

# Data-Driven Discovery of New Eutectic High-Entropy Alloys via Calculation of Phase Diagrams and Machine Learning Integration

Xiaomin Li and Xizhang Chen\*

**Eutectic high-entropy alloys (EHEAs) exhibit unique microstructures and excellent mechanical properties, but their composition discovery has largely relied on trial-and-error methods. Although the Calculation of Phase Diagrams (CALPHAD) method can predict eutectic points accurately, it is computationally intensive for large-scale screening. Meanwhile, machine learning (ML) offers rapid predictions but is limited by the scarcity of high-quality data for EHEAs. To overcome these challenges, a strategy that integrates CALPHAD and ML to efficiently identify eutectic compositions is developed. CALPHAD is used to compute phase equilibria and build a comprehensive dataset combining calculated and experimental data. This dataset trained an ML model capable of rapidly screening candidate compositions. The approach is validated by predicting and fabricating three new eutectic alloys in the Al–Co–Fe–Ni system and one in the Al–Co–Cr–Fe–Ni system. All four are produced via additive manufacturing, and their microstructures match the predicted eutectic structures. Additionally, minor compositional changes result in notable microstructural differences, highlighting the sensitivity of eutectic formation to composition. The results demonstrate the accuracy and efficiency of this CALPHAD-ML hybrid approach, offering a powerful alternative to conventional methods for accelerating EHEA discovery.**

## 1. Introduction

High-entropy alloys (HEAs) are a class of multi-component alloys, typically mixed in equimolar or near-equimolar ratios.<sup>[1,2]</sup> To be considered a true HEA, the alloy should exhibit a configurational entropy greater than  $1.5 R$  ( $\Delta S_{\text{config}} > 1.5 R$ ), where  $R$  is the universal gas constant.<sup>[3]</sup> Due to their high configurational entropy, these alloys can form various random solid solution structures,<sup>[4,5]</sup> including face-centered cubic (FCC), body-centered cubic (BCC), and hexagonal close-packed (HCP) structures. Determining the phase structures of HEAs and their relationships to composition has been a major focus of advanced materials science research.<sup>[6–8]</sup> Extensive studies on HEAs have revealed many promising properties, making them

potential candidates for a wide range of applications—properties that are largely governed by the four core effects: high-entropy effect, severe lattice distortion, sluggish diffusion, and the cocktail effect.<sup>[3]</sup>

In HEAs, FCC-structured alloys generally exhibit excellent ductility and work-hardening ability,<sup>[9,10]</sup> while BCC-structured alloys are known for their high strength and hardness,<sup>[11,12]</sup> making them suitable for applications requiring superior mechanical properties. On the other hand, HCP-structured alloys demonstrate remarkable corrosion resistance and high-temperature creep performance,<sup>[13]</sup> making them ideal for engineering applications in harsh environments.

However, single-phase HEAs often struggle to meet multiple performance requirements simultaneously. Therefore, introducing dual-phase or multiphase structures has emerged as an effective strategy for optimizing their overall properties. The incorporation of a secondary phase in HEAs is typically achieved through two approaches: (1) precipitation induced by aging treatment<sup>[14,15]</sup> and (2) formation of a eutectic structure via solidification control.<sup>[8,16,17]</sup> Among these, eutectic high-entropy alloys (EHEAs) have attracted significant attention due to their in-situ-formed lamellar composite microstructures. Lu et al. first proposed the concept of EHEAs and successfully fabricated the AlCoCrFeNi<sub>2.1</sub> alloy,<sup>[18,19]</sup> which exhibited an excellent combination of strength and ductility, making it a promising candidate for structural applications. Furthermore, eutectic reactions not only enhance the alloy's fluidity but also significantly reduce compositional segregation,<sup>[20]</sup> thereby broadening the potential applications of EHEAs in materials science.

Despite the remarkable properties of EHEAs and the widespread interest they have garnered in the materials science community,<sup>[21–23]</sup> their design remains a significant challenge, particularly in determining eutectic compositions. In the Al–Co–Fe–Ni system, experimental studies have repeatedly demonstrated the formation of eutectic compositions with a lamellar FCC(L1<sub>2</sub>) + BCC(B2) structure,<sup>[24]</sup> offering new opportunities for materials research and engineering applications. Similarly, the Al–Co–Cr–Fe–Ni system also exhibits eutectic compositions.<sup>[19,25]</sup> However, due to the increased complexity of elemental interactions in this system, the eutectic composition range is

X. Li, X. Chen  
School of Mechanical and Electrical Engineering  
Wenzhou University  
Wenzhou, Zhejiang 325035, China  
E-mail: chenxizhang@wzu.edu.cn

The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adem.202500879>.

DOI: 10.1002/adem.202500879

broader, making the screening process considerably more challenging.

Traditional trial-and-error experimental methods have achieved some success in the design of EHEAs, but they remain highly inefficient. For example, Huang et al.<sup>[26]</sup> synthesized an Al<sub>10.33</sub>Co<sub>21.60</sub>Fe<sub>30.45</sub>Ni<sub>27.62</sub> eutectic alloy using a combination of cable wires and pure aluminum wires. However, this method is constrained by fixed elemental ratios and requires extensive experimentation to identify only a few eutectic points, consequently lacking systematic exploration and compositional breadth. Similarly, Jiang et al.<sup>[27]</sup> achieved a transition from a single FCC phase to a eutectic structure containing Laves phases by adjusting the Zr content. However, their approach still relied on repeated experimental trials, making it difficult to efficiently explore the compositional space.

To overcome these limitations, Calculation of Phase Diagrams (CALPHAD) has been introduced for predicting phase diagrams in high-entropy alloys. However, this method requires inputting elemental combinations individually, and its computational complexity increases exponentially with the number of variables, which makes large-scale screening challenging. For example, Hakan Gasan et al.<sup>[28]</sup> computed equilibrium phase diagrams for 324 compositions but identified only 10 potential eutectic compositions, demonstrating low efficiency. Moreover, CALPHAD relies on fixed thermodynamic databases, making it difficult to incorporate updated experimental data, which limits its flexibility in data-driven alloy design.

In this context, ML offers an efficient and systematic solution for designing high-entropy alloy compositions.<sup>[29,30]</sup> In recent years, ML has demonstrated significant potential in HEA research, achieving remarkable progress in phase structure prediction, mechanical property evaluation, and process optimization.<sup>[31,32]</sup> For example, Zhao et al.<sup>[33]</sup> used generative adversarial networks and extreme gradient boosting (XGBoost) models to accurately predict the microhardness of HEAs, while Duan et al.<sup>[34]</sup> integrated thermodynamic parameters with ML to optimize the laser powder bed fusion process for the fabrication of HEAs. However, despite its promise, the application of ML in the precise prediction of EHEA compositions remains challenging, primarily due to the scarcity of relevant experimental data.<sup>[35]</sup>

Therefore, we propose a design strategy that integrates CALPHAD and ML to achieve efficient composition prediction for EHEAs. Taking the Al–Co–Fe–Ni system as the initial research focus, we constructed a fundamental database using CALPHAD calculations and literature data. A machine learning (ML) model was then used to screen compositions, directly predicting eutectic compositions based on alloy composition input. This approach successfully identified three novel eutectic compositions that were not previously reported in the literature, and we experimentally validated these predictions. Building on this foundation, we extended our method to the Al–Co–Cr–Fe–Ni system by incorporating Cr-related data, successfully predicting and fabricating a new eutectic alloy. This further confirmed the strategy's applicability and reliability in more complex systems. Additionally, slight adjustments to some compositions resulted in significant microstructural differences compared to the original eutectic phases, further revealing the influence of composition on eutectic phase stability.

While originally designed to address the challenge of predicting EHEA compositions prediction, this method is not limited to the Al–Co–Fe–Ni and Al–Co–Cr–Fe–Ni systems. By integrating CALPHAD calculations with ML, the approach exhibits strong generalizability and scalability. Its core principle lies in constructing a system-specific database and training a ML model for composition prediction. Therefore, as long as relevant data—such as phase diagram information or experimental results—are available, this strategy can be applied to other high-entropy alloy systems.

## 2. Experimental Section

To effectively identify eutectic compositions in HEAs systems, this study integrates the CALPHAD method with ML techniques. Initially, we focused on the Al–Co–Fe–Ni system and successfully predicted eutectic HEA compositions. After validating the feasibility of this approach in the Al–Co–Fe–Ni system, we extended it to the more complex Al–Co–Cr–Fe–Ni system. Building on the existing Al–Co–Fe–Ni database, we incorporated CALPHAD-calculated data for Cr and relevant literature sources to construct an expanded database for the Al–Co–Cr–Fe–Ni system. Using this expanded database, we predicted eutectic compositions for the Al–Co–Cr–Fe–Ni system and experimentally validated selected predictions.

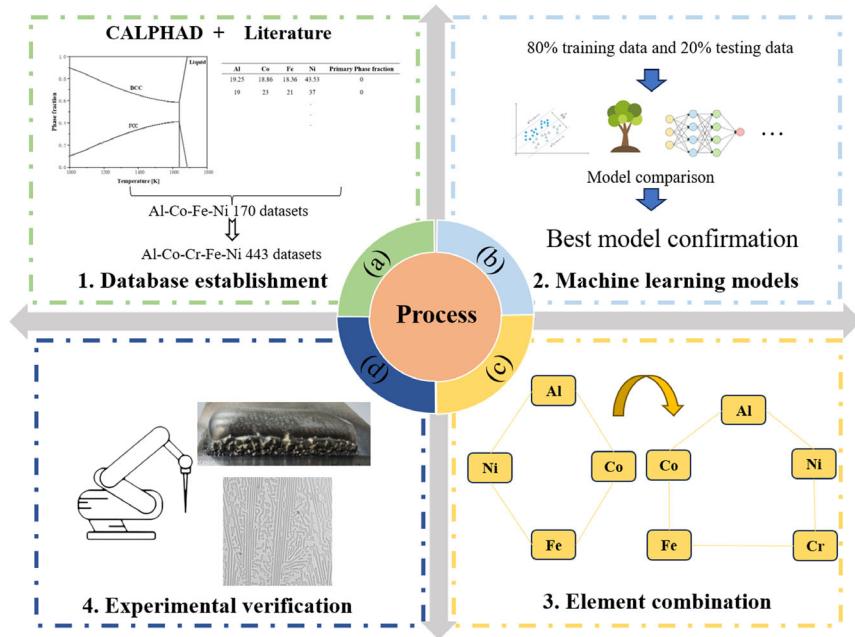
**Figure 1** illustrates the overall strategy for designing EHEAs, which consists of four major steps. 1. Database construction: a targeted database was built by integrating CALPHAD calculations with literature data. 2. ML model training and selection: multiple ML models, including random forest (RF), support vector machines, neural networks, and gradient boosting models, were trained using the database. Their performance was evaluated through cross validation, and the most accurate model was selected as the prediction tool. 3. Eutectic composition prediction: the selected ML model was used to predict potential eutectic compositions by inputting different elemental combinations from the target system. 4. Experimental validation: several predicted eutectic compositions were experimentally synthesized and analyzed to verify the effectiveness of the model.

### 2.1. Database Creation

To predict the eutectic compositions of EHEAs in the target system, an alloy database was constructed to record the elemental compositions and the molar fractions of primary phases. This database combines EHEA composition data from the literature with equilibrium phase diagram data obtained using the CALPHAD method with the TCHEA4 database, thereby providing a data source for ML model development and prediction.

In the Al–Co–Fe–Ni system, the database contains 170 data points, including primary phase data generated from CALPHAD equilibrium phase diagrams and literature data. The model's output variable is defined as the molar fraction of the primary phase, representing the deviation from the eutectic composition:

For EHEAs, the output variable approaches zero. For hypoeutectic and hypereutectic alloys, the molar fraction of the primary FCC(L1<sub>2</sub>) phase is positive, while the molar fraction of the



**Figure 1.** Process diagram for predicting eutectic composition using ML.

primary BCC(B2) phase is negative, with values ranging from -1 to 1.

To further expand the study, the database was expanded to include the Al–Co–Cr–Fe–Ni system by incorporating Cr-related data. The total dataset increased to 443 entries.

**Table 1** presents a selection of literature data referenced in the database.

## 2.2. Model Construction

### 2.2.1. Model Selection

To select the most suitable predictive model, we evaluated eight different ML algorithms, including multiple linear regression (MLR), RF, XGBoost, Gaussian process regression (GPR), K-nearest neighbor regression (KNNR), polynomial kernel support vector regression (PK-SVR), linear kernel support vector regression (LK-SVR), and artificial neural network (ANN). These models were implemented using the Scikit-learn ML library in Python.<sup>[36]</sup> The optimal model for eutectic composition

prediction was selected based on a comparison of  $R^2$  and root mean square error (RMSE) values.

### 2.2.2. Feature Screening

In ML models, the proper selection and combination of input features are critical to prediction accuracy,<sup>[37,38]</sup> as they directly determine the model's ability to capture key patterns in the data. However, feature selection and combination are complex and challenging processes. Improper feature selection not only increases the computational burden but may also significantly reduce prediction accuracy by introducing irrelevant or redundant features.

To address this, Pearson correlation analysis was used as a feature selection tool to systematically identify and select features that are highly correlated with the target variable, thereby optimizing the model's input feature set.

The Pearson correlation coefficient (PCC) is used to measure the linear correlation between two variables,<sup>[39]</sup> and it is calculated using the following formula (1)

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

where  $r$  is the PCC;  $x_i$  and  $y_i$  are the sample values of variables X and Y, respectively;  $\bar{x}$  and  $\bar{y}$  are the sample means of X and Y; and  $n$  is the number of samples.

First, the PCCs for all features were calculated, with 0.9 set as the correlation threshold. To avoid removing too many variables at once, a stepwise elimination strategy was adopted. During correlation analysis, the influence of retained features was excluded, and only the remaining features were analyzed. For highly correlated features ( $r > 0.9$ ), only one of the correlated features was

**Table 1.** Selected eutectic composition data reported in the literature.

Alloy	Al	Co	Cr	Fe	Ni	Reference
Al <sub>19.25</sub> Co <sub>18.86</sub> Fe <sub>18.36</sub> Ni <sub>43.53</sub>	19.25	18.85	0	18.36	43.53	[44]
Al <sub>20.33</sub> Co <sub>21.60</sub> Fe <sub>30.45</sub> Ni <sub>27.62</sub>	20.33	21.60	0	30.45	27.62	[26]
Al <sub>19</sub> Co <sub>23</sub> Fe <sub>21</sub> Ni <sub>37</sub>	19	23	0	21	37	[24]
Al <sub>19</sub> Co <sub>20</sub> Fe <sub>20</sub> Ni <sub>41</sub>	19	20	0	20	41	[45]
Al <sub>18</sub> Co <sub>10</sub> Cr <sub>20</sub> Ni <sub>52</sub>	18	10	20	0	52	[46]
Al <sub>16</sub> Co <sub>41</sub> Cr <sub>15</sub> Fe <sub>10</sub> Ni <sub>18</sub>	16	41	15	10	18	[47]

removed, rather than eliminating all features exceeding the threshold simultaneously. This approach minimized data loss while reducing redundant information in the model.

### 2.2.3. Model Evaluation

To evaluate the performance of the model, two metrics were selected: RMSE and coefficient of Determination ( $R^2$ ). RMSE and  $R^2$  assess the discrepancy between predicted and actual values. A lower RMSE value, closer to 0, indicates better prediction accuracy. The  $R^2$  value ranges between 0 and 1, with values closer to 1 indicating that the predicted values align more closely with the actual values.

The formulas for RMSE and  $R^2$  are as follows

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

where  $n$  is the number of samples,  $y_i$  is the  $i$ -th observed value, and  $\hat{y}_i$  is the  $i$ -th predicted value.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

where  $\bar{y}$  is the mean of the observed values.

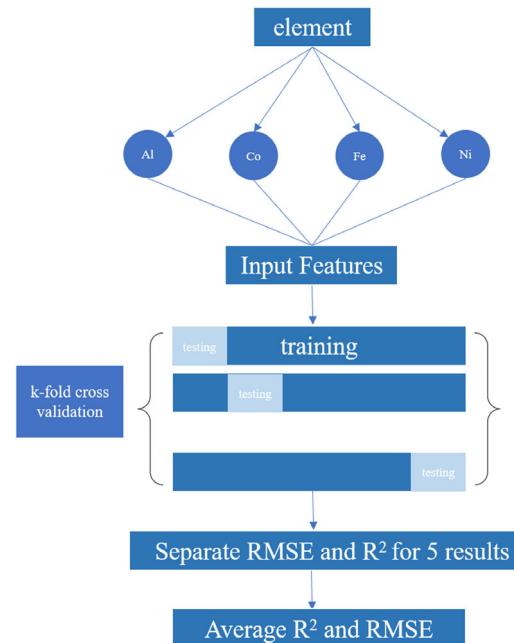
The dataset was randomly split into 80% training data and 20% test data. The training set was used for model training, while the test set was used to evaluate the model's prediction accuracy and generalization ability. To ensure data consistency and reduce biases caused by differences in scale, all data were normalized before splitting.

### 2.2.4. Cross Validation

To comprehensively evaluate model performance and reduce biases caused by specific data subsets, we used 5-fold cross validation. Specifically, the dataset was evenly divided into five subsets, with each round of cross validation selecting one subset as the test set and the remaining four as the training set. This process was repeated five times, ensuring that each subset was used as a test set once, allowing all data points to participate in both model training and validation.

Fivefold cross validation effectively assesses the model's stability across different data splits, enhances its generalization ability, and minimizes biases caused by uneven data partitioning. By evaluating the model's performance across five different train-test combinations, this approach improves the model's predictive accuracy when applied to unseen data.

Taking the Al–Co–Fe–Ni system as an example, the entire process is illustrated in **Figure 2**. The data were first evenly divided into five subsets and then processed using 5-fold cross validation for training and testing. Finally, by aggregating the model performance metrics across different folds, the model with the best generalization capability was selected for subsequent eutectic composition predictions.



**Figure 2.** Cross validation method flow chart.

### 2.3. Composition of Ingredients

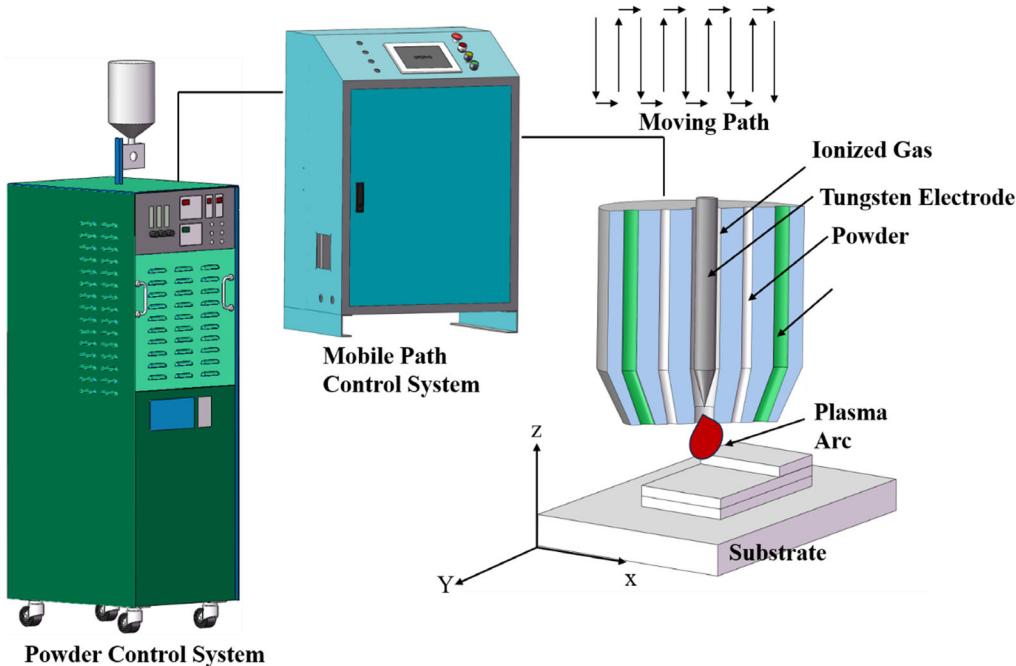
To systematically explore potential EHEA compositions, we used an exhaustive enumeration method to generate all possible element combinations within a predefined composition range, ensuring that the total atomic percentage sums to 100%. Specifically, Python was used to automatically generate all feasible elemental ratios, thereby constructing a dataset that encompasses a broad compositional space.

The generated composition data were subsequently input into the trained ML model for prediction, enabling the identification of alloy compositions with a high probability of forming a eutectic structure.

### 2.4. Sample Preparation

In this study, high-purity (>99.9%) Co, Fe, Ni, and Al elemental powders were used, with particle sizes ranging from 53 to 150  $\mu\text{m}$ . To improve mixing uniformity, all metal powders were first blended in a high-energy ball mill (YXQM12L, MITR Ltd., Hunan, China) for 20 h, followed by vacuum drying in an oven for 100 min. The ball milling process was conducted using 10 mm diameter  $\text{Al}_2\text{O}_3$  ceramic balls, with a ball-to-powder weight ratio of 2:1 and a milling speed of 65  $\text{r min}^{-1}$ . During mixing, the chamber was filled with argon gas as a protective atmosphere to prevent oxidation of the metal powders.

Additive manufacturing was performed using a powder plasma arc additive manufacturing (PPA-AM) system (DML-VO3AD, Duomu Co., Shanghai, China), which consists of a powder control system and a motion path control system, as shown in **Figure 3**. The PPA-AM process used a reciprocating oscillation scanning path. The deposition conditions included a current of



**Figure 3.** Schematic diagram of powder plasma arc additive manufacturing system.

95 A, a scanning speed of 25 mm min<sup>-1</sup>, and a powder feeding gas flow rate (high-purity argon) of 3.0 L min<sup>-1</sup>.

Before the deposition experiment, the substrate (SUS304 stainless steel plate) was cleaned with ethanol to remove surface contaminants and polished using a grinding wheel. Additionally, the substrate was preheated for 4–6 min to ensure a strong bond between the HEA layer and the substrate. To ensure good inter-layer adhesion during the AM process, each deposited layer was ground with a grinding wheel before the next layer was applied.

Subsequently, the microstructure of the samples was thoroughly examined using high-resolution scanning electron microscopy (SEM). Prior to observation, the samples were polished step by step using SiC sandpapers of varying grit sizes (320#, 600#, 800#, 1000#, 1200#, 1500#, and 2000#) to achieve a smooth surface finish. Final polishing was performed using W0.25 diamond spray, ensuring a scratch-free and residue-free surface. After polishing, the samples were etched with aqua regia (a 3:1 mixture of hydrochloric acid and nitric acid) to reveal their metallographic structure.

### 3. Results

#### 3.1. Al–Co–Fe–Ni Eutectic Alloy

##### 3.1.1. Feature Screening Results

In this study, 20 physical parameters related to HEAs were collected from the literature,<sup>[40,41]</sup> (see Table 2). To optimize feature input and enhance the model's generalization ability, we applied the stepwise elimination method described earlier to systematically remove highly correlated features. This approach

minimizes data loss while reducing redundant information in the model.

It is also noteworthy that certain thermodynamic, configurational, and electronic parameters—such as atomic size mismatch ( $\delta$ ), electronegativity difference ( $\Delta\chi$ ), mixing enthalpy ( $\Delta H_{mix}$ ), mixing entropy ( $\Delta S_{mix}$ ), and valence electron concentration (VEC)—play critical roles in the formation of HEA phases.<sup>[3]</sup> These parameters are widely recognized as key indicators influencing phase stability and mechanical behavior. Therefore, incorporating such features is essential for accurate phase prediction in HEAs.

Based on this method, a heatmap including all features was constructed (see Figure 4) to visually represent the correlation distribution among variables.

Through heatmap analysis, we identified high correlations among certain features. To optimize model input and reduce redundancy, a stepwise feature selection method was used. As a result, ten physical parameters with relatively low intercorrelation were retained:  $\Lambda$ ,  $\gamma$ ,  $\eta$ ,  $W$ ,  $D.G$ ,  $\mu$ ,  $\delta r$ , VEC, and  $\Delta H, \Delta S$ . These selected features are visualized in the filtered PCC heat map shown in Figure 5. This selection ensured that the model inputs were both diverse and representative, while minimizing the influence of redundant variables on model performance.

##### 3.1.2. ML Comparison

Figure 6 presents the performance of different ML models in predicting the primary phase fraction of HEAs after 5-fold cross validation, including the average coefficient of determination ( $R^2$ ) and average RMSE. The results indicate that most models achieved relatively high  $R^2$  values, demonstrating both the

**Table 2.** Physical characteristics and calculation formula of high-entropy alloy system.

Physical feature name	Formula
Atomic radius difference	$\delta r = \sqrt{\sum_{i=1}^n C_i * (1 - \frac{r}{r_i})^2}$
Electronegativity difference	$\Delta X = \sqrt{\sum_{i=1}^n C_i (X - X_i)^2}$
Valence electron concentration	$VEC = \sum_{i=1}^n C_i * VEC_i$
Mixing enthalpy	$\Delta H = \sum_{i=1, j>i}^n 4C_i C_j * \Delta H_{ij}^{mix}$
Configurational entropy	$\Delta S = -R \sum_{i=1}^n C_i * \ln(C_i)$
$\Omega$ parameter	$\Omega = Tm * \frac{\Delta S}{ \Delta H }$
$\Lambda$ parameter	$\Lambda = \frac{\Delta S}{\delta r * \delta r}$
$\gamma$ parameter	$\gamma = \frac{\left(1 - \sqrt{\frac{(r+r_{min})^2 - r^2}{(r+r_{min})^2}}\right)}{\left(1 - \sqrt{\frac{(r+r_{max})^2 - r^2}{(r+r_{max})^2}}\right)}$
Local electronegativity mismatch	$D.x = \sum_{i=1}^n \sum_{j=1, j \neq i}^n C_i C_j *  X_i - X_j $
Number of conduction electrons	$e/a = \sum_{i=1}^n C_i * (e_a)_i$
Cohesive energy	$E_c = \sum_{i=1}^n C_i * E_c_i$
Modulus mismatch	$\eta = \sum_{i=1}^n C_i * \frac{2(G_i - G)}{G_i + G} + 0.5 *  C_i * \frac{2(G_i - G)}{G_i + G} $
Local size mismatch	$D.r = \sum_{i=1}^n \sum_{j=1, j \neq i}^n C_i C_j *  r_i - r_j $
Energy term	$A = G * \delta r * (1 + \mu) * (1 - \mu)$
Navarro coefficient	$F = \frac{2G}{1-\mu}$
Work function	$W = (\sum_{i=1}^n C_i * w_i)^6$
Shear modulus	$G = \sum_{i=1}^n C_i G_i$
Shear modulus difference	$\delta G = \sqrt{\sum_{i=1}^n C_i * (1 - \frac{G_i}{G})^2}$
Local modulus mismatch	$D.G = \sum_{i=1}^n \sum_{j=1, j \neq i}^n C_i C_j *  G_i - G_j $
Lattice distortion energy	$\mu = \frac{1}{2} E * \delta r$

effectiveness of integrating ML with the CALPHAD method and the validity of our feature selection approach.

Further analysis of the RMSE results shows that the Extreme Gradient Boosting (XGBoost) model outperformed all other reference models, exhibiting the lowest RMSE value, indicating minimal prediction error. Based on these comprehensive performance metrics, XGBoost was selected as the ML model for this study.

### 3.1.3. Ingredient Exploration

In this section, we used a combination method to generate the composition data for Al–Co–Fe–Ni HEAs, with the formula  $Al_x Co_y Fe_z Ni_{100-x-y-z}$

where x, y, and z represent the molar percentages of Al, Co, and Fe, respectively, with values ranging from 5 to 40% in increments of 0.5%. The proportion of Ni is automatically adjusted based on the amounts of Al, Co, and Fe to ensure that the total composition sums to 100%.

Through this systematic composition generation strategy, we generated over 340 000 distinct Al–Co–Fe–Ni high-entropy alloy compositions and used a ML model to predict their primary phases. Compositions with predicted molar fractions between –0.01 and 0.01 were identified as potential EHEAs. From the

entire dataset, 683 compositions were selected as candidates for further experimental validation.

The elemental frequency statistics for these 683 eutectic HEA compositions are shown in Figure 7. Analysis reveals that eutectic formation is possible only when the Al content is between 13% and 21%. Specifically, Al content is predominantly concentrated in the 13–21% range, while Ni content follows an approximately normal distribution, primarily between 20 and 60%. The Co and Fe contents exhibit a broader range, spanning 5–40%, within which eutectic structures can form. These findings suggest that Al is a key element in eutectic structure formation.<sup>[42,43]</sup>

### 3.1.4. Experimental Verification

In the experimental validation phase of this study, we aimed to assess the accuracy of the ML model in predicting eutectic compositions in HEAs. Based on the model's predictions, three compositions from the Al–Co–Fe–Ni system, which had not been previously reported in the literature, were randomly selected for experimental validation:  $Al_{19}Co_{21}Fe_{21}Ni_{39}$ ,  $Al_{18}Co_{18}Fe_{21}Ni_{43}$ , and  $Al_{18}Co_{15}Fe_{32}Ni_{35}$ . The model predicted that these compositions had a high probability of forming eutectic microstructures.

The selected alloy samples were fabricated using additive manufacturing technology and subsequently analyzed using SEM. As shown in Figure 8, all three samples exhibited a lamellar and rod-like eutectic microstructure, which is consistent with the model's predictions. This heterogeneous lamellar/rod-like microstructure is a typical characteristic of EHEAs.

After confirming the presence of a distinct eutectic microstructure in the samples, X-ray diffraction (XRD) analysis was performed (Figure 9). The XRD spectra revealed a two-phase structure corresponding to the eutectic microstructure, further validating the reliability of the model's predictions.

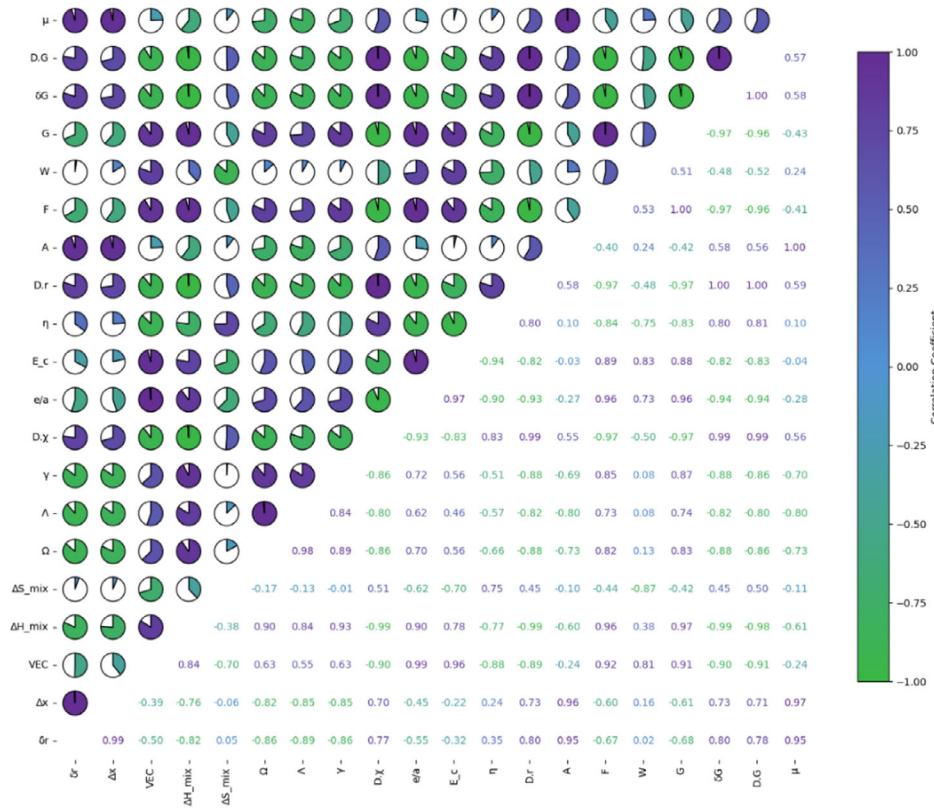
To further verify the accuracy of the model, we fine-tuned the alloy compositions and fabricated two additional alloys for comparative analysis:  $Al_{16}Co_{20}Fe_{21}Ni_{43}$  and  $Al_{20}Co_{20}Fe_{21}Ni_{39}$ .

Microstructural characterization results (Figure 10) revealed that these two compositions exhibited significant differences from the previously identified eutectic microstructure, with the presence of distinct primary phases. The  $Al_{20}Co_{20}Fe_{21}Ni_{39}$  alloy displayed a typical hypoeutectic structure, characterized by a large number of primary FCC dendrites, with lamellar eutectic structures only present in the interdendritic regions. The  $Al_{16}Co_{20}Fe_{21}Ni_{43}$  alloy exhibited a similar microstructure, but its primary phase was an ordered B2 phase.

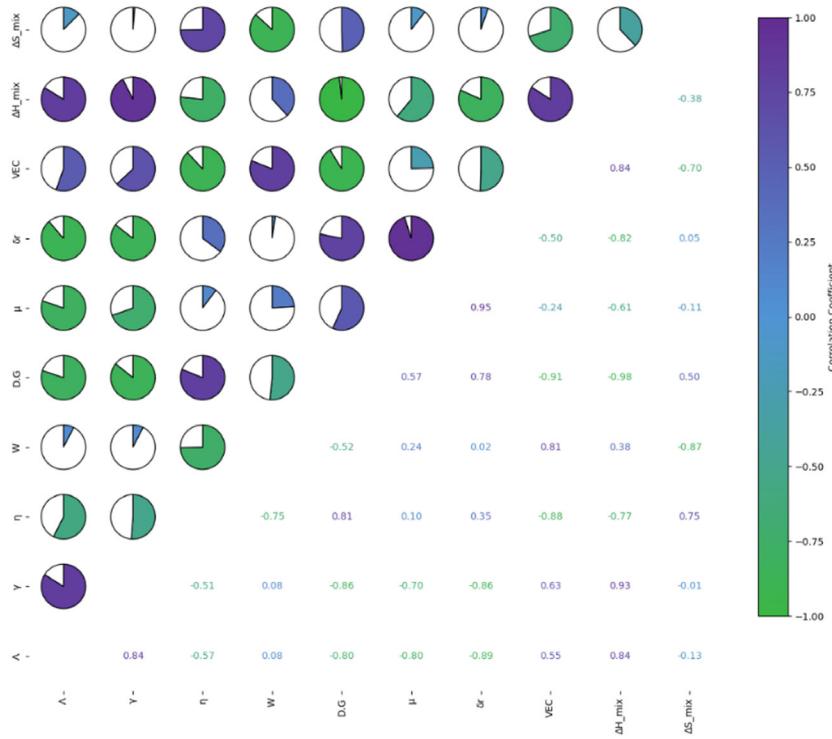
In summary, the experimental results demonstrate that the formation of eutectic microstructures is highly sensitive to alloy composition. Moreover, these findings further validate the effectiveness of the developed ML model in predicting eutectic compositions.

### 3.1.5. Relationship Between Elements and Eutectic Formation

In this study, to explore the effect of alloy element molar percentages on the molar fraction of the primary phase, we identified certain consistent trends within specific composition ranges. Specifically, we investigated the predicted results when



**Figure 4.** PCC heat map of each feature input.



**Figure 5.** PCC heat map of filtered feature input.

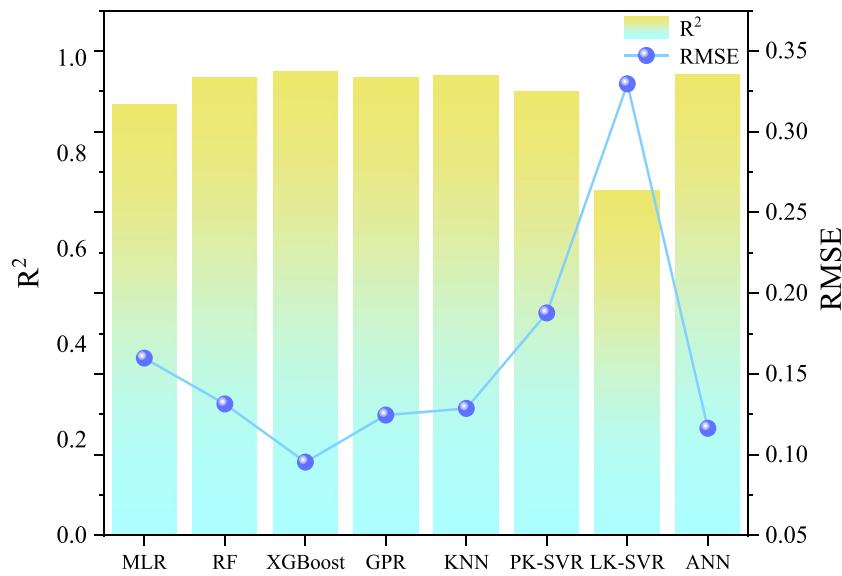


Figure 6.  $R^2$  and RMSE of different ML models after 5-fold cross validation.

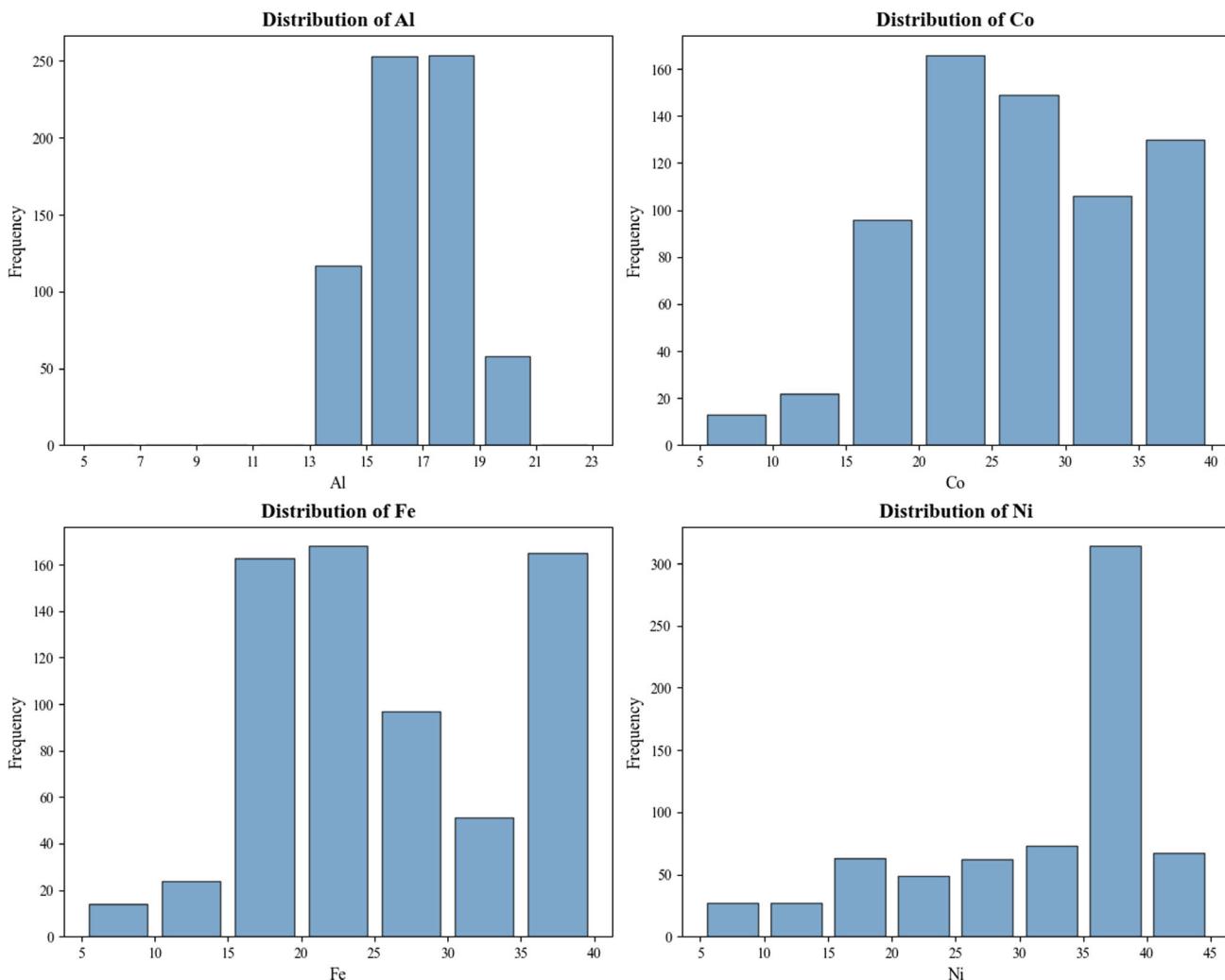
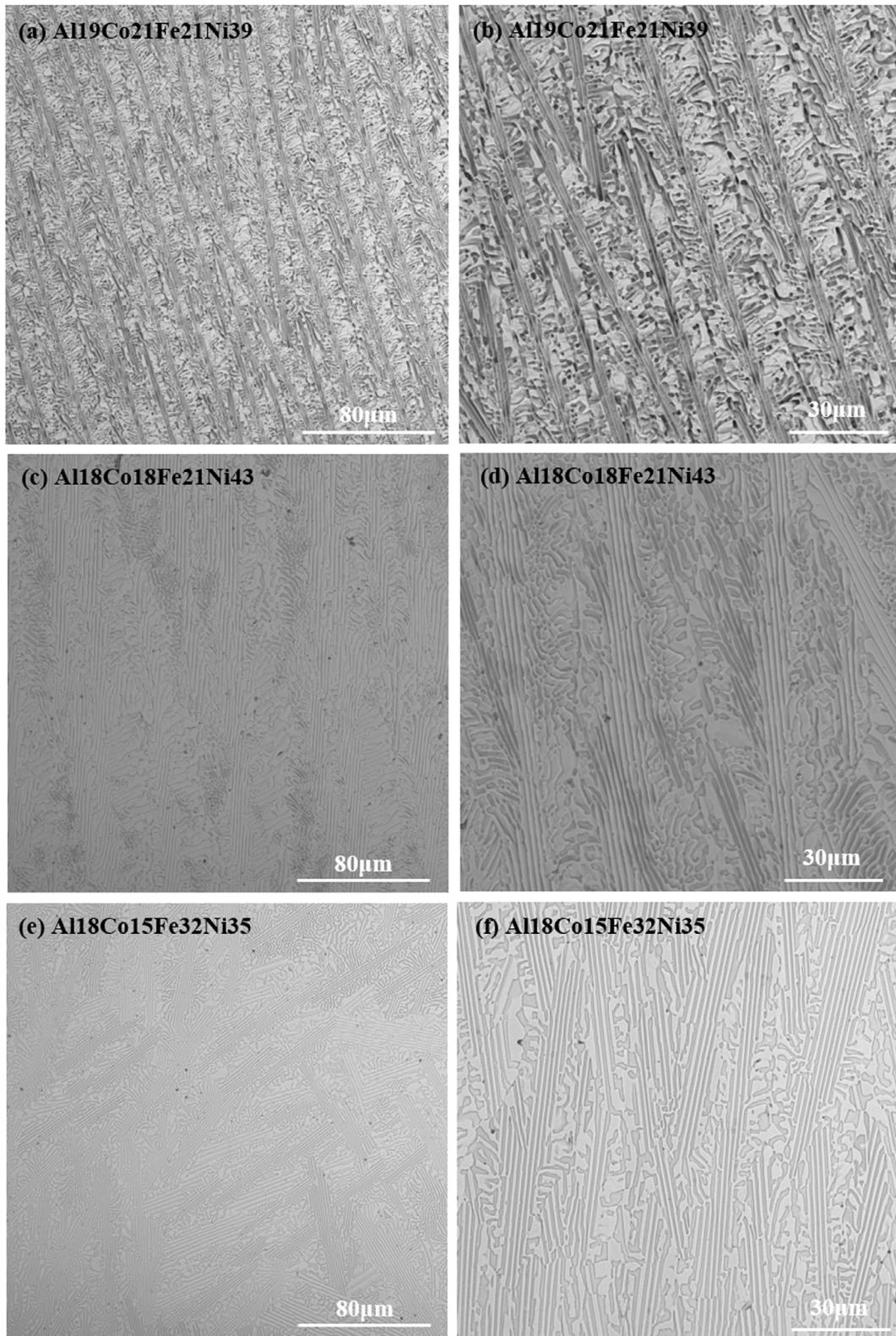


Figure 7. Histogram of eutectic element statistics predicted by the model.



**Figure 8.** Microstructure: a,b)  $\text{Al}_{19}\text{Co}_{21}\text{Fe}_{21}\text{Ni}_{39}$ ; c,d)  $\text{Al}_{18}\text{Co}_{18}\text{Fe}_{21}\text{Ni}_{43}$ ; e,f)  $\text{Al}_{18}\text{Co}_{15}\text{Fe}_{32}\text{Ni}_{35}$ .

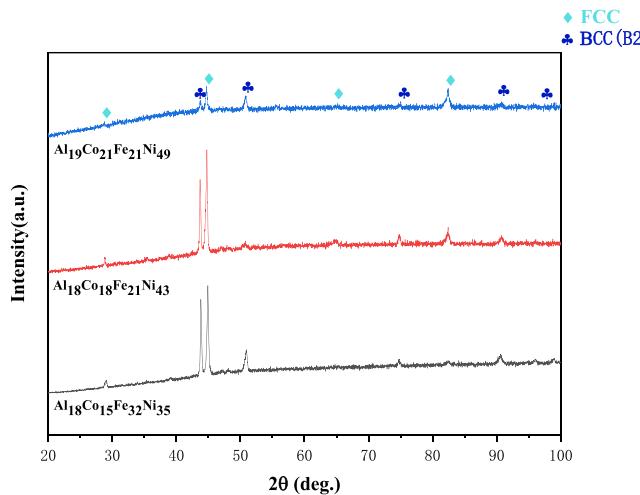


Figure 9. XRD pattern of alloy sample.

substituting Co for Fe and Ni, Fe for Co and Ni, and Ni for Co and Fe under different Al contents.

Taking Co substitution for Fe and Ni as an example, we selected different Al contents (Al<sub>x</sub>, where x = 16, 17, 18, 19) and set different Co contents (Co<sub>y</sub>, where y = 5, 10, 15, 20,

25, 30, 35), while ensuring that the total Fe and Ni content remained 100 - x - y. Through this analysis, we aimed to systematically confirm the specific influence of elemental ratios on eutectic primary phase composition.

We observed that in regions with lower Al content, the primary FCC(L1<sub>2</sub>) phase appeared more frequently, whereas in regions with higher Al content, the primary BCC(B2) phase was more common. Additionally, as Co content increased while Fe and Ni proportions decreased, the fraction of the B2 primary phase increased. Similarly, when Fe content increased while Co and Ni proportions decreased, the B2 primary phase fraction also increased. In contrast, when Ni content increased while Co and Fe proportions decreased, the FCC primary phase fraction increased.

These observations are summarized in Figure 11(a–c), which illustrates the impact of adjusting Co, Fe, and Ni proportions on primary phase formation under different Al contents. The importance of each element in eutectic phase formation follows the order: Al > Ni > Fe > Co. This result aligns well with the statistical distribution of eutectic HEA compositions in Figure 7 and corresponds to the SHAP values presented in Figure 11(d).

SHAP (SHapley Additive exPlanations) values provide a method for explaining model outputs by quantifying the contribution of each feature to the prediction. Based on cooperative game theory and Shapley values, SHAP ensures a fair allocation

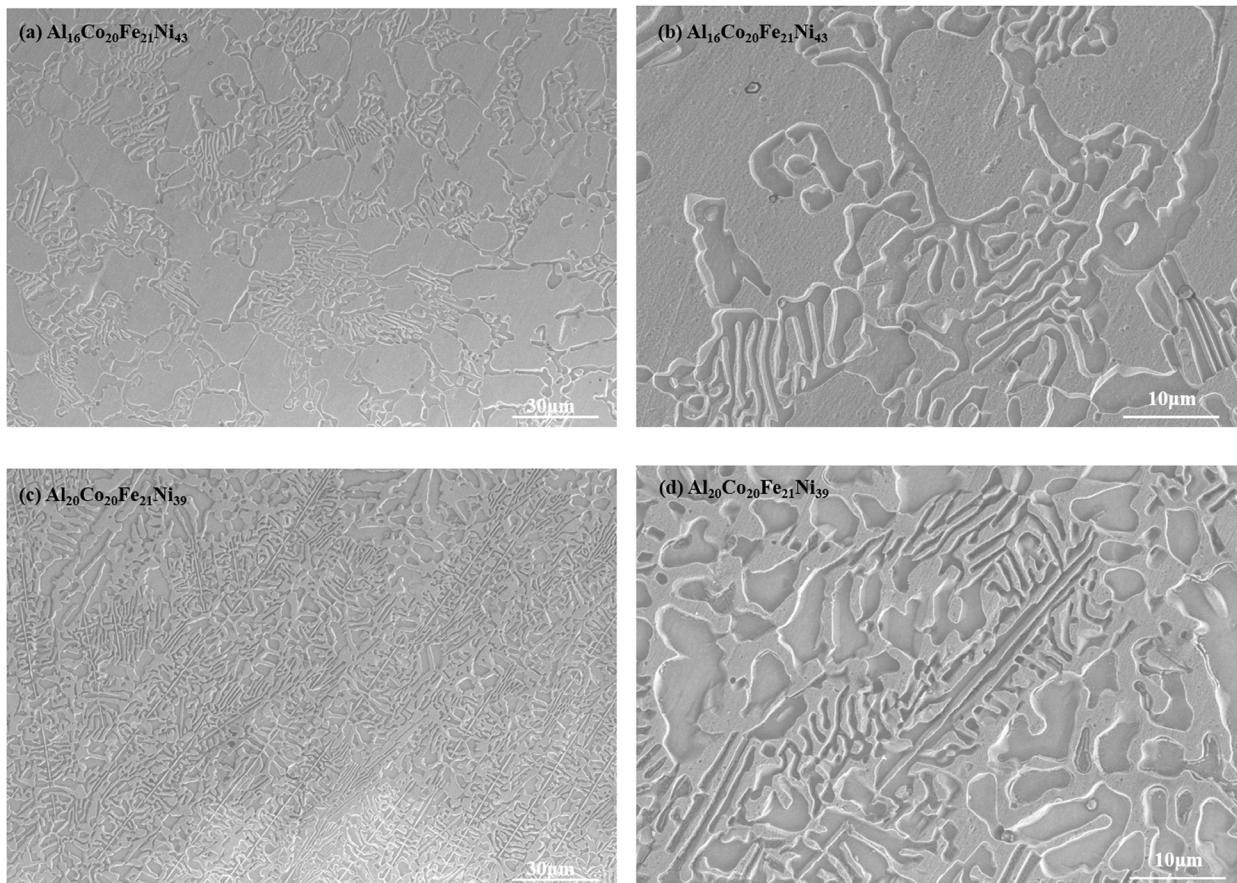
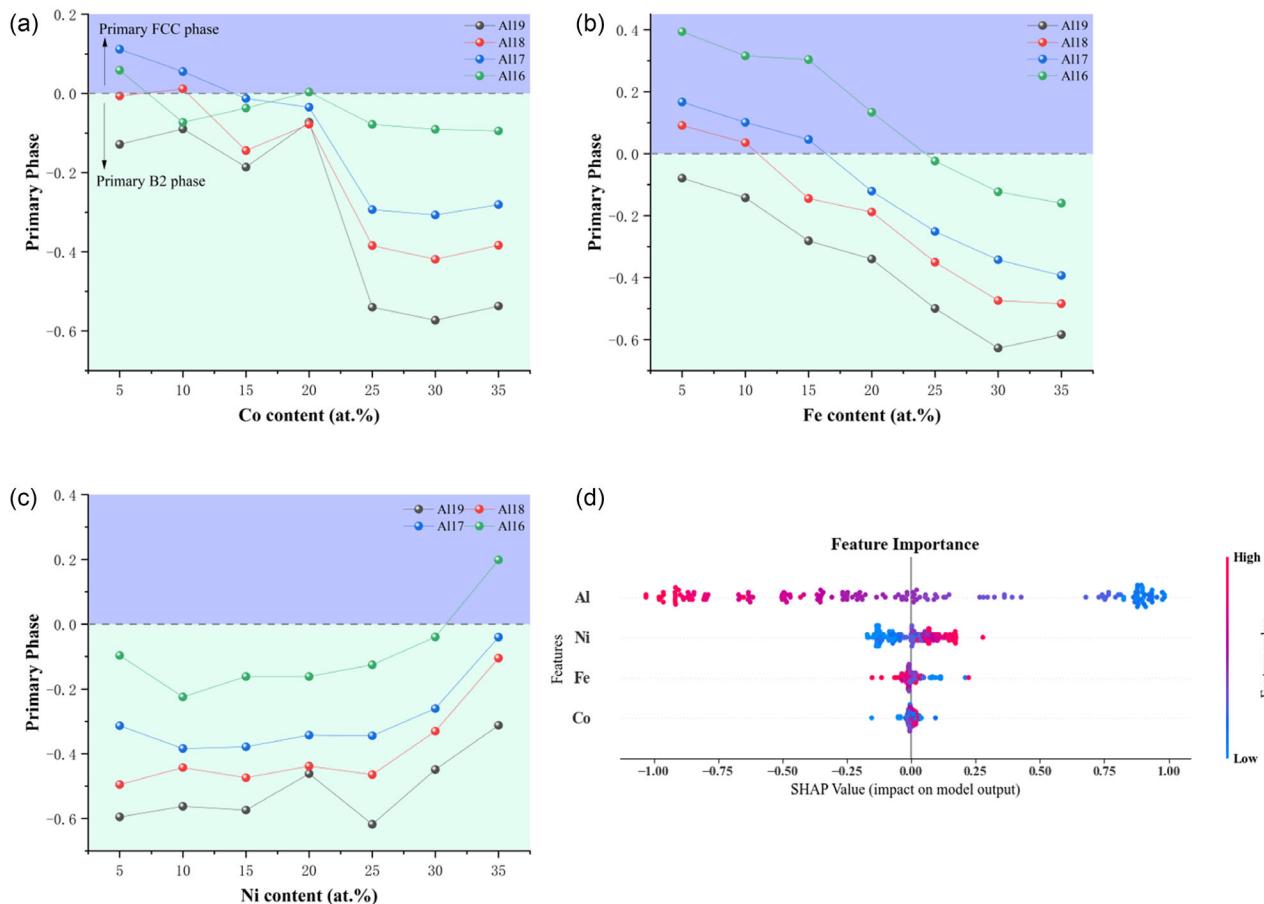


Figure 10. Microstructure. a,b) Al<sub>16</sub>Co<sub>20</sub>Fe<sub>21</sub>Ni<sub>43</sub>. c,d) Al<sub>20</sub>Co<sub>20</sub>Fe<sub>21</sub>Ni<sub>39</sub>.



**Figure 11.** a) Changes in the initial phase mole fraction when Co replaces FeNi in the Al–Co–Fe–Ni system. b) Changes in the initial phase mole fraction when Fe replaces CoNi in the Al–Co–Fe–Ni system. c) Changes in the initial phase mole fraction when Ni replaces CoFe in the Al–Co–Fe–Ni system. d) SHAP value distribution of different elements.

of feature contributions, allowing for a deeper understanding of how the model functions, especially in complex ML models. This method helps us gain insight into the role of each element in predicting eutectic compositions and enhances our understanding of the model's internal mechanisms.

### 3.2. Al–Co–Cr–Fe–Ni Eutectic Alloy

#### 3.2.1. ML Comparison

Using the same ten selected features from the previous feature selection process as input variables, we applied the same eight independent ML algorithms. **Figure 12** presents scatter plots comparing the predicted versus actual values for different ML models. The results indicate that most models still performed well, further demonstrating the effectiveness of our framework for high-entropy alloy prediction.

#### 3.2.2. Ingredient Exploration

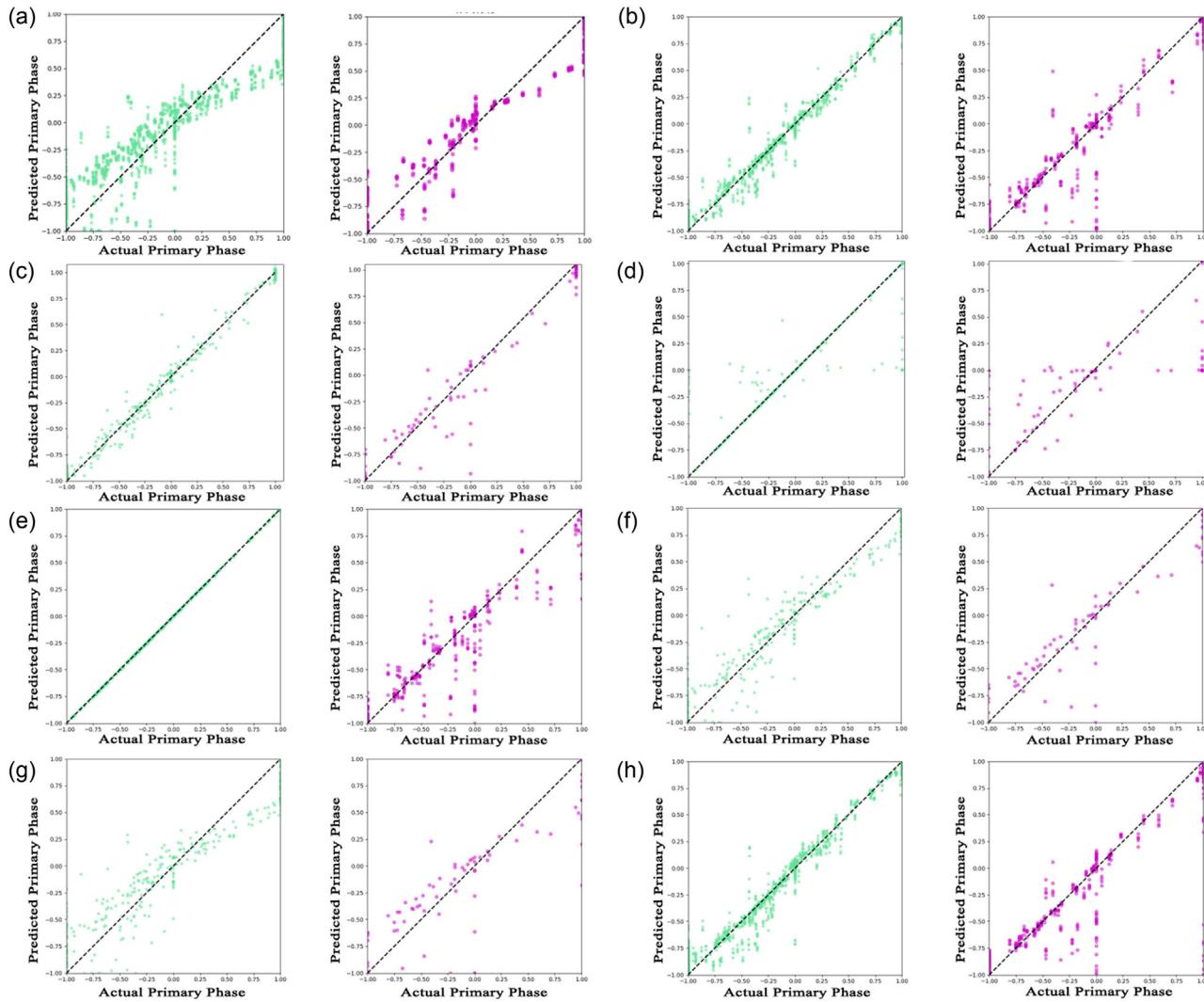
For the Al–Co–Cr–Fe–Ni system, a similar approach was adopted. The composition range for each element was set

between 0 and 35%, ensuring that the total elemental sum remained 100%. The step size for the exhaustive process was set to 0.5%, which effectively controlled the number of combinations while maintaining comprehensive and reasonable composition coverage. A total of 120 208 756 data points were generated,  $\approx$ 120 million compositions.

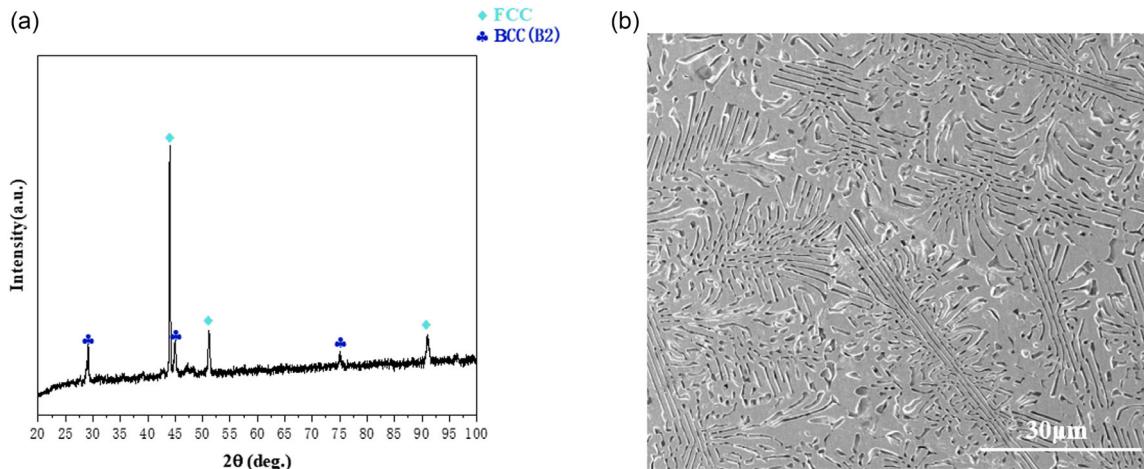
#### 3.2.3. Experimental Verification

In the previous experimental validation phase, the accuracy of the model was fully confirmed. Based on this, a composition predicted as eutectic by the model was selected from the composition exploration dataset: Al<sub>18</sub>Co<sub>18</sub>Cr<sub>19.5</sub>Fe<sub>10.5</sub>Ni<sub>34</sub>.

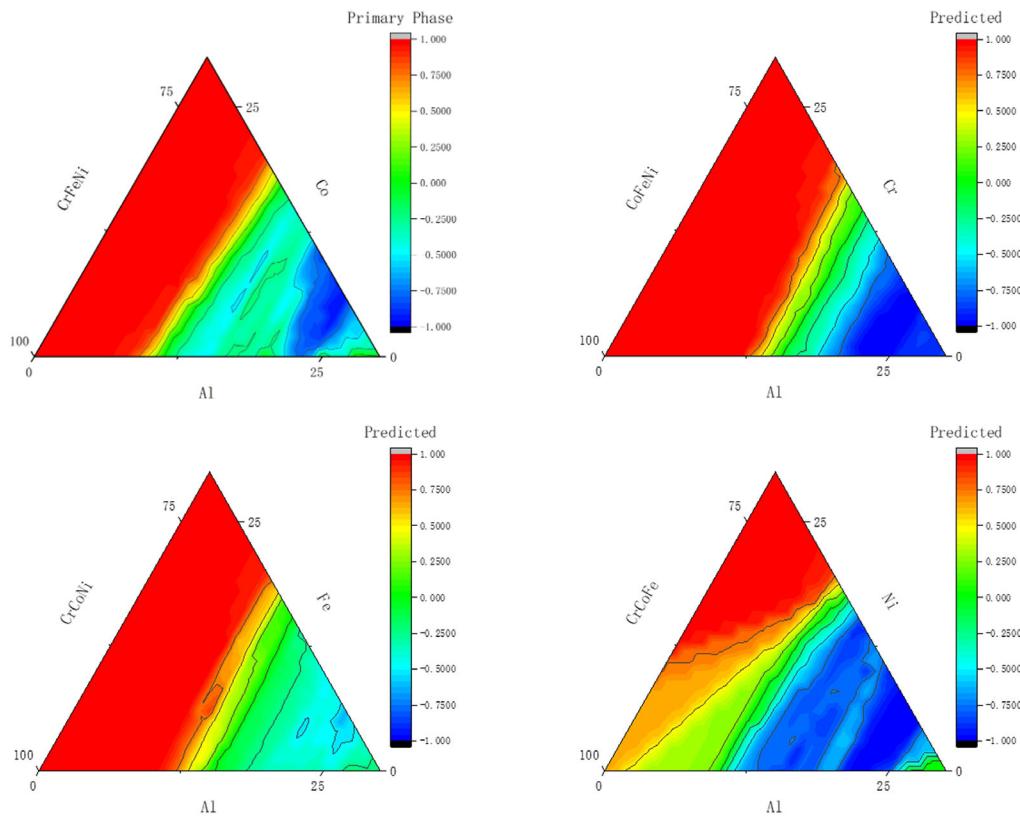
As observed in the SEM image (**Figure 13(b)**), this alloy sample exhibits a typical eutectic microstructure, characterized by uniform distribution and clear phase interfaces, further verifying its eutectic nature. Additionally, XRD analysis (**Figure 13(a)**) shows that the alloy primarily consists of B2 and FCC phases, with no detected intermetallic compound diffraction peaks. This result indicates that its solidification process followed the L  $\rightarrow$  B2 + FCC eutectic reaction pathway, and the alloy maintained a stable dual-phase eutectic structure at room



**Figure 12.** Scatter plot comparison of predicted vs. actual values for different ML models a) MLR; b) RF; c) XGBoost; d) GPR; e) KNN; f) PK-SVR; g) LK-SVR; h) ANN (green: training set; pink: test set).



**Figure 13.**  $\text{Al}_{18}\text{Co}_{18}\text{Cr}_{19.5}\text{Fe}_{10.5}\text{Ni}_{34}$ . a) XRD pattern. b) Microstructure.



**Figure 14.** Contour diagram of the effects of different elements on the primary phase.

temperature, confirming that this composition belongs to the EHEA system.

#### 3.2.4. Relationship Between Elements and Eutectic Formation

**Figure 14** shows the ternary contour maps of Al and other main components (Co, Cr, Fe, and Ni) in the Al–Co–Cr–Fe–Ni system. These contour plots depict the distribution trends of primary phases in EHEAs under various compositional conditions. The results show that the eutectic formation window is mainly distributed in the medium Al content area (about 12 ≈ 20%) and shows an obvious shrinkage trend with the increase of Cr and Ni content, while the addition of Co and Fe has a relatively small effect on the eutectic window and has a certain promoting effect on eutectic formation. This contour analysis clarifies the influence of each element on eutectic formation and provides an important theoretical basis and reference for screening compositions in ML-based eutectic HEA design.

## 4. Conclusion

In this study, the CALPHAD method was integrated with ML techniques to successfully predict the eutectic compositions of Al–Co–Fe–Ni and Al–Co–Cr–Fe–Ni HEA systems. By combining CALPHAD computational data with literature data, a database containing the molar fractions of primary phases was

established, and the XGBoost ML model was used for efficient eutectic composition prediction. A stepwise feature elimination strategy was implemented during modeling to avoid overly complex feature engineering. The main conclusions are as follows. (1) An efficient eutectic composition prediction method. The proposed approach overcomes the time and resource constraints of traditional experimental design, providing a fast and accurate eutectic composition prediction strategy. The incorporation of ML significantly improved prediction efficiency and revealed the influence of key features on model output. (2) Experimental validation confirmed the reliability of model predictions. The predicted alloy compositions were fabricated using additive manufacturing, and their microstructures were analyzed. The results showed a high consistency between the observed microstructures and the model's predictions, demonstrating the effectiveness of this approach in eutectic composition prediction. This finding indicates that the method can be applied to guide the design and optimization of novel HEAs. (3) Influence of elemental ratios on eutectic formation. By fine-tuning the elemental ratios, it was found that small compositional changes significantly affected the eutectic microstructure. This observation highlights both the sensitivity of eutectic structures to composition variations and the accuracy of the model in capturing these effects. (4) The combined framework of ML and CALPHAD has broad applicability. This study demonstrates the great potential of integrating ML with the CALPHAD method for eutectic composition prediction in HEAs.

Additionally, it provides a transferable strategy for exploring other complex alloy systems.

## Acknowledgements

This project was supported by the National Natural Science Foundation of China (grant no. 52261135544) and the Major Science and Technology Innovation Project of Wenzhou City (grant no. ZG2024035)

## Conflict of Interest

The authors declare no conflict of interest.

## Author Contributions

**Xiaomin Li:** conceptualization; investigation; methodology; writing—original (draft). **Xizhang Chen:** conceptualization; supervision; funding acquisition.

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Keywords

CALPHAD, compositional design, high-entropy alloys, machine learning

Received: March 31, 2025

Revised: July 10, 2025

Published online:

- [1] J. Yeh, S. Chen, S. Lin, J. Gan, T. Chin, T. Shun, C. Tsau, S. Chang, *Adv. Eng. Mater.* **2004**, *6*, 299.
- [2] E. P. George, D. Raabe, R. O. Ritchie, *Nat. Rev. Mater.* **2019**, *4*, 515.
- [3] P. Mandal, A. Choudhury, A. B. Mallick, M. Ghosh, *Met. Mater. Int.* **2023**, *29*, 38.
- [4] Z. Li, S. Zhao, R. O. Ritchie, M. A. Meyers, *Prog. Mater. Sci.* **2019**, *102*, 296.
- [5] Y. Zhang, T. T. Zuo, Z. Tang, M. C. Gao, K. A. Dahmen, P. K. Liaw, Z. P. Lu, *Prog. Mater. Sci.* **2014**, *61*, 1.
- [6] Q. Guo, X. Xu, X. Pei, Z. Duan, P. K. Liaw, H. Hou, Y. Zhao, *J. Mater. Res. Technol.* **2023**, *22*, 3331.
- [7] A. Abu-Odeh, E. Galvan, T. Kirk, H. Mao, Q. Chen, P. Mason, R. Malak, R. Arróyave, *Acta Mater.* **2018**, *152*, 41.
- [8] X. Ye, Z. Diao, H. Lei, L. Wang, Z. Li, B. Li, J. Feng, J. Chen, X. Liu, D. Fang, *Mater. Sci. Eng.: A* **2024**, *889*, 145815.
- [9] B. Liu, L. Liu, X. Cao, S. Wang, W. Chen, Q. Jiang, Y. Zhang, F. Wu, J. Shang, R. Zhao, J. Qi, *Mater. Sci. Eng.: A* **2024**, *909*, 146829.
- [10] Y. Yang, T. Chen, L. Tan, J. D. Poplawsky, K. An, Y. Wang, G. D. Samolyuk, K. Littrell, A. R. Lupini, A. Borisevich, E. P. George, *Nature* **2021**, *595*, 245.
- [11] Z. An, S. Mao, Y. Liu, L. Wang, H. Zhou, B. Gan, Z. Zhang, X. Han, *J. Mater. Sci. Technol.* **2021**, *79*, 109.
- [12] F. Liu, P. Liaw, Y. Zhang, *Metals* **2022**, *12*, 501.
- [13] Z. Zhang, E. Axinte, W. Ge, C. Shang, Y. Wang, *Mater. Des.* **2016**, *108*, 106.
- [14] Q. Wang, X. Wu, L. Gu, H. Ding, R. Chen, J. Guo, *Intermetallics* **2024**, *171*, 108365.
- [15] B. Zhang, Y. Cai, W. Mu, W. Guo, *J. Alloys Compd.* **2024**, *970*, 172502.
- [16] X. Jiang, Y. Li, P. Shi, Y. Yang, M. Wang, J. Huang, Y. Qin, Y. Lin, B. Tan, Y. Ruan, X. Wang, B. Zhou, B. Ding, Q. Li, Z. Shen, T. Zheng, C. Liu, P. K. Liaw, Y. Zhong, *J. Mater. Res. Technol.* **2024**, *28*, 4440.
- [17] S. B. Kang, G. Huang, G. Singhal, D. Xie, D. H. Hsieh, Y. Lee, A. A. Kulkarni, J. W. Smith, Q. Chen, K. Thornton, S. Sinha, P. V. Braun, *Adv. Mater.* **2024**, *36*, 2308720.
- [18] Y. Lu, Y. Dong, S. Guo, L. Jiang, H. Kang, T. Wang, B. Wen, Z. Wang, J. Jie, Z. Cao, H. Ruan, T. Li, *Sci. Rep.* **2014**, *4*, 6200.
- [19] Y. Lu, Y. Dong, H. Jiang, Z. Wang, Z. Cao, S. Guo, T. Wang, T. Li, P. K. Liaw, *Scr. Mater.* **2020**, *187*, 202.
- [20] Q. Cai, C. Fang, C. Mendis, I. T. H. Chang, B. Cantor, *J. Alloys Compd.* **2023**, *941*, 168942.
- [21] J. Liu, Z. Li, D. Lin, Z. Tang, X. Song, P. He, S. Zhang, H. Bian, W. Fu, Y. Song, *J. Mater. Sci. Technol.* **2024**, *189*, 211.
- [22] Q. Wu, F. He, J. Li, H. S. Kim, Z. Wang, J. Wang, *Nat. Commun.* **2022**, *13*, 4697.
- [23] Z. Chen, Y. Yang, *J. Mater. Inf.* **2023**, *3*, 10.
- [24] X. Wang, Z. An, J. Cai, C. Jiang, H. Su, X. Luo, Z. Li, S. Wu, L. Yang, H. Long, J. Zhang, S. Mao, Z. Zhang, X. Han, *Mater. Charact.* **2023**, *203*, 113059.
- [25] A. Huang, L. Li, X. Liu, H. Zhang, M. Li, X. Zhang, J. Lu, X. Zhao, *J. Alloys Compd.* **2024**, *992*, 174597.
- [26] Y. Huang, X. Chen, S. Ma, M. Wen, Y. Wang, Y. Wang, *Mater. Lett.* **2024**, *363*, 136312.
- [27] J. Jiang, A. Zhang, J. Han, B. Xin, J. Meng, *Tribol. Int.* **2024**, *191*, 109154.
- [28] H. Gasan, A. Ozcan, *Met. Mater. Int.* **2020**, *26*, 1152.
- [29] C. Zhou, Y. Zhang, J. Stasic, Y. Liang, X. Chen, M. Trtica, *Adv. Eng. Mater.* **2023**, *25*, 2201369.
- [30] B. Steingrimsson, X. Fan, B. Adam, P. K. Liaw, *Adv. Eng. Mater.* **2023**, *25*, 2201903.
- [31] Y. Yan, X. Hu, Y. Liao, Y. Zhou, W. He, T. Zhou, *J. Alloys Compd.* **2025**, *1010*, 177823.
- [32] S. K. Singh, B. K. Mahanta, P. Rawat, S. Kumar, *J. Alloys Compd.* **2024**, *1007*, 176282.
- [33] S. Zhao, B. Jiang, K. Song, X. Liu, W. Wang, D. Si, J. Zhang, X. Chen, C. Zhou, P. Liu, D. Chen, Z. Zhang, P. Ramasamy, J. Tang, W. Lv, K. G. Prashanth, D. Şopu, J. Eckert, *Mater. Des.* **2024**, *238*, 112634.
- [34] R. Duan, J. Xu, Y. Zhao, Q. Zhou, Z. Yan, Y. Xie, P. Dong, L. Xu, K. Feng, Z. Li, X. Liang, U. Ramamurty, *Addit. Manuf.* **2024**, *86*, 104195.
- [35] A. Oñate, J. P. Sanhueza, D. Zegpi, V. Tuninetti, J. Ramirez, C. Medina, M. Melendrez, D. Rojas, *J. Alloys Compd.* **2023**, *962*, 171224.
- [36] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, *J. Mach. Learn. Res.* **2011**, *12*, 2825.
- [37] D. Dai, T. Xu, X. Wei, G. Ding, Y. Xu, J. Zhang, H. Zhang, *Comput. Mater. Sci.* **2020**, *175*, 109618.
- [38] A. Roy, G. Balasubramanian, *Comput. Mater. Sci.* **2021**, *193*, 110381.
- [39] K. H. Zou, K. Tuncali, S. G. Silverman, *Radiology* **2003**, *227*, 617.
- [40] F. Tian, L. K. Varga, N. Chen, J. Shen, L. Vitos, *Intermetallics* **2015**, *58*, 1.
- [41] Y. Zhang, C. Wen, C. Wang, S. Antonov, D. Xue, Y. Bai, Y. Su, *Acta Mater.* **2020**, *185*, 528.

- [42] K. Bai, C. K. Ng, M. Lin, F. Wei, S. Li, S. L. Teo, D. C. C. Tan, P. Wang, D. Wuu, J. J. Lee, Y.-W. Zhang, *Acta Mater.* **2023**, *243*, 118512.
- [43] X. Duan, T. Han, X. Guan, Y. Wang, H. Su, K. Ming, J. Wang, S. Zheng, *J. Mater. Sci. Technol.* **2023**, *136*, 97.
- [44] P. Shi, Y. Li, Y. Wen, Y. Li, Y. Wang, W. Ren, T. Zheng, Y. Guo, L. Hou, Z. Shen, Y. Jiang, J. Peng, P. Hu, N. Liang, Q. Liu, P.K. Liaw, Y. Zhong, *J. Mater. Sci. Technol.* **2021**, *89*, 88.
- [45] Y. Li, P. Shi, M. Wang, Y. Yang, Y. Wang, Y. Li, Y. Wen, W. Ren, N. Min, Y. Chen, Y. Guo, Z. Shen, T. Zheng, N. Liang, W. Lu, P. K. Liaw, Y. Zhong, Y. Zhu, *Mater. Res. Lett.* **2022**, *10*, 602.
- [46] X. Liu, Y. Wu, L. Gao, Z. Dong, H. Zhu, X. Hui, *J. Alloys Compd.* **2024**, *1004*, 175987.
- [47] A. Shafiei, S. Rajabi, *Appl. Phys. A* **2019**, *125*, 783.