Energy and AI 22 (2025) 100606

Contents lists available at ScienceDirect

Energy and AI

journal homepage: www.sciencedirect.com/journal/energy-and-ai

Opportunities and perspectives of artificial intelligence in electrocatalysts design for water electrolysis

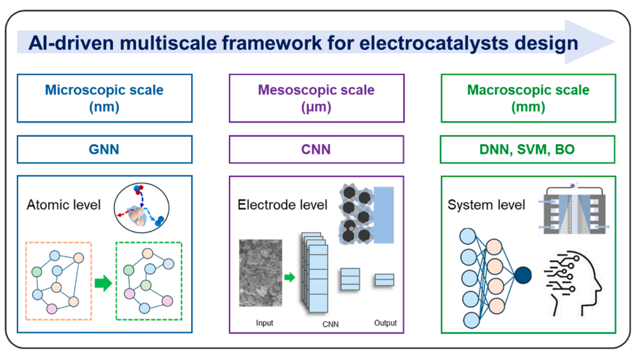
Qing Wang [a], Lizhen Wu [a], Qiang Zheng [b,*], Liang An [a,c,d,*]

[a] Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, China
[b] Zhejiang Key Laboratory of Industrial Intelligence and Digital Twin, Eastern Institute of Technology, Ningbo, Zhejiang 315200, China
[c] Research Institute for Advanced Manufacturing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, China
[d] Research Centre for Carbon-Strategic Catalysis, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, China

HIGHLIGHTS

- A comprehensive AI-driven multiscale design framework was synthesized for electrocatalyst design.
- A cross-scale, data-driven algorithm strategy integrates AI models based on domain-specific features.
- Generative AI and automation experimental techniques enable a shift toward autonomous electrocatalyst discovery.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Artificial intelligence
Electrocatalysts
Water electrolysis
Multiscale design
Automated experimentation

ABSTRACT

As a key pathway for green hydrogen production, water electrolysis is expected to play a central role in the future energy landscape. However, its large-scale deployment is hindered by challenges related to cost, performance, and durability. The emergence of artificial intelligence (AI) has transformed this field by offering powerful and efficient tools for the design and optimization of electrocatalysts. This review outlines an AI-driven multiscale design framework, highlighting its role at the microscopic scale for identifying atomic-level active sites and key descriptors, at the mesoscopic scale for structural and morphological characterization, and at the macroscopic scale for multi-objective optimization and intelligent control. This multiscale framework demonstrates the potential of AI to accelerate the development of next-generation electrocatalysts. In addition, the integration of generative AI and automated experimental techniques is highlighted as promising strategies to further enhance electrocatalyst discovery and promote the practical implementation of water electrolysis technologies.

* Corresponding authors.
E-mail addresses: qzheng@eitech.edu.cn (Q. Zheng), liang.an@polyu.edu.hk (L. An).

https://doi.org/10.1016/j.egyai.2025.100606

Available online 2 September 2025

1. Introduction

Water electrolysis has emerged as a promising and sustainable technology for energy conversion and storage, attracting increasing attention in recent years [1]. It plays a key role in addressing the global energy crisis and facilitating the efficient allocation and utilization of renewable energy sources [2]. Hydrogen, widely regarded as a key energy carrier of the future, can be produced through water electrolysis powered by renewable electricity. This process enables the generation of green hydrogen, offering an effective solution to the intermittency of renewable energy supply [3]. However, the large-scale industrial of electrolyzers still faces major challenges, particularly in terms of durability and cost-effectiveness.

The key reactions in water electrolysis involve two half-reactions: the hydrogen evolution reaction (HER) at the cathode and the oxygen evolution reaction (OER) at the anode. The theoretical Gibbs free energy change (ΔG) for the overall reaction is approximately 237.2 kJ/mol, corresponding to a thermodynamic potential of 1.23 V. Through these electrochemical processes, electrical energy is efficiently converted into chemical energy. Electrocatalysts are critical in facilitating these reactions by lowering the overpotentials and enhancing kinetics. Currently, the most effective HER and OER catalysts are based on precious metal nanomaterials such as platinum (Pt) and iridium (Ir) [2,4,5]. However, the scarcity and high cost of these materials severely limit their practical application in large-scale hydrogen production. Therefore, it is essential to develop electrocatalysts that are not only highly active and stable but also economically viable.

Traditionally, the design of electrocatalysts have relied heavily on trial-and-error approaches and empirical screening methods [6]. Although advances in characterization techniques and theoretical modeling have enabled researchers to explore catalytic activity and reaction mechanisms at the atomic level, significant challenges still remain [7]. These include difficulties in elucidating the true atomic structures of active sites from a multidimensional perspective, limited understanding of structure-property relationships, and the inherent complexity and dynamic nature of catalytic processes. As a result, current materials design strategies often suffer from inefficiencies in both time and resources.

Recent progresses in artificial intelligence (AI) have significantly enhanced research capabilities across a wide range of scientific and engineering disciplines [7], including computer vision, robotics, biomedical science, energy material science, and system engineering [8–10]. Among these developments, machine learning (ML) has emerged as a powerful tool for data-driven principal discovery and analysis. As the most popular and powerful branch of machine learning, deep learning is capable of analyzing large-scale datasets to uncover previously unknown patterns and correlations that may be overlooked by traditional methods [11]. This capability allows for accurate prediction of material properties and behaviors, thereby reducing the reliance on time-consuming and costly experimental and computational procedures. Moreover, machine learning approaches can be utilized to optimize synthesis and process parameters, further accelerating the material development cycle.

Although recent reviews have outlined the progress in applying AI, including ML and other data-driven techniques, to water electrolysis [12–14], there remains a lack of systematic discussion regarding its specific role in electrocatalyst design. Specifically, issues such as integrating information across atomic, mesoscale, and macroscale levels, selecting suitable algorithms for multiscale modeling, and enabling global optimization in electrocatalyst design have not been extensively explored. This underscores the need for a more target and in-depth review in this area.

To address these challenges, this review organized recent advanced researches into an AI-driven multiscale design framework for electrocatalyst design. This framework is structured according to the key objectives at different scales. Herein, at the microscopic atomic scale, the focus is on identifying active sites, reactive mechanisms, and performance descriptors; at the mesoscopic scale, the attention is given to morphological features and interfacial structures; and at the macroscopic scale, the emphasis is the system-level catalyst performance prediction and optimization. By integrating insights across these scales, this review aims to provide a theoretical foundation and guidance for the rational design of next-generation electrocatalysts for water electrolysis.

2. Artificial intelligence approaches

Facing the challenges of complex electrochemical systems, the synergy between AI and existing computational methods is creating new opportunities for more comprehensive and faster theoretical analysis [15]. Traditional first-principles methods, such as density functional theory (DFT), provide quantitatively accurate predictions but incur prohibitively high computational costs. In contrast, interatomic potentials offer substantially reduced computational demands. However, their applicability is constrained by limitations in transferability and inherent nonreactive issues in simulating reactive processes, restricting broader implementation [16]. ML, as a practical AI framework based on computer science and statistics, enables algorithms to autonomously learn from collected data and predict target properties without explicit programming [17]. This data-driven ability allows ML to handle multiscale input variables, from atomic-level descriptors and mesoscopic nanostructures to macroscopic system parameters, facilitating global optimization of electrochemical systems.

Unlike traditional trial-and-error approaches that depend heavily on experience or sequential experimentation, ML not only offers high predictive accuracy, but more importantly, provides exceptional computational efficiency. This efficiency allows rapid learning from and screening of large datasets, enabling comprehensive exploration of complex, high-dimensional design spaces. As a result, ML can identify and learn structure-property relationships from data, and predict properties of newly designed materials with high accuracy and efficiency, thereby overcoming the inherent limitations of traditional trial-and-error approaches [18,19].

As a core branch of ML, deep learning captures complex dependencies and interactions within material structures through multilayer nonlinear transformations. However, different algorithms vary widely in terms of accuracy and applicability, as shown in Table 1. For example, graph neural networks (GNNs) represent atoms and their bonding relationships using graph structures. By incorporating directional information, such as bond angles and torsional interactions, GNNs are capable of modeling atomic-level features and extracting detailed descriptors. However, they are less efficient in representing mesoscopic structures with larger spatial scales. Convolutional neural networks (CNNs), known for their strength in recognizing spatial features, are mainly used in image analysis and for extracting characteristics of mesostructures such as electrode materials. However, they are less suited to disordered systems like amorphous clusters, often requiring preprocessing techniques to extract indirect statistical features. Artificial neural networks (ANNs) are adept at handling complex nonlinear problems and mapping relationships from mesoscopic features to macroscopic performance. They can uncover latent correlations between material properties for performance prediction and multi-objective optimization, but lag in automatic feature extraction. Deep neural networks (DNNs), as representation learning models, transform input features into new spaces through multiple hidden layers and are well-suited for structure-property modeling in large datasets. However, their effectiveness relies heavily on the availability of sufficient high-quality training data.

Classical supervised learning methods also play important roles. Support vector machines (SVMs), for example, use kernel functions to map data into higher-dimensional spaces, enabling material feature classification even in small-sample scenarios, though their general

Table 1
List of various AI approaches for electrocatalysts design.

Algorithm	Typical scenario	Advantages	Disadvantages	Ref.
GNN	Microscale atomic bonding modeling; Prediction of active sites.	Directly process atomic graph structures.	Less efficient for grid-structured data; Slow convergence speed.	[20, 21]
CNN	Mesoscopic morphology analysis.	Extraction of local spatial features; Capturing nonlinear spatial relationships.	Require structured input; Deep networks may cause vanishing gradients.	[22, 23]
ANN	Macro performance prediction; Multi-parameter nonlinear regression.	Fitting highly nonlinear relationships.	Limited automatic feature extractor; Depend on handcrafted input descriptors	[24, 25]
DNN	Multi-objective prediction; Multi-parameter nonlinear regression.	Fitting highly nonlinear relationships; Automatic feature extraction.	Relying on large training data.	[26, 27]
SVM	Descriptor classification; Small-sample learning tasks.	Strong generalization ability on small datasets; Effective with kernel functions.	High reliance on feature engineering.	[28, 29]
BO	High-dimensional parameter tuning; Automated experimental design.	Efficient global optimization.	Sensitive to acquisition function design.	[30, 31]

applicability often depends on manual feature engineering. Gaussian process regression (GPR) provides both performance predictions and uncertainty estimations but is computationally intensive and less practical for large-scale datasets due to the curse of dimensionality. Bayesian optimization (BO) offers efficient hyperparameter tuning in high-dimensional spaces, reduces the number of experimental iterations, and supports multi-objective optimization. However, its performance is highly sensitive to prior distributions and surrogate model hyper-parameters, requiring careful trade-offs based on dimensionality, objective function characteristics, and model choice.

Therefore, given the complementary capabilities of various AI algorithms across different representation levels and application scales, this review summarizes an AI-driven multiscale design framework to support electrocatalyst design, as illustrated in Fig. 1. This framework leverages different algorithms based on their strengths at each scale. GNNs, with their high efficiency in processing graph-structured data, are typically employed for analysis of atomic configurations and local bonding environments at microscopic scale. CNNs, excelling at capturing spatial patterns from imaging data, are applied to morphology characterization at mesoscopic scale. The techniques such as BO, SVM, and DNNs, offering strong capabilities in global optimization and nonlinear regression, are essential for system-level optimization at macroscopic scale. Altogether, this framework provides a structured perspective to guide future research on the data-driven design of advanced electrocatalysts. It is worth noting that these algorithms are not strictly confined to individual scales. They are often used in combination across scales to address complex tasks, reflecting their complementary nature in electrocatalyst design.

3. AI-driven multiscale framework for electrocatalyst design

The catalytic activity of a material is closely linked to its unique atomic structure and physicochemical properties [32]. To overcome performance limitations, researchers have employed various strategies to modify the physicochemical characteristics, electronic structure, and surface properties of electrocatalytic materials, thereby enhancing their activity [33]. To date, a wide range of active sites has been developed through activity-enhancement strategies such as doping engineering, defect engineering, and strain engineering [34]. A common feature of these approaches is their ability to induce synergistic effects among active sites by altering the lattice topology or local distribution within the catalyst, which significantly boosts its intrinsic catalytic performance [35,36]. On the other hand, the practical effectiveness of electrocatalysts is also influenced by the distribution and stability of their nanostructures, as well as by the integration of materials within component architectures. Therefore, this section systematically

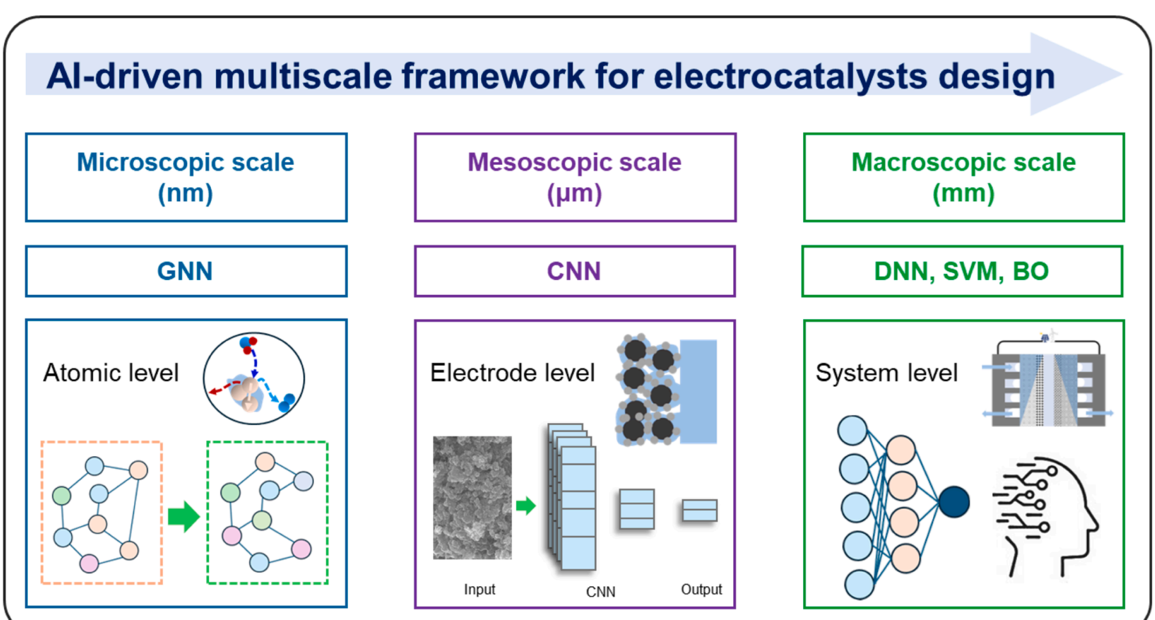


Fig. 1. Schematic of AI-driven multiscale framework for electrocatalysts design.

discusses recent advances in the application of AI for the modeling and optimization of electrocatalysts across three distinct scales: the microscopic scale for atomic structure, the mesoscopic scale for nanostructure configuration, and the macroscopic scale for system integration.

3.1. Microscopic scale: atomic-level active site identification

Understanding the atomic structure of catalytic active sites and analyzing their structure-property relationships is crucial for optimizing catalyst design. One of the advantages of data-driven modeling approaches is their ability to bypass complex and time-intensive procedures, such as solving quantum mechanical equations or elucidating reaction mechanisms, thereby significantly accelerating the research timeline [37,38]. The accuracy and reliability of such modeling approaches depend on three key factors: the computational method, the quality and scale of the dataset, and the selection of appropriate descriptors. Among these, descriptors serve as the foundation of model development and require a deep understanding of materials science to accurately reflect the physical and chemical characteristics of materials. Effective descriptors must comprehensively capture both the fundamental physicochemical properties and the unique structural features of the system under investigation [39].

In the context of electrocatalyst design, descriptors are commonly classified into three categories. The first includes structural descriptors derived from geometric features, such as atomic radius, atomic number, group number, coordination number, and lattice constants. The second category involves electronic descriptors, which describe the electronic properties of materials and typically require first-principles calculations to obtain accurate electronic structures, making them computationally demanding. The third category consists of activity-related descriptors, which represent a material's ability to donate or accept electrons or charged species, such as adsorption energy and electronegativity. In general, ideal descriptors should be simple, readily accessible, and of low dimensionality to improve the efficiency of catalyst screening and performance prediction.

Recent studies have demonstrated the importance of identifying and selecting the most relevant descriptors for modeling catalytic activity in complex systems. As shown in Fig. 2(a) and (b), Jäger et al. evaluated the effectiveness of several advanced structural descriptors, including the smooth overlap of atomic positions, many-body tensor representations, and atom-centered symmetry functions, in predicting hydrogen adsorption energies on nanocluster surfaces [40]. Their results indicated that using a unified predictive model across different nanocluster types, rather than training separate models, substantially reduced the mean absolute error. Moreover, integrating data from multiple sources significantly improved the accuracy of potential energy surface fitting. In another study, Fung et al. proposed a generalized descriptor for oxygen sites in transition metal oxides by adjusting the coordination number [41], successfully establishing a correlation between structural features and catalytic reactivity. Building on this approach, as illustrated in Fig. 2(c) and (d), Tran et al. developed an automated screening framework, integrating predictive modeling with optimization strategies, to identify key descriptors associated with catalytic activity or microkinetic parameters [42]. Their framework screened a dataset of 102 alloy compositions and identified 258 promising candidate surfaces for the hydrogen evolution reaction. Subsequent stability assessments further refined the selection of optimal surfaces. Based on the above cases, these examples highlight the central role of well-designed descriptors in enhancing the predictive accuracy of computational models and advancing the rational design of high-performance catalysts.

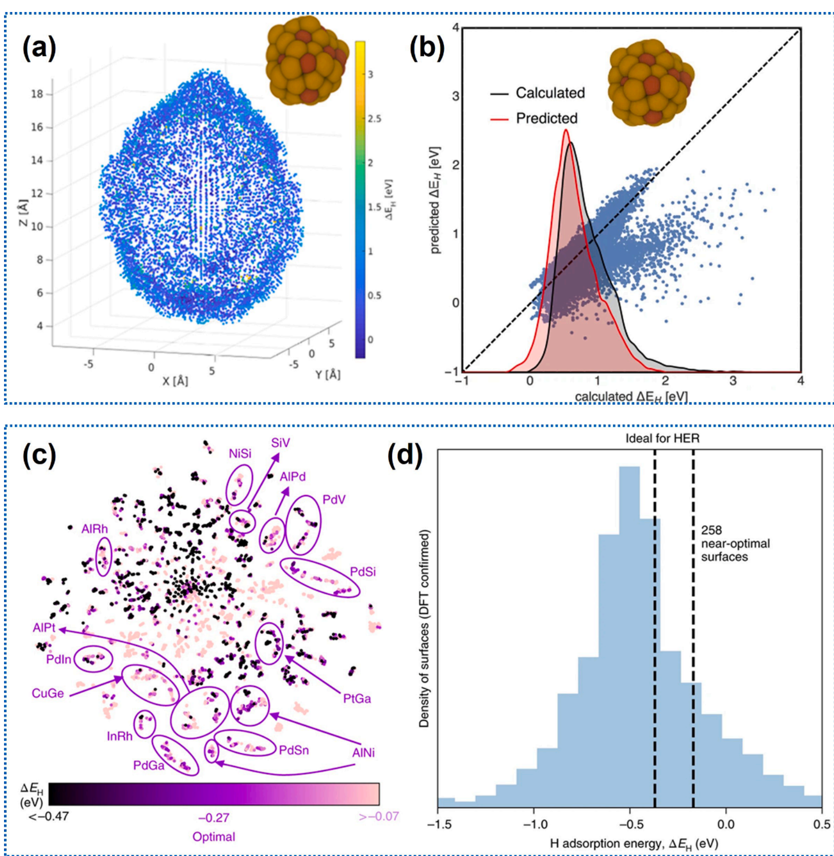


Fig. 2. (a) Hydrogen position scan on the surface of a $Au_{40}Cu_{40}$ cluster; (b) Parity plot of predicted against calculated ΔE_H together with a histogram of predicted (red) and calculated (black) energy distributions. Reproduced with permission from Ref. [40]. (c) t-SNE visualization of all the adsorption sites simulated with DFT; (d) Normalized distribution of low-coverage ΔE_H values calculated by the DFT workflow. Reproduced with permission from Ref. [42].

The application of ANN and GNN, in combination with DFT calculations, provides an effective approach for exploring the intricate relationships between electrocatalytic performance and key material descriptors at the atomic scale [10,34,43]. These methods have proven valuable in identifying active sites and guiding the rational design of advanced catalytic materials. In contrast to traditional trial-and-error strategies that rely heavily on experimental intuition, ANN can capture complex nonlinear correlations between material properties and catalytic behavior by simulating the connectivity patterns of biological neurons [44,45]. This facilitates data-driven catalyst development through multidimensional optimization involving composition tuning, structural control, and performance prediction [46,47], enabling a more precise balance among catalytic activity, stability, and application-specific requirements. GNN, on the other hand, are well suited for representing materials with graph-based structures, such as atomic lattices or molecular frameworks, allowing for effective extraction of spatial and chemical information critical to catalytic function. These models are particularly advantageous for learning from atomistic configurations and for identifying descriptors that are not readily captured by traditional feature sets.

Beyond static prediction, AI approaches also support the investigation of dynamic links between atomic-scale structural characteristics, such as coordination environments and lattice distortions, and macroscopic catalytic outcomes [48,49]. With the development of hierarchical materials databases that incorporate physicochemical properties, synthesis conditions, and structure-property relationships, these modeling tools establish a theoretical foundation for rapidly screening and matching optimal active site configurations to specific catalytic applications. The utility of such methods has been demonstrated in a variety of electrocatalyst studies. As shown in Fig. 3(a), Palkovits et al. [50] employed ANN models in combination with SVM and k-nearest neighbor regression to accurately predict active sites based on a dataset of water oxidation catalysts. In Fig. 3(b), Wang et al. [51] introduced the theory-infused neural network (TinNet) model, which combined deep learning with d-band theory to maintain high prediction accuracy while offering insights into bonding characteristics. For alloy catalysts, a three-layer backpropagation ANN, as illustrated in Fig. 3(c), was used to predict HER activity of previously untested alloy surfaces through iterative weight optimization [52]. These studies demonstrate the capacity of ANN-based models to reveal structure-property relationships with high predictive power. In another example, shown in Fig. 3(d), Zhou et al. [53] developed a multiscale convolutional kernel model informed by topological features, using GNN-based techniques alongside DFT calculations to investigate the HER performance of carbon-supported single-atom catalysts under different coordination environments. This approach significantly enhanced the accuracy of catalytic performance predictions. Collectively, these modeling strategies represent a promising direction for understanding catalytic mechanisms and accelerating the development of novel electrocatalysts from theoretical design to practical implementation [54].

Overall, AI-driven microscopic modeling approaches have shown great promise in elucidating catalytic mechanisms, identifying active sites, and selecting optimal descriptors. However, atomic-level information alone is insufficient to meet the requirements of industrial applications. Therefore, mesoscopic structural and interfacial features which are related to the transport pathways for reactants and products need to be considered to bridge above gap. And descriptors obtained at the atomic scale (such as adsorption energy, d-band center) and active site configuration will serve as input parameters for mesoscopic structure optimization models to guide morphology design and interface engineering.

3.2. Mesoscopic scale: structural characterization and interface engineering

The performance of electrocatalysts in water electrolysis is

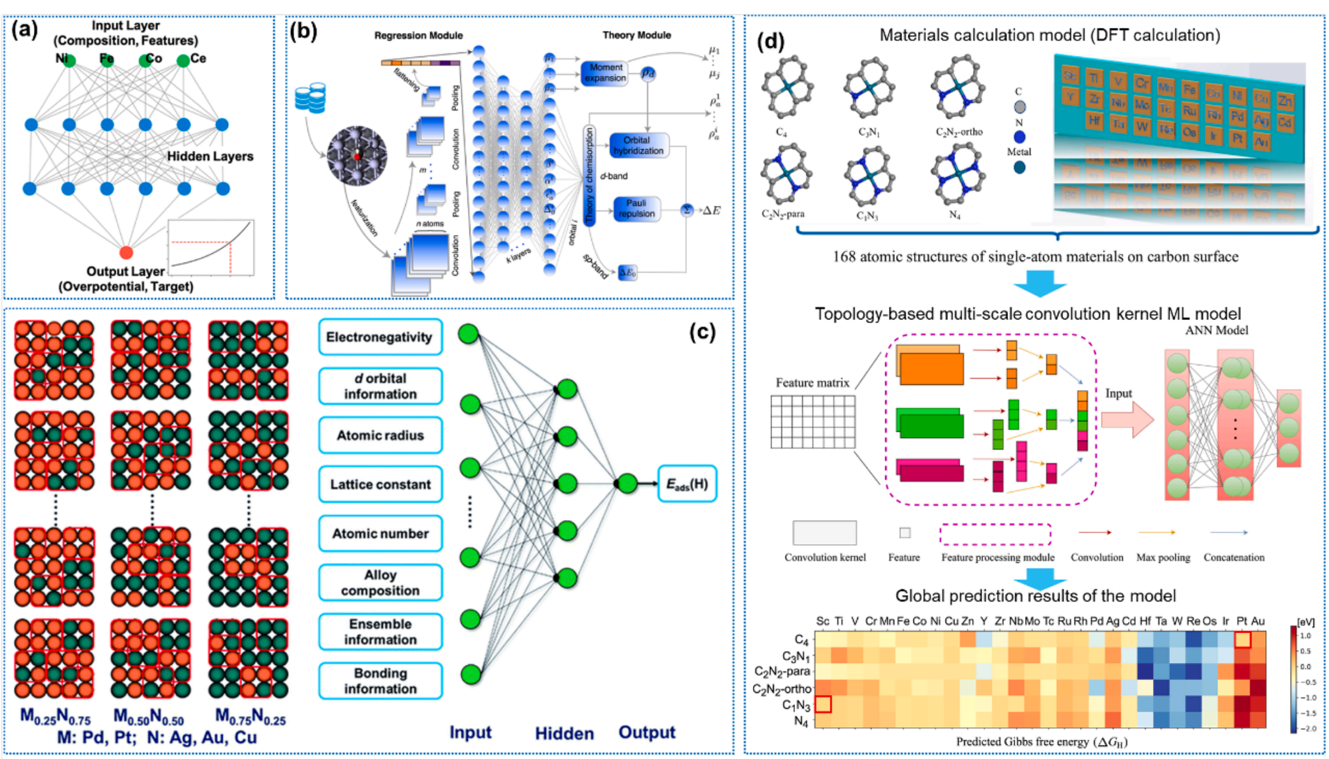


Fig. 3. (a) Artificial neural network with four input nodes (features), two hidden layers with six nodes each, and one output node (target). Reproduced with permission from Ref. [50]. (b) Schematic illustration of the TinNet framework. Reproduced with permission from Ref. [51]. (c) Schematic pictures of the random sampling method and algorithmic architecture of the BPNN model for (100) bimetallic alloys. Reproduced with permission from Ref. [52]. (d) Data-driven rational design of single-atom materials for hydrogen evolution. Reproduced with permission from Ref. [53].

influenced by the interplay of multiscale parameters [55]. Beyond atomic-level actives sites, mesoscopic features such as grain size, pore architecture, and interfacial orientation also play a critical role in determining catalytic activity and stability. In addition, analysis at the mesoscopic scale serves as a bridge between microscopic atomic characteristics and macroscopic device performance. Therefore, the design and optimization of electrocatalysts for practical applications must comprehensively consider key parameters such as the morphology and configuration of catalysts and electrodes, the spatial distribution of components, and mass transport processes in mesoscopic scale.

Image-based learning algorithms have significantly improved the efficiency of structural characterization and optimization of electrodes and catalysts. As the most popular AI models for learning image-like data, CNNs employ multiple processing layers to identify spatial patterns in datasets with grid-like topologies. By taking image pixels as input, CNNs learn and transmit information across layers and eventually output target properties through fully connected neural networks. Liu et al. used CNN as a fast surrogate model for finite element analysis (Fig. 4(a)), establishing a nonlinear mapping between binary anode microstructures and macroscopic performance metrics such as triple-phase boundaries, which substantially reduced computation time and improved model efficiency [56]. The input images for CNNs can be either experimentally obtained scanning electron microscope (SEM) and transmission electron microscope (TEM) images or generated by numerical simulations. Karaca et al. developed a fully automated tool, PoreD^2 (Fig. 4(b)), which quantifies pore and window structures by scanning SEM images, greatly reducing the time required for morphological characterization [57]. For analyzing three-dimensional hierarchical morphologies, X-ray tomography is often required. In the study by Topal et al., region-based CNNs (R-CNNs) were applied to extract feature maps and monitor nanostructure changes (Fig. 4(c)) [58]. This method corrected imaging artifacts in high-resolution scans and compensated for experimental inaccuracies, greatly enhancing the efficiency of catalyst development.

At the mesoscopic scale, the behavior of the catalyst-electrolyte interface also plays a key role in determining performance. Tao et al. proposed a deep neural network enhanced mesoscopic thermodynamic model (DeepMT) based on the DeepONet framework, which maps the microscopic properties of ions to macroscopic field responses and enables rapid prediction of interfacial characteristics [59].

In summary, image-based AI approaches, especially algorithms such as CNNs, have significantly improved the efficiency of structural characterization and optimization of electrode materials at the mesoscopic scale. In addition, by combining X-ray tomography and deep learning technology, researchers can more accurately reconstruct complex pore structures and identify key interface features, providing a solid foundation for future performance prediction and material design. Emerging neural network models also provide the interfacial behavior between

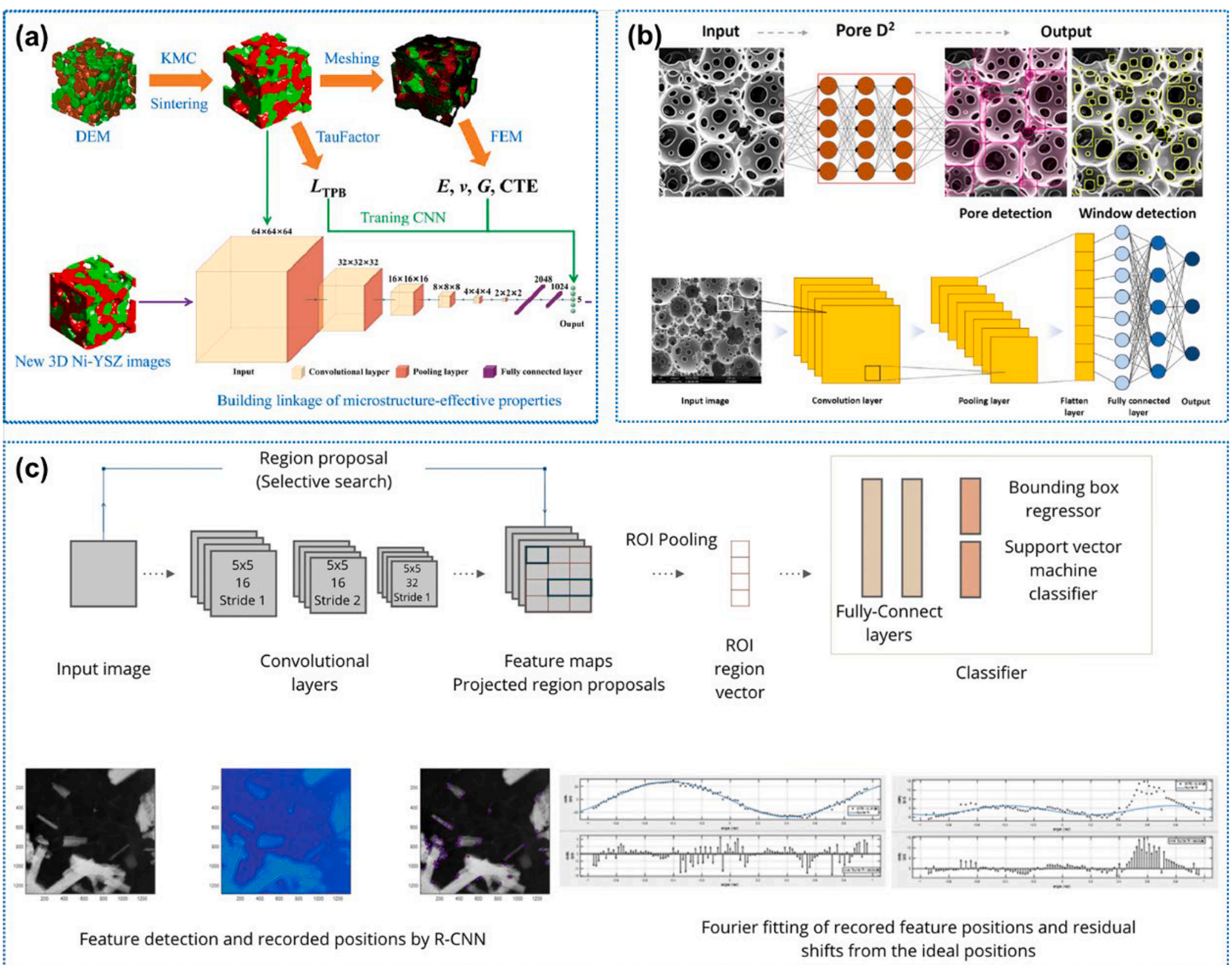


Fig. 4. (a) Deep learning accelerated numerical homogenization framework. Reproduced with permission from Ref. [56]. (b) Evaluation of pore and window size on SEM images via deep learning. Reproduced with permission from Ref. [57]. (c) R-CNN for feature detection and tracking. Reproduced with permission from Ref. [58].

catalysts and electrolytes. The morphology structure and interface properties at the mesoscopic scale not only affect the exposure and stability of active sites, but also directly determine the mass transport behavior and synergistic effects of catalyst in macroscopic systems. Therefore, optimization at the mesoscopic scale provides an important bridge of system-level performance regulation, further promoting AI-driven electrocatalyst design from material discovery to device integration.

3.3. Macroscopic scale: system-level prediction and optimization

The development and design of catalysts must balance several key factors, including catalytic activity, stability, cost, and overall device performance [54,60,61]. However, due to the complex interactions among material components during device integration and the variability introduced by external operational conditions such as temperature and pressure [62,63], traditional experimental and development workflows are often inefficient. These approaches suffer from slow model validation, significant material and time consumption, and are limited in their ability to provide comprehensive insights [64]. At the macroscopic scale, it is essential to establish clear correlations between the meso- and nano- structure of the catalyst layer and the overall performance of the electrochemical device, thereby proposing optimization strategies for system-level integration. Hence, the implementation of intelligent optimization and control strategies at the macroscopic scale is essential [65,66].

ML approaches such as BO can extract nonlinear correlations from experimental and computational datasets, providing a quantitative foundation for multi-objective parameter optimizations. For instance, Hayatzadeh et al. compared the performance of SVM and ANN in predicting parameters of proton exchange membrane electrolyzers [67]. Their study revealed that a genetic algorithm-optimized SVM model not only achieved predictive accuracy comparable to that of ANN, but also improved computational efficiency by approximately 30 %, offering valuable guidance for catalyst optimization. These algorithms are particularly effective in dynamic parameter tuning. As shown in Fig. 5 (a), Jensen et al. employed BO and GPR approaches to dynamically optimize nickel electrodeposition parameters, including current density, temperature, and ligand concentration [68]. The resulting nanostructured hydrogen evolution electrode achieved an overpotential of only −117 mV at 10 mA cm^{-2}, outperforming traditional empirical optimization methods.

Nevertheless, constructing high-throughput optimization platforms typically requires costly specialized equipment for rapid experimentation and characterization. To address this bottleneck, Kodera et al. combined BO with high-throughput experimentation [69] and developed a fully automated robotic system to investigate the effects of catalyst composition on seawater electrolysis selectivity and stability, thereby enabling efficient exploration of complex feature spaces (Fig. 5 (b)). For more complex multi-metallic catalyst systems, Lim et al.

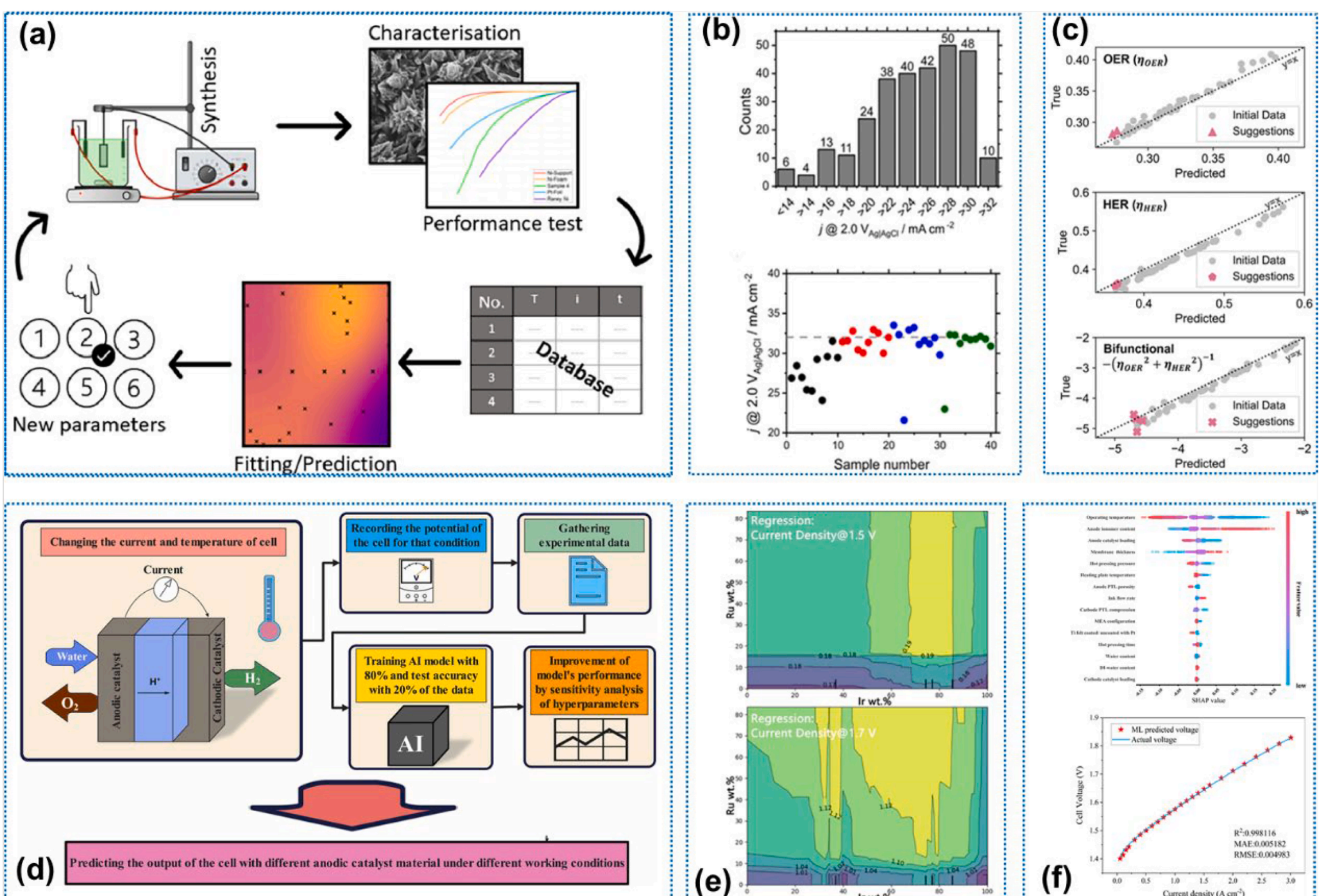


Fig. 5. (a) ML guided development of high-performance nano-structured nickel electrodes. Reproduced with permission from Ref. [68]. (b) Current density histogram for a Co-Mn-Fe-Ni four-element system and an example of Bayesian optimization cycles for current density in the above system. Reproduced with permission from Ref. [69]. (c) True vs. the NN predicted values of the initial dataset alongside the values for the NN-suggested compositions. Reproduced with permission from Ref. [70]. (d) Machine learning-based modeling process. Reproduced with permission from Ref. [71]. (e) 2D PDP interaction plots of Ir wt. % and Ru wt. % in different tasks for modeling. Reproduced with permission from Ref. [72]. (f) SHAP value and the results of machine learning predictions and experimental results based on GA optimization. Reproduced with permission from Ref. [73].

integrated neural network analysis with BO to refine precursor compositions and enhance both catalytic activity and stability (Fig. 5(c)) [70]. System-level performance prediction demands more comprehensive modeling strategies. Bonab et al. optimized spatial distributions using Bayesian optimization and developed a cascaded feedforward neural network model capable of rapidly predicting electrolyzer voltages under various anode materials and operational conditions (Fig. 5(d)) [71]. It is important to note that the complexity of high-dimensional parameter spaces necessitates interpretable ML tools to support decision-making. Ding et al. applied black-box model interpretation techniques to identify key factors contributing to the degradation of membrane electrode assembly performance, providing new insights into material design under multi-objective constraints (Fig. 5(e)) [72]. Zhang et al. employed the shapley additive explanations (SHAP) method to identify dominant variables, reducing the decision space from 15 parameters to 5 while maintaining optimization effectiveness (Fig. 5(f)). This approach also achieved a 67.9 % reduction in computational cost [73].

To overcome the limitations of conventional approaches, data fusion techniques are increasingly employed to integrate theoretical calculations, experimental datasets, and simulation results, enabling the construction of data-driven, multi-scale predictive optimization models. This framework has demonstrated significant value in electrode design and component-level optimization. Bahr et al. optimized the performance of membrane electrode assemblies (MEAs) by simultaneously tuning four key parameters, including cobalt loading and ionomer volume fraction, while balancing current density and cost [74]. Their work highlighted the critical role of multi-parameter synergy in catalyst development and multiscale modeling.

Furthermore, recent studies have explored the enhancement of macroscale electrochemical performance through targeted optimization of mesoscale catalyst structures. As illustrated in Fig. 6(a), Niu et al. developed the GLIDER framework, which integrates generative modeling with surrogate modeling techniques [75]. By using electrochemical performance metrics as feedback, the framework optimizes the nanoscale structure of catalyst layers, providing a novel approach to electrolyzer electrode design. Based on this, Luo et al. combined machine learning with multi-physics simulations, as shown in Fig. 6(b) [76]. Their study employed a random forest regression model to correlate pore structure with mass transport efficiency, revealing the unique value of pore-scale modeling in analyzing dynamic behavior at the triple-phase boundary. These approaches offer a novel strategy for electrode design in electrolyzers, enabling system-level structural

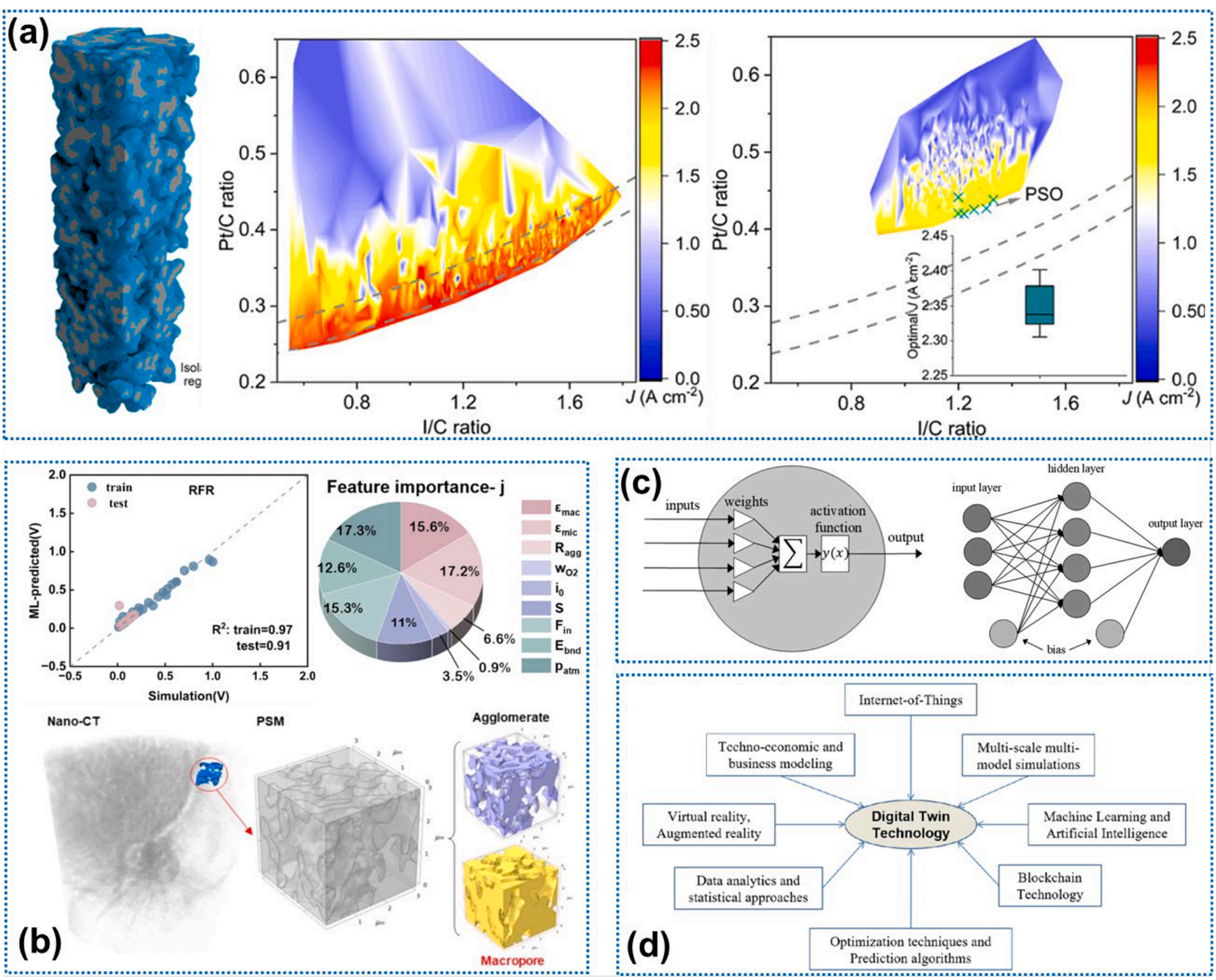


Fig. 6. (a) GLIDER optimization of CL morphology based on Pt/C and I/C ratios and the contour of *J* mapped by the surrogated model under high-fidelity and high-efficiency generation modes, respectively. Reproduced with permission from Ref. [75]. (b) RFR modeling of catalytic layer based on PSM, and the 3D reconstructed mesoscopic pore-scale model of the cathode catalytic layer. Reproduced with permission from Ref. [76]. (c) Structure of a single neuron including the summed-up weights and activation function and architecture of a feed-forward artificial neural network. Reproduced with permission from Ref. [74]. (d) Digital twin technology applications. Reproduced with permission from Ref. [77].

optimization.

DNNs represent a powerful modeling tool capable of optimizing electrolyzer and operational parameters based on training data. As shown in Fig. 6(c), Bahr et al. developed a DNN model that generalizes across various stack designs and control strategies, demonstrating strong adaptability in multi-scale scenarios [74]. In addition, digital twin technology (DTT) has emerged as a transformative approach by constructing virtual replicas of electrolytic systems, as illustrated in Fig. 6 (d) [77]. This method integrates multi-scale physical models with real-time operational data, enabling dynamic monitoring and prediction of system efficiency and durability. DTT-based digital models of catalyst layers can simulate gas diffusion and electron transport pathways under different porosity conditions, offering forward-looking guidance for experimental synthesis and sdevice-level optimization.

In summary, AI approaches demonstrate powerful capabilities in assisting the design and optimization of electrocatalysts at all three scales. At the microscale, AI facilitated the identification of atomic-level active sites and key reaction mechanisms. At the mesoscale, image recognition and structural analysis technologies advance morphology characterization and interface engineering. At the macroscale, ML-based multi-objective optimization methods are improving the efficiency of device-level performance prediction and control. This multiscale design framework not only accelerates the development of electrocatalysts, but also enables the construction of multiscale, high-throughput screening platforms. However, AI-driven electrocatalyst design still faces several challenges. Traditional data-driven approaches struggle to break through the boundaries of known material systems for innovative design. Furthermore, the iterative optimization between computational simulation and experimental verifications is inefficient. To address these challenges, generative AI with its potential for innovative design and automated experimental techniques with high-throughput capabilities will become key directions for overcoming the limitations of the current approaches, as will be discussed in the next section.

4. Generative AI and automated experimental techniques

In recent years, generative modeling strategies have emerged as a promising approach in materials design, significantly accelerating the optimization of various energy-related materials [75]. The core concept of these strategies lies in inverse design, where new candidate structures are generated based on existing knowledge and prior performance feedback, aiming to predict catalyst or electrode architectures with desired properties. Commonly used generative models include variational autoencoders (VAE), diffusion models, autoregressive models, multimodal transformers, and generative adversarial networks (GANs). Lyngby et al. developed a crystal diffusion variational autoencoder (CDVAE) capable of generating two-dimensional materials with high chemical and structural diversity [78]. Many of the generated structures exhibit potential for experimental synthesis, offering new possibilities for autonomous materials discovery. To lower the entry barrier for deploying such models, Manica et al. introduced the generative toolkit for scientific discovery (GT4SD) [79], which simplifies the training, implementation, and customization of generative models, thereby accelerating molecular discovery. In addition, multimodal transformer [80] encoders have shown the ability to capture both local and global features of metal-organic frameworks (MOFs), facilitating their structures design. At the system level, El-Sayad et al. combined the synthetic minority oversampling technique (SMOTE) with GANs to address challenges of data imbalance and nonlinear dependencies in hydrogen production modeling [81]. This integration has improved the reliability and decision-making performance of predictive models in sustainable energy applications. GANs have also been used to generate new alloy structures with specific lattice parameters, contributing to the exploration of novel functional materials [82]. Despite these advances, generative AI tools still face several challenges, including substantial computational requirements, limited availability of high-throughput datasets, and the experimental validation of generated structures and synthesis strategies.

To address these issues, automated experimental platforms, enabled by robotic systems, high-throughput instrumentation, and intelligent control systems, are emerging as powerful tools for validating generative outputs and accelerating the material discovery [69,83,84]. These platforms generate large-scale, high-quality experimental datasets that can significantly enhance model training and validation [85]. In the field of image recognition and processing, in situ electron microscopy enables automated and high-resolution analysis. By integrating deep learning techniques from the field of computer vision, these platforms support real-time interpretation of in-situ experiments, thereby expanding the feasibility of high-throughput screening in materials science. Zeng et al. reported an automated system for kinetic analysis [86], which includes a robotic platform capable of performing electrochemical experiments under varying electrode conditions. This system facilitates the investigation of complex mechanistic behaviors in electrocatalytic processes. Nevertheless, current automated experimental technologies still face limitations, such as high operational costs, complex workflows, and limited generalizability. Further optimization is needed to enhance their scalability and adaptability.

In summary, generative modeling and automated experimentation are often integrated into a closed-loop workflow for iterative catalyst optimization. Based on prior data and design objectives, generative models propose candidate materials, which are then synthesized or characterized via automated high-throughput platforms. The resulting data are fed back into the model to refine subsequent predictions. Through repeated iterations, this closed-loop system enables rapid convergence toward optimal materials that meet targeted performance criteria.

5. Opportunities and prospects

AI presents significant opportunities to accelerate materials discovery and improve the accuracy of performance prediction in the design of electrocatalysts for water electrolysis [87]. This review systematically outlines the role of AI in supporting catalyst development and proposes a multiscale design framework for electrocatalyst design. The core of this framework emphasizes the application of AI algorithms with specialized strengths based on the inherent characteristics of data at different scales, such as atomic-level graph structure interactions, mesoscopic spatial image features, and macroscopic global operational parameters. And it aims to enhance the synergy between datasets, models, and experiments.

Within above paradigm, representative research demonstrates the following typical application paths. At the microscopic level, GNNs facilitate the identification of atomic-scale active sites and key descriptors, such as adsorption energy and d-band center [88–90], enabling faster understanding of catalytic mechanisms and more efficient screening of candidate materials. At the mesoscopic level, CNNs effectively process spatial features including electrode morphology, pore structure, and interface characteristics [91], contributing to improved structural design and characterization. At the macroscopic level, DNN, SVM, and BO methods jointly support the multi-objective optimization and intelligent control of complex parameters (such as composition, structure, and operating conditions). This integrated framework, covering microscopic, mesoscopic, and macroscopic domains, supports a paradigm shift from traditional trial-and-error approaches toward AI-driven electrocatalyst development. In the future, by further integrating the reverse design capabilities of generative AI with high-throughput verification closed loop of the automated experimental platform, the paradigm for electrocatalyst design will be upgraded from assisted optimization to autonomous creation.

Despite recent progress, AI-assisted electrocatalyst design still faces major challenges that need to be addressed. First, data quality and sharing remain key issues. Although high-throughput datasets are essential for machine learning in catalysis, current benchmark databases such as the Open Catalyst Project (OCP) [92] and the Materials Project [93], while valuable, are not yet comprehensive. These datasets support model training and testing, but differences in data collection methods, resolution, and format often lead to inconsistencies. Combining information from experiments and simulations across different scales is limited by poor compatibility. To solve this issue, shared standards and open databases are needed to ensure that results can be compared and repeated across different models and research teams. Building a well-organized and consistent data infrastructure will help AI models learn complex patterns more effectively and improve the speed and accuracy of material discovery.

Second, combining data from different scales is essential to improve the reliability of electrocatalyst design. Using data from only one scale, such as atomic or system level, can miss important relationships and lead to uncertain results. The link between small-scale features and large-scale performance is still not well understood. Connecting experimental data, simulations, and material properties from different levels can help to build a more complete framework. However, most current methods rely on manually selected features and lack automated tools to connect data across scales. Future research should explore the use of graph-based and generative models to build better connections between different types of data, leading to more accurate predictions.

Third, the interpretability of AI models is still limited. While deep learning shows strong performance in prediction tasks, it often lacks transparency. This reduces trust and makes it hard to understand why a model gives a certain result. Improving the clarity of AI models, for example by using simpler structures or adding tools that explain model outputs, will help build confidence in their use for materials science.

Finally, catalyst performance must be evaluated under realistic conditions at the device level. Many AI-generated material structures are promising in theory but their synthetic feasibility has not been fully verified. There is still a gap between computer-based screening and real-world experiments. Electrochemical tests alone cannot capture all important features, and repeated experiments often waste time and resources. Generative AI models should be further developed to support practical testing and work closely with automated, high-throughput experimental systems. These approaches can help optimize electrocatalyst performance in real devices and make large-scale application more achievable.

CRediT authorship contribution statement

Qing Wang: Writing – original draft. **Lizhen Wu:** Writing – original draft. **Qiang Zheng:** Writing – review & editing. **Liang An:** Supervision.

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.

Acknowledgement

This work was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. 15308024), a grant from the Natural Science Foundation of China (Grant No. 42302271), a grant from Ningbo Natural Science Foundation (Grant No. 2024J060), a grant from the Research Institute for Advanced Manufacturing at the Hong Kong Polytechnic University (CDJQ), and a grant from Research Centre for Carbon-Strategic Catalysis at the Hong Kong Polytechnic University (CE2X).

Data availability

Data will be made available on request.

References

[1] Wang J, Wen J, Wang J, Yang B, Jiang L. Water electrolyzer operation scheduling for green hydrogen production: a review. Renew. Sustain. Energy Rev. 2024;203:114779. https://doi.org/10.1016/j.rser.2024.114779.
[2] Ahn SH, Yoo SJ, Kim H-J, Henkensmeier D, Nam SW, Kim S-K, et al. Anion exchange membrane water electrolyzer with an ultra-low loading of Pt-decorated Ni electrocatalyst. Appl. Catal. B. 2016;180:674–9. https://doi.org/10.1016/j.apcatb.2015.07.020.
[3] Chen Y, Liu C, Xu J, Xia C, Wang P, Xia BY, et al. Key components and design strategy for a Proton exchange membrane water electrolyzer. Small Struct. 2023;4:2200130. https://doi.org/10.1002/sstr.202200130.
[4] Hong S, Kim JH, Shin D, Bak G, Jang D, Kim WB, et al. Synergistic effect of Pt-loaded Co N C electrocatalysts for hydrogen evolution reaction in alkaline conditions. Appl. Surf. Sci. 2023;610:155523. https://doi.org/10.1016/j.apsusc.2022.155523.
[5] Chen C, Jin H, Wang P, Sun X, Jaroniec M, Zheng Y, et al. Local reaction environment in electrocatalysis. Chem. Soc. Rev. 2024;53:2022–55. https://doi.org/10.1039/D3CS00669G.
[6] Kreider ME, Yu H, Osmieri L, Parimuha MR, Reeves KS, Marin DH, et al. Understanding the effects of anode catalyst conductivity and loading on catalyst layer utilization and performance for Anion exchange membrane water electrolysis. ACS Catal. 2024;14:10806–19. https://doi.org/10.1021/acscatal.4c02932.
[7] Dmitrieva AP, Fomkina AS, Tracey CT, Romanenko EA, Ayati A, Krivoshapkin PV, et al. AI and ML for selecting viable electrocatalysts: progress and perspectives. J. Mater. Chem. A. 2024;12:31074–102. https://doi.org/10.1039/D4TA04991H.
[8] Iqbal S, Aftab K, Tul Jannat F, Ali Baig M, Kalsoom U. A bibliographic analysis of optimization of hydrogen production via electrochemical method using machine learning. Fuel. 2024;372:132126. https://doi.org/10.1016/j.fuel.2024.132126.
[9] Wang N, Guo Y, Liu L, Shao S. Numerical assessment and optimization of photovoltaic-based hydrogen-oxygen Co-production energy system: a machine learning and multi-objective strategy. Renew. Energy. 2024;227:120483. https://doi.org/10.1016/j.renene.2024.120483.
[10] Xu D, Zhang Q, Huo X, Wang Y, Yang M. Advances in data-assisted high-throughput computations for material design. Mater. Genome Eng. Advanc. 2023;1:e11. https://doi.org/10.1002/mgea.11.
[11] Wang N, Xiang H, Meng L, Tang C, Dong Z, Yang Y, et al. Accelerating the green hydrogen revolution: the synergy of PEMWE and AI. Renew. Energy Syst. Equipment. 2025. https://doi.org/10.1016/j.rese.2024.10.002. S2950208×2400005X.
[12] Ding R, Chen J, Chen Y, Liu J, Bando Y, Wang X. Unlocking the potential: machine learning applications in electrocatalyst design for electrochemical hydrogen energy transformation. Chem. Soc. Rev. 2024;53:11390–461. https://doi.org/10.1039/D4CS00844H.
[13] Wang L, Zhou X, Luo Z, Liu S, Yue S, Chen Y, et al. Review of external field effects on electrocatalysis: machine learning guided design. Adv. Funct. Mater. 2024;34:2408870. https://doi.org/10.1002/adfm.202408870.
[14] Liu X, Fan K, Huang X, Ge J, Liu Y, Kang H. Recent advances in artificial intelligence boosting materials design for electrochemical energy storage. Chemical Eng. J. 2024;490:151625. https://doi.org/10.1016/j.cej.2024.151625.
[15] Adamu H, Abba SI, Anyin PB, Sani Y, Yamani ZH, Qamar M. Tuning OER electrocatalysts toward LOM pathway through the lens of multi-descriptor feature selection by artificial intelligence-based approach. ACS Mater. Lett. 2023;5:299–320. https://doi.org/10.1021/acsmaterialslett.2c00734.
[16] Shangguan Z, Li H, Yang B, Zhao Z, Wang T, Jin L, et al. Optimization of alkaline electrolyzer operation in renewable energy power systems: a universal modeling approach for enhanced hydrogen production efficiency and cost-effectiveness. Int. J. Hydrogen Energy. 2024;49:943–54. https://doi.org/10.1016/j.ijhydene.2023.10.057.
[17] Wang T, Wang J, Zhang C, Wang P, Ren Z, Guo H, et al. Direct operational data-driven workflow for dynamic voltage prediction of commercial alkaline water electrolyzers based on artificial neural network (ANN). Fuel 2024;376:132624. https://doi.org/10.1016/j.fuel.2024.132624.
[18] Tamtaji M, Gao H, Hossain MD, Galligan PR, Wong H, Liu Z, et al. Machine learning for design principles for single atom catalysts towards electrochemical reactions. J. Mater. Chem. A. 2022;10:15309–31. https://doi.org/10.1039/D2TA02039D.
[19] Jerng SE, Park YJ, Li J. Machine learning for CO 2 capture and conversion: a review. Energy AI. 2024;16:100361. https://doi.org/10.1016/j.egyai.2024.100361.
[20] Jiao Z, Liu Y, Wang Z. Application of graph neural network in computational heterogeneous catalysis. J. Chem. Phys. 2024;161:171001. https://doi.org/10.1063/5.0227821.
[21] Li X, Huang W, Lian Y, Tao S. Graph neural network model accelerates biomass adsorption energy prediction on iron-group hydrotalcite electrocatalysts. J. Phys. Chem. Lett. 2024;15:10725–33. https://doi.org/10.1021/acs.jpclett.4c02466.
[22] Wu S, Wang Z, Zhang H, Cai J, Li J. Deep learning accelerates the discovery of two-dimensional catalysts for hydrogen evolution reaction. Energy & Environ Mater. 2023;6:e12259. https://doi.org/10.1002/eem2.12259.
[23] Kim M, Yeo BC, Park Y, Lee HM, Han SS, Kim D. Artificial intelligence to accelerate the discovery of N_2 electroreduction catalysts. Chem. Mater. 2020;32:709–20. https://doi.org/10.1021/acs.chemmater.9b03686.
[24] Ding R, Wang R, Ding Y, Yin W, Liu Y, Li J, et al. Designing AI-aided analysis and prediction models for nonprecious metal electrocatalyst-based proton-exchange

membrane fuel cells. Angewandte. Chemie. 2020;132:19337–45. https://doi.org/10.1002/ange.202006928.
[25] Niknezhad A, Nabian N, Lashkenari MS, Ghorbani M. PANI-ZnO/Pt–Ru electrocatalyst for methanol oxidation: synthesis, characterization, electrocatalytic performance and artificial neural network modeling. Chemic. Eng. Res. Des. 2022; 180:123–33. https://doi.org/10.1016/j.cherd.2022.02.010.
[26] Wu L, Pan Z, Yuan S, Huo X, Zheng Q, Yan X, et al. Optimization of dual-layer flow field in a water electrolyzer using a data-driven surrogate model. Energy AI. 2024; 18:100411. https://doi.org/10.1016/j.egyai.2024.100411.
[27] Araujo RB, Edvinsson T. Supervised AI and deep neural networks to evaluate high-entropy alloys as reduction catalysts in aqueous environments. ACS Catal. 2024;14: 3742–55. https://doi.org/10.1021/acscatal.3c05017.
[28] Hossain SS, Ali SS, Rushd S, Ayodele BV, Cheng CK. Interaction effect of process parameters and Pd-electrocatalyst in formic acid electro-oxidation for fuel cell applications: implementing supervised machine learning algorithms. Intl. J. Energy Res. 2022;46:21583–97. https://doi.org/10.1002/er.7602.
[29] Li J, Wu N, Zhang J, Wu H-H, Pan K, Wang Y, et al. Machine learning-assisted low-dimensional electrocatalysts design for hydrogen evolution reaction. Nano-Micro Lett. 2023;15:227. https://doi.org/10.1007/s40820-023-01192-5.
[30] Yang J, Wang Z, Zhang X, Zhang X, Niu B, Wang C, et al. Active learning driven discovery of novel alloyed catalysts for selective ammonia oxidation. Chemic. Eng. J. 2024;493:152300. https://doi.org/10.1016/j.cej.2024.152300.
[31] Wei S, Baek S, Yue H, Liu M, Yun SJ, Park S, et al. Machine-learning assisted exploration: toward the next-generation catalyst for hydrogen evolution reaction. J. Electrochem. Soc. 2021;168:126523. https://doi.org/10.1149/1945-7111/ac41f1.
[32] Chen H, Liang X, Liu Y, Ai X, Asefa T, Zou X. Active site engineering in porous electrocatalysts. Advanc. Mater. 2020;32:2002435. https://doi.org/10.1002/adma.202002435.
[33] Jiao S, Fu X, Wang S, Zhao Y. Perfecting electrocatalysts *via* imperfections: towards the large-scale deployment of water electrolysis technology. Energy Environ. Sci. 2021;14:1722–70. https://doi.org/10.1039/D0EE03635H.
[34] Abiola A, Manzano FS, Andújar JM. A novel deep reinforcement learning (DRL) algorithm to apply artificial intelligence-based maintenance in electrolysers. Algorithms 2023;16:541. https://doi.org/10.3390/a16120541.
[35] Sun C, Wang L, Zhao W, Xie L, Wang J, Li J, et al. Atomic-level design of active site on two-dimensional MoS_2 toward efficient hydrogen evolution: experiment, theory, and artificial intelligence modelling. Adv. Funct. Mater. 2022;32:2206163. https://doi.org/10.1002/adfm.202206163.
[36] Bensmann B, Rex A, Hanke-Rauschenbach R. An engineering perspective on the future role of modelling in proton exchange membrane water electrolysis development. Curr. Opin. Chem. Eng. 2022;36:100829. https://doi.org/10.1016/j.coche.2022.100829.
[37] Li Q, Ouyang Y, Lu S, Bai X, Zhang Y, Shi L, et al. Perspective on theoretical methods and modeling relating to electro-catalysis processes. Chem. Commun. 2020;56:9937–49. https://doi.org/10.1039/D0CC02998J.
[38] Colliard-Granero A, Gompou KA, Rodenbücher C, Malek K, Eikerling MH, Eslamibidgoli MJ. Deep learning-enhanced characterization of bubble dynamics in proton exchange membrane water electrolyzers. Phys. Chem. Chem. Phys. 2024; 26:14529–37. https://doi.org/10.1039/D3CP05869G.
[39] Liu J, Luo W, Wang L, Zhang J, Fu X, Luo J. Toward excellence of electrocatalyst design by emerging descriptor-oriented machine learning. Adv. Funct. Mater. 2022;32:2110748. https://doi.org/10.1002/adfm.202110748.
[40] Jäger MOJ, Morooka EV, Federici Canova F, Himanen L, Foster AS. Machine learning hydrogen adsorption on nanoclusters through structural descriptors. Npj. Comput. Mater. 2018;4:37. https://doi.org/10.1038/s41524-018-0096-5.
[41] Fung V, Tao FF, Jiang D. General structure–Reactivity relationship for oxygen on transition-metal oxides. J. Phys. Chem. Lett. 2017;8:2206–11. https://doi.org/10.1021/acs.jpclett.7b00861.
[42] Tran K, Ulissi ZW. Active learning across intermetallics to guide discovery of electrocatalysts for CO2 reduction and H2 evolution. Nat. Catal. 2018;1:696–703. https://doi.org/10.1038/s41929-018-0142-1.
[43] Vaz N, Choi J, Cha Y, Kong J, Park Y, Ju H. Multi-objective optimization of the cathode catalyst layer micro-composition of polymer electrolyte membrane fuel cells using a multi-scale, two-phase fuel cell model and data-driven surrogates. J. Energy Chem. 2023;81:28–41. https://doi.org/10.1016/j.jechem.2023.02.027.
[44] Tawalbeh M, Shomope I, Al-Othman A, Alshraideh H. Prediction of hydrogen production in proton exchange membrane water electrolysis via neural networks. Int. J. Thermofluid. 2024;24:100849. https://doi.org/10.1016/j.ijft.2024.100849.
[45] Bilgiç G, Öztürk B, Atasever S, Şahin M, Kaplan H. Prediction of hydrogen production by magnetic field effect water electrolysis using artificial neural network predictive models. Int. J. Hydrogen Energy. 2023;48:20164–75. https://doi.org/10.1016/j.ijhydene.2023.02.082.
[46] Wang M, Zhu H. Machine learning for transition-metal-based hydrogen generation electrocatalysts. ACS Catal. 2021;11:3930–7. https://doi.org/10.1021/acscatal.1c00178.
[47] Ahmed R, Shehab SA, Elzeki OM, Darwish A, Hassanein AE. An explainable AI for green hydrogen production: a deep learning regression model. Int. J. Hydrogen. Energy. 2024;83:1226–42. https://doi.org/10.1016/j.ijhydene.2024.08.064.
[48] Chen X, Rex A, Woelke J, Eckert C, Bensmann B, Hanke-Rauschenbach R, et al. Machine learning in proton exchange membrane water electrolysis — A knowledge-integrated framework. Appl. Energy. 2024;371:123550. https://doi.org/10.1016/j.apenergy.2024.123550.
[49] Zhang X, Tian Y, Chen L, Hu X, Zhou Z. Machine learning: a new paradigm in computational electrocatalysis. J. Phys. Chem. Lett. 2022;13:7920–30. https://doi.org/10.1021/acs.jpclett.2c01710.
[50] Palkovits R, Palkovits S. Using artificial intelligence to forecast water oxidation catalysts. ACS Catal. 2019;9:8383–7. https://doi.org/10.1021/acscatal.9b01985.
[51] Wang S-H, Pillai HS, Wang S, Achenie LEK, Xin H. Infusing theory into deep learning for interpretable reactivity prediction. Nat. Commun. 2021;12:5288. https://doi.org/10.1038/s41467-021-25639-8.
[52] Li H, Xu S, Wang M, Chen Z, Ji F, Cheng K, et al. Computational design of (100) alloy surfaces for the hydrogen evolution reaction. J. Mater. Chem. A. 2020;8: 17987–97. https://doi.org/10.1039/D0TA04615A.
[53] Zhou L, Tian P, Zhang B, Xuan F-Z. Data-driven rational design of single-atom materials for hydrogen evolution and sensing. Nano Res. 2024;17:3352–8. https://doi.org/10.1007/s12274-023-6137-5.
[54] Salehmin MNI, Tiong SK, Mohamed H, Umar DA, Yu KL, Ong HC, et al. Navigating challenges and opportunities of machine learning in hydrogen catalysis and production processes: beyond algorithm development. J. Energy Chem. 2024;99: 223–52. https://doi.org/10.1016/j.jechem.2024.07.045.
[55] Ding R, Chen Y, Rui Z, Hua K, Wu Y, Li X, et al. Machine learning utilized for the development of proton exchange membrane electrolyzers. J. Power Sourc. 2023; 556:232389. https://doi.org/10.1016/j.jpowsour.2022.232389.
[56] Liu X, Zhou S, Yan Z, Zhong Z, Shikazono N, Hara S. Correlation between microstructures and macroscopic properties of nickel/yttria-stabilized zirconia (Ni-YSZ) anodes: meso-scale modeling and deep learning with convolutional neural networks. Energy AI 2022;7:100122. https://doi.org/10.1016/j.egyai.2021.100122.
[57] Karaca I, Aldemir Dikici B. Quantitative evaluation of the pore and window sizes of tissue engineering scaffolds on scanning electron microscope images using Deep learning. ACS Omega. 2024;9:24695–706. https://doi.org/10.1021/acsomega.4c01234.
[58] Topal E, Liao Z, Löffler M, Gluch J, Zhang J, Feng X, et al. Multi-scale X-ray tomography and machine learning algorithms to study MoNi4 electrocatalysts anchored on MoO2 cuboids aligned on Ni foam. BMC Mat. 2020;2:5. https://doi.org/10.1186/s42833-020-00011-0.
[59] Tao H, Wang S, Liu H, Lian C. Deep Neural network enhanced mesoscopic thermodynamic model for unlocking the electrode/electrolyte interface. Angew. Chem. Int. Ed. 2025;64:e202418447. https://doi.org/10.1002/anie.202418447.
[60] Naveena K, Krishnamoorthy M, Karuppiah N, Gouda PK, Hariharan S, Saravanan K, et al. Elevating sustainability with a multi-renewable hydrogen generation system empowered by machine learning and multi-objective optimization. Measurement. 2024;33:101192. https://doi.org/10.1016/j.measen.2024.101192.
[61] Günay ME, Tapan NA. Evaluation of polymer electrolyte membrane electrolysis by explainable machine learning, optimum classification model, and active learning. J. Appl. Electrochem. 2023;53:415–33. https://doi.org/10.1007/s10800-022-01786-8.
[62] Dash BM, Bouamama BO, Pekpe KM, Boukerdja M. Prior knowledge-infused self-supervised learning and explainable AI for fault detection and isolation in PEM electrolyzers. Neurocomputing 2024;594:127871. https://doi.org/10.1016/j.neucom.2024.127871.
[63] Ozdemir SN, Pektezel O. Performance prediction of experimental PEM electrolyzer using machine learning algorithms. Fuel 2024;378:132853. https://doi.org/10.1016/j.fuel.2024.132853.
[64] Chee SW, Lunkenbein T, Schlögl R, Cuenya BR. In situ and operando electron microscopy in heterogeneous catalysis—Insights into multi-scale chemical dynamics. J. Phys. 2021;33:153001. https://doi.org/10.1088/1361-648X/abddfd.
[65] Aydın Ö, Atilla P, Akben SB, Farsak M. Preparation of green hydrogen catalyst with multi-objective optimization. Molecul. Catal. 2022;529:112544. https://doi.org/10.1016/j.mcat.2022.112544.
[66] Zhai F-H, Zhan Q-Q, Yang Y-F, Ye N-Y, Wan R-Y, Wang J, et al. A deep learning protocol for analyzing and predicting ionic conductivity of anion exchange membranes. J. Memb. Sci. 2022;642:119983. https://doi.org/10.1016/j.memsci.2021.119983.
[67] Hayatzadeh A, Fattahi M, Rezaveisi A. Machine learning algorithms for operating parameters predictions in proton exchange membrane water electrolyzers: anode side catalyst. Int. J. Hydrogen Energy. 2024;56:302–14. https://doi.org/10.1016/j.ijhydene.2023.12.149.
[68] Jensen VH, Moretti ER, Busk J, Christiansen EH, Skov SM, Jacobsen E, et al. Machine learning guided development of high-performance nano-structured nickel electrodes for alkaline water electrolysis. Appl. Mater. Today. 2023;35:102005. https://doi.org/10.1016/j.apmt.2023.102005.
[69] Kodera M, Sayama K. An automatic robot system for machine learning–assisted high-throughput screening of composite electrocatalysts. Digit. Discovery. 2023;2: 1683–7. https://doi.org/10.1039/D3DD00116D.
[70] Lim CYJ, I Made R, Khoo ZHJ, Ng CK, Bai Y, Wang J, et al. Machine learning-assisted optimization of multi-metal hydroxide electrocatalysts for overall water splitting. Mater. Horiz. 2023;10:5022–31. https://doi.org/10.1039/D3MH00788J.
[71] Alipour Bonab S, Waite T, Song W, Flynn D, Yazdani-Asrami M. Machine learning-powered performance monitoring of proton exchange membrane water electrolyzers for enhancing green hydrogen production as a sustainable fuel for aviation industry. Energy Report. 2024;12:2270–82. https://doi.org/10.1016/j.egyr.2024.08.028.
[72] Ding R, Chen Y, Rui Z, Hua K, Wu Y, Li X, et al. Guiding the optimization of membrane electrode assembly in a Proton exchange membrane water electrolyzer by machine learning modeling and black-box interpretation. ACS Sustain. Chem. Eng. 2022;10:4561–78. https://doi.org/10.1021/acssuschemeng.1c08522.
[73] Zhang Y, Tan A, Yuan Z, Zhao K, Shi X, Liu P, et al. Data-driven optimization of high-dimensional variables in proton exchange membrane water electrolysis membrane electrode assembly assisted by machine learning. Ind. Eng. Chem. Res. 2024;63:1409–21. https://doi.org/10.1021/acs.iecr.3c03546.

[74] Bahr M, Gusak A, Stypka S, Oberschachtsiek B. Artificial Neural networks for aging simulation of electrolysis stacks. Chemie. Ingenieur. Technik. 2020;92:1610–7. https://doi.org/10.1002/cite.202000089.
[75] Niu Z, Zhao W, Deng H, Tian L, Pinfield VJ, Ming P, et al. Generative artificial intelligence for designing multi-scale hydrogen fuel cell catalyst layer nanostructures. ACS Nano. 2024;18:20504–17. https://doi.org/10.1021/acsnano.4c04943.
[76] Luo L, Liu X, Zhang J, Yao J, Liu B, Zhang J, et al. Insights into the pore structure effect on the mass transfer of fuel cell catalyst layer via combining machine learning and multiphysics simulation. Chem. Eng. Sci. 2025;302:120830. https://doi.org/10.1016/j.ces.2024.120830.
[77] editors. In: Singh P, Agarwal AK, Thakur A, Sinha RK, editors. Challenges and opportunities in green hydrogen production. Singapore: Springer Nature Singapore; 2024. https://doi.org/10.1007/978-981-97-1339-4.
[78] Lyngby P, Thygesen KS. Data-driven discovery of 2D materials by deep generative models. Npj. Comput. Mater. 2022;8:232. https://doi.org/10.1038/s41524-022-00923-3.
[79] Manica M, Born J, Cadow J, Christofidellis D, Dave A, Clarke D, et al. Accelerating material design with the generative toolkit for scientific discovery. Npj. Comput. Mater. 2023;9:69. https://doi.org/10.1038/s41524-023-01028-1.
[80] Kang Y, Park H, Smit B, Kim J. A multi-modal pre-training transformer for universal transfer learning in metal–organic frameworks. Nat. Mach. Intell. 2023;5:309–18. https://doi.org/10.1038/s42256-023-00628-2.
[81] El-Sayad NE, Mohammed KK, Elghamrawy S, Darwish A, Hassanien AE. Green hydrogen production capacity model from water electrolysis and fossil fuel CO2 using generative adversarial network. Fuel 2025;393:134962. https://doi.org/10.1016/j.fuel.2025.134962.
[82] Dahari A, Kench S, Squires I, Cooper SJ. Fusion of complementary 2D and 3D mesostructural datasets using generative adversarial networks. Adv. Energy Mater. 2023;13:2202407. https://doi.org/10.1002/aenm.202202407.
[83] Kamšek AR, Ruiz-Zepeda F, Pavlišič A, Hrnjić A, Hodnik N. Bringing into play automated electron microscopy data processing for understanding nanoparticulate electrocatalysts' structure–property relationships. Curr. Opinion Electrochem. 2022;35:101052. https://doi.org/10.1016/j.coelec.2022.101052.
[84] Jenewein KJ, Akkoc GD, Kormányos A, Cherevko S. Automated high-throughput activity and stability screening of electrocatalysts. Chem. Catal. 2022;2:2778–94. https://doi.org/10.1016/j.checat.2022.09.019.
[85] Horwath JP, Zakharov DN, Mégret R, Stach EA. Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images. Npj. Comput. Mater. 2020;6:108. https://doi.org/10.1038/s41524-020-00363-x.
[86] Zeng JS, Padia V, Chen GY, Maalouf JH, Limaye AM, Liu AH, et al. Nonidealities in CO_2 electroreduction mechanisms revealed by automation-assisted kinetic analysis. ACS Cent. Sci. 2024;10:1348–56. https://doi.org/10.1021/acscentsci.3c01295.
[87] Zhang Q, Yuan Y, Zhang J, Fang P, Pan J, Zhang H, et al. Machine learning-aided design of highly conductive anion exchange membranes for fuel cells and water electrolyzers. Advanc. Mater. 2024:2404981. https://doi.org/10.1002/adma.202404981.
[88] Xu B, Ma W, Wu W, Wang Y, Yang Y, Li J, et al. Degradation prediction of PEM water electrolyzer under constant and start-stop loads based on CNN-LSTM. Energy AI. 2024;18:100420. https://doi.org/10.1016/j.egyai.2024.100420.
[89] Babay M-A, Adar M, Chebak A, Mabrouki M. Dynamics of gas generation in porous electrode alkaline electrolysis cells: an investigation and optimization using machine learning. Energies 2023;16:5365. https://doi.org/10.3390/en16145365.
[90] Madhavan PV, Moradizadeh L, Shahgaldi S, Li X. Data-driven modelling of corrosion behaviour in coated porous transport layers for PEM water electrolyzers. Artifi. Intell. Chem. 2025;3:100086. https://doi.org/10.1016/j.aichem.2025.100086.
[91] Miao L, Jia W, Cao X, Jiao L. Computational chemistry for water-splitting electrocatalysis. Chem. Soc. Rev. 2024;53:2771–807. https://doi.org/10.1039/D2CS01068B.
[92] 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;13:3066–84. https://doi.org/10.1021/acscatal.2c05426.
[93] Liu H, Cheng J, Dong H, Feng J, Pang B, Tian Z, et al. Screening stable and metastable ABO3 perovskites using machine learning and the materials project. Computat. Mater. Sci. 2020;177:109614. https://doi.org/10.1016/j.commatsci.2020.109614.