Zhang et al. adopted different machine-learning algorithms to predict the phase of HEAs, and also found that RF has higher accuracy than DT, LR, and SVM. Their experimental results is in line with the results in this paper [54]. However, RF cannot effectively resolve the inconsistent of prediction results of multiple weak algorithms. While DS evidence theory has excellent conflict-resolution ability. Moreover, accuracies of DT and SVM are higher than LR and KNN. The reason may be because SVM and DT have high adaptability, they require a small amount of training sample, and have good prediction ability in a small dataset [10, 55]. Meanwhile, LR is a simple and linear model and is hard to fit with complex data distribution, thus leading to low accuracy. KNN depends highly on training data, and cannot well deal with high-dimensional data due to huge amount of calculation, thus resulting a poor accuracy [10].

To illustrate the conflict resolution ability of DS evidence theory, partial samples are listed in Table 4. Due to large amounts of samples, only some representative samples are listed.

As observed from the Table 4, the first column lists different HEAs. Columns 2 to 7 are the predicted results of DT, RF, LR, SVM, KNN, and DS evidence theory. Column 8 is the true phase/phases of the samples and the final column refers to the reference of HEAs sample. The samples listed in Table 4 are predicted by different algorithms, the predicted phases of these samples by different algorithms are inconsistent. For example, the true phase of $\text{Al}_{0.8}\text{CrCuFeNi}$ is the SS phase. It is predicted as SS phase by LR and SVM, and as SS+IM phase by other algorithms. Meanwhile, DS evidence theory can successfully resolve these conflict prediction results by its conflict resolution ability,

Table 3
Experiment results of DS evidence and other algorithms.

Material Systems	DT	RF	LR	SVM	KNN	DS
Quinaries	65.6 ± 2.9	70.9 ± 1.7	63.7 ± 2.3	70.3 ± 1.3	62.2 ± 1.9	75.6 ± 2.1
Senaries	63.4 ± 2.1	71.1 ± 1.4	66.3 ± 2.9	67.1 ± 1.5	65.3 ± 1.6	72.3 ± 2.3
Septenaries	66.5 ± 2.4	73.5 ± 1.6	61.7 ± 2.5	68.4 ± 1.9	61.7 ± 1.4	73.8 ± 1.9

Table 4
A snapshot of partial prediction results for HEAs test sample.

Alloy	DT	RF	LR	SVM	KNN	DS	True Phase	Ref.
Al _{0.8} CrCuFeNi	SS+IM	SS+IM	SS	SS	SS+IM	SS	SS	[49]
Al _{0.6} CoFeNiTi _{0.4}	SS+IM	IM	SS	IM	SS+IM	SS	SS+IM	[49]
AlCu _{0.2} Li _{0.5} MgZn _{0.5}	SS+IM	SS+IM	SS	IM	SS+IM	SS+IM	SS+IM	[49]
ZrHfTiCuCo	AM	SS	SS+IM	SS	SS	AM	AM	[48]
FeCoCuNiSn _{0.07}	SS	SS+IM	IM	SS+IM	SS	SS+IM	SS	[50]
Al _{0.02} CoCrFeMnNi	SS+IM	SS	SS+IM	SS	SS+IM	SS	SS	[49]
TiCoCrNiCuAlY _{0.8}	AM	AM	SS+IM	IM	IM	IM	IM	[50]
CuFeNiTiVZr	AM	AM	SS+IM	AM	IM	AM	AM	[51]
Ti _{1.8} CoCrFeNiCuAl _{0.5}	SS+IM	SS+IM	SS+IM	SS	SS	SS+IM	SS+IM	[50]
CoCuFeNiTiVZr	IM	AM	SS+IM	AM	IM	IM	AM	[51]

Table 5
Experiment results between DS evidence and Improved DS.

Material Systems	DS	Improved DS
Quinaries	75.6 ± 2.1	78.2 ± 1.7
Senaries	72.3 ± 2.3	75.8 ± 2.1
Septenaries	73.8 ± 1.9	76.3 ± 1.6

and it can output the correct phase. It is noteworthy that DS evidence theory also resolves inconsistent prediction phases and predicts the correct phase for HEAs samples of AlCu_{0.2}Li_{0.5}MgZn_{0.5}, ZrHfTiCuCo, Al_{0.02}CoCrFeMnNi, TiCoCrNiCuAlY_{0.8}, CuFeNiTiVZr, and Ti_{1.8}CoCrFeNiCuAl_{0.5}. However, DS evidence theory is not always correct. For example, DS evidence theory did not successfully predict the true phase of Al_{0.6}CoFeNiTi_{0.4}: SS+IM. Similarly, DS evidence theory predicts the wrong results for FeCoCuNiSn_{0.07} and CoCuFeNiTiVZr. The reason refers to the high conflict among different algorithms. The DS evidence theory tends to generate counterintuitive results if the conflict is high [41]. The accuracy can be improved if the high-conflict issue is solved effectively.

4.3. Comparison between DS evidence theory and improved DS evidence theory

To solve the high-conflict issue, an improved DS evidence theory based on standard deviation method is proposed. In order to compare the conflict-resolution capability of DS evidence theory and improved DS evidence theory, a set of comparison experiments were performed. The accuracy comparison of DS evidence theory and improved DS evidence theory is shown in Fig. 8.

Details of experiment results between DS evidence theory and improved DS evidence theory are shown in Table 5, where the first column is the order number of different material systems. Column 2 is the experiment results of DS and Column 3 is the experiment results of improved DS.

As shown in Fig. 8 and Table 5, the accuracy of the improved DS evidence theory is higher than the DS evidence theory. This is because the improved DS evidence can assign different weights, which is decided by the algorithm's performance, to minimize conflict. In order to observe the conflict-resolution capability between the DS evidence theory and the improved DS evidence theory, a snapshot of partial prediction results for the HEAs test sample are

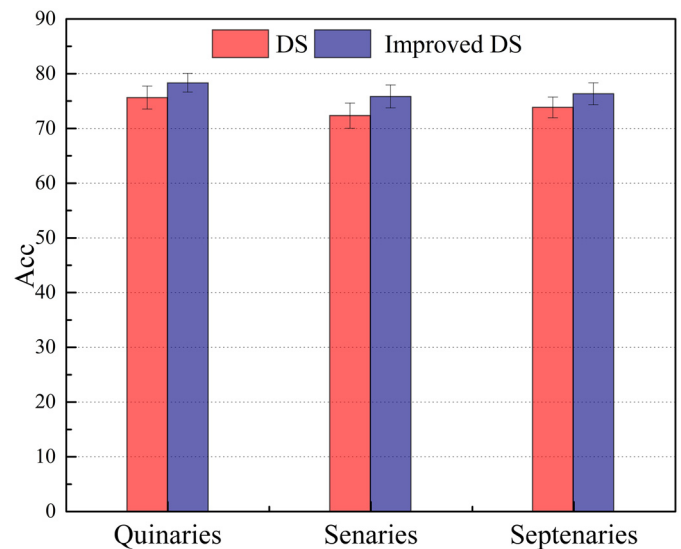


Fig. 8. The accuracy comparison between DS evidence theory and improved DS evidence theory.

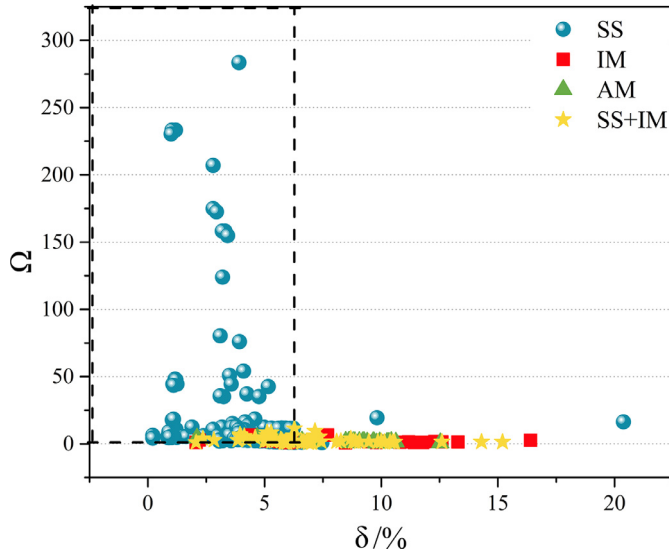
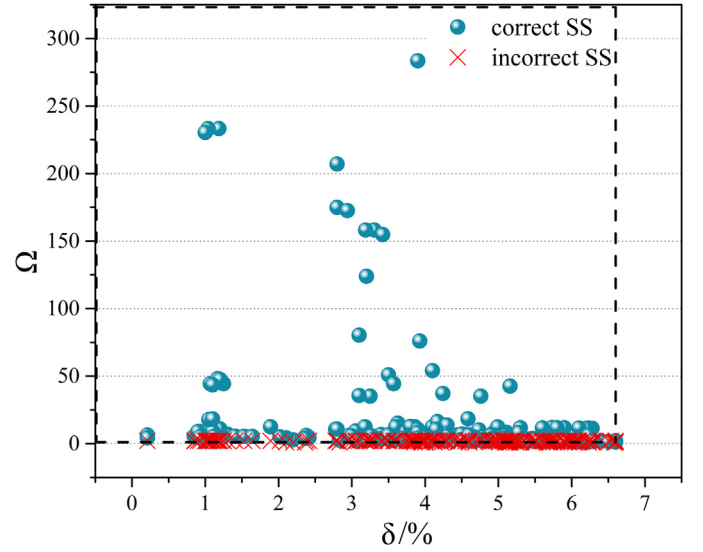
listed in Table 6. Due to the large amounts of samples, only some representative samples are listed in Table 6.

Column 1 is different HEAs, Columns 2 to 8 are the prediction results of DT, RF, LR, SVM, KNN, DS evidence theory and improved DS evidence theory, respectively, and Column 9 is the real phases and the final column refers to the reference of HEAs sample. The predicted results from DT, RF, LR, SVM, and KNN are inconsistent, while the DS evidence theory and the improved DS evidence theory have shown their conflict-resolution capability. It is noteworthy that the DS evidence theory predicts a wrong phase for Al_{0.6}CoFeNiTi_{0.4} due to high conflict among DT, RF, LR, SVM, and KNN, while the predicted result of the improved DS evidence theory is correct. The same phenomenon is also observed in FeCoCuNiSn_{0.07}, Al₁₀Cu₂₀Ni₈Ti₅Zr₅₇, TiZrCuNiBe, TaNbVTiAl_{1.0}, ZrTiVCuNiBe, Ti_{0.5}Co_{1.5}CrFeNi_{1.5}Mo_{0.1}, and CoCuFeNiTiVZr. Experimental results show that the improved DS evidence theory has better conflict-resolution capability and can provide more reliable results.

Table 6

A snapshot of partial prediction results for HEAs test samples.

Alloy	DT	RF	LR	SVM	KNN	DS	Improved DS	True Phase	Ref.
Al _{0.8} CrCuFeNi	SS+IM	SS+IM	SS	SS	SS+IM	SS	SS	SS	[49]
Al _{0.6} CoFeNiTi _{0.4}	SS+IM	IM	SS	IM	SS+IM	SS	SS+IM	SS+IM	[49]
AlCu _{0.2} Li _{0.5} MgZn _{0.5}	SS+IM	SS+IM	SS	IM	SS+IM	SS+IM	SS+IM	SS+IM	[49]
ZrHfTiCuCo	AM	SS	SS+IM	SS	SS	AM	AM	AM	[48]
FeCoCuNiSn _{0.07}	SS	SS+IM	IM	SS+IM	SS	SS+IM	SS	SS	[50]
Al ₁₀ Cu ₂₀ Ni ₈ Ti ₅ Zr ₅₇	AM	AM	SS	AM	AM	IM	AM	AM	[51]
TiZrCuNiBe	AM	SS	SS	AM	SS	SS	AM	AM	[48]
TaNbVTiAl _{1.0}	SS+IM	SS+IM	SS	SS	SS	SS+IM	SS	SS	[48]
Al _{0.02} CoCrFeMnNi	SS+IM	SS	SS+IM	SS	SS+IM	SS	SS	SS	[49]
TiCoCrNiCuAlY _{0.8}	AM	AM	SS+IM	IM	IM	IM	IM	IM	[50]
CuFeNiTiVZr	AM	AM	SS+IM	AM	IM	AM	AM	AM	[51]
ZrTiVCuNiBe	AM	IM	AM	IM	IM	IM	AM	AM	[50]
Ti _{0.5} Co _{1.5} CrFeNi _{1.5} Mo _{0.1}	SS+IM	SS	SS+IM	SS	SS+IM	SS	SS+IM	SS+IM	[50]
Ti _{1.8} CoCrFeNiCuAl _{0.5}	SS+IM	SS+IM	SS+IM	SS	SS	SS+IM	SS+IM	SS+IM	[50]
CoCuFeNiTiVZr	IM	AM	SS+IM	AM	IM	IM	AM	AM	[51]

**Fig. 9.** A $\delta - \Omega$ scatter plot showing the labels of the phases in dataset X.**Fig. 10.** A $\delta - \Omega$ scatter plot showing how samples fall in the scope of empirical knowledge criterion.

4.4. Comparison between the improved DS evidence theory and the HPF

In order to verify the performance of the hybrid prediction frame (HPF) proposed in this paper, a set of comparison experiments between the improved DS evidence theory and the HPF theory are carried out. Before comparing the experiments' results, Fig. 9 shows a scatter plot based on the $\delta - \Omega$ coordinate system that displays the distribution of HEAs phases. The result is consistent with the subplot in Fig. 6 using the same coordinate system.

As shown in Fig. 9, SS phases mostly fall within the upper left corner based on the $\delta - \Omega$ coordinate system, and only small amounts of samples are located elsewhere. This means the $\delta - \Omega$ coordinate system can identify the SS phase. In other words, the $\delta - \Omega$ coordinate system can be considered as the formation criterion for the SS phase. The specific distribution of samples falling within the scope of empirical knowledge are shown in Fig. 10 for further observation.

As shown in Fig. 10, the virtual line draws the scope of empirical knowledge criterion. The 295 samples fall within the scope, 237 of them are SS phase, and the rest are other phases. More than 80% of the samples are labeled as SS phase correctly, which means the accuracy of the empirical knowledge is up to 80%. The results verify the effectiveness of empirical knowledge integrated in the hybrid prediction frame.

Table 7

Accuracy comparison between improved DS and HPF.

Material Systems	Improved DS	HPF
Quinaries	78.2 $\pm$ 1.7	87.8 $\pm$ 2.1
Senaries	75.8 $\pm$ 2.1	86.7 $\pm$ 1.7
Septenaries	76.3 $\pm$ 1.6	83.3 $\pm$ 1.4

A set of comparison experiments between the improved DS evidence theory and the HPF theory are performed. The comparison of improved DS evidence theory and HPF theory is shown in Fig. 11.

The details of experimental results between improved DS evidence theory and HPF theory are shown in Table 7 where the first column, the second column, and the third column are the order number of different material systems, the experiment results of improved DS, and the experiment results of HPF theory, respectively.

As shown in Fig. 11 and Table 7, the HPF theory achieved higher accuracy than the improved DS evidence theory. The reason may be that HPF theory takes advantages of the strong learning ability and conflict-resolution ability of the conflict-resolution model and the high generalization of the empirical knowledge model. The HPF theory can achieve desirable accuracy with less or no training samples. Moreover, the interpretability of HPF is also strengthened

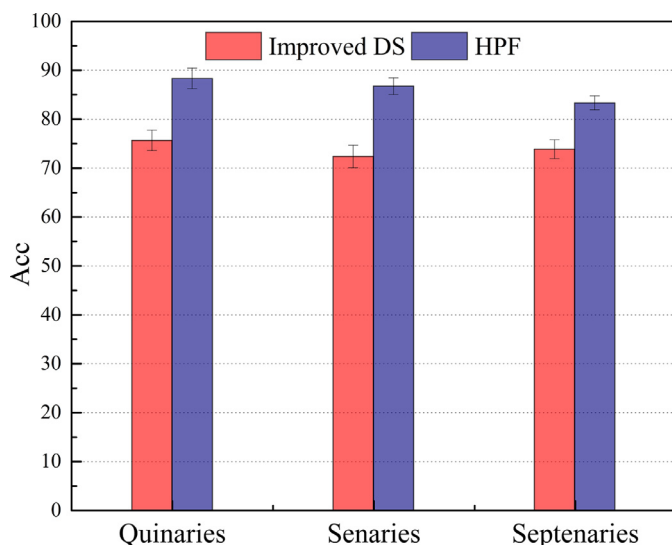


Fig. 11. The accuracy of improved DS evidence theory and HPF.

by integrating the empirical knowledge model. The performance of hybrid model is better than single machine learning method, which also agrees with previous studies in several literatures for the applications in environmental monitoring, automatic control [28, 33, 56].

5. Conclusion

In conclusion, a hybrid prediction frame, which combines a conflict-resolution model and an empirical knowledge model, for the HEAs phase prediction is proposed in this report. A dataset containing 426 HEA samples is utilized to verify the effectiveness of the proposed model. First, the DS evidence theory reached a higher accuracy ($>72.3\%$) than DT, RF, LR, SVM, and KNN. This is because the conflict-resolution capability of the DS evidence theory can effectively resolve the conflictions among different results and provide a result with higher accuracy for each HEAs sample. Second an improved DS evidence theory based on a standard deviation method is proposed to further resolve the high-conflict problem among different results. Experiment results show the improved DS evidence theory has higher accuracy ($>75.8\%$) than the original DS evidence theory. Finally, the proposed HPF method is proposed to take the advantage of strong learning ability of the ML model and high generalization of empirical knowledge. Experiment results show the HPF method can attain higher accuracy ($>83.3\%$) than the improved DS evidence theory, which validates the effectiveness of integrating the empirical knowledge model into the prediction process.

This paper provides solid theoretical basis for the high-conflict challenge in the prediction process of HEAs. The method proposed in this paper cannot only be applied to the HEAs field, but also has certain applicability in modern engineering application, biology, mechanical automation and other fields [30, 33, 34, 56]. This method can also lay a foundation of machine learning in material design.

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 work was supported in part by the National Natural Science Foundation of China (No. 52175455), the National Key Research and Development Program of China (No. 2020YFB1709903), the Science and Technology Innovation Fund of Dalian (No. 2020JJ26GX040), the Equipment Pre-research Foundation (No. 80923010401), the Fundamental Research Funds for the Central Universities (No. DUT20JC19), Science and Technology Planning Project of Hebei Province (21350101D), Iron and Steel Joint Foundation of Hebei Province (E2020402016) and Handan City Science and Technology Research and Development Projects (19422101008-27), National Science Foundation under Award (No. 1943445), General project of Hebei Natural Science Foundation (A2020402013).

References

- [1] C.W. Zhe, Advances of new technologies of materials and modern methods of material testing, PTCA(PART A:Phys. Test.) (2002) 466–472.
- [2] Z. Y. J. Dong, Y.Z. X., Research progress on high entropy alloy with FCC structure, Foundry Technol. (2019) 1008–1011.
- [3] C. Chen, R. Wei, Effect of cooling rate on the phase structure and magnetic properties of Fe_{26.7}Co_{28.5}Ni_{28.5}Si_{4.6}B_{8.7}P₃ high entropy alloy, 2017, 184–186.
- [4] B.T. Eleti, R. L. Zhao, Hot deformation behavior of CoCrFeMnNi FCC high entropy alloy, Mater. Chem. Phys. (2018) 176–186.
- [5] Y.Tian L.C., Y. Shen, Microstructure and corrosion property of CrMnFeCoNi high entropy alloy coating on Q235 substrate via mechanical alloying method, Surf. Interfac. (2019) 135–140.
- [6] W. Zhang, P.K. Liaw, Y. Zhang, Science and technology in high-entropy alloys, Sci. China Mater. (2018) 2–22.
- [7] K. Kaufmann, K.S. Vecchio, Searching for high entropy alloys: a machine learning approach, Acta Mater. (2020) 178–222.
- [8] D. Dai, T. Xu, X. Wei, G. Ding, Y. Xu, J. Zhang, H. Zhang, Using machine learning and feature engineering to characterize limited material datasets of high-entropy alloys, Comput. Mater. Sci. (2020) 1–6.
- [9] J.M. Rickman, H.M. Chan, M.P. Harmer, J.A. Smeltzer, C.J. Marvel, A. Roy, G. Balasubramanian, Materials informatics for the screening of multi-principal elements and high-entropy alloys, Nat. Commun. (2019) 1–10.
- [10] W. Huang, P. Martin, H.L. Zhuang, Machine-learning phase prediction of high-entropy alloys, Acta Mater. (2019) 225–236.
- [11] N. Islam, W. Huang, H.L. Zhuang, Machine learning for phase selection in multi-principal element alloys, Comput. Mater. Sci. (2018) 230–235.
- [12] Q.F. He, Y.F. Ye, Y. Yang, Formation of random solid solution in multicomponent alloys: from Hume-Rothery rules to entropic stabilization, J. Phase Equilibria Diffus. (2017) 416–425.
- [13] T. Kostuchenko, F. Körmann, J. Neugebauer, A. Shapeev, Impact of lattice relaxations on phase transitions in a high-entropy alloy studied by machine-learning potentials, npj Comput. Mater. (2019) 1–7.
- [14] S. Kotha, D. Ozturk, S. Ghosh, Uncertainty-quantified parametrically homogenized constitutive models (UQ-PHMs) for dual-phase α/β titanium alloys, npj Comput. Mater. (2020) 1–20.
- [15] Y. Zhang, Y.J. Zhou, J.P. Lin, Solid-solution phase formation rules for multi-component alloys, Adv. Eng. Mater. (2008) 534–538.
- [16] O.N. Senkov, D.B. Miracle, A new thermodynamic parameter to predict formation of solid solution or intermetallic phases in high entropy alloys, J. Alloys Compd. (2016) 603–607.
- [17] X. Yang, Y. Zhang, Prediction of high-entropy stabilized solid-solution in multi-component alloys, Mater. Chem. Phys. (2012) 233–238.
- [18] Y. Zhang, X. Yang, P.K. Liaw, Alloy design and properties optimization of high-entropy alloys, JOM: J. Miner. Metal. Mater. Soc. (2012) 830–838.
- [19] D.B. Miracle, O.N. Senkov, A critical review of high entropy alloys and related concepts, Acta Mater. (2017) 448–511.
- [20] N. Islam, W. Huang, H.L. Zhuang, Machine learning for phase selection in multi-principal element alloys, Comput. Mater. Sci. (2018) 230–235.
- [21] Y. Li, W. Guo, Machine-learning model for predicting phase formations of high-entropy alloys, Phys. Rev. Mater. (2019) 1–7.
- [22] J.M. Rickman, G. Balasubramanian, C.J. Marvel, H.M. Chan, M.T. Burton, Machine learning strategies for high-entropy alloys, J. Appl. Phys. (2020) 1–11.
- [23] A. Roy, T. Babuska, B. Krick, G. Balasubramanian, Machine learned feature identification for predicting phase and Young's modulus of low-, medium- and high-entropy alloys, Scr Mater (2020) 152–158.
- [24] U. Bhandari, M.R. Rafi, C. Zhang, S. Yang, Yield strength prediction of high-entropy alloys using machine learning, Mater. Today Commun. (2021) 1–17.
- [25] A. Chandra, X. Yao, Evolving hybrid ensembles of learning machines for better generalisation, Neurocomputing (2006) 686–700.
- [26] L. Yu, K.K. Lai, S. Wang, Multistage RBF neural network ensemble learning for exchange rates forecasting, Neurocomputing (2008) 3295–3302.
- [27] Q. Yan, D. Gong, Q. Shi, A. Hengel, C. Shen, I. Reid, Y. Zhang, Attention-guided network for ghost-free high dynamic range imaging, / (2019) 1751–1760.

- [28] Y. Qin, J. Cao, X. Ji, Y. Zhang, Research on video flame detection algorithm based on improved DS evidence theory, *Multimed. Tool. Appl.* (2020) 26747–26763.
- [29] J. Qian, X. Guo, Y. Deng, A novel method for combining conflicting evidences based on information entropy, *Appl. Intell.* (2016) 876–888.
- [30] S. Zhang, J. Lin, L. Su, Z. Zhou, pDHS-DSET: prediction of DNase I hypersensitive sites in plant genome using DS evidence theory, *Anal. Biochem.* (2019) 54–63.
- [31] H. Zeng, B. Yang, X. Wang, J. Liu, D. Fu, RGB-D object recognition using multi-modal deep neural network and DS evidence theory, *Sensors (Basel)* (2019) 1–19.
- [32] X. Liu, C.E. Athanasiou, N.P. Padture, B.W. Sheldon, H. Gao, A machine learning approach to fracture mechanics problems, *Acta Mater.* (2020) 105–112.
- [33] H. Zhang, H.A. Loáiciga, F. Ren, Q. Du, D. Ha, Semi-empirical prediction method for monthly precipitation prediction based on environmental factors and comparison with stochastic and machine learning models, *Hydrol. Sci. J.* (2020) 1928–1942.
- [34] A. Wolanin, G. Camps-Valls, L. Gómez-Chova, G. Mateo-García, C. van der Tol, Y. Zhang, L. Guanter, Estimating crop primary productivity with Sentinel-2 and Landsat 8 using machine learning methods trained with radiative transfer simulations, *Remote Sens. Environ.* (2019) 441–457.
- [35] H. Liu, J. Chen, D. Hissel, H. Su, Remaining useful life estimation for proton exchange membrane fuel cells using a hybrid method, *Appl. Energy* (2019) 910–919.
- [36] C.S. Kong, M. Haverly, H. Simka, S. Shankar, K. Rajan, A fast hybrid methodology based on machine learning, quantum methods, and experimental measurements for evaluating material properties, *Modell. Mater. Sci. Eng.* (2017) 1–11.
- [37] O.N. Senkov, J.D. Miller, D.B. Miracle, C. Woodward, Accelerated exploration of multi-principal element alloys with solid solution phases, *Nat. Commun.* (2015) 1–10.
- [38] Z. L.G., B.X. H., J.G. J., Z. M., P.L. D., Q.M. W., Effect of AL addition on microstructure and wear properties of Alx FeCrCoCuV High-entropy Alloys, *J. Mater. Eng.* (2016) 1–10.
- [39] H.A. L.W., T. T., Research development on the phase generation of multi-principal element highentropy alloys, *Powder Metallur. Ind.* (2016) 64–67.
- [40] K. Yuan, Y. Deng, Conflict evidence management in fault diagnosis, *Int. J. Mach. Learn. Cybernetic.* (2017) 121–130.
- [41] Y. Pan, L. Zhang, X. Wu, M.J. Skibniewski, Multi-classifier information fusion in risk analysis, *Inf. Fusion* (2020) 121–136.
- [42] Y. Chen, A.B. Cremers, Z. Cao, Interactive color image segmentation via iterative evidential labeling, *Inf. Fusion* (2014) 292–304.
- [43] J. Wang, F. Liu, Temporal evidence combination method for multi-sensor target recognition based on DS theory and IFS, *J. Syst. Eng. Electron.* (2017) 1114–1125.
- [44] F. Xiao, An Improved method for combining conflicting evidences based on the similarity measure and belief function entropy, *Int. J. Fuzzy Syst.* (2017) 1256–1266.
- [45] F.A. Faber, L. Hutchison, B. Huang, J. Gilmer, S.S. Schoenholz, G.E. Dahl, O. Vinyals, S. Kearnes, P.F. Riley, O.A. von Lilienfeld, Prediction errors of molecular machine learning models lower than hybrid DFT error, *J. Chem. Theory Comput.* (2017) 5255–5264.
- [46] X. Yang, Y. Zhang, Prediction of high-entropy stabilized solid-solution in multi-component alloys, *Mater. Chem. Phys.* (2012) 233–238.
- [47] E. Sandvol, D. Seber, A. Calvert, M. Barazangi, Grid search modeling of receiver functions: implications for crustal structure in the middle east and North Africa, *J. Geophys. Res.: Solid Earth* 103 (1998) 26899–26917.
- [48] Y.F. Ye, Q. Wang, J. Lu, C.T. Liu, Y. Yang, High-entropy alloy: challenges and prospects, *Mater. Today* (2016) 349–362.
- [49] Z. Leong, Y. Huang, R. Goodall, I. Todd, Electronegativity and enthalpy of mixing Biplots for High Entropy Alloy solid solution prediction, *Mater. Chem. Phys.* (2018) 259–268.
- [50] M.G. Poletti, L. Battezzati, Electronic and thermodynamic criteria for the occurrence of high entropy alloys in metallic systems, *Acta Mater.* (2014) 297–306.
- [51] A.F. Andreoli, J. Orava, P.K. Liaw, H. Weber, M.F. de Oliveira, K. Nielsch, I. Kaban, The elastic-strain energy criterion of phase formation for complex concentrated alloys, *Materialia* (2019) 1–48.
- [52] Y. Jung, Multiple predictingK-fold cross-validation for model selection, *J. Non-parametr. Stat.* (2017) 197–215.
- [53] X. Fan, Y. Guo, Y. Ju, J. Bao, W. Lyu, Multisensor fusion method based on the belief entropy and DS evidence theory, *J. Sensors* (2020) 1–16.
- [54] I. Toda-Caraballo, P.E.J. Rivera-Díaz-del-Castillo, A criterion for the formation of high entropy alloys based on lattice distortion, *Intermetallics* (2016) 76–87.
- [55] B. Pradhan, A comparative study on the predictive ability of the decision, *Comput. Geosci.* (2013) 350–365.
- [56] S. Hou, F. Hua, W. Lv, Z. Wang, Y. Liu, G. Wang, Hybrid modeling of flotation height in air flotation oven based on selective bagging ensemble method, *Math. Probl. Eng.* (2013) 1–9.