ARTICLE OPEN

Check for updates

A general and transferable deep learning framework for predicting phase formation in materials

Shuo Feng [1], Huadong Fu [2✉], Huiyu Zhou[3], Yuan Wu[4], Zhaoping Lu [4] and Hongbiao Dong [1✉]

Machine learning has been widely exploited in developing new materials. However, challenges still exist: small dataset is common for most tasks; new datasets, special descriptors and specific models need to be built from scratch when facing a new task; knowledge cannot be readily transferred between independent models. In this paper we propose a general and transferable deep learning (GTDL) framework for predicting phase formation in materials. The proposed GTDL framework maps raw data to pseudo-images with some special 2-D structure, e.g., periodic table, automatically extracts features and gains knowledge through convolutional neural network, and then transfers knowledge by sharing features extractors between models. Application of the GTDL framework in case studies on glass-forming ability and high-entropy alloys show that the GTDL framework for glass-forming ability outperformed previous models and can correctly predicted the newly reported amorphous alloy systems; for high-entropy alloys the GTDL framework can discriminate five types phases (BCC, FCC, HCP, amorphous, mixture) with accuracy and recall above 94% in fivefold cross-validation. In addition, periodic table knowledge embedded in data representations and knowledge shared between models is beneficial for tasks with small dataset. This method can be easily applied to new materials development with small dataset by reusing well-trained models for related materials.

INTRODUCTION

Machine learning is a powerful tool which has become an important complement to experiment, theory, and modeling[1–6]. It has been widely used in materials research to mine composition–processing–properties relationships: e.g., predicting compound forming energy[7–9], superconductors critical temperature[10,11], alloy's phases[12–16], materials' properties[17–24]. However, many challenges exist when applying machine learning to new materials development[25–27]. It is common that only small dataset is available for a specific task. In materials science it is impossible to assemble big datasets like that in internet and e-commerce, though materials genome initiative[28,29], high-throughput computing, and experiment have increased the speed of generating data by dozens to hundreds of folds[30,31].

With a set of suitable descriptors, conventional machine learning can performance very well even with small dataset. However, the optimal set of descriptors for a specific job in material research is not out-of-shelf. It is selected by trails and errors and adding new pertinent descriptors is always been considered if models' performance is not met requirement[32]. Building new applicable descriptors entails deep understanding of mechanisms, which is very challenging in developing new materials. For example, Ward et al.[7] first used 145 general-purpose Magpie descriptors (descriptive statistics, e.g., average, range, and variance of the constituent elements) in predicting ternary amorphous ribbon alloys (AMRs). Later they used 210 descriptors (including 145 Magpie descriptors and new descriptors derived from physical models and empirical rules developed by amorphous alloys community, e.g., cluster packing efficiency and formation enthalpy) in optimizing Zr-based bulk metallic glass (BMG)[33]. Some descriptors derived from physical models and empirical rules are sensitive to alloying and temperature;

obtaining precise values of them is difficult; using simplified models to calculate them (e.g. utilize ideal solution model in estimating alloy mixing enthalpy instead of Miedema model or experimental results) might weaken the final machine learning models' performance.

How to fully exploit limited data, existing models, and domain expertise is the key to efficiently applying machine learning in materials research, and general and transferrable machine learning frameworks are in urgent need. Transfer learning is a special machine learning technique that enables models to achieve high performance using small datasets through knowledge sharing between modes in related domains[34–36]. Deep learning is an end-to-end learning which combines automatic feature extractors and conventional machine learning models as regressors or classifiers into one model[37]. Deep learning has an advantage over conventional machine learning in exploiting transfer learning for its feature extractors can be easily reused in related tasks.

Predicting the phases of a material, e.g., solid solution phases of simple BCC/FCC/HCP structure, intermetallics of complex structure, metastable amorphous phases, and mixture of different phases, is the basic tasks and fundamental challenges of materials research. AMRs and BMGs extended materials from conventional crystalline (CR) metallic materials to amorphous state materials[38–42]; high-entropy alloys (HEAs), which are also known as multi-principal element alloys (MPEAs) and concentrated solid solution alloys, extended metallic materials from corner and edge regions to the center regions of multi-component phase diagrams[43–47]. Predicting them challenges our classical theory[48]. Researchers have attempted to predict alloys' glass-forming ability (GFA) and HEAs' phases by empirical thermo-physical parameters[49–53], CALPHAD method[54,55], first-principles calculations[56]. Conventional

[1]School of Engineering, University of Leicester, Leicester LE1 7RH, UK. [2]Beijing Advanced Innovation Center of Materials Genome Engineering, University of Science and Technology Beijing, 100083 Beijing, China. [3]School of Informatics, University of Leicester, Leicester LE1 7RH, UK. [4]State Key Laboratory for Advanced Metals and Materials, University of Science and Technology Beijing, 100083 Beijing, China. ✉email: hdfu@ustb.edu.cn; h.dong@le.ac.uk

machine learning was used in these tasks as well[7,13–16,33]. However, developing new amorphous alloys and HEAs by design is still quite challenging, for their mechanisms are still not clear and data are much less than that of conventional materials, e.g., steels, aluminum alloys.

In this work, we propose a general and transferable deep learning (GTDL) framework to predict phase formation in materials with small dataset and unclear transformation mechanism. Case studies on GTDL predictions with a medium-sized dataset (containing 10000+ pieces of data) of GFA and a small dataset (containing only 355 pieces of data) of HEAs demonstrate: GTDL framework outperforms existing models based on manual features, periodic table knowledge embedded in data representations helps to make predictions, and knowledge shared between different models enable prediction with small dataset. The proposed GDTL framework can be easily used in new materials development with small datasets by exploiting trained deep learning models on big dataset of related materials.

RESULTS

GTDL framework

The pipeline of this work and schematics for transfer learning, etc. are shown in Fig. 1a. For deep learning accepts unstructured data, e.g., image, audio, as input, we mapped raw data, e.g., chemistry and processing parameters, to pseudo-images first using some special two-dimensional (2-D) structures, Convolutional neural networks (CNNs) were then utilized to automatically extract features through their hierarchy structure and to make classification/regression. The well-trained feature extractors, i.e., convolutional layers were reused directly for new tasks with small dataset. Here, we used a whole periodic table containing 108 elements for composition mapping (periodic table representation, PTR). In order to bring processing parameters into representation, we mapped them to an unused area in the periodic table (see Supplementary Fig. 1). An example of PTR for alloy Fe$_{73.5}$Cu$_1$Nb$_3$Si$_{13.5}$B$_9$ is given in Fig. 1a. We compared models using different mappings without periodic table structure (see Supplementary Figs 2 and 3), e.g., atom table representation[10], to prove the advantage of the embedded periodic table structure. We also compared our models with conventional machine learning models using manual feature engineering (see the full list of the features in Supplementary Table 1) to validate the convenience of automatic features engineering. The workflow of conventional machine learning is also shown in Fig. 1a. A clear advantage of deep learning framework over conventional machine learning is it can automatically extract features and transfer knowledge.

Many classical CNN structures for image recognition are available now. However, we need to simplify and compress those structures to reduce the risk of overfitting limited data in our tasks. We tested some simplified classical CNNs, e.g., AlexNet[57], VGG[58], GoogLeNet, and Inception module[59]. A VGG-like CNN which is shown in Fig. 1b was used in our work due to its very compact structure and strong power of feature extraction. Our VGG-like CNN has 6274 trainable parameters, only 1% size of atom table CNN[10] (611,465 trainable parameters). Thus, it can reduce the risks of overfitting effectively.

Predicting GFA using GDTL

The GFA of an alloy, i.e., the critical cooling rate below which the alloy melt undergoes nucleation and growth and forms crystal (CR), is a core problem in developing new amorphous alloys. However, it is challenging to measure the critical cooling rate experimentally. Researchers often simplify GFA into three levels: BMG, AMR, and CR, which correspond to strong, weak, and no GFA, respectively[13]. GFA of an alloy can be roughly evaluated through melt-spun (its cooling rate is in the range of $10^6-10^5 \, K \, s^{-1}$) and copper mold casting (its cooling rate is in the range of $10^2-1 \, K \, s^{-1}$): if an alloy forms a crystalline state under melt-spun, it is labeled CR (no GFA); if it forms an amorphous state through melt-spun but forms crystalline state under copper mold casting,
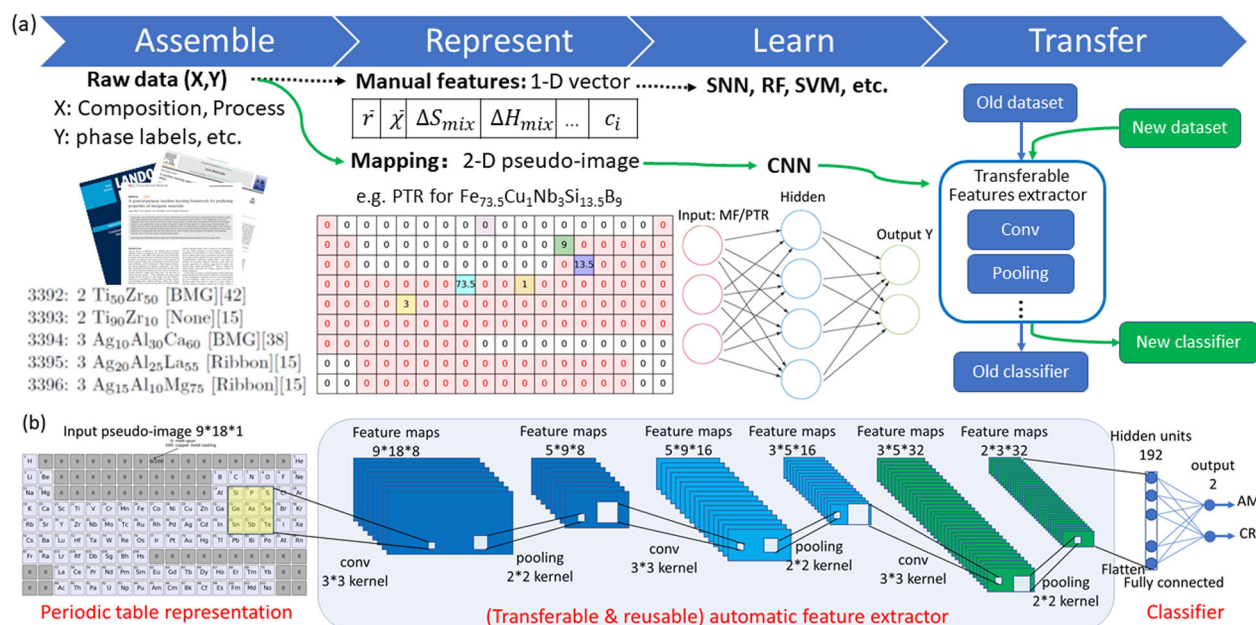
Fig. 1 The workflow of our works. a The workflow of the proposed GTDL framework (in green solid arrows) and conventional machine learning (in black dotted arrows) which does not have the ability of automatically extracting features and knowledge transfer. The schematics for assembling dataset, data representation, machine learning, knowledge transfer, and an example of PTR (periodic table representation) were given. MF, SNN, RF, SVM, and CNN denotes manual features, shallow neural network, random forest, supported vector machine, and convolutional neural network, respectively. In GTDL framework, raw data are mapped to 2-D pseudo-images first, features are then extracted automatically by convolutional layers, knowledge is transferred by sharing the well-trained feature extractors for new tasks with small dataset. **b** The schematics for our VGG-like convolutional neural network.
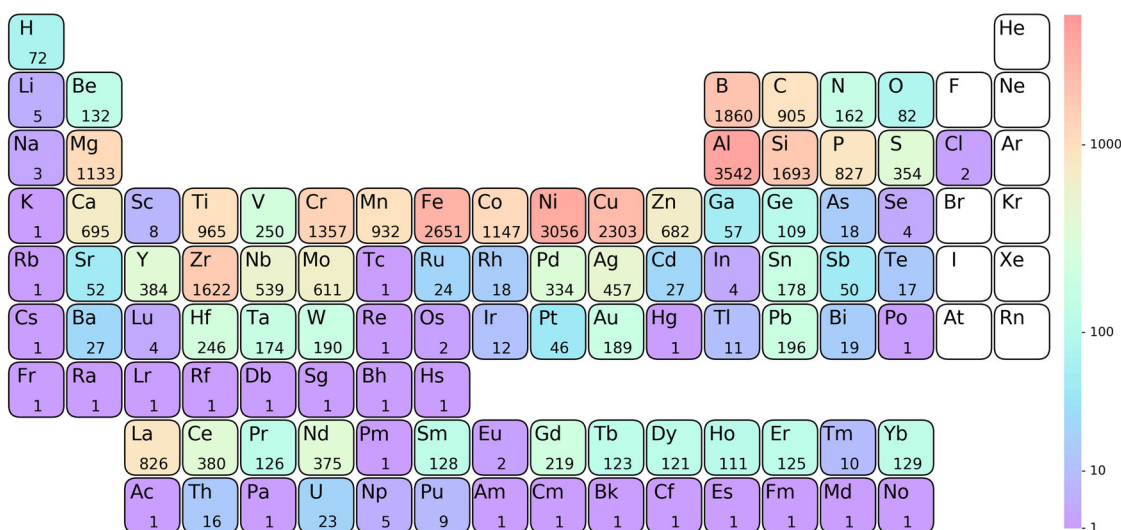
Fig. 2 Statistics of our glass-forming ability dataset. The occurrence numbers of elements in the dataset are given under periodic table background. The blank squares, e.g., squares for noble gases, indicate the elements not in the dataset.

| Table 1. | Comparison of average accuracy among different models under 10-fold cross-validation. | | | | | |
|---|---|---|---|---|---|
| Model | Data representation | Input size | Algorithms | Average accuracy | |
| | | | | Training | Testing |
| Ward's work[7] | Manual features vector generated by Magpie | 145 | Random forest | – | 90% |
| SNN1 | Manual features vector + processing parameter | 13 + 1 | SNN | 89.8% | 89.9% |
| SNN2 | Composition vector + processing parameter | 73 + 1 | SNN | 93.2% | 92.8% |
| SNN3 | Manual features vector + composition vector + processing parameter | 86 + 1 | SNN | 93.9% | 93.5% |
| SNN4 | Manual features vector generated by Magpie + processing parameter | 145 + 1 | SNN | 90.1% | 90.0% |
| CNN1 | Atom table representation | 11 × 11 | CNN | 96.4% | 95.0% |
| CNN2 | Randomized periodic table representation | 9 × 18 | CNN | 96.7% | 94.9% |
| CNN3 | Periodic table representation | 9 × 18 | CNN | 96.4% | 96.3% |

then it is labeled as AMR (weak GFA); if it forms amorphous state under copper mold casting, it is classified as BMG (strong GFA).

In this work, we try to assemble a GFA dataset as large as possible. Our dataset includes Sun's binary alloys GFA dataset (about 3000 entries)[13], Ward's ternary alloys GFA dataset (about 6000 entries)[7], and BMG dataset (about 800 entries)[33], and Miracle's GFA dataset (about 300 entries)[60]. In those datasets, crystalline alloys data are in the minority, because AMRs and BMGs are the focus of research, and crystalline alloys are commonly discarded and unpublished as the failed experimental results. In reality, the number of amorphous alloys is less than that of their crystalline counterparts. To compensate for this weakness and increase the variety of crystalline data in our dataset, we add 800+ pieces of conventional crystalline metallic materials data (including steels, superalloys and Co, Al, Mg, Cu, Zn alloys, etc.) which is extracted from https://www.makeitfrom.com/. Figure 2 shows the statistics of elements distribution in our dataset (for detailed statistics see Supplementary Figs. 5 and 6). Our dataset contains 97 elements in the periodic table, and many of these elements are present simultaneously in entries of CRs, AMRs, and BMGs. Considering that some AMRs in our dataset are actually BMGs (due to incomplete record and experiment), we did not simply treat it as a (CR/AMR/BMG) ternary classification problem. Instead, a processing parameter (0 represents rapid solidification melt-spun, and 100 represent copper mold casting of normal cooling rate) was added into this problem to convert the ternary classification problem into a (AM/CR) binary classification problem

(AM represents forming amorphous state, and CR represents forming crystalline state). The size of our original dataset is 10,440, and the size of dataset after conversion is 16,250.

Table 1 shows the average training and testing accuracies of four shallow neural networks (SNNs) and three CNNs in 10-fold cross-validation. SNN1 (using only 14 features derived from empirical rules of BMGs community) and SNN4 (using 145 general-purpose Magpie descriptors[7]) show the lowest testing accuracy of about 90%. They show a marginal difference in accuracy with Ward's random forest models (89.9% vs. 90%). SNN4 and Ward's random forest model used 145 general-purpose Magpie features, and SNN1 only used 14 features (including one processing parameter, mixing entropy, the statistical information of atomic radius, Pauling electronegativity, bulk modulus, and work function). We found increasing features or even using the full list of features (see Supplementary Table 1) did not improve accuracy. SNN2 only used composition vector as input, but it showed higher accuracy than SNN1 and SNN4. SNN3 used manual features vector plus composition vector as input and it improved the accuracy further. Due to our limited understanding of the GFA's physical mechanisms and lack of precise property data as input (e.g., ideal solution model and Miedema model were used to calculate alloy mixing entropy and mixing enthalpy, respectively), improving the model accuracy by adding more pertinent features is impracticable. All four SNNs show lower accuracies than three CNNs. Besides CNNs' accuracy advantage over SNNs, it is also quite convenient to use CNNs, for they only need compositions
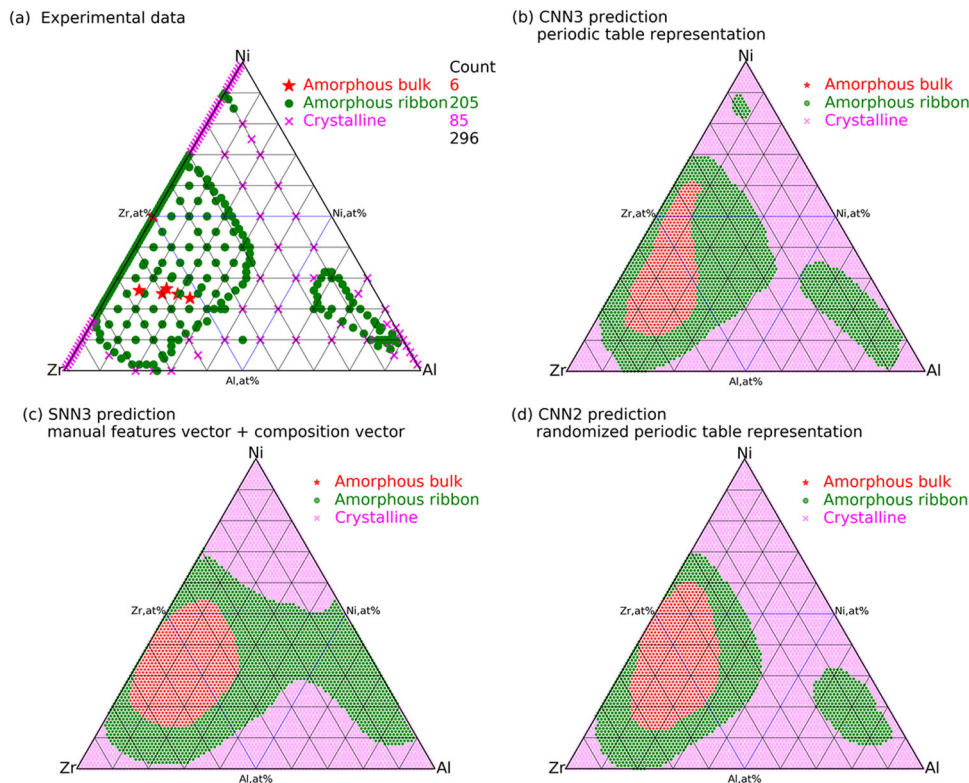
Fig. 3 Comparison of experimental data and predictions. a Experimental data points of Al–Ni–Zr ternary system in our dataset and the predictions of **b** CNN3, **c** SNN3, and **d** CNN2.

and processing parameters as input, and they automatically extract features through convolutional layers. CNN3 which refers to a CNN with PTR shows the highest testing accuracy of 96.3%. The only difference among three CNNs is in that the data representations of CNN1 and CNN2 did not have periodic table structure. The advantage of CNN3 over CNN2 and CNN1 is not evident (only 1.3% higher). However, we will demonstrate that CNN3 has more obvious advantages over other models in predicting unseen alloys, i.e., better generalization.

The Al–Ni–Zr ternary system has 296 entries (include 186 entries from the Al–Ni–Zr ternary system, and 110 entries from Al–Ni, Al–Zr, Ni–Zr binary systems) in our dataset and the distribution of data points is relatively uniform in composition space, see Fig. 3a the ground truth of the Al–Ni–Zr system. So, the Al–Ni–Zr system is quite suitable to validate and compare models. Figure 3b–d shows the GFA prediction of CNN3, SNN3, and CNN2. CNN3 successfully predicted three amorphous composition areas, and the shapes and boundaries of these areas are satisfied when compared with the ground truth. Other models did not predict all three areas. SNN3 did not predict the crystalline area between two amorphous composition areas, i.e., the GFA of the area was overestimated. CNN2 successfully predicted two amorphous composition areas but missed the small amorphous composition areas near Ni corner. All models correctly predicted the five BMGs in ground truth and the predicted BMGs cover certain area (not some discrete points) around the ground truth points. It is reasonable, researchers commonly reported the optimal BMGs only, and BMG candidates (especially before the appearance of BMGs) are archived as AMRs data. This sparse and ununiform distribution of BMG data points usually induces BMG data points buried by surrounding densely distributed AMR data points and omitted as a noise (see Fig. 3a). That is why we adjusted the ternary classification into binary classification, i.e., ternary classification easily underestimates alloys' GFA.

Table 2. Comparison of models' prediction accuracy on unseen alloy systems.

| Alloy system	SNN3	SNN4	CNN1	CNN2	CNN3
Leave-one system-out test on 160 ternary systems	–	76.5%	–	77.3%	83.8%
Leave-one system-out test on Al–Ni–Zr system	62.3%	63.5%	68.4%	66.4%	80.3%
Ir–Ni–Ta–(B), Mg–Cu–Yb, S-bearing BMGs	17/28	9/28	18/28	23/28	27/28
RE$_6$Fe$_{72}$B$_{22}$	5/13	8/13	7/13	8/13	10/13
Binary alloys outliers	8/18	16/18	13/18	13/18	16/18

To validate predicting ability of models on unseen alloy systems, we carried out a leave-one-system-out (LOSO, like the leave-one-cluster-out cross-validation used by Meredig et al.[61]) cross-validation on 160 ternary systems which has over 40 entries in our dataset. In LOSO cross-validation for a ternary system A–B–C, entries of A–B, A–C, B–C binary alloys and A–B–C ternary alloys were hold out as testing dataset. Models were trained with the remaining dataset. The average testing accuracies of SNN4, CNN2, and CNN3 under LOSO cross-validation are shown in Table 2. CNN3 outperforms CNN2 and SNN4 in predicting unseen alloy systems by about 7%.

Table 2 also show the LOSO cross-validation results for the Al–Ni–Zr system. Here, we used Al–Ni–Zr AMR results (5151 composition points in total, not the 296 Al–Ni–Zr entries in our dataset) in Fig. 3b as ground truth to calculate prediction accuracy. The predictions of CNN2 and CNN3 are shown in Fig. 4. We can see CNN3 shows accuracy advantage over other CNNs and SNNs by at least 12% when no Al–Ni–Zr data are in training dataset.

To further validate the generalization of the models, we collected some newly reported BMG alloys and some specially
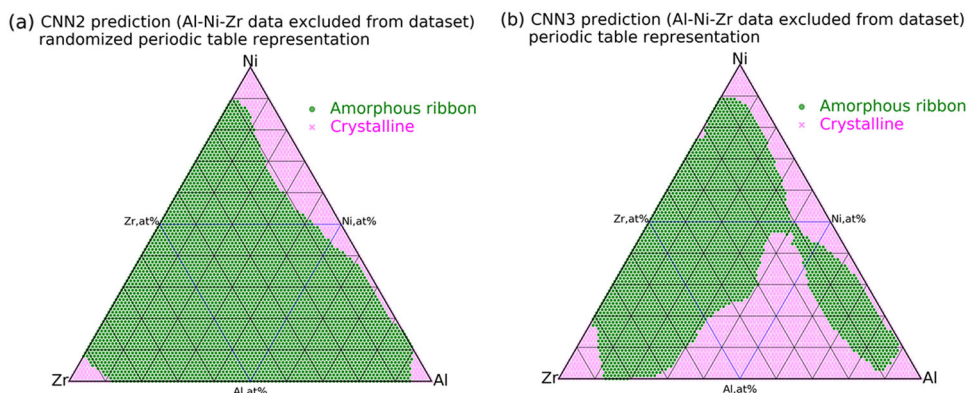
Fig. 4 Predictions for Al–Ni–Zr ternary system by the re-trained models. a CNN2, **b** CNN3, using dataset in which data about Al–Ni, Al–Zr, Ni–Zr binary alloys, and Al–Ni–Zr-containing multi-component alloys were removed.
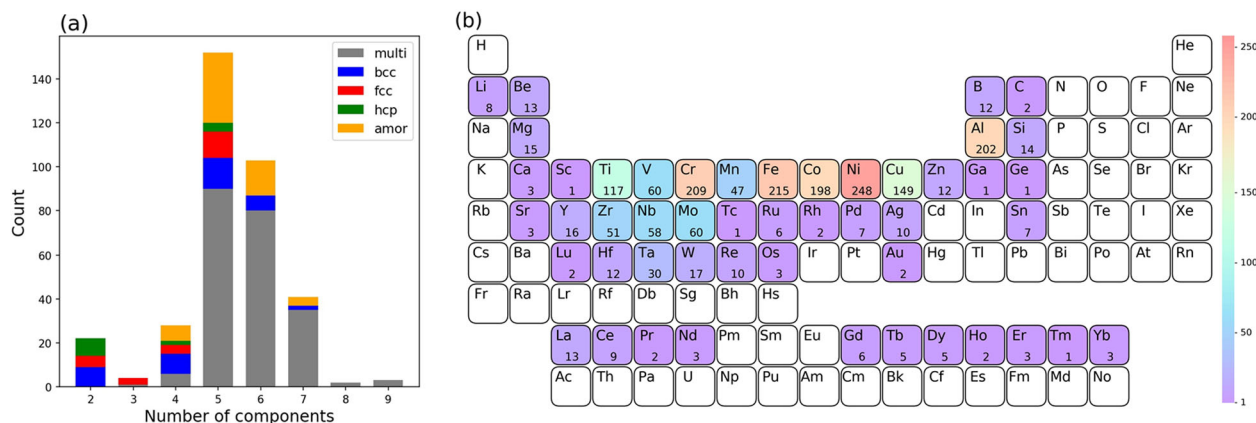
Fig. 5 Statistics of the 355 HEAs in the dataset. a The numbers of binary to nonary HEAs and the proportions of different phases. **b** The occurrence numbers of elements in the dataset are given under periodic table background. The blank squares, e.g., squares for noble gases, signify the elements not in the dataset.

selected alloys that outside our dataset, e.g., high-temperature Ir–Ni–Ta–(B) BMGs[62], Mg–Cu–Yb BMGs[63], sulfur-bearing BMGs[64], RE-bearing alloys $RE_6Fe_{72}B_{22}$ (RE: Sc, Y, La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er)[65], and 18 binary alloys outliers according to empirical criteria[66]. Our dataset only has one ternary AMR data point about Ir–Ni–Ta–(B) system and does not have any data about Mg–Cu–Yb system and sulfur-bearing AMRs and BMGs. Rare earth elements have close physical and chemical properties. However, experimental results show the simple substitution of rare earth elements causes the GFA variations of $RE_6Fe_{72}B_{22}$ alloys. Louzguine-Luzgin reported 18 binary alloys outliers which should be good glass-formers according to empirical criteria, but they cannot form an amorphous state even in rapid solidification. Table 2 shows that CNN3 with PTR plus automatic feature engineering attained the highest prediction ability and SNN3 and SNN4 based on manual feature engineering performed the worst. The performances of CNN1 and CNN2, which use automatic feature engineering but do not have periodic table structure in data representation, are between SNN3 and CNN3. The detailed comparisons are shown in Supplementary Tables 5 and 6. These rigorous tests strongly verified CNN3 can be used to predict the GFA in unassessed alloy systems.

Overall, when dataset is large enough (e.g. the Al–Ni–Zr system), the benefit of adding periodic table structure (domain expertise) to representation is not obvious. When data are insufficient or no data are available, domain expertise is vital. Periodic table structure plus CNN, like CNN3, brings the convenience of

automatic feature engineering and improves the generalization by introducing background knowledge.

Transfer learning of HEAs with small dataset using GTDL

The well-trained deep learning models for GFA can be reused in predictions of related materials e.g., HEAs. All previous machine learning studies on HEAs used manual feature engineering plus conventional machine learning models, e.g. supported vector machine[14] and SNN[15]. These models need sophisticated features as input and can only distinguish BCC from FCC, or differentiate intermetallics from solid solutions. Tasks like predicting HEAs of HCP structure is rather difficult due to limited data. The two machine learning tasks, i.e. predicting GFA and predicting phases of HEAs, have different output domain (amorphous/crystalline binary classification in GFA prediction and five phases labels in HEAs prediction) and highly correlated (or overlapped) input domain from the point of transfer learning: Figs 2 and 5b show common elements in those alloys are similar; some amorphous alloys are also HEAs; the descriptors developed in conventional machine learning for GFA and HEAs can be shared[7,13–16,33] (e.g. atomic size difference, mixing enthalpy, mixing entropy, difference in Pauling electronegativities, and valence electron concentration). So, we believe that the automatic feature extractors of the well-trained CNNs, which have outperformed known manual features in GFA prediction, will work in HEAs prediction too. Based on the features, we built a high-performance model with a small dataset which can discriminate five types of phases (BCC, FCC, HCP, amorphous, mixture of multiple phases) in HEAs in one go.

Table 3. The scores of transfer learning on high-entropy alloys dataset under fivefold cross-validation with three data representations.

Data representation	Transfer from	Accuracy	Precision	Recall	F1 score
Periodic table representation	CNN3	0.935	0.940	0.935	0.934
Randomized periodic table representation	CNN2	0.854	0.837	0.854	0.831
Atom table representation	CNN1	0.884	0.888	0.884	0.875

Here, we used the dataset from Gao's review on HEAs[51] where experimentally synthesized 355 HEAs data are collected. Therein, 41 samples have single BCC phase, 24 samples have single FCC phase, 14 samples have single HCP phase, and 59 samples have single amorphous phase. The remaining 217 samples with multiple phases. Numbers of binary to nonary MPEAs and the proportions of BCC, FCC, HCP, amorphous and multiple phases are shown in Fig. 5a. Most of the samples consist of five or six elements and the single-phase HEAs only account for a small fraction. There are 50 elements in the dataset and their occurrence frequencies are shown in Fig. 5b. Elements Fe, Ni, Cr, Co, and Al occur in more than 190 samples, while Sc, Tc, Ga, Ge, and Tm only occur once. It is rather difficult to build machine learning models using such small dataset, with so many elements and unbalanced data distribution.

In transfer learning from GFA to HEAs, the 2-D representations of HEAs' compositions were fed into the well-trained CNN1, CNN2, CNN3, and the intermediate results (high-dimensional features yielded from convolutional layers) of theses CNNs were extracted. Then these features were used in new classifier (here we used random forest for its good interpretability, and it need very little hyperparameters optimization) as input. Stratified data division strategy (to ensure training and testing dataset have similar data distribution) and Sklearn package were used in training. Table 3 shows the average scores of our transfer learning models on HEAs dataset under fivefold cross-validation. Our model without resorting to any manual features engineering is capable of distinguishing BCC, FCC, HCP, amorphous, and multiple-phase mixture with fivefold cross-validation scores (average accuracy/recall/precision/F1 on testing datasets) over 94% after training and test. We should bear in mind that when labels' distribution is unbalanced like that of our HEAs data, achieving high recall, high precision, and high accuracy at the same time is very difficult. We can see model transferred from CNN3 has the highest scores which indicates that PTR is also beneficial for transfer learning. Our previous results and some research[67] show that if dataset is not big enough, domain knowledge is important for model's performance. Though raw data (alloy compositions) are the same for CNN1, CNN2, and CNN3, the direct input (data representations) for them and the information extracted by corresponding feature extractors are different. Domain knowledge (periodic table structure) was embedded in CNN3's input and embodied in the features extracted, while CNN1 and CNN2 do not have access to this knowledge. The proposed transfer learning model is an upgrade for conventional machine learning relying on manual feature engineering and could serve as an effective guide for designing new HEAs.

DISCUSSIONS

To explain why PTR and transfer learning is effective, we illustrated the information that is automatically extracted from different representations by CNNs. Visualizing the high-dimensional features extracted by convolutional layers, i.e., the intermediate results of CNNs, is a good way to explore the extracted features. However, finding the visual and intuitive relationship between elements from these high-dimensional features (see Supplementary Fig. 10) is still very challenging: dimensionality reduction is necessary. Those high-dimensional features were compressed by principal component analysis and the first two/four principal components were visualized.

Figure 6a illustrates the knowledge of 108 elements extracted by CNN with PTR and it shows apparent periodic trends: elements from 18 groups, lanthanide (group 19), and actinides (group 20) are clustered in different regions (marked with different colors); group 1 to group 18 distribute along a semicircle in sequence; elements from lanthanide and actinides distribute in two semicircles with atomic number sequence; elements in one group distributes from semicircle's inside to outside according to ascending atomic number. More than half the elements in periodic table have limited data in our dataset, and halogens (group 17), noble gases (group 18), etc. are absent in our dataset, but the trends of them are consistent and reasonable. It indicates the PTR transfers the knowledge of periodic table to the GFA knowledge, i.e., background knowledge was absorbed by the machine learning models. Figure 6b illustrates the knowledge extracted by CNN from representation without periodic table background: randomized periodic table structure embedded in data representation was learned by model.

Periodic table has abundant physical and chemical knowledge (see Supplementary Fig. 4). Atomic radius, Pauli electronegativity, valence electrons density, and other physical chemistry properties display periodic variations in periodic table. When developing new amorphous alloys, periodic table is often used as a map. Similar atom substitution and column substitution are common strategies for improving GFA. The spatial information or elements' relative position information is difficult to be fully described by manual features engineering. The solution is keeping the periodic table structure in representation. Materials properties originated from electrons' behaviors. The periodic characteristic of element properties in periodic table originated from electrons configuration. The electron configuration of an element can be inferred given its position in periodic table. The abscissa and ordinate of an element in PTR correspond to group number (outer shell electrons number) and its period number (the number of electron shells). CNN exacts the spatial (or co-ordinates) information of pixels in 2-D representation through convolutional layers. So, the knowledge of each element's group number, period number, and electronic configuration in PTR can be transferred to the features that CNN automatically extracted. Element properties (such as atomic radius and Pauli electronegativity) are not explicitly provided in PTR. However, the periodic characteristic of element properties in rows (period) and columns (group) is embedded in PTR. The element properties that CNN3 (PTR) learned from GFA dataset vary with atomic number periodically (see Supplementary Fig. 11). In contrast, the element properties learned by CNN2 (randomized PTR) did not show periodic characteristic (see Supplementary Fig. 12).

It explained why CNN3 shows better performance in predicting new data than CNN1 and CNN2: we provided different expertise to CNN1, CNN2, and CNN3; and domain knowledge is helpful for machine learning models with small dataset[67]. Adding periodic table structure into data representations affords models the ability to infer useful information from the periodic table when direct data are insufficient.

The features of 355 HEAs generated by GFA model are shown in Fig. 7. We can see that alloys of the same phases tend to cluster in the diagrams. Based on the first and second principal features, we
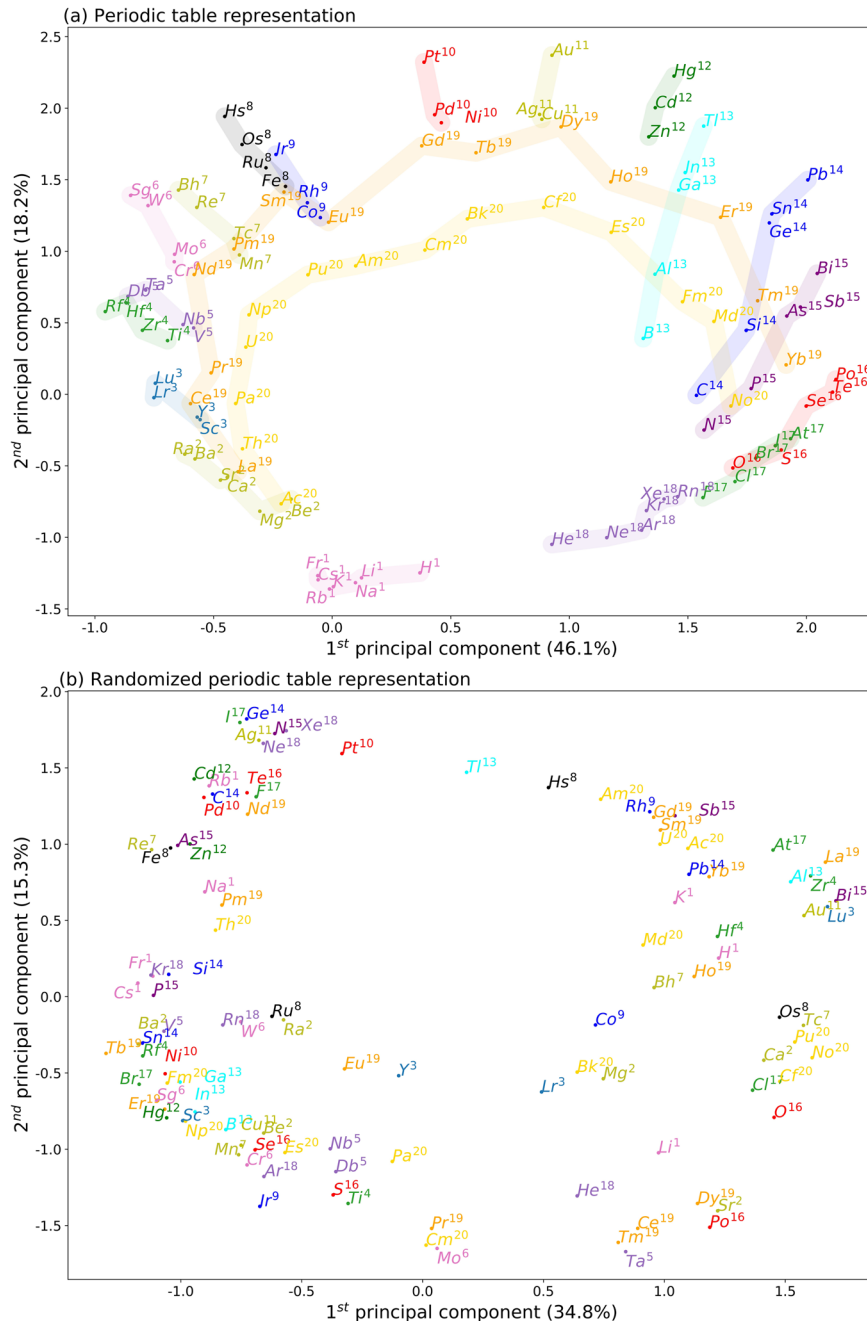
Fig. 6 **Features analysis of the GFA prediction model after PCA.** Projection of the feature vectors of 108 elements onto the plane spanned by the first and second principal axes. The percentage represents the ratio of the variance on the principal axis direction. Elements are colored according to their elemental groups. **a** Periodic table representation. **b** Randomized periodic table representation. The superscript 1–18 on element symbol represents the element's group number; superscript 19 and 20 represent lanthanide and actinides, respectively.

can intuitively distinguished stable BCC, FCC, HCP, and multi-phase alloys. Most alloys of metastable amorphous phases can be discriminated from other alloys of stable phases by third and fourth principal features visually. It indicates transfer learning from GFA to HEA is successful and justifies the high scores of our model for HEAs.

In sum, CNNs get domain knowledge (e.g., periodic table knowledge) embedded in 2-D representation through learning. Periodic table knowledge and PTR is beneficial for machine learning models with small dataset. The feature extractor of CNN for GFA can generate appropriate features for HEAs prediction brings the success of transfer learning.

METHODS

Data representations

Raw data need to be converted into one-dimensional (1-D) vector of features by manual feature engineering for conventional machine learning. This is a process of refining information and adding expertise to data representation. The performance of final models relies on the quality of data representations. The 1-D vector of features (attributes/descriptors) used as input for this work include (a) statistics information of components' properties, e.g., the maximum/minimum/average atomic radius, Pauling electronegativity, elemental bulk modulus, elemental work function, melting point, etc.; (b) composition vector; (c) parameters derived from empirical criteria, e.g., mixing entropy ΔS_{mix}, mixing enthalpy ΔH_{mix}, the
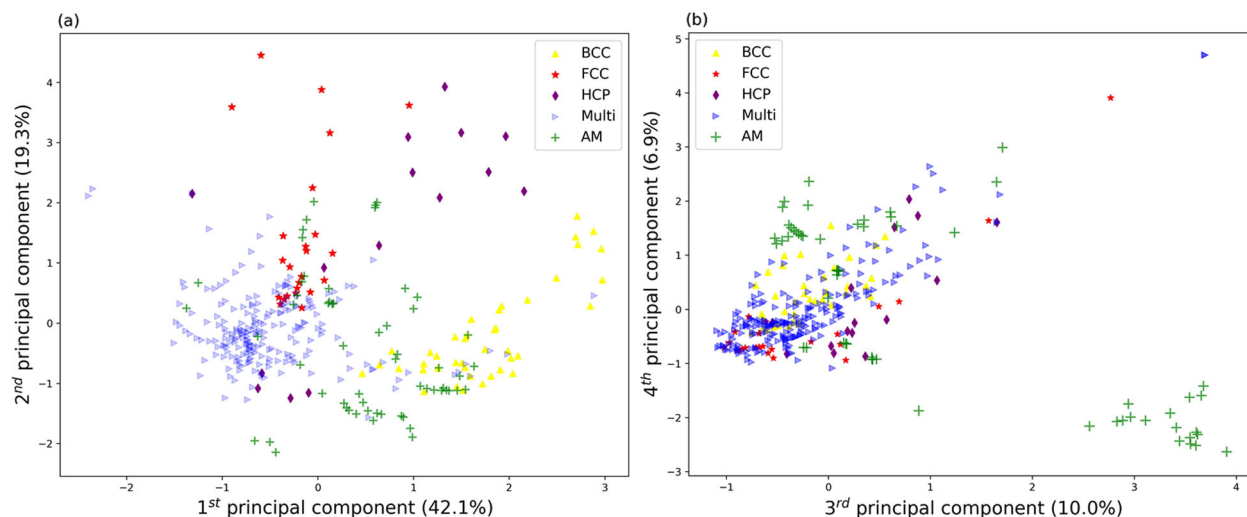
Fig. 7 The high-entropy alloys' first four principal features generated by glass-forming ability model. Alloys are colored according to their phases. The percentage represents the ratio of the variance on the principal axis direction.

atomic size difference ΔR, the electronegativity difference $\Delta \chi$, valence electron concentration VEC, etc.

$$\Delta S_{\text{mix}} = -R \sum_{i=1}^{n} c_i \ln c_i \tag{1}$$

$$\Delta H_{\text{mix}} = \sum_{i=1, i \neq j}^{n} 4\Delta_{\text{mix}}^{AB} c_i c_j \tag{2}$$

$$\Delta R = \sqrt{\sum_{i=1}^{n} c_i \left(1 - \frac{r_i}{\bar{r}}\right)^2}, \bar{r} = \sum_{i=1}^{n} c_i r_i \tag{3}$$

$$\Delta \chi = \sqrt{\sum_{i=1}^{n} c_i (\chi_i - \bar{\chi})^2}, \bar{\chi} = \sum_{i=1}^{n} c_i \chi_i \tag{4}$$

$$\text{VEC} = \sum_{i=1}^{n} c_i (\text{VEC})_i \tag{5}$$

where c_i is the atomic fraction of the ith component; Δ_{mix}^{AB} is the mixing enthalpy of alloy A–B; r_i is the Miracle's atomic radius of the ith component; χ_i is the electronegativity of the ith component; $\bar{r}$ is the average atoms radius of the components in the alloy; $\bar{\chi}$ is the average electronegativity of the components in the alloy; $(\text{VEC})_i$ is the valence electron concentration of the ith component; VEC is the average valence electron concentration of the components in the alloy. The r_i were taken from Miracle's paper[68]; Δ_{mix}^{AB} were taken from Takeuchi's paper[69]; Pauling electronegativity, elemental bulk modulus, elemental work function, etc. were taken from Guo's paper[52].

A schematic diagram for our PTR for alloy composition and preparation process used in CNN3 is shown in Supplementary Fig. 1. PTR mimics digital images. Alloy composition and preparation processes are mapped to a 2-D pseudo-image of 9 pixels × 18 pixels (162 pixels in total). Each square represents a pixel. The 108 blue squares correspond to 108 elements in the periodic table, e.g., the first pixel/square in the first row is used to store the atomic percentage of element hydrogen in an alloy. The 54 gray squares are the unused area in the periodic table. The alloy composition (in atomic percentage) is mapped to the corresponding blue squares, and the preparation process (0 represents melt-spun and 100 represents copper mold casting) is mapped to a gray square (we arbitrarily chose the ninth pixel/square in the first row in this work). The rest pixels/squares are set to 0. The randomized PTR used in CNN2 is almost the same with PTR except 108 elements were randomly placed in the periodical table area (see Supplementary Fig. 2). The atom table representation used in CNN1 are square images of 11 × 11 pixels, elements are placed in an atom table from left to right and from top to bottom according to the atomic number of elements (see Supplementary Fig. 3). The preparation process is mapped to the last pixel in the atom table and the rest unused pixels are set to 0.

CNN structure
A VGG-like CNN was used in automatically extracting features and making classification. The structure of our VGG-like CNNs (see its schematics in Fig. 1b) is as follows: the size of convolutional filters was 3 × 3 for all the three convolutional layers, and the stride was set at 1. The channel number in convolutional layer doubles from 8 to 16 to 32. Padding was used for the input of convolutional layers by adding zeros around the border, i.e., a zero-padding of one, to preserve as much information as possible. The most common type of convolution with a ReLU filter was used, and the value of each filter was learned during the training process. The CNN consists of two parts. One part is the feature extractor involving the first three pairs of convolutional layers, pooling layers, and ReLU (Rectified Linear Unit) layers which have a nonlinear activation function $f(x) = \max(0, x)$. The other part is the classifier with one full connection layer and one softmax classification layer. The details of the VGG-like CNN are shown in Supplementary Table 4 and Supplementary Fig. 7. Due to the limit of dataset size and small input images, our CNNs have much fewer layers, channels, and trainable parameters (about 6000) than the well-known VGG-16 (ref. [58]) (about 133 million).

Training details
In prediction of GFA, all models were created and tested using the Keras with Tensorflow as its backend. The full list of manual features used in SNNs is shown in Supplementary Table 1. All possible combinations of manual features were tested, and the optimal combination which achieved the best accuracy was chosen. Hyperparameters, e.g., neuron number, were also optimized. SNNs of 20 neurons in the hidden layer were used in this work. In the training phase, the output of the SNN and CNN fitted the ground truth, and the categorical cross-entropy was used as the loss function to evaluate the fitness. The training epoch was set to 2000 (loss values almost remain unchanged), and 10-fold cross-validation (the dataset was split into 10 parts, each time 1 part was hold out as testing dataset, the remaining parts were used in training models, no validation dataset and early stop was used in training, 10 models were created after cross-validation) was used to evaluate the training/testing accuracy. In prediction of new alloys' GFA, the results of a committee consisted of 10 models were utilized.

DATA AVAILABILITY
The dataset used to generate the results in this work are available at https://github.com/sf254/glass-froming-ability-prediction.

CODE AVAILABILITY
The codes pertaining to the current work are available at https://github.com/sf254/glass-froming-ability-prediction.

REFERENCES

1. Mater, A. C. & Coote, M. L. Deep learning in chemistry. *J. Chem. Inf. Model.* **59**, 2545–2559 (2019).
2. Vasudevan, R. K. et al. Materials science in the artificial intelligence age: high-throughput library generation, machine learning, and a pathway from correlations to the underpinning physics. *MRS Commun.* **9**, 821–838 (2019).
3. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. & Walsh, A. Machine learning for molecular and materials science. *Nature* **559**, 547–555 (2018).
4. Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A. & Kim, C. Machine learning in materials informatics: recent applications and prospects. *npj Comput. Mater.* **3**, 54 (2017).
5. Wei, J. et al. Machine learning in materials science. *InfoMat* **1**, 338–358 (2019).
6. Raccuglia, P. et al. Machine-learning-assisted materials discovery using failed experiments. *Nature* **533**, 73–76 (2016).
7. Ward, L., Agrawal, A., Choudhary, A. & Wolverton, C. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Comput. Mater.* **2**, 16028 (2016).
8. Zheng, X., Zheng, P. & Zhang, R. Machine learning material properties from the periodic table using convolutional neural networks. *Chem. Sci.* **9**, 8426–8432 (2018).
9. Zhang, J. et al. Robust data-driven approach for predicting the configurational energy of high entropy alloys. *Mater. Des.* **185**, 108247 (2020).
10. Zeng, S. et al. Atom table convolutional neural networks for an accurate prediction of compounds properties. *npj Comput. Mater.* **5**, 84 (2019).
11. Stanev, V. et al. Machine learning modeling of superconducting critical temperature. *npj Comput. Mater.* **4**, 29 (2018).
12. Jha, D. et al. Elemnet: deep learning the chemistry of materials from only elemental composition. *Sci. Rep.* **8**, 17593 (2018).
13. Sun, Y. T., Bai, H. Y., Li, M. Z. & Wang, W. H. Machine learning approach for prediction and understanding of glass-forming ability. *J. Phys. Chem. Lett.* **8**, 3434–3439 (2017).
14. Li, Y. & Guo, W. Machine-learning model for predicting phase formations of high-entropy alloys. *Phys. Rev. Mater.* **3**, 095005 (2019).
15. Zhou, Z. et al. Machine learning guided appraisal and exploration of phase design for high entropy alloys. *npj Comput. Mater.* **5**, 128 (2019).
16. Huang, W., Martin, P. & Zhuang, H. L. Machine-learning phase prediction of high-entropy alloys. *Acta Mater.* **169**, 225–236 (2019).
17. Chang, Y.-J., Jui, C.-Y., Lee, W.-J. & Yeh, A.-C. Prediction of the composition and hardness of high-entropy alloys by machine learning. *JOM* **71**, 3433–3442 (2019).
18. Kondo, R., Yamakawa, S., Masuoka, Y., Tajima, S. & Asahi, R. Microstructure recognition using convolutional neural networks for prediction of ionic conductivity in ceramics. *Acta Mater.* **141**, 29–38 (2017).
19. Avery, P. et al. Predicting superhard materials via a machine learning informed evolutionary structure search. *npj Comput. Mater.* **5**, 89 (2019).
20. Xiong, J., Zhang, T.-Y. & Shi, S.-Q. Machine learning prediction of elastic properties and glass-forming ability of bulk metallic glasses. *MRS Commun.* **9**, 576–585 (2019).
21. Wen, C. et al. Machine learning assisted design of high entropy alloys with desired property. *Acta Mater.* **170**, 109–117 (2019).
22. Kostiuchenko, T., Körmann, F., Neugebauer, J. & Shapeev, A. Impact of lattice relaxations on phase transitions in a high-entropy alloy studied by machine-learning potentials. *npj Comput. Mater.* **5**, 55 (2019).
23. Ziletti, A., Kumar, D., Scheffler, M. & Ghiringhelli, L. M. Insightful classification of crystal structures using deep learning. *Nat. Commun.* **9**, 2775 (2018).
24. Isayev, O. et al. Universal fragment descriptors for predicting properties of inorganic crystals. *Nat. Commun.* **8**, 1–12 (2017).
25. Feng, S., Zhou, H. & Dong, H. Using deep neural network with small dataset to predict material defects. *Mater. Des.* **162**, 300–310 (2019).
26. Ghiringhelli, L. M., Vybiral, J., Levchenko, S. V., Draxl, C. & Scheffler, M. Big data of materials science: critical role of the descriptor. *Phys. Rev. Lett.* **114**, 105503 (2015).
27. Zhang, Y. & Ling, C. A strategy to apply machine learning to small datasets in materials science. *npj Comput. Mater.* **4**, 25 (2018).
28. Jain, A. et al. Commentary: The materials project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
29. Kirklin, S. et al. The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies. *npj Comput. Mater.* **1**, 15010 (2015).
30. Curtarolo, S. et al. The high-throughput highway to computational materials design. *Nat. Mater.* **12**, 191–201 (2013).
31. Kube, S. A. et al. Phase selection motifs in high entropy alloys revealed through combinatorial methods: large atomic size difference favors BCC over FCC. *Acta Mater.* **166**, 677–686 (2019).
32. Saal, J. E., Oliynyk, A. O. & Meredig, B. Machine learning in materials discovery: confirmed predictions and their underlying approaches. *Annu. Rev. Mater. Res.* **50**, 49–69 (2020).
33. Ward, L. et al. A machine learning approach for engineering bulk metallic glass alloys. *Acta Mater.* **159**, 102–111 (2018).
34. Pan, S. J. & Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **22**, 1345–1359 (2010).
35. Hutchinson, M. L. et al. Overcoming data scarcity with transfer learning. Preprint at https://arxiv.org/abs/1711.05099 (2017).
36. Torrey, L. & Shavlik, J. In: *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques.* (ed. Olivas, E. S.) 242–264 (IGI Global, 2010).
37. Lecun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
38. Inoue, A. & Takeuchi, A. Recent development and application products of bulk glassy alloys. *Acta Mater.* **59**, 2243–2267 (2011).
39. Greer, A. L. Confusion by design. *Nature* **366**, 303–304 (1993).
40. Laws, K. J., Miracle, D. B. & Ferry, M. A predictive structural model for bulk metallic glasses. *Nat. Commun.* **6**, 8123 (2015).
41. Wang, W. H. Bulk metallic glasses with functional physical properties. *Adv. Mater.* **21**, 4524–4544 (2009).
42. Perim, E. et al. Spectral descriptors for bulk metallic glasses based on the thermodynamics of competing crystalline phases. *Nat. Commun.* **7**, 12315 (2016).
43. Li, Z., Zhao, S., Ritchie, R. O. & Meyers, M. A. Mechanical properties of high-entropy alloys with emphasis on face-centered cubic alloys. *Prog. Mater. Sci.* **102**, 296–345 (2019).
44. Yeh, J.-W. Alloy design strategies and future trends in high-entropy alloys. *JOM* **65**, 1759–1771 (2013).
45. George, E. P., Raabe, D. & Ritchie, R. O. High-entropy alloys. *Nat. Rev. Mater.* **4**, 515–534 (2019).
46. Zhang, W., Liaw, P. K. & Zhang, Y. Science and technology in high-entropy alloys. *Sci. China Mater.* **61**, 2–22 (2018).
47. Huang, E., Liaw, P. K. & Editors, G. High-temperature materials for structural applications: new perspectives on high-entropy alloys, bulk metallic glasses, and nanomaterials. *MRS Bull.* **44**, 847–853 (2019).
48. Miracle, D. B. & Senkov, O. N. A critical review of high entropy alloys and related concepts. *Acta Mater.* **122**, 448–511 (2017).
49. Lu, Z. P. & Liu, C. T. A new glass-forming ability criterion for bulk metallic glasses. *Acta Mater.* **50**, 3501–3512 (2002).
50. Takeuchi, A. & Inoue, A. Classification of bulk metallic glasses by atomic size difference, heat of mixing and period of constituent elements and its application to characterization of the main alloying element. *Mater. Trans.* **46**, 2817–2829 (2005).
51. Gao, M. C. et al. Thermodynamics of concentrated solid solution alloys. *Curr. Opin. Solid State Mater. Sci.* **21**, 238–251 (2017).
52. Guo, S. & Liu, C. T. Phase stability in high entropy alloys: formation of solid-solution phase or amorphous phase. *Prog. Nat. Sci. Mater. Int* **21**, 433–446 (2011).
53. Troparevsky, M. C., Morris, J. R., Kent, P. R. C. R. C., Lupini, A. R. & Stocks, G. M. Criteria for predicting the formation of single-phase high-entropy alloys. *Phys. Rev. X* **5**, 011041 (2015).
54. Senkov, O. N., Miller, J. D., Miracle, D. B. & Woodward, C. Accelerated exploration of multi-principal element alloys for structural applications. *Calphad* **50**, 32–48 (2015).
55. Abu-Odeh, A. et al. Efficient exploration of the high entropy alloy composition-phase space. *Acta Mater.* **152**, 41–57 (2018).
56. Lederer, Y., Toher, C., Vecchio, K. S. & Curtarolo, S. The search for high entropy alloys: a high-throughput ab-initio approach. *Acta Mater.* **159**, 364–383 (2018).
57. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90 (2017).
58. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. In: *Third International Conference on Learning Representations, {ICLR} 2015*, San Diego, CA, USA, Conference Track Proceedings. (eds Bengio, Y. & LeCun, Y.) https://iclr.cc/archive/www/2015.html (2015).
59. Szegedy, C. et al. Going deeper with convolutions. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.* 1–9 (Boston, USA, 2015).
60. Miracle, D. B., Louzguine-Luzgin, D. V., Louzguina-Luzgina, L. V. & Inoue, A. An assessment of binary metallic glasses: correlations between structure, glass forming ability and stability. *Int. Mater. Rev.* **55**, 218–256 (2010).
61. Meredig, B. et al. Can machine learning identify the next high-temperature superconductor? Examining extrapolation performance for materials discovery. *Mol. Syst. Des. Eng.* **3**, 819–825 (2018).

62. Li, M.-X. et al. High-temperature bulk metallic glasses developed by combinatorial methods. *Nature* **569**, 99–103 (2019).
63. Shamlaye, K. F., Laws, K. J. & Löffler, J. F. Exceptionally broad bulk metallic glass formation in the Mg–Cu–Yb system. *Acta Mater.* **128**, 188–196 (2017).
64. Kuball, A., Gross, O., Bochtler, B. & Busch, R. Sulfur-bearing metallic glasses: A new family of bulk glass-forming alloys. *Scr. Mater.* **146**, 73–76 (2018).
65. Lin, C.-Y., Tien, H.-Y. & Chin, T.-S. Soft magnetic ternary iron-boron-based bulk metallic glasses. *Appl. Phys. Lett.* **86**, 162501 (2005).
66. Louzguine-Luzgin, D. V. et al. Role of different factors in the glass-forming ability of binary alloys. *J. Mater. Sci.* **50**, 1783–1793 (2015).
67. Murdock, R., Kauwe, S., Wang, A. & Sparks, T. Is domain knowledge necessary for machine learning materials properties? Preprint at https://chemrxiv.org/articles/preprint/Is_Domain_Knowledge_Necessary_for_Machine_Learning_Materials_Properties_/11879193/1 (2020).
68. Miracle, D. B. Efficient local packing in metallic glasses. *J. Non Cryst. Solids* **342**, 89–96 (2004).
69. Takeuchi, A. & Inoue, A. Calculations of mixing enthalpy and mismatch entropy for ternary amorphous alloys. *Mater. Trans. JIM* **41**, 1372–1378 (2000).

AUTHOR CONTRIBUTIONS

S.F., H.F., and H.D. planned the work. S.F. assembled and analyzed data. S.F. wrote the manuscript and H.F., H.Z., Y.W., Z.L., and H.D. revised the manuscript. All authors discussed and commented on the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at https://doi.org/10.1038/s41524-020-00488-z.

Correspondence and requests for materials should be addressed to H.F. or H.D.

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

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