

Full length article

Distance-insensitive Graph Neural Networks and multi-task learning for accurate prediction of adsorption energies on alloy nanoclusters

Koki Otsuka ^a, Anh Khoa Augustin Lu ^{a,b}, Koji Shimizu ^c, Satoshi Watanabe ^a ^{*}^a Department of Materials Engineering, The University of Tokyo, 113-8656, Tokyo, Japan^b Research Center for Materials Nanoarchitectonics, National Institute for Materials Science, 305-0044, Tsukuba, Japan^c Materials DX Research Center, National Institute of Advanced Industrial Science and Technology, 305-8568, Tsukuba, Japan

ARTICLE INFO

Dataset link: [Dataset for Distance-Insensitive Graph Neural Networks and Multi-Task Learning for Accurate Prediction of Adsorption Energies on Alloy Nanoclusters \(Original data\)](#)

Keywords:

Adsorption energy prediction
Graph neural networks
Multi-task learning
Alloy nanoclusters
Catalyst screening

ABSTRACT

Developing methodologies for accurate yet computationally efficient prediction of adsorption energies is important for accelerating alloy catalyst screening. However, current data-driven approaches still face challenges when applied to finite alloy nanoclusters, where domain-matched adsorption-energy datasets remain limited and robust screening methods that do not strongly rely on fully relaxed geometries are desirable. To address these issues, we developed a distance-insensitive, adsorbate-oriented orbital graph neural network model (BE-OGCNN) that incorporates orbital interaction features without using explicit interatomic distance or angle descriptors. On a 44-atom alloy-cluster benchmark, the proposed model achieved a mean absolute error of 0.042 eV, outperforming the other distance-insensitive architectures examined here and achieving accuracy comparable to distance-sensitive reference models without using explicit geometric descriptors. We further examined whether bulk-derived auxiliary labels can improve adsorption-energy prediction through multi-task learning with d-band center and total energy auxiliary tasks. The multi-task approach yielded consistent improvement for larger 85-atom clusters, suggesting that physically motivated auxiliary tasks may improve data efficiency in adsorption-energy prediction for alloy nanoparticles.

1. Introduction

The design of new high-performance catalysts is paramount for addressing global challenges towards the realization of clean energy, sustainable chemical synthesis, and environmental remediation [1,2]. Rational catalyst design, guided by computational methods, has emerged as a powerful strategy to accelerate this process, moving beyond traditional trial-and-error experimentation [3,4]. Among these methods, density functional theory (DFT) has become an indispensable tool for calculating the electronic structure of materials and predicting key catalytic properties, such as the adsorption energies of reaction intermediates, which often serve as crucial descriptors for catalytic activity [5]. However, the predictive power of DFT comes at a significant computational cost, which becomes prohibitively expensive when screening a vast number of candidate materials. This limitation is particularly severe for nanoparticle catalysts, which contain diverse low-coordination sites and finite-size effects that are not fully represented by ideal flat surfaces [6,7].

To overcome the computational bottleneck of DFT, machine learning (ML) has emerged as a promising alternative for rapidly predicting material properties [8]. The relevance of ML-based surrogate modeling

extends across a wide range of energy-materials problems; recent reviews have highlighted AI applications involving interfacial energetics and transport, including CO₂ reduction electrocatalysts and battery materials where ionic transport and electrolyte stability govern performance [9]. In heterogeneous catalysis, recent reviews have summarized the rapid development of ML models for predicting adsorption energies on surfaces and nanostructured catalysts [10]. Graph Neural Networks (GNNs) are especially suitable for atomistic systems because they can represent a material as a graph and learn structure–property relationships directly from atomic environments [11]. The pioneering Crystal Graph Convolutional Neural Network (CGCNN) demonstrated the feasibility of this approach for crystalline materials [12]. Subsequent models have expanded both accuracy and applicability: the Bond-type Embedded CGCNN (BE-CGCNN) introduced bond-type information for adsorption-energy prediction on catalytic nanoparticles [13], the improved CGCNN (i-CGCNN) incorporated higher-order local-environment information [14], and the Orbital Graph CNN (OGCNN) introduced orbital-interaction features to provide a more chemically informed description of bonding [15].

^{*} Corresponding author.E-mail address: watanabe@cello.t.u-tokyo.ac.jp (S. Watanabe).

Recent progress has further broadened the scope of atomistic ML beyond single-property GNNs. Multi-task learning (MTL) has been increasingly used to exploit correlations among related material properties. Recent examples include AEGCNN-MTL for simultaneous prediction of band gaps and work functions [16], surface-emphasized MTL-CGCNN (SEM-CGCNN) for surface energies and work functions of intermetallics [17], and other MTL frameworks for material property prediction [18,19]. In parallel, large-scale training and pretraining have rapidly advanced atomistic ML. The Open Catalyst Project, including OC20 and OC22, established large benchmark datasets for surface-adsorbate and oxide electrocatalyst systems [20,21]. More recently, large atomic models such as DPA-2 have shown the value of multi-task pretraining across heterogeneous atomistic datasets [22], while catalyst-oriented models such as element-based machine learning potentials (EMLP) and multimodal language-graph learning have explored large-scale or pretrained representations for heterogeneous catalysis and adsorption configurations [23,24].

These advances provide important context for the present work, but they also highlight two remaining issues for applying ML to alloy nanoparticle catalyst screening. First, many state-of-the-art atomistic models rely explicitly on interatomic distances, angles, forces, or fully relaxed three-dimensional coordinates. This is advantageous for accurate potential-energy-surface modeling, but it can be less convenient for early-stage screening when relaxed structures are not yet available for all candidate nanoparticles. Previous work has also recognized this issue; for example, Back et al. developed a convolutional neural-network model to predict CO and H binding energies from atomic surface structures using initial adsorption configurations as inputs [25]. Second, although large adsorption-energy datasets such as OC20 and OC22 are invaluable, their coverage does not fully match the target domain considered here. They are primarily based on periodic slab models and therefore do not directly represent finite nanoparticles with edges, vertices, curved surfaces, and size-dependent electronic structures. In addition, their catalyst compositions are mostly limited to relatively simple elemental combinations and do not systematically cover highly multicomponent alloy nanoparticles such as high-entropy-alloy-like catalysts. Expanding adsorption-energy datasets to such multicomponent nanoparticle surfaces would require a very large number of costly surface relaxation and adsorption calculations.

One practical way to mitigate this data limitation is to exploit information from more readily available materials datasets. Bulk crystal calculations are generally more standardized and less costly than exhaustive adsorption calculations over many nanoparticle surface sites, and large bulk datasets already cover broader composition spaces than current adsorption-energy datasets. For example, the Alexandria dataset contains diverse transition-metal alloy structures, including highly multicomponent systems with five or more elements. In a related direction, transfer learning has already been shown to reduce data requirements within adsorption-related tasks. Zhang et al. first trained a neural-network model with local-environment descriptors using Li-polysulfide binding energies on MoSe_2 , and then fine-tuned the model for the structurally related WSe_2 host using substantially fewer DFT training data [26]. This study motivates a further question: whether information learned from bulk structures can also be transferred to adsorption-energy prediction. If successful, such bulk-to-adsorption transfer would provide a practical route to improve label efficiency for catalyst nanoparticle screening without requiring adsorption-energy calculations for every auxiliary structure.

In this work, we therefore focus on a complementary geometry-light approach for adsorption-energy prediction on alloy nanoclusters together with auxiliary learning from bulk data. First, we enhance the BE-CGCNN backbone by incorporating architectural concepts inspired by i-CGCNN and OGCNN. Through a systematic comparison, we identify the Bond-type Embedded Orbital Graph Convolutional Neural Network (BE-OGCNN) as the most effective model among

the distance-insensitive architectures examined here for representing surface-adsorbate interactions on alloy clusters.

Second, we introduce an MTL framework designed to transfer information from bulk transition-metal data to the target adsorption-energy task. The model is trained on the primary task of H adsorption-energy prediction together with auxiliary tasks for d-band center and total energy. The d-band center is a well-established electronic descriptor related to chemisorption on transition-metal surfaces [27]. More broadly, electronic-structure information from the density of states has been shown to provide useful features for machine-learning prediction of adsorption energies [28]. Total energy provides complementary information on structural stability and chemical environments. In this way, the present work examines whether bulk-derived auxiliary labels can improve the data efficiency and transferability of a distance-insensitive adsorption-energy model.

Overall, this work does not aim to replace distance-sensitive universal potentials or large pretrained catalyst models. Rather, it complements them by examining how far a geometry-light, distance-insensitive GNN can be pushed for adsorption-energy prediction on alloy nanoclusters, and how auxiliary bulk-property learning can help exploit broader and more diverse materials data for this target task.

2. Method

2.1. Datasets

The dataset used in this study consists of three main components: an in-house cluster dataset created to train and evaluate the GNN backbone for adsorption energy prediction, and two public datasets, OpenCatalyst 20 (OC20) [20] and Alexandria [29], which were used for the multitask learning framework.

2.1.1. In-house cluster dataset

We constructed an in-house cluster dataset based on first-principles calculations. The computational methodology is described below. It is mainly based on the methods described in the previous work by Miyamoto and coworkers [30].

Density Functional Theory (DFT) calculations were performed using the Vienna Ab initio Simulation Package (VASP) [31]. The Perdew–Burke–Ernzerhof (PBE) [32] exchange–correlation functional was adopted. The projector-augmented-wave (PAW) [33] method was used to represent the interaction of valence electrons with inner-shell electrons and nuclei. A plane-wave energy cut-off of 500 eV was used. Brillouin-zone integrations were performed using Monkhorst–Pack k-point sampling [34]: $3 \times 3 \times 3$ grids were used for 19- and 44-atom cluster models, and a $1 \times 1 \times 1$ grid was used for 85-atom cluster models. A vacuum spacing of at least 15 Å was introduced to avoid interactions between periodic images. The electronic convergence criterion was 10^{-4} eV, and the atomic configurations were relaxed until the atomic forces were below 0.01 eV/Å.

This dataset consists of hydrogen (H) adsorption energies (ΔE_{H}) on octahedral transition-metal cluster models of varying sizes. Nine catalytically relevant transition metal elements were selected for this study: Ag, Au, Co, Cu, Ir, Ni, Pd, Pt, and Rh. These elements were chosen because they are representative late transition metals that are widely studied as electrocatalysts and heterogeneous catalysts, including for hydrogen-related reactions such as the hydrogen evolution reaction. In addition, their partially filled or nearly filled d bands make them suitable for analysis within the d-band framework, which is central to the auxiliary-task design used in this work. Across the three cluster sizes, unary, binary, and ternary compositions were generated from these elements as described below. The octahedral cluster motif was adopted because it contains characteristic nanoscale adsorption environments, such as terrace-, edge-, and vertex-like sites, within a simple and well-defined geometry. The three cluster sizes were selected to examine different roles in model construction and validation:

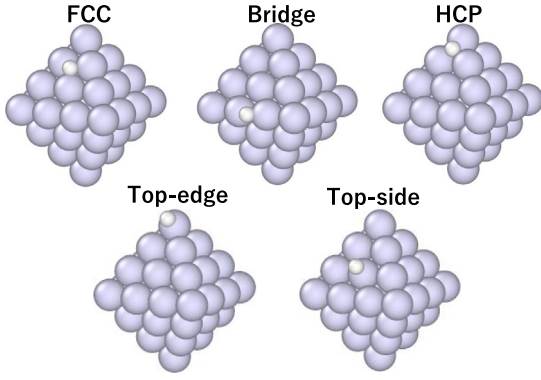


Fig. 1. Sample structures from the in-house cluster dataset (44 atoms). Small white spheres denote adsorbed H atoms.

Table 1

Composition of the in-house cluster dataset.

Number of atoms	19	44	85
Elements	1 ~ 3	1 ~ 2	1 ~ 3
Data	3820	4254	115

19-atom clusters provide compact structures with a wide range of low-coordination sites, 44-atom clusters provide the main benchmark size for evaluating adsorption-energy prediction accuracy, and 85-atom clusters (approximately 1.7 nm in diameter) represent larger nanoparticles that are closer to experimentally accessible nanoparticle sizes and were therefore used to test size transferability. Specifically, as listed in Table 1, the dataset includes 19-atom models composed of up to 3 elements selected from the nine candidates (3820 data points), 44-atom models composed of 1 to 2 elements (4254 data points), and 85-atom models composed of 1 to 3 elements (115 data points). Ternary 44-atom clusters were not included because systematically extending the main 44-atom benchmark to ternary compositions would greatly increase the number of elemental arrangements, adsorption sites, and associated DFT relaxations. We therefore limited this benchmark to unary and binary clusters to keep dataset generation computationally tractable; the limited ternary subsets at 19 and 85 atoms were used for complementary composition and size coverage rather than as an exhaustive benchmark. For each cluster model, as shown in Fig. 1, various high-symmetry sites (including atop, bridge, and hollow sites) were considered for H adsorption to capture a diverse range of adsorption environments. In this study, we specifically focused on the hydrogen adsorption energy (ΔE_H) as the primary prediction target, because hydrogen is one of the most fundamental adsorbates and plays a crucial role in various catalytic reactions.

The H adsorption energy (ΔE_H) was calculated as:

$$\Delta E_H = E_{H*} - E_* - \frac{1}{2}E_{H_2} \quad (1)$$

where E_{H*} , E_* , and E_{H_2} denote the total energies of the H-adsorbed cluster, the clean cluster, and an isolated H_2 molecule, respectively.

2.1.2. Public datasets for multi-task learning

Data for the MTL framework was sourced from existing large-scale public databases. As the d-band center is primarily applicable to transition metals, we filtered the public datasets. Specifically, the collected structures consist exclusively of pure transition metals and their alloys; no oxides or structures containing non-metallic elements are included. To avoid introducing an element-selection bias towards the nine elements used in the in-house cluster dataset, the OC20 subset was not restricted to these nine elements. Instead, we extracted H-adsorption structures composed only of transition metals excluding Group 12 elements, resulting in coverage of 37 transition-metal elements. Group

12 elements (e.g., Zn, Cd, and Hg) were explicitly excluded because their d bands are fully occupied and lie deep below the Fermi level, rendering the classical d-band center theory inapplicable for describing their physical and catalytic properties. The details of public databases used in this study are as follows.

- **OpenCatalyst 20 (OC20) Dataset:** We extracted 1316 data points from the OC20 database. This subset consists of H-adsorption structures whose catalyst components contain only transition metals excluding Group 12 elements, covering 37 transition-metal elements in total. The OC20 calculations were performed using VASP with PAW potentials, the RPBE functional, a 350 eV plane-wave cutoff, and Monkhorst-Pack k-point meshes, without spin polarization. Structural relaxation was continued until the atomic forces were below 0.03 eV/Å, with an electronic convergence criterion of 10^{-4} eV [20]. For the selected structures, the H adsorption energies were recalculated from the energies of the H-adsorbed and corresponding clean slabs together with the H reference energy used in the preprocessing procedure. The adsorption energies were used for the primary task, while the corresponding total energies, converted to eV per atom, were used for the total-energy auxiliary task.
- **Alexandria Dataset:** Bulk crystal structures containing only transition metals were extracted from the PBE release of the Alexandria dataset dated July 2, 2025. Total energies were converted to eV per atom before use as auxiliary labels. Alexandria was selected because it provides a large and compositionally diverse collection of DFT-calculated bulk structures, including multicomponent transition-metal alloys beyond simple unary, binary, and ternary systems. According to the published Alexandria protocol, geometry optimizations and total-energy calculations were performed using VASP with the PBE functional, PAW potentials, a 520 eV plane-wave cutoff, and uniform Γ -centered k-point grids with a density of 1000 k-points per reciprocal atom; structures were relaxed until the forces were below 0.005 eV/Å [29]. Site-resolved d-band centers were extracted from the archived `vasprun.xml` files using the `CompleteDos.get_band_center` method in `pymatgen` with `sites=[site]` and `band=OrbitalType.d`, and no structure-level averaging was applied. The calculation included the full d-projected DOS, comprising both occupied and unoccupied states, used the Fermi level as the zero-energy reference, and integrated over the full energy range available in the DOS data. Structures for which the required output files were missing or could not be processed consistently were excluded from the dataset. The availability of both total-energy labels and post-processable electronic-structure information was a key reason for using Alexandria as the source of bulk auxiliary data [35].

The three datasets were therefore generated using different DFT protocols, most notably PBE for the in-house and Alexandria datasets and RPBE for OC20, as well as different plane-wave cutoffs and convergence settings. Total-energy labels were converted consistently to eV per atom before training. Nevertheless, residual differences arising from the underlying DFT protocols may affect learning from the combined datasets and constitute a limitation of the present MTL setup.

2.2. Model architectures

Our work is founded upon the BE-CGCNN framework [13]. A defining characteristic of this framework, which we retain and extend in all our models, is its deliberately distance- and angle-insensitive design. This design is intended for geometry-light catalyst screening, where fully relaxed atomic coordinates may not be available for every candidate structure at the initial screening stage. While distance-sensitive models can achieve high accuracy when well-defined geometries are

available, their applicability in early-stage screening depends on how initial structures are constructed. In practice, such geometries are not uniquely defined and may vary across datasets, making consistent input definition challenging. In this context, distance-insensitive representations provide a more robust alternative for integrating heterogeneous data sources. In this study, “distance-insensitive” means that explicit interatomic distance and bond-angle descriptors are not used as input features or in the graph convolution operations. Instead, the model uses atom features, nearest-neighbor connectivity, and categorical bond-type features after graph construction. Specifically, the framework categorizes the interactions between connected atoms into four distinct bonding types (e.g., metallic, covalent, chemisorption, and non-bonded interactions). These categorical bond types are explicitly encoded as one-hot vectors to formulate the bond features, denoted as $u_{(i,j)}^k$. Because $u_{(i,j)}^k$ depends only on connectivity and bond category, the model does not require explicit geometric descriptors after the connectivity graph is defined. Therefore, small coordinate changes that preserve the nearest-neighbor graph do not change the input representation of BE-CGCNN-based models, which is advantageous for screening based on approximate structures. This should be distinguished from a fully coordinate-independent representation: atomic coordinates are still used to construct the initial nearest-neighbor graph, and model predictions may be affected if structural relaxation changes the bonding connectivity.

This distance-insensitive formulation may offer a practical computational advantage because it avoids explicit radial and angular basis expansions and equivariant tensor operations. However, the present study was not designed as a controlled computational-cost benchmark, and we therefore do not claim that the proposed model is universally faster or less computationally demanding than the distance-sensitive reference models. The actual computational cost depends on the model configuration, graph size, hardware, and screening workload. Moreover, the extended variants are not intended to minimize the parameter count relative to the original BE-CGCNN: BE-ICGCNN introduces additional bond-update operations, while BE-OGCNN includes orbital-interaction descriptors and an encoder-decoder module. These components increase model capacity while preserving the distance-insensitive input representation.

While the original BE-CGCNN was developed for single-element Pt nanoparticles, our study focuses on predicting adsorption energies on multi-element alloy clusters or slabs. To accommodate this, we first established a generalized criterion for graph construction. Specifically, two atoms are considered nearest neighbors, and thus connected by an edge in the graph, if their interatomic distance is less than or equal to the sum of their respective covalent radii. This chemically intuitive method allows for robust graph construction for clusters with varying elemental compositions, which was confirmed to be effective for prediction on alloy catalysts. In practice, the determination of nearest neighbors was performed using the Atomic Simulation Environment (ASE) [36].

In this work, we developed two BE-CGCNN-based variants by incorporating architectural concepts from i-CGCNN and OGCNN: BE-ICGCNN and BE-OGCNN, respectively.

2.2.1. Model incorporating i-CGCNN: BE-ICGCNN

We developed BE-ICGCNN by incorporating two key mechanisms from the i-CGCNN [14] into our BE-CGCNN framework. Importantly, these extensions were implemented using only topological information to preserve the distance-insensitive nature of the model.

Firstly, we implemented a dynamic bond update mechanism. Instead of relying on fixed bond embeddings determined solely by bond type, the bond features $u_{(i,j)}^k$ in the BE-ICGCNN are updated in each convolutional layer (t). This is performed by incorporating the current feature vectors of the atoms $v_i^{(t)}$ and $v_j^{(t)}$ at both ends of the bond into the bond vector. Specifically, we define a concatenated vector $z_{(i,j)}^{(t)}$:

$$z_{(i,j)}^{(t)} = v_i^{(t)} \oplus v_j^{(t)} \oplus u_{(i,j)}^{(t)} \quad (2)$$

where $\oplus$ denotes concatenation. The bond feature is then updated as follows:

$$u_{(i,j)}^{(t+1)} = u_{(i,j)}^{(t)} + \sigma \left(z_{(i,j)}^{(t)} W_{ef}^{(t)} + b_{ef}^{(t)} \right) \odot g \left(z_{(i,j)}^{(t)} W_{es}^{(t)} + b_{es}^{(t)} \right) \quad (3)$$

where σ is the sigmoid function, g is the softplus function, $\odot$ is the element-wise multiplication, and W and b are the learnable weights and biases. This iterative update process allows the model to optimize the bond representation as it passes through the layers, enabling it to reflect the specific chemical bonding states derived from the elemental characteristics more accurately.

Secondly, we introduced explicit three-body correlations. While the standard BE-CGCNN updates atomic features by aggregating pairwise interactions between a central atom and its neighbors, our BE-ICGCNN additionally considers interactions involving the central atom and pairs of neighboring atoms. This is achieved by defining a three-body concatenated vector $z_{(i,j,l)}^{(t)}$ for a central atom i and two distinct neighbors j and l connected by bonds k and k' :

$$z_{(i,j,l)}^{(t)} = v_i^{(t)} \oplus v_j^{(t)} \oplus v_l^{(t)} \oplus u_{(i,j)}^{(t)} \oplus u_{(i,l)}^{(t)} \quad (4)$$

The node features are then updated by aggregating both pairwise and three-body interaction terms:

$$v_i^{(t+1)} = v_i^{(t)} + \sum_j \sigma \left(z_{(i,j)}^{(t)} W_f^{(t)} + b_f^{(t)} \right) \odot g \left(z_{(i,j)}^{(t)} W_s^{(t)} + b_s^{(t)} \right) + \sum_{j,l} \sigma \left(z_{(i,j,l)}^{(t)} W_{3f}^{(t)} + b_{3f}^{(t)} \right) \odot g \left(z_{(i,j,l)}^{(t)} W_{3s}^{(t)} + b_{3s}^{(t)} \right) \quad (5)$$

By aggregating these three-body terms for all neighboring pairs, the model can capture non-additive effects and complex local environments that pairwise interactions alone cannot resolve.

2.2.2. Model incorporating OGCNN: BE-OGCNN

In this model, we adopted two central ideas of the Orbital Graph Convolutional Neural Network (OGCNN) [15]. First, the outer product of valence electron orbital vectors was used as a rich feature descriptor that explicitly encodes orbital-orbital interactions. Second, we incorporated the encoder-decoder network architecture employed in the OGCNN. This module, designed as a multilayer perceptron, enables the model to effectively select and learn the most significant latent features from the high-dimensional input space combining basic atomic attributes and orbital interaction data.

However, our implementation diverges from the original OGCNN methodology in a critical aspect regarding feature aggregation. In the original OGCNN, the contribution of each neighbor atom's orbital features is weighted based on the solid angle of its corresponding Voronoi polyhedron. By contrast, we replaced this Voronoi-based weighting with a weighting factor defined as the inverse of the coordination number ($1/CN$) for each neighbor to maintain our distance-insensitive architecture while accounting for variations in local atomic density.

Specifically, the orbital feature matrix X^c for a central atom c is computed by aggregating the orbital vectors $\vec{O}^c$ of the central atom and $\vec{O}^n$ of its CN_c neighboring atoms as follows:

$$X^c = \vec{O}^c \otimes \left(\sum_{n=1}^{CN_c} \frac{1}{CN_n} \vec{O}^n \right) \quad (6)$$

where $\otimes$ denotes the outer product. This modification allows us to normalize the contribution of orbital interactions without requiring precise, relaxation-dependent geometry information.

To integrate the orbital information with the node features, the two-dimensional orbital interaction matrix X^i is first reshaped into a one-dimensional vector. Subsequently, this flattened orbital feature matrix, denoted as $\text{Flatten}(X^i)$, is concatenated with the basic atomic attributes v_{basic}^i and passed through the encoder-decoder multilayer perceptron network. The encoder first extracts the most significant latent features h_i from the high-dimensional input:

$$h_i = \text{MLP}_{\text{encoder}} \left(v_{\text{basic}}^i \oplus \text{Flatten}(X^i) \right) \quad (7)$$

The decoder then projects this latent representation to the desired hidden dimension to generate the initial node feature $v_i^{(0)}$:

$$v_i^{(0)} = \text{MLP}_{\text{decoder}}(h_i) \quad (8)$$

This enriched initial representation is then utilized in the subsequent graph convolutional layers.

2.3. Multi-task learning setup

We implemented a hard parameter sharing MTL framework [18] to enhance the data efficiency and generalization capability of the best-performing GNN backbone. The primary task is the prediction of H adsorption energy (ΔE_{H}) on surfaces and nanoclusters. The auxiliary tasks are the prediction of the d-band center (ϵ_d) and the total energy (E_{Total}). These auxiliary tasks were selected because they are physically related to adsorption energetics: the d-band center reflects the electronic structure of transition-metal surfaces relevant to chemisorption, while the total energy provides information on the stability and chemical environment of the underlying structure. Although atomic forces are also important quantities and could be useful auxiliary targets, force supervision was not adopted in this work because the present adsorption-energy task is evaluated for relaxed structures, where residual forces are not the primary target information. The total energy prediction task was applied for all of the structures, while d-band center prediction task was applied for only structures from Alexandria. The model shared the same feature extraction layers (GNN layers) across all tasks, corresponding to hard parameter sharing, while only the final output layers (MLP layers) are specific to each task. The total loss function is a weighted sum of the normalized individual task losses, defined as

$$\mathcal{L}_{\text{total}} = w_{\Delta E_{\text{H}}} \frac{\text{MAE}_{\Delta E_{\text{H}}}}{s_{\Delta E_{\text{H}}}} + w_{\epsilon_d} \frac{\text{MAE}_{\epsilon_d}}{s_{\epsilon_d}} + w_{E_{\text{Total}}} \frac{\text{MAE}_{E_{\text{Total}}}}{s_{E_{\text{Total}}}}, \quad (9)$$

where MAE_t is the mean absolute error for task t and s_t is the standard deviation of the corresponding target values in the training data. This normalization makes the loss terms dimensionless before applying task weights. Therefore, the equal weights used in this work ($w_{\Delta E_{\text{H}}} = w_{\epsilon_d} = w_{E_{\text{Total}}} = 1$) correspond to equal weighting of normalized losses, not raw losses with different physical units. Task-weight optimization may further improve performance, but was not performed here to keep the setting simple and reproducible.

3. Results and discussion

3.1. Comparison of GNN architectures

We first compared the performance of the original BE-CGCNN and two variants inspired by i-CGCNN and OGCNN for adsorption-energy prediction on alloy clusters. In addition, we evaluated two distance-sensitive reference architectures: Allegro [37], a state-of-the-art equivariant GNN model that explicitly uses geometric information, and ALIGNN [38], which incorporates bond distances and angles through atomistic line-graph message passing. For the predictions presented in this section, we utilized DFT-relaxed structures as inputs for all models to enable a fair comparison of model performance under consistent conditions. For the 44-atom cluster dataset used here, structures that collapsed during DFT relaxation were excluded during dataset construction. We further confirmed that the nearest-neighbor connectivity graphs of all remaining 44-atom structures were identical before and after DFT relaxation. Therefore, for the BE-CGCNN-based distance-insensitive models, which use connectivity and bond-type information rather than explicit geometric descriptors, the choice between initial and relaxed coordinates is expected to have little effect on this benchmark. Nevertheless, for more complex structures where relaxation can substantially change bonding connectivity, predictions from the present distance-insensitive models should be interpreted with care.

All models were trained, validated, and tested on our 44-atom cluster dataset. The dataset was randomly partitioned into training, validation, and test sets, and the exact split files are publicly available at https://github.com/modelna/BE_OGCNN. The split ratio was 8:1:1. Because of this random split, chemically similar compositions, elemental arrangements, or adsorption motifs may appear across different subsets. Therefore, the present 44-atom benchmark should be interpreted mainly as an interpolation test within the sampled chemical and adsorption-site space, rather than as a strict composition- or motif-disjoint extrapolation test. However, the same training, validation, and test data were used for all models compared in this section, so the benchmark remains appropriate for evaluating their relative performance under identical data conditions. Hyperparameters for the BE-CGCNN-based models and Allegro are listed in the Supporting Information.

The resulting parity plots are shown in Fig. 2. The original BE-CGCNN and the BE-iCGCNN variant showed similar performance, achieving test MAEs of 0.065 eV (Fig. 2(a), Fig. 2(b)). In contrast, the BE-OGCNN model achieved a distinctly lower test MAE of 0.042 eV (Fig. 2(c)), showing the benefit of incorporating orbital-interaction features into the bond-type-embedded BE-CGCNN framework with an adsorbate-oriented prediction scheme. Among the distance-sensitive reference models, Allegro achieved an MAE of 0.044 eV (Fig. 2(d)), while ALIGNN achieved the lowest MAE of 0.038 eV (Fig. 2(e)). Thus, BE-OGCNN was the best-performing distance-insensitive model examined here and achieved nearly the same accuracy as distance-sensitive reference models: its MAE was slightly lower than that of Allegro and only 0.004 eV higher than that of the angle-aware ALIGNN model. This result indicates that the proposed distance-insensitive representation can retain competitive accuracy without explicit distance or angle descriptors, although geometric information remains beneficial when fully relaxed structures are available.

The better accuracy of the BE-OGCNN compared to the other distance-insensitive models can be understood from the fundamental physics of adsorption. Adsorption energy is intrinsically governed by the hybridization between the adsorbate molecular orbitals and the valence orbitals of the underlying metal surface [39]. In BE-CGCNN and BE-iCGCNN, the distance-insensitive representation limits the ability to capture this hybridization effect through detailed geometric information. By contrast, the explicit embedding of valence electron configurations in BE-OGCNN enables the model to capture key aspects of chemical bonding directly from elemental identity, even without precise structural relaxation information. The difference between BE-OGCNN and the other BE-CGCNN-based variants can be seen more clearly in the validation MAE during training, as shown in Figure S1 in the Supporting Information. In addition, comparison with the original graph-level OGCNN formulation shown in Figure S2 in the Supporting Information further supports the importance of combining orbital-interaction features with bond-type embedding and an adsorbate-oriented prediction scheme.

For a screening-oriented assessment, we also evaluated the classification performance for identifying adsorption configurations with $|\Delta E_{\text{H}}| \leq 0.2$ eV. BE-OGCNN achieved a precision of 0.826, a recall of 0.864, and an F1 score of 0.844; the complete results are provided in Table S2 of the Supporting Information.

In conclusion, the BE-OGCNN architecture compensates for the absence of explicit geometric descriptors through physically relevant orbital-interaction features and an adsorbate-environment-focused readout. Although ALIGNN achieved the lowest MAE by explicitly incorporating geometric information, BE-OGCNN provided the best accuracy among the distance-insensitive models and remained competitive with the distance-sensitive references. We therefore adopted BE-OGCNN as the GNN backbone for all subsequent investigations in this study.

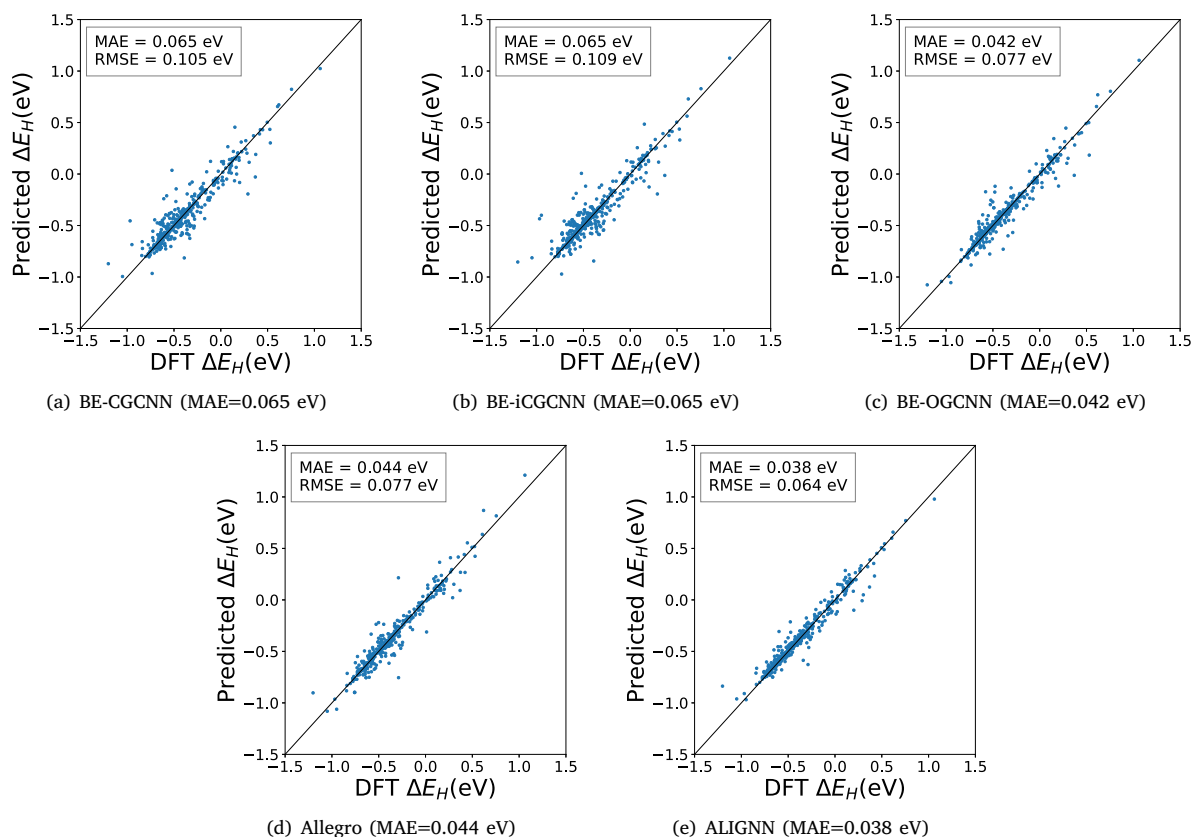


Fig. 2. Parity plots of adsorption energies with respect to DFT calculation for (a) BE-CGCNN, (b) BE-iCGCNN, (c) BE-OGCNN, (d) Allegro, and (e) ALIGNN on the 44-atom cluster test set.

3.2. Improved data efficiency via multi-task learning

To investigate the impact of our MTL framework on the model's data efficiency for adsorption-energy prediction, we compared several multi-task configurations with the corresponding single-task model using varying amounts of training data (from 100 to 1000 samples). The filtered OC20 subset, consisting of H-adsorption structures composed only of transition metals excluding Group 12 elements, was used as the adsorption-energy dataset for this analysis. For each trial, the validation and test sets were kept fixed.

The single-task baseline was trained exclusively to predict adsorption energy (ΔE_H). In contrast, the multi-task models were additionally trained to predict total energy (E_{Total}) and/or d-band center (ϵ_d) using a fixed set of 1000 bulk structures from the Alexandria dataset together with the same adsorption-energy data. The number of auxiliary bulk structures was fixed at 1000 throughout Section 3.2 and was not scaled with the number of adsorption-energy training samples. We evaluated three multi-task configurations with different combinations of auxiliary tasks: (1) total energy only (ΔE_H & E_{Total}), (2) d-band center only (ΔE_H & ϵ_d), and (3) both properties (ΔE_H & ϵ_d & E_{Total}).

To ensure statistical reliability, each experiment was repeated 20 times with different random seeds. The prediction results are summarized in Fig. 3. Fig. 3(a) compares the adsorption-energy MAE of all four model configurations across varying training data sizes, where data points are slightly shifted along the x-axis for visual clarity. Fig. 3(b) highlights the learning curves of the single-task baseline and the best-performing multi-task model (ΔE_H & ϵ_d & E_{Total}). In both panels, the error bars and the shaded region represent the 95% confidence intervals derived from the 20 independent runs. As shown in the figure, all multi-task models consistently exhibit lower MAEs than the single-task baseline across the entire range of training data sizes, demonstrating

the benefit of the multi-task approach. Among the multi-task variants, the model utilizing both auxiliary tasks (ΔE_H & ϵ_d & E_{Total}) achieved the highest accuracy, particularly at larger data sizes. Notably, the configuration including the d-band center (ΔE_H & ϵ_d) performed almost as well as the full model and outperformed the model using total energy alone (ΔE_H & E_{Total}). This suggests that the d-band center acts as an important electronic descriptor, helping the model capture adsorption physics more effectively than structural energy alone.

To further validate the significance of these improvements, we performed paired t-tests between the single-task model and the best-performing multi-task model (ΔE_H & ϵ_d & E_{Total}) across the 20 independent runs. The resulting p -values for each training-data size are summarized in Table 2. In all evaluated cases, the p -values were below 0.05, indicating that the performance improvements obtained with the MTL framework were statistically significant. As a control experiment, we also evaluated a multi-task model using atomic numbers as an auxiliary task instead of physically meaningful auxiliary properties. As shown in Figure S3 in the Supporting Information, this atomic-number auxiliary task did not provide comparable improvement over the single-task model. These results suggest that the observed improvement is not explained solely by the addition of an auxiliary output branch or extra supervised labels, but is related to auxiliary properties physically connected to adsorption energetics, such as the d-band center and total energy. This result further suggests that bulk auxiliary data without adsorption-energy labels (ΔE_H) can still provide useful transferable information for adsorption-energy prediction when physically correlated properties are learned simultaneously.

3.3. Performance validation on larger cluster structures

To assess the transferability of our model to more realistic catalyst sizes, we evaluated its prediction accuracy on 85-atom clusters. These

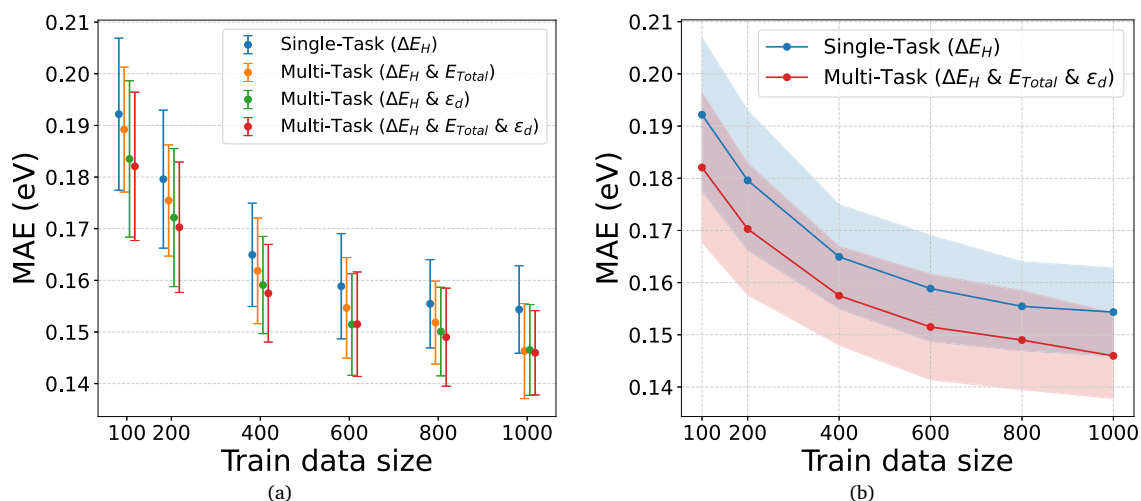


Fig. 3. Comparison of validation mean absolute error (MAE) among the Single-Task model and Multi-Task models. (a) The MAEs of the Single-Task model and three Multi-Task model configurations with different combinations of auxiliary tasks (ϵ_d and E_{Total}) evaluated at various training data sizes. (b) Learning curves of the Single-Task model and the best-performing Multi-Task ($\Delta E_H \epsilon_d$ and E_{Total}) model. The error bars and shaded regions indicate the 95% confidence intervals over 20 independent runs.

Table 2

Statistical significance of the improvement by the Multi-Task model ($\Delta E_H \epsilon_d$ & E_{Total}) over the Single-Task model.

Train data size	p-value (Paired t-test)
100	< 0.001
200	0.00126
400	< 0.001
600	< 0.001
800	0.0029
1000	< 0.001

structures represent a domain not included in the training set, which consists only of smaller clusters (19 and 44 atoms) and slab surfaces (OC20).

We compared the performance of the single-task model against the multi-task model after 100 epochs of training. For the MTL framework, we augmented the training set with 2000 bulk structures from the Alexandria dataset and utilized both the d-band center and total energy as auxiliary tasks. The same transition-metal filtering criterion described in Section 2.1.2 was applied to the public datasets, rather than restricting them to the nine elements used in the in-house cluster dataset.

The parity plots for the 85-atom cluster test set are presented in Fig. 4. The single-task model achieved an MAE of 0.122 eV and a root mean square error (RMSE) of 0.152 eV (Fig. 4(a)). In comparison, the multi-task model achieved an MAE of 0.105 eV and an RMSE of 0.140 eV (Fig. 4(b)). Although the absolute improvement is modest, the reduction in both MAE and RMSE suggests that the MTL framework improves not only the average prediction accuracy but also the frequency of larger prediction errors.

These results suggest that MTL can modestly improve model reliability when applied to previously unseen cluster sizes. The improvement should not be interpreted as a dramatic performance gain; rather, it indicates a consistent reduction in prediction errors across complementary metrics. Two factors may contribute to this improvement. First, the model may benefit from learning bulk-like environments. Large nanoclusters contain internal atomic environments more similar to bulk crystals than those found in smaller clusters. By learning bulk-derived properties such as the d-band center and total energy, the model may better describe these environments. Second, simultaneous learning of auxiliary tasks may provide a regularization effect that

improves generalization beyond the specific structural patterns present in the smaller-cluster training data.

An important limitation of the present validation is that all in-house clusters adopt an octahedral motif. This controlled motif provides multiple cluster sizes, terrace, edge, and vertex environments, and atop, bridge, and hollow adsorption sites, but it does not represent the full morphological diversity of catalytic nanoparticles. In particular, irregular, reconstructed, defect-rich, or thermally generated structures may contain adsorption environments absent from the present dataset. Consequently, the reported accuracy should not be assumed to transfer directly to arbitrary nanoparticle morphologies. Evaluating the model on randomly generated clusters and structures sampled by molecular dynamics is an important direction for future work.

4. Conclusions

We investigated a dual strategy for improving the accuracy and data efficiency of GNN-based adsorption-energy prediction on alloy nanoclusters. First, we showed that the distance-insensitive BE-OGCNN architecture, which incorporates orbital-interaction features and an inverse coordination-number weighting scheme, was the best-performing distance-insensitive model examined in this work. The model achieved an MAE of 0.042 eV for hydrogen adsorption-energy prediction on 44-atom unary and binary alloy clusters containing elements selected from nine species (Ag, Au, Co, Cu, Ir, Ni, Pd, Pt, and Rh). This accuracy was comparable to that of the distance-sensitive Allegro model (MAE = 0.044 eV) and close to that of the angle-aware ALIGNN model (MAE = 0.038 eV) under the present benchmark conditions. These results indicate that orbital-interaction features, bond-type embedding, and an adsorbate-oriented prediction scheme can largely compensate for the absence of explicit geometric descriptors in this task, while explicit geometric information can still provide additional accuracy benefits when relaxed structures are available.

Second, we implemented an MTL framework in which the model was co-trained on adsorption energy, d-band center, and total energy. The results suggest that bulk-derived auxiliary labels can improve the data efficiency of adsorption-energy prediction and can modestly enhance transferability to clusters larger than those used in training. These findings indicate that combining distance-insensitive graph representations with bulk-property auxiliary learning is a promising direction for data-efficient catalyst nanoparticle screening, although further validation across broader compositions, adsorbates, and nanoparticle morphologies will be necessary to establish general applicability.

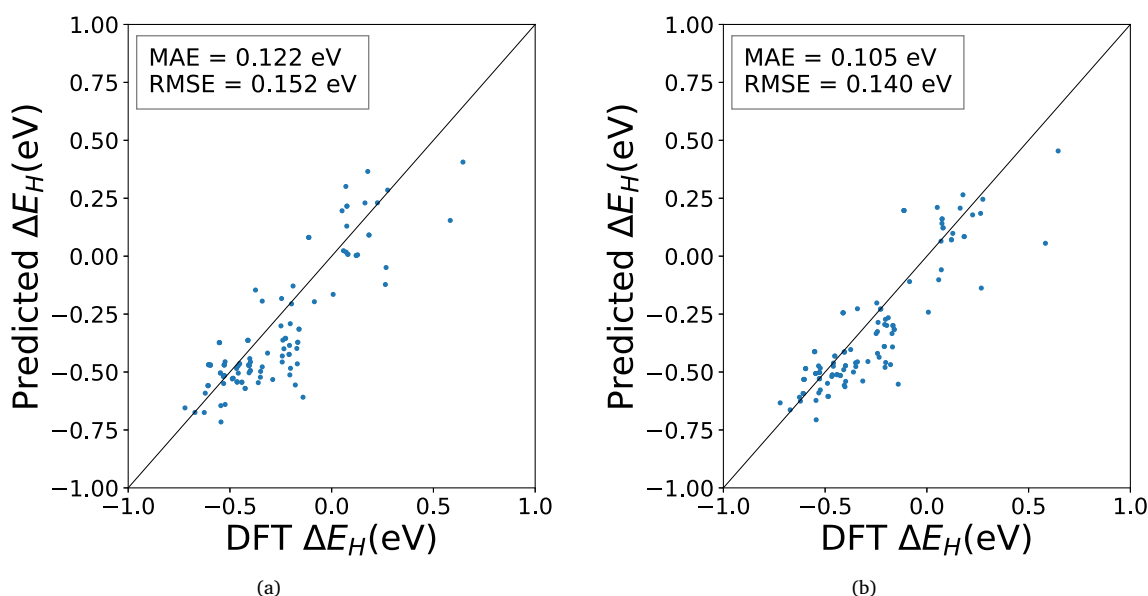


Fig. 4. Comparison of the prediction accuracy between (a) Single-Task and (b) Multi-Task models on large, 85-atom clusters.

CRedit authorship contribution statement

Koki Otsuka: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Anh Khoa Augustin Lu:** Writing – review & editing. **Koji Shimizu:** Writing – review & editing, Investigation, Funding acquisition, Data curation. **Satoshi Watanabe:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the authors used ChatGPT to improve language clarity and readability in the manuscript. After using this tool, the authors carefully reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

This work was supported by the Japan Science and Technology Agency (JST) as part of SICORP (Grant Number JPMJSC21E2), and by the Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant Number 24K01284).

Code availability

The implemented machine learning model code is available at https://github.com/modelna/BE_OGCNN.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Part of the calculations were performed using the supercomputers at the Information Technology Center and the Institute for Solid State Physics (ISSP), the University of Tokyo.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.commatsci.2026.114987>.

Data availability

Research Link Provided

[Dataset for Distance-Insensitive Graph Neural Networks and Multi-Task Learning for Accurate Prediction of Adsorption Energies on Alloy Nanostructures \(Original data\)](#) (github)

References

- [1] M.S. Akhtar, W. Zaman, Sustainable catalysis for green chemistry and energy transition, *Catalysts* 15 (2025) 773, <http://dx.doi.org/10.3390/catal15080773>.
- [2] V. Van Pham, S. Park, Catalysis for renewable energy and sustainable development, *Top. Catal.* 67 (2024) 1053–1054, <http://dx.doi.org/10.1007/s11244-024-02000-z>.
- [3] H. Hosseini, C.J. Herring, C.F. Nwaokorie, G.A. Sulley, M.M. Montemore, Computational design of catalysts with experimental validation: Recent successes, effective strategies, and pitfalls, *J. Phys. Chem. C* 128 (2024) 18144–18157, <http://dx.doi.org/10.1021/acs.jpcc.4c04949>.
- [4] M.G. Quesne, F. Silveri, N.H. de Leeuw, C.R.A. Catlow, Advances in sustainable catalysis: A computational perspective, *Front. Chem.* 7 (2019) 182, <http://dx.doi.org/10.3389/fchem.2019.00182>.
- [5] V. Butera, Density functional theory methods applied to homogeneous and heterogeneous catalysis: A short review and a practical user guide, *Phys. Chem. Chem. Phys.* 26 (2024) 7950–7970, <http://dx.doi.org/10.1039/D4CP00266K>.
- [6] A. Kordatos, K. Mohammed, R. Vakili, H. Manyar, A. Goguet, E. Gibson, M. Carravetta, P. Wells, C.-K. Skylaris, Bridging the size gap between experiment and theory: Large-scale DFT calculations on realistic sized pd particles for acetylene hydrogenation, *RSC Adv.* 14 (2024) 27799–27808, <http://dx.doi.org/10.1039/D4RA03369H>.
- [7] L.G. Verga, J. Arons, M. Sarwar, D. Thompsett, A.E. Russell, C.-K. Skylaris, DFT calculation of oxygen adsorption on platinum nanoparticles: Coverage and size effects, *Faraday Discuss.* 208 (2018) 497–522, <http://dx.doi.org/10.1039/C7FD00218A>.
- [8] L. Fiedler, K. Shah, M. Bussmann, A. Cangi, Deep dive into machine learning density functional theory for materials science and chemistry, *Phys. Rev. Mater.* 6 (2022) 040301, <http://dx.doi.org/10.1103/PhysRevMaterials.6.040301>.
- [9] Z. Wang, A. Chen, K. Tao, Y. Han, J. Li, MatGPT: A vane of materials informatics from past, present, to future, *Adv. Mater.* 36 (2024) 2306733, <http://dx.doi.org/10.1002/adma.202306733>.

- [10] L. Cao, Recent advances in the application of machine-learning algorithms to predict adsorption energies, *Trends Chem.* 4 (2022) 347–360, <http://dx.doi.org/10.1016/j.trechm.2022.01.012>.
- [11] P. Reiser, M. Neubert, A. Eberhard, L. Torresi, C. Zhou, C. Shao, H. Metni, C. van Hoesel, H. Schopmans, T. Sommer, P. Friederich, Graph neural networks for materials science and chemistry, *Commun. Mater.* 3 (2022) 93, <http://dx.doi.org/10.1038/s43246-022-00315-6>.
- [12] T. Xie, J.C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, *Phys. Rev. Lett.* 120 (2018) 145301, <http://dx.doi.org/10.1103/PhysRevLett.120.145301>.
- [13] K. Bang, D. Hong, Y. Park, D. Kim, S.S. Han, H.M. Lee, Machine learning-enabled exploration of the electrochemical stability of real-scale metallic nanoparticles, *Nat. Commun.* 14 (2023) 3004, <http://dx.doi.org/10.1038/s41467-023-38758-1>.
- [14] C.W. Park, C. Wolverton, Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery, *Phys. Rev. Mater.* 4 (2020) 063801, <http://dx.doi.org/10.1103/PhysRevMaterials.4.063801>.
- [15] M. Karamad, R. Magar, Y. Shi, S. Siahrostami, I.D. Gates, A.B. Farimani, Orbital graph convolutional neural network for material property prediction, *Phys. Rev. Mater.* 4 (2020) 093801, <http://dx.doi.org/10.1103/PhysRevMaterials.4.093801>.
- [16] Y. Lu, M. Chen, Q. Zhang, J. Zhang, C. Zhang, S. Xu, Q. Bi, Adaptive edge-aware graph convolutional with multi-task learning for simultaneous prediction of material properties, *Npj Comput. Mater.* 12 (2026) 49, <http://dx.doi.org/10.1038/s41524-025-01917-7>.
- [17] G. Shi, Y. Wang, K. Yang, Y. Qiu, H. Zhu, X. Zeng, A surface emphasized multi-task learning framework for surface property predictions: A case study of magnesium intermetallics, *J. Magnes. Alloys* 14 (2026) 49, <http://dx.doi.org/10.1016/j.jma.2024.12.005>.
- [18] C.M.A. Rahman, N.B. Alam, A.S. El-Wakeel, J. Ryu, P.K. Gyawali, Multi-task learning for material property prediction, in: *2024 Int. Conf. Mach. Learn. Appl., ICMLA, 2024*, pp. 1065–1070.
- [19] F. Conrad, H. Wiemer, S. Ihlenfeldt, Leveraging multi-task learning regressor chains for small and sparse tabular data in materials design, *Mach. Learn.: Sci. Technol.* 6 (2025) 015045, <http://dx.doi.org/10.1088/2632-2153/adae53>.
- [20] 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, Open catalyst 2020 (OC20) dataset and community challenges, *ACS Catal.* 11 (2021) 6059–6072, <http://dx.doi.org/10.1021/acscatal.0c04525>.
- [21] R. Tran, J. Lan, M. Shuaibi, B.M. Wood, S. Goyal, A. Das, J. Heras-Domingo, A. Kolluru, A. Rizvi, N. Shoghi, A. Sriram, F. Therrien, J. Abed, O. Voznyy, E.H. Sargent, Z.W. Ulissi, C.L. Zitnick, The open catalyst 2022 (OC22) dataset and challenges for oxide electrocatalysts, *ACS Catal.* 13 (2023) 3066–3084, <http://dx.doi.org/10.1021/acscatal.2c05426>.
- [22] D. Zhang, X. Liu, X. Zhang, C. Zhang, C. Cai, H. Bi, Y. Du, X. Qin, A. Peng, J. Huang, B. Li, Y. Shan, J. Zeng, Y. Zhang, S. Liu, Y. Li, J. Chang, X. Wang, S. Zhou, J. Liu, X. Luo, Z. Wang, W. Jiang, J. Wu, Y. Yang, J. Yang, M. Yang, F.-Q. Gong, L. Zhang, M. Shi, F.-Z. Dai, D.M. York, S. Liu, T. Zhu, Z. Zhong, J. Lv, J. Cheng, W. Jia, M. Chen, G. Ke, W. E, L. Zhang, H. Wang, DPA-2: A large atomic model as a multi-task learner, *Npj Comput. Mater.* 10 (2024) 293, <http://dx.doi.org/10.1038/s41524-024-01493-2>.
- [23] C. Yang, C. Wu, W. Xie, D. Xie, P. Hu, General reactive element-based machine learning potentials for heterogeneous catalysis, *Nat. Catal.* 8 (2025) 891–904, <http://dx.doi.org/10.1038/s41929-025-01398-3>.
- [24] J. Ock, S. Badrinarayanan, R. Magar, A. Antony, A. Barati Farimani, Multimodal language and graph learning of adsorption configuration in catalysis, *Nat. Mach. Intell.* 6 (2024) 1501–1511, <http://dx.doi.org/10.1038/s42256-024-00930-7>.
- [25] S. Back, J. Yoon, N. Tian, W. Zhong, K. Tran, Z.W. Ulissi, Convolutional neural network of atomic surface structures to predict binding energies for high-throughput screening of catalysts, *J. Phys. Chem. Lett.* 10 (2019) 4401–4408, <http://dx.doi.org/10.1021/acs.jpclett.9b01428>.
- [26] H. Zhang, Z. Wang, J. Ren, J. Liu, J. Li, Ultra-fast and accurate binding energy prediction of shuttle effect-suppressive sulfur hosts for lithium-sulfur batteries using machine learning, *Energy Storage Mater.* 35 (2021) 88–98, <http://dx.doi.org/10.1016/j.ensm.2020.11.009>.
- [27] B. Hammer, J.K. Nørskov, Why gold is the noblest of all the metals, *Nature* 376 (1995) 238–240, <http://dx.doi.org/10.1038/376238a0>.
- [28] V. Fung, G. Hu, P. Ganesh, B.G. Sumpter, Machine learned features from density of states for accurate adsorption energy prediction, *Nat. Commun.* 12 (2021) 88, <http://dx.doi.org/10.1038/s41467-020-20342-6>.
- [29] J. Schmidt, N. Hoffmann, H.-C. Wang, P. Borlido, P.J.M.A. Carriço, T.F.T. Cerqueira, S. Botti, M.A.L. Marques, Machine-learning-assisted determination of the global zero-temperature phase diagram of materials, *Adv. Mater.* 35 (2023) 2210788, <http://dx.doi.org/10.1002/adma.202210788>.
- [30] K. Miyamoto, K. Shimizu, A.K.A. Lu, S. Watanabe, Construction of machine learning potentials toward the exploration of alloy cluster catalysts, *E-J. Surf. Sci. Nanotechnol.* 23 (2025) 188–192, <http://dx.doi.org/10.1380/ejssnt.2025-028>.
- [31] G. Kresse, J. Hafner, Ab initio molecular dynamics for liquid metals, *Phys. Rev. B* 47 (1993) 558–561, <http://dx.doi.org/10.1103/PhysRevB.47.558>.
- [32] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.* 77 (1996) 3865–3868, <http://dx.doi.org/10.1103/PhysRevLett.77.3865>.
- [33] P.E. Blöchl, Projector augmented-wave method, *Phys. Rev. B* 50 (1994) 17953–17979, <http://dx.doi.org/10.1103/PhysRevB.50.17953>.
- [34] H.J. Monkhorst, J.D. Pack, Special points for Brillouin-zone integrations, *Phys. Rev. B* 13 (1976) 5188–5192, <http://dx.doi.org/10.1103/PhysRevB.13.5188>.
- [35] L. Sbailò, Á. Fekete, L.M. Ghiringhelli, M. Scheffler, The NOMAD artificial-intelligence toolkit: Turning materials-science data into knowledge and understanding, *Npj Comput. Mater.* 8 (2022) 250, <http://dx.doi.org/10.1038/s41524-022-00935-z>.
- [36] A.H. Larsen, et al., The atomic simulation environment—a python library for working with atoms, *J. Phys.: Condens. Matter.* 29 (2017) 273002, <http://dx.doi.org/10.1088/1361-648X/aa680e>.
- [37] A. Musaelian, S. Batzner, A. Johansson, L. Sun, C.J. Owen, M. Kornbluth, B. Kozinsky, Learning local equivariant representations for large-scale atomistic dynamics, *Nat. Commun.* 14 (2023) 579, <http://dx.doi.org/10.1038/s41467-023-36329-y>.
- [38] K. Choudhary, B. DeCost, Atomistic line graph neural network for improved materials property predictions, *Npj Comput. Mater.* 7 (2021) 185, <http://dx.doi.org/10.1038/s41524-021-00650-1>.
- [39] B. Hammer, J.K. Nørskov, Electronic factors determining the reactivity of metal surfaces, *Surf. Sci.* 343 (1995) 211–220, [http://dx.doi.org/10.1016/0039-6028\(96\)80007-0](http://dx.doi.org/10.1016/0039-6028(96)80007-0).