

# Convolutional Neural Networks and Volcano Plots: Screening and Prediction of Two-Dimensional Single-Atom Catalysts for CO<sub>2</sub> Reduction Reactions

Haoyu Yang<sup>1</sup>, Juanli Zhao<sup>2</sup>, Qiankun Wang<sup>3</sup>, Bin Liu<sup>2</sup>, Wei Luo<sup>4</sup>,  
Ziqi Sun<sup>1,\*</sup>, and Ting Liao<sup>5,\*</sup>

<sup>1</sup>*School of Chemistry and Physics, Queensland University of  
Technology, George Street, Brisbane, QLD 4000, Australia*

<sup>2</sup>*School of Materials Science and Engineering, Shanghai University,  
Shanghai, China*

<sup>3</sup>*School of Computer Science and Engineering, Southeast  
University, Nanjing, China*

<sup>4</sup>*College of Materials Science and Engineering, Institute of  
Functional Materials, Donghua University, Shanghai, China*

<sup>5</sup>*School of Mechanical Medical and Process Engineering,  
Queensland University of Technology, George Street, Brisbane,  
QLD 4000, Australia*

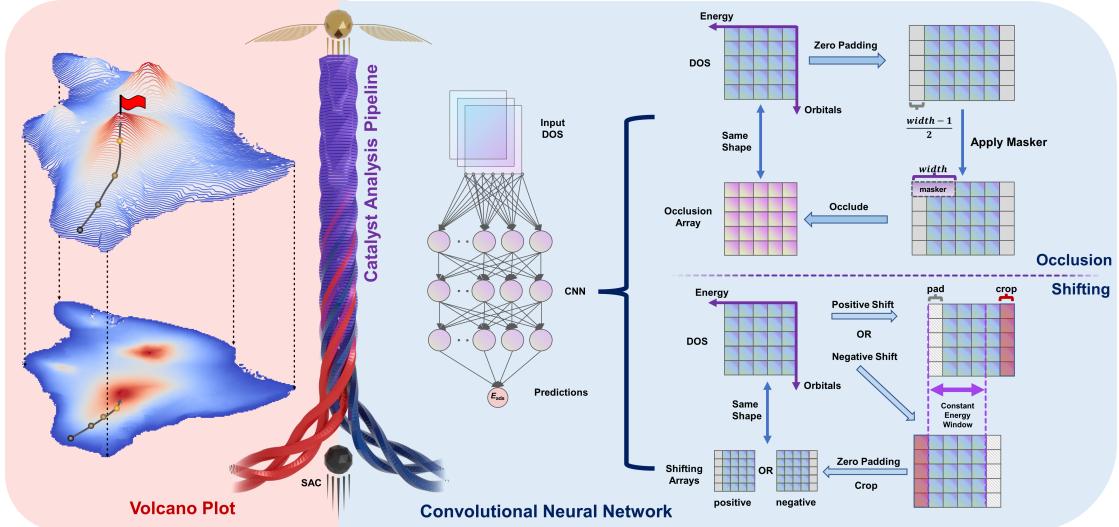
\*Corresponding author: ziqi.sun@qut.edu.au

\*Corresponding author: t3.liao@qut.edu.au

## Abstract

Single-atom catalysts (SACs) have emerged as pivotal frontiers for catalyzing a myriad of chemical reactions, yet the diverse combinations of active elements and support materials, the nature of coordination environments, elude traditional methodologies in searching optimal SAC systems with superior catalytic performance. Herein, by integrating multi-branch Convolutional Neural Network (CNN) analysis models to hybrid descriptor based activity volcano plot, two-dimensional (2D) SAC system composed of diverse metallic single atoms anchored on six type of 2D supports, including graphitic carbon nitride ( $\text{g-C}_3\text{N}_4$ ), nitrogen-doped graphene, graphene with dual-vacancy, black phosphorous, boron nitride (BN), and  $\text{C}_2\text{N}$ , are screened for efficient  $\text{CO}_2\text{RR}$ . Starting from establishing a correlation map between the adsorption energies of intermediates and diverse electronic and elementary descriptors, sole singular descriptor lost magic to predict catalytic activity, including d-band center. Deep learning method utilizing multi-branch CNN model therefore was employed, using 2D electronic density of states (eDOS) as input to predict adsorption energies. Hybrid-descriptor enveloping both C- and O-types of  $\text{CO}_2\text{RR}$  intermediates was introduced to construct volcano plots and limiting potential periodic table, aiming for intuitive screening of catalyst candidates for efficient  $\text{CO}_2$  reduction to  $\text{CH}_4$ . The eDOS occlusion experiments were performed to unravel individual orbital contribution to adsorption energy. To explore the electronic scale principle governing practical engineering catalytic  $\text{CO}_2\text{RR}$  activity, orbitalwise eDOS shifting experiments based on CNN model were employed. The study involves examining the adsorption energy and, consequently, catalytic activities while varying supported single atoms. This work offers a tangible framework to guide both theoretical screening and experimental synthesis, thereby paving the way for systematically designing efficient SACs.

# 1 Introduction



**Figure 1: Catalyst performance analysis pipeline integrating volcano plots and CNN.** The figure demonstrates the integrated usage of volcano plots for predictive assessment of existing catalysts, and the utility of CNNs to predict and modulate adsorption energies  $E_{\text{ads}}$  using eDOS as input. eDOS occlusion and shifting experiments are conducted to extract chemical information from CNN model.

Electrochemical CO<sub>2</sub>RR stands as a promising avenue for long-term seasonal energy storage [13] and mitigating excess CO<sub>2</sub> in the atmosphere.

Despite decades of research, metallic copper remains a leading CO<sub>2</sub>RR electrocatalyst for hydrocarbon production [47], albeit burdened by substantial overpotential and low faradic efficiency and poor operational stability [8, 39]. In this regard, supported SACs have displayed remarkable activity and selectivity across various catalytic reactions [58, 64], leveraging their unique electronic properties and coordination environment. To date, experimental studies have primarily focused on two-electron reduction products due to the facile release of gaseous CO [5, 26, 51]. Consequently, there is a pressing need to modulate the electronic structure of single-atom sites to enable hydrocarbon production at reduced overpotential.

The challenge towards effective catalysts design lies in understanding the reaction mechanisms. Experimental determination of these mechanisms remains challenging due to the complexity of *in situ* spectroscopic detection of reaction intermediates [69]. Computational modeling, particularly methods based on quantum mechanics (QM) such as density functional theory (DFT), allows for the profiling

the energetics of intermediates and unveils structure-activity relationships for identified active sites [15, 6], contributing to understanding of reaction mechanisms. However, the high computational cost of QM methods restricts the accessible catalyst space [25, 9, 18].

In this context, researchers have explored descriptor-based methods to circumvent QM calculations and establish direct correlations between target properties and electronic factors, commonly referred to as descriptors or features. Traditional methods of crafting descriptors, as seen in the d-band model of Hammer and Nørskov [20], necessitate meticulous selection of chemical properties based on chemical knowledge and are inherently case-dependent [27]. Particularly in the case of SACs, the relationship is intricate, rendering manual crafting inadequate [22, 56]. To streamline this featurization process, we employed CNNs, which autonomously derive a hierarchical representation directly from the eDOS [57, 53, 35]. This approach accurately maps to the activity space, negating the need for manual crafting of descriptors and providing a more efficient method to understand and establish the structure-mechanism-activity relationship for electrochemical reactions.

In this study, we introduced a catalyst performance analysis pipeline (Figure 1) integrating a customized CNN model and volcano plots. Our CNN model accurately estimated the adsorption energies, the activity descriptor, from the eDOS of 2D materials supported single metal catalysts. Evaluations on the adsorption energies of nine intermediates along CO<sub>2</sub>RR and competing hydrogen evolution reaction (HER) as well yielded prediction mean absolute errors (MAEs) at the level of 0.1 eV, which is well aligned with the precision of conventional DFT calculations [29, 37, 62]. This result validated the capability of current CNN model to accurately capture intricate spatial information from eDOS. Building on this, we introduced an orbitalwise eDOS occlusion experiment was adopted afterwards inspired from the occlusion technique widely used in computer vision and pattern recognition applications. Systematic eDOS experiment was performed to elucidate orbitalwise contributions to the adsorption energies at different energy levels, which produced consistent outcome from Crystal Orbital Hamilton Population (COHP) analysis and real-space wavefunction visualization. All these results justified the proficiency of CNN model on identifying the dominant orbitals for adsorbate-substrate interactions. Furthermore, orbital-wise eDOS shifting experiment, exploring the impact of orbital-wise eDOS shifting on the adsorption behavior of studied catalysts were performed.

Simultaneously, we incorporated the volcano plots into the catalysis performance analysis pipeline. Besides offering intuitive analysis of catalyst activities (represented by the elevation of volcanoes), volcano plots points to the directions on how effort can be exerted to identify promising catalyst candidates. Built upon

scaling relations [50], we introduced a hybrid-descriptor scheme incorporating both C- and O-centered species. The scaling coefficient of determination ( $R^2$ ) was increased by 0.1387 using hybrid descriptor, which demonstrate an improved prediction accuracy compared to the well documented single descriptor based method.

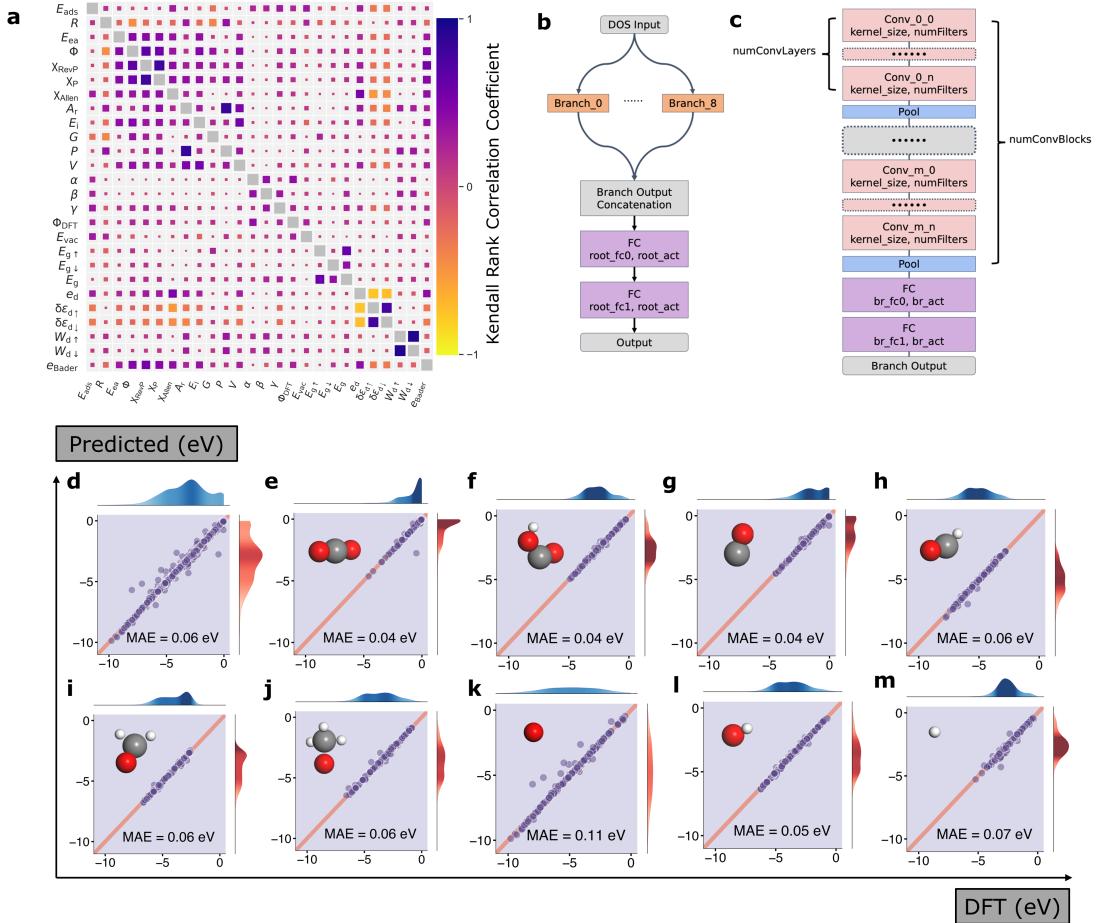
In summary, we introduced a catalyst performance analysis pipeline by integrating CNN and volcano plot analysis, using CO<sub>2</sub>RR as an example proof of concept. The CNN, which correlates eDOS with activity descriptors, establishes a robust electronic structure-activity relationship. Notably, it is capable to predict the influence of eDOS disturbances, caused by either alloying, strain, or defects [44, 3, 36], to modulate the catalytic activity of materials, which will offer insights into the intrinsic mechanism for rationalized catalyst optimization strategies. Advancing the current volcano plot method, the involvement of CNN models in catalyst performance analysis pipeline holds enormous promise for advanced exploration of catalysts with superior performance.

## 2 Results

### 2.1 Adsorption energy prediction from electronic density of states with convolutional neural network

Originating from the pioneering work of Hammer and Nørskov [21], the d-band center model has emerged as a successful paradigm for employing descriptor to understand and predict the adsorption behavior on transition metal surfaces. To investigate alternative electronic and elementary descriptors that may serve as more suitable substitutes for the d-band center model in the context of SACs, we performed Kendall rank correlation analysis to account for both linearity and non-linearity [28]. The correlation coefficient map, as presented in Figure 2a, highlighted the d-band center as the most informative descriptor with a Kendall's  $\tau$  coefficient of -0.33, and followed by the vacuum level and electronegativity. However, d-band center alone does not suffice for accurate predictions of intermediate adsorption energy as illustrated in the scatter plot Supplementary Figure 18, and therefore, refinements are imperative. Unlike bulk transition metal surfaces, where the variations in the d-band centre account for the majority of the adsorption energy variations [44, 55], the unique structure of SACs introduces more variables rather than d-band center that are intrinsically correlated with the variations of adsorption energy. This finding also interpreted that in some reported prior studies, indicating that depending solely on the d-band center may be insufficient for accurate adsorption energy predictions in the context of SACs [54, 17, 12, 65, 23]. As a result, the development of a method tailored specifically for catalysts like SACs becomes essential.

In our study, we employ CNN, a renowned deep learning method proficient in capturing spatial hierarchy from matrix-like data, such as the eDOS. Notably, CNN's autonomy from prior knowledge and human intervention during feature extraction renders it more efficient than traditional machine-learning methods. Our approach harnesses the entire 2D eDOS with a multi-branch CNN model. Figure 2b and Figure 2c delineate the model and its branches. The workflow involves processing the entire 2D eDOS as input, segregating orbitals into branches, and directing the concatenated outputs through another convolutional block, which generates the adsorption energy prediction. Sharing learnable parameters across branches optimizes computational efficiency. By leveraging the CNN model, we predicted adsorption energies from the eDOS of supported single atoms, and validated through parity plots (Figure 2d-m). The model predicts adsorption energies across nine species within the CO<sub>2</sub>RR process and competing HER side reaction, attaining an averaged MAE of 0.06 eV. This MAE aligns with the established precision of DFT methods, which typically range between 0.1-0.2 eV concerning experimental measurement [62, 29, 37]. The full list of prediction errors is available



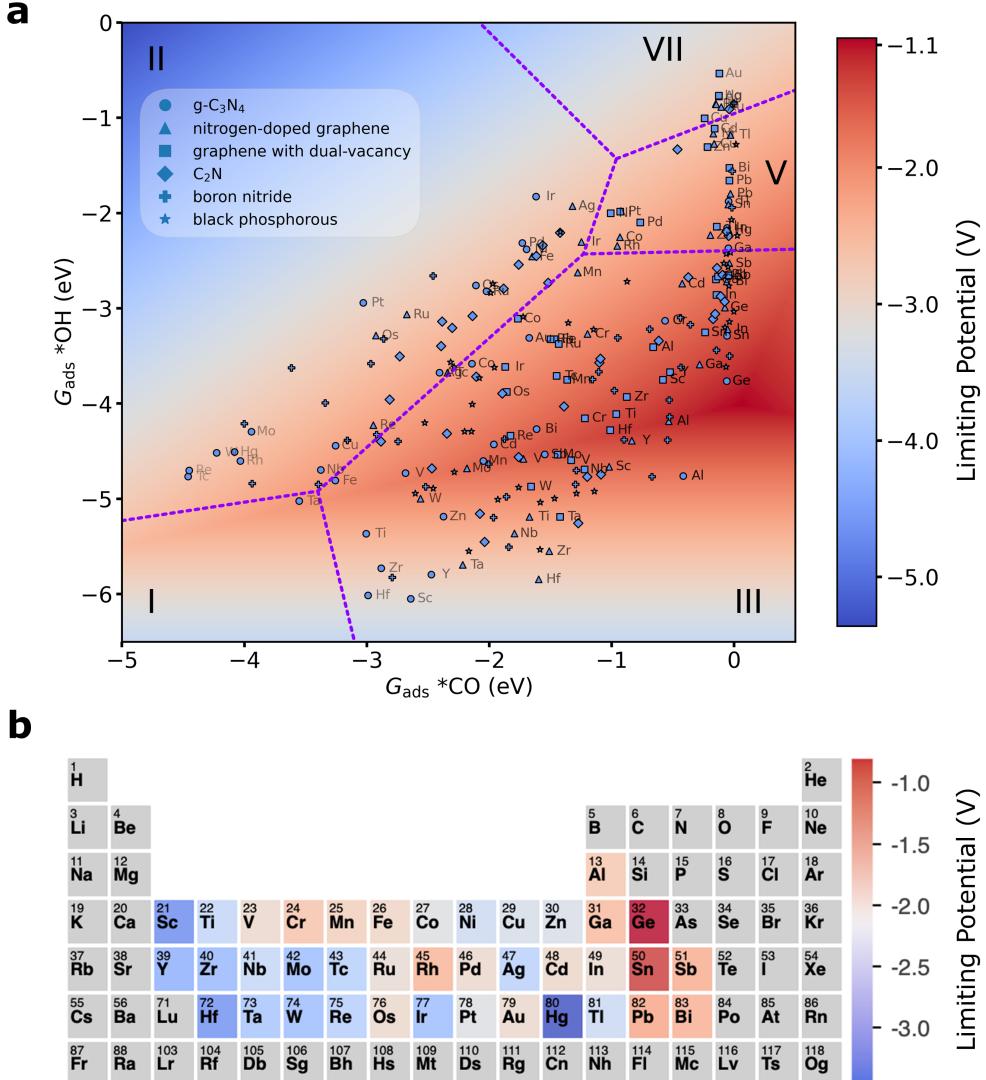
**Figure 2: CNN for electronic density of states.** **a.** Kendall rank correlation coefficient map illustrating the correlation between adsorption energy and other electronic and elementary descriptors. Definitions of the notations are provided in Supplementary section 2.2. **b-c.** Architectural of **(b)** the root and **(c)** each branch of the CNN model. The architecture stays consistent across individual branches. **d-m.** Parity plots demonstrating the performance of the CNN model across nine adsorbates: **(d)** all adsorbates, **(e)** CO<sub>2</sub>, **(f)** COOH, **(g)** CO, **(h)** CHO, **(i)** CH<sub>2</sub>O, **(j)** OCH<sub>3</sub>, **(k)** O, **(l)** OH, and **(m)** H. The augmentation data are not displayed in these plots, nor included in MAE evaluations. Kernel density estimate results are shown along the top and right sides of the main plot, representing DFT calculations and CNN model prediction distributions, respectively.

in Supplementary section 2.4. These results confirm the CNN model’s ability to predict adsorption energies from eDOS input. Moreover, they suggest the potential for predicting adsorption energies for SACs using eDOS, although this would demand more sophisticated feature engineering techniques.

In terms of adsorption energy prediction, the CNN model maintained consistency across diverse elemental compositions and varied adsorption strengths. It handled species containing oxygen, hydrogen, and carbon atoms equally well, spanning an adsorption energy range from 0 to 10 eV. Significantly, after training the CNN model, making predictions (inferences) scales only linearly with system size and demands considerably less computational resources compared to DFT calculations [7]. Conversely, unlike QM based methods, which could generate reasonably precise energy estimation for individual candidates, CNN relies on extensive datasets to ensure reliable predictions. In our case, the original dataset comprised 2052 samples, leading to a CNN model with a validation MAE of 0.3725 eV. In pursuit of enhanced prediction accuracy, we proposed a data augmentation method detailed in the Methods section. This augmentation resulted in an expanded dataset of 12312 samples, enhancing prediction performance with a reduced validation MAE of 0.1736 eV. Comparatively, our CNN model diverges from prior attempts, notably Fung’s pioneer work [16], by requiring only the eDOS of supported single metal atom alone and adsorbate. This revision eliminated the necessity for a complete set of substrate and adsorbate eDOS and significantly reduced computational cost. We foresee the potential in the applicability of our method to diverse species and electrocatalytic reactions, given the availability of comprehensive datasets. The strength of our CNN model lies in its freedom from element- or adsorbate-specific parameters, relying solely on eDOS as input. This adaptability opens up the possibility for its utilization in a wide array of electrocatalytic studies.

## 2.2 Hybrid descriptors enabled volcano plot for catalyst analysis

Catalyst activity volcano plots, built upon the scaling relations proposed by Peterson and Nørskov’s [50], goes beyond assessment of candidatures’ energetics only but to provide informative guidance on discovering catalysts candidates with superior performance [2, 11]. Our endeavor to improve the accuracy of volcano plots leads to the development of the hybrid-descriptor, which encompasses both C- and O- bonding types of CO<sub>2</sub>RR intermediates within the scaling relations. This adaptation arose from the intuition that most CO<sub>2</sub>RR intermediates are featured with both bonding types rather than any single one in dominance. By integrating both bonding featured descriptors into the hybrid approach, we found the  $R^2$



**Figure 3: Limiting potential volcano plot and periodic table.** **a.** Limiting potential volcano plot computed from hybrid-descriptor-based linear scaling relations at potential of  $-0.17$  V vs SHE. In the plot, the circle, triangle, square, diamond, plus, and star symbols represent metals supported on  $\text{g-C}_3\text{N}_4$ , nitrogen-doped graphene, graphene with dual-vacancy, black phosphorous, BN, and  $\text{C}_2\text{N}$ , respectively. Dotted purple lines delineate regions representing rate-determining steps, with roman numerals denoting the respective RDS for each region. The reaction steps are defined in Equation (1) to Equation (8). **b.** Limiting potential periodic table illustrating the theoretical activity of single metal atom catalysts supported on  $\text{g-C}_3\text{N}_4$ .

was increased by 0.1387. Detailed information about the scaling parameters and performance enhancements are provided in Supplementary section 1.5 and Supplementary Figure 10. This improvement to some extent underscored our premise that CO<sub>2</sub>RR intermediates typically exhibit chemical similarities encompassing both O- and C-type descriptors. After taking this dual nature into consideration, our hybrid-descriptor method can evaluate the intermediates adsorption energies more accurately. Importantly, the introduction of hybrid descriptors is computationally affordable compared to DFT methods, as it involves only a set of linear regressions to identify the optimal mixing ratio, as elaborated in the Methods section.

Expanding upon the revised scaling relations, we introduce an activity volcano plot and a rate-determining step (RDS) volcano plot for the CO<sub>2</sub>RR process, shown in Figure 3a and Supplementary Figure 11. Given that HER is a key competing reaction with CO<sub>2</sub>RR [19], we also incorporated a selectivity volcano plot, presented in Supplementary Figure 12, to offer a comprehensive analysis of catalyst candidates. In adherence to Peterson and Nørskov's original framework [50], our choice of x and y axes descriptors aligned with  $G_{\text{ads}*\text{CO}}$  and  $G_{\text{ads}*\text{OH}}$ . Further elaboration on how the limiting potentials were evaluated across the volcano plot can be found in the Methods section. Among the candidatures explored, Ge@g-C<sub>3</sub>N<sub>4</sub> emerged as the most promising catalyst, demonstrating theoretical activity at a limiting potential of -1.2126 eV. Notably, the observed selectivity trend paralleled the activity trend, suggesting that catalysts exhibiting high theoretical activity are also resistant to HER side reactions. The RDS volcano plot (Supplementary Figure 11) indicated protonation of \*CO as the rate-determining step for nearly optimal candidates. Moreover, the activity volcano plot predicted an optimal catalyst featuring a theoretical limiting potential of -1.0516 V, at  $G_{\text{ads}*\text{CO}}$  of 0.0589 eV and  $G_{\text{ads}*\text{OH}}$  of -4.0075 eV, positioned at the right-center of the volcano plot. Results from RDS volcano plot indicate that protonation of \*CO likely serves as the RDS for CO<sub>2</sub>RR. Consequently, it should be the focal point for future catalyst designs. The activity and selectivity volcano plots imply that enhancing theoretical activity could be achieved by decreasing CO affinity while increasing OH affinity with the catalysts. This result suggests a perspective: weak binding of CO to the catalyst surface might facilitate the rotation of C-O backbone from the “upright” configuration in \*CO to the “horizontal” orientation in \*CHO [49], and therefore facilitating dynamics and protonation of the \*CO species.

The volcano plot offers an intuitive analysis of the reaction energetics and provides insights into the mechanisms of the reaction. While the volcano plot provides a broad overview, it lacks the ability to delve into lower-level explanations rooted in electronic structures, where detailed electronic interactions and specific bonding mechanisms remain beyond its scope. Despite its limitations, the volcano

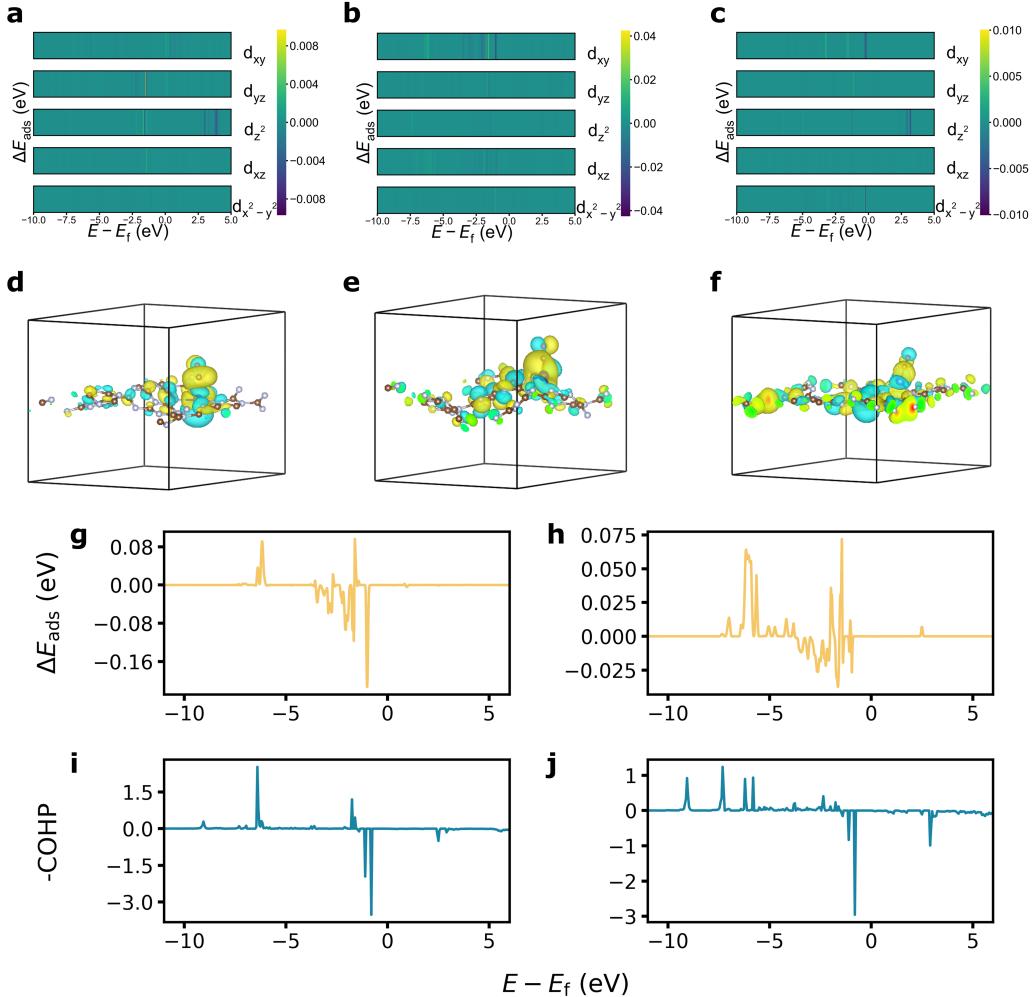
plot serves as a useful tool for identifying focal elementary steps in the reaction pathway that demand closer attention during catalyst design. It also highlights directions for further catalyst refinement, and can also be integrated with the shifting experiment discussed later to optimize existing catalysts.

Additionally, we mapped the theoretical limiting potentials of the analyzed catalyst candidates onto periodic tables, shown in Figure 3b and Supplementary Figure 13 to Supplementary Figure 17. Our analyses do not reveal consistent trends within groups or periods, nor did they exhibit evident patterns across substrates. These observations suggest intricate interactions between supported metal atoms and substrates. The absence of straightforward trends underscores the need for nuanced explorations in understanding these interactions like the CNN model.

### 2.3 Validation of CNN predictions

Historically, interpretability has posed challenges for deep learning methods, often rendering their inner workings akin to black boxes [68, 67, 52]. The occlusion technique, widely utilized in Computer Vision [66, 31, 59], involves systematically masking sections of the input matrix to assess their impact on the final prediction score. In our work, we introduce an orbitalwise occlusion method to discern individual orbital contribution to adsorption energy within the CNN model, drawing inspiration from Fung’s pioneer integration of this technique into eDOS analysis [16]. In this work, occlusion experiments involve orbitalwise masking specific eDOS, which are then fed into the CNN model to observe resulting disturbance in adsorption energy. A visual representation of this process can be found in Supplementary Figure 21. Conceptually, this experiment simulates the effect of shielding electronic states, allowing investigation of the potential contribution to adsorption energy from these states. Additionally, we attempted to extract the planewave coefficients of Kohn-Sham orbital for wavefunction visualization. This validation step was undertaken to confirm the spatial distribution of the orbitals pinpointed by the occlusion experiments.

To validate our method, we conducted occlusion experiments on Cr@g-C<sub>3</sub>N<sub>4</sub>, Co@g-C<sub>3</sub>N<sub>4</sub> and Cu@g-C<sub>3</sub>N<sub>4</sub>. Each of these supported metals possesses 3d subshells with different electronic configurations: half-filled, partially filled and fully filled, allowing us to test the universality of our approach. The results of the occlusion experiment on the supported Cr single atom in Cr@g-C<sub>3</sub>N<sub>4</sub> are given in Figure 4a. The  $d_{x^2-y^2}$  orbital emerges as dominant in CO interaction. Shielding electrons centered around -2 eV significantly alter the adsorption energy, indicating the potency of this orbital. The real-space wavefunction visualization in Figure 4d corroborates these findings, highlighting the interaction between the  $d_{x^2-y^2}$  orbital of single metal atom and the adsorbate. This suggests the  $d_{x^2-y^2}$  orbital of supported Co predominantly influence the interaction with CO, and thus shielding



**Figure 4: Validation of CNN predictions.** **a-c,** The eDOS occlusion experiment results for (a) Cr of Cr@g-C<sub>3</sub>N<sub>4</sub>, (b) Co of Co@g-C<sub>3</sub>N<sub>4</sub> and (c) Cu of Cu@g-C<sub>3</sub>N<sub>4</sub>. eDOS occlusion experiments were conducted on the initial states of CO adsorption process. **d-f,** Real-space wavefunction visualization for (d) Cr  $d_{x^2-y^2}$  orbital of Cr@g-C<sub>3</sub>N<sub>4</sub>, (e) Co  $d_{xy}$  orbital of Co@g-C<sub>3</sub>N<sub>4</sub> and (f) Cu  $d_{xy}$  orbital of Cu@g-C<sub>3</sub>N<sub>4</sub>. Wavefunction visualizations were performed on the final states of CO adsorption process. The yellow and aqua colors represent areas of positive and negative Kohn-Sham orbital coefficients, respectively. **g-j,** Occlusion experiment results for (g)  $d_{xy}$  orbital and (h)  $d_{xz}$  orbital for Co of Co@g-C<sub>3</sub>N<sub>4</sub>. COHP analysis for (i)  $d_{xy}$  orbital and (j)  $d_{xz}$  orbital for Co of Co@g-C<sub>3</sub>N<sub>4</sub>, performed between Co and all adsorbate atoms in the final states of CO adsorption process.

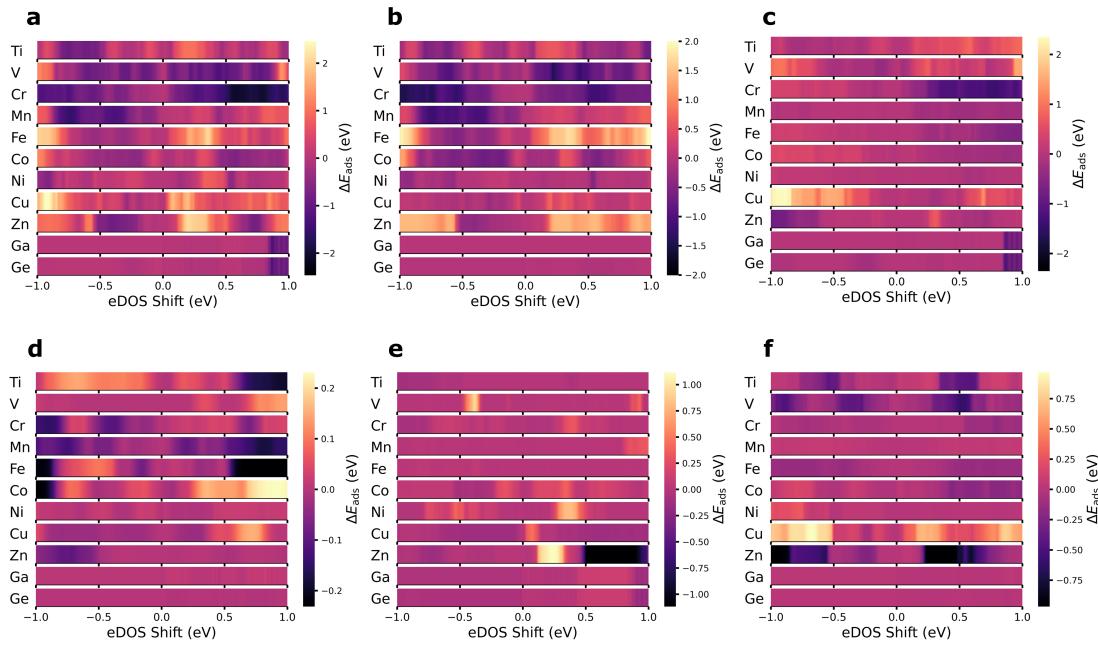
electrons from this orbital markedly disturb the adsorption energy. Additionally, occlusion experiments were performed on the Co  $d_{xy}$  orbital of Co@g-C<sub>3</sub>N<sub>4</sub> (Figure 4b) and the Cu $d_{xy}$  orbital of Cu@g-C<sub>3</sub>N<sub>4</sub> (Figure 4c). In both cases, the  $d_{xy}$  orbital was identified as the dominant contributor. While identifying the dominant orbital is crucial, these findings also suggest that interactions with adsorbates occurring out-of-plane are influenced by in-plane orbitals as well. Importantly, it becomes evident that d orbitals do not universally impact interactions with adsorbates, emphasizing the need for case-specific discussions.

To validate the chemical relevance of predictions made by CNN model, we conducted COHP analysis, a theoretical bond-detecting tool used to identify bonding and anti-bonding contributions to band-structure energy [10]. In this research, interactions between the supported metal atom and atoms within a 5 Å radius were calculated to understand its bonding environment, as detailed in the Methods section. The results of the occlusion experiment on the Co  $d_{xy}$  orbital of Co@g-C<sub>3</sub>N<sub>4</sub>, identified as the dominant orbital via occlusion experiments, were compared with the COHP analysis, as shown in Figure 4g and Figure 4i. The occlusion experiment pinpointed strong metal-adsorbate interactions at energy levels of -1 eV and -6.3 eV, which were confirmed by the COHP analysis. We then investigated the non-dominant Co  $d_{xz}$  orbital, and the COHP analysis highlighted interactions at -1 eV and -6.5 eV, aligning with the findings from the occlusion experiments.

In summary, orbitalwise occlusion experiments provide chemical interpretations into deep learning models, enhancing their interpretability. Furthermore, these occlusion results are readily understandable to chemists, enabling the identification of orbitalwise contributions to the adsorption process.

## 2.4 eDOS shifting experiments for prediction of better catalysts

While CNN models excel in prediction of the activity of catalyst candidates directly from eDOS, volcano plots point out directions leading to better catalysts. Exploring potential disturbances to adsorption energies and aiming to shift existing catalyst candidates closer to the peak of the activity volcano plot presents an intriguing avenue of investigation. In this work, we introduce an orbitalwise shifting experiment to explore the impact of orbitalwise eDOS shifting on adsorption behavior. The experiment involved shifting eDOS along the energy axis, and the CNN model was deployed to predict the resultant disturbance in adsorption energy for each shifting operation. Focusing on all Period 4 metal elements examined in this study, we analyzed the effects of eDOS shifting. To induce a moderate disturbance, we selected the energy range of -1 to 1 eV, and a shifting resolution of 0.005 eV/step was chosen to align with the eDOS calculation resolution.



**Figure 5: eDOS shifting experiments.** Impact of orbitalwise eDOS shifting on CO adsorption energy, as predicted by the CNN model, for single metal atom catalysts supported on g-C<sub>3</sub>N<sub>4</sub>. The disturbances caused by shifting of (a) entire  $d$ , (b)  $d_{xy}$ , (c)  $d_{yz}$ , (d)  $d_{z^2}$ , (e)  $d_{xz}$  and (f)  $d_{x^2-y^2}$  orbitals are illustrated. The shifting step size corresponds to the energy resolution of the eDOS, and is 0.005 eV in this study. A positive eDOS shift indicates a shift towards higher energy levels, and vice versa.

Figure 5 illustrates that supported Ga and Ge, the sole p-block metals in our study, exhibited no response to d orbital shifting. This observation is consistent with their lack of d-electrons in their valence shells, reinforcing the reliability of our shifting experiments. We also investigated potential disturbance from p orbital shifting, as shown in Supplementary Figure 31 and found no significant response. This implies that Ge@g-C<sub>3</sub>N<sub>4</sub> holds promise as a CO<sub>2</sub>RR electrocatalyst, given its high activity and resilience to electronic structure disturbances. Nevertheless, fine-tuning its performance through eDOS modulation might pose a challenge. In contrast, supported Fe, Cu and Zn display strong sensitivity to eDOS shifting. Interestingly, the adsorption energy consistently shifted towards more positive values, indicating a weakening of the catalyst-adsorbate interaction, regardless of the shifting direction within the investigated energy range. This underscores the intricate interplay between the supported metal atom and the substrate, emphasizing the necessity of a tool like the CNN model to predict such disturbances. Supported Cr exhibited a contrary response to eDOS shifting, strengthening the interaction with the CO adsorbate regardless of the shifting direction. Supported Ti, V, Co and Ni showed weak response to eDOS shifting, suggesting the need for alternative methods beyond manipulating their eDOS to regulate their interactions with CO adsorbate.

In summary, shifting experiments offer a valuable means to predict the potential disturbances in adsorption behavior resulting from eDOS shifting, a task that is challenging to accomplish directly within the DFT framework. However, it's important to note that achieving the predicted shifting in actual theoretical or experimental scenarios may not be easily accessible. Skillful manipulation of the electronic structure of catalyst candidates would be required, presenting a practical challenge. In contrast to prior studies on bulk metals [16], wherein the adsorption energy exhibits a continuous shift with eDOS variations, the response of SACs is notably discrete. These responses manifest as isolated peaks throughout the shifting range, attributed to the distinctive structure of SACs. Despite this challenge, having a tool that can predict the potential impact of eDOS shifting is advantageous for the theoretical design of SACs.

## 3 Methods

### 3.1 DFT calculations

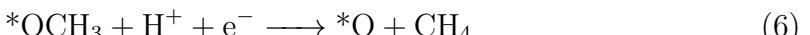
Spin-polarized density functional theory calculations were performed using the Vienna Ab initio Simulation Package (VASP) version 5.4.4 [33, 32, 34] to optimize geometries and determine the electronic structure. The Perdew-Burke-Ernzerhof (PBE) functional [33] within the generalized gradient approximation (GGA) [48] was utilized to describe the electron-ion interactions and electronic exchange correlations. Kohn-Shan equations were approached employing a plane-wave basis set with a cut-off energy of 440 eV. Geometry optimizations were performed using  $\Gamma$ -Centered K-point meshes of  $(2 \times 2 \times 1)$  until the force on each atom fell below  $0.02 \text{ eV \AA}^{-1}$ . Electronic structure calculations utilized denser K-point meshes of  $(3 \times 3 \times 1)$  and employed self-consistent field (SCF) methods.

### 3.2 Catalyst models

Supported SACs were modelled by anchoring individual metal atoms onto various 2D substrates, including g-C<sub>3</sub>N<sub>4</sub>, nitrogen-doped graphene, graphene with dual-vacancy, black phosphorous, boron nitride, and monolayer C<sub>2</sub>N. The specific single metal atoms studied are outlined in Supplementary Figure 1. To prevent interactions between adjacent images along the z-axis, vacuum layers of 15 Å were introduced.

### 3.3 CO<sub>2</sub>RR pathway

The C1 production CO<sub>2</sub>RR pathway, as initially described by Peterson et al. [49] and later confirmed by Jiao et al. [24] for g-C<sub>3</sub>N<sub>4</sub> based SACs, involves the following elementary steps. Asterisks (\*) indicate species adsorbed on the SAC surfaces:



### 3.4 Free energy calculations

Free energies were computed from electronic energies by considering thermal corrections, incorporating normal mode analysis of all degrees of freedom of adsorbates. The free energy  $G$  was calculated using the formula:

$$G = E + E_{\text{ZPE}} - TS \quad (9)$$

where  $G$  represents the free energy,  $E$  denotes electronic energy,  $E_{\text{ZPE}}$  is the zero-point energy,  $T$  represents temperature in Kevin, and  $S$  stands for entropy. Detailed information regarding these corrections can be found in Supplementary section 1.3.

The computational hydrogen electrode (CHE) model [49, 45] was employed to introduce potential dependence. At the reference electrode surface, the reaction



is in equilibrium, allowing the calculation of the chemical potential of the electron-proton pair as:

$$\mu(\text{H}^+) + \mu(\text{e}^-) = \frac{1}{2}\mu(\text{H}_{2(\text{g})}) - eU \quad (11)$$

For the external potential  $U$  applied, the change in free energy ( $\Delta G$ ) was determined using the equation:

$$\Delta G_n(U) = \Delta G_n(U = 0) + neU \quad (12)$$

Here,  $\mu$  represents chemical potential,  $U$  signifies the external potential applied,  $n$  represents the number of proton-electron pairs transferred, and  $e$  denotes the elementary charge.

### 3.5 Hybrid descriptor for volcano plots

Hybrid descriptors incorporate both descriptors within a single equation, enhancing the accuracy of scaling relationship with minimal computational overhead, as demonstrated in Supplementary Figure 10 and Supplementary section 1.5. Comprehensive analyses are presented Supplementary section 1.5. This leads to the formulation of the scaling relation:

$$G_{\text{ads}}Z = k[\theta^*G_{\text{ads}}\text{CO} + (1 - \theta^*)G_{\text{ads}}\text{OH}] + c_Z \quad (13)$$

Here,  $\theta$  represents the mixing ratio of the CO descriptor,  $Z$  denotes any adsorbate, and  $k$  and  $c_Z$  are adsorbate-specific scaling parameters. The optimal mixing ratio

$\theta$  is determined through iterative exploration across the range, utilizing a step length of 0.01.

13 could be simplified to a more general form:

$$G_{\text{ads}}Z = a_Z G_{\text{ads}}\text{CO} + b_Z G_{\text{ads}}\text{OH} + c_Z \quad (14)$$

Here,  $a_Z$ ,  $b_Z$ ,  $c_Z$  are adsorbate-specific parameters.

Consequently, the limiting potential  $U_L$  is entirely determined by two descriptors through the following equation:

$$U_L = -\frac{\max\{\Delta G_i\}}{e} \quad (15)$$

Where  $\Delta G_i$  represents the free energy change of reaction step i.

This approach simplifies the theoretical performance assessment of any catalyst within the scope of our research to only two descriptors:  $G_{\text{ads CO}}$  and  $G_{\text{ads OH}}$ . This method offers a visual representation of the high-dimensional optimization challenge. To aid visualization, a 2D mesh grid is established, enabling vectorized assessments of limiting potentials across the descriptor value ranges. At each grid point, the limiting potential is computed using Equation (15).

### 3.6 CNN architecture and hyperparameter tuning

We designed our CNN architecture drawing inspiration from its successful implementation in tasks involving electrocardiogram signal classification [61], as illustrated in Figure 2b and Figure 2c. This multi-branched CNN comprises nine branches tailored to handle eDOS orbitals while ensuring a reduced memory footprint.

The effectiveness of neural network models is intricately tied to the choice of hyperparameters. To identify the optimal set of hyperparameters, we employed Hyperband [38], a modern and parallelizable bandit-based optimization algorithm. The search space for hyperparameters is detailed in Supplementary section 2.4, and the resulting optimal hyperparameters are presented in Supplementary section 2.4.

### 3.7 Data augmentation and CNN model training

To enhance the generalizability and robustness of the our CNN model, and achieve comprehensive coverage of the chemical space [63], we applied data augmentation techniques to the initial dataset using the following procedures: - In addition to the designated “initial state” of the adsorption process, where the adsorbate positioned 6.5 Å above the supported single metal atom, five intermediate images were interpolated. These images were distributed between the initial and final states at

a spacing of 0.5 Å, starting from the initial state side. - A diverse set of smearing techniques was employed to determine of partial occupancies  $f_{nk}$ , including the tetrahedron method [4] and the Gaussian method. For the Gaussian method, the smearing width was varied within the range of 0.01 eV to 0.1 eV. This careful selection of methods was made to guarantee reliable and consistent predictions of adsorption energies from the model, regardless of the smearing technique used.

The augmented dataset comprises 12,312 samples, with adsorption energies having a standard deviation of 1.8117 eV and a variance of 3.2821 eV<sub>2</sub>. Twenty percent of these samples were allocated to the validation set. Throughout the training process, a batch size of 64 was employed to ensure a balanced learning dynamic.

### 3.8 eDOS occlusion experiments

Supplementary Figure 21 visually illustrates the eDOS occlusion experiment methodology. To maintain consistent shapes, zero-padding was applied to the input eDOS arrays. Focusing on understanding the influence of metals on adsorption behavior, a masker with dimensions [width, 1, 1] was employed along the metal channel to eliminate cross-orbital interactions. Here, “width” denotes the masker width, and detailed explanations for determining the width are provided in Supplementary section 2.5. Through recursive application of these maskers, variations in predicted adsorption energies from the CNN model were recorded, generating an occlusion array.

### 3.9 eDOS shifting experiments

Supplementary Figure 30 illustrates the implemented eDOS shifting experiment protocol. In this study, the initial electronic density of states underwent controlled shifting along the energy axis. This protocol ranges from -1 eV to 1 eV, with a step length of 0.005 eV per image. To maintain a uniform energy window, each image underwent cropping and padding procedures. The resulted output arrays took the form of [400, 1] arrays for simultaneous orbital shifting and [400, 9] arrays for individual orbital shifting. The energy range for shifting was chosen to induce a moderate disturbance to adsorption energy.

### 3.10 Chemical bond analysis and real-space wavefunction visualization

COHP analysis was performed using the Local-Orbital Basis Suite Towards Electronic-Structure Reconstruction (LOBSTER) package [10, 30, 43, 40, 14] version 4.1.0,

with the GGA-PBE wavefunctions fitted by S. Maintz [30, 41]. A Gaussian smearing for energy integration was employed with a broadening width of 0.05 eV. The quantification of bonding strength between the single metal atom and the substrate was achieved through the summation of interactions with neighboring atoms within a range of 0.5 Å to 5.0 Å from the single metal atom. Orbitals of interest were confirmed through band-structure analysis, and real-space wavefunctions were subsequently visualized using VASPKIT [60] and VESTA [42].

### 3.11 Machine learning and data analysis environment

The CNN model was implemented using TensorFlow [1], and hyperparameter optimization was conducted using the Hyperband [38] algorithm via KerasTuner [46]. For a more detailed overview of our machine learning setup and data analysis environment, please refer to Supplementary section 3.

## A Code and Data Availability

Essential source code and data required to replicate the study’s results—including VASP input files, structure files of SACs, eDOS arrays and the pretrained CNN model—are available on GitHub repository at <https://github.com/DanielYang59/cnn4dos>, and can be provided by the corresponding authors upon request. Due to their substantial size in terabytes, the complete source data (complete VASP output files and such) is not currently hosted on publicly available repositories, but can be obtained from the corresponding authors upon request.

## B Acknowledgements

We acknowledge funding support from the Australian Research Council (FT160100281, FT180100387, DP200103568, DP230101625). The QUT eResearch Office provided computational resources and services used in this work. This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), supported by the Australian Government. It was also supported by resources provided by the Pawsey Supercomputing Research Centre with funding from the Australian Government and the Government of Western Australia. Insights into deep learning methodologies were contributed by Mr. Yanwei Guan from Chongqing University and Mr. Zhipeng He from Assoc. Prof. Chun Ouyang’s team at QUT. The machine learning and data analysis implementation benefited from resources provided by the GitHub, TensorFlow and Stack Overflow communities. Dr. Ryky Nelson from RWTH Aachen University offered valuable guidance on Crystal Orbital Hamilton Population analysis.

## C Author Contributions

H.Y. performed DFT calculations, implemented Python code for machine learning methodologies and data analysis, and drafted and refined the manuscript. J.Z. engaged in valuable discussions on DFT calculations. Q.W. provided advice on the computer code revisions. B.L. provided guidance on DFT calculations and offered research supervision. W.L. facilitated resource acquisition and offered research supervision. Z.S. initiated the research concept, contributed to manuscript revisions, and supervised the overall research. T.L. guided DFT calculations and data analysis, offering continuous supervision throughout the research and manuscript drafting.

## **D Competing Interests**

The authors declare no competing interests in relation to this research.

## References

- [1] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. {TensorFlow}: a system for {Large-Scale} machine learning. In *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, pages 265–283, 2016.
- [2] AA Balandin. Modern state of the multiplet theor of heterogeneous catalysis. *Advances in catalysis*, 19:1–210, 1969.
- [3] Bapi Bera, Arup Chakraborty, Tathagata Kar, Pradipkumar Leuaa, and Manoj Neergat. Density of states, carrier concentration, and flat band potential derived from electrochemical impedance measurements of n-doped carbon and their influence on electrocatalysis of oxygen reduction reaction. *The Journal of Physical Chemistry C*, 121(38):20850–20856, 2017.
- [4] Peter E Blöchl, Ove Jepsen, and Ole Krogh Andersen. Improved tetrahedron method for brillouin-zone integrations. *Physical Review B*, 49(23):16223, 1994.
- [5] Yanming Cai, Jiaju Fu, Yang Zhou, Yu-Chung Chang, Qianhao Min, Jun-Jie Zhu, Yuehe Lin, and Wenlei Zhu. Insights on forming n, o-coordinated cu single-atom catalysts for electrochemical reduction co<sub>2</sub> to methane. *Nature Communications*, 12(1):586, 2021.
- [6] Emily A Carter. Challenges in modeling materials properties without experimental input. *Science*, 321(5890):800–803, 2008.
- [7] Anand Chandrasekaran, Deepak Kamal, Rohit Batra, Chiho Kim, Lihua Chen, and Rampi Ramprasad. Solving the electronic structure problem with machine learning. *npj Computational Materials*, 5(1):22, 2019.
- [8] Yalu Chen, Yufeng Huang, Tao Cheng, and William A Goddard III. Identifying active sites for co<sub>2</sub> reduction on dealloyed gold surfaces by combining machine learning with multiscale simulations. *Journal of the American Chemical Society*, 141(29):11651–11657, 2019.
- [9] Beatriz Roldan Cuenya and Farzad Behafarid. Nanocatalysis: size-and shape-dependent chemisorption and catalytic reactivity. *Surface Science Reports*, 70(2):135–187, 2015.
- [10] Volker L Deringer, Andrei L Tchougréeff, and Richard Dronskowski. Crystal orbital hamilton population (cohp) analysis as projected from plane-wave basis sets. *The journal of physical chemistry A*, 115(21):5461–5466, 2011.

- [11] Olaf Deutschmann, Helmut Knözinger, Karl Kochloefl, and Thomas Turek. Heterogeneous catalysis and solid catalysts, 1. fundamentals. *Ullmann's encyclopedia of industrial chemistry*, 2000.
- [12] Giovanni Di Liberto, Luis A Cipriano, and Gianfranco Pacchioni. Universal principles for the rational design of single atom electrocatalysts? handle with care. *ACS Catalysis*, 12(10):5846–5856, 2022.
- [13] Cao-Thang Dinh, Thomas Burdyny, Md Golam Kibria, Ali Seifitokaldani, Christine M Gabardo, F Pelayo García de Arquer, Amirreza Kiani, Jonathan P Edwards, Phil De Luna, Oleksandr S Bushuyev, et al. Co<sub>2</sub> electroreduction to ethylene via hydroxide-mediated copper catalysis at an abrupt interface. *Science*, 360(6390):783–787, 2018.
- [14] Richard Dronskowski and Peter E Blöchl. Crystal orbital hamilton populations (cohp): energy-resolved visualization of chemical bonding in solids based on density-functional calculations. *The Journal of Physical Chemistry*, 97(33):8617–8624, 1993.
- [15] Jeremy T Feaster, Chuan Shi, Etosha R Cave, Toru Hatsukade, David N Abram, Kendra P Kuhl, Christopher Hahn, Jens K Nørskov, and Thomas F Jarillo. Understanding selectivity for the electrochemical reduction of carbon dioxide to formic acid and carbon monoxide on metal electrodes. *Acs Catalysis*, 7(7):4822–4827, 2017.
- [16] Victor Fung, Guoxiang Hu, Panchapakesan Ganesh, and Bobby G Sumpter. Machine learned features from density of states for accurate adsorption energy prediction. *Nature communications*, 12(1):88, 2021.
- [17] Victor Fung, Guoxiang Hu, Zili Wu, and De-en Jiang. Descriptors for hydrogen evolution on single atom catalysts in nitrogen-doped graphene. *The Journal of Physical Chemistry C*, 124(36):19571–19578, 2020.
- [18] Bryan R. Goldsmith, Jacques Esterhuizen, Jin-Xun Liu, Christopher J. Bartel, and Christopher Sutton. Machine learning for heterogeneous catalyst design and discovery. *AIChe Journal*, 64(7):2311–2323, 2018.
- [19] Akansha Goyal, Giulia Marcandalli, Vladislav A Mints, and Marc TM Koper. Competition between co<sub>2</sub> reduction and hydrogen evolution on a gold electrode under well-defined mass transport conditions. *Journal of the American Chemical Society*, 142(9):4154–4161, 2020.
- [20] BJKN Hammer and Jens K Nørskov. Electronic factors determining the reactivity of metal surfaces. *Surface science*, 343(3):211–220, 1995.

- [21] Bjørk Hammer and Jens Kehlet Nørskov. Theoretical surface science and catalysis-calculations and concepts. In *Advances in catalysis*, volume 45, pages 71–129. Elsevier, 2000.
- [22] Zhong-Kang Han, Debalaya Sarker, Runhai Ouyang, Aliaksei Mazheika, Yi Gao, and Sergey V Levchenko. Single-atom alloy catalysts designed by first-principles calculations and artificial intelligence. *Nature communications*, 12(1):1833, 2021.
- [23] Hai-Cai Huang, Yang Zhao, Jing Wang, Jun Li, Jing Chen, Qiang Fu, Yu-Xiang Bu, and Shi-Bo Cheng. Rational design of an efficient descriptor for single-atom catalysts in the hydrogen evolution reaction. *Journal of Materials Chemistry A*, 8(18):9202–9208, 2020.
- [24] Yan Jiao, Yao Zheng, Ping Chen, Mietek Jaroniec, and Shi-Zhang Qiao. Molecular scaffolding strategy with synergistic active centers to facilitate electrocatalytic co<sub>2</sub> reduction to hydrocarbon/alcohol. *Journal of the American Chemical Society*, 139(49):18093–18100, 2017.
- [25] Ryosuke Jinnouchi and Ryoji Asahi. Predicting catalytic activity of nanoparticles by a dft-aided machine-learning algorithm. *The journal of physical chemistry letters*, 8(17):4279–4283, 2017.
- [26] Wen Ju, Alexander Bagger, Guang-Ping Hao, Ana Sofia Varela, Ilya Sinev, Volodymyr Bon, Beatriz Roldan Cuenya, Stefan Kaskel, Jan Rossmeisl, and Peter Strasser. Understanding activity and selectivity of metal-nitrogen-doped carbon catalysts for electrochemical reduction of co<sub>2</sub>. *Nature communications*, 8(1):944, 2017.
- [27] Seiji Kajita, Nobuko Ohba, Ryosuke Jinnouchi, and Ryoji Asahi. A universal 3d voxel descriptor for solid-state material informatics with deep convolutional neural networks. *Scientific reports*, 7(1):16991, 2017.
- [28] Maurice G Kendall. A new measure of rank correlation. *Biometrika*, 30(1/2):81–93, 1938.
- [29] Scott Kirklin, James E Saal, Bryce Meredig, Alex Thompson, Jeff W Doak, Muratahan Aykol, Stephan Rühl, and Chris Wolverton. The open quantum materials database (oqmd): assessing the accuracy of dft formation energies. *npj Computational Materials*, 1(1):1–15, 2015.
- [30] Toshikatsu Koga, Katsutoshi Kanayama, Shinya Watanabe, and Ajit J Thakkar. Analytical hartree–fock wave functions subject to cusp and asymptotic constraints: He to xe, li<sup>+</sup> to cs+, h<sup>-</sup> to i<sup>-</sup>. *International journal of quantum chemistry*, 71(6):491–497, 1999.

- [31] Adam Kortylewski, Qing Liu, Huiyu Wang, Zhishuai Zhang, and Alan Yuille. Combining compositional models and deep networks for robust object classification under occlusion. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 1333–1341, 2020.
- [32] Georg Kresse and Jürgen Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational materials science*, 6(1):15–50, 1996.
- [33] Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical review B*, 54(16):11169, 1996.
- [34] Georg Kresse and Jürgen Hafner. Ab initio molecular dynamics for liquid metals. *Physical review B*, 47(1):558, 1993.
- [35] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012.
- [36] Kohei Kusada, Dongshuang Wu, Tomokazu Yamamoto, Takaaki Toriyama, Syo Matsumura, Wei Xie, Michihisa Koyama, Shogo Kawaguchi, Yoshiki Kubota, and Hiroshi Kitagawa. Emergence of high orr activity through controlling local density-of-states by alloying immiscible au and ir. *Chemical science*, 10(3):652–656, 2019.
- [37] Kurt Lejaeghere, Gustav Bihlmayer, Torbjörn Björkman, Peter Blaha, Stefan Blügel, Volker Blum, Damien Caliste, Ivano E Castelli, Stewart J Clark, Andrea Dal Corso, et al. Reproducibility in density functional theory calculations of solids. *Science*, 351(6280):aad3000, 2016.
- [38] Lisha Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar. Hyperband: A novel bandit-based approach to hyperparameter optimization. *Journal of Machine Learning Research*, 18(185):1–52, 2018.
- [39] Guiji Liu, Michelle Lee, Soonho Kwon, Guosong Zeng, Johanna Eichhorn, Aya K Buckley, F Dean Toste, William A Goddard III, and Francesca M Toma. Co<sub>2</sub> reduction on pure cu produces only h<sub>2</sub> after subsurface o is depleted: theory and experiment. *Proceedings of the National Academy of Sciences*, 118(23):e2012649118, 2021.
- [40] Stefan Maintz, Volker L Deringer, Andrei L Tchougréeff, and Richard Dronskowski. Analytic projection from plane-wave and paw wavefunctions and

- application to chemical-bonding analysis in solids. *Journal of computational chemistry*, 34(29):2557–2567, 2013.
- [41] Stefan Maintz, Volker L Deringer, Andrei L Tchougréeff, and Richard Dronskowski. Lobster: A tool to extract chemical bonding from plane-wave based dft, 2016.
  - [42] Koichi Momma and Fujio Izumi. Vesta: a three-dimensional visualization system for electronic and structural analysis. *Journal of Applied crystallography*, 41(3):653–658, 2008.
  - [43] Ryky Nelson, Christina Ertural, Janine George, Volker L Deringer, Geofroy Hautier, and Richard Dronskowski. Lobster: Local orbital projections, atomic charges, and chemical-bonding analysis from projector-augmented-wave-based density-functional theory. *Journal of Computational Chemistry*, 41(21):1931–1940, 2020.
  - [44] Jens K Nørskov, Frank Abild-Pedersen, Felix Studt, and Thomas Bligaard. Density functional theory in surface chemistry and catalysis. *Proceedings of the National Academy of Sciences*, 108(3):937–943, 2011.
  - [45] Jens Kehlet Nørskov, Jan Rossmeisl, Ashildur Logadottir, LRKJ Lindqvist, John R Kitchin, Thomas Bligaard, and Hannes Jonsson. Origin of the overpotential for oxygen reduction at a fuel-cell cathode. *The Journal of Physical Chemistry B*, 108(46):17886–17892, 2004.
  - [46] Tom O’Malley, Elie Bursztein, James Long, François Chollet, Haifeng Jin, Luca Invernizzi, et al. Kerastuner. <https://github.com/keras-team/keras-tuner>, 2019.
  - [47] Silvio Osella and William A Goddard III. Co<sub>2</sub> reduction to methane and ethylene on a single-atom catalyst: A grand canonical quantum mechanics study. *Journal of the American Chemical Society*, 145(39):21319–21329, 2023.
  - [48] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.
  - [49] Andrew A Peterson, Frank Abild-Pedersen, Felix Studt, Jan Rossmeisl, and Jens K Nørskov. How copper catalyzes the electroreduction of carbon dioxide into hydrocarbon fuels. *Energy & Environmental Science*, 3(9):1311–1315, 2010.
  - [50] Andrew A Peterson and Jens K Nørskov. Activity descriptors for co<sub>2</sub> electroreduction to methane on transition-metal catalysts. *The Journal of Physical Chemistry Letters*, 3(2):251–258, 2012.

- [51] Wenhao Ren, Xin Tan, Wanfeng Yang, Chen Jia, Shumao Xu, Kaixue Wang, Sean C Smith, and Chuan Zhao. Isolated diatomic ni-fe metal–nitrogen sites for synergistic electroreduction of co<sub>2</sub>. *Angewandte Chemie International Edition*, 58(21):6972–6976, 2019.
- [52] Neil Savage. Breaking into the black box of artificial intelligence, 2022.
- [53] Richard Socher, Brody Huval, Bharath Bath, Christopher D Manning, and Andrew Ng. Convolutional-recursive deep learning for 3d object classification. *Advances in neural information processing systems*, 25, 2012.
- [54] Zhongti Sun, Zhilong Song, and Wan-Jian Yin. Going beyond the d-band center to describe co<sub>2</sub> activation on single-atom alloys. *Advanced Energy and Sustainability Research*, 3(2):2100152, 2022.
- [55] Ichigaku Takigawa, Ken-ichi Shimizu, Koji Tsuda, and Satoru Takakusagi. Machine-learning prediction of the d-band center for metals and bimetals. *RSC advances*, 6(58):52587–52595, 2016.
- [56] Hari Thirumalai and John R Kitchin. Investigating the reactivity of single atom alloys using density functional theory. *Topics in Catalysis*, 61:462–474, 2018.
- [57] Du Tran, Lubomir Bourdev, Rob Fergus, Lorenzo Torresani, and Manohar Paluri. Learning spatiotemporal features with 3d convolutional networks, 2015.
- [58] Aiqin Wang, Jun Li, and Tao Zhang. Heterogeneous single-atom catalysis. *Nature Reviews Chemistry*, 2(6):65–81, 2018.
- [59] Angtian Wang, Yihong Sun, Adam Kortylewski, and Alan L Yuille. Robust object detection under occlusion with context-aware compositionalnets. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12645–12654, 2020.
- [60] Vei Wang, Nan Xu, Jin-Cheng Liu, Gang Tang, and Wen-Tong Geng. Vasptkit: A user-friendly interface facilitating high-throughput computing and analysis using vaspt code. *Computer Physics Communications*, 267:108033, 2021.
- [61] Kuba Weimann and Tim OF Conrad. Transfer learning for ecg classification. *Scientific reports*, 11(1):5251, 2021.
- [62] Jess Wellendorff, Trent L. Silbaugh, Delfina Garcia-Pintos, Jens K. Nørskov, Thomas Bligaard, Felix Studt, and Charles T. Campbell. A benchmark

- database for adsorption bond energies to transition metal surfaces and comparison to selected dft functionals. *Surface Science*, 640:36–44, 2015. Reactivity Concepts at Surfaces: Coupling Theory with Experiment.
- [63] Yutong Xie, Ziqiao Xu, Jiaqi Ma, and Qiaozhu Mei. How much of the chemical space has been covered? measuring and improving the variety of candidate set in molecular generation. *CoRR*, abs/2112.12542, 2021.
  - [64] Hong Bin Yang, Sung-Fu Hung, Song Liu, Kaidi Yuan, Shu Miao, Liping Zhang, Xiang Huang, Hsin-Yi Wang, Weizheng Cai, Rong Chen, et al. Atomically dispersed ni (i) as the active site for electrochemical co<sub>2</sub> reduction. *Nature energy*, 3(2):140–147, 2018.
  - [65] Hao Yuan, Zhenyu Li, Xiao Cheng Zeng, and Jinlong Yang. Descriptor-based design principle for two-dimensional single-atom catalysts: carbon dioxide electroreduction. *The journal of physical chemistry letters*, 11(9):3481–3487, 2020.
  - [66] Matthew D Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part I* 13, pages 818–833. Springer, 2014.
  - [67] Quan-shi Zhang and Song-Chun Zhu. Visual interpretability for deep learning: a survey. *Frontiers of Information Technology & Electronic Engineering*, 19(1):27–39, 2018.
  - [68] Quanshi Zhang, Ying Nian Wu, and Song-Chun Zhu. Interpretable convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8827–8836, 2018.
  - [69] Qing Zhao, John Mark P Martirez, and Emily A Carter. Revisiting understanding of electrochemical co<sub>2</sub> reduction on cu (111): competing proton-coupled electron transfer reaction mechanisms revealed by embedded correlated wavefunction theory. *Journal of the American Chemical Society*, 143(16):6152–6164, 2021.