Atomic Structure-free Representation of Active Motifs for Expedited Catalyst Discovery

*Dong Hyeon Mok and Seoin Back**

Department of Chemical and Biomolecular Engineering, Institute of Emergent Materials, Sogang University, Seoul 04107, Republic of Korea

AUTHOR INFORMATION

Corresponding Author

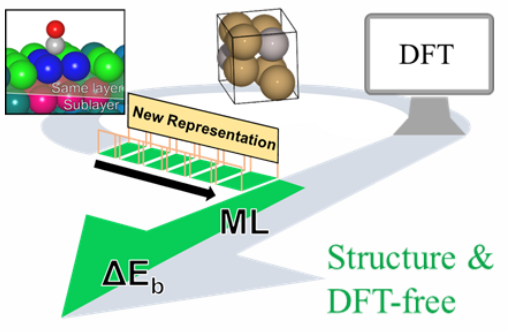
E-mail: sback@sogang.ac.kr (SB)

ABSTRACT

To discover new catalysts using density functional theory (DFT) calculations, binding energies of reaction intermediates are considered as descriptors and calculated to predict catalytic activities. Recently, machine learning methods have been developed to reduce the number of computationally intensive DFT calculations for high-throughput screening. These methods require several steps such as bulk structure relaxations, surface structure modelling, and active sites identification, which could be time-consuming as the number of new candidate materials increases. To bypass these procedures, in this work, we report atomic structure-free representation of active motifs to predict binding energies on catalyst surfaces. We identify binding site atoms and their nearest neighboring atoms positioned in the same layer and the sublayer, and their atomic properties are collected to construct fingerprints. Our method enabled a quicker training (~200 sec) compared to the previous deep-learning models, and predicted CO and H binding energies with mean absolute errors of 0.120 and 0.105 eV, respectively. Our method is also capable of creating all possible active motifs without performing any DFT calculations and the predicted binding energy distributions can suggest promising candidates to accelerate catalyst discovery.

KEYWORDS Binding Energy, High-throughput Screening, Machine Learning, Heterogeneous Catalysis, Density Functional Theory Calculations

TOC GRAPHICS



In computational catalysis field, binding energies of reaction intermediates are central for predicting catalytic properties from first-principles calculations. They can be used to predict binding energies of other reaction intermediates based on scaling relations[1-3] and estimate transition state energies via Brønsted−Evans−Polanyi (BEP) relation[4] without performing expensive nudged elastic band (NEB) calculations.[5, 6] By coupling scaling relations and BEP relations to perform a microkinetic analysis, volcano plots can be constructed to obtain a desired range of binding energies to achieve the maximum reaction rates.[7, 8] In this sense, binding energies are considered as descriptors for predicting catalytic activity and selectivity, which can be utilized to discover new catalysts by performing high-throughput screening for diverse classes and compositions of materials.[9-15]

A procedure of the conventional high-throughput screening is summarized as follows[16, 17]: (1) bulk structures are collected from inorganic materials database such as The Materials Project[18] or The Open Quantum Materials Database (OQMD)[19], (2) lattice parameters and a shape of the bulk unit cell are optimized using density functional theory (DFT) calculations, (3) surface atomic structures (e.g., (111) or (211) facets) are modelled and (4) binding energies at all unique sites are calculated. As the number of candidate materials increases, the number of required DFT calculations would significantly increase. To tackle this challenge, various approaches have been developed to date to reduce the number of required DFT calculations.

For example, Hammer and Nørskov established *d*-band theory, which correlates adsorbate binding energies and *d*-band center of surface atoms[20], which was further improved by considering other properties of *d*-band such as *d*-band upper edge[21, 22] and 0−4 th moment characteristics (filling, center, width, skewness and kurtosis).[23] With regard to structural effects on binding energies, Calle-Vallejo et al. developed a "generalized" coordination number, which is an weighted

average of surface coordinations.[24] Recently, machine learning (ML) has been applied to the computational catalysis to alleviate formidable computational cost issues of DFT calculations. To predict binding energies on catalyst surfaces, Tran and Ulissi prepared fingerprints using information of coordination sites such as coordination numbers of binding sites, electronegativity, atomic numbers and median binding energies of the first and second nearest neighboring (NN) atoms, and their regression models predicted ΔE_{CO*} and ΔE_{H*} reasonably well (mean absolute error (MAE) ~ 0.2 eV) using ~20,000 DFT calculated data of intermetallic alloys for each adsorbate.[25] Using the same dataset, Back et al.[9] reported convolutional neural network (CNN) on top of a graph representation of surface atomic structures, the modified version of CGCNN originally developed by Xie and Grossman[26], to predict ΔE_{CO*} and ΔE_{H*} with MAE of 0.13 eV for both adsorbates. Recently, Gu et al.[15] demonstrated that labeled site CGCNN (LS-CGCNN), which labels the binding site atoms of the unrelaxed bare surface geometry, greatly improved the prediction accuracy and resulted in MAE of 0.116 eV and 0.085 eV for ΔE_{CO*} and ΔE_{H*}, respectively. However, we note that the approaches reported to date require bulk optimizations, surface modelling and active sites identification followed by numerical transformations of the entire surface atomic structures into machine-readable inputs, implying that one needs to carry out all the processes for new materials. This could be time-consuming if one wishes to predict binding energies of multiple new materials that have not been uploaded to the database yet, necessitating a novel strategy for efficient high-throughput screening.

In this work, we report atomic structure-free representation of active motifs to predict binding energies on catalyst surfaces without performing any DFT calculations. In our method, we first identify binding site atoms and their nearest neighboring atoms positioned in the same layer and the sublayer, and prepare fingerprints using their tabulated atomic properties. This model achieved

test set MAEs of 0.120 and 0.105 eV for predicting ΔE_{CO} and ΔE_{H*}, respectively, comparable to the previously reported graph-based deep-learning models using the same dataset,[9, 15] with the elapsed time for training the model being much faster (~200 sec) compared to the deep-learning models (a few hours). As our method does not require any DFT calculations for prediction, we further apply the method to collect binding energy distributions of bimetallic catalysts by predicting binding energies of 3,780 unique active motifs. Our method successfully identified compositions of several high-performing CO_2 electroreduction catalysts reported to date, and suggest promising combinations requiring future experimental validations.

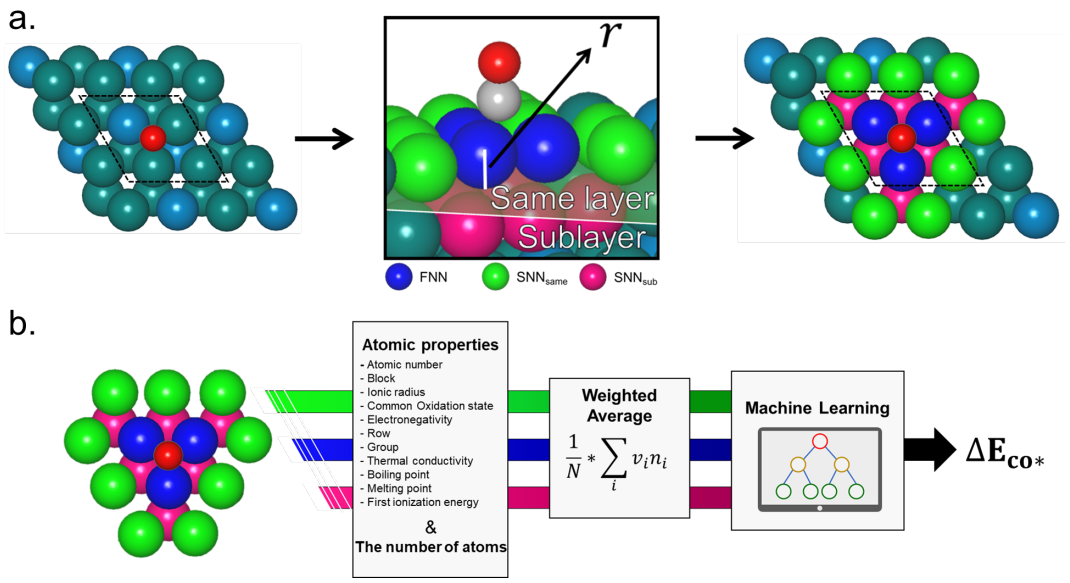


Figure 1. The overall procedure of our method. (a) Identification of FNN, SNN_{same} and SNN_{sub} atoms. (b) Preparation of fingerprints to train models. Weighted average values of 11 atomic properties and the number of atoms comprising the site for FNN, SNN_{same}, and SNN_{sub} are concatenated to construct fingerprints. For each type of the sites (FNN, SNN_{same} and SNN_{sub}), v_i, n_i and N correspond to an atomic property, the number of element i and the number of total atoms, respectively.

Compared to the previous approaches,[9, 15] which converted entire atomic structures into graph representations, we focus on active motifs including binding sites and their nearest neighbors, which correspond to the first and second NN of adsorbates, respectively (referred to as FNN and SNN henceforth). The previous results indeed suggested that contributions from the third and fourth NN to binding energies are negligible.[9] Using the optimized structures from *GASpy* database[25] (16,097 for CO* and 18,362 for H* adsorbed metal alloy surfaces, see **Supplementary Note: Data preprocessing**), we employed Voronoi algorithm implemented in pymatgen[27] to determine the FNN and SNN of adsorbates. We then converted atomic coordinates of the FNN and SNN into fractional coordinates. When a difference between the fractional *z* coordinate of SNN atoms and that of FNN is less than a fractional atomic radius (atomic radius normalized by the cell length in *z* direction) of binding site atoms, SNN atoms are assigned to be in the same layer (SNN_{same}). Otherwise, they are assigned to be in the sublayer (SNN_{sub}) (**Figure 1**). We note that fractional coordinates are necessary to correctly classify atoms in the same layer and sublayer for parallelepiped unit cells (**Figure S1**). As a result, we obtained three types of sites that could affect binding energies, i.e., FNN, SNN_{same} and SNN_{sub}. We then collected 11 tabulated atomic properties (atomic number, block, ionic radius, common oxidation state, electronegativity, row, group, thermal conductivity, boiling point, melting point, first ionization energy) from "periodic table module" implemented in pymatgen.[27] Finally, we prepared fingerprints of active motifs consisting of 36 columns in total, where we have three types of sites and each type has 12 columns consisting of weighted averages of atomic properties and the number of atoms in the type (**Figure 1**). The prepared fingerprints were then standardized using the standard scaler implemented in scikit-learn.[28] An example of fingerprint constructions can be found in **Figure S2.**

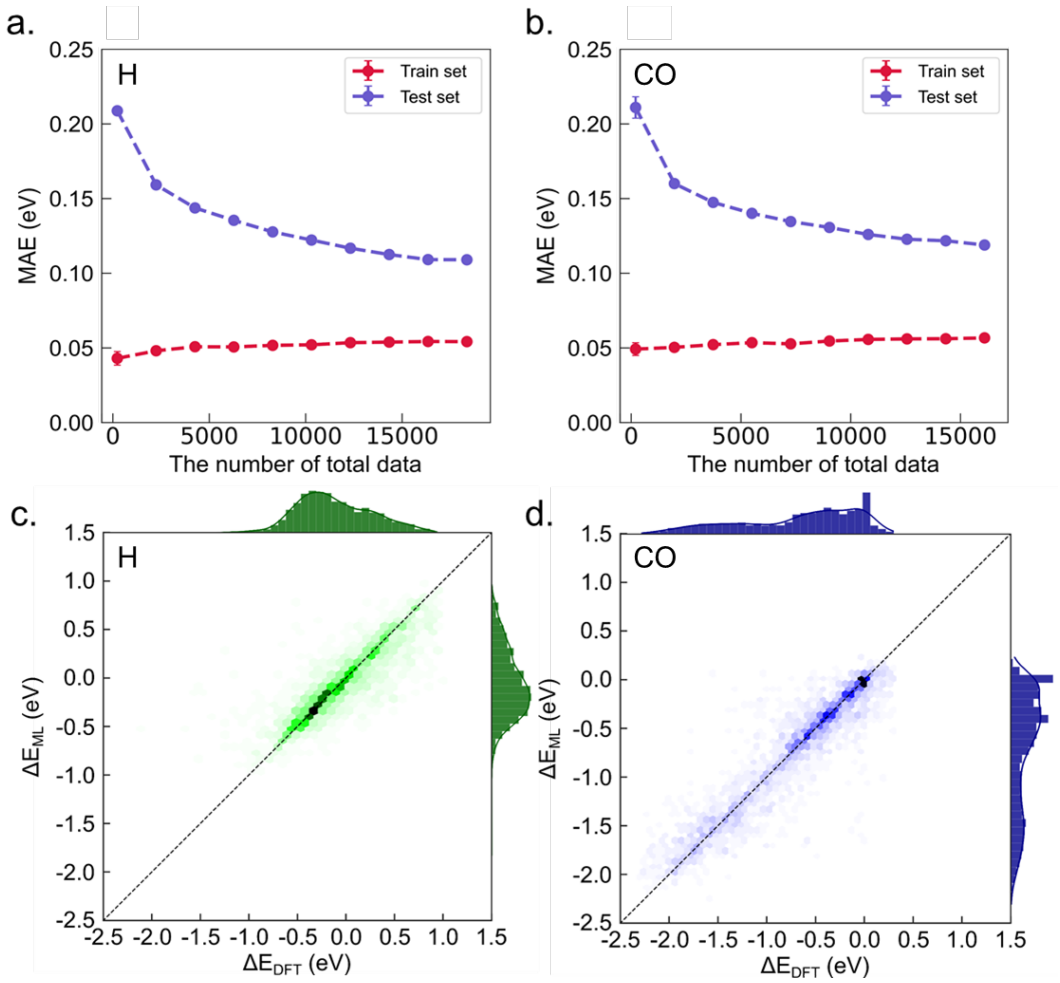


Figure 2. MAEs of test and train set with respect to the number of total data for (a) ΔE_{H*} and (b) ΔE_{CO*} predictions, where error bars refer to standard deviations of MAEs obtained from 5-fold cross validations. Two-dimensional histograms of ML predicted and DFT calculated (c) ΔE_{H*} and (d) ΔE_{CO*}.

Based on our representations, we evaluated performances of various ML algorithms implemented in scikit-learn,[28] including random forest regressor (RFR)[29], extra-trees regressor (ETR)[30], light gradient boosting regressor (LGBR)[31], gradient boosting regressor (GBR)[32], least absolute shrinkage and selection operator (LASSO)[33] and support vector machine regressor (SVR)[34], where the optimal hyperparameters were selected by Bayesian optimization[35]. We

performed 5-fold cross validations by 100 times with random splits of train (80 %) and test data (20 %) to avoid overfitting and data selection bias. In addition, we applied an ensemble approach, where a set of trained models was used to provide statistics of predictions, which was reported to further improve a model prediction accuracy.[15] Prediction performances of various ML algorithms are summarized in **Table S1**.

Generally, ensemble-based methods such as RFR, ETR and GBR predicted binding energies more accurately compared to other algorithms, with the GBR achieving the lowest MAEs for predicting ΔE_{CO*} (0.120 eV) and ΔE_{H*} (0.105 eV), comparable to the results of LS-CGCNN by Gu et al. (0.116 eV (ΔE_{CO*}) and 0.085 eV (ΔE_{H*}))[15] or better than the results of CGCNN by Back et al. (0.13 eV for both ΔE_{CO*} and ΔE_{H*}).[9] From learning curves, we observed a remarkable decrease in test set MAEs as the number of data increases up to ~10,000 total data (**Figure 2a** and **2b**). We highlight that our method achieved reasonable accuracy (MAE ~ 0.15 eV) using only 3,000 data, i.e., 16.5 % of total data, which is noticeable considering that similar MAEs were achieved using 10,000 data in the case of CGCNN by Back et al.[9] To rationalize the origin of improvements of our model's prediction accuracy, we tested the effect of not distinguishing atomic positions of SNN. In this representation, we only have two types of sites (FNN and SNN), thus SNN atoms in the same layer (SNN_{same}) and the sublayer (SNN_{sub}) are considered to be identical. MAEs of test set were found to be 0.128 eV and 0.121 eV for ΔE_{CO*} and ΔE_{H*}, respectively, highlighting the importance of considering relative positions of SNN.

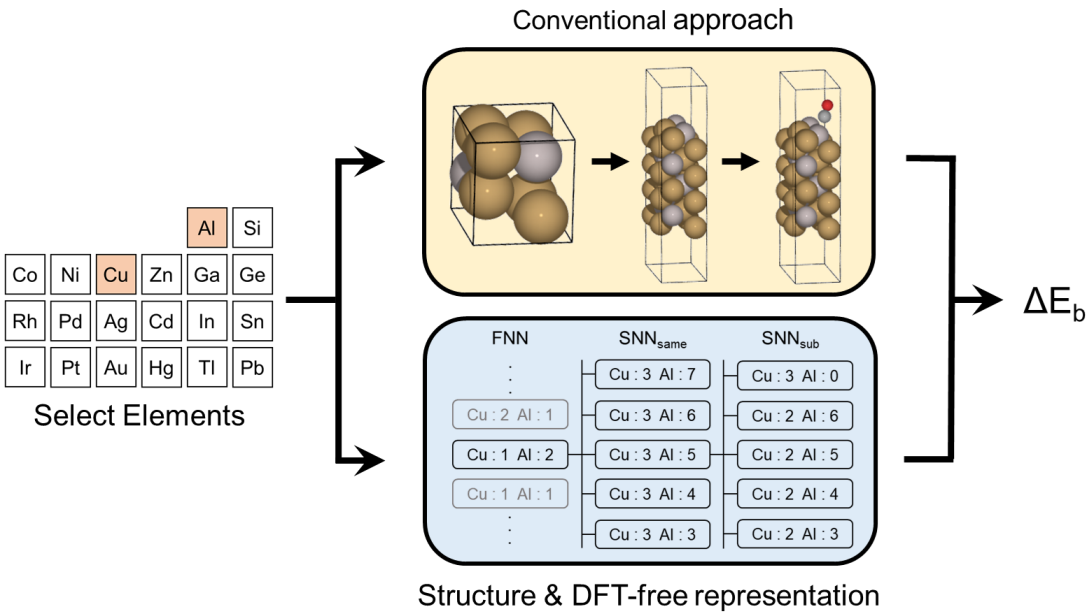


Figure 3. Two approaches to generate fingerprints for binding energy predictions. The conventional approach is to relax bulk materials by DFT calculations, build surface structures, find unique active sites and construct fingerprints. Our representation enabled new approach, where we select elements, place those elements in the FNN, SNN_{same} and SNN_{sub} to enumerate any active motifs for binding energy predictions without performing any DFT calculations.

We highlight two merits of our method compared to the previously reported graph-based deep-learning models trained using the same dataset.[9, 15] First, our method achieved comparable accuracies for ΔE_{CO*} and ΔE_{H*} predictions, considering only the FNN and SNN instead of entire atomic structures and employing simple machine learning algorithms. Consequently, the time required for training was noticeably short compared to the deep-learning models, 196.448 seconds using 1 core of Intel(R) Xeon(R) Gold 5222 CPU @ 3.80 GHz, while the deep-learning models required a few hours. Second, our method can predict ΔE_{CO*} and ΔE_{H*} of any active motifs without performing any DFT calculations. One can either (1) build atomic surface structures to generate

fingerprints of active motifs (the conventional approach) or (2) enumerate any active motifs by placing elements in the FNN, SNN_{same} and SNN_{sub} (**Figure 3**). We note that the previous models and other ML methods[9, 15, 36, 37] for binding energy predictions required DFT optimizations of bulk structures, modelling of atomic surface structures, or density of states calculations to construct fingerprints. In the following, we demonstrate a practical application of our method to accelerate catalysts discovery.

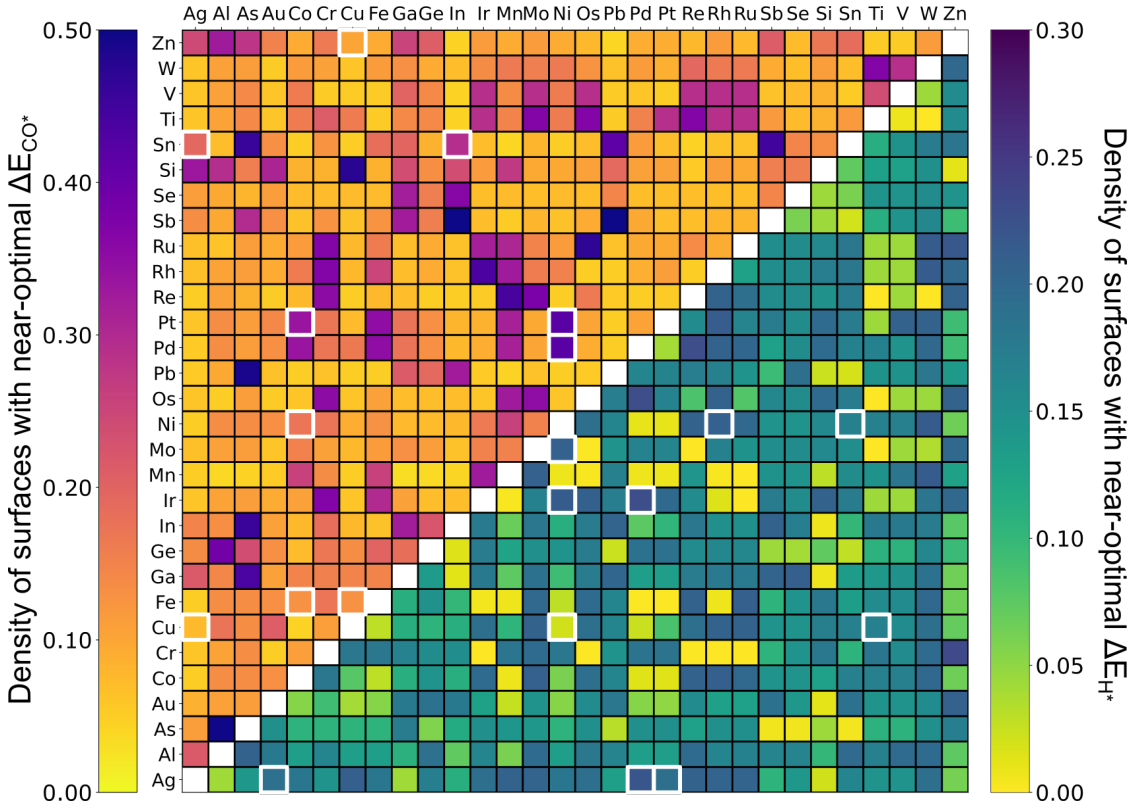


Figure 4. Heat map visualizing densities of possible surfaces of bimetallic catalysts, which have optimal ΔE_{CO*} (-0.67 eV ± 0.1 eV) and ΔE_{H*} (-0.27 eV ± 0.1 eV)[25]. White squares refer to experimentally identified bimetallic combinations for CO_2 reduction[38, 39] (upper left) and H_2 evolution (lower right)[40].

We enumerated 3,780 unique active motifs for many combinations of two elements and predicted their ΔE_{CO*} and ΔE_{H*} using the trained model (See **Figure S5** for details on the fingerprint enumeration), and their densities close to the optimal ΔE_{CO*} and ΔE_{H*} for CO_2 reduction and H_2 evolution, respectively, are plotted in the heat map (**Figure 4**). Bimetallic combinations having high density of surfaces with near-optimal binding energies can be considered as promising candidates for the CO_2 reduction and H_2 evolution. Interestingly, our method found some of experimentally observed catalysts. For the CO_2 reduction, which uses ΔE_{CO*} as an activity descriptor, Cu[41-43], Ni[44, 45], Pt[46] and In[47]-based bimetallic catalysts are commonly reported to be active. Particularly, Pt-Ni[48], Pd-Ni[49], Cu-Si[41] and Pt-Co[46] were found to have high densities of near-optimal CO binding energies. In addition, we also found that Ir-Pd[50], Ir-Ni[51, 52], Ag-Pd[53], known to be active for H_2 evolution, have high densities of near-optimal ΔE_{H*}. We expect some bimetallic combinations with high densities of the near optimal ΔE_{CO*} (e.g., In-Sb, Ag-Sn) and ΔE_H (e.g., Pd-Re, Co-Rh) can be considered as promising candidates for CO_2 reduction and for H_2 evolution, respectively, requiring further experimental validations.

In summary, we developed atomic structure-free representation for binding energy predictions. Compared to the conventional representation requiring entire atomic structures and DFT calculations, we only considered the first nearest neighbors and the second nearest neighbors in the same and sublayer. Their atomic properties and the number of atoms in the type of sites (FNN, SNN_{same}, SNN_{sub}) were used to construct fingerprints, and GBR algorithm was used for training. Our method enabled a quick training (~200 sec using CPU vs. a few hours in the case of CGCNN using GPU), and achieved accurate predictions of ΔE_{CO*} (0.120 eV MAE) and ΔE_{H*} (0.105 eV MAE) comparable to the CGCNN methods.[9, 15] Particularly, the number of data required to achieve reasonable accuracy (0.15 eV MAE) was significantly smaller than the CGCNN method.[9]

Furthermore, as our representation does not require any DFT calculations and surface structure modelling to predict binding energies of new materials, all possible active motifs can be enumerated at once and their binding energy distributions can be predicted. Given fast training/prediction with high accuracies, this approach could be used to accelerate catalyst discovery by suggesting promising elemental combinations for experimental validations.

ASSOCIATED CONTENT

Supporting Information.

The following files are available free of charge.

Tables of prediction accuracies of various algorithms and previously reported methods, an example of fingerprint generation, data preprocessing procedure, fingerprint enumeration details and graphical representation of difference between absolute and fractional coordinates.

(file type, i.e., PDF)

brief description (file type, i.e., PDF)

AUTHOR INFORMATION

Corresponding Author

Seoin Back−Department of Chemical and Biomolecular Engineering, Institute of Emergent Materials, Sogang University, Seoul 04107, Republic of Korea; Email: sback@sogang.ac.kr

Author

Dong Hyeon Mok−Department of Chemical and Biomolecular Engineering, Institute of Emergent Materials, Sogang University, Seoul 04107, Republic of Korea; Email: ahrehd0506@sogang.ac.kr

Notes

Any additional relevant notes should be placed here.

The authors declare no competing financial interests.

ACKNOWLEDGMENT

This research was supported by C1 Gas Refinery Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (2015M3D3A1A01064929) and the core KRICT project from the Korea Research Institute of Chemical Technology (SI2151-10). SB acknowledges supercomputing resources from Korea Institute of Science and Technology Information (KISTI) (KSC-2021-CRE-0060).

REFERENCES

1. Jones, G.; Bligaard, T.; Abild-Pedersen, F.; Nørskov, J. K., Using scaling relations to understand trends in the catalytic activity of transition metals. *J. Condens. Matter Phys.* **2008,** *20* (6), 064239.

2. 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* (9), 1311-1315.

3. Kulkarni, A.; Siahrostami, S.; Patel, A.; Nørskov, J. K., Understanding catalytic activity trends in the oxygen reduction reaction. *Chem. Rev.* **2018,** *118* (5), 2302-2312.

4. Wang, S.; Petzold, V.; Tripkovic, V.; Kleis, J.; Howalt, J. G.; Skulason, E.; Fernández, E.; Hvolbæk, B.; Jones, G.; Toftelund, A., Universal transition state scaling relations for (de) hydrogenation over transition metals. *Phys. Chem. Chem. Phys.* **2011,** *13* (46), 20760-20765.

5. Henkelman, G.; Uberuaga, B. P.; Jónsson, H., A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **2000,** *113* (22), 9901-9904.

6. Henkelman, G.; Jónsson, H., Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Phys.* **2000,** *113* (22), 9978-9985.

7. 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* (1), 206-217.

8. Nørskov, J. K.; Abild-Pedersen, F.; Studt, F.; Bligaard, T., Density functional theory in surface chemistry and catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **2011,** *108* (3), 937-943.

9. Back, S.; Yoon, J.; Tian, N.; Zhong, W.; Tran, K.; Ulissi, Z. W., Convolutional neural network of atomic surface structures to predict binding energies for high-throughput screening of catalysts. *J. Phys. Chem. Lett.* **2019,** *10* (15), 4401-4408.

10. Zheng, J.; Nash, J.; Xu, B.; Yan, Y., Perspective—Towards Establishing Apparent Hydrogen Binding Energy as the Descriptor for Hydrogen Oxidation/Evolution Reactions. *J. Electrochem. Soc.* **2018,** *165* (2), H27-H29.

11. 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* (7), 1159-1165.

12. Porosoff, M. D.; Kattel, S.; Li, W.; Liu, P.; Chen, J. G., Identifying trends and descriptors for selective CO2 conversion to CO over transition metal carbides. *ChemComm* **2015,** *51* (32), 6988-91.

13. Jacobsen, C. J.; Dahl, S.; Clausen, B. S.; Bahn, S.; Logadottir, A.; Nørskov, J. K., Catalyst design by interpolation in the periodic table: bimetallic ammonia synthesis catalysts. *J. Am. Chem. Soc.* **2001,** *123* (34), 8404-8405.

14. Feaster, J. T.; Shi, C.; Cave, E. R.; Hatsukade, T.; Abram, D. N.; Kuhl, K. P.; Hahn, C.; Nørskov, J. K.; Jaramillo, T. F., Understanding selectivity for the electrochemical reduction of carbon dioxide to formic acid and carbon monoxide on metal electrodes. *ACS Catal.* **2017,** *7* (7), 4822-4827.

15. Gu, G. H.; Noh, J.; Kim, S.; Back, S.; Ulissi, Z.; Jung, Y., Practical deep-learning representation for fast heterogeneous catalyst screening. *J. Phys. Chem. Lett.* **2020,** *11* (9), 3185-3191.

16. Tran, K.; Palizhati, A.; Back, S.; Ulissi, Z. W., Dynamic workflows for routine materials discovery in surface science. *J. Chem. Inf. Model.* **2018,** *58* (12), 2392-2400.

17. Back, S.; Na, J.; Ulissi, Z. W., Efficient Discovery of Active, Selective, and Stable Catalysts for Electrochemical H2O2 Synthesis through Active Motif Screening. *ACS Catal.* **2021,** *11* (5), 2483-2491.

18. 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* (1), 011002.

19. Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C., The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies. *Npj Comput. Mater.* **2015,** *1* (1), 1-15.

20. Hammer, B.; Nørskov, J. K., Electronic factors determining the reactivity of metal surfaces. *Surf. Sci.* **1995,** *343* (3), 211-220.

21. Xin, H.; Vojvodic, A.; Voss, J.; Nørskov, J. K.; Abild-Pedersen, F., Effects of d-band shape on the surface reactivity of transition-metal alloys. *Phys. Rev. B* **2014,** *89* (11), 115114.

22. Vojvodic, A.; Nørskov, J.; Abild-Pedersen, F., Electronic structure effects in transition metal surface chemistry. *Top. Catal.* **2014,** *57* (1-4), 25-32.

23. Li, Z.; Achenie, L. E.; Xin, H., An adaptive machine learning strategy for accelerating discovery of perovskite electrocatalysts. *ACS Catal.* **2020,** *10* (7), 4377-4384.

24. Calle-Vallejo, F.; Tymoczko, J.; Colic, V.; Vu, Q. H.; Pohl, M. D.; Morgenstern, K.; Loffreda, D.; Sautet, P.; Schuhmann, W.; Bandarenka, A. S., Finding optimal surface sites on heterogeneous catalysts by counting nearest neighbors. *Science* **2015,** *350* (6257), 185-189.

25. 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* (9), 696-703.

26. Xie, T.; Grossman, J. C., Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **2018,** *120* (14), 145301.

27. 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.

28. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V., Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011,** *12*, 2825-2830.

29. Breiman, L., Random forests. *Mach. Learn.* **2001,** *45* (1), 5-32.

30. Geurts, P.; Ernst, D.; Wehenkel, L., Extremely randomized trees. *Mach. Learn.* **2006,** *63* (1), 3-42.

31. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y., Lightgbm: A highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process. Syst.* **2017,** *30*, 3146-3154.

32. Friedman, J. H., Greedy function approximation: a gradient boosting machine. *Ann. Stat.* **2001**, 1189-1232.

33. Tibshirani, R., Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Methodol.* **1996,** *58* (1), 267-288.

34. Cortes, C.; Vapnik, V., Support-vector networks. *Mach. Learn.* **1995,** *20* (3), 273-297.

35. Eggensperger, K.; Feurer, M.; Hutter, F.; Bergstra, J.; Snoek, J.; Hoos, H.; Leyton-Brown, K. In *Towards an empirical foundation for assessing bayesian optimization of hyperparameters*, NIPS workshop on Bayesian Optimization in Theory and Practice, 2013; p 3.

36. Ma, X.; Li, Z.; Achenie, L. E.; Xin, H., Machine-learning-augmented chemisorption model for CO2 electroreduction catalyst screening. *J. Phys. Chem. Lett.* **2015,** *6* (18), 3528-3533.

37. 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* (23), 5152-5159.

38. Gnanamani, M. K.; Jacobs, G.; Hamdeh, H. H.; Shafer, W. D.; Liu, F.; Hopps, S. D.; Thomas, G. A.; Davis, B. H., Hydrogenation of carbon dioxide over Co–Fe bimetallic catalysts. *ACS Catal.* **2016,** *6* (2), 913-927.

39. Luc, W.; Collins, C.; Wang, S.; Xin, H.; He, K.; Kang, Y.; Jiao, F., Ag–Sn bimetallic catalyst with a core–shell structure for CO2 reduction. *J. Am. Chem. Soc.* . **2017,** *139* (5), 1885-1893.

40. Lu, Q.; Hutchings, G. S.; Yu, W.; Zhou, Y.; Forest, R. V.; Tao, R.; Rosen, J.; Yonemoto, B. T.; Cao, Z.; Zheng, H., Highly porous non-precious bimetallic electrocatalysts for efficient hydrogen evolution. *Nat. Commun.* **2015,** *6* (1), 1-8.

41. Birhanu, M. K.; Tsai, M. C.; Kahsay, A. W.; Chen, C. T.; Zeleke, T. S.; Ibrahim, K. B.; Huang, C. J.; Su, W. N.; Hwang, B. J., Copper and Copper-Based Bimetallic Catalysts for Carbon Dioxide Electroreduction. *Adv. Mater. Interfaces* **2018,** *5* (24), 1800919.

42. Nie, X.; Wang, H.; Janik, M. J.; Chen, Y.; Guo, X.; Song, C., Mechanistic insight into C–C coupling over Fe–Cu bimetallic catalysts in CO_2 hydrogenation. *J. Phys. Chem. C* **2017,** *121* (24), 13164-13174.

43. Su, X.; Sun, Y.; Jin, L.; Zhang, L.; Yang, Y.; Kerns, P.; Liu, B.; Li, S.; He, J., Hierarchically porous Cu/Zn bimetallic catalysts for highly selective CO_2 electroreduction to liquid C2 products. *Appl. Catal. B* **2020,** *269*, 118800.

44. Bian, Z.; Das, S.; Wai, M. H.; Hongmanorom, P.; Kawi, S., A review on bimetallic nickel-based catalysts for CO_2 reforming of methane. *ChemPhysChem* **2017,** *18* (22), 3117-3134.

45. Takanabe, K.; Nagaoka, K.; Nariai, K.; Aika, K.-i., Titania-supported cobalt and nickel bimetallic catalysts for carbon dioxide reforming of methane. *J. Catal.* **2005,** *232* (2), 268-275.

46. Porosoff, M. D.; Chen, J. G., Trends in the catalytic reduction of CO_2 by hydrogen over supported monometallic and bimetallic catalysts. *J. Catal.* **2013,** *301*, 30-37.

47. Li, J.; Zhu, M.; Han, Y. F., Recent Advances in Electrochemical CO_2 Reduction on Indium-Based Catalysts. *ChemCatChem* **2021,** *13* (2), 514-531.

48. Liu, D.; Li, Y.; Kottwitz, M.; Yan, B.; Yao, S.; Gamalski, A.; Grolimund, D.; Safonova, O. V.; Nachtegaal, M.; Chen, J. G., Identifying dynamic structural changes of active sites in Pt–Ni bimetallic catalysts using multimodal approaches. *ACS Catal.* **2018,** *8* (5), 4120-4131.

49. Lee, J. H.; Kattel, S.; Jiang, Z.; Xie, Z.; Yao, S.; Tackett, B. M.; Xu, W.; Marinkovic, N. S.; Chen, J. G., Tuning the activity and selectivity of electroreduction of CO_2 to synthesis gas using bimetallic catalysts. *Nat. Commun.* **2019,** *10* (1), 1-8.

50. Şen, B.; Aygün, A.; Şavk, A.; Akocak, S.; Şen, F., Bimetallic palladium–iridium alloy nanoparticles as highly efficient and stable catalyst for the hydrogen evolution reaction. *Int. J. Hydrog. Energy* **2018,** *43* (44), 20183-20191.

51. De, S.; Zhang, J.; Luque, R.; Yan, N., Ni-based bimetallic heterogeneous catalysts for energy and environmental applications. *Energy Environ. Sci.* **2016,** *9* (11), 3314-3347.

52. Papaderakis, A.; Pliatsikas, N.; Patsalas, P.; Tsiplakides, D.; Balomenou, S.; Touni, A.; Sotiropoulos, S., Hydrogen evolution at Ir-Ni bimetallic deposits prepared by galvanic replacement. *J. Electroanal. Chem.* **2018,** *808*, 21-27.

53. Nazir, R.; Fageria, P.; Basu, M.; Pande, S., Decoration of carbon nitride surface with bimetallic nanoparticles (Ag/Pt, Ag/Pd, and Ag/Au) via galvanic exchange for hydrogen evolution reaction. *J. Phys. Chem. C* **2017,** *121* (36), 19548-19558.