Article

Application of Machine Learning to Predict Grain Boundary Embrittlement in Metals by Combining Bonding-Breaking and Atomic Size Effects

Xuebang Wu [1],* , **Yu-xuan Wang [1,2]**, **Kan-ni He [1,2]**, **Xiangyan Li [1]**, **Wei Liu [1]**, **Yange Zhang [1]**, **Yichun Xu [1]** and **Changsong Liu [1],***

[1] Key Laboratory of Materials Physics, Institute of Solid State Physics, Chinese Academy of Sciences, Hefei 230031, China; wyx811@mail.ustc.edu.cn (Y.-x.W.); kangnh@mail.ustc.edu.cn (K.-n.H.); xiangyanli@issp.ac.cn (X.L.); wliu@issp.ac.cn (W.L.); yangezhang@issp.ac.cn (Y.Z.); xuyichun@issp.ac.cn (Y.X.)

[2] Department of Materials Science and Engineering, University of Science and Technology of China, Hefei 230026, China

* Correspondence: xbwu@issp.ac.cn (X.W.); csliu@issp.ac.cn (C.L.)

Abstract: The strengthening energy or embrittling potency of an alloying element is a fundamental energetics of the grain boundary (GB) embrittlement that control the mechanical properties of metallic materials. A data-driven machine learning approach has recently been used to develop prediction models to uncover the physical mechanisms and design novel materials with enhanced properties. In this work, to accurately predict and uncover the key features in determining the strengthening energies, three machine learning methods were used to model and predict strengthening energies of solutes in different metallic GBs. In addition, 142 strengthening energies from previous density functional theory calculations served as our dataset to train three machine learning models: support vector machine (SVM) with linear kernel, SVM with radial basis function (RBF) kernel, and artificial neural network (ANN). Considering both the bond-breaking effect and atomic size effect, the nonlinear kernel based SVR model was found to perform the best with a correlation of $r^2 \sim 0.889$. The size effect feature shows a significant improvement to prediction performance with respect to using bond-breaking effect only. Moreover, the mean impact value analysis was conducted to quantitatively explore the relative significance of each input feature for improving the effective prediction.

Keywords: grain boundary embrittlement; machine learning; strengthening energy; support vector machine; artificial neural network

1. Introduction

Segregation-induced changes in grain boundary (GB) cohesion are often the controlling factor limiting the mechanical properties of metallic alloys. A small amount of solute atoms may significantly alter fracture toughness and corrosion of metallic alloys by orders of magnitude [1–3]. To evaluate the strengthening or weakening effect of segregants on GB cohesion, one prevalent approach is to calculate the so-called strengthening energy or embrittling potency, ΔE_{SE}, of a particular segregated impurity, which is the segregation energy difference between a GB and a fracture free surface (FS) using the Rice-Wang model [4,5]. The value of ΔE_{SE} plays a key role in the GB embrittlement or strengthening because there is a positive correlation between ΔE_{SE} and the experimental shift of ductile-to-brittle transition temperature (DBTT) that could be used for the design of new alloys [6]. During the last few decades, based on accurate first-principle calculations, an intensive effort has been focused on quantification of the segregation-induced changes of GB cohesion and a large amount of quantitative

computational data have been accumulated in different materials, such as Fe [7–13], Al [5,14,15], Ni [1,8,16–18], W [19], and Mo [20–22].

On the other hand, a few phenomenological models have been developed to understand and predict the solute-induced changes in GB cohesion. An earlier simple bond-breaking model was proposed by Seah to describe ΔE_{SE} of different solutes in Fe GBs [23]. Geng et al. related ΔE_{SE} in Fe and Ni GBs via a modified bond-breaking model with an added elastic mismatch term, and proved the predictions with rigorous first-principles results [8]. A recent study by Gibson and Schuh reviewed the existing data of 400 calculations and found that the simple bond-breaking model is robust to describe the solute-induced changes in GB cohesion [24]. The values of ΔE_{SE} show a strong and positive correlation with the difference in cohesive energies (ΔC), while the contribution to embrittlement from other factors such as atomic size effects and charge transfer are secondary [24]. In addition, Lejcek et al. also reviewed the interfacial segregation and GB embrittlement in Ni and Fe and found that ΔE_{SE} could be determined by the difference of sublimation energies of the solute and solvent (ΔH), which is in accordance with the analytical work of Seah [25,26]. The strengthening energy generally increases with increasing ΔH albeit in individual cases some exceptions can be found [25,26]. Very recently, Gibson and Schuh developed a quantitative model for ΔE_{SE} under conditions of equilibrium segregation and proposed a GB cohesion map to predict whether a given solute–solvent pair will exhibit weakening or strengthening of GBs [27]. Except the similar common feature of ΔC or related quantities with previous modes, the ratio of bonding energies between the solute and solvent, captured by the ratio of their surface energies (RS), was emphasized in their model [27]. Instead of the traditional one-factor bond-breaking model that relates ΔE_{SE} with relative cohesive energy, Tran et al. used a simple two-factor linear model described by the relative metallic radii and the relative difference in cohesive energy, and found it is able to account for most of the variations in the ΔE_{SE} with a value of $r^2 > 0.79$ [28].

The above semi-empirical models and accurate first-principles calculations significantly advance the understanding of solute-induce changes of GB cohesion undoubtedly. However, these methods are limited either in terms of their accuracy or high computational cost. Furthermore, the identification of key features in determining ΔE_{SE} is far from trivial. Thus, the underlying mechanism in ΔE_{SE} of solutes in different metallic GBs is not well understood. As there exists a quite extended amount of the accurate values of ΔE_{SE} in the literature from density functional theory (DFT) calculations, it could enable us to make such a broad, quantitative analysis and prediction of ΔE_{SE} using machine learning methods. The machine learning is one kind of statistical analysis method to capture the complex internal relationships by learning from empirical data [29] and gradually becomes a significant tool in the physics and material research [30]. Some common methods of machine learning include Support Vector Machine (SVM) [31,32], Artificial Neural Network (ANN) [33,34], Random Forest [35], and so on. The machine learning has been widely used to predict fundamental properties such as the formation enthalpy [36], solid solubility [37], solute diffusion [38], and lattice thermal conductivity [39] from DFT calculations and experimental vales [40]. Recently, the structures and energies of clean GBs in Cu and Al by using machine learning methods [41–43]. Zhu et al. reveal new ground states and multiple GBs in Cu by unsupervised machine learning post-processing analysis [41]. Gomberg et al. applied machine learning to connect the GB macro degrees of freedom and energy in asymmetric tilt GBs in Al and the models show a good prediction for GB energies [42]. Tamura et al. proposed a new scheme based on machine learning to predict atomic energies and GB energies in Al symmetric tilt GBs [43]. However, to date, no report has addressed the prediction of GB embrittlement in metals by solute segregation using machine learning methods.

In this work, we apply machine learning to predict ΔE_{SE} of solutes in different metallic GBs including Ni, Fe, Al, W, Mo, W, and Zr, using easily available atomic and elemental properties of the constituting atoms, known as features or descriptors. Three machine learning algorithms are considered, including support vector machine (SVM) with linear kernel, SVM with radial basis function (RBF) kernel, and artificial neural network (ANN). We use standard statistical analysis methods to

determine the factors and predict ΔE_{SE} of solutes in different host metals. Our purpose is to develop fast and accurate models for prediction of ΔE_{SE} and uncover the key features.

2. Methods

The dataset we use for training and testing contains 142 data points by DFT calculations collected from literature [8,19,25] and first principles calculations, see Supplementary Table S1. The set contains 5 hosts which are Ni, Fe, W, Al, and Mo. To build accurate and reliable machine learning models, it is important to include relevant features that collectively capture the trends in the ΔE_{SE} across the different metals. Based on the input parameters of the above models and by taking into account the accessibilities of the parameters, we select the difference of cohesive energies (ΔC) between the host and the segregated solute atoms, the ratio of their surface energies (RS), and the difference of sublimation enthalpies (ΔH) as chemical input parameters. For structural input parameters, the difference of atomic radii (ΔR) between the hosts and solutes is adopted. The values of chemical and structural features are provided in Table S1 of the Supplementary Material. Note that all the chemical and structural parameters can be obtained without performing any first principles calculations. A scatter-plot of the relationship between ΔE_{SE} and ΔC, RS, ΔH, and ΔR is shown in Figure 1. It is clear from the figure that ΔE_{SE} is positively correlated with ΔC, RS, and ΔH, indicating that the bond-breaking effect plays an important role in GB embrittlement. In addition, the atomic size effect also plays a non-negligible role because it is also well correlated with ΔE_{SE}.
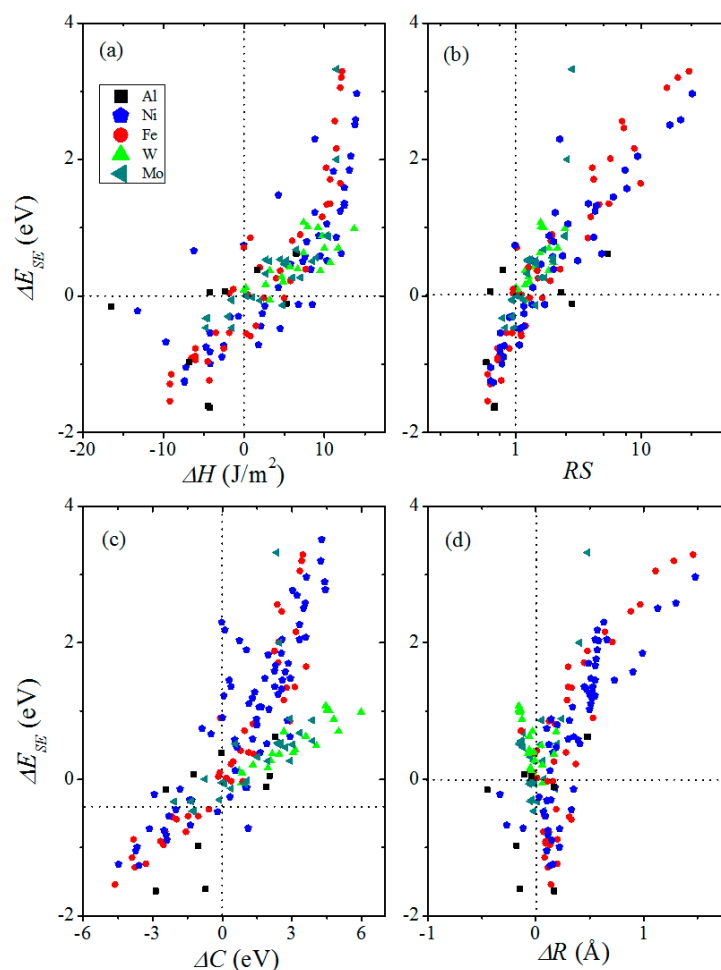
Figure 1. Trend of strengthening energies ΔE_{SE} plotted against (**a**) difference of sublimation enthalpies ΔH, (**b**) ratio of the surface energies RS, (**c**) difference of cohesive energies ΔC, and (**d**) difference of atomic radii ΔR between the host and the segregated solute atoms. Data are from the aggregated data set.

As discussed above, the correlation between bond-breaking effect and size effect in the context of GB embrittlement motivate the use of machine learning models to analyze the extent to which each of these phenomena is correlated with ΔE_{SE}, once the other is taken into account. To asses this question, and to see if other variables are associated with ΔE_{SE}, three machine learning models are constructed including SVM with linear kernel, SVM with RBF kernel, and ANN [31,44]. For SVM, there are many types of kernel function that can affect the performance of predictions and what we use in this work are two common ones, linear and radial basis function (RBF) [31,32]. To realize the SVM modeling, we consider the ε-SVR method using the LIBSVM package in MATLAB (version 9.3.0.713579, R2017b, MathWorks, Inc., Natick, MA, USA) [45]. For ANN, the type of network model we use in this work is the common one based on the back-propagation learning algorithm with Bayesian regularization [46,47], and operated via the toolbox called "nntool" in MATLAB.

The prediction performance of these methods is evaluated by four metrics, which are generally used as the error statistical parameters, mean absolute error (MAE), root mean square error (RMSE), standard deviation of error (SDE), and square correlation coefficient (r^2). MAE is defined as

$$MAE \;=\; \frac{1}{l} \sum_{l}^{i=1} \left| f(x_i) - y_i \right|. \tag{1}$$

RMSE is defined as

$$RMSE = \sqrt{\frac{1}{l} \sum_{i=1}^{l} (f(x_i) - y_i)^2}. \tag{2}$$

SDE is defined as

$$SDE \;=\; \sqrt{\frac{1}{l} \sum_{i=1}^{l} \left(\left| f(x_i) - y_i \right| - MAE \right)^2}. \tag{3}$$

Square correlation coefficient r^2 is defined as

$$r^2 = \left(\frac{\sum_{i=1}^{l} \left(f(x_i) - \overline{f(x)} \right) (y_i - \overline{y})}{\sqrt{\sum_{i=1}^{l} \left(f(x_i) - \overline{f(x)} \right)^2} \sqrt{\sum_{i=1}^{l} (y_i - \overline{y})^2}} \right)^2, \tag{4}$$

where l is data size, y_i is the value of DFT calculations, and $f(x_i)$ is the predicted value of corresponding input. $\overline{f(x)} = \frac{1}{l} \sum_{i=1}^{l} f(x_i)$ (analogously for $\overline{y}$) and r is Pearson correlation coefficient [48].

3. Results and Discussion

Figure 2 shows the fitted results from each of three machine learning methods compared against the DFT calculations using the above four features. These results are from using the entire dataset as the training data for each machine learning method. For the prediction using SVM with linear kernel in Figure 2a, a grid optimization method is used to optimize C, and the values of C range from 2^{-10} to 2^{10} and the step size is 0.5. The SVM model of the best prediction performance was found when C = $2^{3.5}$. The linear regression results are given by $\Delta E_{SE} = -0.238980475 + 0.024468378\,\Delta H + 0.033997136\,RS + 0.223070248\,\Delta C + 1.05643494\,\Delta R$. For the prediction using SVM with RBF kernel in Figure 2b, two parameters, C and γ, can be changed to optimize the model. Similarly, the grid optimization method is adopted to optimize C and γ ranging from 2^{-10} to 2^{10}, respectively. The best parameters are found when C = 2^4 and $\gamma = 2^{0.5}$. For the prediction using ANN in Figure 2c, we use one input layer, one hidden layer including 4 hidden units and the output layer through the pre-test. The results of RMSE and r^2 for three models are summarized in Table 1. It can be seen that the SVM model with RBF kernel and the ANN model show better performances than the SVM model with a linear kernel.
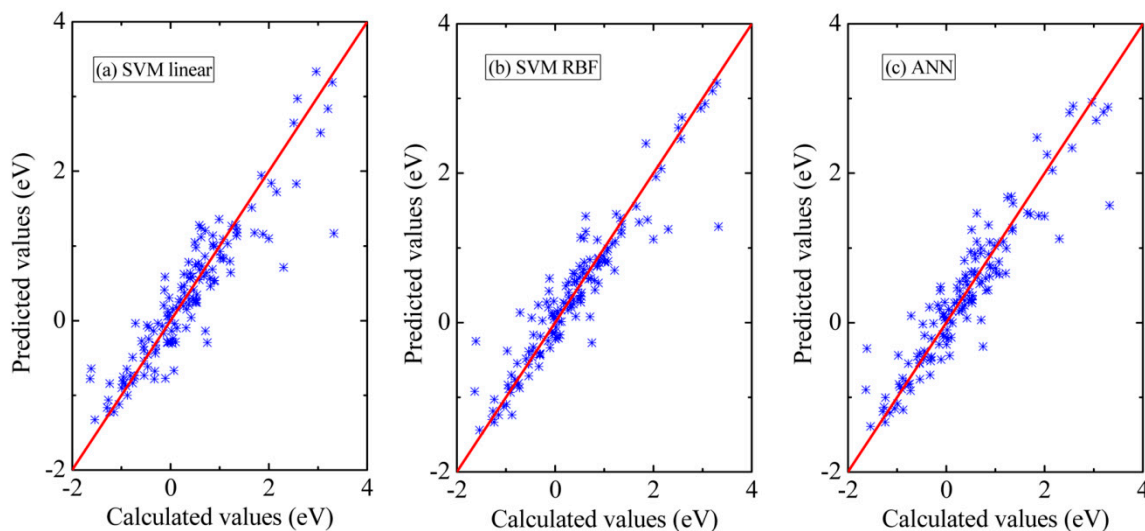
Figure 2. Comparison of ΔE_{SE} from the density functional theory (DFT) calculations and the full-fit results from the three machine learning models with four input features. (**a**) Support vector machine (SVM) model with linear kernel, (**b**) SVM model with radial basis function (RBF) kernel, and (**c**) artificial neural network (ANN).

Table 1. Values of mean absolute error (MAE), root mean square error (RMSE), square correlation coefficient (SDE) and r^2 from full fit and 10-fold cross validation predictions of three machine learning models with four input features. SVM: support vector machine; RBF: radial basis function; ANN: artificial neural network.

Methods	Metrics	SVM with Linear Kernel	SVM with RBF Kernel	ANN
Full fitting	MAE (eV)	0.286	0.233	0.265
	RMSE (eV)	0.406	0.359	0.367
	SDE (eV)	0.288	0.274	0.254
	r^2	0.843	0.876	0.870
10-fold CV	MAE (eV)	0.300	0.280	0.288
	RMSE (eV)	0.424	0.414	0.409
	SDE (eV)	0.300	0.305	0.290
	r^2	0.827	0.835	0.839

Note that a good prediction for the training dataset does not suggest that the model has good prediction ability for the unknown dataset. Thus, to increase reliability of method, the data is randomly divided into 10 groups and during 10 iterations one group is set as the testing dataset, while the remaining nine groups are set as the training dataset. This process is called the 10-fold cross validation that can eliminate the chance of over-fitting [48,49]. For the SVM models with linear kernel and RBF kernel, we follow the similar optimization of model parameters for every group predictions. The values of parameters C and γ are listed in Table 2. For the stochastic nature of ANN model, the prediction procedure of every group is repeated 30 times and an average of 30 times predictions is taken. The prediction results of three models with all four features are shown in Figure 3. A comparison of MAE, RMSE, SDE, and r^2 from the three models is also displayed in Table 1. Generally, the error of cross-validation procedure should always be higher than the error of fitted procedure and this is this case here. It can be seen from Table 1 that, for all three models, the values of MAE, RMSE, SDE, and r^2 are in the range of 0.280–0.300 eV, 0.409–0.424 eV, 0.290–0.300 eV, and 0.827–0.839, respectively. This indicates that the cross-validation errors from three methods using the four features are comparable. Moreover, to increase the accuracy of the models, the 14-fold cross validation tests are also performed, and the

prediction results are shown in the Supplementary Material. Generally, the prediction results are similar to that using 10-fold cross validation, suggesting that the latter is sufficient to ensure accuracy.
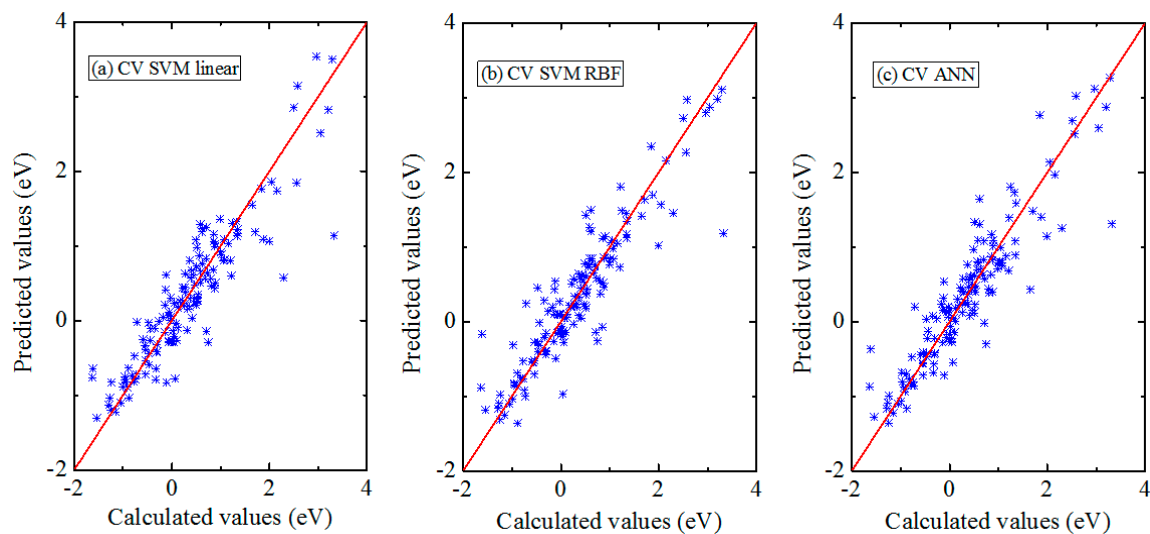

Figure 3. Comparison of ΔE_{SE} from the DFT calculations and the 10-fold cross validation prediction results using (**a**) SVM model with linear kernel, (**b**) SVM model with RBF kernel, and (**c**) ANN.

Table 2. The values of parameter C for SVM model with linear kernel and parameters C and γ for SVM model with an RBF kernel.

Group	SVM with Linear Kernel	SVM with RBF Kernel	
	C	C	γ
G1	11.3137085	11.3137085	22.627417
G2	11.3137085	16	2
G3	8	8	2
G4	64	16	11.3137085
G5	11.3137085	2	4
G6	32	11.3137085	2
G7	362.038672	2.82842712	4
G8	16	11.3137085	2
G9	11.3137085	11.3137085	11.3137085
G10	4	32	8

To analyze the effect of features on the ΔE_{SE}, the SVM models with RBF kernel are used with different combinations of input features. For each set of features, the SVM models are optimized with full fitting. The values of RMSE and r^2 of the SVM models with different input features are displayed in Figure 4. The nonlinear SVM model including ΔH, ΔC, and ΔR is found to be the best description for ΔE_{SE}, with RMSE of 0.339 eV and r^2 of 0.889, while the worst performance is obtained using ΔH and RS, with RMSE of 0.423 eV and r^2 of 0.828, respectively. The correlation among ΔH, RS, and ΔC can be inferred from the fact that the values of RMSE and r^2 do not add to one another when the regressors are combined in the same model. However, the bond breaking effects reflected by ΔH, RS, and ΔC are clearly demonstrated as significant in GB cohesion. It should be noted that, when removing the feature ΔR from the groups of ΔH + RS + ΔC + ΔR, ΔH + ΔC + ΔR, ΔH + RS + ΔR, and RS + ΔC + ΔR, the performance becomes poorer in different degrees. The analysis shows that ΔR is a significant feature to the contribution of ΔE_{SE}. Therefore, the statistical conclusion shows that bond-breaking and atomic size effects are independent and substantial contributors to GB cohesion.
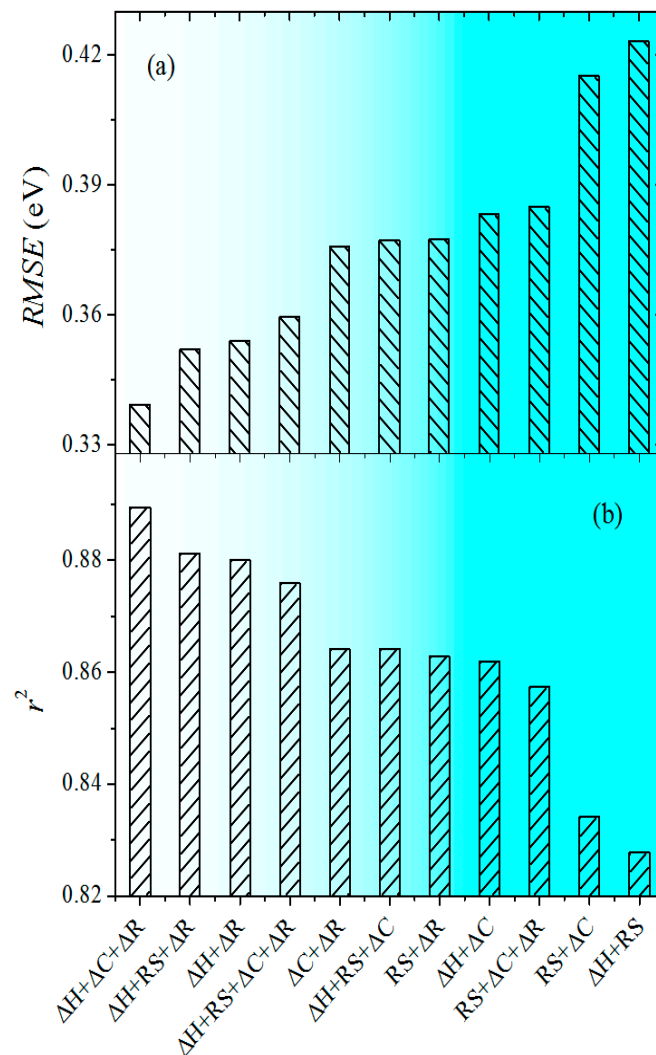
Figure 4. Comparison of values of (**a**) root mean square error (RMSE) and (**b**) r^2 of the SVM models with RBF kernel for different input features.

To quantitatively explore the relative significance of each input variable for improving the prediction performance, the mean impact value (MIV) analysis was conducted using the similar method previously by Jiang et al. [50] and Liu et al. [51]. The MIV values for each input variable on each output variable are calculated and shown in Figure 5. It can be found that the important sequence of the factors for the strengthening energies is $\triangle C > \triangle R > \triangle H > RS$. $\triangle C$ and $\triangle R$ are the two most important factors to influence the strengthening energies. This may explain why the two-factor model is able to account for most of the variation in the strengthening energies in Mo GBs [28]. As shown in Figure 5, the features $\triangle H$, $\triangle C$ and $\triangle R$ show a positive correlation with grain boundary embrittlement, while RS appears a negative correlation. According to the MIV analysis, it can be found that $\triangle H$, $\triangle C$, and $\triangle R$ have a high correlation with grain boundary embrittlement. This is consistent with the above conclusion that the SVM model with these three features yields the best results amongst all the test models.
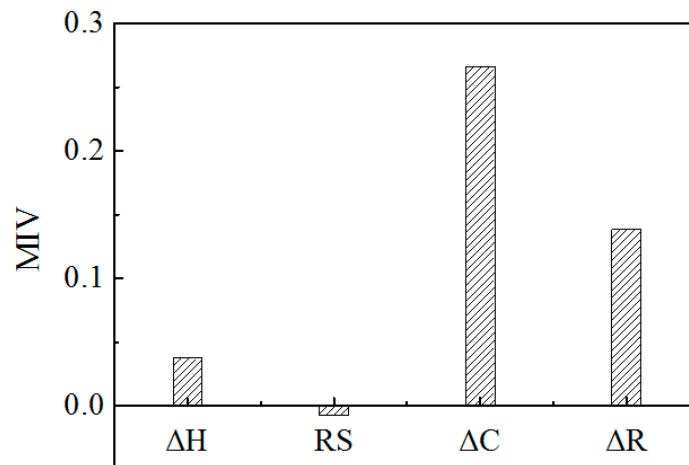
Figure 5. The mean impact values (MIV) for input features ΔH, RS, ΔC, and ΔR.

The aggregation of ΔE_{SE} in different host metals from so many sources should be expected to lead to significant scatter due to the different computations such as potential selection and relaxation standard [52]. Here, the relative good performance of nonlinear SVM model using ΔH, ΔC, and ΔR is thus interpreted to be highly physically meaningful. The present model could not only be used to understand GB strengthening or embrittlement and its underlying physical origins, but also serves as a quantitative prediction of the strengthening energies by solute segregation in other systems that have not yet been studied experimentally or computationally, which is helpful for the rational design and screening of novel materials with desired mechanical properties for specific applications.

4. Conclusions

In this work, we have trained three separate machine learning models to infer the driving forces for GB embrittlement and predict strengthening energies for impurities in different host metals. The readily accessible features (i.e., difference of sublimation enthalpies ΔH, ratio of surface energies RS, difference of cohesive energies ΔC and difference of atomic radii ΔR) are chosen as descriptors. It was shown that the energetics of embrittlement is quantitatively described by two simple effects: bond-breaking and atomic size. A nonlinear kernel based support vector regression model with features ΔH, ΔC, and ΔR shows the best performance prediction for the aggregated set of strengthening energies, with RMSE of 0.339 eV and r^2 of 0.889. The feature of the size effect was found to exhibit considerable importance to the model prediction. Additionally, the MIV based analysis has been carried out for evaluating the features' importance. Results show that the important sequence of the factors for the GB embrittlement is: $\Delta C > \Delta R > \Delta H > RS$. The methods employed in the present work, i.e., clarifying the physical mechanism and extracting the key feature quantities as descriptors by first principles calculations and then predicting material properties via machine learning, can be extended to predict other material properties.

Author Contributions: Conceptualization, X.W.; methodology, Y.-x.W. and K.-n.H.; investigation, X.W., Y.-x.W., K.-n.H. and X.L.; writing-original draft preparation, X.W., Y.-x.W. and K.-n.H.; writing-review and editing, X.L., W.L., Y.Z. and Y.X.; supervision, X.W. and C.L. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Všianská, M.; Šob, M. The effect of segregated sp-impurities on grain-boundary and surface structure, magnetism and embrittlement in nickel. *Prog. Mater. Sci.* **2011**, *56*, 817–840. [CrossRef]
2. Rogers, H. Hydrogen Embrittlement of Metals: Atomic hydrogen from a variety of sources reduces the ductility of many metals. *Science* **1968**, *159*, 1057–1064. [CrossRef]
3. King, A.; Johnson, G.; Engelberg, D.; Ludwig, W.; Marrow, J. Observations of intergranular stress corrosion cracking in a grain-mapped polycrystal. *Science* **2008**, *321*, 382–385. [CrossRef]
4. Rice, J.R.; Wang, J.-S. Embrittlement of interfaces by solute segregation. *Mater. Sci. Eng. A* **1989**, *107*, 23–40. [CrossRef]
5. Lu, G.-H.; Zhang, Y.; Deng, S.; Wang, T.; Kohyama, M.; Yamamoto, R.; Liu, F.; Horikawa, K.; Kanno, M. Origin of intergranular embrittlement of Al alloys induced by Na and Ca segregation: Grain boundary weakening. *Phys. Rev. B* **2006**, *73*. [CrossRef]
6. Yamaguchi, M. First-principles study on the grain boundary embrittlement of metals by solute segregation: Part I. iron (Fe)-solute (B, C, P, and S) systems. *Metall. Mater. Trans. A* **2011**, *42*, 319–329. [CrossRef]
7. Wachowicz, E.; Ossowski, T.; Kiejna, A. Cohesive and magnetic properties of grain boundaries in bcc Fe with Cr additions. *Phys. Rev. B* **2010**, *81*. [CrossRef]
8. Geng, W.; Freeman, A.J.; Olson, G.B. Influence of alloying additions on grain boundary cohesion of transition metals: First-principles determination and its phenomenological extension. *Phys. Rev. B* **2001**, *63*. [CrossRef]
9. Bauer, K.-D.; Todorova, M.; Hingerl, K.; Neugebauer, J. A first principles investigation of zinc induced embrittlement at grain boundaries in bcc iron. *Acta Mater.* **2015**, *90*, 69–76. [CrossRef]
10. Wu, R.; Freeman, A.; Olson, G. First principles determination of the effects of phosphorus and boron on iron grain boundary cohesion. *Science* **1994**, *265*, 376–380. [CrossRef] [PubMed]
11. Kim, M.; Geller, C.B.; Freeman, A. The effect of interstitial N on grain boundary cohesive strength in Fe. *Scr. Mater.* **2004**, *50*, 1341–1343. [CrossRef]
12. Meslin, E.; Fu, C.-C.; Barbu, A.; Gao, F.; Willaime, F. Theoretical study of atomic transport via interstitials in dilute Fe−P alloys. *Phys. Rev. B* **2007**, *75*. [CrossRef]
13. Huang, X.; Janisch, R. Partitioning of Interstitial Segregants during Decohesion: A DFT Case Study of the Σ3 Symmetric Tilt Grain Boundary in Ferritic Steel. *Materials* **2019**, *12*, 2971. [CrossRef]
14. Zhang, S.; Kontsevoi, O.Y.; Freeman, A.J.; Olson, G.B. Sodium-induced embrittlement of an aluminum grain boundary. *Phys. Rev. B* **2010**, *82*. [CrossRef]
15. Zhang, S.; Kontsevoi, O.Y.; Freeman, A.J.; Olson, G.B. First principles investigation of zinc-induced embrittlement in an aluminum grain boundary. *Acta Mater.* **2011**, *59*, 6155–6167. [CrossRef]
16. Yamaguchi, M.; Shiga, M.; Kaburaki, H. Grain boundary decohesion by impurity segregation in a nickel-sulfur system. *Science* **2005**, *307*, 393–397. [CrossRef]
17. Razumovskiy, V.I.; Lozovoi, A.; Razumovskii, I. First-principles-aided design of a new Ni-base superalloy: Influence of transition metal alloying elements on grain boundary and bulk cohesion. *Acta Mater.* **2015**, *82*, 369–377. [CrossRef]
18. Kang, J.; Glatzmaier, G.C.; Wei, S.-H. Origin of the Bismuth-Induced Decohesion of Nickel and Copper Grain Boundaries. *Phys. Rev. Lett.* **2013**, *111*. [CrossRef]
19. Wu, X.; You, Y.-W.; Kong, X.-S.; Chen, J.-L.; Luo, G.-N.; Lu, G.-H.; Liu, C.; Wang, Z. First-principles determination of grain boundary strengthening in tungsten: Dependence on grain boundary structure and metallic radius of solute. *Acta Mater.* **2016**, *120*, 315–326. [CrossRef]
20. Janisch, R.; Elsässer, C. Segregated light elements at grain boundaries in niobium and molybdenum. *Phys. Rev. B* **2003**, *67*. [CrossRef]
21. Tahir, A.; Janisch, R.; Hartmaier, A. Ab initio calculation of traction separation laws for a grain boundary in molybdenum with segregated C impurites. *Model. Simul. Mater. Sci. Eng.* **2013**, *21*. [CrossRef]

22. Kumar, A.; Eyre, B. Grain boundary segregation and intergranular fracture in molybdenum. *Proc. R. Soc. Lond. A Math. Phys. Eng. Sci.* **1980**, 431–458. [CrossRef]
23. Seah, M. Adsorption-induced interface decohesion. *Acta Metall.* **1980**, *28*, 955–962. [CrossRef]
24. Gibson, M.A.; Schuh, C.A. A survey of ab-initio calculations shows that segregation-induced grain boundary embrittlement is predicted by bond-breaking arguments. *Scr. Mater.* **2016**, *113*, 55–58. [CrossRef]
25. Lejček, P.; Šob, M.; Paidar, V. Interfacial segregation and grain boundary embrittlement: An overview and critical assessment of experimental data and calculated results. *Prog. Mater. Sci.* **2017**, *87*, 83–139. [CrossRef]
26. Lejček, P.; Šob, M. An analysis of segregation-induced changes in grain boundary cohesion in bcc iron. *J. Mater. Sci.* **2014**, *49*, 2477–2482. [CrossRef]
27. Gibson, M.A.; Schuh, C.A. Segregation-induced changes in grain boundary cohesion and embrittlement in binary alloys. *Acta Mater.* **2015**, *95*, 145–155. [CrossRef]
28. Tran, R.; Xu, Z.; Zhou, N.; Radhakrishnan, B.; Luo, J.; Ong, S.P. Computational study of metallic dopant segregation and embrittlement at molybdenum grain boundaries. *Acta Mater.* **2016**, *117*, 91–99. [CrossRef]
29. Murphy, K.P. *Machine Learning: A Probabilistic Perspective*; MIT Press: Cambridge, MA, USA, 2012.
30. Council, N.R. *Integrated Computational Materials Engineering: A Transformational Discipline for Improved Competitiveness and National Security*; National Academies Press: Washington, DC, USA, 2008.
31. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin, Germany, 2013.
32. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
33. Bishop, C.M. *Neural Networks for Pattern Recognition*; Oxford University Press: Oxford, UK, 1995.
34. Zurada, J.M. *Introduction to Artificial Neural Systems*; West Publishing Company: St. Paul, MN, USA, 1992; Volume 8.
35. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
36. Ubaru, S.; Miȩdlar, A.; Saad, Y.; Chelikowsky, J.R. Formation enthalpies for transition metal alloys using machine learning. *Phys. Rev. B* **2017**, *95*. [CrossRef]
37. Li, S.; Zhang, H.; Dai, D.; Ding, G.; Wei, X.; Guo, Y. Study on the factors affecting solid solubility in binary alloys: An exploration by Machine Learning. *J. Alloy. Compd.* **2019**, *782*, 110–118. [CrossRef]
38. Wu, H.; Lorenson, A.; Anderson, B.; Witteman, L.; Wu, H.; Meredig, B.; Morgan, D. Robust FCC solute diffusion predictions from ab-initio machine learning methods. *Comput. Mater. Sci.* **2017**, *134*, 160–165. [CrossRef]
39. Chen, L.; Tran, H.; Batra, R.; Kim, C.; Ramprasad, R. Machine Learning Models for the Lattice Thermal Conductivity Prediction of Inorganic Materials. *Comput. Mater. Sci.* **2019**, *170*. [CrossRef]
40. Schleder, G.R.; Padilha, A.C.; Acosta, C.M.; Costa, M.; Fazzio, A. From DFT to machine learning: Recent approaches to materials science–a review. *J. Phys. Mater.* **2019**, *2*. [CrossRef]
41. Zhu, Q.; Samanta, A.; Li, B.; Rudd, R.E.; Frolov, T. Predicting phase behavior of grain boundaries with evolutionary search and machine learning. *Nat. Commun.* **2018**, *9*, 467. [CrossRef]
42. Gomberg, J.A.; Medford, A.J.; Kalidindi, S.R. Extracting knowledge from molecular mechanics simulations of grain boundaries using machine learning. *Acta Mater.* **2017**, *133*, 100–108. [CrossRef]
43. Tamura, T.; Karasuyama, M.; Kobayashi, R.; Arakawa, R.; Shiihara, Y.; Takeuchi, I. Fast and scalable prediction of local energy at grain boundaries: Machine-learning based modeling of first-principles calculations. *Model. Simul. Mater. Sci. Eng.* **2017**, *25*. [CrossRef]
44. HKDH, B. Neural networks in materials science. *ISIJ Int.* **1999**, *39*, 966–979.
45. Chang, C.-C.; Lin, C.-J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 27. [CrossRef]
46. Hecht-Nielsen, R. Theory of the backpropagation neural network. In *Neural Networks for Perception*; Elsevier: Amsterdam, NL, USA, 1992; pp. 65–93.
47. Burden, F.; Winkler, D. Bayesian regularization of neural networks. In *Artificial Neural Networks*; Humana Press: Totowa, NJ, USA, 2008; pp. 23–42.
48. Agrawal, A.; Deshpande, P.D.; Cecen, A.; Basavarsu, G.P.; Choudhary, A.N.; Kalidindi, S.R. Exploration of data science techniques to predict fatigue strength of steel from composition and processing parameters. *Integr. Mater. Manuf. Innov.* **2014**, *3*, 8. [CrossRef]
49. Refaeilzadeh, P.; Tang, L.; Liu, H. Cross-validation. *Encycl. Database Syst.* **2009**, 532–538. [CrossRef]

50. Jiang, X.; Hu, J.; Jia, M.; Zheng, Y. Parameter matching and instantaneous power allocation for the hybrid energy storage system of pure electric vehicles. *Energies* **2018**, *11*, 1933. [CrossRef]

51. Liu, M.; Yu, Z.; Zhang, Y.; Wu, H.; Liao, H.; Deng, S. Prediction and analysis of high velocity oxy fuel (HVOF) sprayed coating using artificial neural network. *Surf. Coat. Technol.* **2019**, *378*. [CrossRef]

52. Gibson, M.A. Segregation and Embrittlement in Metallic Interfaces: Bounds, Models, and Trends. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2016.