Supplementary Information

A general machine-learning framework for high-throughput screening for stable and efficient RuO_2-based acidic oxygen evolution reaction catalysts

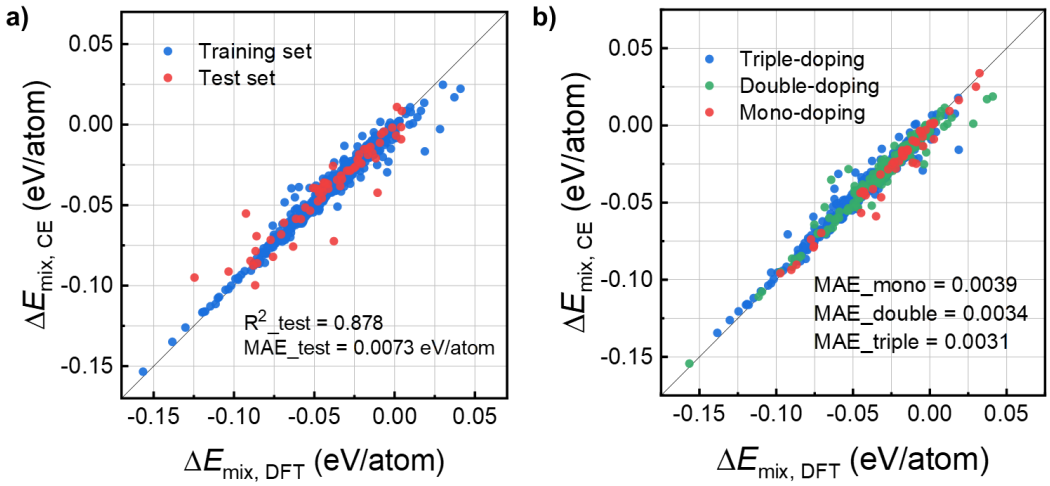
Zhe Shang[1], Qian Dang[1], Fengmei Wang[2], Xiaoming Sun[2], Hui Li[1*]

[1]Beijing Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology, Beijing 100029, China

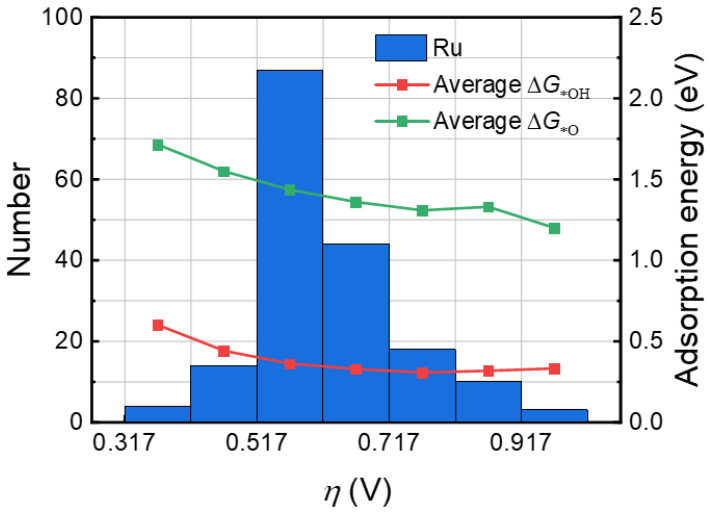
[2]State Key Laboratory of Chemical Resource Engineering, Beijing University of Chemical Technology, Beijing 100029, China

*E-mail: hli@mail.buct.edu.cn

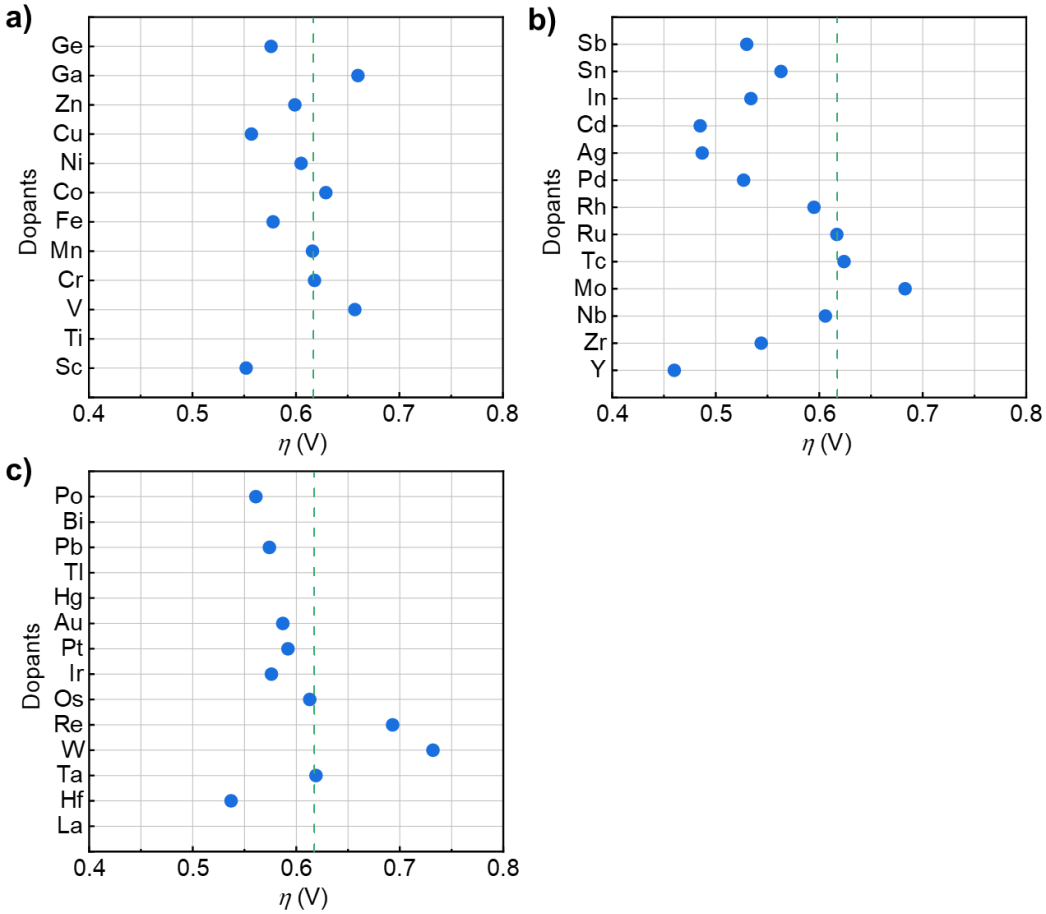
Contents



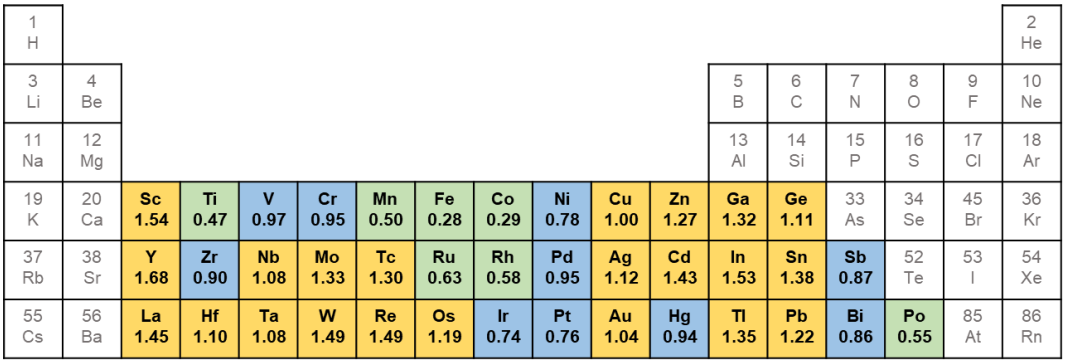
Supplementary Figure 1. The prediction performance of CE models. (a) The correlation between model-predicted and DFT-calculated ΔE_{mix}. The model covered all of doping elements. (b) The prediction performance of the CE model on mono-, double-, and triple-doping catalysts.



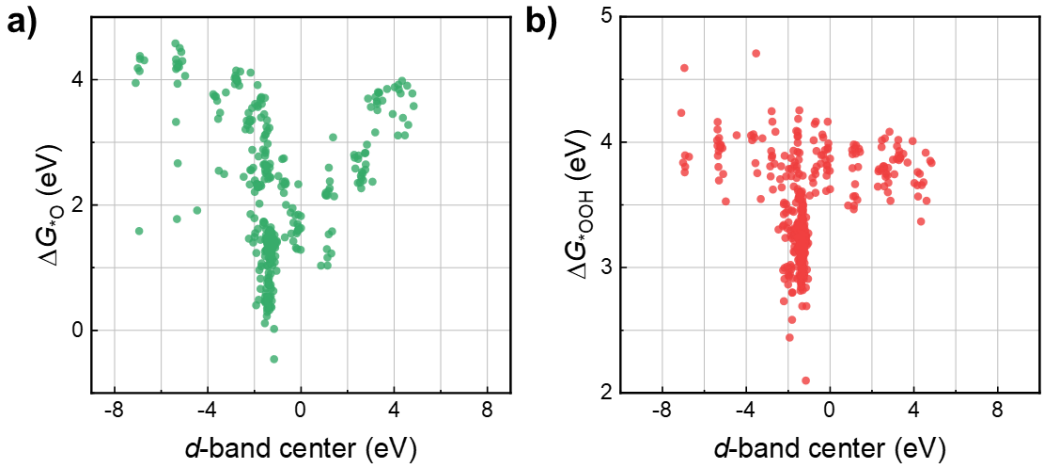
Supplementary Figure 2. The overpotential distribution of Ru sites and average ΔG_{*OH} and ΔG_{*O} in the corresponding range.



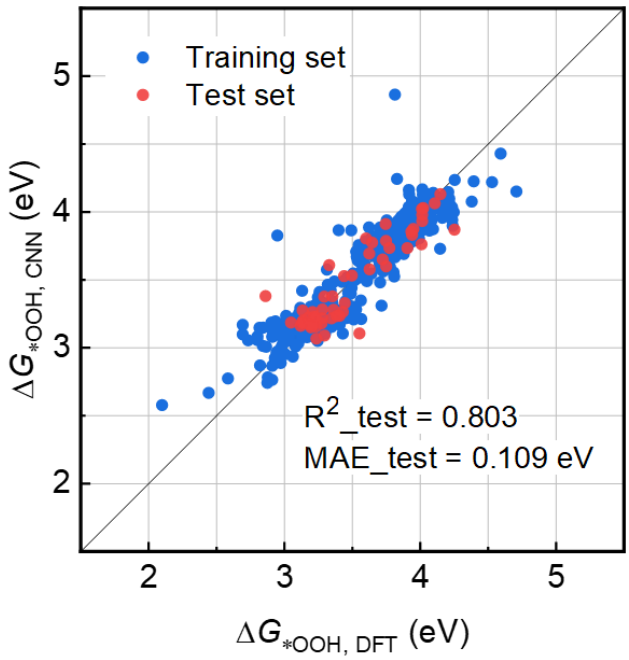
Supplementary Figure 3. The overpotential distribution of Ru sites under the influence of different dopants. The absent data indicates that not fully optimized intermediates exist for the M-doped RuO_2.



Supplementary Figure 4. The average overpotential of each metal active center, green color means lower average overpotential than Ru, blue color represents average overpotential in the range of 0.7~1.0 V, and orange color represents average overpotential more than 1.0 V.

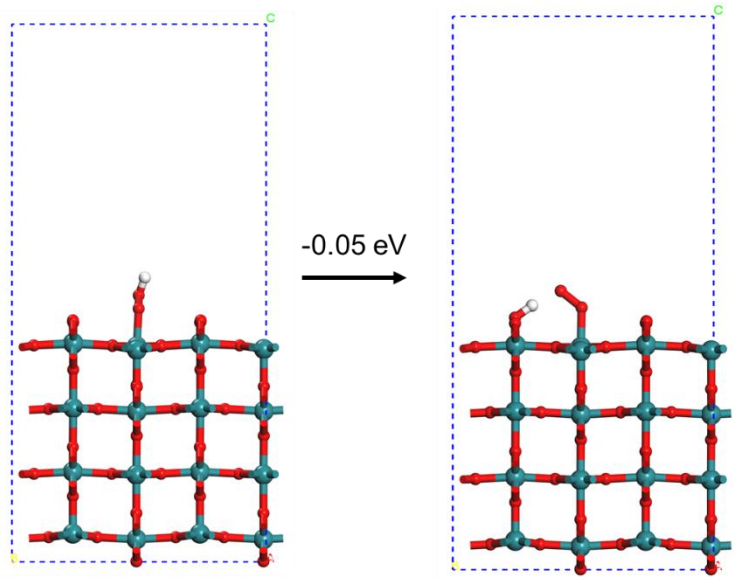


Supplementary Figure 5. The correlations between (a) ΔG_{*O}, (b) ΔG_{*OOH} and *d*-band center of active centers.

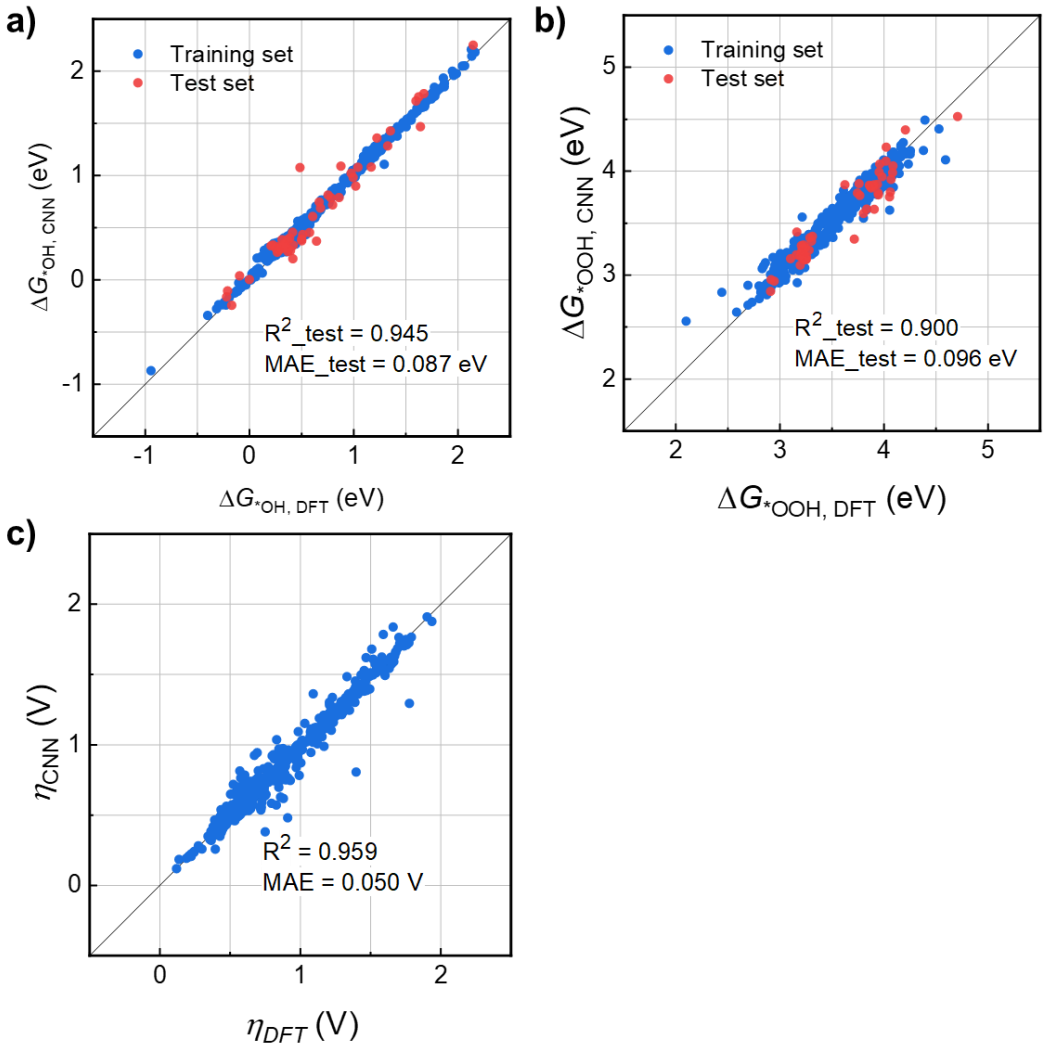


Supplementary Figure 6. The correlation between model-predicted and DFT-calculated ΔG_{*OOH}.

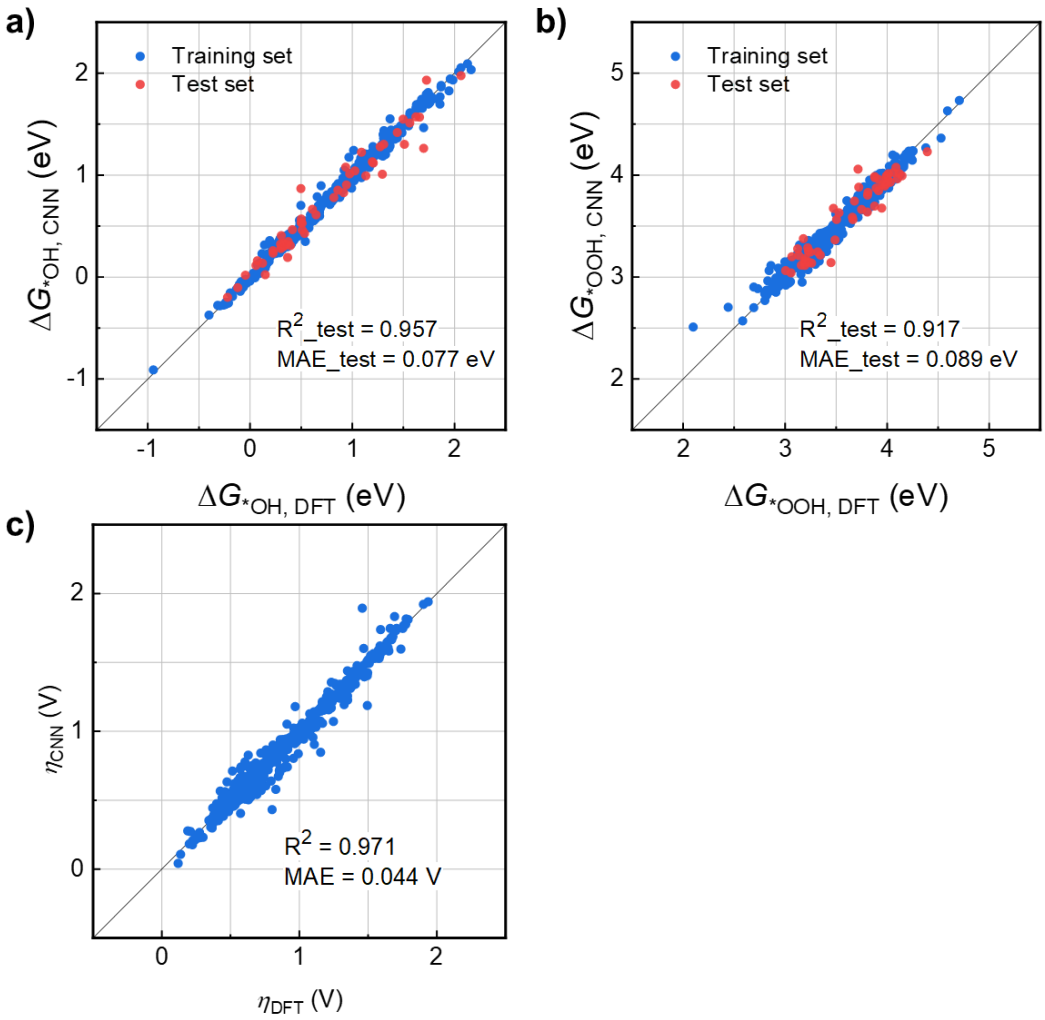
The CNN model was trained only based on DOS of active sites.



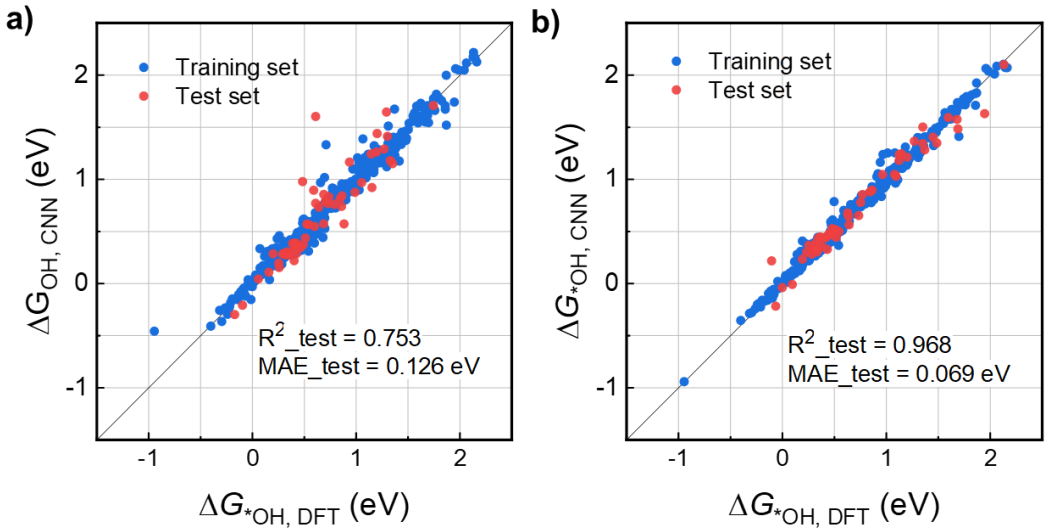
Supplementary Figure 7. Two optimized structures of *OOH on RuO_2 surface.



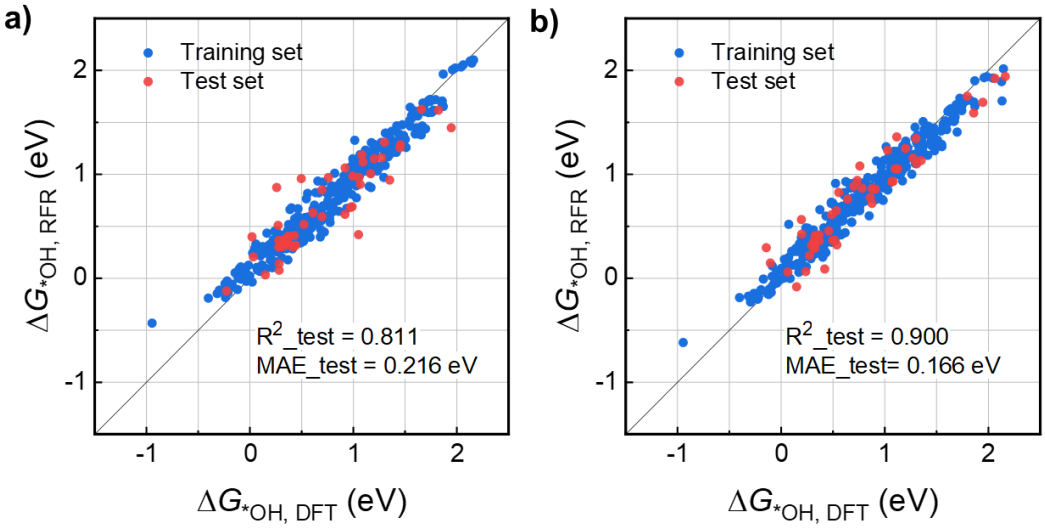
Supplementary Figure 8. The correlations between ΔG_{*OH}, ΔG_{*OOH} and η calculated by DFT and predicted by the CNN models. (a) and (b) show the prediction accuracy of ΔG_{*OH} and ΔG_{*OOH}, respectively. (c) shows the prediction accuracy of η. These CNN models were trained based on DOS of adsorbed oxygen in *O intermediates.



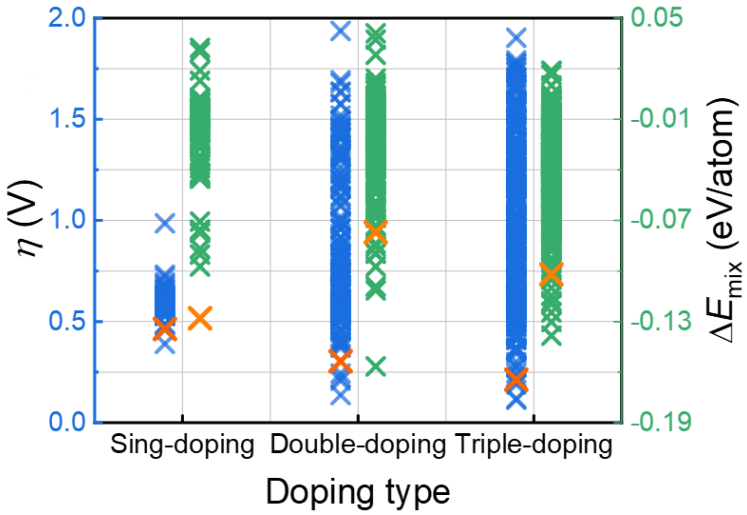
Supplementary Figure 9. The correlations between ΔG_{*OH}, ΔG_{*OOH} and η calculated by DFT and predicted by the CNN models. (a) and (b) show the prediction accuracy of ΔG_{*OH} and ΔG_{*OOH}, respectively. (c) shows the prediction accuracy of η. The CNN models were trained based on both DOS of active center and adsorbed oxygen in *O intermediates.



Supplementary Figure 10. The correlations between ΔG_{*OH} calculated by DFT and predicted by CNN. (a, b) The CNN modes were trained based on *s*, *p*-DOS and *d*-DOS of active centers, respectively.

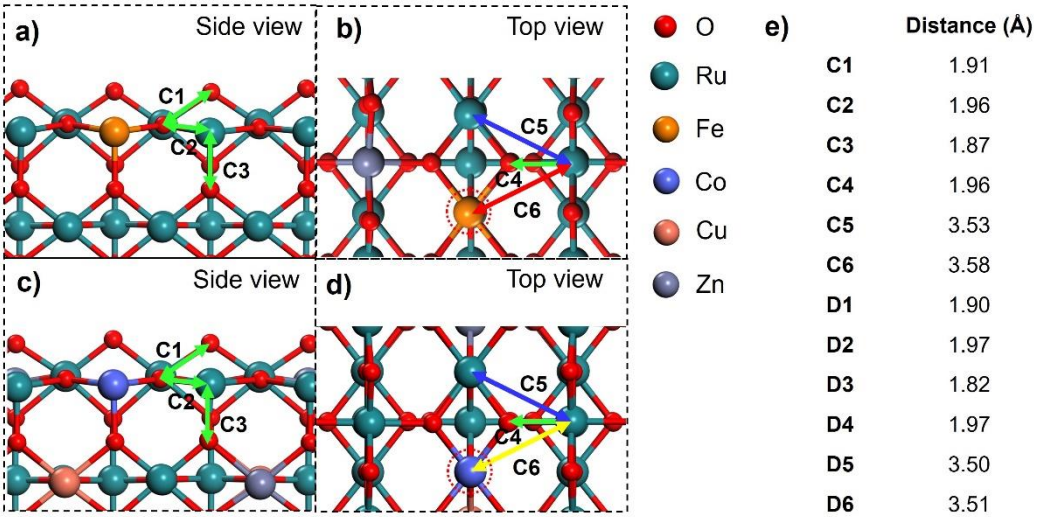


Supplementary Figure 11. The correlations between ΔG_{*OH} calculated by DFT and predicted by RFR. (a, b) The RFR models were trained based on the top five principal components of original *d*-DOS and convolutionally processed *d*-DOS of active center, respectively.

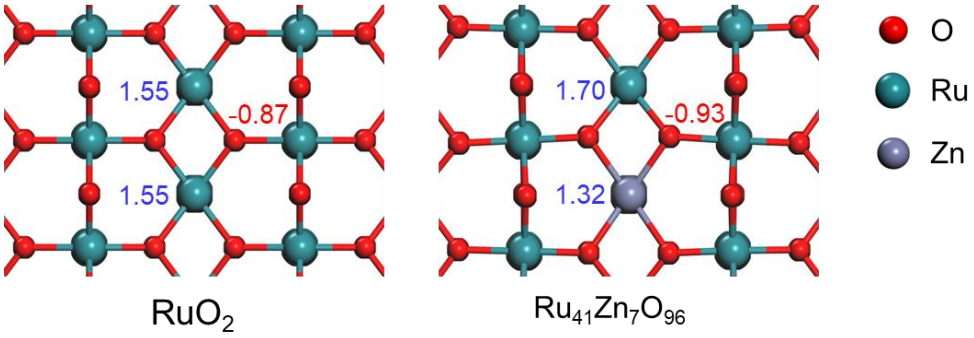


Supplementary Figure 12. The η and ΔE_{mix} distribution of mono-, double-, triple-doped RuO_2.

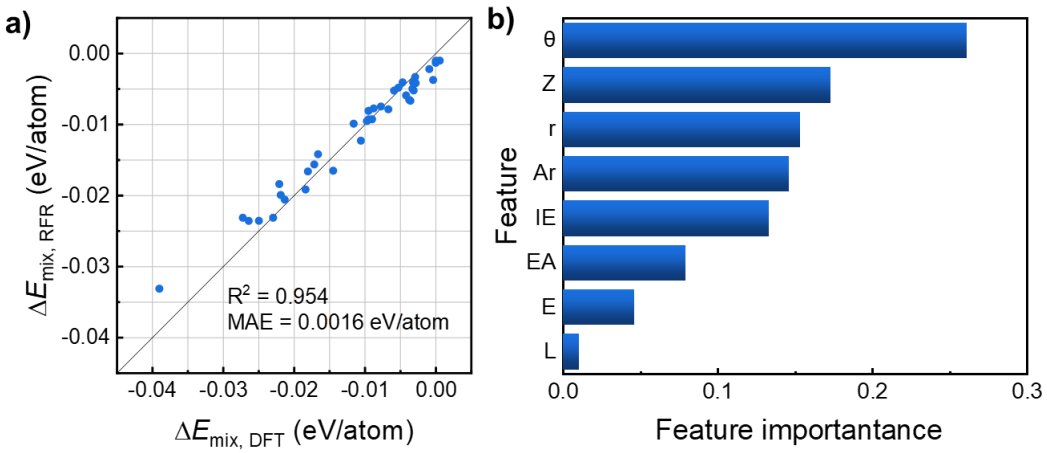
The orange cross marks indicate screened doping structures with both good stability and activity.



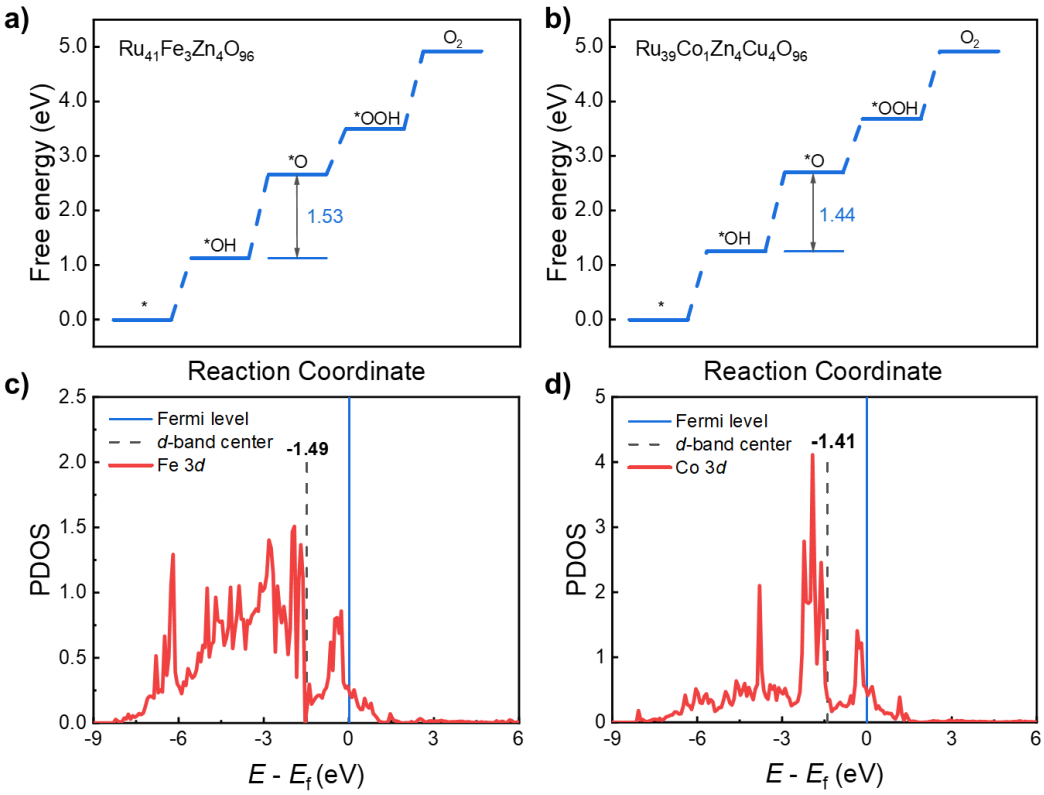
Supplementary Figure 13. Schematic diagrams of distance values of Ru-O, Ru···Ru, Ru···Fe, Ru···Co on the surfaces of $Ru_{41}Zn_4Fe_3O_{96}$ and $Ru_{39}Zn_4Cu_4Co_1O_{96}$. Green, blue, red and yellow double arrows indicate Ru-O, Ru···Ru, Ru···Fe, and Ru···Co, respectively.



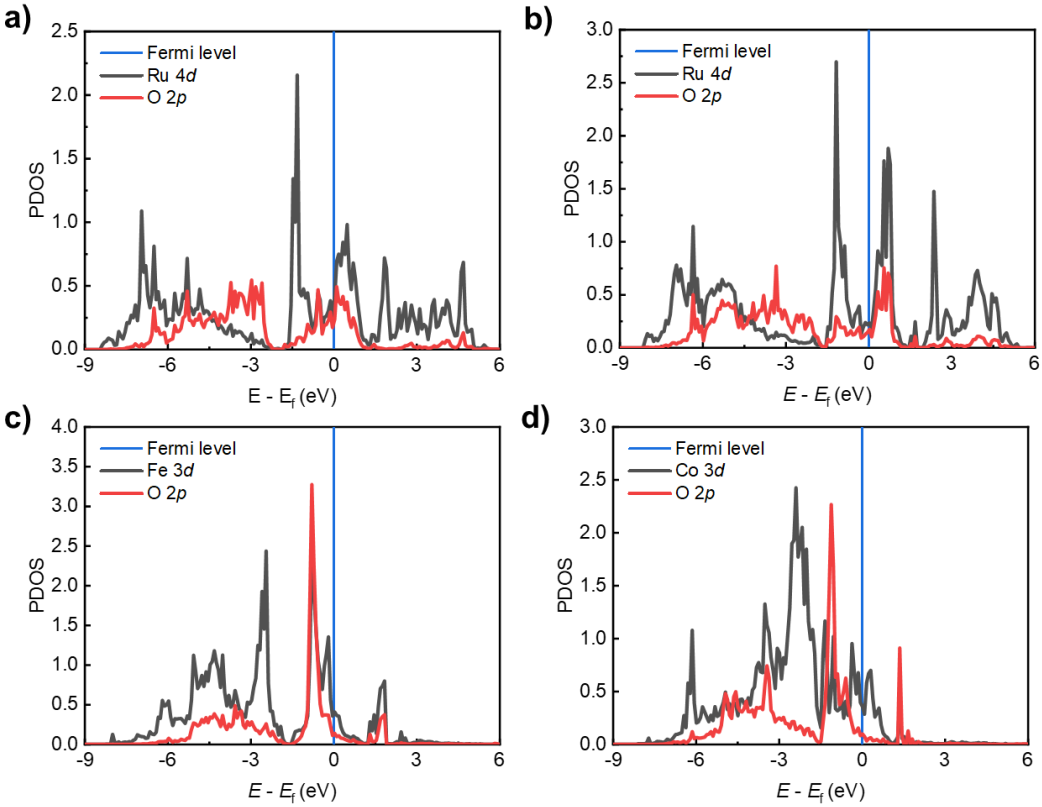
Supplementary Figure 14. Bader charge analysis for Ru, Zn, and O atoms on the surfaces of RuO_2 and $Ru_{41}Zn_7O_{96}$.



Supplementary Figure 15. (a) The correlation between RFR-predicted ΔE_{mix} and DFT-calculated ΔE_{mix}. (b) The relative importance of features based on the RFR model of ΔE_{mix}.

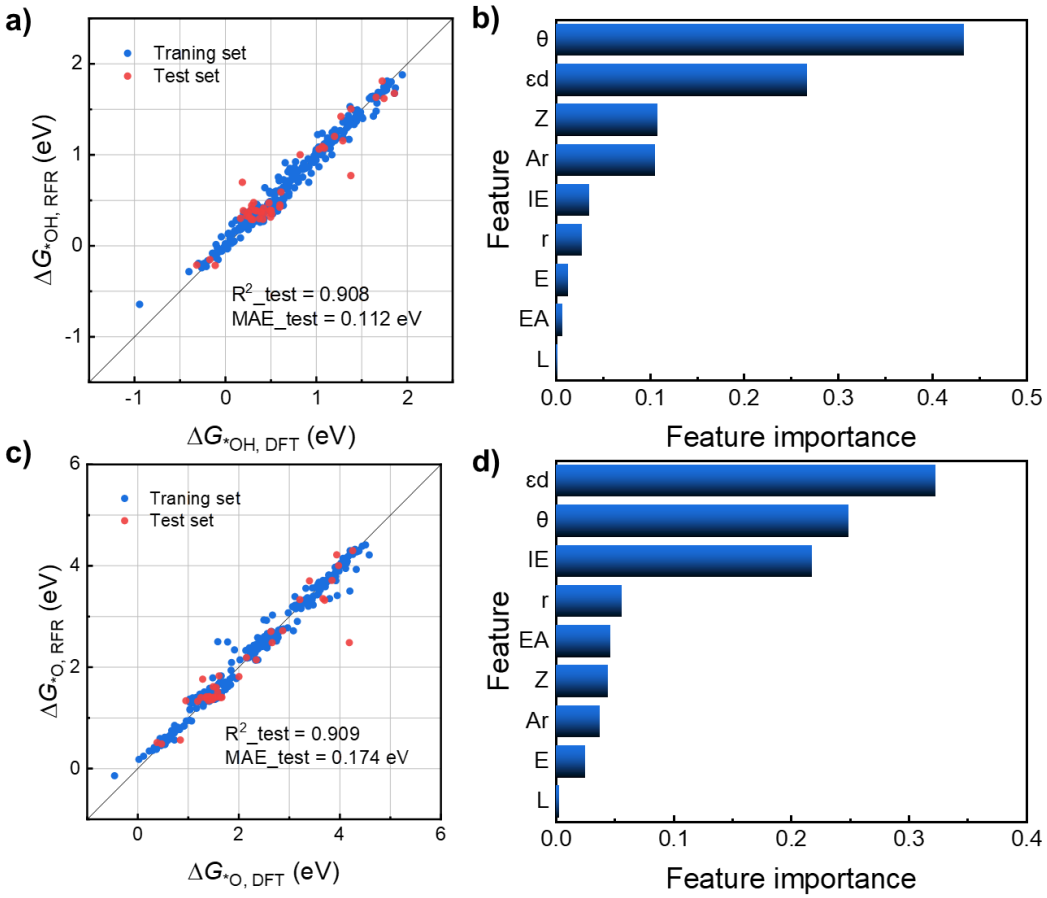


Supplementary Figure 16. The free energy diagrams for AEM path on the surfaces of (a) $Ru_{41}Zn_4Fe_3O_{96}$ and (b) $Ru_{39}Zn_4Cu_4Co_1O_{96}$. (c) and (d) display projected *d*-DOS of active sites corresponding to (a) and (b), respectively.

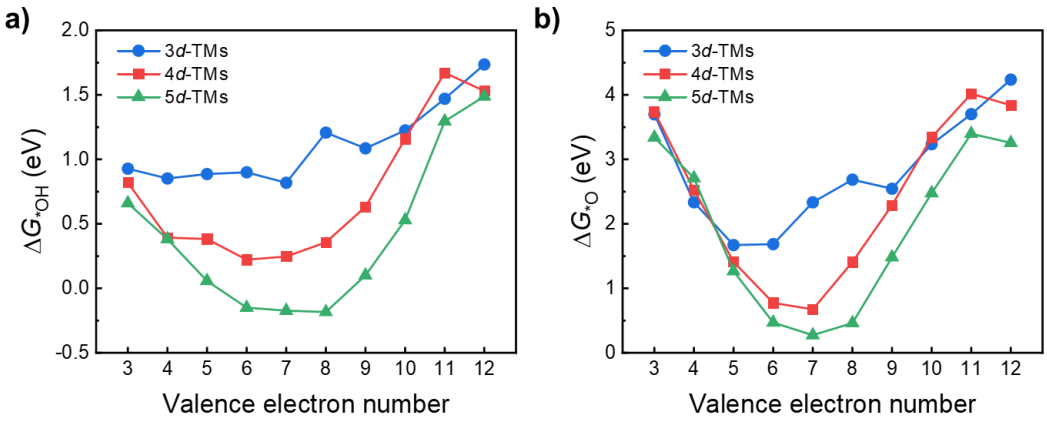


Supplementary Figure 17. Projected DOS of active site and adsorbed oxygen in *O intermediates.

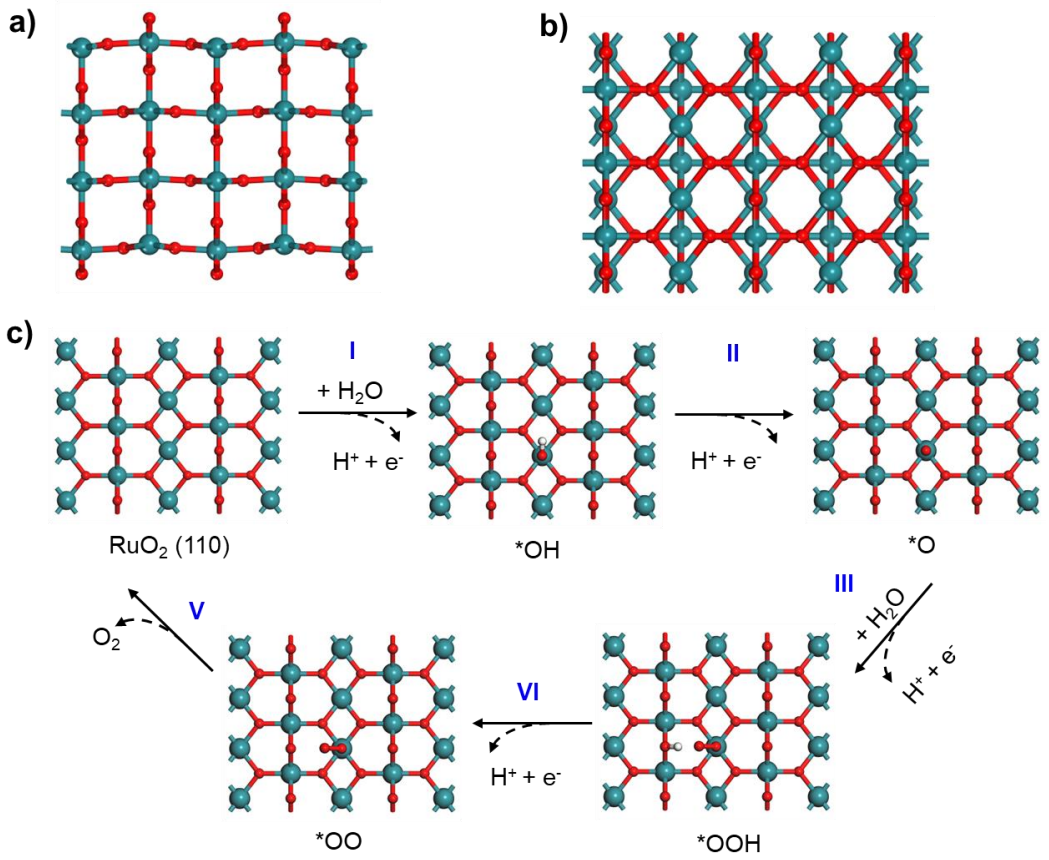
(a-d) correspond to RuO_2, $Ru_{41}Zn_7O_{96}$, $Ru_{41}Zn_4Fe_3O_{96}$ and $Ru_{39}Zn_4Cu_4Co_1O_{96}$, respectively.



Supplementary Figure 18. (a) The correlation between RFR-predicted ΔG_{*OH} and DFT-calculated ΔG_{*OH}. (b) The relative importance of features based on the RFR model of ΔG_{*OH}. (c) The correlation between RFR-predicted ΔG_{*O} and DFT-calculated ΔG_{*O}. (d) The relative importance of features based on the RFR model of ΔG_{*O}. All non-TMs are excluded in above results.

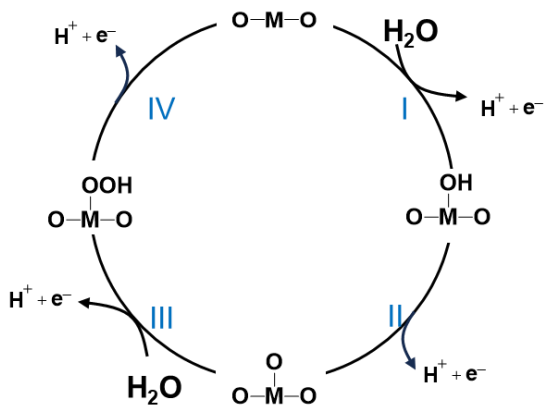


Supplementary Figure 19. The changes of (a) ΔG_{*OH} and (b) ΔG_{*O} across the valence electron number.



Supplementary Figure 20. (a) and (b) show the side view and top view of RuO_2 slab model, respectively. 15 Å vacuum space was added on the top of the slab. (c) AEM path on RuO_2 (110) surface.

a)



b)

$$* + H_2O \rightarrow *OH + H^+ + e^- \qquad \Delta G_1 = \Delta G_{*OH} \qquad (1)$$

$$*OH \rightarrow *O + H^+ + e^- \qquad \Delta G_2 = \Delta G_{*O} - \Delta G_{*OH} \qquad (2)$$

$$*O \rightarrow *OOH + H^+ + e^- \qquad \Delta G_3 = \Delta G_{*OOH} - \Delta G_{*O} \qquad (3)$$

$$*OOH \rightarrow O_2 + H^+ + e^- \qquad \Delta G_4 = 4.92 - \Delta G_{*OOH} \qquad (4)$$

$$\Delta G_{*OH} = G_{*OH} + 1/2G_{H_2} - G^* - G_{H_2O} \qquad (5)$$

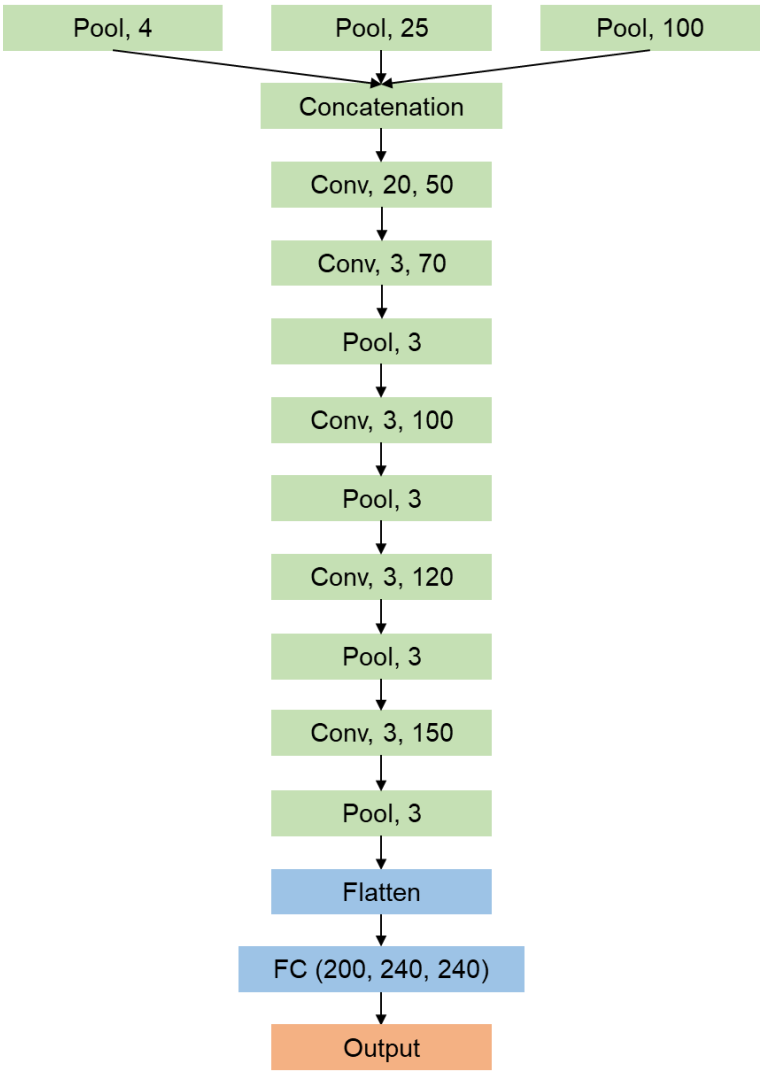
$$\Delta G_{*O} = G_{*O} + G_{H_2} - G^* - G_{H_2O} \qquad (6)$$

$$\Delta G_{*OOH} = G_{*OOH} + 3/2G_{H_2} - G^* - G_{H_2O} \qquad (7)$$

$$\Delta G_{O_2} = G_{H_2O} + 2G_{H_2} - 4 \times 1.23(eV) \qquad (8)$$

$$\eta = \max(\Delta G_1, \Delta G_2, \Delta G_3, \Delta G_4)/e - 1.23(eV) \qquad (9)$$

Supplementary Figure 21. a The schematic diagram of adsorbate evolution mechanism (AEM) and (**b**) theoretic calculation formulas of free energy barrier of every step and overpotential.



Supplementary Figure 22. The architecture of CNN model. Pool represents average pooling and the number on the right is pooling size. Conv represents convolution and numbers on the right are convolutional kernel size and the number of filters. FC represents fully connected network.

Supplementary Table 1. The prediction performance of CNN models.

Input	Output	Dataset	R^2	MAE (eV)	RMSE (eV)
DOS of active sites	ΔG_{*OH}	Training Set	0.992	0.033	0.047
		Test Set	0.968	0.074	0.104
	ΔG_{*O}	Training Set	0.991	0.070	0.108
		Test Set	0.968	0.142	0.189
	ΔG_{*OOH}	Training Set	0.971	0.055	0.072
		Test Set	0.924	0.082	0.105
DOS of adsorbed oxygens	ΔG_{*OH}	Training Set	0.994	0.029	0.039
		Test Set	0.945	0.087	0.126
	ΔG_{*OOH}	Training Set	0.946	0.066	0.091
		Test Set	0.900	0.096	0.131
Total DOS of active sites and adsorbed oxygens	ΔG_{*OH}	Training Set	0.994	0.026	0.040
		Test Set	0.957	0.077	0.116
	ΔG_{*OOH}	Training Set	0.971	0.051	0.069
		Test Set	0.917	0.089	0.111

Supplementary Table 2. The comparison of predicted and calculated ΔE_{mix} of filtered structures.

Structures	Predicted ΔE_{mix} (eV/atom)	Calculated ΔE_{mix} (eV/atom)
$Ru_{41}Zn_7O_{96}$	-0.119	-0.124
$Ru_{41}Zn_4Fe_3O_{96}$	-0.085	-0.077
$Ru_{39}Co_1Cu_4Zn_4O_{96}$	-0.107	-0.102

Supplementary Table 3. The comparison of predicted and calculated adsorption energies and overpotentials of filtered structures.

Structures	ΔG_{ads} or η	Predicted values	Calculated values
$Ru_{41}Zn_7O_{96}$	ΔG_{*OH} (eV)	0.467	0.587
	ΔG_{*O} (eV)	1.647	1.573
	ΔG_{*OOH} (eV)	3.370	3.267
	η (V)	0.493	0.464
$Ru_{41}Zn_4Fe_3O_{96}$	ΔG_{*OH} (eV)	1.135	1.126
	ΔG_{*O} (eV)	2.742	2.660
	ΔG_{*OOH} (eV)	3.529	3.495
	η (V)	0.377	0.304
$Ru_{39}Co_1Cu_4Zn_4O_{96}$	ΔG_{*OH} (eV)	1.188	1.252
	ΔG_{*O} (eV)	2.617	2.695
	ΔG_{*OOH} (eV)	3.670	3.677
	η (V)	0.199 V	0.213

Supplementary Table 4. The DFT-calculated η of pure RuO_2 and screened catalysts based on PBE and RPBE.

Structures	PBE (V)	RPBE (V)
RuO_2	0.617	0.698
$Ru_{41}Zn_7O_{96}$	0.493	0.487
$Ru_{41}Zn_4Fe_3O_{96}$	0.304	0.175
$Ru_{39}Co_1Cu_4Zn_4O_{96}$	0.213	0.078