Available online at www.keaipublishing.com/en/journals/journal-of-magnesium-and-alloys/

ScienceDirect

Journal of Magnesium and Alloys xxx (xxxx) xxx

www.elsevier.com/locate/jma

Full Length Article

A surface emphasized multi-task learning framework for surface property predictions: A case study of magnesium intermetallics

Gaoning Shi[a,1], Yaowei Wang[b,1], Kun Yang[b], Yuan Qiu[b], Hong Zhu[a,*], Xiaoqin Zeng[c,*]

[a] *University of Michigan - Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, Shanghai 200240, China*
[b] *Engineering Technology Center, Shenyang Aircraft Corp., Shenyang, Liaoning 110850, China*
[c] *State Key Laboratory of Metal Matrix Composites, Shanghai Jiao Tong University, Shanghai 200240, China*

Received 14 September 2024; received in revised form 24 November 2024; accepted 1 December 2024
Available online xxx

Abstract

Surface properties of crystals are critical in many fields, including electrochemistry and photoelectronics, the efficient prediction of which can expedite the design and optimization of catalysts, batteries, alloys etc. However, we are still far from realizing this vision due to the rarity of surface property-related databases, especially for multicomponent compounds, due to the large sample spaces and limited computing resources. In this work, we present a surface emphasized multi-task crystal graph convolutional neural network (SEM-CGCNN) to predict multiple surface properties simultaneously from crystal structures. The model is evaluated on a dataset of 3526 surface energies and work functions of binary magnesium intermetallics obtained through first-principles calculations, and obvious improvements are observed both in efficiency and accuracy over the original CGCNN model. By transferring the pre-trained model to the datasets of pure metals and other intermetallics, the fine-tuned SEM-CGCNN outperforms learning from scratch and can be further applied to other surface properties and materials systems. This study could be a paradigm for the end-to-end mapping of atomic structures to anisotropic surface properties of crystals, which provides an efficient framework to understand and screen materials with desired surface characteristics.

Keywords: Graph neural networks; Multi-task learning; Surface energy; Work function; Intermetallic compounds; Mg alloy.

1. Introduction

Surface properties, including surface energy and work function, are crucial in determining materials' behavior in various applications, particularly in electrochemistry and photoelectronics [1,2]. Surface energy, governing the stability of different crystal facets, is important in understanding the surface structure/reconstruction, the crystal's equilibrium shape, and the adsorption, catalysis, or corrosion reactions [3]. Similarly, the work function, which represents the energy required to remove an electron from the material's surface, is a critical factor for charge injection, electron transfer, and photovoltaic performance [4]. Surface energy and work function, as two fundamental and facet-dependent properties, are essential for understanding phenomena like catalytic activity, surface hydrophobicity, crystal growth, and electronic junction behavior [5–8], which can be effectively obtained from density functional theory (DFT) calculations on surface slabs with varying crystal orientations and terminations. However, in contrast with the rapid increase in available computational data structured in open-source materials databases, surface properties like the work function are extremely rare as each bulk material typically has dozens of distinct low-index crystalline surfaces and terminations [9]. To the best of our knowledge, while the largest computational surface catalysis database, OC20, provides extensive adsorption data, it does not include work function or surface energy information [10]. The Materials Project database contains surface energy and work function data for about 100 elemental crystals [11], while the JARVIS-DFT

* Corresponding authors.
E-mail addresses: hong.zhu@sjtu.edu.cn (H. Zhu), xqzeng@sjtu.edu.cn (X. Zeng).
[1] Co-first authors, these authors contributed equally to this work.

https://doi.org/10.1016/j.jma.2024.12.005

database and the C2DB database only have work functions calculated for 2D materials [12,13].

When it comes to Mg alloys, the lightest structural materials with high specific strengths [14–16] and excellent biocompatibilities [17,18], the surface properties of solid solution and intermetallics are strongly correlated with their corrosion behavior and mechanical performance [19–21]. For corrosion protection, as the alloying effects on corrosion behavior continue to be extensively studied experimentally and theoretically to enhance its intrinsic corrosion resistance [19,22–25], detailed information regarding surface features is critical. In this context, surface energy representing the stability of surfaces is closely related to the corrosion rate by affecting both anodic and cathodic reaction. Specifically, a more comprehensive evaluation of some surface properties (e.g. hydrogen adsorption energy [24], work function [26], vacancy formation energy [27]) can be achieved at nanometric scale based on the Wulff shape constructed from surface energies. Work function describes the degree of difficulty for charge transfer in anodic dissolution [28], where a high work function corresponds to better corrosion resistance by increasing the corrosion potential [29,30]. However, despite the importance of these surface properties, relevant datasets of Mg intermetallics are particularly scarce due to the large sample spaces and limited computing resources, with the largest one contains 150 pieces of the work function of Mg-based solid solutions and Mg-containing intermetallics [31].

Over the past few years, machine learning (ML) has established itself as a potent tool within the realm of materials science [32–36], which is particularly effective in aiding the screening and designing processes within extensive material systems [34–36]. For example, it is important to improve the intrinsic corrosion resistance of Mg alloys by suppressing the dissolution and hydrogen evolution reactions [16,37], and the latest progress has successfully demonstrated that machine learning can serve as an effective tool for screening corrosion-resistant Mg alloys with proper solid solution and intermetallic compound phases [1,24,31]. Furthermore, considering the necessity of featurizing slabs for predicting surface properties, recent studies have focused on predicting surface properties such as electron affinity, ionization energy, surface energy, and work function within the framework of compounds in their respective fields [9,36,38]. Nevertheless, these studies primarily focus on individual surface property predictions, and rarely intersect with Mg intermetallic systems.

Among the various ML techniques, data-driven modeling enables inexpensive and accurate predictions of material properties, facilitating rapid screening of large material spaces to identify potential candidates with desired properties [39–41]. For instance, graph convolutional network (GCN) model [42] automatically selects features, facilitating end-to-end mapping of atomic structures to target attributes. Base on GCN, Xie and Grossman [43] have developed crystal graph convolutional neural network (CGCNN) by treating atoms as graph nodes and chemical bonds as edges. Recently, some variants have been developed by including bond angles [44], crystallographic information [45], global state [46], or bond type [47] in the model to improve predictive accuracy. Besides, multi-task learning (MTL) has been combined with GCN models to strengthen the transfer learning (TL) property along with the ability to inject physics knowledge and reduce the computational effort [48–50]. Nevertheless, to the best of our knowledge, such models haven't been used for surface property predictions.

In this work, we present a surface emphasized multi-task CGCNN, SEM-CGCNN, to predict surface energies and work functions simultaneously and demonstrate its predictive ability for binary intermetallics of magnesium alloys. In this model, GradNorm [51] is used to balance the gradient loss of each task, and an atomic feature for surface/internal atoms is embedded to emphasize the contribution of surface atoms to surface properties based on empirical physical cognition. We train SEM-CGCNN on a DFT dataset of surface energies and work functions for 3526 surfaces of binary Mg intermetallics, and demonstrate its interpretability by extracting the elemental embedding vectors. Finally, the pre-trained model is utilized for parameter initialization in transfer learning to demonstrate its generalization ability in open-source surface property datasets.

2. Computational methods

2.1. Generation of the training dataset

The DFT simulations of crystal unit cell and surface slab models were performed with the projector augmented wave (PAW) [52] method as implemented in the Vienna Ab Initio Simulation Package (VASP) [53]. The exchange-correlation functional of Perdew-Burke-Ernzerhof (PBE) generalized gradient approximation (GGA) was applied [54]. The cut-off energy of plane wave was set at 480 eV, and k-point mesh was automatically generated by pymatgen codes [11]. The convergence criteria of energy and force were set to 10^{-4} eV and 0.02 eV/Å, respectively. All slab models are symmetrical and thicker than 10 Å. The atoms 3Å away from the surface were fixed during the structural optimization, and a vacuum layer of 15 Å was employed along Z direction. The work function was computed as the difference between Fermi energy and the local vacuum level away from the surface. The calculation method of surface energy and work function can be found in the Supplementary Material.

With the high-throughput workflow demonstrated in our previous work [55], 188 binary Mg intermetallics from Materials Project [11], OQMD [56] and AFLOW [57] were screened out as listed in Table S1. All the low-index surfaces (Miller indices up to (111)) with different terminations were considered and 3574 surface energies and work functions were obtained. After filtering out samples with negative surface energy and significant surface reconstructions during relaxation processes, the dataset contains 3526 surface energies and work functions, and their distributions are shown in Fig. 1a and Fig. S1.

Please cite this article as: G. Shi, Y. Wang, K. Yang et al., A surface emphasized multi-task learning framework for surface property predictions: A case study of magnesium intermetallics, Journal of Magnesium and Alloys, https://doi.org/10.1016/j.jma.2024.12.005

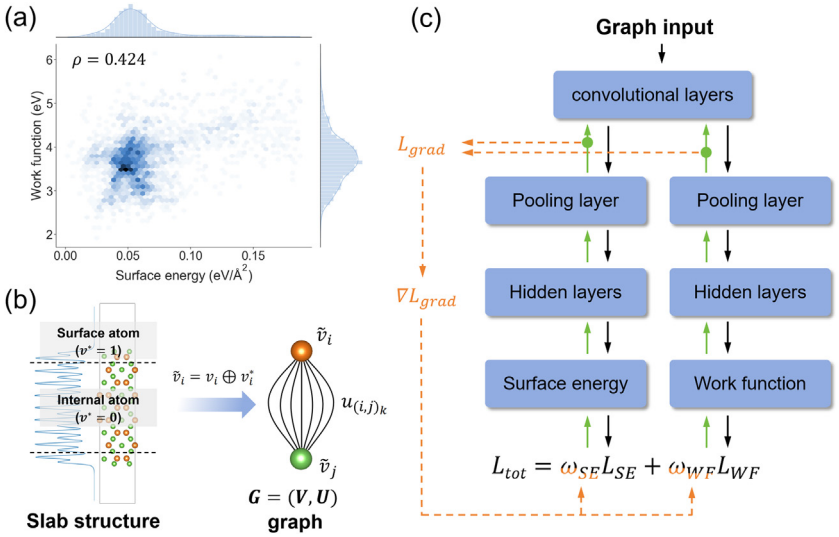


Fig. 1. Architecture of SEM-CGCNN: (a) joint distribution density of DFT calculated surface energy and work function in the dataset of Mg intermetallics; (b) representation of feature embedding for surface/internal atoms, where $\tilde{v}_i$ denotes surface embedded atom feature; (c) SEM-CGCNN network architecture for DFT calculated surface energy and work function. Black arrows represent forward pass, green arrows represent backward pass and orange lines represent balancing ω_{SE} and ω_{WF} with GradNorm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.2. Architecture of SEM-CGCNN

SEM-CGCNN is a CGCNN-based model with multiple heads in fully connected (FC) layers for MTL to achieve better generalization performance, with an extra atom feature embedded for surface/internal atoms to emphasize the contribution of surface atoms. SEM-CGCNN inputs crystal structure and converts it into a graph, and outputs multiple surface properties (surface energy and work function in this work).

2.2.1. Crystal graph convolutional neural network (CGCNN)

As one of the earliest and most established GCN models for predicting crystal properties, CGCNN has been widely validated in its performance and scalability, and hence chosen as our benchmark model. The work by Xie and Grossman [43] developed a generalized crystal graph convolutional network to represent the crystals and to predict their properties with accuracy of *ab initio* physics models. It transforms crystal structure into an undirected multigraph $G = (V, U)$, regarding atoms as graph nodes V and chemical bonds as edges U. In this work, each node i is represented by one feature vector v_i and each edge $(i, j)_k$ is represented by one feature vector $u_{(i,\ j)_k}$ for kth bond connecting atom i and atom j. The original graph is constructed by embedded atom feature v_i along with bond feature $u_{\ (i,\ j)_k}$. In convolutional layers, v_i is iteratively updated by convolving with surrounding bonds $u_{(i,\ j)_k}$ through a nonlinear graph convolutional function, which can be expressed in a simple way:

$$v_i^{(t+1)} = \mathrm{Conv}(v_i^{(t)}, v_j^{(t)}, u_{(i,j)_k}), (i, j)_k \in G, \tag{1}$$

where t denotes the iteration number of the convolutional function. Typically, the form of convolutional function in Eq. (1) has the largest impact on prediction performance with the information of symmetries and invariances captured [58]. To differentiate the interactions between neighbors, Xie and Grossman define $z_{(i,j)k}^{(t)} = v_i^{(t)} \oplus v_j^{(t)} \oplus u_{(i,j)k}$ and perform convolution by:

$$v_i^{(t+1)} = v_i^{(t)} + \sum_{j,k} \sigma\left(z_{(i,j)k}^{(t)} W_f^{(t)} + b_f^{(t)}\right) \odot sp\left(z_{(i,j)k}^{(t)} W_s^{(t)} + b_s^{(t)}\right), \tag{2}$$

where $\oplus$ denotes concatenation and $\odot$ denotes element-wise multiplication. σ denotes a sigmoid function and sp denotes a softplus function. As expected, $\sigma(\cdot)$ functions as factors to differentiate interactions between neighbors while $sp(\cdot)$ functions as local environment of central atoms, and $W_f^{(t)}$, $W_s^{(t)}$, $b_f^{(t)}$, $b_s^{(t)}$ are the corresponding weight matrices and biases. After convolutions, the network learns the feature vector v_i for each atom by iteratively incorporating information from its surrounding environment. These vectors are then average pooled to obtain an overall vector representation v_c for the crystal:

$$v_c = \mathrm{Pool}(v_0, v_1, ..., v_N). \tag{3}$$

The pooling layer ensures permutational invariance with respect to atom indexing and size invariance with respect to the choice of unit cell. v_c is then regarded as an input of FC layer to capture the complex mapping between crystal structure and a target property of p:

$$\hat{y}_p = \mathrm{A_L}(sp(\mathrm{A_{L-1}}(sp(...sp(\mathrm{A_1}(v_c))...)))), \tag{4}$$

where $\mathrm{A_L}$ (with $L = 1, 2, \ldots, \mathrm{L}$) are affine mappings. Using backpropagation and optimizers like stochastic gradient descent (SGD) or adaptive moment estimation (Adam), we can solve the following optimization problem by iteratively

updating the weights W with DFT calculated data:

$$\min_{W} L_p := \frac{1}{n}\sum_{i=1}^{n}\left(y_{p,i} - \hat{y}_{p,i}\right)^2, \tag{5}$$

where L_p is the loss function of mean square error calculated from prediction values $\hat{y}_{p,i}$ and sample values $y_{p,i}$. In our MTL models, L_p will be further used to compose the total loss function later.

2.2.2. Feature embedding for surface atoms

Surface atomic structures are critical in determining the surface properties, which could vary significantly even for the same material with various Miller indices and terminations. We exhibited the surface effect on some slabs from two perspectives in Fig. S2. From deep learning, the self-attention mechanism [36] is applied and shows that the atoms on the surface tend to have higher contribution to the prediction result. From physical characteristics, the charge density is obviously different between surface and interior of slabs, which is a reflection of Friedel oscillations [59]. Therefore, in expectation of highlighting the contribution of surface atoms, the original atom features v_i are concatenated with vectors v_i^*, which labels the outermost atoms on surface, as shown in Fig. 1b, where the operation $\oplus$ denotes concatenation. The modified atom features were further substituted into Eq. (1) after being embedded. The outermost atoms were defined as those within the threshold depth of surface, and that threshold depth was demonstrated as atreated as a hyperparameter. Based on our test results, the model performs best when it takes 2Å as shown in Fig. S3.

2.2.3. Multi-task learning (MTL)

The fundamental motivation for multi-task learning is to achieve better generalization performance [48], and there are two main architectures for MTL in the deep learning context (hard parameter sharing and soft parameter sharing). Here we use hard parameter sharing, which is the most widely used MTL approach and suitable for handling tasks with strong relevance. It has shared convolutional layers across all tasks and task-specific output layers for each task, as shown in Fig. 1c. In MTL, each task has an individual loss function generated from predicted values and the total loss function for the network is the weighted linear sum of individual losses of each task. A common setup for the total loss function of the multi-task problem is [48,50]:

$$\mathcal{L}(t) = \sum_{p\in p} \omega_p(t) L_p(t), \tag{6}$$

where $\mathcal{L}$ is the total loss of the network, L_p are individual losses from each of the task-specific layers and ω_p are the weights for the individual losses. The values of the ω_p in Eq. (5) are determined by GradNorm (details in next subsection) in SEM-CGCNN. The multiple target quantities in MTL are interpreted as mutual inductive biases [50] because the error of a single quantity acts as a regularizer with respect to the loss functions of other quantities.

2.2.4. Gradient normalization

For a multitask loss function Eq. (2), the aims of GradNorm algorithm are two-fold [51]: (1) to place gradient norms for different tasks on a common scale, enabling us to reason about their relative magnitudes, and (2) to dynamically adjust gradient norms to ensure that different tasks train at similar rates. Specifically,

$$L_{grad}(t; w_i(t)) = \sum_i \left| G_W^{(i)}(t) - \overline{G}_W(t) \times [r_i(t)]^{\alpha} \right|_1, \tag{7}$$

where $G_W^{(i)}(t)$ denotes the L_2 norm of the gradient of the last shared layer with respect to the weights W, $\bar{G}_W(t)$ is the average gradient norm across all tasks at training step t, $r_i(t)$ is the relative inverse learning rate of task i, and α is a hyperparameter which determines the strength of balancing tasks (a high α value penalizes tasks with rapid loss reduction, instead favoring those with slower loss decreases). GradNorm L_{grad} presents the gap between actual and target gradient norms on all tasks, and it can be differentiated with respect to w_i, which gives a weight of corresponding task directly. Then the computed $\nabla_{w_i} L_{grad}$ is applied to updated each w_i iteratively (Fig. 1c).

2.3. Model construction and training

Pytorch [60] is utilized for implementation of our model, serving as a robust neural network library and a performance portability layer for running on multiple hardware architectures with various optimization algorithms. Thus SEM-CGCNN can be run on CPUs and GPUs, from laptops to supercomputers. In this work, the crystal graphs are constructed by searching the nearest 12 neighbors within a cutoff radius of 12 Å, and Adam optimizer is utilized to update parameters in the training process with an initial learning rate of 0.001, which is reduced by 90 % after 100 epochs for tighter convergence. The embedding dimension is set to 64, and the number of both convolutional layers and hidden layers are set to 3 by default. To choose the best model, we apply a train-validation scheme to optimize the prediction. All datasets containing unrelaxed structures and their corresponding properties are randomly divided into 60 % training set, 20 % validation set, and 20 % testing set, and the model performs best on the validation set is selected for final testing, which reduce the likelihood of overfitting.

The optimal hyperparameters are obtained through the Tree-structured Parzen Estimator (TPE)-based Bayesian optimization method [61] in a train-validation process, implemented with the open-source Python library hyperopt [62]. This approach offers several advantages, particularly in efficiently searching high-dimensional and complex hyperparameter spaces. We set a maximum of 200 trails (redundant searches are avoided) for each model to balance computational resources with the need for thorough optimization. The search space used for the optimization process is detailed in Table S2. With the optimal hyperparameters, a 5-fold cross validation is also conducted to further validate the stability and generalization capability of the models.

Please cite this article as: G. Shi, Y. Wang, K. Yang et al., A surface emphasized multi-task learning framework for surface property predictions: A case study of magnesium intermetallics, Journal of Magnesium and Alloys, https://doi.org/10.1016/j.jma.2024.12.005

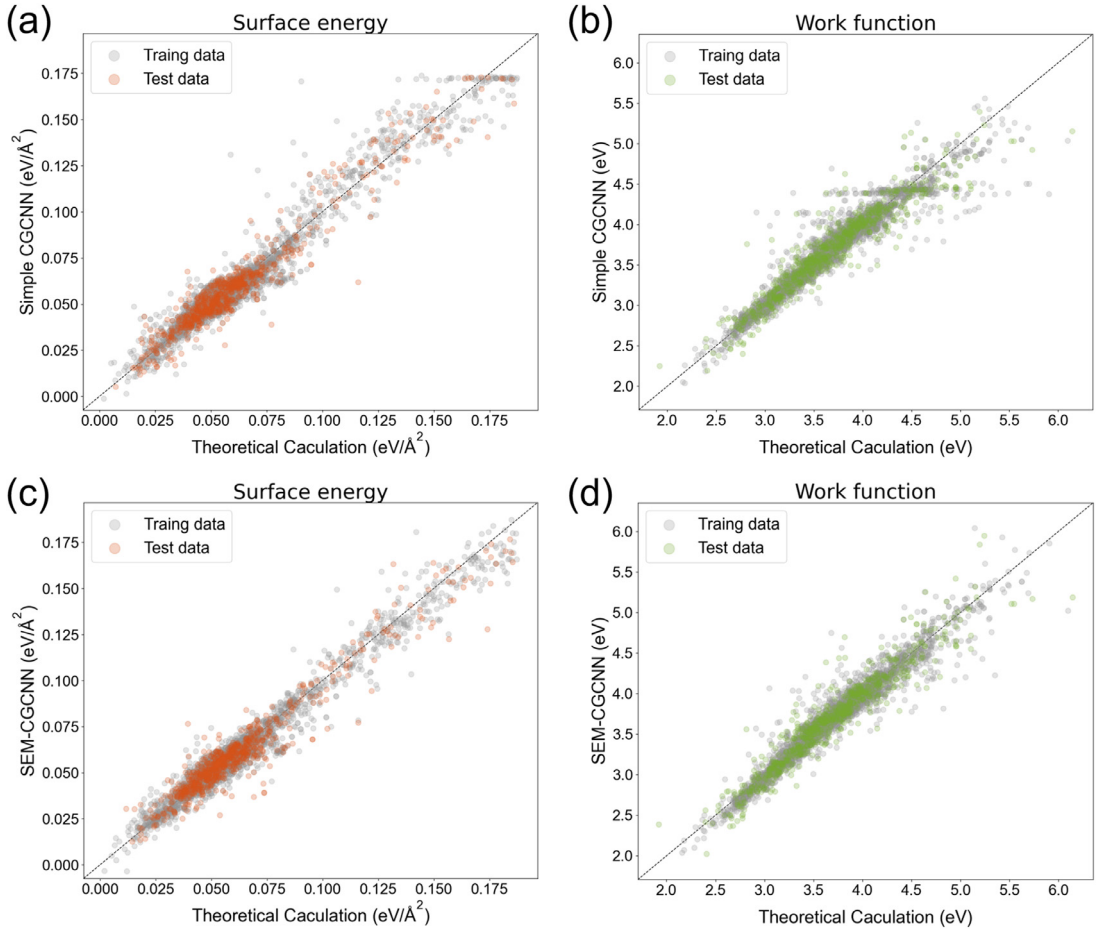


Fig. 2. 2D histogram representing the predicted surface energy and work function against DFT calculations for facets of binary Mg intermetallics, using the model of (a)(b) simple CGCNN and (c)(d) SEM-CGCNN. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The pre-trained model on the dataset of binary Mg intermetallics is further used in transfer learning to improve the performance of models on the datasets of pure metals in the Materials Virtual Lab [63] (containing surface energy and work function) and intermetallics by Palizhati et al. [38] (surface energy only). We implement scratch (SC) models and transfer learning (TL) models respectively. For SC models, the models are trained directly on the target dataset from scratch without imparting any prior knowledge from the source data to the model. For TL models, we perform fine-tuning which uses the weights from the pre-trained model as the initial weights for the network (which shares the same architecture as that used during the training of the pre-trained model) and these initial weights are then refined using a smaller dataset.

3. Results and discussion

3.1. Model accuracy and reliability

In Table 1, the performance of CGCNN based single-task learning (STL) model and modified multi-task learning (MTL) models are compared in predicting surface energy and work function of binary Mg intermetallics, individually or simultaneously. The hyperparameters are shared in these models with default values listed in Table S3. We compute not only the accuracy (MAE and R^2) of each model, but also the total loss optimized (for STL, the individual losses of different tasks are added up as a total loss). This simple metric gives a comprehensive quantification for different models. GradNorm algorithm and surface embedding are evaluated to be helpful, and the model of SEM-CGCNN combining them together performs the best. The results show that the multi-headed SEM-CGCNN model effectively leverages the strong correlations among physical properties to mitigate prediction uncertainties, and the computational time is comparable to that of a single-headed CGCNN model. This trend is basically the same after conducting a Bayesian optimalization on hyperparameters and the following 5-fold cross validation, as shown in Table S3 and S4. Notably, for work functions, which have higher prediction errors compared to surface energies, the MTL models significantly reduce the prediction errors compared to the STL model.

The parity plots for the predictions by simple CGCNN against DFT calculated surface energy and work function for

Please cite this article as: G. Shi, Y. Wang, K. Yang et al., A surface emphasized multi-task learning framework for surface property predictions: A case study of magnesium intermetallics, Journal of Magnesium and Alloys, https://doi.org/10.1016/j.jma.2024.12.005

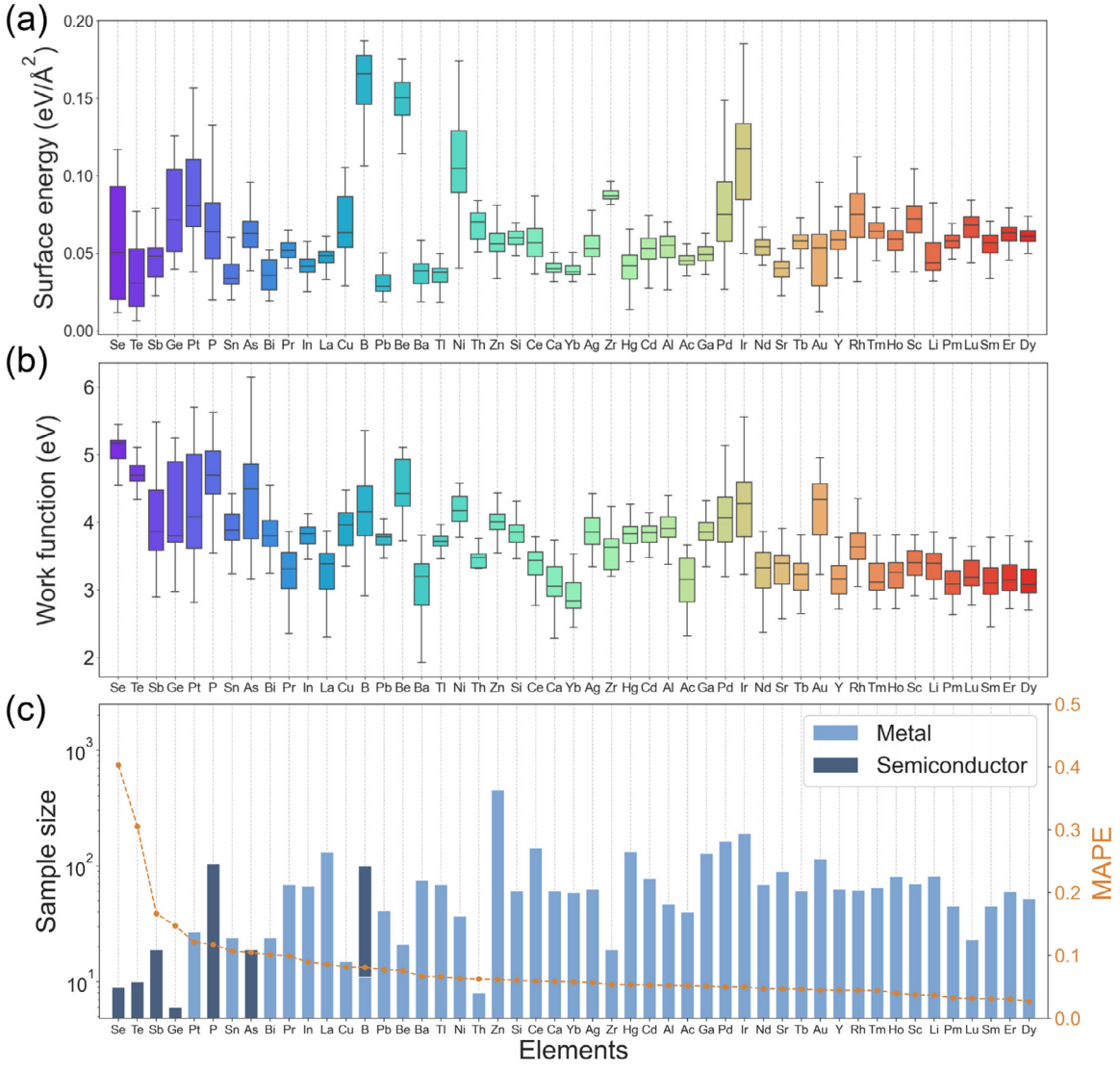


Fig. 3. Distribution analysis of samples and prediction errors for chemical elements of binary Mg intermetallics: (a)(b) box plots of the surface energy and work function for various elements, with outliers hidden for clarity; (c) distribution of samples sizes and prediction errors for various elements. The mean absolute percent error (MAPE) is used to evaluate prediction error for each chemical element, and it's averaged from surface energy and work function. The blue and dark blue bars represent metals and semiconductors, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
The predictive accuracy of CGCNN based models on test set.

Models	Model description	Surface energy (eV/Å^2)		Work function (eV)		Total loss
		MAE	R^2	MAE	R^2	
STL	Surface energy	0.0055	0.9056	–	–	0.1727
	Work function	–	–	0.1253	0.8578	
MTL	Simple	0.0059	0.8807	0.1261	0.8843	0.1836
	GradNorm	**0.0053**	0.9081	0.1288	0.8802	0.1708
	Surface embedded	0.0056	0.8986	0.1211	0.8843	0.1776
	SEM-CGCNN	**0.0053**	**0.9348**	**0.1180**	**0.9072**	**0.1570**

STL, single-task learning; MTL, multi-task learning.

Please cite this article as: G. Shi, Y. Wang, K. Yang et al., A surface emphasized multi-task learning framework for surface property predictions: A case study of magnesium intermetallics, Journal of Magnesium and Alloys, https://doi.org/10.1016/j.jma.2024.12.005

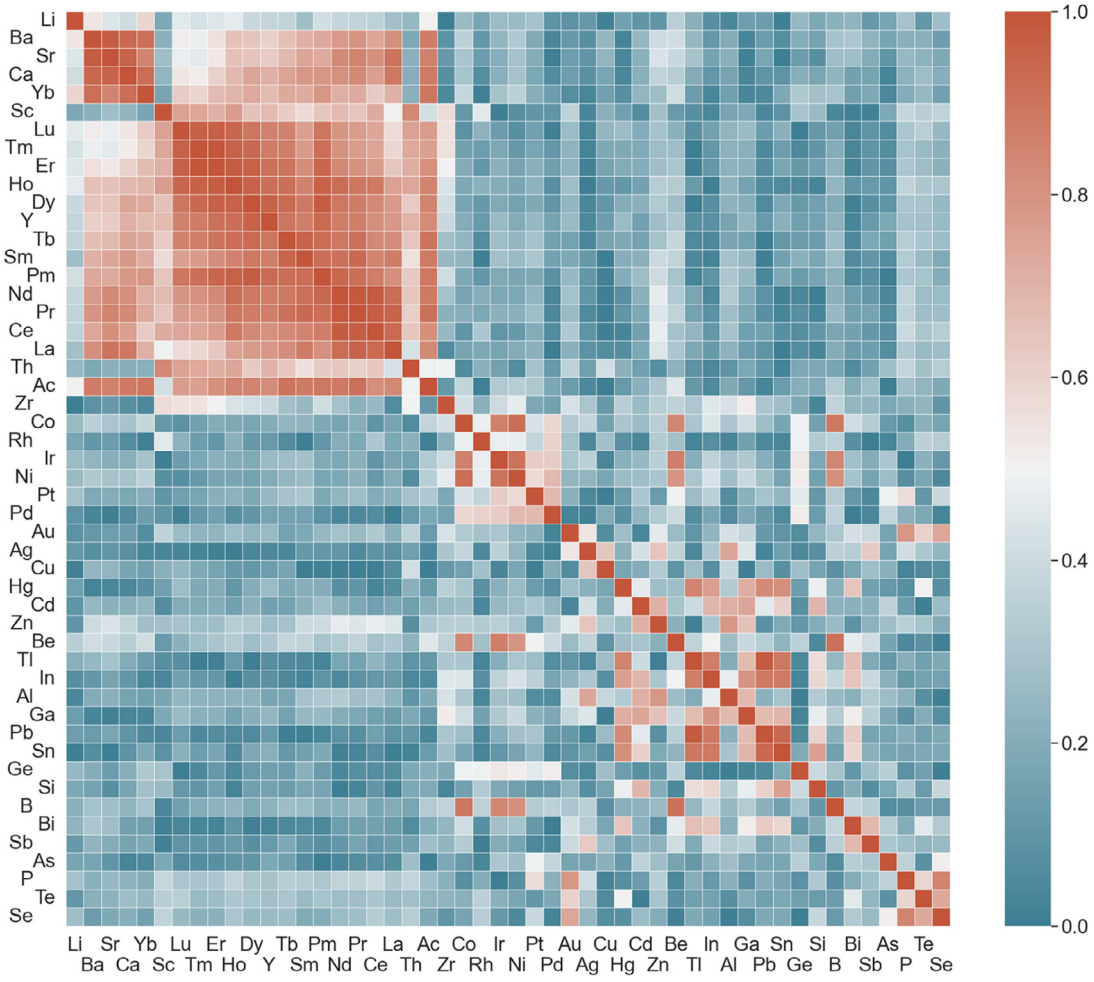


Fig. 4. Pearson correlations between elemental embedding vectors of 50 elements in 188 binary Mg intermetallics, arranged in order of increasing Mendeleev number [65] for easier visualization of trends.

training data (gray dots) and test data (orange and green dots) are shown in Fig. 2a,b respectively. The majority of data points cluster near the diagonal lines, which indicates that the surface energy and work function predicted by CGCNN considerably align with the corresponding theoretical values from first-principles calculations. Nevertheless, it is notable that some points exhibit horizontal distribution, which are corresponding to different facets of a single compound (primarily MgB_7, MgB_{13}, $MgAu_3$). This observation suggests that the original CGCNN model fails to effectively differentiate facets with varying orientations. However, this limitation is apparently mitigated in the SEM-CGCNN model, which leverage the strong correlations among physical properties to reduce predictions uncertainties, as evidenced by the parity plots depicted in Fig. 2c, d. From the perspective of data distribution, the data points for surface energy are densely clustered near zero, contrasting with the more uniformly distributed work function data. And as anticipated, this non-uniform concentration yields more accurate predictions in regions abundant with data, resulting in higher accuracy for surface energy predictions. However, the lack of samples exhibiting high surface energy may potentially hinder the model's generalization ability.

To better analyze the model performance, and make clear the limitations of our model, the distribution of all binary Mg intermetallics in our dataset and prediction errors for chemical elements involved are shown in Fig. 3. Among all elements evaluated, the predictive errors of their respective Mg intermetallics indicate that for 46 elements, the prediction errors remain below 15 %, with 40 elements exhibiting prediction errors below 10 %. It's noteworthy that the elements with prediction errors higher than 10 % are in or near the region of metalloids (excepting Pt) with weaker metallicity, and it can be seen more clearly in Fig. S4. The exception of Pt could be attributed to its wider sample distribution range with smaller sample size. The metallicity for each chemical element is also shown in Fig. 3. It's evident that the Mg compounds associated with Se, Te, and Sb (namely $MgSe_2$, $MgTe_2$, Mg_3Sb_2), which exhibit the highest prediction errors, are semiconductors. Generally, those elements with high prediction errors usually have smaller sample sizes and wider sample distribution range, and most of them are semicon-

Please cite this article as: G. Shi, Y. Wang, K. Yang et al., A surface emphasized multi-task learning framework for surface property predictions: A case study of magnesium intermetallics, Journal of Magnesium and Alloys, https://doi.org/10.1016/j.jma.2024.12.005

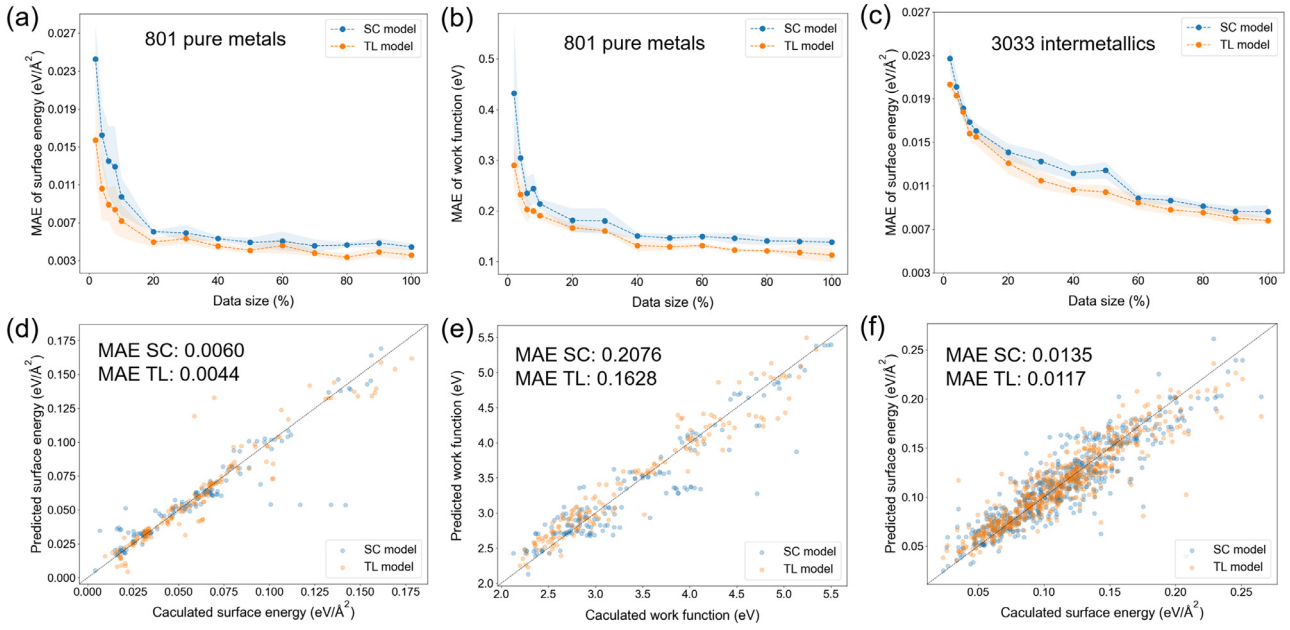


Fig. 5. Transfer learning performance of SEM-CGCNN: (a)(b)(c) learning curve for predicting surface energy (work function) in datasets of 801 pure metals by Materials Virtual Lab[63] and 3033 intermetallics by Palizhati et al. [38] with different training data sizes on a fixed test set; (d)(e)(f) prediction error analysis of scratch (SC) and transfer learning (TL) model in two datasets, each with a data size of only 20 % allocated for training. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ducting intermetallics, which are minority among all samples and significantly deviate from metals in terms of their coordination numbers and binding energy. This result aligns with previous work by Palizhati et al. [38], who demonstrated that CGCNN may perform inadequately for covalent materials.

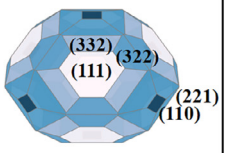
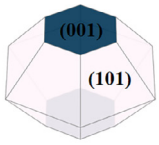
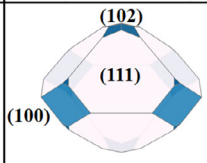
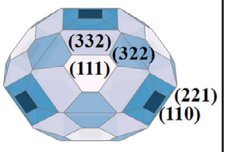
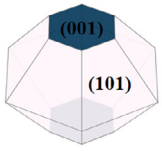
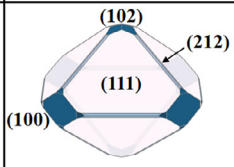
3.2. Interpretability and transfer learning

Compared to the prevailing approach within the materials ML community of constructing single-purpose models for each quantity, the advantage of deep learning models lies in their typically superior generalization capability. This stems from their ability to automatically extract features from raw data, thereby acquiring more comprehensive information and higher-level feature representations, provided that the data volume requirements are met. One of the primary objectives of MTL is also to enhance feature interpretability and generalization capability, which becomes particularly crucial in the context of relatively small datasets available in materials science. In applications of chemistry and materials science, a desirable feature should possess interpretability of known chemical intuition [46,64]. To this end, we extracted elemental embedding vectors of element X in Mg-X intermetallics from the pre-trained model and demonstrated their Pearson correlations, as shown in Fig. 4. In fact, after convolutional layers, the node features which are atom features of X formerly, have not only contained information about the X atoms, but also information about the Mg atoms within the receptive field. Obviously, this would diminish the discriminability of node features for different elements. Nevertheless, the correlations between the elemental embeddings still correctly reproduce the trends in the periodic table of elements, and it also reflects the phenomenon of weight condensation throughout the convolutional layers especially for rare earth elements (Fig. S5), which is a positive indication of the captured information on symmetry. It is noteworthy that the extracted trends faithfully replicate well-known “exceptions” in the periodic arrangement of atoms. The observation that Yb does not conform to the lanthanoids but rather aligns closer with alkaline earth elements is consistent with chemical intuition and matches well with the structure maps by Pettifor [65].

To check the robustness of the proposed framework, we built a pre-trained model using SEM-CGCNN on the dataset of binary Mg intermetallics, by using both surface energy and work function as the source materials property. This pre-trained model was then used to perform transfer learning on two datasets [63,38]. The first one contains 801 facets of 75 pure metals with the surface energy and work function, while the second one contains 3033 facets of 367 intermetallics (19 unitary, 260 binary and 88 ternary) with the surface energy. We performed a training size-based analysis to examine the performance difference between scratch (SC) and transfer learning (TL) models as shown in Fig. 5. The validation and test set are both fixed to 20 % of the total data size, with the remaining data proportionally allocated for training. Fig. 5a-c show that TL models always outperform SC models for all the training data sizes in surface energy (work function) prediction. The convergence of the model’s performance is significantly better in the dataset of pure metals, and it’s evident that there is a lack of samples for the model to featurize the large structural differences in the dataset of intermetallics. Fig. 5d-f compare the prediction results of two models when

Please cite this article as: G. Shi, Y. Wang, K. Yang et al., A surface emphasized multi-task learning framework for surface property predictions: A case study of magnesium intermetallics, Journal of Magnesium and Alloys, https://doi.org/10.1016/j.jma.2024.12.005

(a)

Crystal	Pt (mp-126)	$MgCo_2$ (mp-864931)	$CuFePt_2$ (mp-3702)
DFT			
SEM-CGCNN			
$\bar{\gamma}$ (eV/Å²)	0.1015 (0.0977)	0.0963 (0.0937)	0.1203 (0.1213)
$\bar{\Phi}$ (eV)	5.435 (5.516)	3.809 (3.783)	-

(b)

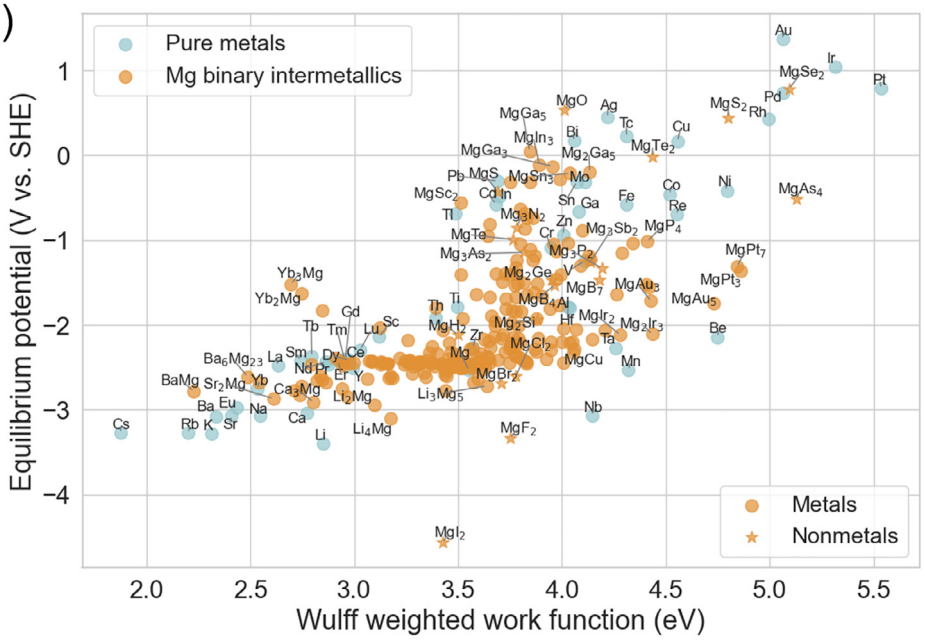


Fig. 6. Prediction results of Wulff shape weighted surface properties: (a) comparison of Wulff structures constructed using DFT calculated and SEM-CGCNN predicted surface energy. The source data of Pt, $MgCo_2$ and $CuFePt_2$ are from Materials Virtual Lab [63], this work and Palizhati et al. [38], with facets up to a maximum absolute Miller index of 3, 1, 2 respectively. The facets with area ratio larger than 1 % are labeled, and darker shades correspond to relatively higher surface energy. $\bar{\gamma}$ and $\bar{\Phi}$ are weighted surface energy and work function based on Wulff shape respectively. The values above are predicted, while those within parentheses below are calculated via DFT; (b) scatter plot showing the relationship between dissolution equilibrium potential and Wulff weighted work function for pure metals and Mg-based binary intermetallics. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

trained with a data size of 20 % (equivalent to 12 % of the whole dataset), providing snapshots under small dataset conditions. Additionally, for the dataset of intermetallics, the best MAE of 0.0073 eV/Å from TL model is comparable with the original work [38], despite using fewer than half of its parameters in our work.

Further we evaluated the prediction uncertainty of Wulff construction [66] using the leave-one-out method, which was designed to predict new intermetallic combinations or new pure metals not seen both in training set and validation set. As shown in Fig. 6a, the Wulff structures were constructed for Pt (mp-126), $MgCo_2$ (mp-864931), and $CuFePt_2$ (mp-3702) by DFT calculated and SEM-CGCNN predicted surface energy from three datasets. The predicted results closely resemble the computationally obtained Wulff structures. The Wulff weighted surface energies and work functions of each crystal are also listed for SEM-CGCNN's predictions and DFT results. With the facet-dependent error analysis listed in Table

S5, S6 and S7, we noticed that the Wulff weighted surface energies and work functions exhibit lower errors than the MAE of surface energies and work functions across all surfaces, which can be interpreted as lower-index surfaces occupying a larger area ratio in Wulff structures while also exhibiting relatively smaller prediction errors. Under this premise, it is believed that the model can provide reasonably accurate predictions for these two properties in practical applications, given that the standard errors for DFT estimated facet-dependent surface energy and work function for elemental crystals have been reported as up to 0.0168 eV/Å [63] and 0.246 eV [26], respectively.

Furthermore, we predict the Wulff shape weighted work function of 246 Mg-based binary intermetallics using the aforementioned method and compare them with those of 56 pure metals, as shown in Fig. 6b. It is important to note that while the Wulff shape from surface energies provides useful estimates of surface properties at the nanometric scale, its application to macroscopic properties should be done with caution (the morphology of crystal is also affected by surface defects/adsorptions, growth condictions and interface energy) [66]. For polycrystalline materials, however, such work function predictions have been shown to be relatively reliable [26]. The dissolution equilibrium potentials are also calculated to evaluate their stability in neutral solutions, with the concentration of dissolved ions set to 10^{-6} mol/L (details can be found in the Supplementary Material). From the perspective of corrosion protection, second phases with equilibrium potential slightly higher than that of Mg are thermodynamically desirable [67], and semiconductors like Mg_2Ge and Mg_3As_2 are known as cathodic "poisons" which can form a Schottky barrier with magnesium matrix [68,69]. From a photocatalytic perspective, certain Mg-based intermetallic semiconductors with both high equilibrium potentials and work functions (e.g., MgS_2, $MgAs_4$, and $MgSe_2$) are potential candidates for constructing heterojunctions with noble metals [70]. Additionally, intermetallics with low work functions (e.g., BaMg, Ba_6Mg_{23}, Yb_2Mg, Yb_3Mg) have potential applications in electron emission devices [71], high-brightness photocathodes [72], and thermionic energy converters [73].

Though SEM-CGCNN has offered a valid implementation for our multi-task learning framework in surface property predictions and can be readily adapted to open-source surface property database to achieve a fine-tuned model for specific materials systems, several challenges still remain unresolved in this study, ranging from DFT calculation to model prediction, highlighting the need for further investigation and innovation. Firstly, the DFT calculated surface energies have inherent errors brought by surface reconstruction, finite size effect and various approximations. Secondly, since the predictions are based on the unrelaxed slabs, the effect of relaxation on surface properties might be underestimated. Recent works [9] attempted to calculate and predict surface energy and work function under non-relaxation conditions, but primarily for rough screening purposes. A modified bond-breaking model [74] shows promise in mitigating this issue by refining unrelaxed surface structures based on Hook's law, leading to more accurate representations of the atomic arrangements. Furthermore, current GCN models still face the challenges in adequately capturing underlying patterns and key influential factors in graph data with complex structures and features, and their generalization ability remains limited in diverse sample space. To further improve the predictive performance, future work could focus on better addressing the influence of surface relaxation and integrating additional factors. Furthermore, the flexibility of our MTL framework allows for alignment with evolving state-of-the-art GCN models, ensuring continued accuracy and efficiency.

4. Conclusions

In this study, we introduce a multi-task learning framework for predicting surface properties from crystal structure graphs, trained on 3526 binary Mg intermetallic surface structures especially for surface energies and work functions. This model demonstrates advantages in efficiency, flexibility, and applicability, making it a valuable tool for surface-related research. The main conclusions are as follows:

(1) The SEM-CGCNN model, with its flexibility in multi-task learning, demonstrates higher efficiency than single-task CGCNN models and conventional electronic structure calculations, while effectively leveraging correlations among surface properties to reduce prediction uncertainties.
(2) The model has demonstrated its generalization ability in predicting facet-dependent surface energies and work functions for out-of-sample crystals, underscoring its potential for applications in surface property analysis.
(3) The framework provides a powerful tool for estimating surface properties as a surrogate for theoretical and experimental studies in electrochemistry, photoelectronics, and interface engineering.

Data availability

The source code and dataset used in our work are freely accessible at:
[https://gitlab.com/iamds/SEM-CGCNN]

Declaration of competing interest

X.Q. Zeng is an associate editor for Journal of Magnesium and Alloys and was not involved in the editorial review or the decision to publish this article. All authors declare that there are no competing interests.

CRediT authorship contribution statement

Gaoning Shi: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yaowei Wang:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Kun Yang:** Supervision. **Yuan Qiu:** Supervision. **Hong Zhu:** Writing – review & editing, Supervision, Resources. **Xiaoqin Zeng:** Supervision, Resources.

Please cite this article as: G. Shi, Y. Wang, K. Yang et al., A surface emphasized multi-task learning framework for surface property predictions: A case study of magnesium intermetallics, Journal of Magnesium and Alloys, https://doi.org/10.1016/j.jma.2024.12.005

Acknowledgments

The research was financially supported by the National Key R&D Program (No. 2021YFB3501002) supported by the Ministry of Science and Technology of China, National Natural Science Foundation of China (No. 51825101, 52127801). Computations were carried out on the π 2.0 cluster supported by the Center for High Performance Computing at Shanghai Jiao Tong University.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jma.2024.12.005.

References

[1] J.A. Yuwono, X. Li, T.D. Doležal, A.J. Samin, J.Q. Shi, Z. Li, N. Birbilis, Npj Mater. Degrad. 7 (2023) 87, doi:10.1038/s41529-023-00409-7.
[2] Y. Wang, C. Xu, Y. Zhou, J. Lee, Q. Chen, H. Chen, Small 20 (2024) 2308897, doi:10.1002/smll.202308897.
[3] D. Li, Q. Chen, J. Chun, K. Fichthorn, J. De Yoreo, H. Zheng, Chem. Rev 123 (2023) 3127–3159, doi:10.1021/acs.chemrev.2c00700.
[4] C. Zhao, C.G. Tang, Z.-L. Seah, Q.-M. Koh, L.-L. Chua, R.-Q. Png, P.K.H. Ho, Nat. Commun. 12 (2021) 2250, doi:10.1038/s41467-021-22358-y.
[5] N.-C. Lai, G. Cong, Z. Liang, Y.-C. Lu, Joule 2 (2018) 1511–1521, doi:10.1016/j.joule.2018.04.009.
[6] A. Valsesia, C. Desmet, I. Ojea-Jiménez, A. Oddo, R. Capomaccio, F. Rossi, P. Colpo, Commun. Chem. 1 (2018) 1–11, doi:10.1038/s42004-018-0054-7.
[7] M. Wu, Z. Zhang, X. Xu, Z. Zhang, Y. Duan, J. Dong, R. Qiao, S. You, L. Wang, J. Qi, D. Zou, N. Shang, Y. Yang, H. Li, L. Zhu, J. Sun, H. Yu, P. Gao, X. Bai, Y. Jiang, Z.-J. Wang, F. Ding, D. Yu, E. Wang, K. Liu, Nature 581 (2020) 406–410, doi:10.1038/s41586-020-2298-5.
[8] Y. Liu, J. Guo, E. Zhu, L. Liao, S.-J. Lee, M. Ding, I. Shakir, V. Gambin, Y. Huang, X. Duan, Nature 557 (2018) 696–700, doi:10.1038/s41586-018-0129-8.
[9] P. Schindler, E.R. Antoniuk, G. Cheon, Y. Zhu, E.J. Reed, Adv. Funct. Mater. (2024) 2401764, doi:10.1002/adfm.202401764.
[10] L. Chanussot, A. Das, S. Goyal, T. Lavril, M. Shuaibi, M. Riviere, K. Tran, J. Heras-Domingo, C. Ho, W. Hu, A. Palizhati, A. Sriram, B. Wood, J. Yoon, D. Parikh, C.L. Zitnick, Z. Ulissi, ACS Catal 11 (2021) 6059–6072, doi:10.1021/acscatal.0c04525.
[11] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K.A. Persson, APL Mater 1 (2013) 011002, doi:10.1063/1.4812323.
[12] K. Choudhary, K.F. Garrity, A.C.E. Reid, B. DeCost, A.J. Biacchi, A.R. Hight Walker, Z. Trautt, J. Hattrick-Simpers, A.G. Kusne, A. Centrone, A. Davydov, J. Jiang, R. Pachter, G. Cheon, E. Reed, A. Agrawal, X. Qian, V. Sharma, H. Zhuang, S.V. Kalinin, B.G. Sumpter, G. Pilania, P. Acar, S. Mandal, K. Haule, D. Vanderbilt, K. Rabe, F. Tavazza, Npj Comput. Mater. 6 (2020) 173, doi:10.1038/s41524-020-00440-1.
[13] M.N. Gjerding, A. Taghizadeh, A. Rasmussen, S. Ali, F. Bertoldo, T. Deilmann, N.R. Knøsgaard, M. Kruse, A.H. Larsen, S. Manti, T.G. Pedersen, U. Petralanda, T. Skovhus, M.K. Svendsen, J.J. Mortensen, T. Olsen, K.S. Thygesen, 2D Mater 8 (2021) 044002, doi:10.1088/2053-1583/ac1059.
[14] T. Xin, Y. Zhao, R. Mahjoub, J. Jiang, A. Yadav, K. Nomoto, R. Niu, S. Tang, F. Ji, Z. Quadir, D. Miskovic, J. Daniels, W. Xu, X. Liao, L.-Q. Chen, K. Hagihara, X. Li, S. Ringer, M. Ferry, Sci. Adv. 7 (2021) eabf3039, doi:10.1126/sciadv.abf3039.
[15] Z.-Z. Jin, M. Zha, S.-Q. Wang, S.-C. Wang, C. Wang, H.-L. Jia, H.-Y. Wang, J. Magnes. Alloys 10 (2022) 1191–1206, doi:10.1016/j.jma.2022.04.002.
[16] W. Xu, N. Birbilis, G. Sha, Y. Wang, J.E. Daniels, Y. Xiao, M. Ferry, Nature Mater 14 (2015) 1229–1235, doi:10.1038/nmat4435.
[17] A. Fattah-alhosseini, R. Chaharmahali, K. Babaei, M. Nouri, M.K. Keshavarz, M. Kaseem, J. Magnes. Alloys 10 (2022) 2354–2383, doi:10.1016/j.jma.2022.09.002.
[18] X. Gu, Y. Zheng, Y. Cheng, S. Zhong, T. Xi, Biomaterials 30 (2009) 484–498, doi:10.1016/j.biomaterials.2008.10.021.
[19] J. Wang, Y. Yuan, T. Chen, L. Wu, X. Chen, B. Jiang, J. Wang, F. Pan, J. Magnes. Alloys 10 (2022) 1786–1820, doi:10.1016/j.jma.2022.06.015.
[20] N. Singh, U. Batra, K. Kumar, N. Ahuja, A. Mahapatro, Bioactive Materials 19 (2023) 717–757, doi:10.1016/j.bioactmat.2022.05.009.
[21] X. Chen, J. Venezuela, Z. Shi, L. Wang, M. Dargusch, Nano Energy 122 (2024) 109269, doi:10.1016/j.nanoen.2024.109269.
[22] W. Wang, Q. Pan, X. Wang, B. Liu, Corros. Sci. 226 (2024) 111695, doi:10.1016/j.corsci.2023.111695.
[23] M. Deng, L. Wang, D. Höche, S.V. Lamaka, C. Wang, D. Snihirova, Y. Jin, Y. Zhang, M.L. Zheludkevich, Mater. Horiz. 8 (2021) 589–596, doi:10.1039/D0MH01380C.
[24] Y. Wang, Q. Tang, X. Xu, P. Weng, T. Ying, Y. Yang, X. Zeng, H. Zhu, Acta Mater 255 (2023) 119063, doi:10.1016/j.actamat.2023.119063.
[25] H. Sun, G. Su, Y. Zhang, J.-C. Ren, X. Chen, H. Hou, Z. Ding, T. Zhang, W. Liu, Acta Mater 244 (2023) 118562, doi:10.1016/j.actamat.2022.118562.
[26] R. Tran, X.-G. Li, J.H. Montoya, D. Winston, K.A. Persson, S.P. Ong, Surf. Sci. 687 (2019) 48–55, doi:10.1016/j.susc.2019.05.002.
[27] D.A. Tompsett, S.C. Parker, M.S. Islam, J. Am. Chem. Soc. 136 (2014) 1418–1426, doi:10.1021/ja4092962.
[28] W. Li, D.Y. Li, Appl. Surf. Sci. 240 (2005) 388–395, doi:10.1016/j.apsusc.2004.07.017.
[29] S. Mosleh-Shirazi, G. Hua, F. Akhlaghi, X. Yan, D. Li, Sci. Rep. 5 (2015) 18154, doi:10.1038/srep18154.
[30] Q. Li, H. Lu, J. Cui, M. An, D. Li, RSC Adv. 6 (2016) 97606–97612, doi:10.1039/C6RA19563F.
[31] X. Wei, J. Wang, C. Wang, S. Zhu, L. Wang, S. Guan, J. Mater. Res. 37 (2022) 3792–3802, doi:10.1557/s43578-022-00752-6.
[32] A. Merchant, S. Batzner, S.S. Schoenholz, M. Aykol, G. Cheon, E.D. Cubuk, Nature 624 (2023) 80–85, doi:10.1038/s41586-023-06735-9.
[33] V. Gupta, K. Choudhary, B. DeCost, F. Tavazza, C. Campbell, W. Liao, A. Choudhary, A. Agrawal, Npj Comput. Mater. 10 (2024) 1, doi:10.1038/s41524-023-01185-3.
[34] Y. Zhou, Y. Ouyang, Y. Zhang, Q. Li, J. Wang, J. Phys. Chem. Lett. 14 (2023) 2308–2316, doi:10.1021/acs.jpclett.2c03288.
[35] Z. Wang, X. Lin, Y. Han, J. Cai, S. Wu, X. Yu, J. Li, Nano Energy 89 (2021) 106337, doi:10.1016/j.nanoen.2021.106337.
[36] S. Kiyohara, Y. Hinuma, F. Oba, J. Am. Chem. Soc. 146 (2024) 9697–9708, doi:10.1021/jacs.3c13574.
[37] J. Jiang, X. Geng, X. Zhang, J. Magnes. Alloys 11 (2023) 1906–1930, doi:10.1016/j.jma.2023.05.011.
[38] A. Palizhati, W. Zhong, K. Tran, S. Back, Z.W. Ulissi, J. Chem. Inf. Model. 59 (2019) 4742–4749, doi:10.1021/acs.jcim.9b00550.
[39] S. Lu, Q. Zhou, Y. Ouyang, Y. Guo, Q. Li, J. Wang, Nat. Commun. 9 (2018) 3405, doi:10.1038/s41467-018-05761-w.
[40] R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T.D. Hirzel, D. Duvenaud, D. Maclaurin, M.A. Blood-Forsythe, H.S. Chae, M. Einzinger, D.-G. Ha, T. Wu, G. Markopoulos, S. Jeon, H. Kang, H. Miyazaki, M. Numata, S. Kim, W. Huang, S.I. Hong, M. Baldo, R.P. Adams, A. Aspuru-Guzik, Nat. Mater. 15 (2016) 1120–1127, doi:10.1038/nmat4717.
[41] H. Xiao, R. Li, X. Shi, Y. Chen, L. Zhu, X. Chen, L. Wang, Nat. Commun. 14 (2023) 7027, doi:10.1038/s41467-023-42870-7.
[42] T.N. Kipf, M. Welling, Semi-Supervised Classification with Graph Convolutional Networks, arXiv preprint arXiv:1609.02907, (2016).
[43] T. Xie, J.C. Grossman, Phys. Rev. Lett. 120 (2018) 145301, doi:10.1103/PhysRevLett.120.145301.

[44] K. Choudhary, B. DeCost, Npj Comput. Mater. 7 (2021) 185, doi:10.1038/s41524-021-00650-1.
[45] C.W. Park, C. Wolverton, Phys. Rev. Mater. 4 (2020) 063801, doi:10.1103/PhysRevMaterials.4.063801.
[46] C. Chen, W. Ye, Y. Zuo, C. Zheng, S.P. Ong, Chem. Mater. 31 (2019) 3564–3572, doi:10.1021/acs.chemmater.9b01294.
[47] K. Bang, D. Hong, Y. Park, D. Kim, S.S. Han, H.M. Lee, Nat. Commun. 14 (2023) 3004, doi:10.1038/s41467-023-38758-1.
[48] S. Sanyal, J. Balachandran, N. Yadati, A. Kumar, P. Rajagopalan, S. Sanyal, P. Talukdar, MT-CGCNN: Integrating Crystal Graph Convolutional Neural Network with Multitask Learning for Material Property Prediction, arXiv preprint arXiv:1811.05660, (2018).
[49] S. Ruder, An Overview of Multi-Task Learning in Deep Neural Networks, arXiv preprint arXiv:1706.05098, (2017).
[50] M.Lupo Pasini, P. Zhang, S.Temple Reeve, J.Youl Choi, Mach. Learn.: Sci. Technol. 3 (2022) 025007, doi:10.1088/2632-2153/ac6a51.
[51] Z. Chen, V. Badrinarayanan, C.-Y. Lee, A. Rabinovich, Proceedings of the 35th International Conference on Machine Learning, PMLR, 80, 2018, pp. 794–803 https://proceedings.mlr.press/v80/chen18a.
[52] P.E. Blöchl, Phys. Rev. B 50 (1994) 17953–17979, doi:10.1103/PhysRevB.50.17953.
[53] J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 77 (1996) 3865–3868, doi:10.1103/PhysRevLett.77.3865.
[54] G. Kresse, J. Furthmüller, Comput. Mater. Sci. 6 (1996) 15–50, doi:10.1016/0927-0256(96)00008-0.
[55] Y. Wang, T. Xie, Q. Tang, M. Wang, T. Ying, H. Zhu, X. Zeng, J. Magnes. Alloys 12 (2024) 1406–1418, doi:10.1016/j.jma.2021.12.007.
[56] S. Kirklin, J.E. Saal, B. Meredig, A. Thompson, J.W. Doak, M. Aykol, S. Rühl, C. Wolverton, Npj Comput. Mater 1 (2015) 15010, doi:10.1038/npjcompumats.2015.10.
[57] S. Curtarolo, W. Setyawan, G.L.W. Hart, M. Jahnatek, R.V. Chepulskii, R.H. Taylor, S. Wang, J. Xue, K. Yang, O. Levy, M.J. Mehl, H.T. Stokes, D.O. Demchenko, D. Morgan, Comput. Mater. Sci. 58 (2012) 218–226, doi:10.1016/j.commatsci.2012.02.005.
[58] W E., J. Han, L. Zhang, Integrating Machine Learning with Physics-Based Modeling, arXiv preprint arXiv:2006.02619, (2020).
[59] P. Staikov, Talat.S. Rahman, Phys. Rev. B 60 (1999) 15613–15616, doi:10.1103/PhysRevB.60.15613.
[60] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, PyTorch: An imperative style, high-performance deep learning library, Proc. Adv. Neural Inf. Process. Syst. 32 (2019) 8026–8037, doi:10.5555/3454287.3455008.
[61] J.S. Bergstra, R. Bardenet, Y. Bengio, B. Kégl, Algorithms for hyper-parameter optimization, Proc. Adv. Neural Inf. Process. Syst. 24 (2011) 2546–2554, doi:10.5555/2986459.2986743.
[62] J. Bergstra, D. Yamins, D.D. Cox, in: Proceedings of the 30th International Conference on Machine Learning, PMLR, 28, 2013, pp. 115–123. https://proceedings.mlr.press/v28/bergstra13.
[63] R. Tran, Z. Xu, B. Radhakrishnan, D. Winston, W. Sun, K.A. Persson, S.P. Ong, Sci. Data 3 (2016) 160080, doi:10.1038/sdata.2016.80.
[64] Q. Zhou, P. Tang, S. Liu, J. Pan, Q. Yan, S.-C. Zhang, Proc. Natl. Acad. Sci. U.S.A. 115 (28) (2018) E6411–E6417, doi:10.1073/pnas.1801181115.
[65] D.G. Pettifor, Mater. Sci. Technol. 4 (8) (1988) 675–691, doi:10.1179/mst.1988.4.8.675.
[66] E. Ringe, R.P. Van Duyne, L.D. Marks, Nano Lett 11 (2011) 3399–3403, doi:10.1021/nl2018146.
[67] G.L. Song, Corros. Prevent. Magn. Alloy (2013) 3–37, doi:10.1533/9780857098962.1.3.
[68] R.L. Liu, M.F. Hurley, A. Kvryan, G. Williams, J.R. Scully, N. Birbilis, Sci. Rep. 6 (2016) 28747, doi:10.1038/srep28747.
[69] N. Birbilis, G. Williams, K. Gusieva, A. Samaniego, M.A. Gibson, H.N. McMurray, Electrochem. Commun. 34 (2013) 295–298, doi:10.1016/j.elecom.2013.07.021.
[70] L. Yuan, Z. Geng, J. Xu, F. Guo, C. Han, Adv. Funct. Mater. 31 (2021) 2101103, doi:10.1002/adfm.202101103.
[71] R.K. Barik, A. Bera, R.S. Raju, A.K. Tanwar, I.K. Baek, S.H. Min, O.J. Kwon, M.A. Sattorov, K.W. Lee, G.-S. Park, Appl. Surf. Sci. 276 (2013) 817–822, doi:10.1016/j.apsusc.2013.04.004.
[72] E.R. Antoniuk, Y. Yue, Y. Zhou, P. Schindler, W.A. Schroeder, B. Dunham, P. Pianetta, T. Vecchione, E.J. Reed, Phys. Rev. B 101 (2020) 235447, doi:10.1103/PhysRevB.101.235447.
[73] J.W. Schwede, T. Sarmiento, V.K. Narasimhan, S.J. Rosenthal, D.C. Riley, F. Schmitt, I. Bargatin, K. Sahasrabuddhe, R.T. Howe, J.S. Harris, N.A. Melosh, Z.-X. Shen, Nat. Commun. 4 (2013) 1576, doi:10.1038/ncomms2577.
[74] Y. Zhang, W.Y. Wang, P. Li, K. Ren, Y. He, X. Gao, H. Kou, J. Wang, Y. Wang, H. Song, X. Liang, J. Li, Scripta Mater 244 (2024) 116026, doi:10.1016/j.scriptamat.2024.116026.

Please cite this article as: G. Shi, Y. Wang, K. Yang et al., A surface emphasized multi-task learning framework for surface property predictions: A case study of magnesium intermetallics, Journal of Magnesium and Alloys, https://doi.org/10.1016/j.jma.2024.12.005