Convolutional neural network of atomic surface structures to predict binding energies for high-throughput screening of catalysts

Seoin Back,[†,‡] Junwoong Yoon,[†,‡] Nianhan Tian,[†] Wen Zhong,[†] Kevin Tran,[†] and Zachary W. Ulissi[*,†]

†*Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA*

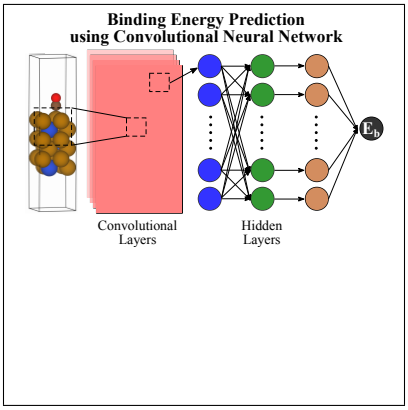
‡*These authors contributed equally to this work.*

E-mail: zulissi@andrew.cmu.edu

Abstract

High-throughput screening of catalysts can be performed using density functional theory calculations to predict catalytic properties, often correlated with adsorbate binding energies. However, more complete investigations would require an order of 2 more calculations compared to the current approach, making the computational cost a bottleneck. Recently developed machine-learning methods have been demonstrated to predict these properties from hand-crafted features but have struggled to scale to large composition spaces or complex active sites. Here, we present an application of deep-learning convolutional neural network of atomic surface structures using atomic and Voronoi polyhedra-based neighbor information. The model effectively learns the most important surface features to predict binding energies. Our method predicts CO and H binding energies after training with 12,000 data for each adsorbate with a mean absolute error of 0.15 eV for diverse chemical space. Our method is also capable of creating saliency maps that determine atomic contributions to binding energies.

Graphical TOC Entry



Keywords

machine-learning, high-throughput screening, density functional theory calculations, heterogeneous catalysts, convolutional neural network

Understanding atomic and molecular interactions with solid surfaces is a basis for predicting catalytic properties in heterogeneous catalysis. Density functional theory (DFT) calculations have played an important role in understanding these interactions and estimating binding strengths accurately with respect to the corresponding experimental measurements (MAE = $\sim$0.2 eV).[1,2] Binding energies can be used as descriptors to predict catalytic properties such as electrochemical onset potentials,[3,4] turn over frequencies,[5] product selectivity[6] based on scaling relations between binding energies of reaction intermediates,[3,7,8] and Brønsted-Evans-Polanyi relation between reaction energies and kinetic barriers.[8,9]

In the past decades, high-throughput screenings of catalysts have been performed mainly using simple metal/alloy systems, and the most stable and second most stable facets such as FCC (111) and (100) have been used due to their low surface energies.[10,11] In addition, one descriptor has been usually used to predict catalytic activities—e.g., H, OH and CO binding energies to predict theoretical overpotentials of H_2 evolution/oxidation reactions, O_2 reduction reaction and CO_2 reduction reaction, respectively.[11–14] To more completely study catalytic properties of materials, however, we need to reconsider those two assumptions. First, the most stable facet is not always most active. For example, concave Au nanoparticles consisting of a large portion of high index facets showed much higher catalytic activity and selectivity compared to thin film consisting of stable facets.[15] Therefore, it is essential to consider several possible facets and a number of unique active sites on those facets,[16] increasing the computational cost by a factor of 20, assuming we consider 5 facets and 4 possible active sites. Second, one may want to extend a chemical space to explore, but it has been frequently observed that the scaling relations do not hold anymore for systems other than metal/alloy systems. Examples include single metal atom embedded in two-dimensional materials, p-block atom embedded in metals, and metal dichalcogenides.[17–20] Therefore, we need to consider all reaction intermediates, increasing the computational cost by a factor of 5$\sim$10.[4,21] Altogether, the required computational cost becomes an order of 2 more expensive than the current high-throughput screenings, making the computational cost of DFT a main

bottleneck. Thus, it is essential to develop strategies to alleviate the computational cost issue.

Recent developments of machine-learning (ML) techniques and their applications to catalysis have suggested that the ML could substantially facilitate the high-throughput screenings. Based on ML predicted binding energies, one could reduce the number of DFT calculations by excluding unpromising candidates.[16,22,23] For example, using fingerprints based on selected chemical and physical features of active sites and their neighbor information, many artificial neural network (ANN) models have been successful in achieving an acceptable accuracy (MAE = $\sim$0.2 eV) with respect to the DFT calculated binding energies. Unfortunately, many models are based on electronic features of active surface atoms, such as d-band characteristics, requiring *ab-initio* level computational cost to prepare them. Therefore, ML models that use readily available data and that achieve high accuracy with the minimal number of training DFT data are strongly desirable to pursue an efficient catalyst search over the broad chemical space.

Recently, Xie and Grossman applied a deep-learning convolutional neural network (CNN) on top of a graph representation of bulk crystals to predict various properties[24,25] and reached DFT accuracy after training with $\sim$30,000 DFT calculated data. The graph representation of the crystals includes information of atomic features and interatomic distances, and iterative convolutions extract neighbor information to update atomic feature vector. Note that none of expensive electronic structure information was used during the process; only the crystal structures and basic atomic features were used.

In this work, we use a modified form of this CNN method to predict CO and H binding energies on diverse surfaces of pure metals, metal alloys, and intermetallic surfaces. In our method, we collect the graph connections based on Voronoi polyhedra to take into account additional adsorbate atoms. We report state-of-the-art MAEs of 0.15 eV for both CO and H binding energy predictions using 12,000 training data for each adsorbate across a far larger collection of surface composition and structure than previously possible. We further show

that our method can be used to create saliency maps to determine atomic contributions to binding energies and to automatically detect failed DFT calculations.

We modified the bulk crystal prediction code to better represent surface features. In the previous effort of the CNN applications to bulk property predictions, the authors encoded atomic information and interatomic bonding interactions into node and edge vectors, respectively,[24] where the node vectors include the basic atomic information and the edge vectors contain atomic indices of nearest neighbor atoms and their bonding distances from the center atom. The main difference between bulk structures and our atomic surface structures is that the latter contain adsorbate atoms, where their initial positions are arbitrarily set by users. Since a goal of our method is to predict binding energies using initial structures as inputs, the arbitrariness of the initial adsorbate position should be addressed. We note that there is no such arbitrariness in the graph representation of bulk crystals, since the final relaxed structures were used.[24] To resolve this arbitrariness, we modified the code so that we can incorporate neighbor information based on Voronoi polyhedra as implemented in Pymatgen[26] as "VoronoiConnectivity" (Figure 1b). Voronoi polyhedra, also known as Wigner-Seitz cell,[27] of atoms in surface structures could provide local environment information as a solid angle. Each atomic polyhedron encompasses space in which distance to the center atom is less than or equal to the distance to other atoms. Solid angles between the plane shared by polyhedra of two adjacent atoms and center atoms are calculated, and values normalized to the maximum solid angle that belongs to the center atom were used to represent local environments of the center atom. We note that larger solid angles correspond to stronger interaction between the center and the neighbor atoms. To prepare atomic feature vectors, we used the identical tabulated atomic information as in the original code.[24]

In some cases, the binding site changes from the initially set site to a more stable site during structure relaxation—e.g., from a top site to a bridge site. Since binding site changes could affect a prediction performance, we tested this effect by including a connectivity distance information from adsorbate atoms to all atoms in the surface structure taken from

the final relaxed structures (Figure 1). We compared the prediction performance of training with initial structures only, final structures only, and initial structures with the connectivity distance information of adsorbates obtained from the final structures (Figure S1). Obviously, the prediction performance of training with final structures was best (MAE = 0.13 eV) and that with initial structures was worst (MAE = 0.19 eV), while that with initial structures and the adsorbate distance information lies in between (MAE = 0.15 eV). For the rest of this work, we present the results trained with initial structures and the adsorbate distance information, but we highlight that training with initial structures only already demonstrated a reasonable prediction accuracy.

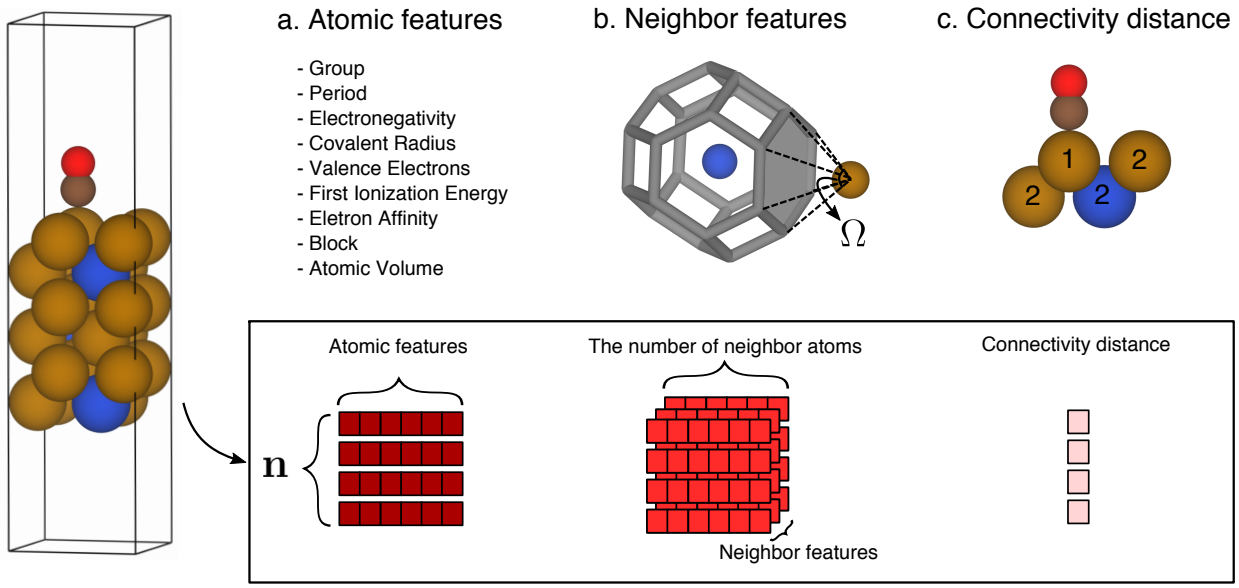


Figure 1: A graphical representation of converting an atomic structure containing **n** atoms into numerical inputs for the convolutional neural network. (a) 9 basic atomic properties are presented by one-hot encoding[24] to prepare the atomic feature vectors. (b) Neighbor information was encoded using the solid angle (Ω) based on the Voronoi polyhedron. The grey skeleton indicates the Voronoi polyhedron of Cu atom, and the solid angle between Fe atom and the shared plane of Cu and Fe polyhedrons is marked. Other nearest neighbor atoms were omitted for simplicity. Color code for atoms: Cu (blue), Fe (brown), C (grey), O (red). (c) The connectivity distances from adsorbate to all atoms in surface structures are counted. A side view of the surface structure with adsorbate CO molecule, surface atoms up to the second layer and their corresponding connectivity distances are presented as an example.

Once we converted atomic structures into graph representations, we concatenated atomic

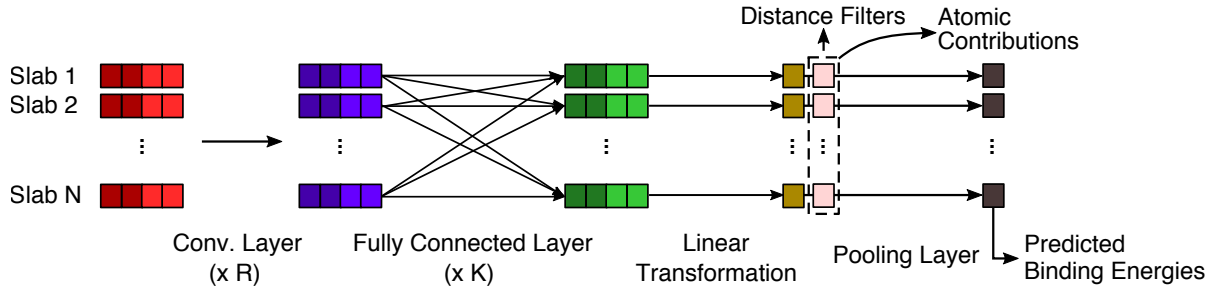


Figure 2: Schematic illustration of the convolutional neural network on top of the graph representation of atomic surface structures. Dark red and red blocks correspond to atomic and neighbor feature vectors as in Figure 1. Pink blocks correspond to the connectivity distance vectors used as a filter to exclude atoms that are too far from adsorbates (higher than the connectivity distance of 2). More details are elaborated in the main text.

feature vectors of atom **i** and **j** ($\mathbf{v}_i$ and $\mathbf{v}_j$) and neighbor feature vectors between atom **i** and **j** ($\mathbf{u}_{(i,j)}$), i.e., $\mathbf{z}_{(i,j)} = \mathbf{v}_i \oplus \mathbf{v}_j \oplus \mathbf{u}_{(i,j)}$. We then utilized the same convolutional layer (eqn (1)) as described in Ref.[24]

$$\mathbf{v}_i^t = \mathbf{v}_i^{t-1} + \sum_j \sigma\left(\mathbf{z}_{(i,j)}^{t-1}\mathbf{W}_f^{t-1} + \mathbf{b}_f^{t-1}\right) \odot g\left(\mathbf{z}_{(i,j)}^{t-1}\mathbf{W}_s^{t-1} + \mathbf{b}_s^{t-1}\right) \tag{1}$$

where $\odot$ denotes an element-wise multiplication, σ denotes a sigmoid function, g denotes non-linear activation functions ("Leaky ReLu" in this study), **W** and **b** denote weights and biases of the neural networks, respectively. After **R** convolutional layers, resulting vectors are then fully connected via **K** hidden layers, followed by a linear transformation to scalar values. Distance filters were applied to exclude contributions of atoms that are too far from the adsorbates. At this stage, (1) atomic contributions of each atom to binding energies could be predicted and (2) mean pooling layer is applied to predict DFT calculated binding energies (Figure 2).

The neural network is trained to minimize the loss function (MAE) between the DFT calculated and predicted binding energies using Adam optimizer with decoupled weight decay (L_2 Regularization coefficient: 10^{-5}) and warm restart.[28] We tested two pooling functions (sum and mean) and four activation functions (Sigmoid, Softplus, leaky ReLu, ReLu) and the comparison results are presented in Figure S2. We divided DFT results into 20 % and

80 % of test and training sets, respectively. To prevent the overfitting, the validation set (25 % of training sets) was used to optimize hyperparameters (Table S1).

Table 1: A comparison of our DFT data and the number of parameters of our model to literature values, where the goal was to predict binding energies. We counted all unique components (element, stoichiometry, space group, facet) in literature and our DFT data. For the previous studies that did not explicitly mention the number of total DFT data or parameters, we estimated based on the presented results. The number of parameters indicate how complicated the model is. Specific parameter components of our model are summarized in Table S2.

	Element	Stoichiometry	Space group	Facet	Total DFT data	Parameters	Methods
Calle-Vallejo et al.[29]	1	1	1	22	132	2	GCN[a]
Abild-Pederson et al.[30]	13	1	3	6	165	2	SR[b]
Dickens et al.[31]	45	3	3	3	901	2	ED[c]
Noh et al.[32]	28	3	1	1	263	$10^1 \sim 10^2$	ML[d]
Andersen et al.[33]	9	3	1	4	884	$\sim 10^1$	ML[d]
Li et al.[22]	24	4	1	1	1,032	$\sim 10^2$	ML[d]
Batchelor et al.[34]	5	1	1	1	1,869	15 (O), 55 (OH)	ML[d]
This work	37	96	110	41	43,247	4,938 (CO), 6,738 (H)	ML[d]

[a]generalized coordination number, [b]scaling relations, [c]electronic descriptors, [d]machine-learning.

We first compare how our DFT datasets and method are different from the previous approaches[29–34] (Table 1). First, the DFT data used in most previous studies covered only the limited chemical space—e.g., materials consisting of up to two elements based mainly on transition metals, fixed elemental compositions for alloys (1:1, 1:3) and low-index facets of the most common crystal structures (FCC (111), FCC (100), HCP (0001)). On the other hand, our DFT datasets include materials consisting of up to four elements using 37 elements, 96 stoichiometries, 110 space groups and 41 facets. Second, the number of parameters in our CNN method is substantially more compared to the previous approaches. For binding energy predictions through scaling relations,[30] generalized coordination numbers[29] or electronic descriptors,[31] only two parameters (slope and bias) are optimized through a linear regression. ANNs consisting of an input layer, several hidden layers and an output layer are reported to have $\sim 10^2$ parameters.[22,33,34] In our deep-learning CNN, there are thousands of parameters to be optimized during the training so that it can predict binding energies on a variety of surfaces of catalysts. However, we note that simpler methods are useful in specific

cases, which will be discussed in the last part.

Figure 3 shows the performance of our method on predicting CO and H binding energies. We note that a difference in the prediction accuracy between CO and H using the small number of data could be originated from their data distribution; the standard deviations of CO and H DFT binding energies are 0.65 and 0.42 eV, respectively. In both cases, we observed a systematic decrease in the error metric (MAE) as the number of training atomic structures increased. Interestingly, we observed a near convergence of the prediction accuracy at 5,000 and 8,000 training data for CO and H, respectively, similar to our previous work.[16] For comparison, we also plotted MAE obtained using 12,000 training data from our previous work,[16] where fingerprint-based surrogate models were trained. We note that exact comparison is not available since training/validation/test data splitting was not used previously, but training/test splitting was used instead. The previously reported values are, thus, the lower-limit accuracy of the model. The best MAEs we achieved for the CO and H binding energy prediction are 0.15 eV in both cases, and 86 % of test data are within the accuracy of 0.25 eV. We highlight that MAEs for CO and H binding energy prediction in this study are lower by 0.05 eV compared to our previous model. Given a reasonable accuracy of our model and facile preparation of input data, we expect this method could be applied for the high-throughput catalyst screening to remove unpromising candidates, thus, reducing the number of DFT jobs.

After the convolutional layers and the fully connected layers, the resulting output vectors are linearly transformed to scalar values for each atom, which are then filtered by the connectivity distance criterion. We only considered atoms with the connectivity distance up to 2 as they are expected to strongly interact with adsorbates. These values are then averaged in the pooling layer and the resulting value corresponds to the predicted binding energies. Thus, the linearly transformed scalar values represent the contribution of corresponding atoms to the predicted binding energies. We note that the prediction accuracy change was negligible (< 0.01 eV) even we included all the atoms, indicating that the first and the sec-

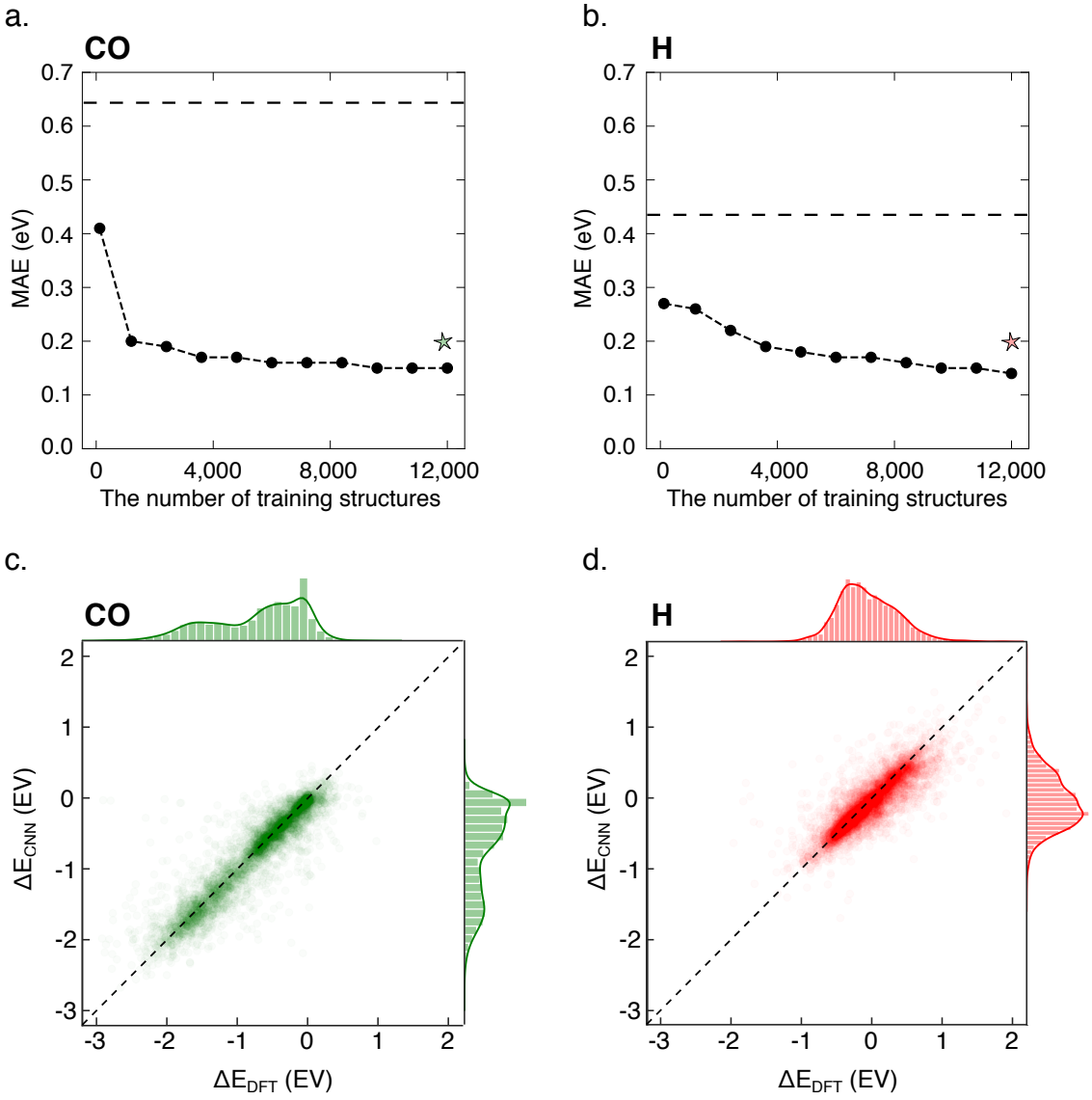


Figure 3: Mean absolute error (MAE) of test set with respect to the number of training atomic structures for binding energy predictions of (a) CO and (b) H. The horizontal dashed lines correspond to the standard deviation of DFT binding energies. As a comparison, we added the prediction error from our previous study (green and red star).[16] Two-dimensional histograms of DFT-calculated and CNN-predicted binding energies of (c) CO and (d) H.

ond nearest-neighbor of adsorbates mainly contribute to binding energies. In Figure 4, we present a graphic representation of each atomic contribution to binding energies using CO adsorption at top and bridge sites of Cu (211) surface as an example. We normalized the atomic contribution into [0,1] range by mapping the minimum and the maximum values of the contributions to 0 and 1, respectively. Clearly, we observe that surface atoms directly interacting with adsorbates have the highest contribution to the binding energies. We also visualized atomic contributions using the connectivity distance up to 4, and, as expected, their contributions are found to be negligible (Figure S3). This result rationalizes the use of the first and second nearest neighbor to extract neighbor information of active sites for the preparation of fingerprints in our previous model.[16] An example of bimetallic surface is shown in Figure S4. We note that the conventional DFT calculations are not capable of isolating the calculated binding energies into each atomic contribution.

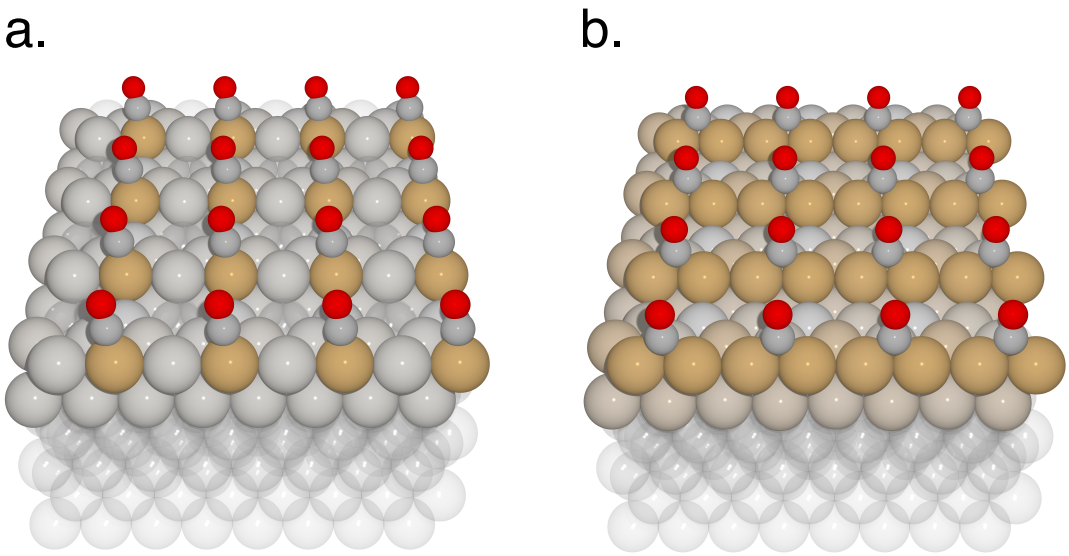


Figure 4: Graphic representations of atomic contributions to CO binding energies at (a) top and (b) bridge site of Cu (211) surface. Darker spheres indicate higher contributions. Only atoms within the connectivity distance of 2 are considered and transparent spheres indicate atoms with connectivity distances higher than 2.

Another feature of our method is to automatically detect wrong DFT calculations in datasets, which may take lots of time by human analysis. Since our workflow in *GASpy* automatically performs DFT calculations (see the Computational Methods for details of cal-

culations), there could be a few incorrect atomic structures and binding energies, where the final relaxed structures significantly differ from the initial structures due to surface reconstructions or unphysical initial atomic structures (Figure S5). Since our model uses neighbor information based on the Voronoi polyhedra, it cannot accurately predict binding energies of structures where atomic structural change is significant. After training, we considered data outside of $\pm 3\sigma$ as outliers and this process was repeated two times (Figure S6 and S7). Out of 20,771 and 22,361 data for CO and H binding, 622 and 628 outliers were detected.

The deep-learning CNN results presented in this work suggest that, given sufficient training data, it could be effective in reducing computational cost issue in the high-throughput screenings. However, we highlight that simpler methods may be more effective for specific cases or smaller datasets. Details are discussed in the following:

- For systems consisting of single metal element, generalized coordination number (GCN) is the best choice. The GCN analysis on low and high index Pt facets in various sizes of Pt nanoparticles predicted OH* binding energies very accurately (MAE $\sim$0.056 eV) using only 20 DFT data.[29] In addition, one can easily count the GCN of active sites. However, an extension of this method to alloy or other materials has not been reported.

- Scaling relations and electronic descriptors, such as d-band center, are suitable for a dozens of metal/alloy DFT data. The drawback of the scaling relation is that it is adsorbate-specific and a model should be optimized for each adsorbate of interest.[30] For d-band center analysis, one should perform DFT optimization and post density-of-states analysis to calculate d-band center values. Further, a preparation of inputs requires DFT calculations for new systems to obtain binding energies of atomic species or d-band center values.

- Artificial neural networks (ANN) have also shown a similar accuracy to the scaling relations or electronic descriptors,[22,32–34] and they are suitable for hundreds/thousands of DFT data of similar metal/alloy systems. They require a careful preprocessing of

fingerprints which could considerably affect the prediction accuracy. A preparation of inputs for a binding energy prediction of new systems is usually fast as most of the fingerprints are determined beforehand.

- The deep-learning methods, CNN as in this study, could predict binding energies for a variety of surfaces without a preparation of fingerprints as they collect all the information automatically. However, the main drawback of the deep-learning is that it requires lots of data ($>$ 10,000 DFT data). As we want to include diverse materials for the high-throughput screenings, the deep-learning methods would be ideally suited to this purpose.

In summary, we modified the original crystal graph CNN code by Xie and Grossman[24] to collect neighbor information using Voronoi polyhedra for the application in predicting binding energies on heterogeneous catalyst surfaces. Our method predicted CO and H binding energies with 0.15 eV MAE for a variety of materials which have never been considered previously. Given the reasonable accuracy of our model, we expect that binding energy prediction from our method could help to rule out unpromising candidates and reduce the number of DFT jobs in the high-throughput catalyst screening to effectively facilitate a catalyst discovery. Furthermore, our method successfully partitioned the calculated binding energies into each atomic contribution, rationalizing the use of the first and second nearest neighbor atoms for the preparation of fingerprints. It also automatically detected wrong DFT calculations generated during the high-throughput calculations, which will be useful to remove those outliers from the datasets.

Computational Methods

We collected DFT calculated results of CO and H binding on surfaces of pure metals, metal alloys and intermetallic alloys from our open-source *GASpy* database ($\sim$20,000 data for each adsorbate). In the following, the calculation procedure is elaborated and more details can

be found in our previous publication.[16]

Our DFT calculated data were automatically generated as follows; 1) 1,499 bulk structures from the Materials Project[35] are relaxed. 2) Using the optimized bulk structures, unique surfaces with Miller indices between -2 and 2 are enumerated, resulting in 17,507 surfaces. 3) All unique active sites are identified, 1,684,908 sites in total. 4) Using "Materials Id" in the Materials Project as an user input, all the required jobs (structure relaxation with and without adsorbates) are automatically submitted to supercomputers. 4) The calculated binding energies and initial/final structures of successfully completed jobs with a residual force less than 0.05 eV/Å are uploaded to our *GASpy* database.

We used Pymatgen[26] to enumerate various surfaces and to find unique active sites. We used Luigi[36] and Fireworks[37] to manage high-throughput DFT calculations across many clusters. DFT calculations were performed using VASP code[38,39] with the revised Perdew-Burke-Ernzerhof (RPBE) functional[2] and projector augmented wave (PAW) pseudopotential.[40]

Acknowledgement

This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Supporting Information Available

The optimized hyperparameters of the convolutional neural networks. Various test results to optimize the convolutional neural networks. A summarized table of trained parameters. Graphic representations of atomic contributions using different cutoff values of connectivity distance. Examples of outliers.

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Hensley, A. J.; Ghale, K.; Rieg, C.; Dang, T.; Anderst, E.; Studt, F.; Campbell, C. T.; McEwen, J.-S.; Xu, Y. DFT-based method for more accurate adsorption energies: an adaptive sum of energies from RPBE and vdW density functionals. *J. Phys. Chem. C* **2017**, *121*, 4937–4945.

(2) Hammer, B.; Hansen, L. B.; Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Phys. Rev. B* **1999**, *59*, 7413.

(3) Man, I. C.; Su, H.-Y.; Calle-Vallejo, F.; Hansen, H. A.; Martínez, J. I.; Inoglu, N. G.; Kitchin, J.; Jaramillo, T. F.; Nørskov, J. K.; Rossmeisl, J. Universality in oxygen evolution electrocatalysis on oxide surfaces. *ChemCatChem* **2011**, *3*, 1159–1165.

(4) Peterson, A. A.; Abild-Pedersen, F.; Studt, F.; Rossmeisl, J.; Nørskov, J. K. How copper catalyzes the electroreduction of carbon dioxide into hydrocarbon fuels. *Energy & Environ. Sci.* **2010**, *3*, 1311–1315.

(5) Singh, S.; Li, S.; Carrasquillo-Flores, R.; Alba-Rubio, A. C.; Dumesic, J. A.; Mavrikakis, M. Formic acid decomposition on Au catalysts: DFT, microkinetic modeling, and reaction kinetics experiments. *AIChE J.* **2014**, *60*, 1303–1319.

(6) Yoo, J. S.; Abild-Pedersen, F.; Nørskov, J. K.; Studt, F. Theoretical analysis of transition-metal catalysts for formic acid decomposition. *ACS Catal.* **2014**, *4*, 1226–1233.

(7) Jones, G.; Bligaard, T.; Abild-Pedersen, F.; Nørskov, J. K. Using scaling relations to understand trends in the catalytic activity of transition metals. *J. Phys.: Condens. Matter* **2008**, *20*, 064239.

(8) Wang, S.; Petzold, V.; Tripkovic, V.; Kleis, J.; Howalt, J. G.; Skulason, E.; Fernandez, E.; Hvolbæk, B.; Jones, G.; Toftelund, A. Universal transition state scaling relations for (de) hydrogenation over transition metals. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20760–20765.

(9) Bligaard, T.; Nørskov, J. K.; Dahl, S.; Matthiesen, J.; Christensen, C. H.; Sehested, J. The Brønsted–Evans–Polanyi relation and the volcano curve in heterogeneous catalysis. *J. Catal.* **2004**, *224*, 206–217.

(10) Montemore, M. M.; Medlin, J. W. Scaling relations between adsorption energies for computational screening and design of catalysts. *Catal. Sci. Technol.* **2014**, *4*, 3748–3761.

(11) Greeley, J.; Stephens, I.; Bondarenko, A.; Johansson, T. P.; Hansen, H. A.; Jaramillo, T.; Rossmeisl, J.; Chorkendorff, I.; Nørskov, J. K. Alloys of platinum and early transition metals as oxygen reduction electrocatalysts. *Nat. Chem.* **2009**, *1*, 552.

(12) Greeley, J.; Nørskov, J. K. Combinatorial density functional theory-based screening of surface alloys for the oxygen reduction reaction. *J. Phys. Chem. C* **2009**, *113*, 4932–4939.

(13) Greeley, J.; Nørskov, J. K. Large-scale, density functional theory-based screening of alloys for hydrogen evolution. *Surf. Sci.* **2007**, *601*, 1590–1598.

(14) Hansen, H.; Shi, C.; Lausche, A.; Peterson, A.; Nørskov, J. Bifunctional alloys for the electroreduction of CO_2 and CO. *Phys. Chem. Chem. Phys.* **2016**, *18*, 9194–9201.

(15) Lee, H.-E.; Yang, K. D.; Yoon, S. M.; Ahn, H.-Y.; Lee, Y. Y.; Chang, H.; Jeong, D. H.; Lee, Y.-S.; Kim, M. Y.; Nam, K. T. Concave rhombic dodecahedral Au nanocatalyst with multiple high-index facets for CO_2 reduction. *ACS Nano* **2015**, *9*, 8384–8393.

(16) Tran, K.; Ulissi, Z. W. Active learning across intermetallics to guide discovery of electrocatalysts for CO_2 reduction and H_2 evolution. *Nat. Catal.* **2018**, *1*, 696.

(17) Gani, T. Z.; Kulik, H. J. Understanding and Breaking Scaling Relations in Single-Site Catalysis: Methane to Methanol Conversion by Fe^{IV}=O. *ACS Catal.* **2018**, *8*, 975–986.

(18) Chan, K.; Tsai, C.; Hansen, H. A.; Nørskov, J. K. Molybdenum sulfides and selenides as possible electrocatalysts for CO_2 reduction. *ChemCatChem* **2014**, *6*, 1899–1905.

(19) Back, S.; Lim, J.; Kim, N.-Y.; Kim, Y.-H.; Jung, Y. Single-atom catalysts for CO_2 electroreduction with significant activity and selectivity improvements. *Chem. Sci.* **2017**, *8*, 1090–1096.

(20) Lim, H.-K.; Shin, H.; Goddard III, W. A.; Hwang, Y. J.; Min, B. K.; Kim, H. Embedding covalency into metal catalysts for efficient electrochemical conversion of CO_2. *J. Am. Chem. Soc.* **2014**, *136*, 11355–11361.

(21) Back, S.; Tran, K.; Ulissi, Z. Towards a design of active oxygen evolution catalysts: Insights from automated density functional theory calculations and machine learning. *ChemRxiv* **2019**, DOI: 10.26434/chemrxiv.7926869.

(22) Li, Z.; Wang, S.; Chin, W. S.; Achenie, L. E.; Xin, H. High-throughput screening of bimetallic catalysts enabled by machine learning. *J. Mater. Chem. A* **2017**, *5*, 24131–24138.

(23) Li, Z.; Ma, X.; Xin, H. Feature engineering of machine-learning chemisorption models for catalyst design. *Catal. Today* **2017**, *280*, 232–238.

(24) Xie, T.; Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **2018**, *120*, 145301.

(25) Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S. P. Graph Networks as a Universal

Machine Learning Framework for Molecules and Crystals. *Chem. Mater.* **2019**, *31*, 3564–3572.

(26) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci* **2013**, *68*, 314–319.

(27) Wigner, E.; Seitz, F. On the constitution of metallic sodium. *Phys. Rev.* **1933**, *43*, 804.

(28) Loshchilov, I.; Hutter, F. Fixing weight decay regularization in adam. *arXiv* **2017**, preprint arXiv:1711.05101.

(29) Calle-Vallejo, F.; Martínez, J. I.; García-Lastra, J. M.; Sautet, P.; Loffreda, D. Fast prediction of adsorption properties for platinum nanocatalysts with generalized coordination numbers. *Angew. Chem. Int. Ed.* **2014**, *53*, 8316–8319.

(30) Abild-Pedersen, F.; Greeley, J.; Studt, F.; Rossmeisl, J.; Munter, T.; Moses, P. G.; Skulason, E.; Bligaard, T.; Nørskov, J. K. Scaling properties of adsorption energies for hydrogen-containing molecules on transition-metal surfaces. *Phys. Rev. Lett.* **2007**, *99*, 016105.

(31) Dickens, C. F.; Montoya, J. H.; Kulkarni, A. R.; Bajdich, M.; Nørskov, J. K. An electronic structure descriptor for oxygen reactivity at metal and metal-oxide surfaces. *Surf. Sci.* **2019**, *681*, 122–129.

(32) Noh, J.; Back, S.; Kim, J.; Jung, Y. Active learning with non-ab initio input features toward efficient CO_2 reduction catalysts. *Chem. Sci.* **2018**, *9*, 5152–5159.

(33) Andersen, M.; Levchenko, S.; Scheffler, M.; Reuter, K. Beyond scaling relations for the description of catalytic materials. *ACS Catal.* **2019**, *9*, 2752–2759.

(34) Batchelor, T. A.; Pedersen, J. K.; Winther, S. H.; Castelli, I. E.; Jacobsen, K. W.; Rossmeisl, J. High-entropy alloys as a discovery platform for electrocatalysis. *Joule* **2019**, *3*, 834–845.

(35) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **2013**, *1*, 011002.

(36) Bernhardsson, E.; Freider, E.; Rouhani, A. Luigi, a Python package that builds complex pipelines of batch jobs. 2012; `https://github.com/spotify/luigi`, (bithub, 2012).

(37) Jain, A.; Ong, S. P.; Chen, W.; Medasani, B.; Qu, X.; Kocher, M.; Brafman, M.; Petretto, G.; Rignanese, G.-M.; Hautier, G. et al. FireWorks: a dynamic workflow system designed for high-throughput applications. *Concurrency Computat.: Pract. Exper.* **2015**, *27*, 5037–5059.

(38) Kresse, G.; Hafner, J. Ab initio molecular dynamics for open-shell transition metals. *Phys. Rev. B* **1993**, *48*, 13115.

(39) Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6*, 15–50.

(40) Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **1999**, *59*, 1758.