Engineering 52 (2025) 172–182

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Research
AI for Process Manufacturing—Article

Theoretical High-Throughput Screening of Single-Atom CO_2 Electroreduction Catalysts to Methanol Using Active Learning

Honghao Chen [a], Jun Yin [b], Jiali Li [b], Xiaonan Wang [a,c,*]

[a] Department of Chemical Engineering, Tsinghua University, Beijing 100084, China
[b] Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore 119077, Singapore
[c] State Key Laboratory of Chemical Engineering and Low-Carbon Technology, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Article history:
Received 29 July 2024
Revised 7 January 2025
Accepted 18 March 2025
Available online 5 August 2025

Keywords:
CO_2 electrochemical reduction
Machine learning
Active learning
Catalyst
Decarbonization

ABSTRACT

Industrial decarbonization is critical for achieving net-zero goals. The carbon dioxide electrochemical reduction reaction (CO_2RR) is a promising approach for converting CO_2 into high-value chemicals, offering the potential for decarbonizing industrial processes toward a sustainable, carbon-neutral future. However, developing CO_2RR catalysts with high selectivity and activity remains a challenge due to the complexity of finding such catalysts and the inefficiency of traditional computational or experimental approaches. Here, we present a methodology integrating density functional theory (DFT) calculations, deep learning models, and an active learning strategy to rapidly screen high-performance catalysts. The proposed methodology is then demonstrated on graphene-based single-atom catalysts for selective CO_2 electroreduction to methanol. First, we conduct systematic binding energy calculations for 3045 single-atom catalysts to identify thermodynamically stable catalysts as the design space. We then use a graph neural network, fine-tuned with a specialized adsorption energy database, to predict the relative activity and selectivity of the candidate catalysts. An autonomous active learning framework is used to facilitate the exploration of designs. After six learning cycles and 2180 adsorption calculations across 15 intermediates, we develop a surrogate model that identifies four novel catalysts on the Pareto front of activity and selectivity. Our work demonstrates the effectiveness of leveraging a domain foundation model with an active learning framework and holds potential to significantly accelerate the discovery of high-performance CO_2RR catalysts.

© 2025 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Given the intensifying climate impacts, such as escalating heatwaves and extreme precipitation events, the urgency to mitigate greenhouse gas emissions is increasingly critical. This urgency is underscored by the contemporary warming of 1.2 °C above pre-industrial levels—a situation that now demands a rigorous and immediate response to avert severe ecological and socio–economic consequences, as agreed upon in the Paris Agreement [1]. In this context, carbon dioxide electrochemical reduction reaction (CO_2RR) technologies are gaining substantial attention, as they can convert carbon dioxide (CO_2) into valuable chemicals and renewable fuels [2]. Among various products, the conversion of CO_2 into carbon monoxide (CO) has undergone significant advancements and is regarded as a promising pathway for industrial adoption [3]. While other processes are still under development, they present potential options for achieving a sustainable future [4].

Currently, the industrial application of CO_2RR is primarily constrained by techno–economic factors, which are influenced by various aspects such as system design, electrolyzer design, catalyst design, and electrolyte design [5]. Among these factors, catalyst development is central to industrial applications of the CO_2RR process; it significantly influences the current density, Faradaic efficiency, energy efficiency, and stability, all of which have a substantial impact on operating costs. The lack of high-performance catalysts for producing high-value-added chemicals (e.g., alcohol and other products) further limits the application potential of the CO_2RR process. Therefore, the advancement of high-performance electrocatalysts that can produce desirable

---

* Corresponding author.
E-mail address: wangxiaonan@mail.tsinghua.edu.cn (X. Wang).

https://doi.org/10.1016/j.eng.2025.03.039
2095-8099/© 2025 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

high-value chemicals is necessary for the industrial adoption of CO_2RR [6]. Among various kinds of candidate catalysts, single-atom catalysts (SACs) have attracted special attention. With their unique electronic structure, low-coordination metal atoms, strong metal-carrier interactions, and high atomic-utilization efficiency, SACs exhibit excellent catalytic performance characterized by low cost, adjustable structure, high energy efficiency, and low overpotential, making them a promising candidate for advancing CO_2RR technologies [7–9].

Nevertheless, discovering catalysts with outstanding performance for products beyond CO remains a significant challenge [7]. Given the vast and seemingly limitless catalyst design options, machine learning has emerged as a suitable approach to expedite the effective exploration of such design spaces [10–12]. Machine learning has been widely adopted at different design stages for various materials in the carbon capture and utilization value chain and has also been used to aid CO_2RR catalyst design [13]. Such attempts have been demonstrated in alloy nanoparticles [14,15], metal oxides [16], SACs [17], dual-atom catalysts [18], and metal/alloy designs [19]. For example, Chen et al. [17] utilized an extreme gradient boosting regression model to quickly predict hydrogen evolution reaction (HER) and CO_2RR activity. Wan et al. [18] used a similar method and extended it to screen phthalocyanine dual-metal-site catalysts, decreasing the computational resource requirement by 6.87 times. Tamtaji et al. [20] used the support vector regression (SVR) algorithm and properties of the metal, substrate, and intermediates to develop a general prediction model for the Gibbs energies of various reaction intermediates involved in oxygen reduction reaction (ORR), nitrogen reduction reaction (NRR), HER, and CO_2RR. Their work showed a promising future in applying machine learning methods for CO_2RR electrocatalyst screening.

However, most studies on CO_2RR catalysts using machine learning approaches have primarily focused on determining the adsorption energy of CO and hydrogen (H) as the indicator for CO_2RR activity and CO_2RR/HER selectivity [14,18–20]. Although this research has revealed underlying principles for designing catalysts, it is still incomplete. Thus, it is crucial for researchers to shift their attention toward studying individual products, pathways, and intermediates beyond CO to understand the selectivity and dominant reaction pathways associated with these catalysts, which will require more information from machine learning models [19]. However, current machine-learning-based research in this field typically uses features specially designed according to the specific structures [17,18,20]. Although features derived from structural observations and atomic properties allow for increased physical understanding, they may present challenges in system variations and depend on human expertise for redesign, limiting the transferability of these machine learning models [19].

To address the challenges in exploring the catalytic activity and selectivity of CO_2RR catalysts, we propose a comprehensive, open-source computational workflow that combines universal graph neural networks with density functional theory (DFT) in an active learning loop (Fig. 1). This automated framework efficiently explores a vast catalyst space, leveraging deep active learning for broad applicability across diverse catalyst types without the need to redesign descriptors, thereby increasing transferability once an atom-level model is established. By focusing on individual intermediates across the full reaction network, our method provides insights into selectivity and dominant pathways in CO_2 reduction, ultimately aiming to improve catalytic efficiency. The nature of this workflow allows it to support and guide future experimental efforts, establishing an adaptable platform for quick catalyst screening and design.

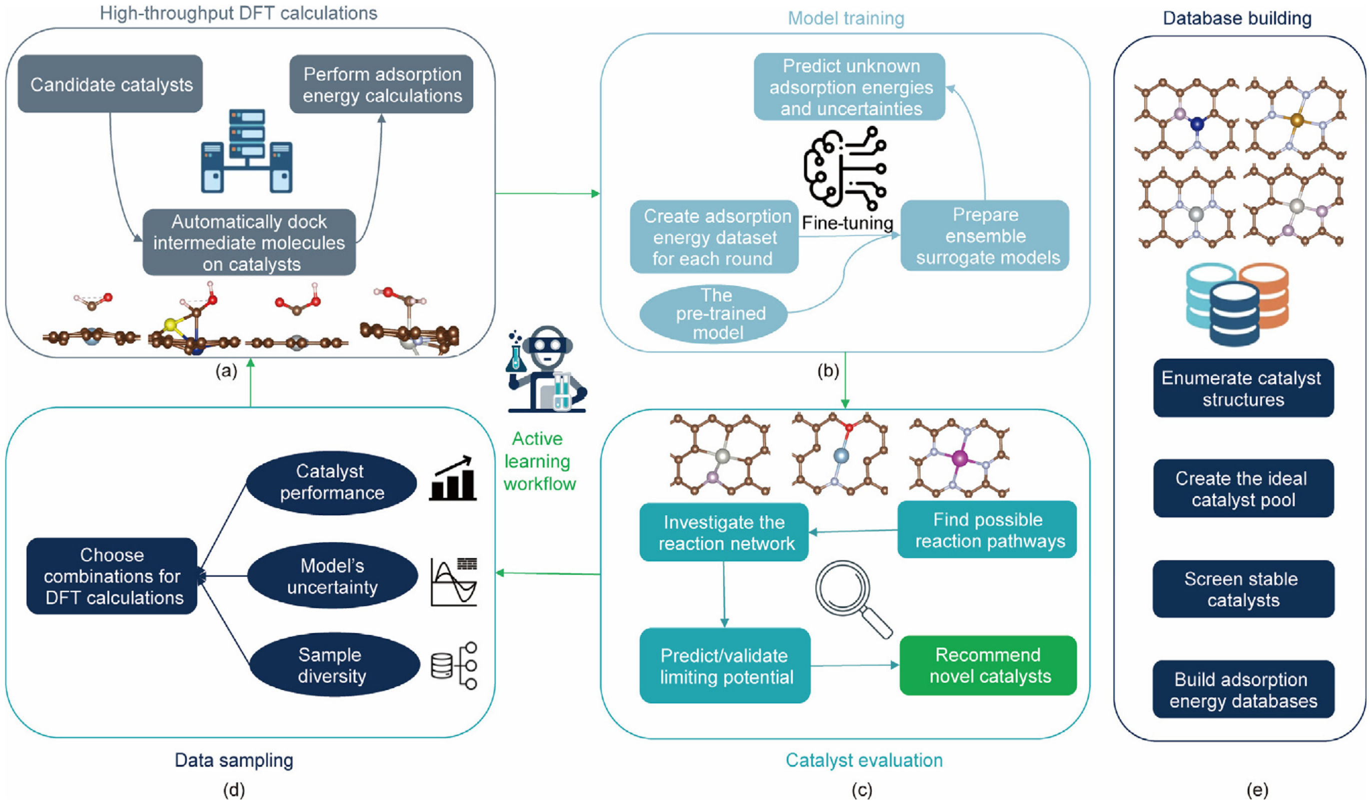
High-throughput DFT calculations
Candidate catalysts
Perform adsorption energy calculations
Automatically dock intermediate molecules on catalysts
(a)
Model training
Predict unknown adsorption energies and uncertainties
Fine-tuning
Create adsorption energy dataset for each round
Prepare ensemble surrogate models
The pre-trained model
(b)
Database building
Enumerate catalyst structures
Create the ideal catalyst pool
Screen stable catalysts
Build adsorption energy databases
Active learning workflow
Catalyst performance
Choose combinations for DFT calculations
Model's uncertainty
Sample diversity
Data sampling
(d)
Investigate the reaction network
Find possible reaction pathways
Predict/validate limiting potential
Recommend novel catalysts
Catalyst evaluation
(c)
(e)

Fig. 1. The overall workflow for exploring novel SACs for CO_2RR. (a–d) The main components in the active learning framework; (a) performing high-throughput automated DFT calculations to acquire data for training a surrogate model; (b) utilizing a universal pre-trained model to fine-tune our dataset and perform a quick prediction of adsorption energies; (c) using the predicted adsorption energy to evaluate the performance of candidate catalysts in the whole space, with all existing DFT-calculated energies checked to validate the model recommendations; (d) integrating a sampling strategy to determine the data points to be added to the database. (e) Alongside the active learning process, structures and high-throughput calculations are documented for further analysis.

2. Methods

2.1. DFT calculations

In this research, we perform the DFT calculations using the Vienna *ab-initio* simulation package (VASP) quantum chemistry software package from University of Vienna, Austria [21]. Based on previous studies [14,17,18,22–24] exploring pure DFT and machine learning approaches, and considering the balance of computational cost, we use the classic computational hydrogen electrode (CHE) model [25] to estimate the activity and selectivity of the catalysts, and the DFT calculation data are used as the ground truth for deep learning models. However, we must emphasize that, because the CHE model is a simplified model, a discrepancy can exist between the theoretical and experimental results. Nevertheless, it is helpful in capturing the trend of catalyst performance.

We use the projector-augmented wave (PAW) method [26,27] and the Perdew–Burke–Ernzerhof (PBE) [28] functional to calculate the electronic exchange energy. The DFT-D3 correction [29] with Becke–Johnson damping is employed to describe the medium- and long-range interactions. All calculations in this work are performed with spin polarization correction [30]. The plane wave cutoff energy is set at 520 eV, and a vacuum layer of 15 Å is introduced to separate the interactions between the adsorbed surfaces.

The calculations for slabs are conducted using the Gaussian smearing method with a width of 0.05 eV. Here, we use a 5 × 5 graphene supercell with lattice parameters $a = b = 12.34$ Å to generate the catalyst surface used in this work, and the k-point grid is sampled with a density of 3 × 3 × 1. The convergence criteria are set to 10^{-6} eV for electronic self-consistency, and the forces on each atom are relaxed to below 0.02 eV·$Å^{-1}$ for geometry optimization.

In this work, we use the following equation to calculate the binding energy of catalysts:

$$E_{\text{bind}} = E_{M+\text{vac}} - E_M - E_{\text{vac}} \tag{1}$$

where E_{bind} is the binding energy, $E_{M+\text{vac}}$ is the total energy of the geometry where the specified metal atom M is anchored on the defect with type vac calculated by DFT, E_M is the average energy per atom in the stable metal crystal, and E_{vac} is the total energy of the defect with type vac.

We employ an additional thermal correction, as shown in Eq. (2), to our calculation results using VASPKIT [31] under the conditions of 298.15 K and 101.3 kPa (1 atm) for the gas molecules and the adsorbate on the catalyst. Additionally, due to the inherent error of the PBE functional, we employ another semi-empirical method [32] for the gas molecules. Details are provided in Appendix A Tables S1–S3.

$$\mu = E_{\text{DFT}} + E_{\text{ZPE}} + \text{H} - \text{TS} \tag{2}$$

Here, μ is Gibbs free energy for the system. The E_{DFT} is the energy from the raw DFT calculation, E_{ZPE} is the zero-point energy, $\text{H} - \text{TS}$ is the correction term derived from VASPKIT output for enthalpy and entropy terms. For the partial density of state (PDOS) analysis, we employ a more fine-grained approach in our electronic structure calculations, utilizing a 5 × 5 × 1 k-point mesh and a tetrahedron method with Bloch corrections and enabling the projection of wave functions onto spherical harmonics.

2.2. Preparation of the dataset

Input and output formats should be determined for machine learning models in order to build a database for supervised machine learning tasks. In our work, the enumerated atomic structure in crystallographic information file (CIF) format is chosen as our dataset input using Python Materials Genomics (PyMatGen) [33]. We performed massive high-throughput calculations automatically using a self-written code kit to prepare the necessary data for the machine learning models (more details in Appendix A Figs. S1 and S2). A vast combination of ideal catalysts can be created by substituting the central metal atoms and surrounding atoms of the catalysts. For this work, 29 candidate metal centers and five doping elements (nitrogen (N), sulfur (S), oxygen (O), carbon (C), and phosphorus (P)) were chosen, with systematic substitution resulting in 1015 and 2030 catalysts for the M-X_3 and M-X_4 catalysts, respectively (Table S4 in Appendix A); the relevant codes are available in Appendix A.

Before screening for potential catalysts, we used the binding energy as our main criterion to filter out stable catalysts. The binding energy is used to describe the trend of cohesion of metal atoms for SACs and has been found to be one of the main factors influencing the stability of SACs [34]. When building the candidate catalyst pool for SACs, we screened out catalysts with a positive binding energy or a vertical displacement greater than 2 Å, as most of these were observed to have unreasonable structures for a stable catalyst (Fig. 2(a)).

For output, we performed a post-processing procedure on the DFT calculation results. As a common strategy in this field [14,17,35], the adsorption energy is calculated using Eq. (3) and adopted to construct the dataset as the main prediction target.

$$\Delta G_{\text{ads},X^*} = \mu_{X^*} - \mu_X - \mu_* \tag{3}$$

where μ_{X^*}, μ_X, and μ_* represent the total Gibbs energy of the surface with adsorbed intermediate molecule X , the Gibbs energy of the pure intermediate molecule X , and the Gibbs energy of the empty surface, respectively. $\Delta G_{\text{ads},X^*}$ is the adsorption energy for the molecule X. This processing step ensures that the predicted targets are roughly within the same order of magnitude, which is beneficial for predicting machine learning models compared with directly predicting the total energy from DFT calculations. When constructing the dataset, only those calculations that converge were used for the dataset. However, when examining the initial dataset, we still found some samples with unreasonable adsorption energies due to convergence to abnormal structures. Therefore, any sample with an adsorption energy greater than an absolute value of 10 eV was excluded from the dataset to avoid such unexpected sample types.

2.3. Catalyst performance evaluation

To apply the adsorption energy prediction results from the machine learning models or DFT calculations to predict catalyst performance, we further transformed them into the limiting potential for each catalyst–product pair. The limiting potential refers to the potential at which all reaction steps are downhill in the free energy change diagram; it is a common indicator for evaluating catalyst performance. It can be calculated using Eqs. (4) and (5) [36].

$$U_{\text{L}}(\text{product}) = \min_{p_m \in P} \{\Delta G_{\max}^{p_1}, \Delta G_{\max}^{p_2}, \ldots\} \tag{4}$$

$$\Delta G_{\max}^{p_m} = \max_{r_i \in p_m} \{\Delta G_{r_1}(U_{\text{app}}), \Delta G_{r_2}(U_{\text{app}}), \ldots\} \tag{5}$$

where U_{L} is the limiting potential. P denotes all possible reaction pathways from a reactant to a specified product, $\Delta G_{\max}^{p_m}$ represents the maximum reaction energy barrier along one specified pathway p_m ($p_m \in P$), and $\Delta G_{r_i}(U_{\text{app}})$ represents the energy barrier of each elementary step r_i under the applied potential U_{app}. The actual reaction pathway is the one with the lowest energy barrier, where the energy barrier of each pathway is determined by the elementary step with the highest energy barrier. The absolute value of the limiting potential is commonly used to describe the activity of an electrochemical catalyst toward a specific product, while the difference

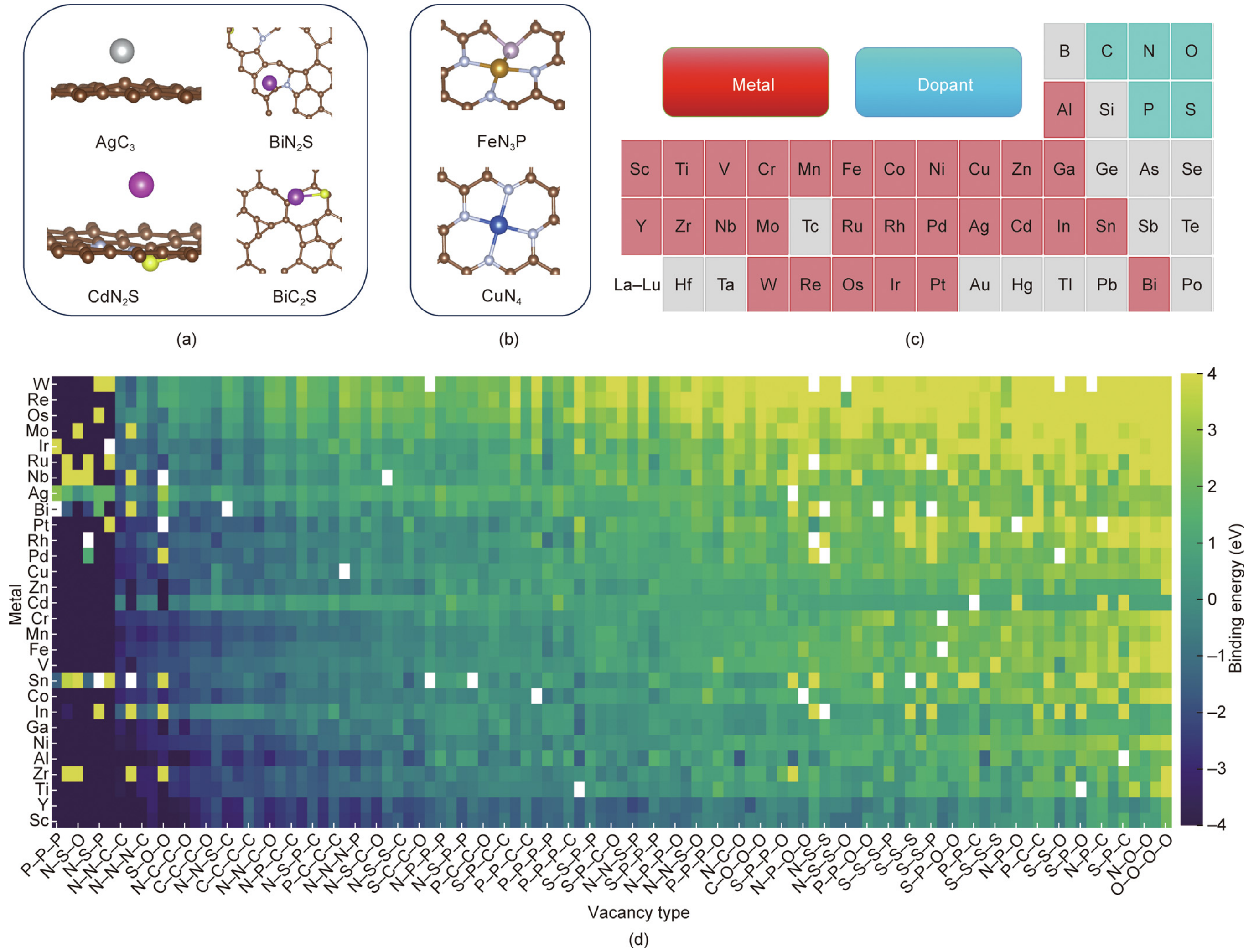

Fig. 2. Binding energy calculations for SACs. (a, b) Illustration of (a) the candidate catalysts in the design space and (b) the unstable catalysts expelled from the design space. (c) The elements chosen for the central atom and dopants. (d) Heatmap for the binding energy of SACs, where the lower left region indicates relative stability and the upper right region shows relative instability. Missing values, representing non-converged data, are excluded from the design space.

between two products can be used as an indicator for product selectivity; thus, the limiting potential for all products could be used to establish suitable theoretical descriptors to evaluate the comprehensive performance (activity/selectivity) of a catalyst.

For both electrochemical steps happening on the surface and steps involving adsorption and desorption, the free energy change for each reaction step is calculated based on the CHE model [37]:

$$\text{Reaction}: A^{*} + n\left(\mathrm{H}^{+} + \mathrm{e}^{-}\right) \rightarrow B^{*} \tag{6}$$

$$\begin{aligned} \Delta G_{A \rightarrow B} &= \mu_{B^{*}} - \mu_{A^{*}} - nG(\mathrm{e}) \\ &= \left(\Delta G_{\mathrm{ads},B^{*}} + \mu_{B}\right) - \left(\Delta G_{\mathrm{ads},A^{*}} + \mu_{A}\right) - nG(\mathrm{e}) \end{aligned} \tag{7}$$

$$G(\mathrm{e}) = \mu_{\mathrm{H}_2}/2 - \mathrm{e}U_{\mathrm{app}} - 0.0592\mathrm{pH} \tag{8}$$

where $\Delta G_{A \rightarrow B}$ is the Gibbs energy change for an elementary step. $G(\mathrm{e})$ is the energy for proton/electron pair; $\mu_{B^{*}}$ and $\mu_{A^{*}}$ are the Gibbs energies of the surface with adsorbed intermediate molecules B and A; μ_B, μ_A, and μ_{H_2} are the Gibbs energies of the pure intermediate molecule B, molecule A, and hydrogen gas molecule, respectively; n is the electron transfer count involved in the reaction; $\Delta G_{\mathrm{ads},B^{*}}$ and $\Delta G_{\mathrm{ads},A^{*}}$ are the adsorption energies for the molecules B and A, respectively.

In this study, we used pH = 7 and U_{app} = −0.38 V versus SHE for all calculations to evaluate catalyst performance, while using the limiting potential as a metric to simulate performance under neutral conditions and equilibrium potential. Since all the chemical elementary steps under consideration are one-electron reactions, the Gibbs energy will shift uniformly, and the applied potential will not affect the relative differences between steps along the reaction path or the limiting potential between different electrochemical steps. However, a few steps involve non-electron-transfer processes, such as CO_2 adsorption and CO desorption, which could alter the relative performance ranking. To address this, we also evaluated two additional potentials, −0.88 V versus SHE and −1.38 V versus SHE, following the final iteration of the machine learning process.

2.4. Active learning loop

In a typical active learning process, the preparation of new data points, the training model, and determining the following sampling data points are major steps performed in sequence in each active learning loop, as shown in Fig. 1 [38]. Recent studies have demonstrated that pre-training strategies such as transfer learning and self-supervised learning can significantly improve model performance. Our framework integrates transfer learning from a foundational model to capitalize on these benefits [39].

When building the initial database, we randomly selected 400 catalyst–intermediate combinations and used these data to construct the initial dataset for active learning iterations. After constructing the initial dataset, about 200 data points were randomly held out as a fixed test dataset and were used to evaluate the model alongside the iterations. In each active learning iteration, we used a five-fold cross-validation to train the model for the rest of the data. All the data were randomly divided into five equal parts, with one part serving as the validation set and the remaining parts as the training set. The validation set was changed sequentially in each iteration.

For the training procedure, we utilized EquiformerV2 as our main model, as it achieves state-of-the-art performance on large-scale benchmarks, underscoring its capability to efficiently learn intricate patterns in data [40]. We utilized a fine-tuning paradigm leveraging the EquiformerV2 model with 31 million parameters pre-trained on the OC20+22 dataset (Tables S5 and S6 in Appendix A). This dataset is focused on similar adsorption energy prediction tasks but has a much broader scope of catalyst materials and adsorbates [41,42]. Our fine-tuning effort focused on our high-standard specialized adsorption energy dataset, where we aimed to refine the model's ability to predict the adsorption energy in our system with high precision. This was undertaken by implementing a full fine-tuning procedure, employing a significantly reduced learning rate to adjust the pre-trained weights. This approach ensured the preservation of the generic features learned from the larger dataset while making precise adjustments to optimize performance for specific adsorption energy predictions, increasing the model's applicability and accuracy in our targeted study.

After training, we used the average of the five model predictions to evaluate the model's performance and to predict the adsorption energies of all candidate catalyst–adsorbate systems. For model uncertainty estimation, the variance of the ensemble model was adopted to represent the uncertainty. The prediction and uncertainty can be expressed as follows:

$$\text{prediction} = \frac{1}{N}\sum_{t} y_t \tag{9}$$

$$\text{uncertainty} = \sqrt{\frac{1}{N}\sum_{t}\left(y_t - \bar{y}\right)^2} \tag{10}$$

where N, y_t, and $\bar{y}$ represent the total number of predictions, the prediction at the tth time, and the mean prediction value, respectively. This method is also called k-fold cross-validation model averaging. While retaining the simplicity of the training procedure, this method makes it possible to obtain a more stable prediction and better uncertainty estimation [43]. Samples with the highest uncertainty indicate that the model has less knowledge about those systems; thus, choosing the samples helps improve the model's performance.

In this study, we employed a mixed data-sampling strategy during each round of active learning, incorporating catalyst performance, data value, and data diversity to determine subsequent datapoints for DFT calculations. Catalyst performance was evaluated using the Pareto front concept, jointly optimizing activity (quantified by the absolute value of the limiting potential) and selectivity (represented by the energy difference between two competing product pathways). We assume two catalysts, CatA and CatB with the following relation:

$$\begin{aligned} &\text{Selectivity(CatA)} \leq \text{Selectivity(CatB)} \\ &\text{Activity(CatA)} \leq \text{Activity(CatB)} \end{aligned} \tag{11}$$

where CatB dominates CatA with at least one strict inequality, and the Pareto front is the collection of such non-dominated solutions. To identify promising candidates, we adopted an "onion-like" peeling approach: For each iteration, after determining the Pareto front, we removed the identified catalysts from the candidate pool and recalculated the front, iterating until 16 promising catalysts were selected. The Pareto front concept has been successfully applied in various self-driving laboratory platforms for experiment-based catalyst discovery [44,45], demonstrating that multi-objective optimization and on-the-fly validation can expedite the identification of novel, high-performance catalyst systems. After selecting the top catalysts, we continued sampling by choosing the 80 slab-adsorbate systems with the greatest uncertainty, identifying them as the most informative to the model, and randomly selecting 80 additional slab-adsorbate systems from the unexplored space to increase diversity. Altogether, these choices yielded approximately 400 new datapoints for DFT calculations. Finally, as discovered by Chen et al. [46], pre-optimizing geometries can improve predictive performance. Consequently, after completing the active learning iterations, data from prior DFT calculations were used to train a machine learning force field, leveraging the EquiformerV2 31M model to further improve model performance.

2.5. Model evaluation

The main evaluation metrics for the model include the coefficient of determination (R^2), the mean absolute error (MAE), and the root mean squared error (RMSE), which are defined in Eqs. (12)–(14):

$$R^2 = 1 - \frac{\sum_s (y_s - \hat{y}_s)^2}{\sum_s \left(\hat{y}_s - \bar{y}_s\right)^2} \tag{12}$$

$$\text{MAE} = \frac{1}{N}\sum_{s} |y_s - \hat{y}_s| \tag{13}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{s} (y_s - \hat{y}_s)^2}{N}} \tag{14}$$

where $\hat{y}_s$ represents the actual value of the sample s, y_s represents the prediction value of sample s, and $\bar{y}_s$ represents the average actual value of sample s. The MAE and RMSE provide an intuitive estimate of the magnitude of the model's error. At the same time, the coefficient of determination evaluates the model's goodness of fit, with values closer to 1 indicating a better fit between the model and the actual values and thus indicating better model performance.

3. Results and discussion

3.1. Design space determination

M-X_3 and M-X_4 catalysts are graphene-based SACs. Their structures usually consist of a single central metal atom surrounded by three or four coordinated environment metals located on graphene defects. In this work, 29 studied metals [9,47] were selected as the central metals for SACs, combined with coordinate C, N, O, S, and P atoms, resulting in 3045 catalysts as the ideal catalyst space. However, such a design space would produce numerous unstable structures due to the structural mismatch between the metal atom and the doped vacancy, and calculations for these catalysts would eventually produce unreasonable results, contaminating the dataset used for training purposes. Therefore, before building the initial dataset, we performed a screen for unstable catalysts by calculating their binding energies. By filtering out the catalysts with a positive binding energy and those with significant displacement after geometry relaxation, we were able to build a relatively reasonable design space and avoid unnecessary calculations in later active learning loops (Figs. 2(a) and (b)).

The results showed that over half of the enumerated catalysts were unlikely to be thermodynamically stable because of positive binding energies. Among 3007 converged calculations, only 994 catalysts had a negative binding energy (Figs. 2(c) and (d)). After further screening out those with significant displacement, only 956 catalysts were reserved in our design space (Figs. S3–S5 in Appendix A). However, we must emphasize that this screening procedure is only a primary screening for potentially suitable candidate catalysts that makes our design space more reasonable. The judgment of stability is a more complicated problem that requires an investigation beyond simple binding energy criteria.

3.2. Model performance evolution through active learning cycles

Due to the high computational cost of DFT calculations, the amount of data in the database will be limited. Effective active learning strategies can significantly increase the amount of valuable information contained in the collected data, thereby addressing the challenges posed by the high cost of data acquisition [38]. Fig. 1 illustrates the active learning process used in our framework. The general idea is to train the model on the existing dataset, select the next set of samples to be labeled using a query function in active learning, incorporate the newly labeled data into the dataset, and then train the machine learning model through retraining to improve the model's performance. This process is iteratively repeated until the final model is obtained. Along with this process, promising catalysts are recommended iteratively and validated on the fly, so the learning progress can be terminated at any time if the budget reaches its limitation.

To balance model universality with performance, we employed the state-of-the-art EquiformerV2 model in this work (Fig. 3(a)) [40]. It is a graph-neural-network-based machine learning model that accepts the atom-level structure as the input. In our study, which focuses on SAC screening, the inherent challenge posed by small-dataset scenarios is significant, particularly in achieving high predictive accuracy and robust model performance. To address this challenge, we employed a fine-tuning paradigm that capitalizes on the extensive knowledge embedded within large pre-existing datasets.

In our study, a *k*-fold model averaging method is implemented within the active learning iteration process, with the aim of providing a dual output: a prediction of the adsorption energy and an uncertainty estimate for each prediction. This method involves using the average output of an ensemble model as the definitive model prediction. To validate the robustness of our model and the validity of the active learning procedure, an invariant test set was held out comprising approximately 200 data points; this test set remained untouched during the training phases of all active learning iterations. Specifically, the OC20 and OC22 datasets [41,42], which encompass comprehensive data on catalyst adsorption energies, set the foundation for our model fine-tuning.

As shown in Fig. 3(b), during the active learning process, as the model progressively assimilates more detailed information regarding the structures and adsorption energies, a noticeable performance improvement is observed in the whole process, with a decrease from 0.5 to 0.3 eV in MAE (Fig. S6 in Appendix A). By initializing our model with weights pre-trained on the large dataset, we leveraged the already learned diverse and complex patterns, allowing the model to perform better in our small dataset scenario (Figs. S7 and S8 in Appendix A). In our research, we found that this strategy significantly impacted the model performance, with an improvement in performance, with the average MAE from 0.45 to 0.30 eV, tested with the final round of the dataset (Figs. 3 (c) and (d)). After completing six iterations, we reach the limit of our computational budget. The final assessment of the model shows a prediction accuracy with an R^2 value of 0.872 and MAE and RMSE values of 0.308 and 0.540 eV, respectively. The results highlight our commitment to increasing the accuracy and reliability of adsorption energy predictions from deep learning models through the iterative active learning framework. The final dataset distribution is shown in Appendix A Figs. S9–S11.

3.3. Reaction network construction and active catalyst exploration

After gaining a trained machine learning model, the prediction of adsorption energies can be done quickly for all candidate materials. In this study, we investigated the relevant literature and propose the reaction pathway network [7,48,49] shown in Fig. 4(a), which includes pathways for four major products of CO_2RR: carbon dioxide, formic acid, methanol, and methane. Furthermore, we consider another side reaction: the HER, which often occurs when developing CO_2 reduction catalysts [50].

In this work, we use the limiting potential as the indicator for the activity and selectivity for the target product, as it has been used in computational screening works [24,36]. The limiting potential is defined as the lowest potential when an electrochemical reaction takes place and is often used to evaluate the performance of catalysts from a thermodynamic perspective. By utilizing the trained adsorption energy machine learning model, the limiting potential for all products of all candidate catalysts could be predicted within just minutes, making it possible to explore the vast candidate space. Usually, a catalyst with a lower limiting potential is favored when selecting a specific catalyst [18,47,51]. The difference in limiting potentials among various target products can also indicate the selectivity of the catalyst [24], thus guiding catalyst screening with the desired target product and avoiding side reactions. More details about the screening process are provided in Section 2.

Using the machine learning model, we predict the candidate catalyst limiting potential for methanol as an example (Fig. 4(b)), to illustrate the activity of catalysts toward electroreduction to methanol. Predictions for other products can be found in Appendix A Figs. S12–S16. It should be noted that quite a few catalysts showed a similar limiting potential for various products and, due to the prediction error in adsorption energies when applying machine learning methods, this diagram might be inaccurate when used alone. In the future, a strategy to utilize the machine learning prediction results should be further developed, especially considering the prediction errors. When designing novel materials, such prediction trends could guide researchers toward an initial guess for the dopant metals in catalysts; further DFT validation should be performed before experiments.

3.4. Catalyst discovery with deep learning models

Considering the inherent errors associated with machine learning predictions, our research strategically employs on-the-fly validation in conjunction with DFT calculations. This makes it possible to iteratively refine and explore potential catalysts in order to ensure the material properties are screened effectively and efficiently, even if the prediction precision from the machine learning model is insufficient to make accurate predictions. In prior studies, SACs—specifically those with a CuN_4 local environment composition—have demonstrated significant efficacy in CO_2RR and have successfully produced methanol [52]. Building upon these foundational insights, our work employs the well-characterized CuN_4 SAC as a baseline. By focusing on the limiting potential of methanol as

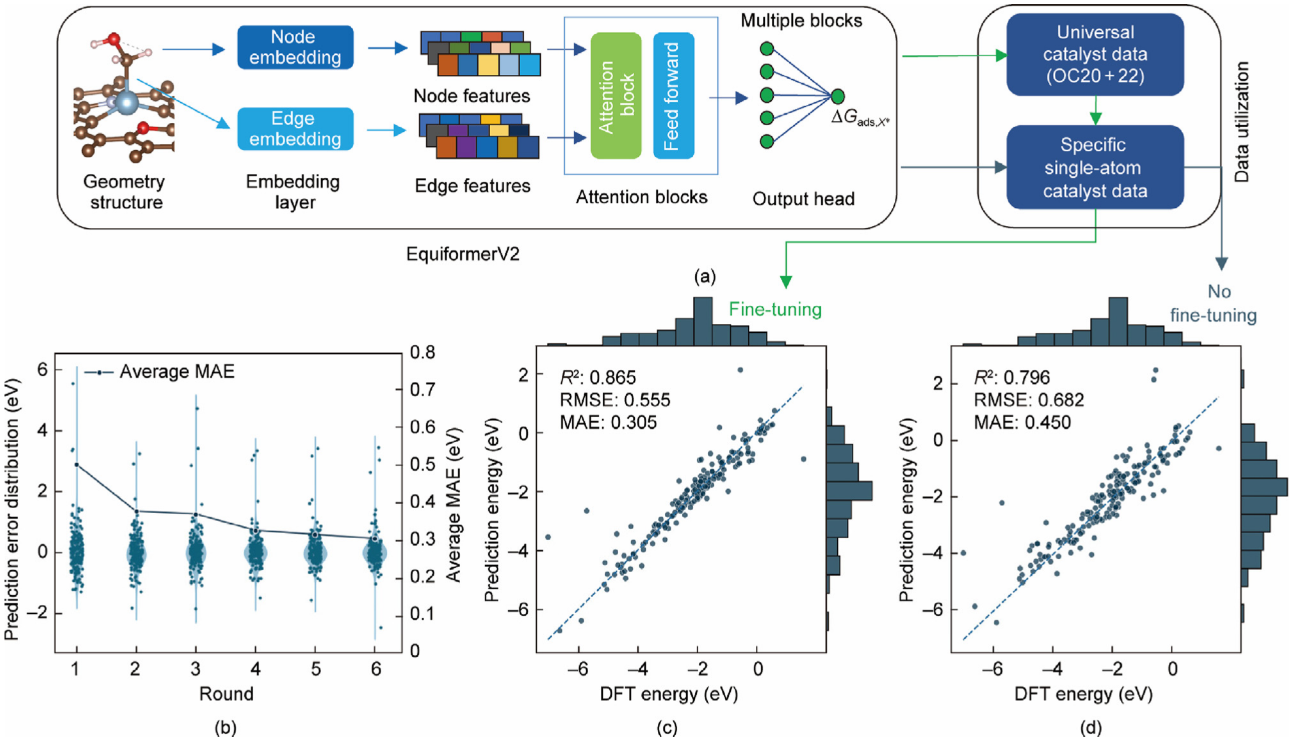

Fig. 3. Active learning and model performance. (a) Illustration of the EquiformerV2 architecture, where catalyst structure is encoded by atom types and positions and processed through equivariant graph attention to generate a new representation from which adsorption energy is predicted. (b) Changes in the performance and distribution of prediction errors throughout active learning. (c, d) Performance of the ensemble model in the final active learning round, and a comparison of (c) fine-tuning versus (d) no fine-tuning.

the indicator for activity and the differential limiting potentials between methanol and other products as the indicator for selectivity, we use these criteria to evaluate the catalyst performance found by our machine learning models, and raise an optimization problem involving two competing objectives. To solve this problem, we utilize the multi-objective optimization concept of the Pareto front, ensuring that improvements in one objective do not come at the expense of another, to identify catalysts that offer an optimal balance between activity and selectivity. By focusing on non-dominated catalysts across the multiple performance metrics recommended by the machine learning models, we can effectively push the boundaries of what is feasible within current catalytic systems.

After five rounds of validation calculations, one of our best candidate catalysts, Pd-C_3P, shows an improvement of approximately 1 V in its limiting potential for the desired product; this is coupled with a simultaneous decrease of 0.3 V in the potential difference to the primary side product in the DFT simulation (Fig. 5(a)). Additionally, we identified three additional novel catalysts in the Pareto front—Al-N_2O, Ga-C_3S, and Cd-N_4—that also demonstrated relatively smoother reaction steps in their elementary reactions. This significant increase in selectivity and activity underscores the efficacy of our machine-learning-guided approach. Furthermore, we used the machine learning predictions to recommend catalyst candidates at $U_{app} = -0.88$ V and $U_{app} = -1.38$ V. From the newly recommended top-10 catalysts, seven and six candidates, respectively, converged in the subsequent DFT calculations, and the results confirmed their performance improvements. (Figs. 5(b)–(e)). These findings demonstrate that machine learning is an effective preliminary tool for predicting trends in catalyst optimization, highlighting its potential in the iterative optimization process.

However, we note differences between the DFT simulations and the experimental results (Fig. S17 and Table S7 in Appendix A). Recent theoretical investigations have emphasized the critical importance of incorporating solvation effects, potential-dependent adsorption, and the dynamic evolution of catalysts under operational conditions, which are not considered in this work that uses the traditional CHE model. Addressing these factors in future studies will significantly increase the precision and applicability of the DFT method, thereby aligning the theoretical predictions more closely with experimental observations [53,54]. Crucially, bridging the gap between simplified DFT predictions and experimental data can be further facilitated by integrating multiple data sources into advanced machine learning frameworks, such as a two-stage hierarchical approach, where the predicted adsorption energies serve as features for a second model trained against the experimental targets [55]. In parallel, conceptual validation of the proposed catalysts can be pursued by benchmarking simpler representative models, continuing comparisons within established reaction networks for CO_2RR process, and conducting detailed microkinetic analyses of reaction pathways. These approaches could offer indirect support for predictions involving more structurally intricate or novel catalysts. To summarize, advancements in more accurate simulation methods to achieve chemical precision, together with machine learning models that capture the complex relationship between structure and performance, will be crucial for bridging the gap between simulation and experimental data. These combined efforts will pave the way for next-generation catalyst design driven by artificial intelligence methods.

3.5. Uncovering design principles through subgroup discovery

Considering the complex relation of electronic structure to catalyst performance (Figs. S18 and S19 in Appendix A), using the database constructed through active learning, we applied

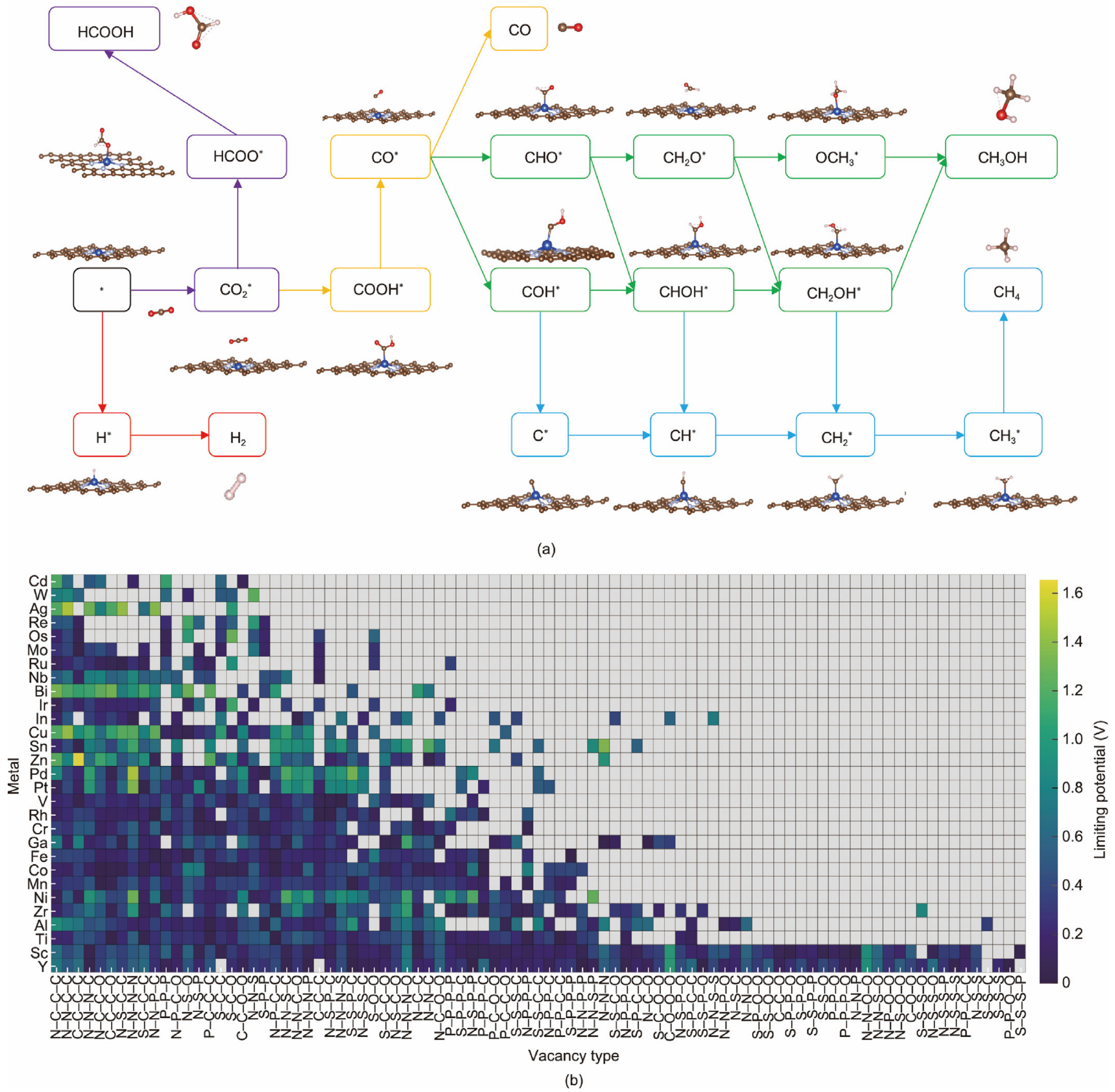

Fig. 4. Utilizing the model obtained from active learning to predict the limiting potential. (a) Illustration of the reaction network analyzed in this study. (b) Prediction of limiting potentials for methanol production with the model from the last round of active learning. All values presented are calculated under U_{app} = −0.38 V.

subgroup discovery (SGD) to identify the optimal range of adsorption energies for catalyst design. For the activity, determined by the limiting potential to CH_3OH, we found that the optimal adsorption energy ranges were between −1.76 and −1.31 eV for COOH*, and between −2.31 and −1.66 eV for HCOO*. The subgroup identified by these ranges had an average limiting potential of 0.364 V, compared with the dataset average of 0.651 V, indicating improved catalytic performance. For selectivity, the optimal adsorption energy ranges were between −3.64 and −3.26 eV for CH*, and less than −2.44 eV for H*. The selectivity—defined by the potential difference to the primary side product—was 0.051 V for the subgroup, compared with 0.277 V for the entire dataset (Fig. 6). These findings provide a general design principle for adsorption energy tuning, thereby guiding the rational design of graphene-based SACs with increased activity and selectivity for CO_2 electroreduction to methanol. Moreover, we furthered this smaller dataset within the optimal range of adsorption energy to fine-tune our model (Fig. S20 in Appendix A). Experiments with this smaller dataset suggested that predicting within the subgroup can reduce prediction errors. However, determining the optimal adsorption energy range still requires substantial prior effort.

4. Conclusions

In conclusion, this work introduced a method that integrates DFT calculations, deep learning models, and active learning

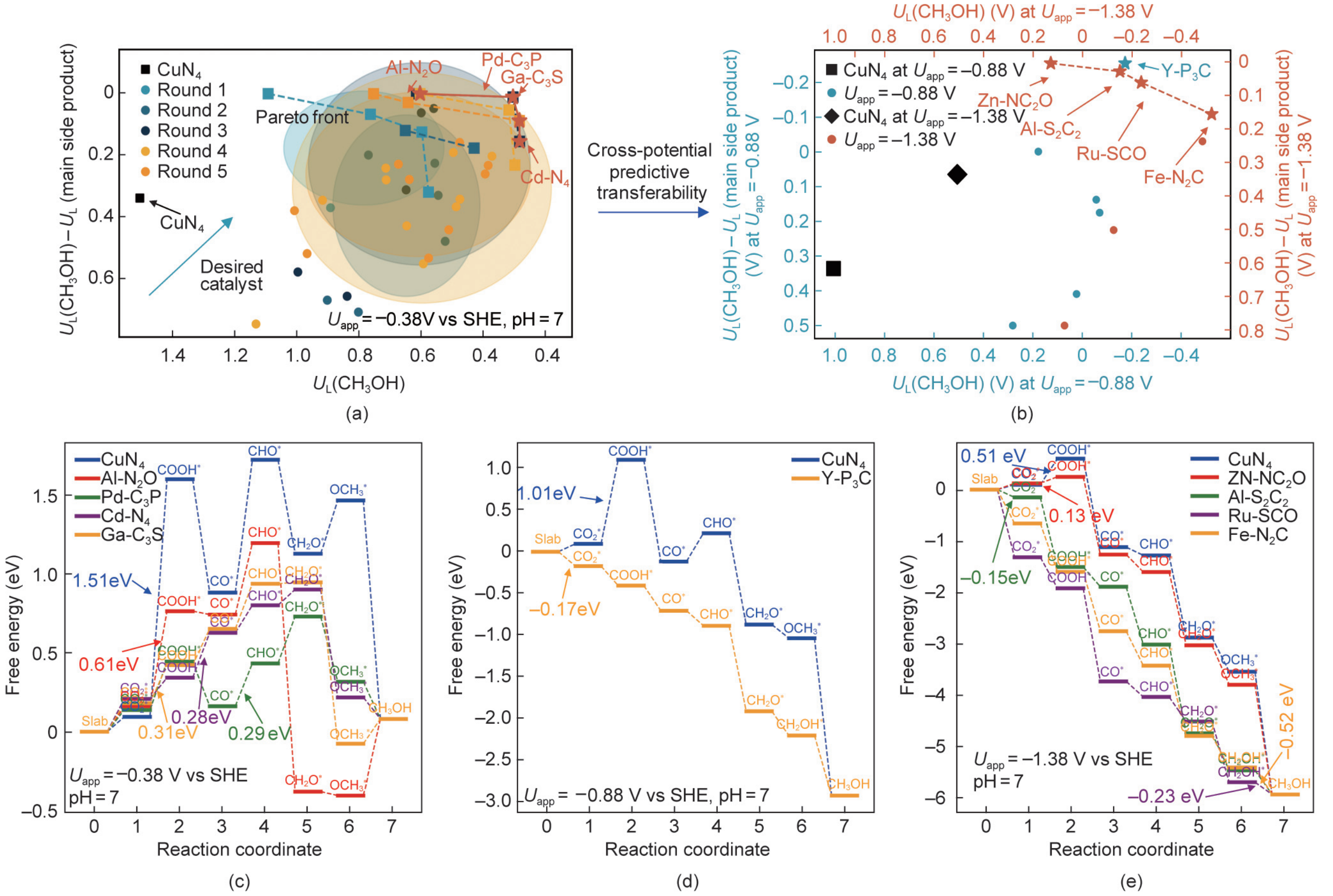

Fig. 5. Catalyst discovery via machine learning. (a) Visualization of the iterative refinement process in machine-learning-driven catalyst discovery under U_{app} = −0.38 V, showing the Pareto boundary and prediction range for each iteration. The experimentally validated CuN$_4$ catalyst and those identified on the Pareto front are denoted as a square, while the aggregate Pareto front across all iterations is denoted by stars. (b) DFT validation of the extrapolated prediction result under different potentials (U_{app} = −0.88 V and U_{app} = −1.38 V, respectively). (c–e) Gibbs energy profile for the machine learning model's recommended catalyst versus the CuN$_4$ catalyst.

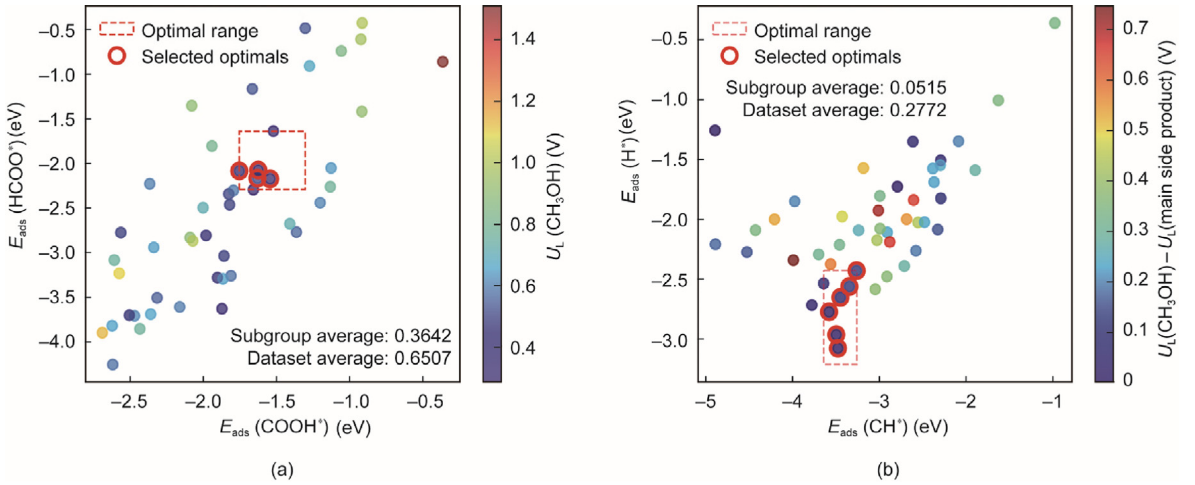

Fig. 6. SGD result for the optimal range of adsorption energies under U_{app} = −0.38 V. (a) SGD result for activity; (b) SGD result for selectivity. E_{ads}: adsorption energy for the molecule.

strategies to rapidly screen high-performance CO$_2$RR catalysts. By employing deep active learning, we were able to efficiently explore a broad catalyst space, facilitating the targeted design of CO$_2$RR catalysts. Developing a comprehensive database encompassing various catalyst structures and intermediate species enabled fast predictions with acceptable accuracy in adsorption energies and limiting potentials. This database also provides flexibility for application across diverse structures and facilitates detailed mechanistic investigations using machine learning models; it is made publicly available alongside this work. The optimally fine-tuned model demonstrated robust performance, achieving an R^2 of 0.865 and an MAE of 0.305 eV. It successfully identified several novel catalysts that improve the limiting potential by approximately 1 V for the desired product while reducing it by 0.3 V for the main side product. The predictions results were also validated under different applied potentials, showing the effectiveness of the method presented in this work when extrapolating results into different scenarios. Furthermore, by utilizing the SGD method, we uncovered an optimal range of adsorption energies, provided valuable insights into possible design strategies, and pointed to areas for future improvement.

Moving forward, several key challenges must be addressed to increase the predictive accuracy and reliability of the models used. Developing more precise computational methods—including accurate theoretical models that account for potential dependency, solvation effects, and structural evolution—is essential to bridge the gap between theoretical predictions and experimental results. These factors can sometimes be decisive and must not be overlooked, even during the primary screening stages. For machine learning models, constructing highly accurate frameworks that can investigate transition states and energy barriers and integrate experimental data to bridge the gap between simulation and experimental data is crucial for a more nuanced understanding of catalytic mechanisms. Moreover, a stability evaluation of SACs, which is more complex than the preliminary screening performed here, requires deeper investigation.

Finally, a major obstacle we face lies in further reducing the errors brought by the machine learning models, particularly in complex and broad catalyst-adsorbate landscapes. One critical challenge involves effectively handling out-of-distribution (OOD) samples, which arise when certain structures or adsorbate configurations deviate substantially from the training set. By integrating strategies such as uncertainty-based sampling, distribution-focused sampling, and structure-aware modeling, future work can more systematically identify and incorporate these rare and complex configurations into the training process. Moreover, expanding the training dataset—especially via active learning that prioritizes OOD or high-uncertainty samples—promises to further drive down prediction errors. Increased data diversity not only reduces overfitting but also better reflects the real-world variety of catalytic surfaces and intermediates. Hyperparameter tuning is another key facet of model improvement. Although larger and deeper network architectures can yield performance gains, they also have higher computational costs, making careful resource–benefit trade-offs essential. Alongside these technical optimizations, integrating expert knowledge into machine learning pipelines can meaningfully augment performance. Embedding domain-specific insights—such as regarding electronic structures, geometrical symmetry, or known catalytic descriptors—helps guide the model toward more physically accurate predictions [56]. In particular, multi-task and transfer learning approaches can leverage relationships among various material properties to improve model generality and robustness [57]. Additionally, mixture-of-experts (MoE) models offer a framework in which specialized sub-networks address distinct sub-tasks (e.g., dissociative versus non-dissociative adsorbates), thereby increasing both accuracy and interpretability [58]. Despite the challenges described above, our approach offers a valuable universal framework for accelerating the discovery of high-performance CO_2RR catalysts and opens new avenues for the rational design of new catalysts.

CRediT authorship contribution statement

Honghao Chen: Visualization, Software, Data curation, Writing – original draft, Validation, Investigation, Methodology, Formal analysis, Conceptualization. **Jun Yin:** Investigation, Methodology. **Jiali Li:** Methodology. **Xiaonan Wang:** Resources, Funding acquisition, Supervision, Conceptualization, Writing – review & editing, Methodology.

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

This work was supported by the National Key Research and Development Program of China (2022ZD0117501), the Scientific Research Innovation Capability Support Project for Young Faculty (ZYGXQNJSKYCXNLZCXM-E7), and the Tsinghua University Initiative Scientific Research Program and the Carbon Neutrality and Energy System Transformation (CNEST) Program led by Tsinghua University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eng.2025.03.039.

References

[1] Meinshausen M, Lewis J, McGlade C, Gütschow J, Nicholls Z, Burdon R, et al. Realization of Paris Agreement pledges may limit warming just below 2 °C. Nature 2022;604(7905):304–9.
[2] Gao W, Liang S, Wang R, Jiang Q, Zhang Y, Zheng Q, et al. Industrial carbon dioxide capture and utilization: state of the art and future challenges. Chem Soc Rev 2020;49(23):8584–686.
[3] Jin S, Hao Z, Zhang K, Yan Z, Chen J. Advances and challenges for the electrochemical reduction of CO_2 to CO: from fundamentals to industrialization. Angew Chem 2021;133(38):20795–816.
[4] Xu D, Li K, Jia B, Sun W, Zhang W, Liu X, et al. Electrocatalytic CO_2 reduction towards industrial applications. Carbon Energy 2023;5(1):e230.
[5] Lin R, Guo J, Li X, Patel P, Seifitokaldani A. Electrochemical reactors for CO_2 conversion. Catalysts 2020;10(5):473.
[6] Kibria MG, Edwards JP, Gabardo CM, Dinh CT, Seifitokaldani A, Sinton D, et al. Electrochemical CO_2 reduction into chemical feedstocks: from mechanistic electrocatalysis models to system design. Adv Mater 2019;31(31):1807166.
[7] Zhang J, Cai W, Hu FX, Yang H, Liu B. Recent advances in single atom catalysts for the electrochemical carbon dioxide reduction reaction. Chem Sci 2021;12(20):6800–19.
[8] Tan X, Yu C, Ren Y, Cui S, Li W, Qiu J. Recent advances in innovative strategies for the CO_2 electroreduction reaction. Energy Environ Sci 2021;14(2):765–80.
[9] Li M, Wang H, Luo W, Sherrell PC, Chen J, Yang J. Heterogeneous single-atom catalysts for electrochemical CO_2 reduction reaction. Adv Mater 2020;32(34):2001848.
[10] Wei J, Chu X, Sun XY, Xu K, Deng HX, Chen J, et al. Machine learning in materials science. InfoMat 2019;1(3):338–58.
[11] Cai J, Chu X, Xu K, Li H, Wei J. Machine learning-driven new material discovery. Nanoscale Adv 2020;2(8):3115–30.
[12] Duan C, Nandy A, Kulik HJ. Machine learning for the discovery, design, and engineering of materials. Annu Rev Chem Biomol Eng 2022;13(1):405–29.
[13] Yan Y, Borhani TN, Subraveti SG, Pai KN, Prasad V, Rajendran A, et al. Harnessing the power of machine learning for carbon capture, utilisation, and storage (CCUS)—a state-of-the-art review. Energy Environ Sci 2021;14(12):6122–57.
[14] Zhong M, Tran K, Min Y, Wang C, Wang Z, Dinh CT, et al. Accelerated discovery of CO_2 electrocatalysts using active machine learning. Nature 2020;581(7807):178–83.
[15] Chen Y, Huang Y, Cheng T, Goddard WA III. Identifying active sites for CO_2 reduction on dealloyed gold surfaces by combining machine learning with multiscale simulations. J Am Chem Soc 2019;141(29):11651–7.
[16] Cheng D, Zhao ZJ, Zhang G, Yang P, Li L, Gao H, et al. The nature of active sites for carbon dioxide electroreduction over oxide-derived copper catalysts. Nat Commun 2021;12:395.
[17] Chen A, Zhang X, Chen L, Yao S, Zhou Z. A machine learning model on simple features for CO_2 reduction electrocatalysts. J Phys Chem C 2020;124(41):22471–8.
[18] Wan X, Zhang Z, Niu H, Yin Y, Kuai C, Wang J, et al. Machine-learning-accelerated catalytic activity predictions of transition metal phthalocyanine dual-metal-site catalysts for CO_2 reduction. J Phys Chem Lett 2021;12(26):6111–8.

[19] Hu E, Liu C, Zhang W, Yan Q. Machine learning assisted understanding and discovery of CO_2 reduction reaction electrocatalyst. J Phys Chem C 2023;127(2):882–93.

[20] Tamtaji M, Chen S, Hu Z, Goddard III WA, Chen G. A surrogate machine learning model for the design of single-atom catalyst on carbon and porphyrin supports towards electrochemistry. J Phys Chem C 2023;127(21):9992–10000.

[21] Kresse G, Furthmüller J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. Phys Rev B 1996;54(16):11169–86.

[22] Wang Y, You L, Zhou K. Origin of the N-coordinated single-atom Ni sites in heterogeneous electrocatalysts for CO_2 reduction reaction. Chem Sci 2021;12(42):14065–73.

[23] Hossain MD, Huang Y, Yu TH, Goddard III WA, Luo Z. Reaction mechanism and kinetics for CO_2 reduction on nickel single atom catalysts from quantum mechanics. Nat Commun 2020;11:2256.

[24] Wang S, Li L, Li J, Yuan C, Kang Y, Hui KS, et al. High-throughput screening of nitrogen-coordinated bimetal catalysts for multielectron reduction of CO_2 to CH_4 with high selectivity and low limiting potential. J Phys Chem C 2021;125(13):7155–65.

[25] Nørskov JK, Rossmeisl J, Logadottir A, Lindqvist L, Kitchin JR, Bligaard T, et al. Origin of the overpotential for oxygen reduction at a fuel-cell cathode. J Phys Chem B 2004;108(46):17886–92.

[26] Kresse G, Joubert D. From ultrasoft pseudopotentials to the projector augmented-wave method. Phys Rev B 1999;59(3):1758–75.

[27] Blöchl PE. Projector augmented-wave method. Phys Rev B 1994;50(24):17953–79.

[28] Perdew JP, Burke K, Ernzerhof M. Generalized gradient approximation made simple. Phys Rev Lett 1996;77(18):3865–8.

[29] Grimme S, Antony J, Ehrlich S, Krieg H. A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. J Chem Phys 2010;132(15):154104.

[30] Perdew JP, Wang Y. Pair-distribution function and its coupling-constant average for the spin-polarized electron gas. Phys Rev B 1992;46(20):12947–54.

[31] Wang V, Xu N, Liu JC, Tang G, Geng WT. VASPKIT: a user-friendly interface facilitating high-throughput computing and analysis using VASP code. Comput Phys Commun 2021;267:108033.

[32] Granda-Marulanda LP, Rendón-Calle A, Builes S, Illas F, Koper MTM, Calle-Vallejo F. A semiempirical method to detect and correct DFT-based gas-phase errors and its application in electrocatalysis. ACS Catal 2020;10(12):6900–7.

[33] Ong SP, Richards WD, Jain A, Hautier G, Kocher M, Cholia S, et al. Python Materials Genomics (PyMatGen): a robust, open-source python library for materials analysis. Comput Mater Sci 2013;68:314–9.

[34] Di Liberto G, Giordano L, Pacchioni G. Predicting the stability of single-atom catalysts in electrochemical reactions. ACS Catal 2024;14(1):45–55.

[35] Tran K, Ulissi ZW. Active learning across intermetallics to guide discovery of electrocatalysts for CO_2 reduction and H_2 evolution. Nat Catal 2018;1(9):696–703.

[36] Kulkarni A, Siahrostami S, Patel A, Nørskov JK. Understanding catalytic activity trends in the oxygen reduction reaction. Chem Rev 2018;118(5):2302–12.

[37] Peterson AA, Abild-Pedersen F, Studt F, Rossmeisl J, Nørskov JK. How copper catalyzes the electroreduction of carbon dioxide into hydrocarbon fuels. Energy Environ Sci 2010;3(9):1311–5.

[38] Lookman T, Balachandran PV, Xue D, Yuan R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. npj Comput Mater 2019;5:21.

[39] Cui T, Tang C, Su M, Zhang S, Li Y, Bai L, et al. Geometry-enhanced pretraining on interatomic potentials. Nat Mach Intell 2024;6(4):428–36.

[40] Liao YL, Wood B, Das A, Smidt T. EquiformerV2: improved equivariant transformer for scaling to higher-degree representations. 2023. arXiv:2306.12059.

[41] Chanussot L, Das A, Goyal S, Lavril T, Shuaibi M, Riviere M, et al. Open catalyst 2020 (OC20) dataset and community challenges. ACS Catal 2021;11(10):6059–72.

[42] Tran R, Lan J, Shuaibi M, Wood BM, Goyal S, Das A, et al. The open catalyst 2022 (OC22) dataset and challenges for oxide electrocatalysts. ACS Catal 2023;13(5):3066–84.

[43] Dutschmann TM, Kinzel L, ter Laak A, Baumann K. Large-scale evaluation of *k*-fold cross-validation ensembles for uncertainty estimation. J Cheminform 2023;15:49.

[44] Suvarna M, Zou T, Chong SH, Ge Y, Martín AJ, Pérez-Ramírez J. Active learning streamlines development of high performance catalysts for higher alcohol synthesis. Nat Commun 2024;15:5844.

[45] Bennett JA, Orouji N, Khan M, Sadeghi S, Rodgers J, Abolhasani M. Autonomous reaction Pareto-front mapping with a self-driving catalysis laboratory. Nat Chem Eng 2024;1(3):240–50.

[46] Chen L, Tian Y, Hu X, Yao S, Lu Z, Chen S, et al. A universal machine learning framework for electrocatalyst innovation: a case study of discovering alloys for hydrogen evolution reaction. Adv Funct Mater 2022;32(47):2208418.

[47] Wang X, Niu H, Liu Y, Shao C, Robertson J, Zhang Z, et al. Theoretical investigation on graphene-supported single-atom catalysts for electrochemical CO_2 reduction. Cat Sci Technol 2020;10(24):8465–72.

[48] Nitopi S, Bertheussen E, Scott SB, Liu X, Engstfeld AK, Horch S, et al. Progress and perspectives of electrochemical CO_2 reduction on copper in aqueous electrolyte. Chem Rev 2019;119(12):7610–72.

[49] Todorova TK, Schreiber MW, Fontecave M. Mechanistic understanding of CO_2 reduction reaction (CO_2RR) toward multicarbon products by heterogeneous copper-based catalysts. ACS Catal 2020;10(3):1754–68.

[50] Zhang YJ, Sethuraman V, Michalsky R, Peterson AA. Competition between CO_2 reduction and H_2 evolution on transition-metal electrocatalysts. ACS Catal 2014;4(10):3742–8.

[51] Back S, Lim J, Kim NY, Kim YH, Jung Y. Single-atom catalysts for CO_2 electroreduction with significant activity and selectivity improvements. Chem Sci 2017;8(2):1090–6.

[52] Yang H, Wu Y, Li G, Lin Q, Hu Q, Zhang Q, et al. Scalable production of efficient single-atom copper decorated carbon membranes for CO_2 electroreduction to methanol. J Am Chem Soc 2019;141(32):12717–23.

[53] Bai X, Zhao X, Zhang Y, Ling C, Zhou Y, Wang J, et al. Dynamic stability of copper single-atom catalysts under working conditions. J Am Chem Soc 2022;144(37):17140–8.

[54] Cao H, Zhang Z, Chen JW, Wang YG. Potential-dependent free energy relationship in interpreting the electrochemical performance of CO_2 reduction on single atom catalysts. ACS Catal 2022;12(11):6606–17.

[55] Abed J, Kim J, Shuaibi M, Wander B, Duijf B, Mahesh S, et al. Open catalyst experiments 2024 (OCx24): bridging experiments and computational models. 2024. arXiv:2411.11783.

[56] Deng B, Zhong P, Jun K, Riebesell J, Han K, Bartel CJ, et al. CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. Nat Mach Intell 2023;5(9):1031–41.

[57] Zhang D, Liu X, Zhang X, Zhang C, Cai C, Bi H, et al. DPA-2: a large atomic model as a multi-task learner. npj Comput Mater 2024;10:293.

[58] Chang R, Wang YX, Ertekin E. Towards overcoming data scarcity in materials science: unifying models and datasets with a mixture of experts framework. npj Comput Mater 2022;8:242.