Chemical Engineering Journal 490 (2024) 151625

Contents lists available at ScienceDirect

Chemical Engineering Journal

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

Review

Recent advances in artificial intelligence boosting materials design for electrochemical energy storage

Xinxin Liu [c,d,1], Kexin Fan [e,1], Xinmeng Huang [f,1], Jiankai Ge [a,*], Yujie Liu [b,*], Haisu Kang [a,*]

[a] Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[b] Department of Materials Science and Engineering, University of Michigan, Ann Arbor, MI 48109, USA
[c] Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA 19104, USA
[d] Department of Materials Science and Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA
[e] Andrew and Peggy Cherng Department of Medical Engineering, Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125, USA
[f] Graduate Group in Applied Mathematics and Computational Science, University of Pennsylvania, Philadelphia, PA 19104, USA

ARTICLE INFO

Keywords:
Artificial intelligence
Batteries
Fuel cells
Supercapacitors
Material design

ABSTRACT

In the rapidly evolving landscape of electrochemical energy storage (EES), the advent of artificial intelligence (AI) has emerged as a keystone for innovation in material design, propelling forward the design and discovery of batteries, fuel cells, supercapacitors, and many other functional materials. This review paper elucidates the burgeoning role of AI in materials from foundational machine learning (ML) techniques to its current pivotal role in advancing the frontiers of materials science for energy storage, including enhancing the performance, durability, and safety of battery technologies, fuel cell efficiency and longevity, and the materials fine-tuning in supercapacitors to achieve superior energy storage capabilities. Collectively, we present a comprehensive overview of the recent AI advancements that have significantly accelerated the development of next-generation materials for EES, offering insights into future research trajectories and the potential for AI to unlock new horizons in materials science.

1. Introduction

The growth of energy consumption greatly increases the burden on the environment [1]. To address this issue, it is critical for human society to pursue clean energy resources, such as wind, water, solar and hydrogen [2]. Developing electrochemical energy storage devices has long been considered as a promising topic in the clean energy field, as it is environment-friendly, high-efficient, and portable [3–6]. Researchers have been developing different electrochemical energy storage devices for over a hundred years, including batteries, fuel cells, and supercapacitors, etc. [7–10]. Among these electrochemical energy storage devices, materials play a vital role in promoting the ability, capacity, and duality [11–13]. Therefore, a systematic design of materials for electrochemical devices is needed, which usually contains designs of electrodes, electrolytes, catalysts, etc. [14–16].

However, the current landscape of materials design, particularly in the context of electrochemical energy storage, faces notable challenges. Traditionally, the development of new materials relies heavily on trial-and-error methods, which are time-consuming, labor-intensive, and often lack precision. Computational methods, including density functional theory (DFT), molecular dynamics (MD), and Monte Carlo (MC), etc., provide a sophisticated way of designing materials.[236,239,240] Though computational designs of materials have reached notable achievements in past decades, they are limited by the accuracy of methods, timescale/lengthscale limitation, and heavy computational cost. These often lead to oversimplified models that fail to capture the behaviors of materials in practical scenario. As a result, the pace of discovering and optimizing materials for applications like batteries, fuel cells, and supercapacitors has been slower than desired, hindering the rapid advancement needed in the clean energy sector.

In recent years, the advent of AI has opened a new era in materials design, offering transformative solutions to these longstanding challenges. AI, particularly machine learning and deep learning techniques, can analyze vast datasets, uncovering patterns and relationships that

* Corresponding authors.
E-mail addresses: jiankai2@illinois.edu (J. Ge), yujl@umich.edu (Y. Liu), haisuk2@illinois.edu (H. Kang).
[1] These authors contributed equally.

https://doi.org/10.1016/j.cej.2024.151625
Received 16 February 2024; Received in revised form 5 April 2024; Accepted 23 April 2024
Available online 24 April 2024

elude traditional methods. This capability enables AI to predict material properties and behaviors with remarkable accuracy, significantly reducing the reliance on costly and time-consuming experiments and computational methods. Additionally, AI can optimize the synthesis and processing parameters, leading to more efficient material development cycles. The integration of AI in materials design not only accelerates the discovery of novel materials but also enhances our understanding of material science, paving the way for more innovative and effective electrochemical energy storage solutions. The synergy of AI with existing computational methods also creates new opportunities for more comprehensive, rapid, and accurate modeling, such as deep learning potential in MD simulations, generative diffusion models for transition states, ChatGPT for materials design.[86,238,241,242] The existence of AI greatly facilitates the formation of innovative and predictive approaches to materials science.

In this review, we summarized theoretical basis and recent progress of materials design for electrochemical energy storage with the assistance of AI. Starting from introducing basic concepts of AI toolkit, we discussed classical methods like machine learning, deep learning, and reinforce learning, and most recent AI techniques like generative diffusion models and large language models that assist materials design. Then, focusing on applications, we first introduced how AI could benefit the development of battery materials, which divided into electrode materials, electrolyte materials, multiscale modeling, and experimental planning and synthesis. Second, we discussed the challenge and applications of machine learning in fuel cell systems. Lastly, we discussed how AI could help design supercapacitors, such as predicting and optimizing the property, finding descriptors for carbon-based supercapacitors, and helping characterizations. To wrap up, we pointed out perspectives, and proposed the future direction of AI accelerating materials design.

2. AI toolkit for materials design

We offer a summary of classical ML methods in Section 2.1 and then introduce more recent advancements in artificial intelligence and their potential applications in energy storage materials.

2.1. Classic Machine learning

ML harnesses algorithms and models to acquire knowledge, make inferences based on historical data and existing domain knowledge, and represents a central domain within the broader field of AI. ML can be further stratified into distinct categories, chiefly supervised learning, and unsupervised learning, according to the availability of labeled data. Each of these domains harbors a collection of seminal algorithms.

The premise of supervised learning revolves around a feature variable denoted as x, coupled with an associated label variable y. This label variable y can be either continuous or discrete. Supervised learning algorithms aim to predict the label variable provided with a feature variable, also called covariates. Depending on the discreteness of label variables, supervised learning can be further subdivided into two distinct categories, namely regression tasks and classification tasks [25–27]. Within the regression task, well-established approaches encompass but are not limited to linear regression with regularization (e.g., the ℓ_2 and ℓ_1-norm regularization) [24,28], Gaussian process regression [29], time series regression [30]. On the other hand, common classification algorithms include logistic regression [31], random forest regression [32], support vector machines [33], the naive Bayes approach [34].

Conversely, unsupervised learning is uniquely tailored for exploring and revealing concealed patterns inherent within the features, often with little or no emphasis on the label variables. The main unsupervised learning tasks are clustering and dimensionality reduction. Prominent clustering techniques encompass the K-means clustering [35], Gaussian mixture models [36], and self-organizing maps [37]. Concurrently, seminal dimensionality reduction techniques include the principle components analysis [38] and independent component analysis [39].

Over the past two decades, ML algorithms have found extensive application in data-driven materials science [40,41]. Noteworthy examples include the application of linear regression for the predictive screening of candidate materials and predicting specific properties [17,42]. Gaussian process regression has proven invaluable for forecasting the performance of various energy materials characterized by intricate structural compositions, interfaces, and formulations in the context of solar cells and lithium-ion batteries [43]. Logistic regression, in turn, plays a pivotal role in the pursuit of energy materials through high-throughput computational screening techniques [18]. More specifically, Sodeyama et al. [17] propose a comprehensive machine-learning strategy with three linear regression techniques, showing that linear regression facilitates the most precise predictions of the properties of electrolyte liquids. Building upon logistic regression, Sendek et al. [18] introduce a novel method for screening to pinpoint high-potential materials suitable for solid-state electrolytes. Shandiz and Gauvin [44] utilize eight different clustering algorithms to delve into how the crystal structure impacts the efficiency of battery electrodes, revealing that the highest accuracy in predictions can be achieved with random forests. Pilania et al. [45] develop SVM classification models to predict new perovskite halides, leading to the discovery of several new compositions of ABX_3 with perovskite structures. More detailed ML applications in material science are introduced in the sections below. A comprehensive comparison of classical ML methods is detailed in Table 1.

2.2. Deep learning

Neural networks, often referred to as deep learning models, represent a revolutionary paradigm and building block in modern machine learning. The simplest neural network model is the perceptron [46]. The perceptron comprises three layers: the input layer, the hidden layer, and the output layer. The input layer aligns with the feature variable x and thus shares the same dimension as features. After being fed the feature variable, the hidden layer processes the input through a linear weighted combination and applies an entry-wise non-linear transformation. The non-linear transformation is also referred to as activations in deep learning and can be flexibly adjusted in various scenarios [47]. The output layer corresponds to the prediction, which serves as the primary target of supervision. While the perceptron offers ease of use and interpretability, it may falter in modeling intricate data patterns. The multilayer perceptron (MLP), also recognized as the fully connected deep neural network, enhances the vanilla perceptron's expressiveness by incorporating multiple hidden layers. Consequently, MLPs demonstrate remarkable aptitude in capturing intricate relationships [48,49].

Despite the deep and full connection, MLPs easily suffer from overfitting due to over-parametrization. The prowess of deep neural networks (DNNs) lies in their diverse internal connection architectures. Specialized internal connection layers enable a more effective acquisition of hierarchical data representations that abstract increasingly complex features as information traverses. One representative DNN model is the convolutional neural network (CNN), renowned for its proficiency in visual tasks [50]. CNNs deploy convolutional layers, each associated with a small set of shared weights, to detect features such as edges, textures, and shapes while reducing spatial dimensions. Subsequently, shallow, and fully connected layers are used to facilitate high-level abstraction and classification. Notable CNN examples include AlexNet [51], VGG [52], ResNet [53]. Another seminal architectural marvel is the recurrent neural net- work (RNN) [54]. specifically designed to handle sequential data types, including time series, speech, and text. RNNs employ recurrent connections to capture temporal dependencies, making them indispensable for speech recognition and language modeling applications. Representative RNN models include GRU [55], LSTM [56], and Seq2Seq [57]. The architectural representations of the Perceptron, MLP, CNN, and RNN are depicted in Fig. 1.

Table 1
Comparison of common classic machine learning algorithms.

Algorithms	Task	Advantages	Disadvantages	Applications
Linear regression with regularization	Regression	Simple implementation, model interpretability	Vulnerability to outliers, poor expressiveness	Feature selection, material screening[17]
Logistic regression	Classification	Simple implementation, uncertainty quantification	Incapability to capture complex relationships	Material screening, property prediction[18]
Support vector machine	Classification	Effective in high-dimensions, robust to outliers	Sensitive to hyperparameters, high computation cost	Material screening, property prediction[19]
Naive Bayes	Classification	Simple implementation, good interpretability	Vulnerable to correlation, limited expressiveness:	Optimization[20]
Random forest	Regression, classification	Accurate predictions, robust to overfitting	High computation cost, correlated predictions	Material screening, property prediction[21]
Gaussian mixture model	Regression	Robust to outliers, soft clustering	Unsalable in high dimension, sensitive to initialization	Visualization, feature engineering [22]
K-means/median	Clustering	Simple implementation, good interpretability	Sensitive to initialization, vulnerable to data scaling	Visualization, feature engineering [23]
Principle component analysis	Dimension reduction	Decorrelation, good visualization	Linear restriction, incapable for category data	Feature selection [24]

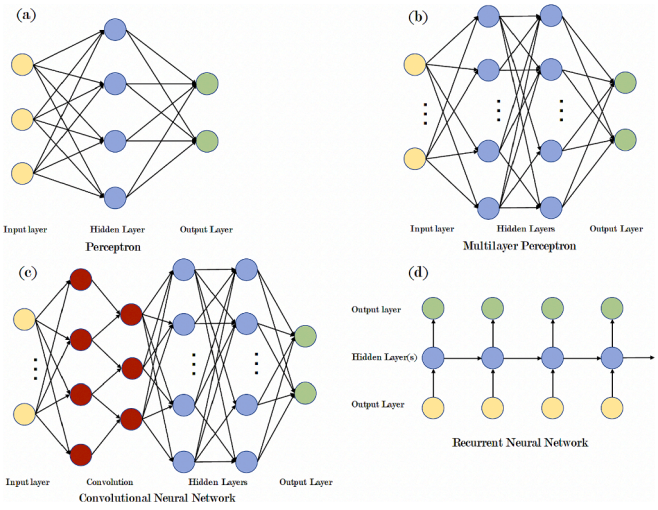

Fig. 1. Architectures of four common neural networks: (a) Perceptron. (b) Multilayer Perceptron. (c) Convolutional neural network. (d) Recurrent neural network. (citation).

The remarkable capacity of DNN models to discern intricate patterns and dependencies within data renders them invaluable in supervised and unsupervised learning setups. In materials science, DNNs find extensive applications, notably in candidate screening and predicting material properties encompassing vital characteristics like bandgaps, conductivity, and crystal structures [58,59]. Specifically, CNNs have emerged as powerful tools for automating tasks such as microstructure analysis, crystallography, and detecting material defects in images and electron microscopy data [60]. Furthermore, DNNs facilitate exploring intricate connections between material properties and crystal structures, elucidating structural motifs and correlations between composition and properties [61]. Finally, neural networks are indispensable in handling substantial datasets, offering promising avenues to ease density functional theory (DFT) calculations and thus drive the development of new energy materials.

2.3. Reinforcement learning

Reinforcement learning (RL) stands as an emerging subfield of artificial intelligence, focused on training intelligent agents to make sequential decisions within dynamic and interactive environments [62,63]. Unlike supervised or unsupervised learning, RL emphasizes acquiring knowledge through iterative interaction and feedback. This dynamic process unfolds through the interaction between one or multiple agents and an environment, where the agent, typically under the control of DNN models, observes an environmental state s_t such from the environment at each time step t. The state is a sufficient statistic of the environment and comprises all the necessary information for the agent to take the best action, such as the configurations of its actuators and sensors. The autonomous agent then interacts with the environment by executing an action at, following the observation of state s_t. The environment, in response, transits to a new state s_{t+1} based on the current state and the action taken and sends a reward (or sometimes penalty) r_{t+1} to the agent. The agent's objective is to learn an optimal policy π—a mapping from states to actions— to maximize the expected cumulative reward over the long term. The agent-environment-interaction loop is illustrated in Fig. 2.

One predominant algorithm in RL is the *Q*-learning algorithm [64]. The algorithm aims to learn the *Q*-function—an optimal real-valued function with action-state input—that quantifies the expected cumulative reward for taking a particular action in each given state and following the policy thereafter. Deep reinforcement learning (DRL) has gained significant prominence in recent years, thanks to the successful integration of deep neural networks with RL [65,66]. Depending on the module learned using deep neural networks, DRL has diversified into various branches, each tailored to the distinct function learned using deep neural networks, including deep *Q*-networks (DQNs) and policy gradient methods [67,68].

RL is witnessing burgeoning applications in material science. For instance, RL can be employed in material discovery and optimization by endowing agents the capacity to select material compositions and the fine-tuning of processing parameters, such as temperature and pressure, to obtain predefined material properties or the optimization of chemical reactions [70,71]. Moreover, RL is also applicable in catalyst design and the optimization of chemical reactions. In this regard, agents can learn optimal reaction conditions, catalyst compositions, and reaction pathways to improve the efficiency and selectivity of chemical processes [72,73]. In addition, RL has demonstrated its utility to optimize experiments in materials science. By strategically selecting the most informative experiments or measurements, RL endeavors to maximize scientific discovery while reducing the experimental costs [74].

2.4. Generative diffusion models

Generative diffusion models (abbreviated as diffusion models) have been recently recognized as the most potent image-generation approach, providing an innovative approach to modeling intricate data distributions [75]. Diffusion models are inspired by diffusion processes rooted in physics and probability theory [76]. In diffusion models, the generative process starts with a basic distribution, typically Gaussian, and progressively refines it to approximate a specific empirical data

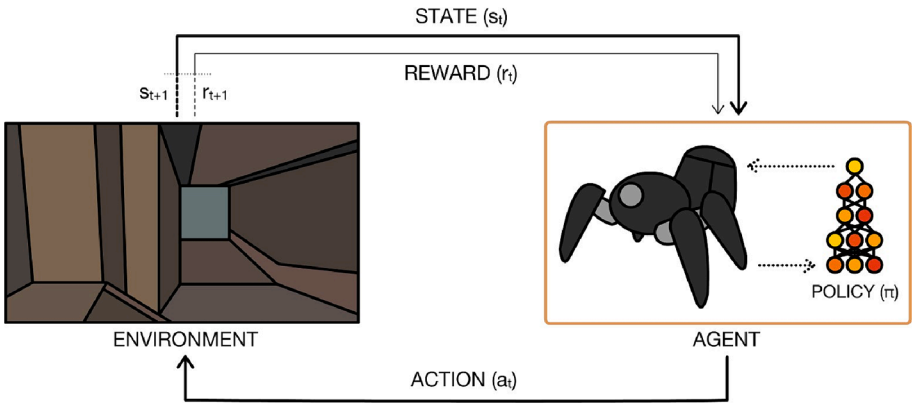

Fig. 2. The agent-environment-interaction loop in RL. At time t, the agent receives state s_t from the environment. The agent obeys its policy π to choose an action at. Once the action is executed, the environment transits to the next state s_{t+1} and provides the agent the reward r_{t+1}. The agent uses knowledge of state transitions $(s_t, a_t, s_{t+1}, r_{t+1})$ to improve its policy. Reproduced from Ref. [69]. Copyright IEEE Signal Processing Magazine.

distribution. This intricate process is achieved through a series of orchestrated transformations, often referred to as "denoising steps". Each denoising step applies a learned transformation to the data, progressively reshaping it to resemble the true data distribution more closely. Notably, the training of diffusion models follows a reversed trajectory, where the data distribution is consistently transformed into a Gaussian distribution by iteratively injecting noise. An underpinning concept of diffusion models is the score function [77], defined as the gradient of the logarithm of the probability density $\nabla \ln(p(x))$. This score function is a vector field that directs the probability density to grow the most rapidly. The workflow of diffusion models is elucidated in Fig. 3.

Diffusion models find primary utility in generating high-quality samples and performing image denoising. This category encompasses representative models such as denoising diffusion probabilistic models (DDPMs) [79], score-based generative models (SGMs) [80,81], and stochastic differential equations (Score SDEs) [82]. All of them are internally modeled with DNNs and show remarkable performance in generating realistic images. Beyond image generation, diffusion models have showcased impressive capabilities in density estimation, facilitating accurate likelihood assessment for a given data point. This ability is particularly invaluable when detecting anomaly detection or quantifying uncertainty.

The application horizon of diffusion models has expanded to material science. Within the context, diffusion models are deployed for generating, simulating, and optimizing materials characterized by desired properties. Notably, diffusion models play a pivotal role in the generation of novel material structures and compositions, significantly enhancing the exploration of chemical space and the creation of materials tailored to specific requirements, such as superconductors [83,84]. Furthermore, these models are harnessed to simulate the properties of materials, including their electronic structures and bandgaps. This capability enables predicting material properties without necessitating resource-intensive and costly experimental endeavors [85]. Also, diffusion models find application in generating synthetic material images, which serve as invaluable resources for training and evaluating image analysis algorithms. These synthetic images faithfully replicate intricate microstructures and defects within materials [21]. Most recently, Duan et al. [86] applied diffusion models to transition state (TS) research, which could greatly accelerate finding TS between reactants and products.

2.5. Large language models

Large Language Models (LLMs) represent a groundbreaking advancement of artificial intelligence. LLMs are initially designed to predict the next word in a sentence but later manifest a deep understanding of grammar, semantics, and general world knowledge. LLMs have an immense scale of billions or even trillions of parameters and incorporate an exceptional architecture named transformer into neural networks, which has proven highly effective in capturing the intricacies of natural language [87]. At the core of the transformer lies the attention module, denoted as Attention $(Q, K, V) = \text{softmax}(QK^{\top})V$, composed of three parameterized matrices: Q (queries), K (keys), and V (values). This module is central to the model's ability to process and understand textual data. An illustrative depiction of single dot-product attention and multi-head attention (a parallel ensemble version) is presented in Fig. 4.

Training LLMs demands immense computational resources, including specialized hardware and a vast corpus of text data encompassing billions or trillions of words. The field of LLMs has witnessed several seminal models. GPT-4 (Generative Pre-trained Transformer by

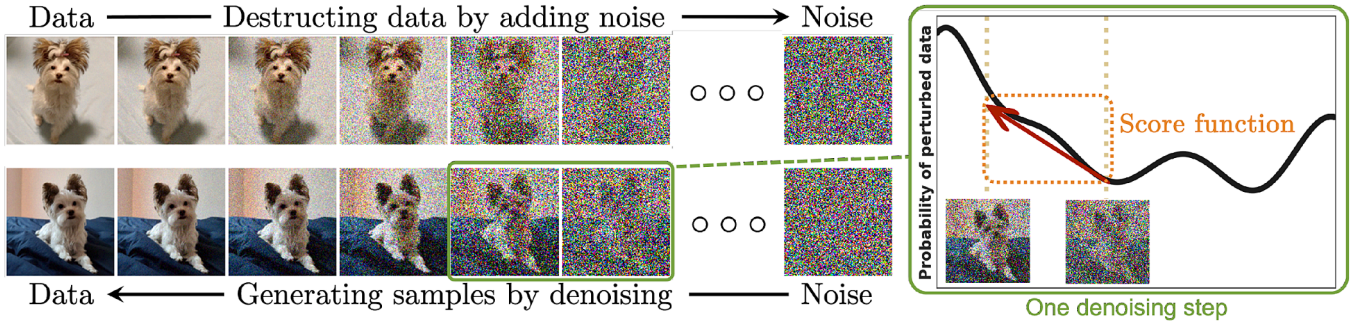

Fig. 3. Diffusion models continually perturb data by injecting noise, then reverse this process to generate new data from randomly sampled noise. Each denoising step in the reverse process requires leveraging the score function, i.e., the directions of data distribution pointing to higher likelihood. Reproduced from Ref. [78]. Copyright ACM Computing Surveys.

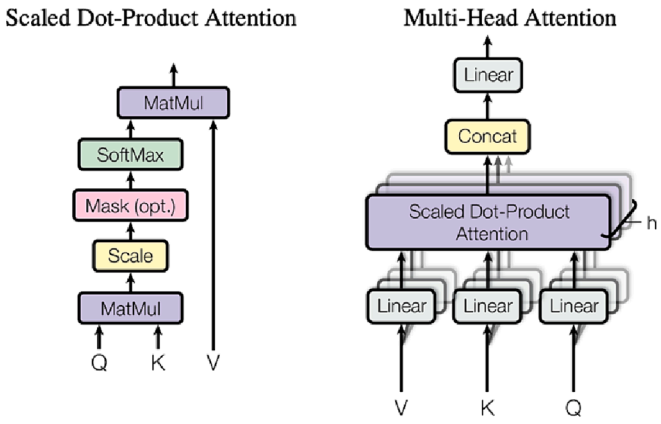

Fig. 4. Left: Scaled dot-product attention. Right: Multi-head attention consists of several attention layers running in parallel. Reproduced from Ref. [87]. Copyright Neural Information Processing Systems.

OpenAI) [88] has achieved a milestone with a staggering 100 trillion parameters. Other examples include Google's BERT [89] and Facebook's RoBERTa [90]. Despite the substantial cost of their training, LLMs manifest exceptional performance in natural language processing. As they are exposed to more data, LLMs grasp an innate understanding of diverse topics and concepts. This understanding empowers them to generate human-like text and answer questions spanning various subjects, ranging from literature and science to history and pop culture. LLMs also evolve through in-context learning, enabling them to adapt to specific tasks or domains without re- training [91]. This adaptability equips LLMs to function as domain experts, responding to domain-specific queries and offering tailored recommendations. An in-depth comparison of LLMs with deep learning and classical ML is referred to Table 2.

LLMs have great potential to advance research in the field of material science. They can assist in predicting material properties and speed up materials discovery. LLMs can serve as natural language interfaces for accessing materials, databases, and tools, allowing researchers to conveniently retrieve information using conversational queries. LLMs can also predict chemical reactions based on input descriptions or constraints, facilitating the design of new materials with desired properties and optimizing chemical processes [92,93]. Furthermore, by analyzing extensive datasets and scientific literature, language models can uncover trends, patterns, and correlations in material science and even generate hypotheses for further experimental investigation [94,95]. A high-level workflow for leveraging language models to enhance chemistry and material science research is as in Fig. 5.

3. AI for battery materials

3.1. Electrode materials

3.1.1. Optimization techniques

Battery performance is significantly influenced by both the crystal systems of the materials used and the voltages of the electrodes. Crystal systems impact the battery's longevity during charge–discharge cycles due to differences in structural stability, ionic conductivity, and electrochemical stability. On the other hand, electrode voltages play a role in determining theoretical capacity, energy efficiency, and safety. Higher voltage materials generally offer greater energy density but may be more susceptible to unstable chemical reactions under extreme conditions. So, achieving a balance between crystal systems and electrode voltages is essential in battery research and development to meet various application requirements, considering factors such as energy and power density, stability, and safety.

Machine learning is significantly improving lithium-ion battery technology by optimizing electrode materials. Techniques like Artificial Neural Networks (ANN), Support Vector Machines (SVM), Random Forest (RF), K-Nearest Neighbors (KNN), and Extremely Randomized Trees (ERT) effectively predict cathode material properties. These algorithms focus on specific features like space group, formation energy, and density to accurately determine crystal systems and electrode voltages, key for battery performance [96]. RF and ERT are particularly effective in analyzing crystal volume and site number. Additionally, techniques like ab initio calculations and Partial Least Squares (PLS) analysis [97] are crucial for understanding lithium-ion intercalation in cathodes. The effectiveness of these methods depends on having comprehensive datasets and selecting relevant descriptors. The successful use of these varied algorithms and feature selection techniques demonstrates machine learning's vital role in advancing lithium-ion battery electrode materials.

Table 2
Comparison of Classic ML, Deep Learning, and LLMs.

Comparison	Classic ML	Deep Learning	LLMs
Training Data Size	Moderate	Large	Very large
Input Datatype	Numeric	Numeric	Multimodal
Human Interaction	No	No	Yes
Feature Engineering	Manual	Automatic	Automatic
Model Complexity	Simple	Complex	Very Complex
Interpretability	Good	Poor	Poorer
Performance	Moderate	High	Highest
Hardware Requirements	Low	High	Very High

3.1.2. Performance and durability analysis

The analysis of performance and durability of electrode materials in lithium-ion batteries increasingly relies on advanced machine learning methods. These methods play a key role in predicting and enhancing the properties of the materials. For analyzing electrode material behavior, techniques like Regression Analysis, Support Vector Machines (SVM), and Random Forest (RF) are employed. They predict how electrode materials perform under various conditions, forecasting their response to stress, degradation, and long-term usage. This is vital for the reliability and lifespan of the batteries. In durability assessment, machine learning is crucial for understanding and predicting material degradation over time. This involves examining capacity fade, mechanical stress, and chemical stability under different charging and discharging cycles [45,98,99]. Additionally, algorithms model the long-term behavior of electrode materials, helping in selecting materials that balance performance and durability. This aims to improve their performance and lifespan. There is also a continuous effort to refine the machine learning models themselves to ensure their accuracy and effectiveness as electrode materials and battery technologies evolve. Overall, machine learning is essential in the performance and durability analysis of electrode materials, driving the development of more efficient and longer-lasting lithium-ion batteries. These techniques foster a deeper understanding of material behavior, advancing battery technology.

3.1.3. Cathode

Various machine learning (ML) approaches have been employed to gain insights and optimize their properties. For Ni-rich NMC cathode materials, Min et al. utilized ML-aided analysis to determine optimal synthesis parameters, incorporating input variables such as calcination temperatures, Ni content, and particle size, while assessing output variables like initial capacity, cycle life, and residual Li [100]. Among the ML models employed, Elastic Random Tree (ERT) performed the best, and the optimal experimental parameters suggested by ML were validated through additional experiments.

In the case of Li-rich layered oxide cathodes, Kireeva and Pervov employed a Support Vector Machine (SVM) model to uncover synthesis-property relationships, considering input variables such as composition,

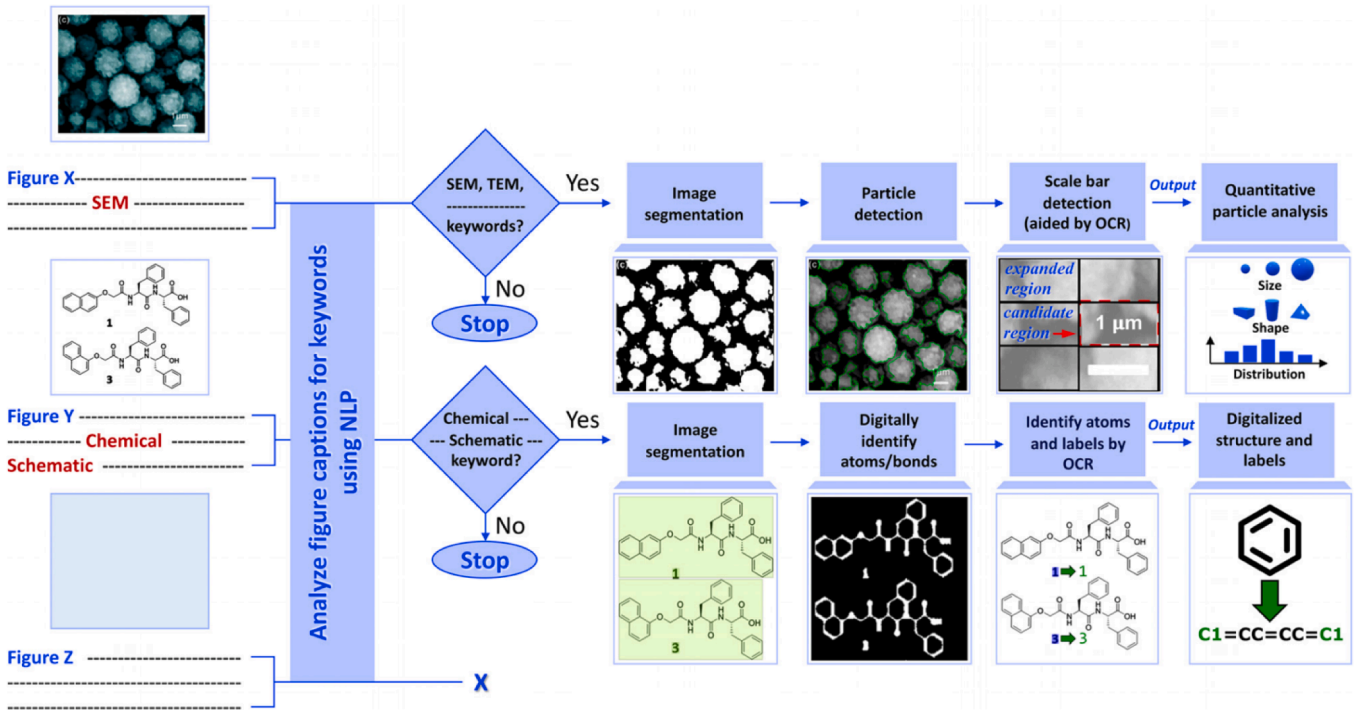

Fig. 5. Image extraction schematic including examples derived from microscopy images or molecular structures. Reproduced from Ref. [94]. Copyright Applied Physics Reviews.

synthesis method, and temperatures, with outputs including discharge capacity and Coulombic efficiency [101]. The ML analysis facilitated the identification of critical parameters influencing tailored cathode characteristics.

In the study of low-strain cathode materials, Wang et al. constructed a Partial Least Squares (PLS) model using lattice constants calculated through Density Functional Theory (DFT) to predict volume changes in cathode materials [97]. Key descriptors for accurate predictions were the radius of transition metal ions and their octahedron distortion.

Joshi et al. focused on voltage profiles for various battery chemistries, utilizing data from the Materials Project. Descriptors included intercalation cation nature and lattice type, leading to the identification of potential new electrode materials for Na- and K-ion batteries [102].

In the context of charge-informed atomistic modeling, CHGNet, a machine-learning interatomic potential (MLIP), was published. It models the universal potential energy surface, pre-trained on a decade's worth of density functional theory calculations from the Materials Project Trajectory Dataset. CHGNet (Fig. 6) stands out for its ability to explicitly include magnetic moments, enhancing its capability to represent both atomic and electronic degrees of freedom accurately. This feature allows it to provide deeper insights into ionic systems, demonstrating applications in areas like charge-informed molecular dynamics and Li diffusion. CHGNet's integration of charge information marks a significant advancement in simulating complex electron interactions in solid-state materials [103].

For organic electrode materials, Allam et al. created a DFT-based database for various organic molecules and employed an Artificial Neural Network (ANN) to predict redox potentials [104]. Input variables encompassed electronic properties and geometrical information, revealing the effectiveness of this approach in predicting redox potentials by considering important variables such as electron affinity, HOMO, LUMO, etc.

ML aids in enhancing the electrochemical properties of organic radical materials, crucial for improving the performance of electrode materials in redox flow batteries [105]. In the classification of data sets, Shandiz *et al.* explored various algorithms to classify Li-ion silicate-based cathodes [96]. ML models including Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Sparse Discriminant Analysis (SDA), ANN, SVM, kNN, RF, and ERT were evaluated. RF and ERT classifiers emerged as the most effective, with relevant descriptors such as crystal volume, number of sites, formation energy, energy above hull, and band gap identified in the classification process. These studies collectively highlight the diverse applications of ML in optimizing and understanding various aspects of battery cathode materials.

3.2. Electrolyte Materials

3.2.1. Ionic solid electrolyte

Machine learning (ML) techniques have effectively optimized ion mobility, conductivity, and dendrite growth suppression.

Through the utilization of computational data, partial least squares (PLS) analysis, and density functional theory (DFT), Min et al. [97] employed NEB with DFT to compute Li^+ migration energies, identifying low-ionic-conductivity solid electrolytes with olivine-type oxide structures. Similarly, Jalem et al. [106] employed artificial neural networks (ANN) to discover low-migration-energy compositions within Tavorite-Type $LiMTO_4F$ solid electrolytes, highlighting ML's capacity for identifying fast ion conductors.

Studies by Kireeva and Pervov on garnet-type oxide solid electrolytes utilized support vector machine (SVM) models to predict ionic conductivity, thereby enriching the understanding of material properties [102]. Fujimura et al. [107] integrated theoretical and experimental data with SVM to evaluate $Li_8 - cAaBbO_4$ LISICONs, pinpointing compositions with enhanced ionic conductivity. Zhang et al.'s application of clustering methods to categorize materials based on Li-ion conductivity underscores the practicality of such ML approaches [108].

Sendek et al. [109] used classification LR to distinguish superionic from non-superionic Li-containing solid electrolytes, emphasizing multi-descriptor schemes' effectiveness. Ahmad et al. screened inorganic solids for dendrite suppression using DFT and trained CGCNN and KRR models. They identified dendrite-suppressing candidates as soft, anisotropic, and with high mass density.

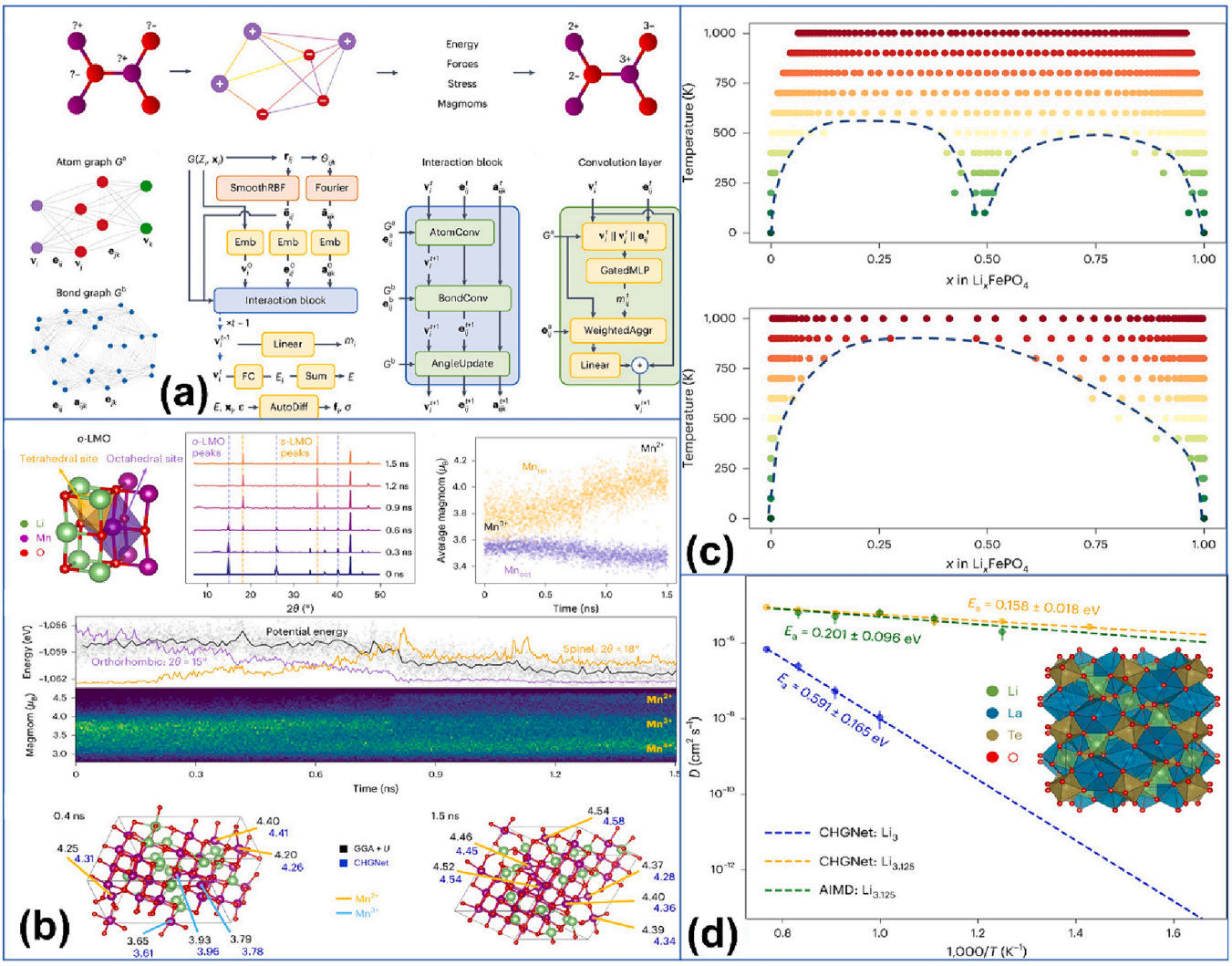

Fig. 6. (a) CHGNet model architecture. (b) Li0.5MnO2 phase transformation and charge disproportionation (c) LixFePO4 phase diagrams from CHGNet (d) Li diffusivity in garnet Li3La3Te2O12. Reproduced from Ref. [103]. Copyright 2023, Nature Portfolio.

Polymer-based solid electrolytes have diverse applications, with studies primarily focusing on dendrite growth inhibition and the integration of ceramic composite fillers. Notably, polyethylene oxide (PEO) stands out due to its ability to establish a stable and resilient solid electrolyte interface (SEI) layer on lithium metal anodes, thereby extending battery lifespan. This phenomenon is attributed to the formation of lithium fluoride (LiF), which forms a protective barrier on the anode surface. Other significant polymers include polyacrylonitrile (PAN), aliphatic polycarbonates (APCs), polyvinylidene fluoride (PVDF), and polyacrylates (PAs), which can serve as co-polymers or crosslinking agents alongside PEO. Moreover, polymers can function as fillers in composites, synergizing with lithium-ion conductors or ceramics such as NASICON [110–113].

3.2.2. *Liquid electrolytes*

ML studies are relatively less common due to liquid electrolytes' highly disordered nature. However, ML methods still offer utility in various aspects of studying liquid electrolytes.

Nakayama et al. focused on predicting cation-solvent interaction energies within liquid electrolytes for lithium-ion batteries (LIBs) using an exhaustive search with a Gaussian Process (GP) method [114]. These interaction energies are crucial for Li-ion transport and play a significant role in solvation and desolvation processes at interfaces.

Sodeyama *et al.* utilized an exhaustive search with linear regression (ES-LiR) method to incorporate the melting point as a target property for liquid electrolytes. This approach struck a balance between prediction accuracy and computational cost [115].

ML-enhanced molecular dynamics (MD) simulations have been applied to simulate extreme types of liquid electrolytes, particularly those with significant dipole polarization, such as highly concentrated electrolytes and ionic-liquid-based electrolytes [116,117]. Neural Networks and Deep Tensor Neural Networks are employed to handle atom-wise system descriptions and polarization terms.

The study of Zn^{2+} in water involved applying ANN to learn an effective physical potential for highly disordered systems. Challenges persist when extending this method to more complex systems with anions, cations, and solvents.

Preliminary results from the Johansson Group and MIT Group indicated the potential efficacy of using ML techniques on highly concentrated electrolytes (HCE) of LiTFSI in ACN, with a small mean average error (MAE) observed [17,118].

Furthermore, ML is not limited to computational analysis but can also enhance experimental studies of electrolytes. Applications range from using ANN to interpret spectroscopy data to the concept of fully automated laboratories [119], offering efficient ways to address complex simulation and experimental data comparison challenges.

In conclusion, while ML studies on liquid electrolytes face challenges due to their disordered nature, still many successful applications and

potential areas where ML can contribute, including predicting interaction energies, aiding in simulations, and enhancing experimental studies. These efforts address the challenges in understanding and optimizing liquid electrolytes for battery technologies.

3.2.3. Property prediction and analysis

ML is revolutionizing the prediction of electrolyte properties in LIBs. While solid electrolytes are extensively studied, liquid electrolytes present challenges due to their disordered structures. The key is understanding ion transport, especially the properties of lithium-ion and solvent interactions. Researchers use ML techniques like multiple linear regression and Gaussian processes to predict these properties more accurately [114,120,121]. Additives in electrolytes, crucial for performance, are also analyzed using ML, focusing on their redox potential. Solid electrolytes are highlighted for their potential in safer, high-performance batteries, with ML aiding in predicting important properties like ion conductivity and mechanical strength. The main challenge in applying ML here is the availability of comprehensive data for more accurate predictions.

3.2.4. Design and optimization approaches

Solid and liquid electrolytes in LIBs present contrasting challenges. Solid electrolytes, extensively researched, differ from liquid electrolytes, which are complex due to their disordered structures. Understanding ion transport, particularly lithium-ion and solvent interactions, is key in this domain. Machine Learning (ML) techniques like multiple linear regression and Gaussian processes are increasingly used to predict essential electrolyte properties more accurately [114,120]. Additionally, ML is employed to analyze electrolyte additives, focusing on their redox potential, a critical factor for battery performance and longevity. Solid electrolytes are promising for creating safer, high-performance batteries. ML helps predict vital properties like ion conductivity and mechanical strength, essential for their practical use. However, applying ML effectively in electrolyte research requires comprehensive, high-quality data. Accurate predictions hinge on extensive datasets covering various electrolyte properties and compositions. Therefore, gathering and curating quality data is vital for maximizing ML's potential in advancing LIB technology. In summary, ML's integration into electrolyte research represents a significant step toward optimizing and innovating LIBs. With ongoing research and data enhancement, ML is poised to greatly improve our understanding and development of advanced battery systems.

3.3. Multiscale modeling

Traditional first-principles methods, such as DFT, offer high accuracy but are computationally demanding, making them most suitable for small systems. In contrast, interatomic potentials, while less computationally intensive, are limited by transferability and nonreactivity issues, restricting their application to specific systems.

3.3.1. Model accuracy

To overcome these limitations, ML-assisted approaches have emerged as valuable tools. ML techniques enhance the representation of many-body interatomic potentials [122,123], as demonstrated by Gaussian approximation potentials (GAP) and ANN potentials. Several notable case studies illustrate the impact of ML in molecular dynamics simulations:

Deringer and Csanyi's GAP Model leverages ML to simulate amorphous carbon materials, significantly improving prediction accuracy for complex compounds [124]. Li et al.'s work demonstrates the effectiveness of an ANN potential trained on DFT-computed structures of Li_3PO_4, enhancing potential models for more accurate material property predictions [125]. Deng *et al.*'s eSNAP Model combines the SNAP formalism with electrostatic interactions, improving property prediction in materials like Li3N, particularly for complex interactions [126]. Shao et al. utilize an ANN potential to simulate ionic conductivity in alkaline electrolyte solutions, showcasing ML's ability to reduce computational costs while maintaining high accuracy, a significant advantage for complex material systems [127].

Specific ML applications encompass cluster expansion methods for estimating local properties, DeepDFT for charge density prediction in large systems (e.g., NMC cathodes and liquid electrolytes), and unsupervised learning methods like density-based clustering of trajectories (DCT) for elucidating complex diffusion mechanisms in solid electrolytes [128,129].

In addition, the integration of CHGNet, a graph neural network-based MLIP, demonstrates a significant advancement in multiscale modeling. CHGNet's ability to accurately represent both atomic and electronic degrees of freedom, thanks to its pretraining on extensive DFT calculations, offers improved predictions in complex systems, particularly in solid-state materials. Its inclusion exemplifies ML's capability to handle intricate electron interactions in multiscale simulations, further contributing to the development of sophisticated battery technologies [105].

Overall, ML plays a pivotal role in advancing battery technology by enabling more precise and efficient simulations, addressing the intricate physical and chemical properties of battery materials. This integration is crucial for developing superior battery technologies that cater to the diverse nature of battery materials and their behaviors. ML also addresses computational challenges, facilitating large-scale simulations of complex compounds, including amorphous materials. However, ML potentials require substantial reference data, especially for disordered structures.

3.3.2. Computational efficiency

ML methods, such as Gaussian approximation potentials (GAP) and artificial neural network ANN potentials, are applied in multiscale modeling for lithium-ion battery research. These ML techniques enhance the representation of many-body interatomic potentials, offering a more computationally feasible alternative to first-principles methods like Density Functional Theory (DFT). ML facilitates simulations of complex compounds, including those with amorphous structures. This is beneficial for modeling battery material behaviors that influence performance and longevity. ML also increases sampling efficiency in cluster expansion methods [105,130], which is useful for connecting atomic scale interactions with larger scale phenomena in battery materials. In summary, ML in multiscale modeling for lithium-ion batteries contributes to more efficient and accurate simulations, aiding in the development of improved battery technologies.

In conclusion, ML plays a pivotal role in advancing multiscale modeling of battery materials. It bridges the gap between atomic and larger length and time scales, benefiting particularly the study of disordered systems. ML models significantly enhance computational efficiency without compromising accuracy, addressing the complexities of battery materials and contributing to the development of improved battery technologies.

3.4. Experimental planning and synthesis

3.4.1. Methodologies and techniques

In the field of design for LIBs, ML is employed in various capacities to enhance battery performance and design:

One key application is in Electrolyte Design, where ML is utilized to identify organic solvents with lower reduction potentials [131]. Typically employing supervised learning algorithms, this approach improves the reduction and oxidation stability in electrolytes, thereby contributing to the expansion of electrochemical windows.

ML also plays a vital role in Cathode Material Improvements. Through regression analysis, ML helps predict structural changes and voltage behaviors in NMC cathodes [97]. This predictive capability facilitates the optimization of cathode designs with the aim of enhancing

reversible capacities and reducing voltage decay.

Additionally, ML models, potentially incorporating classification algorithms, are instrumental in Anode Materials and Coatings [132,133]. These models pinpoint effective materials for interfacial coatings, leading to improved cyclability and overall battery performance.

In the realm of All-Solid-State Batteries, unsupervised learning techniques like clustering are employed to classify solid-state Li-ion conductors. This classification aids in the discovery of new, fast Li-conductors with improved conductivity. These diverse applications of ML underscore its significant and multifaceted role in advancing the development of sophisticated materials for LIBs.

3.4.2. *Case studies and practical applications*

Liu *et al.* used ML to explore the stability of doped garnet LLZO against Li metal [135]. Ibhahim *et al.* applied a Bayesian neural network for modeling conductivity in polyethylene oxide-based electrolytes [136,137]. Hatakeyama-Sato *et al.* (Fig. 7) utilized Gaussian process models to discover high-conductivity Li-ion conductive polymers [134]. Joshi et al. employed ML models to predict voltage profiles in Na- and K-ion battery cathodes [104]. Wang et al. studied cathode material volume changes during de-lithiation [138]. Zhang *et al.* modeled lithium polysulfide adsorption energy on layered sulphides [108]. Shree's framework combining reinforcement learning with an ML surrogate objective, trained using quantum chemistry simulations, exemplifies the advanced computational strategies employed for these studies [106]. Natarajan and Van der Ven, and Hochins and Visvanathan, used neural networks for analyzing electrode materials' disorder during lithiation/de-lithiation [123,139].

Also, machine learning tools could be utilized to analyze SEI. Soleymanibrojeni et al. [140] presents an active learning workflow coupled with a kinetic Monte Carlo (kMC) model for SEI formation, integrating statistical tools to reduce model uncertainty and providing a new perspective on the reactions controlling SEI formation in Li-ion batteries, crucial for enhancing battery performance and lifespan.

3.4.3. *High-Throughput methods*

Traditional experimental processes pose significant challenges in materials discovery for batteries, as they are slow, costly, and heavily reliant on chemical intuition and trial-and-error approaches. This limitation hampers the rapid exploration of novel materials.

To address these challenges, high-throughput (HT) synthesis methods have emerged as a promising alternative (Fig. 8).[141] These methods, often combinatorial in nature, facilitate the exploration of vast compositional spaces and synthesis conditions for battery materials. Examples include the investigation of Si − M thin films, the study of $LiNi_{1/3}Mn_{1/3}Co_{1/3}O_2$ preparation conditions, and the mapping of phase diagrams for various systems [142]. However, the adoption of HT synthesis requires alternative automated setups and a shift in sample handling, characterization, and data analysis paradigms.

ML plays a significant role in this context by enabling the analysis of large and complex experimental datasets, uncovering correlations between synthesis conditions and sample properties. Despite its significant potential, ML has yet to be widely applied in the battery field.

Another promising avenue is the integration of robotics with AI or ML for autonomous material synthesis and electrode/cell optimization, although this approach remains underreported in the battery field. Several case studies illustrate the practical application of these concepts. For instance, Beal and co-workers produced thin-film sample libraries of Li-ion electrolyte materials using HT physical vapor deposition (PVD) [143] and employed various analysis techniques, including LA-ICP-MS, spectroscopic ellipsometry, impedance, and X-ray diffraction (XRD) [144,145]. They applied techniques such as PCA, multivariate curve resolution with alternating least squares (MCR-ALS), recursive partition analysis, and ANN to analyze XRD data [143,146,147].

In another example, Ceder's group conducted literature data mining to examine synthesis conditions for metal oxides and sulfides across thousands of articles. They employed parse trees and neural network word labeling to extract synthesis step parameters and identified relationships between synthesis temperature and compositional complexity. However, challenges such as standardization in data presentation and the lack of negative result reports remain significant hurdles for text mining.

The key takeaways from these developments include the integration of HT experimental approaches with ML analysis, enabling the generation of large experimental datasets and more efficient discovery of battery materials. Looking ahead, further advancements in AI-aided approaches, robotics, and literature data mining have the potential to significantly enhance the efficiency and effectiveness of materials discovery in the battery field. In summary, the application of HT methods, combined with ML and AI techniques, represents a significant advancement in the experimental planning, screening, and synthesis of battery materials, offering the potential for more rapid and cost-effective development of new battery materials [143].

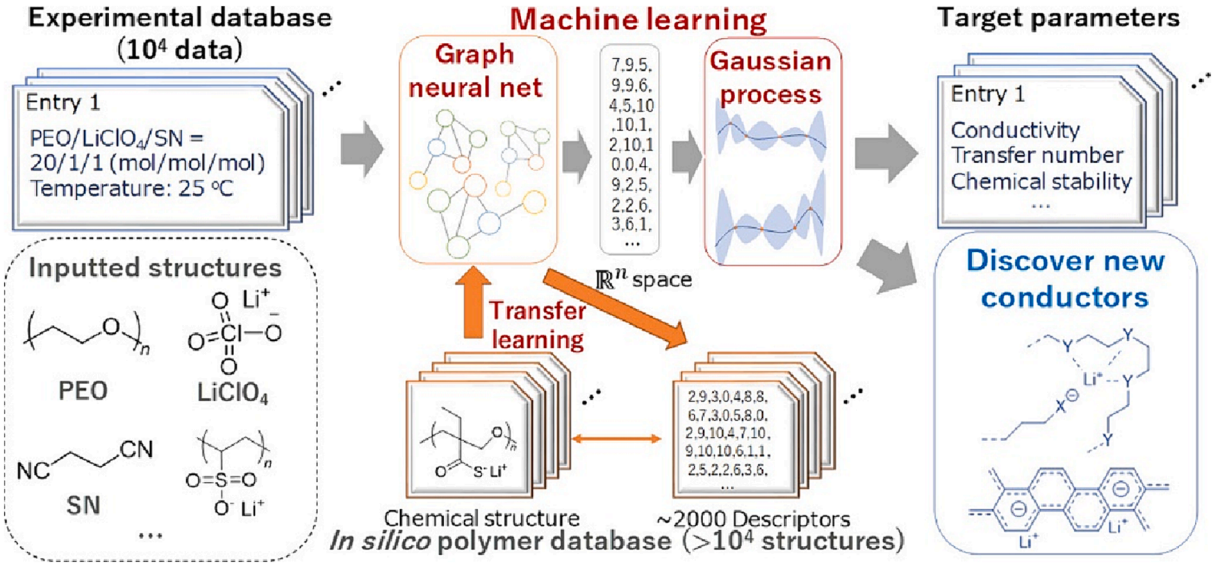

Fig. 7. AI-based prediction scheme for solid polymer electrolyte properties, specifically conductivity, Reproduced from Ref. [134]. Copyright 2020, American Chemical Society.

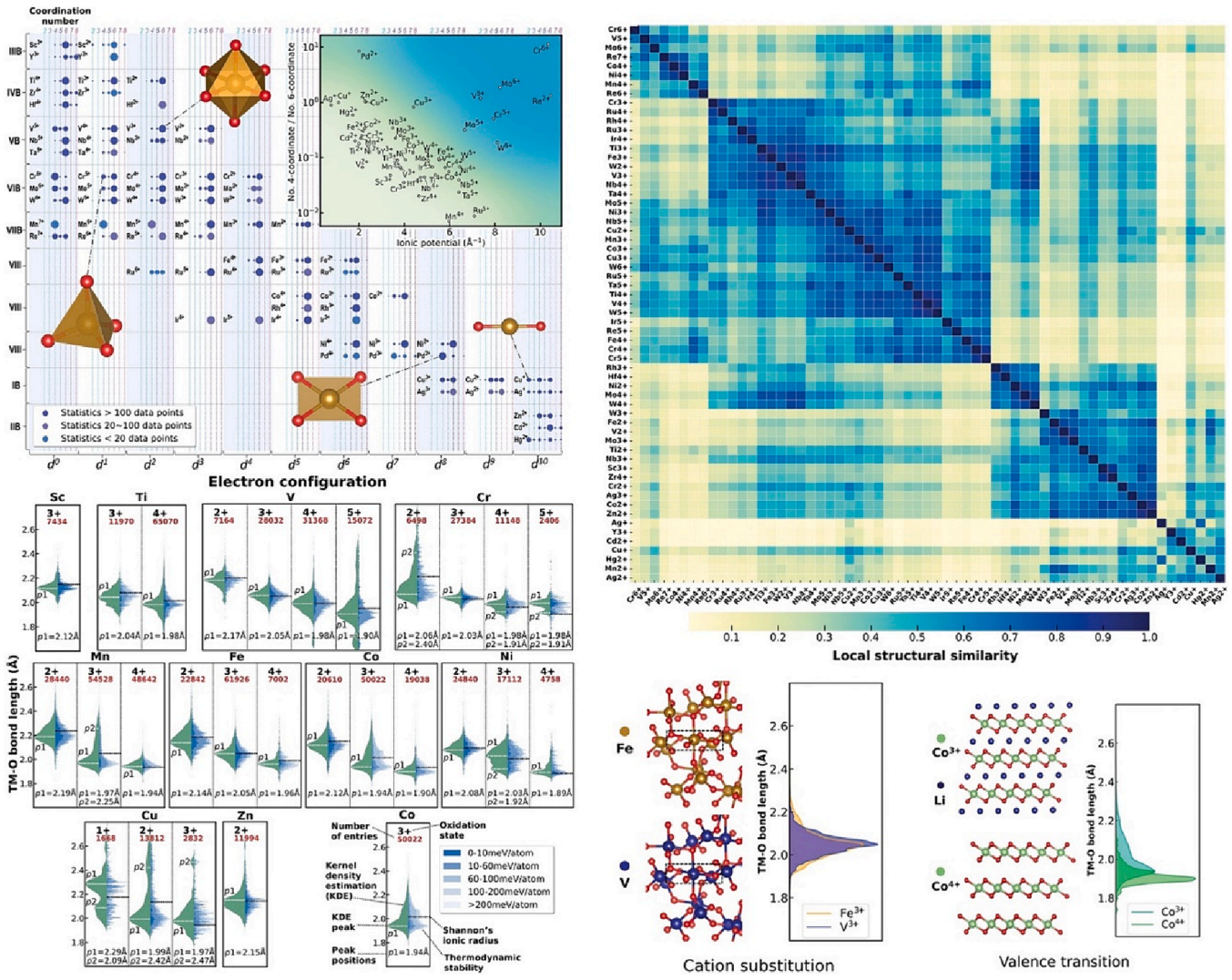

Fig. 8. High-throughput methods investigating transition metals' ionic properties, Reproduced from Ref. [141]. Copyright 2022, Wiley.

4. AI in fuel cells

In the contemporary landscape of the world economy, the integration and efficient management of sustainable electrical energy generation are paramount. Traditional methods of energy production have exerted a profound impact on the global environment, contributing significantly to climate change [148]. A stark illustration of this is the revelation from the International Energy Agency, which warns of "Energy-related greenhouse gas emissions leading to considerable climate degradation with an average 6 °C global warming" [24]. This alarming situation underscores the urgency for alternatives, and clean energy emerges as a viable solution to mitigate these challenges.

Fuel cells (FCs) represent a significant advancement in sustainable energy technology. As electrochemical devices, they convert chemical energy, most commonly from hydrogen, directly into electrical energy through an electrochemical reaction with oxygen [149,150,237]. This process is intrinsically efficient and environmentally friendly, with water often being the only by-product, starkly contrasting with the emissions from conventional combustion-based energy generation [151–153]. The operation of FCs circumvents the significant environmental impacts associated with traditional energy sources, aligning with the global imperative for green and sustainable energy solutions.

The benefits of FCs extend to high energy conversion efficiency, often exceeding 60 %, surpassing many traditional energy generation methods [154]. This efficiency, combined with the modularity of FCs, enables their use in diverse applications, ranging from portable devices to large-scale power systems. FCs are also known for their reliability and quiet operation, making them suitable for a variety of environments, including residential and sensitive industrial areas [155]. The FC market has responded to these advantages with significant growth, expanding from $630 million in 2013 to $2.54 billion in 2018 [24]. This expansion reflects the broadening application of FCs across multiple sectors, including residential power, transportation, and industrial uses [156].

FCs are differentiated by electrolyte type—acidic or alkaline—and operating temperature. Alkaline Fuel Cells (AFCs) use aqueous KOH and work at cooler temperatures up to 100 °C, suitable where pure hydrogen and oxygen are used [157]. Proton Exchange Membrane Fuel Cells (PEMFCs) and Direct Methanol Fuel Cells (DMFCs) employ acidic polymers and operate at similar mild temperatures, making them ideal for portable applications due to quick start-up times [158]. Phosphoric Acid Fuel Cells (PAFCs) operate at intermediate temperatures (150–220 °C) and can handle impure fuels, while Molten Carbonate Fuel Cells (MCFCs) work at higher temperatures (550–700 °C) suitable for stationary power generation with the capability of CO_2 recycling [159,160]. Solid Oxide Fuel Cells (SOFCs), operating at the highest temperatures (600–1000 °C), allows for the use of various fuels and non-precious metal catalysts, thus reducing costs and enhancing fuel flexibility for large-scale applications [161].

4.1. Current challenges

The path to mainstream adoption of FC technology is fraught with challenges, particularly in the domains of cost reduction, efficiency optimization, and durability improvement [40]. The complexity inherent in FC systems arises from the delicate interplay of phenomena such as mass transport, heat transfer, and electrochemical reactions, which are influenced by the materials used, the design of the system, and the operating conditions [162].

The development of FCs demands breakthroughs in material science, a deeper understanding of the fundamental processes, and enhancements in analytical and experimental methodologies. For instance, the use of non-platinum group metal catalysts and advanced membrane materials, as well as strategies for reducing platinum loading, are critical for cost management [163]. Additionally, improved water and thermal management techniques, coupled with sophisticated system health monitoring, are key for maintaining high performance and extending the lifespan of FCs [164].

ML has the potential to tackle these multifaceted challenges by leveraging its strengths in data analysis, pattern recognition, and predictive modeling. ML algorithms can process vast amounts of experimental and operational data to unearth insights that can inform material selection, FC design, and operational optimization. These insights can lead to the identification of cost-effective materials and designs that do not compromise performance or durability. ML can also assist in real-time control and diagnostics of FCs, optimizing their operation under dynamic conditions and predicting maintenance needs before failures occur. This proactive approach to maintenance can significantly enhance the durability and reliability of FC systems.

Furthermore, unlike traditional physical models, ML models can efficiently handle the complex, multi-physical interactions within a FC without explicitly programming the underlying phenomena [165]. This approach is advantageous when physics is not fully understood or too complex to model directly. Once trained, ML models can rapidly evaluate scenarios and simulate conditions for large, multi-cell systems, which would be computationally intensive with physics-based models. This efficiency is crucial for the iterative design and testing cycles needed to propel FC technology forward.

In conclusion, ML stands as a transformative tool in the FC industry, offering a route to overcome current technical and economic barriers. It can accelerate the development of FCs by enhancing our understanding of their operation, improving design and material selection, and optimizing performance and longevity, thus propelling FCs toward competitive commercialization.

4.2. Applications of ML in fuel Cells

This section methodically examines the application of ML techniques in FC technology, encompassing aspects of design, performance evaluation, and durability prediction. The design and performance evaluation of FCs focus on areas such as material selection and chemical reaction modeling, which directly impact the polarization or current–voltage (I-V) curves. Durability prediction, an essential aspect of FC technology, involves assessing the State of Health (SoH), estimating the Remaining Useful Life (RUL) of the FCs, and conducting fault diagnostics. SoH is a vital measure of an FC's condition, indicating its level of degradation and capacity to carry out particular discharge activities. By monitoring SoH, one can enhance performance consistency and prolong the lifespan of the FC, thereby diminishing running expenses. Fault prediction is aimed at identifying potential issues before they lead to system failure. Fault diagnosis in FCs involves determining the nature and location of faults, such as identifying occurrences of flooding or other abnormal conditions.

4.2.1. Fuel cell design and performance evaluation

ML enhances FC design by efficiently processing complex data that would be formidable to handle through manual methods. It streamlines material selection by scouring comprehensive databases to pinpoint optimal candidates for FC components like membranes and catalysts, taking into account a range of properties and performance metrics. This approach is becoming increasingly prevalent in the fields of chemistry and materials science for the discovery of new material properties and the development of advanced materials [166]. Traditional approaches to material property characterization, such as experimental measurements and theoretical calculations, are often resource-intensive, demanding considerable cost, time, and computational power. Furthermore, the myriad of factors influencing material properties complicates the quest for suitable materials using conventional techniques alone. ML provides a significant advantage in this realm by leveraging existing data to predict material properties, even when the underlying physical phenomena are not fully understood, thus enhancing the search for materials and expediting the development process with increased efficiency.

In an innovative study focused on optimizing the performance of PEMFCs, Wang et al. [167] developed a novel artificial intelligence (AI) framework to enhance the catalyst layer (CL), the core electrochemical reaction region of PEMFCs. Existing experimental and modeling methods have proven insufficient for the in-depth optimization of the CL composition. This limitation is mainly due to the complexity involved in accurately modeling and experimentally testing the numerous variables that influence the CL's performance. Such variables include the concentration and distribution of the catalyst, ionomer, and pore structures within the CL. Traditional methods often fail to capture the intricate interactions between these variables, leading to suboptimal CL designs that do not fully exploit the potential of the PEMFCs. To address these problems, this study introduced a combination of a data-driven surrogate model and a stochastic optimization algorithm. This approach aimed at achieving global optimization of multiple variables to improve the maximum power density of PEMFCs. The foundation of this optimization framework was a database generated from simulation results of a three-dimensional computational fluid dynamics (CFD) PEMFC model (Fig. 9a), coupled with a CL agglomerate model (Fig. 9b). This database served to train the surrogate model, which was based on the Support Vector Machine (SVM) algorithm. The surrogate model's prediction performance was impressive, with a squared correlation coefficient (R-square) of 0.9908 and a mean percentage error of just 3.3375 % in the test set. This indicated that the surrogate model's accuracy was comparable to that of the physical CFD model but was far more efficient in terms of computation resources. For instance, the surrogate model could calculate a polarization curve in about a second, a task that could take hundreds of processor-hours using the physical CFD model. The efficiency of the surrogate model was further leveraged by incorporating it into a Genetic Algorithm (GA) to determine the optimal CL composition. To validate this approach, the optimal composition identified by the surrogate model was tested against the physical model. The comparison revealed a minimal percentage error of only 1.3950 % between the surrogate model's predictions and the physical model's simulations of maximum power densities under the optimal CL conditions (Fig. 9c).

In another study addressing the complexities of PEMFCs, Ding et al. [168] showcased the efficacy of explainable AI in enhancing the design of high-performance membrane electrode assemblies (MEAs). The study's workflow began with the development of a multiphysics model, experimentally validated to reliably predict current–voltage (IV) curves under various conditions. The researchers then utilized 11 key experimental parameters to construct surrogate models through eight different ML algorithms. A critical aspect of this study was the application of interpretative methods such as feature ranking, partial dependence plots (PDP), and SHAP analysis on these AI models. This analysis was instrumental in understanding how each input parameter influenced the output performance. A notable finding was the need for a tradeoff in catalyst layer thickness, balancing chemical reaction and mass transfer. The ANN model, with an impressive R^2 of 0.99854, emerged as the best-

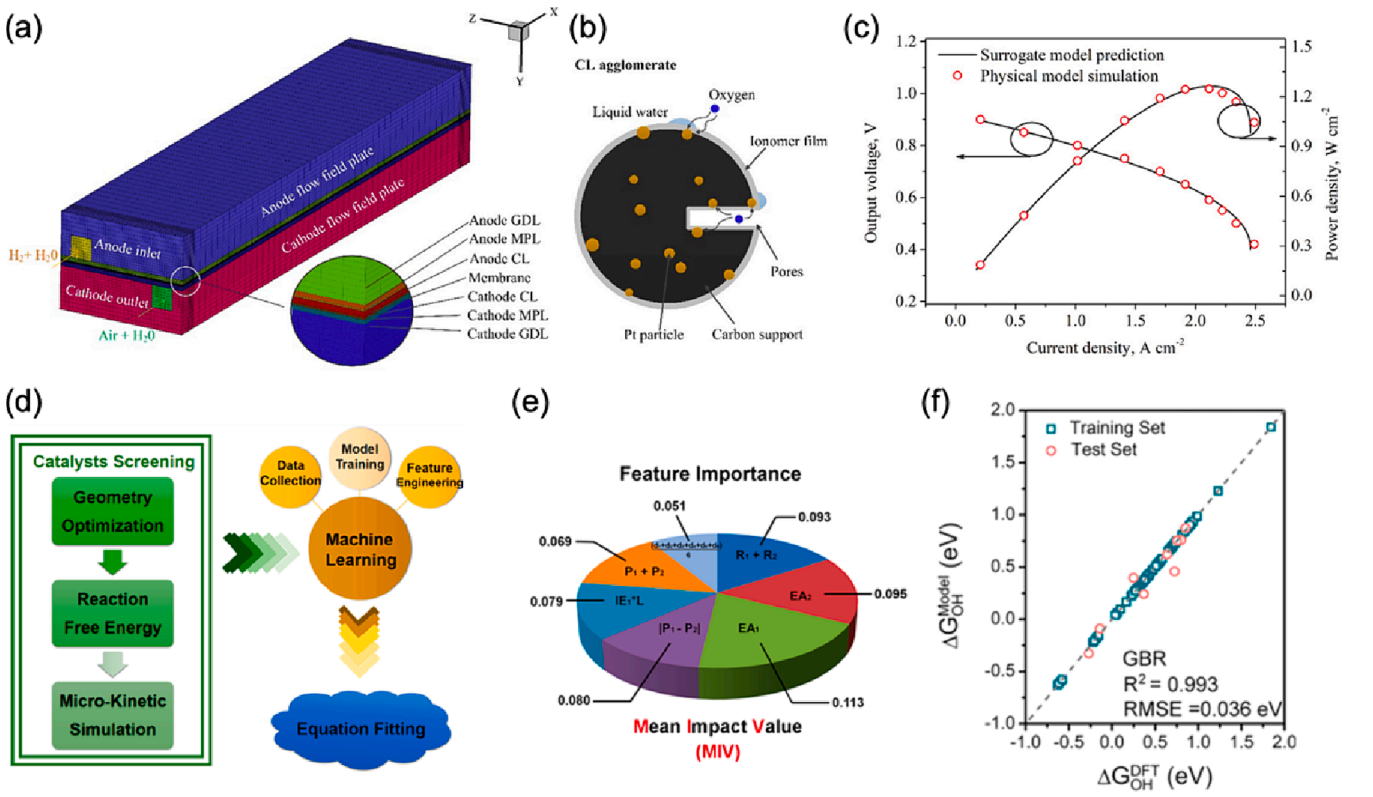

Fig. 9. (a) Schematic of the simulated PEMFC computational domain of the physical model. Reproduced from Ref. [167]. Copyright 2020, Elsevier. (b) Schematic of the CL single agglomerate. Reproduced from Ref. [167]. Copyright 2020, Elsevier. (c) Surrogate model predicted and physical model simulated polarization curves under the optimal CL composition. Reproduced from Ref. [167]. Copyright 2020, Elsevier. (d) Design Framework of DMSCs. Reproduced from Ref. [169]. Copyright 2019, American Chemical Society. (e) Seven features that are most related to the catalytic performance of DMSCs in the order of importance, namely, the electron affinity of two metal atoms (EA_1 and EA_2), the sum of the vdW radius of two metal atoms ($R_1 + R_2$), the difference in the Pauling electronegativity between two metal atoms ($|P_1 - P_2|$), the product of ionization energy of TM_1 (IE_1) and the distance between two metal atoms ($IE_1 \times L$), the sum of Pauling electronegativity of two metal atoms ($P_1 + P_2$), and the average distance between TM_1/TM_2 atoms and the surrounding N atoms ($\frac{d_1+d_2+d_3+d_4+d_5+d_6}{6}$). Reproduced from Ref. [169]. Copyright 2019, American Chemical Society. (f) Comparison of DFT-computed ΔG_{OH}* values with those predicted by GBR algorithm. Reproduced from Ref. [169] Copyright 2019, American Chemical Society.

performing model. It was subsequently used as the basis for optimization via a genetic algorithm, leading to the prediction of an optimal catalyst layer thickness of 4.88 μm. The experimental results closely aligned with the AI predictions, especially in areas such as ionomer content, a crucial determinant of FC performance. The optimized MEA achieved remarkable results, with a maximum power density of 1.27 W/cm^2 and a current density of 1.27 A/cm^2 at 0.7 V, significantly outperforming commercial products with 3.2 times the regularized Pt utilization.

For the optimization of oxygen reduction reaction (ORR) catalyst, Zhu *et al.* [169] explored the activity origin and design principles for dual-metal-site catalysts (DMSCs) using a combination of DFT computations and ML (Fig. 9d). The study identified that the catalytic activity of DMSCs, crucial for the ORR in PEMFCs, is primarily determined by several fundamental factors, including the electron affinity, electronegativity, and radii of the embedded metal atoms (Fig. 9e). The researchers developed predictor equations using ML to quantitatively describe the ORR activity of DMSCs, revealing that several experimentally unexplored DMSCs could exhibit outstanding ORR activity, surpassing even platinum. The ML study utilized gradient boosted regression (GBR) for its high accuracy in training, with a root mean square error (RMSE) of 0.036 and coefficient of determination (R^2) values of 0.993 indicating a strong correlation with DFT-computed data (Fig. 9f). This approach allowed for efficient screening of highly active catalysts from a vast pool of candidates, significantly enhancing the process of material selection for ORR catalysts. The study's results have implications for the broader field of FC technology, providing a framework for designing and evaluating DMSCs with high ORR activity.

ML is also increasingly being utilized in the chemical reaction modeling of FCs. A typical ANN structure used for predicting cell voltage comprises three layers: an input layer (feeding in operating conditions such as current, temperature, pressure, flow rate, and humidity), hidden layers, and an output layer (outputting cell voltage). Various studies exemplify the application of this approach.

Su *et al.* [170] developed a three-layer ANN-based model to predict cell voltage uniformity in a 140-cell, 60 kW PEMFC stack, using a backpropagation algorithm for training (Fig. 10a). The model, exhibiting remarkable accuracy with less than 2 mV error in various conditions, identified improved voltage uniformity and performance with increased cathode pressure, and a decrease in performance at higher operating currents. This model, focusing on steady-state conditions, addresses a crucial research gap in predicting cell voltage distributions, challenging for physics-based models due to complex geometry and computational demands.

Li *et al.* [171] proposed a novel deep learning framework integrating Long Short-Term Memory (LSTM) and ANN to enhance the prediction accuracy of dynamic output voltage in PEMFC (Fig. 10b). This framework uniquely combines LSTM's ability to extract temporal information from past PEMFC states—determined using autocorrelation and partial autocorrelation functions—with the ANN's capability to learn from current system inputs like stack temperature, fluid temperature, and pressure. The fusion of LSTM and ANN outputs, leveraging both historical and current data, enables more accurate predictions of PEMFC's dynamic output voltage. Validated with data from a laboratory-scale PEMFC system, this model was compared against existing methods

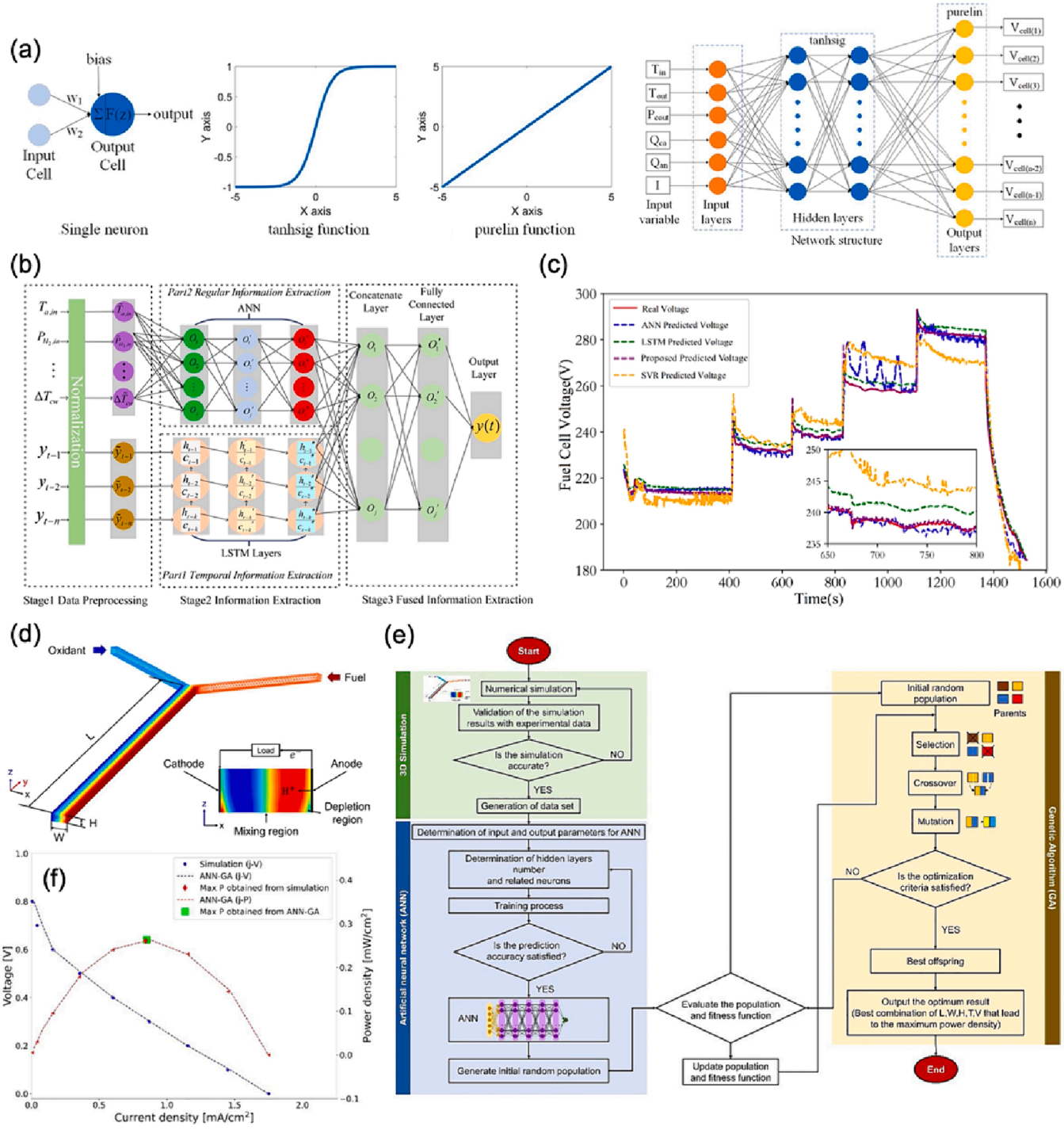

Fig. 10. (a) The ANN model structure. Reproduced from Ref. [170]. Copyright 2022, Elsevier. (b) Structure of the LSTM and ANN fusion model. Reproduced from Ref. [171]. Copyright 2023, IEEE. (c) Experiment validation: output voltages with LSTM and ANN fusion model and real values. Reproduced from Ref. [171]. Copyright 2023, IEEE. (d) Schematic image of the Y-shaped membraneless microfluidic FC operation. (L: Channel length; W: Channel width; H: Channel height). Reproduced from Ref. [174]. Copyright 2022, Elsevier. (e) Overall optimization flowchart. Reproduced from Ref. [174]. Copyright 2022, Elsevier. (f) Comparison of the polarization curves predicted by the ANN-GA and calculated by the 3D multiphysics simulation under the GA-indicated optimal operating condition that leads to the maximum power density (j: current density; V: cell voltage; P: power density). Reproduced from Ref. [174]. Copyright 2022, Elsevier.

like SVR, standalone ANN, and LSTM (Fig. 10c). The LSTM-ANN fusion model demonstrated superior performance, achieving the lowest mean square error of 1.303 in predicting the dynamic output voltage.

Sayed *et al.* [172] introduced a hybrid method that integrates ANN and a Forensic-Based Investigation algorithm (FBI) to optimize microalgae-based microbial FCs (MMFC). Focusing on the use of yeast (Saccharomyces cerevisiae) as a bioanode and microalgae as a cathodic biocatalyst in treating tofu wastewater, their approach involved a two-phase process. Initially, they developed an ANN model to accurately predict operational parameters of the MMFC, notably enhancing the coefficient-of-determination for both power density and COD removal efficiency beyond traditional ANOVA methods. This was followed by the application of the FBI method to optimize concentrations of yeast and wastewater, aiming to simultaneously maximize power density and COD

removal. Compared with an optimized performance by response surface methodology and measured data, the integration between ANN and FBI increased the performance of MMFC is by 2.24 %.

Le *et al.* [173] proposed an innovative approach using ANN to model and optimize the performance of direct urea-hydrogen peroxide FCs (DUHPFCs). They predicted the DUHPFC's voltage based on key design and operational parameters, including anode catalyst properties, urea concentration, KOH concentration, temperature, and feeding flow rate. The most effective ANN model utilized a two-hidden layer structure with a 7–10-6–1 topology, leveraging the Levenberg-Marquardt algorithm, logistic sigmoid function, and 80 % of the data for training. This model demonstrated excellent prediction capability, evidenced by a mean squared error (MSE) of 0.51×10^{-4} and an R-value of 0.9993. Furthermore, to identify the optimal process parameters for the DUHPFCs, bio-inspired algorithms (BIAs) such as Particle Swarm Optimization (PSO) and GA were applied. Both algorithms yielded similar optimal results, though the ANN-PSO combination was faster than ANN-GA in performance. Under optimal conditions identified by these methods—using $Ni_{0.2}Co_{0.8}$/Ni-foam as the anode catalyst, urea concentration of 1.4 M, KOH concentration of 6.2 M, temperature of 70 °C, and a flow rate of 5.9 mL/min—the calculated maximum power density reached 45.6 mW/cm^2.

Nguyen et al. [174] developed a novel approach to optimize membraneless microfluidic FC (MMFC) performance (Fig. 10d), combining an ANN with a Genetic Algorithm (GA) (Fig. 10e). They utilized a three-dimensional multiphysics model ($R^2 = 0.976$) to generate training data for the ANN, which accurately simulated the MMFC performance ($R^2 = 0.999$) with significantly reduced computation time (0.041 s). The ANN-GA model identified optimal MMFC design and operational parameters, achieving a maximum power density of 0.263 mW/cm^2. The ANN-GA and 3D multiphysics simulations show a strong concordance in predicting power density and current density under optimal operating conditions (Fig. 10f).

4.2.2. Durability prediction

This section of the review addresses the use of ML methods in predicting the durability of FCs, with a particular emphasis on the SoH and Remaining Useful Life (RUL). ML methods offer significant benefits in monitoring the SoH of FCs, thereby improving performance stability and extending service life. The fast response and robustness of ML methods in predicting the RUL are particularly advantageous, especially under varying FC system loads which can lead to operational challenges such as thermoelectric oscillations, insufficient fuel, carbon deposition in reformers, and system temperature overruns, which is due to the discrepancy in response times between the electrical characteristics of the FC (milliseconds) and the slower responses of gas delivery (seconds) and temperature (minutes), creating a significant lag between gas delivery and thermal response [175].

Based on FC degradation, various durability tests and prognostics models are proposed. Generally, the present prognostics approaches can be concluded as model-based, data-driven, and hybrid methods. The model-based method is based on physical models to study the aging prediction of FCs. Shen et al. [176] employed a model-predictive-based fuzzy control method to solve a nonlinear optimization problem aimed at improving fuel economy, maintaining battery charge, and reducing FC degradation. The novel approach integrates a SoH estimator and an energy storage system scheduler to optimize performance amidst uncertainties. By utilizing a Takagi-Sugeno (T-S) fuzzy model, the system effectively addresses nonlinear powertrain challenges and concurrently considers battery sustenance, FC durability, and fuel economy. Bressel et al. [177] introduced an aging tolerant control strategy and prognostics for PEMFC using a time-varying model structured in the Energetic Macroscopic Representation (EMR) formalism (Fig. 11). This strategy generates load current references considering SoH and calculates gas input flows through an inverted energetic model. A Maximum Power Point Tracking algorithm identifies the maximum power deliverable and forecasts RUL. The research validates a Degradation Tolerant Control (DTC) method that adjusts inputs based on SoH to reduce aging effects, moving beyond traditional SoH thresholds to a power-based end-of-life indicator.

The data-driven method overcomes the drawbacks of the physical model by learning directly from the historical experimental data to make the prediction [70]. The traditional neural network-based approach as well as newly emerging approaches like ML and deep learning are some typical frameworks for the data-driven prediction method. Nagulapati et al. [178] monitored the SoH of PEMFCs using data-driven ML techniques. The study utilized a dataset from 1000 h of PEMFC operation under dynamic load conditions, the research analyzed the impact of training data volume on the prediction accuracies of three ML models: Gaussian Process Regression (GPR), SVM, and ANN. The models were assessed using indicators such as cell voltages, current density distribution, and impedance spectra. The dataset was divided to represent various degradation stages with training and testing ratios of 30:70, 50:50, and 70:30. The results indicated that GPR excelled in predicting nonlinear time series data with a lower RMSE, particularly in the early and middle stages of degradation. SVM showed robust performance in smaller, more complex datasets, while ANN demonstrated high accuracy and convergence, particularly with larger training sets reflecting advanced degradation. Mao and Jackson also compared various traditional ML approaches (ANN, ANFIS, PF) for SoH prediction, with ANN showing the best prediction effect despite high computational requirements. ANFIS provided accurate predictions with lower computational costs, and PF was suitable for complex situations like FC faults. Zuo et al. [179] introduced an attention-enhanced Recurrent Neural Network (RNN) model, combining it with LSTM and Gated Recurrent Unit (GRU) to improve the durability prognostics of PEMFC in PHM systems (Fig. 12). This model notably increased prediction accuracy for PEMFC output voltage degradation, achieving lower root mean square errors than conventional LSTM and GRU models.

Deep Reinforcement Learning (DRL) is distinct from traditional methods like SVM and ANN due to its approach to long-term operational strategies. DRL is capable of dynamically managing power output and system efficiency by learning from interactions with the FC system environment [180]. While SVM and ANN rely on predefined data and may have limitations in adapting to variable conditions, DRL's continuous learning process allows it to adjust to changing scenarios, offering a different approach to optimizing FC performance and longevity. Zhang et al. [181] proposed a data-driven DRL-based energy management strategy (EMS) to optimize the performance and lifespan of both lithium-ion batteries (LIB) and PEMFC in FC vehicles (FCVs). This strategy is unique in its synchronous consideration of the SoH and thermal effects on the LIB and PEMFC stack under real-world driving scenarios. Employing the Soft Actor-Critic (SAC) algorithm for its stability and efficiency in continuous action spaces, the proposed EMS dynamically allocates power to maximize a multi-objective reward, effectively balancing hydrogen consumption rate and minimizing transient costs related to SoH descent and temperature penalties. The EMS was trained and validated using long-term driving data from Chongqing city, China. Results showcased the proposed SAC-based EMS's superiority, enhancing overall SoH of the powertrain by up to 14.01 % compared to other DRL approaches, while effectively regulating maximum temperatures. This not only addresses the common shortcomings in value-based RL and deterministic policy-based DRL methods but also prioritizes the durability and thermal stability of the powertrain system. In parallel, Chen et al. [182] also proposed a DRL-based EMS for FC vehicles but emphasizing the health degradation of both the FC system and the power battery. They improved the standard SAC algorithm by introducing a Beta policy to solve the high bias issue caused by the inconsistency between the infinite support of stochastic policy and the bounded physics constraints (Fig. 13). This study aimed to find a balance between driving cost and charge margin, considering health constraints as a crucial part of the strategy. The EMS showed a

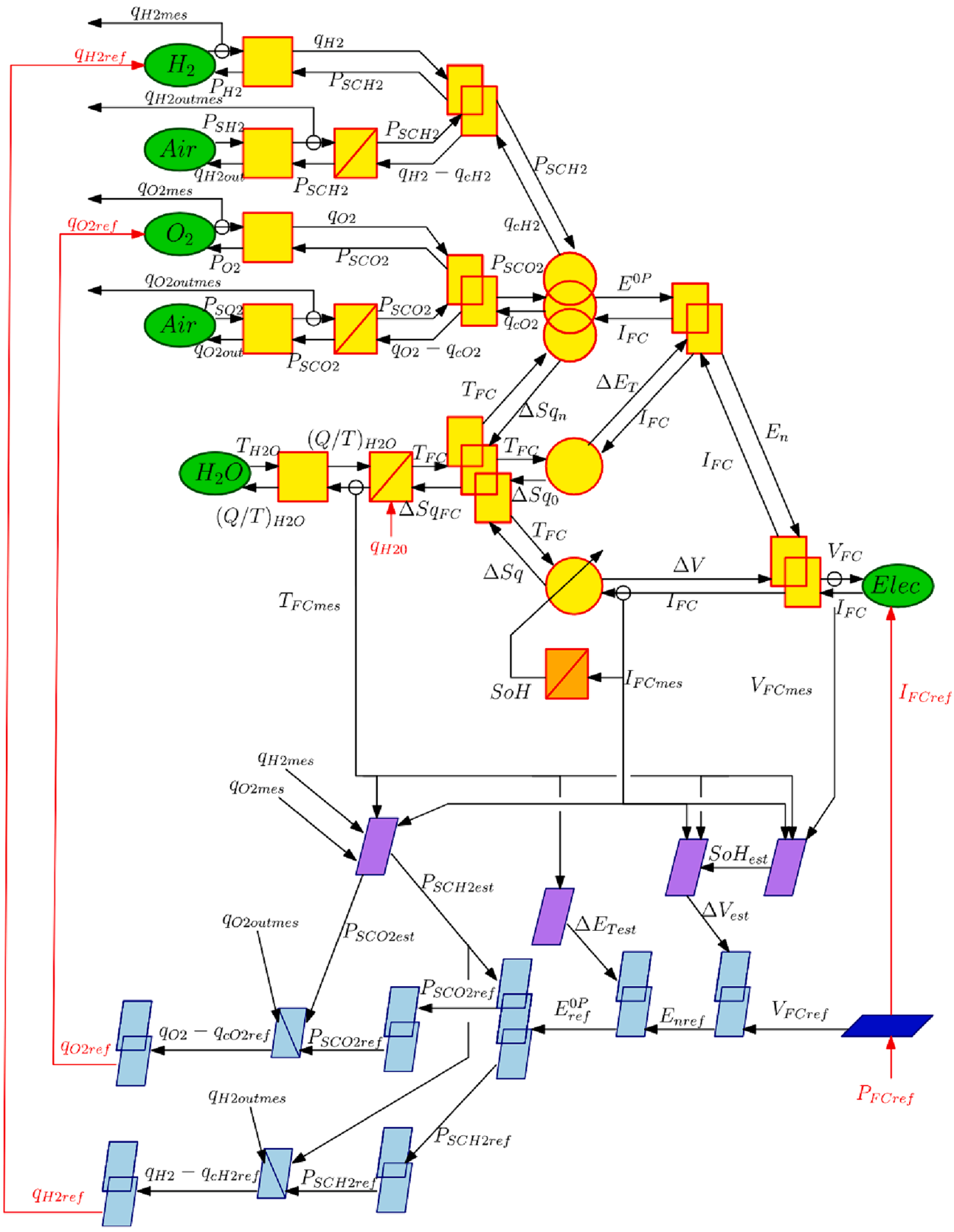

Fig. 11. Time varying parameters EMR model of PEMFC and associated control. Reproduced from Ref. [177]. Copyright 2020, Elsevier.

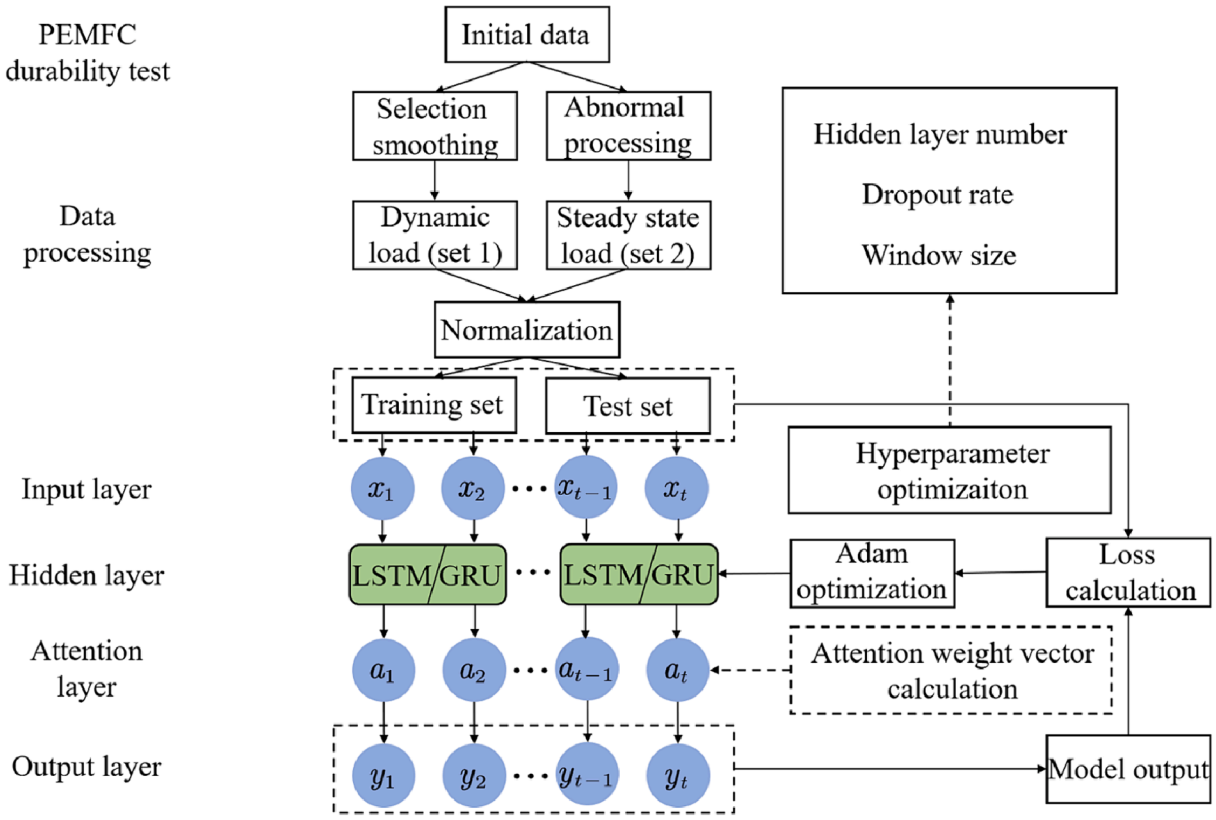

Fig. 12. Diagrams of the attention-based LSTM and GRU RNN model. Reproduced from Ref. [179]. Copyright 2021, Elsevier.

performance gap of 5.12 % with dynamic programming-based EMS regarding money cost but was 4.72 % better concerning equivalent hydrogen consumption.

Song *et al.* [183] developed a DRL-based EMS for FC hybrid electric vehicles (FCHEVs) using a fully-continuous Deep Deterministic Policy Gradient (DDPG) algorithm (Fig. 14). This novel approach optimizes real-time power distribution while addressing thermal and aging behaviors of the hybrid power sources, marking a first in enhancing FCHEV safety and lifespan. The strategy effectively reduced temperature spikes in lithium-ion batteries and hybrid system degradation, leading to improved thermal safety, operating performance, and economy. Validated under typical road conditions, the DDPG-based EMS outperformed traditional methods by focusing not only on fuel economy and charge balance but also on mitigating the detrimental effects of power demands and frequent starts-stops on the power supply system.

Jia *et al.* [184] developed a health-aware DRL EMS for FC hybrid electric buses (FCHEB), addressing the limitations of traditional DRL algorithms such as overestimation and poor training stability. Recognizing that DRL-based EMSs often fail to generalize beyond specific optimization objectives, the researchers utilized the Twin Delayed Deep Deterministic (TD3) policy gradient algorithm, combined with high-quality learning experience (HQLE) obtained from actual city bus driving cycles and an advanced model predictive control (MPC) strategy (Fig. 15). This approach was aimed at improving the optimization capability of the strategy and ensuring the health of the vehicular energy systems. The proposed TD3-HQLE-based EMS significantly improved training efficiency by 61.85 % and fuel economy by 7.45 %, while also extending the FC's life by 4 % and the battery's life by 19.4 % compared to the conventional TD3-based EMS. These results highlight the efficacy of the health-aware DRL approach in addressing the inherent drawbacks of DRL algorithms, enhancing fuel economy, and monitoring and extending the SoH of the vehicle's energy systems.

Fault diagnostics in FCs, essential for balancing system lifespan, manufacturing costs, and performance, categorize faults into internal (e.g., electrode degradation, catalyst poisoning) and external (e.g., failures in fuel, air supply, and heating units). ML-based fault diagnosis, using historical data or knowledge for training classifiers or expert systems, has become more prevalent due to the complexity and computational intensity of model-based methods. For example, Zheng *et al.* [185] developed a data-driven online fault diagnosis method for solid oxide FC systems, focusing on air leakage and fuel starvation using PCA and SVM. Tian *et al.* [186] introduced a data-driven method for diagnosing hydrogen leakage in FC vehicles, overcoming limitations of traditional hydrogen concentration sensors (HCSs). Using Fisher discrimination analysis, linear least squares for preprocessing, and relevance vector machine for pattern recognition, their method achieved over 95 % accuracy in diagnosing leaks larger than 2 g/s and consistently outperformed conventional HCS-based methods, especially when leaks were distant from sensors. Xing *et al.* [187] developed a data-driven fault diagnosis approach for PEMFCs using sensor pre-selection and ANN. By analyzing sensor data in time and frequency domains, they filtered out less responsive sensors and used the remaining data to build an ANN model, trained with various algorithms. The method achieved 99.2 % diagnostic accuracy and 98.3 % recall.

FC fault diagnosis research has seen the application of traditional ML methods like SVM, ANNs, and Bayesian networks, which perform well even with a small amount of training data. In contrast, DL methods, which have become increasingly popular, require a large volume of training data. Zhang *et al.* [188] tackled simultaneous fault diagnosis in solid oxide FC systems using a stacked sparse autoencoder (SSAE), equipped with K-binary classifiers. This method efficiently extracts crucial features from raw system data, significantly enhancing fault diagnosis by utilizing both labeled and unlabeled samples. Zuo et al. [189] developed a deep learning-based method to diagnose flooding faults in PEMFC under varying load conditions. They used indicators based on water transport mechanisms and visualized time-series data as

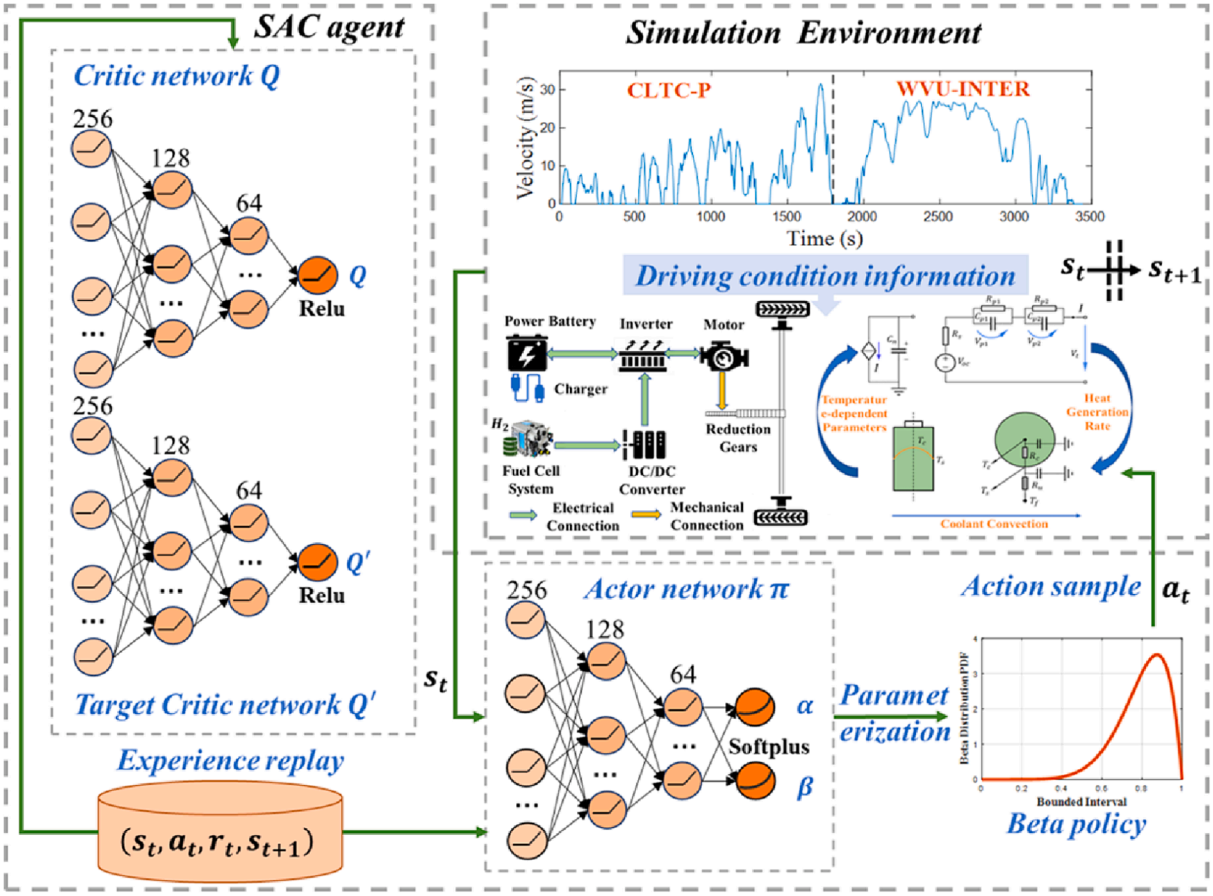

Fig. 13. The overall architecture of the proposed SAC-EMS. Reproduced from Ref. [182]. Copyright 2023, Elsevier.

2-D graphs for analysis with a CNN enhanced by batch normalization for better generalization. Validated with real PEMFC data, their model effectively identified flooding faults with over 99 % accuracy. Zhang and Guo [190] introduced BPNN-InceptionNet, a fault diagnosis method for PEMFC systems in FC trams, utilizing information fusion and deep learning (Fig. 16). The method employs a back propagation neural network (BPNN) to extract features from data and then uses an InceptionNet-based CNN for fault classification. Experiments revealed a kappa coefficient of 0.9884 for BPNN-InceptionNet, outperforming traditional BPNN, BPNN-VGG, and SVM classifiers. Yuan et al. [191] developed a novel fault diagnosis framework for PEMFCs using a hybrid deep learning network combining Residual Network (ResNet) and LSTM. This approach effectively diagnoses 25 types of faults, including membrane drying and flooding, with an impressive 99.632 % accuracy, outperforming conventional models.

In summary, whether using traditional ML or DL methods, fault diagnosis of FCs typically requires extensive data collection for training. Current fault diagnosis mainly focuses on external and simple internal faults, with limited research on complex internal faults like catalyst poisoning. Combining data-driven ML models with fault mechanism/simulation models could enhance interpretability and improve prediction performance, offering a promising avenue for future research in FC diagnostics.

5. AI for supercapacitor

Supercapacitors (SCs), or ultracapacitors, which can store up to 100 times more energy per unit volume/mass compared to conventional electrolytic capacitors, are one of the most advanced electrical energy storage technologies for the past few years [192,193]. Typical applications of SCs include automotive fields, public transportation, wearables, telecommunications, etc. Meanwhile, with the rapid evolving AI developments, computational-driven material science is able to make significant progress, providing guidelines and interpretations to experiments and giving further accurate predictions for either new materials and extraordinary properties [194–196]. This synergy, when coming to the realm of supercapacitors between AI and material science, heralds a new era of innovation, where the intelligent analysis and processing of vast datasets can identify potential materials and structural designs that could revolutionize supercapacitor technology. This multidisciplinary approach will not only accelerate the development of high-performance SCs but also enhances our understanding of the underlying electrochemical processes, setting the stage for the next generation of energy storage solutions.

5.1. Performance predictions and optimizations of general SCs

A significant kind of electrochemical charge storage mechanism is electrostatic double-layer capacitors (EDLCs), which store energy through the electrostatic attraction of ions on the surface involving the formation of a Helmholtz double-layer charges with certain separations at the interface between the electrode surface and the electrolyte, rather than through chemical reactions, as is the case similar to batteries but with extremely higher power density in the order of 15 kW kg^{-1} [197,198]. Based on these treasurable device foundations, characterizing their capacitance becomes a necessity, not only for choosing the most appropriate electrode materials, but also to help with the optimization of EDLC design.

Mediocre modeling of EDLCs is restricted on many aspects, such as limitations on just few electrolytes and electrodes, incompleteness of traditional theories [199], manifest heavier computational cost when going for higher accuracy or larger system size [200,201]. Conversely,

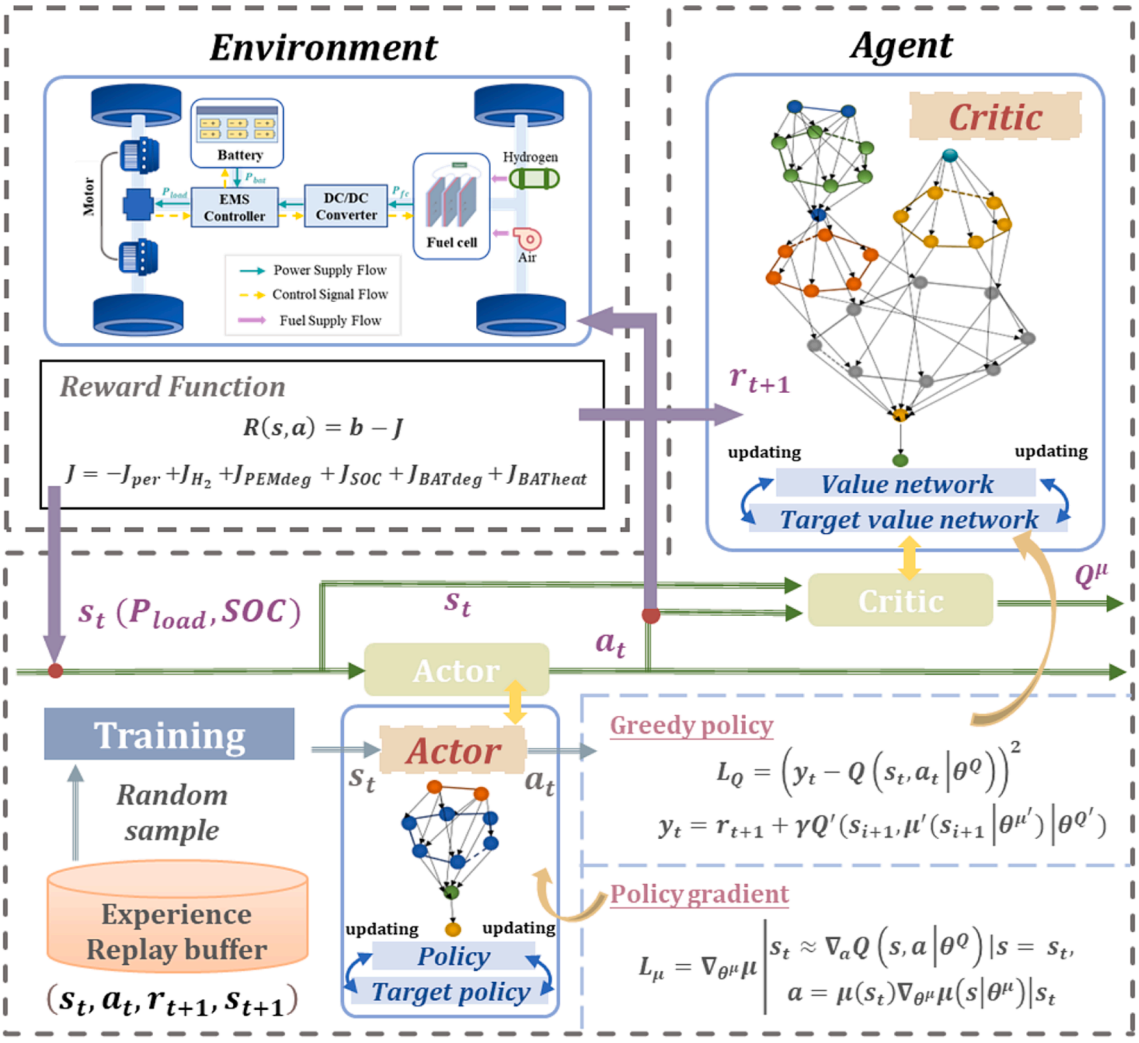

Fig. 14. The schematic frameworks of DDPG-based control system. Reproduced from Ref. [183]. Copyright 2023, IEEE.

the current widespread availability of open-source literature has transformed AI into a potent tool for managing experimental data pertaining to EDLCs, enabling the execution of comprehensive ML analyses across various EDLC models, and facilitating the development of a holistic perspective on optimized EDLCs. Enormous research has been conducted on using data-driven ML methods, for instance, SVM, neural networks (NNs), to predict properties of batteries [202,203]. Specifically, for general EDLCs, Nanda *et al.* [204] build datasets from 400 + published literatures including essential features of electrolytes, electrodes, and device that could potentially affect capacity retention, then use ML algorithms like RF and Random Subspace (RS) for prediction models for accomplishing attribute prioritization to determine highest-ranked attributes. See Fig. 17(a)–(b) for details. It turns out that carbon/polymer, type of electrodes, and testing current density are top ones, which is consistent with the theoretical validations. Furthermore, Su *et al.* [205] dive into the importance of solvent effects for capacity performance by combining ML approaches and experimental data and select the size of solvent molecule (d_{or}, d_{mv}) and the dielectric constant (ε_r) as variables with most contributions according to the best ML model, further improving the conventional Helmholtz model [206] by adding more terms to give clues to classical density functional theory simulations as shown in Fig. 17(c)–(e). Other than these basic ML routes, more advanced technologies are being applied to investigate EDLCs. Li *et al.* [207] and Zhou *et al.* [208] employ extreme learning machine (ELM) model and hybrid genetic algorithm (HGA) in view of the introductory workflow as Fig. 17 (f) for the prediction of EDLC life cycles with lower MSE, less estimation time, and insufficient data, indicating strong AI-based approaches for real-time forecasting of supercapacitor aging in industrial settings, significantly outperforming existing models and potentially revolutionizing the prediction of the remaining useful life for both supercapacitors and secondary ion batteries with its precision and robustness.

In addition to some specific physical properties, the general potential for electrochemical energy storage in SCs [209], such as charge/voltage relation, can be predicted via the above-mentioned ML methods, for example, SVM and NNs from *Jha et al.* [210], SVR and RF from Shariq *et al.* [211], extreme gradient boosting (XGBoost) from Liu *et al.* [212]. All of these works show great alignment comparing with experiments.

5.2. Carbon-based SCs descriptions

Known as the most common electrode materials, carbon is definitely crucial to SCs given the advantages of controllable porosity, high thermal stability, and lower cost [213]. Nevertheless, conventional carbon materials are way too far to fulfill the requirements of SC device, so researchers start to focus on the carbon nanomaterials [214,215], for example, activated carbon (AC), graphene, carbon nanotubes, three-dimensional carbon (3DC) to look for high-performance SCs with certain designing mechanisms and fabrications [216]. As we delve deeper into the potential of these nanomaterials, it becomes clearer that

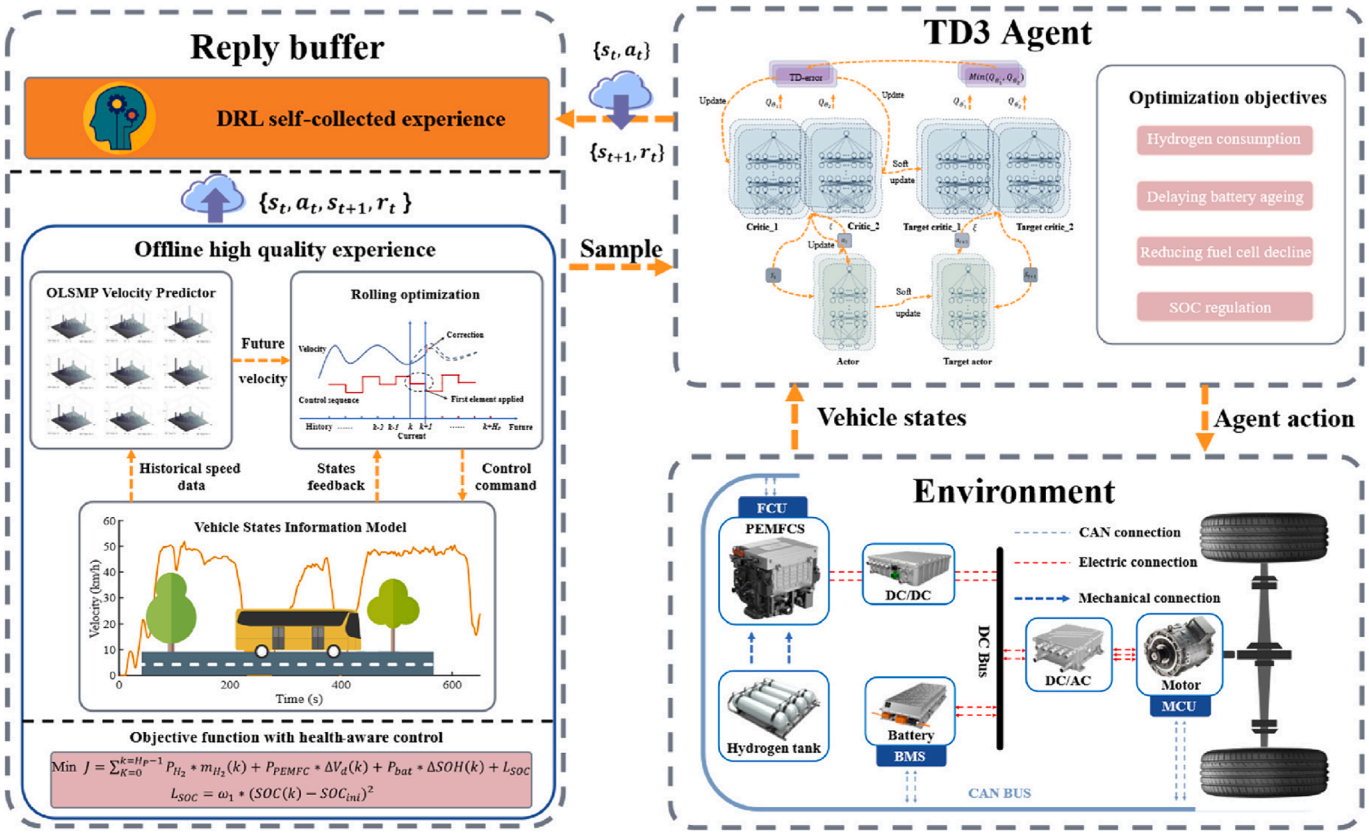

Fig. 15. The framework of TD3-HQLE-based EMS. Reproduced from Ref. [184]. Copyright 2023, Elsevier.

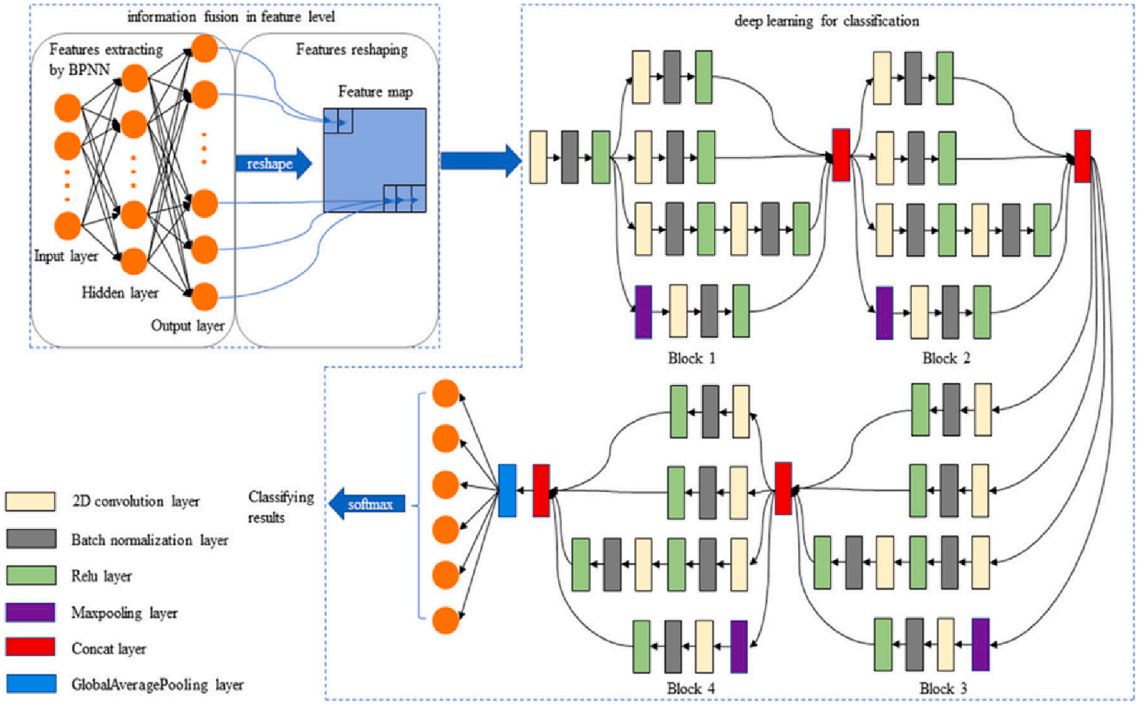

Fig. 16. The proposed BPNN-InceptionNet framework. Reproduced from Ref. [190]. Copyright 2021, Elsevier.

the integration of AI stands as the next logical step. The advanced analytics and predictive modeling further enhance the development of carbon-based nanomaterials, offering unprecedented precision in the design and optimization of the next generation of supercapacitors, thereby bridging the gap between current limitations and future possibilities.

First, for general carbon materials, Zhu *et al.* [217] and other researchers [218–220] apply various models like artificial NN and extreme

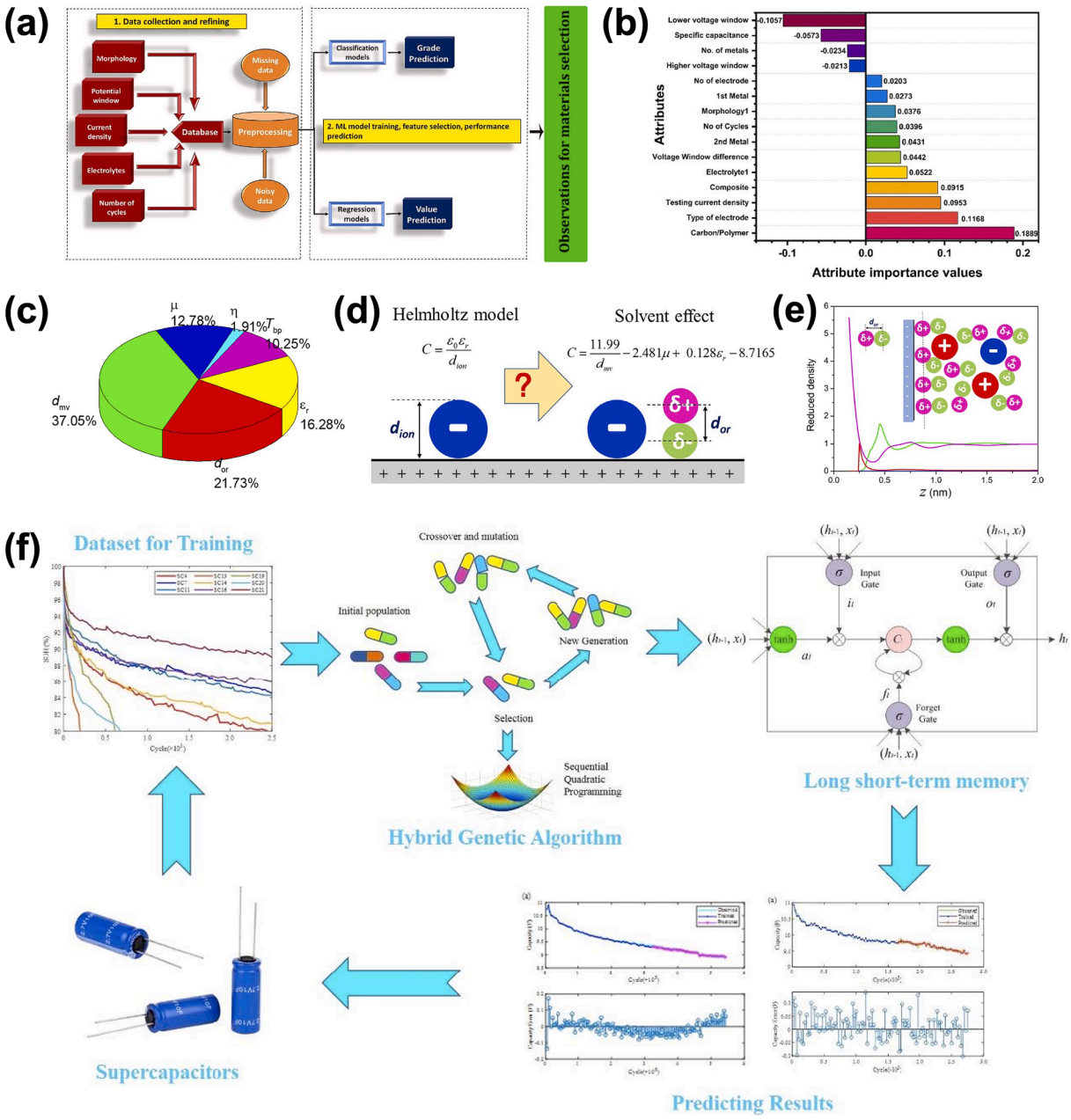

Fig. 17. (a) Schematic overview of database building and ML modeling of EDLC features aiming for supercapacitor performances in terms of cyclic stability. (b) Attribute prioritization from the database in (a). The attribute importance values have been mentioned against each attribute in the graph with analyzed importance of each feature. Reproduced from Ref. [204]. Copyright 2022, Elsevier. (c) The relative contribution of solvent variables to the predictive EDLC capacitance of the multilayer perceptions model. (d) Comparison between the traditional Helmholtz model and the model from the ML model. (e) Calculated EDLC structure of the ionic liquid electrolytes with solvents from classical density functional theory. Reproduced from Ref. [205]. Copyright 2019, Elsevier. (f) Workflow of sequential quadratic programming based on the combination of a recurrent neural network method called long short-term memory and HGA for predicting supercapacitors charging at steady state mode and estimating their remaining useful life. Reproduced from Ref. [208]. Copyright 2021, Elsevier.

learning machine to predict the capacitance of different carbon-based SCs, which takes into account of features like specific surface area (SSA), I_D/I_G ratio, pore volume (PV), pore size (PS), and so on, illustrated in Fig. 18(a-b). Furthermore, from the ML technical perspective, Ahmed *et al.* [221] utilize hyperparameter tuning to refine the efficiency of normal ML algorithms like RF and RS to better predict the various in-operando performance metrics of carbon-based supercapacitors. More specifically on the material side, porous carbon-active materials notably boost the specific capacitance of carbon-based SCs, which can be successfully predicted by ML methods [222,223]. Rahimi *et al.* [224] develop a multilayer perceptron (MLP)-NN model to investigate in-depth relations between the physio-chemical and operational properties of N/O co-doped, or named as heteroatom-doped [225], ACs and AC-based electrodes of EDLCs with the textural parameters of microstructural properties including SSA and PV, operational properties, and N/O-functional groups. Taking advantage of sensitivity analysis, they report that the performance is most significantly influenced by micropore surface area and micropore volume. Functional groups containing nitrogen and oxygen, such as carboxyl (OIII), pyridinic (N-6), pyrrolic (N-5), and various oxidized nitrogen configurations (N-X), contribute substantially to the electrochemical performance of EDLCs. See Fig. 18(c) for the best design. Similarly, Mishra *et al.* [226] and Reddy *et al.* [227] focus on the physiochemical properties using regression and NN models, extraordinarily highlighting the presence of

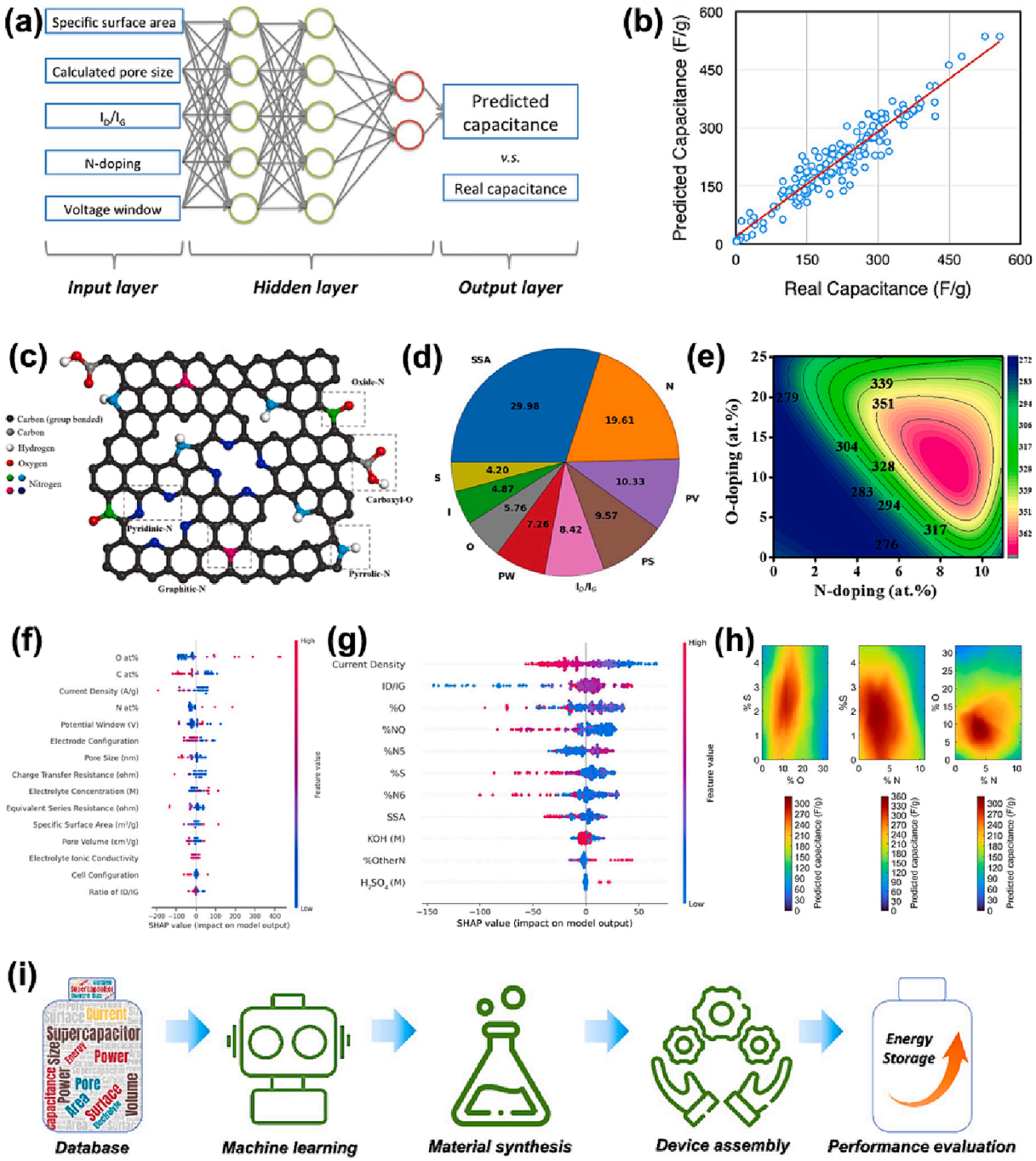

Fig. 18. (a) Illustration of artificial NN model. (b) Comparison of the predicted capacitance and the real data in published papers. Reproduced from Ref. [217]. Copyright 2018, Elsevier. (c) Schematic surface of the optimal N/O dual-doped AC-based electrode (N = 5.86 at%, Oxygen = 11.27 at%, N-6 = 1.36 at%, N-5 = 2 at %, N-Q = 1.94 at%, NX = 0.83 at%, OIII = 7.91 at%, SSA = 2400 m^2g^{-1}, S_{mic} = 1400 m^2g^{-1}, V_t = 1.61 cm^3g^{-1} and V_{mic} = 1.38 cm^3g^{-1}) predicted by generic algorithm-MLP. The best specific capacitance obtained 550F/g at 1 Ag^{-1} in 6 M KOH with potential window of 0.8 V. Reproduced from Ref. [224] Copyright 2022, Elsevier. (d) Feature analysis for the subset of data having 6 M KOH electrolyte. Reproduced from Ref. [226]. Copyright 2023, Nature Portfolio. (e) Contour plots of the specific capacitance (F/g) of supercapacitors about N-doping and O-doping. Reproduced from Ref. [227]. Copyright 2023, Elsevier. (f) SHAP values summary diagram illustrating the impact of features on the ANN model. Reproduced from Ref. [228]. Copyright 2023, Elsevier. (g) SHAP variable importance plot on full dataset (The features are ranked vertically based on its importance (the sum of the absolute SHAP values) with the highest on the top. The horizontal location of each dot shows how much they contribute to the capacitance with the color referring to the feature value of a particular sample.). (h) The contour plot showing the capacitance of the co-doped graphene with combination of all nitrogen species. Left: oxygen and sulphur. Middle: nitrogen and sulphur. Right: nitrogen and oxygen. Reproduced from Ref. [229]. Copyright 2023, IOP. (i) Schematic diagram of data-driven designing high-performance supercapacitors process. Reproduced from Ref. [230]. Copyright 2023, Elsevier.

nitrogen doping and nitrogen/oxygen (at. %) as necessary descriptors for specific capacitance (F/g) in Fig. 18(d-e). In addition, for graphene-based ones, Saad *et al.* [228] examine though artificial NN model and reveal the great effect of the atomic percentage of nitrogen and oxygen doped graphene, and Chenwittayakhachon *et al.* [229] further demonstrate the synergistic effect arising from co-doping nitrogen, sulphur, and locate the optimized region for different co-doping elemental groups in Fig. 18(h) with high capacitance and high retention rate. Both are

judged by the SHAP values which assign importance values for each feature presented in Fig. 18(f-g). Last but not least, Wang *et al.* [230] construct a framework for data-driven design of SC device, starting from the synthesis of 3DC with nitrogen source and etching to the choices of electrolyte type and processing methods, summarized in Fig. 18(i). Jamaluddin *et al.* [231] test biomass carbon materials in SCs and the dominant factors are quite consistent with pure carbon-based ones. Lu *et al.* [232] propose a ML strategy using carbonized metal–organic frameworks (MOFs) and employ theoretical calculations for adsorption and binding energy and verifies the outstanding performance of specific MOFs.

5.3. Novel materials characterizations in SCs

With the increasing influence of AI and ML, explorations for new materials for device applications have been emerging these years. Coincidentally, demands for new electrode and electrolyte materials are rapidly growing as the advanced studies of various carbon-based SCs. Many review papers have covered on the functional designs of SCs leveraging data science, which greatly assists new materials development in this area [38,39]. Here, we present a thorough summary on novel materials used in current-stage SCs and the indispensable role of AI though this process.

Firstly, Lu *et al.* [233] focus on layered double hydroxides (LDHs) in SCs benefiting from their unique layered structures and rich REDOX sites, where ML models like extreme gradient boosting aid the discovery of LDHs with large basal spacing ($d_{spacing}$) for the enhancement of F/g, as displayed in Fig. 19(a), the predictions of new LDHs ($Co_{0.67}Fe_{0.33}[Fe(CN)_6]_{0.11}$)•$(OH)_2$ with the largest $d_{spacing}$ of 12.40 Å, increasing by 10.91 % compared to the maximum $d_{spacing}$ of $Mg_{0.67}Al_{0.33}[Fe(CN)_6]_{0.08}$•$(OH)_2$ with $d_{spacing}$ of 11.18 Å that has ever been reported. Additionally, Ghosh *et al.* [234] choose Cerium oxynitride as a SC electrode to perform the data-driven predictions based on typical SC features we have mentioned before to help with the effectivity, practicality, and feasibility of this kind of novel material for SC device applications. See Fig. 19(b) for successful validations on the experimental data of CeO_xN_y. Moreover, Jha *et al.* [235] predict specific capacitance variation and retention of plant-derived lignin-based SCs listing as SC1, 2, and 3 acronyms referring to different weight percentages of lignin/ NiWo4 nanoparticles/ polyvinylidene fluoride combinations using five ML methods for the most proper modeling (Fig. 19(c)), which shows great alignment with experiments in Fig. 19(d).

6. Challenges and perspectives

In the realm of AI assisted energy storage devices research, the quality of data is critical, directly impacting the performance of various models. Without sufficient data, models presented less accurate results and lost physical insights. To navigate these challenges, it is essential to enhance data quality in terms of precision, variety, and volume, and to

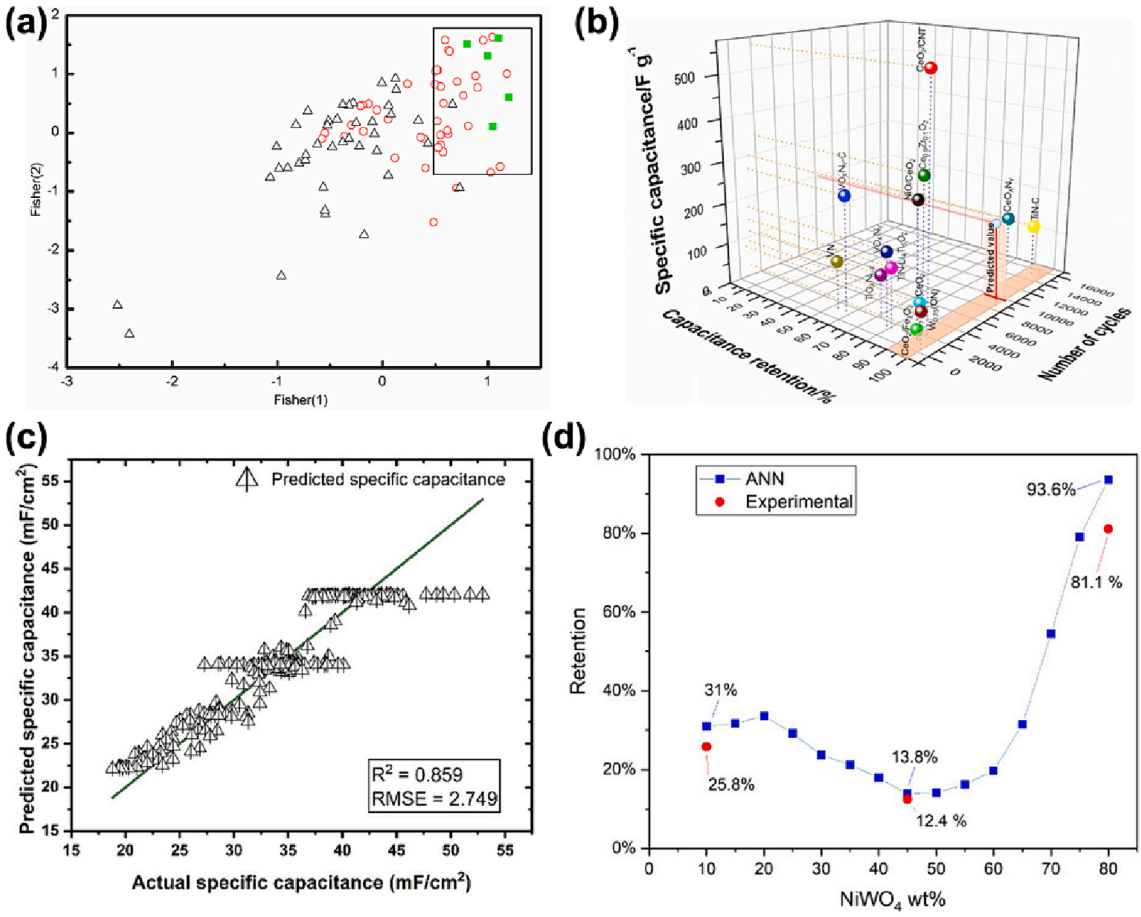

Fig. 19. (a) Materials pattern recognition of different types of samples by using fisher discriminant analysis. Reproduced from Ref. [233]. Copyright 2021, Elsevier. (b) Material mapping with the variables such as capacitance retention percentage (%), specific capacitance (F/g), and the number of cycles operated. The map shows cerium oxynitride's position on it with reference to the predicted value and other comparable materials present in literature. Reproduced from Ref. [234]. Copyright 2021, Elsevier. (c) Comparison of actual specific capacitance and test set predictions by artificial NN trained on a dataset of 1000 cycles of SC3 using a train/test split of 80/20. (d) Predicted supercapacitor retention after 600 cycles by artificial NN compared to experimental values. Reproduced from Ref. [235]. Copyright 2021, American Chemical Society.

optimize AI models through improved structures and learning strategies. AI and ML are instrumental in efficiently exploring chemical, formulation, and operational spaces, significantly reducing the need for exhaustive experimental and computational efforts. This efficiency is further enhanced by emerging techniques like hybrid algorithms and advanced learning strategies. These approaches could not only improve the research of electrochemical devices but also have broader applications in material simulation, supporting the development of efficient force fields and contributing to multiscale modeling.

Advanced ML techniques are crucial in energy storage R&D, as it could accelerate the research cycle and provide additional insights to traditional systems. These include identifying reaction mechanisms from electrochemical results and potential applications in tomography image segmentation for Non-Destructive Testing (NDT), development of computational and experimental databases, high-throughput experimentation, and text mining from literature. However, the journey of ML in electrochemical energy storage devices research is fraught with core challenges, including the selection of appropriate descriptors, data scarcity, error determination, and the establishment of standards and user-friendly tools. Furthermore, there is a need to bridge scales in modeling and foster collaboration between AI specialists and battery experts. As we move from current "weak AI" applications towards the potential of "strong AI," ethical considerations become increasingly important, particularly in ensuring effective human–machine collaboration and the robustness of algorithms, which is closely tied to product safety. Addressing these challenges and harnessing the full potential of AI and ML will be pivotal in advancing the field of energy storage research and development.

CRediT authorship contribution statement

Xinxin Liu: Writing – original draft, Visualization, Investigation, Formal analysis, Data curation, Conceptualization. **Kexin Fan:** Writing – original draft, Visualization, Validation, Formal analysis, Conceptualization. **Xinmeng Huang:** Writing – original draft, Visualization, Validation, Formal analysis, Data curation, Conceptualization. **Jiankai Ge:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Formal analysis, Conceptualization. **Yujie Liu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Formal analysis, Data curation. **Haisu Kang:** Writing – review & editing, Validation, Supervision, Formal analysis, Conceptualization.

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.

Data availability

No data was used for the research described in the article.

Acknowledgements

J.G. acknowledges the support of the Mavis Future Faculty Fellows program in the Grainger College of Engineering, University of Illinois, Urbana–Champaign.

References

[1] S. Bilgen, Structure and environmental impact of global energy consumption, Renew. Sustain. Energy Rev. 38 (2014) 890–902.
[2] I. Dincer, C. Acar, A review on clean energy solutions for better sustainability, Inter. J. Energy Res. 39 (5) (2015) 585–606.
[3] Z. Yang, J. Zhang, M.C. Kintner-Meyer, X. Lu, D. Choi, J.P. Lemmon, J. Liu, Electrochemical energy storage for green grid, Chem. Rev. 111 (5) (2011) 3577–3613.
[4] T.S. Mathis, N. Kurra, X. Wang, D. Pinto, P. Simon, Y. Gogotsi, Energy storage data reporting in perspective—guidelines for interpreting the performance of electrochemical energy storage systems, Adv. Energy Mater. 9 (39) (2019) 1902007.
[5] W. Lv, Z. Li, Y. Deng, Q.-H. Yang, F. Kang, Graphene-based materials for electrochemical energy storage devices: opportunities and challenges, Energy Storage Mater. 2 (2016) 107–138.
[6] J.B. Goodenough, Energy storage materials: a perspective, Energy Storage Mater. 1 (2015) 158–161.
[7] C.-Y. Wang, Fundamental models for fuel cell engineering, Chem. Rev. 104 (10) (2004) 4727–4766.
[8] C. Zhong, Y. Deng, W. Hu, J. Qiao, L. Zhang, J. Zhang, A review of electrolyte materials and compositions for electrochemical supercapacitors, Chem. Soc. Rev. 44 (21) (2015) 7484–7539.
[9] M. Winter, B. Barnett, K. Xu, Before Li ion batteries, Chem. Rev. 118 (23) (2018) 11433–11456.
[10] Y. Liang, Y.J.N.R.M. Yao, Designing modern aqueous batteries, Nat. Rev. Mater. 8 (2) (2023) 109–122.
[11] M.S. Whittingham, Lithium batteries and cathode materials, Chem. Rev. 104 (10) (2004) 4271–4302.
[12] G. Wang, L. Zhang, J. Zhang, A review of electrode materials for electrochemical supercapacitors, Chem. Soc. Rev. 41 (2) (2012) 797–828.
[13] X. Yu, T. Zhou, J. Ge, C. Wu, Recent advances on the modulation of electrocatalysts based on transition metal nitrides for the rechargeable Zn-air battery, ACS Mater. Lett. 2 (11) (2020) 1423–1434.
[14] Y.G. Guo, J.S. Hu, L.J. Wan, Nanostructured materials for electrochemical energy conversion and storage devices, Adv. Mater. 20 (15) (2008) 2878–2887.
[15] F.K. Kessler, Y. Zheng, D. Schwarz, C. Merschjann, W. Schnick, X. Wang, M. J. Bojdys, Functional carbon nitride materials—design strategies for electrochemical devices, Nat. Rev. Mater. 2 (6) (2017) 1–17.
[16] H. Yang, Q. He, Y. Liu, H. Li, H. Zhang, T. Zhai, On-chip electrocatalytic microdevice: an emerging platform for expanding the insight into electrochemical processes, Chem. Soc. Rev. 49 (10) (2020) 2916–2936.
[17] K. Sodeyama, Y. Igarashi, T. Nakayama, Y. Tateyama, M. Okada, Liquid electrolyte informatics using an exhaustive search with linear regression, Phys. Chem. Chem. Phys. 20 (35) (2018) 22585–22591.
[18] A.D. Sendek, Q. Yang, E.D. Cubuk, K.-A.-N. Duerloo, Y. Cui, E.J. Reed, Holistic computational structure screening of more than 12000 candidates for solid lithium-ion conductor materials, Energy Environ. Sci. 10 (1) (2017) 306–320.
[19] V. Sundararaghavan, N.J.C.M.S. Zabaras, Classification and reconstruction of three-dimensional microstructures using support vector machines, Comput. Mater. Sci. 32 (2) (2005) 223–239.
[20] M. Nakayama, K. Kanamori, K. Nakano, R. Jalem, I. Takeuchi, H. Yamasaki, Data-driven materials exploration for Li-ion conductive ceramics by exhaustive and informatics-aided computations, Chem. Rec. 19 (4) (2019) 771–778.
[21] R.K. Luu, M. Wysokowski, M.J. Buehler, Generative discovery of de novo chemical designs using diffusion modeling and transformer deep neural networks with application to deep eutectic solvents, Appl. Phys. Lett. (2023) 122 (23).
[22] H. Mami-Zadeh, R. Solgi, J.F. Carrier, H. Ghadiri, Material classification based on Dual-Energy Micro-CT images by the Gaussian mixture model, J. Instrum. 17 (02) (2022) P02001.
[23] R. Cohn, E. Holm, Unsupervised Machine Learning Via Transfer Learning and k-Means Clustering to Classify Materials Image Data, Integrating Materials and Manufacturing Innovation 10 (2) (2021) 231–244.
[24] R. Tibshirani, Regression Shrinkage and Selection via the Lasso, J. Roy. Stat. Soc.: Ser. B (Methodol.) 58 (1) (1996) 267–288.
[25] N.R. Draper, H. Smith, Applied Regression Analysis, Vol. 326, John Wiley & Sons: , New Jersey, 1998.
[26] S. Chatterjee, A.S. Hadi, Regression Analysis by Example, John Wiley & Sons, New Jersey, 2013.
[27] S.B. Kotsiantis, I. Zaharakis, P. Pintelas, Supervised machine learning: a review of classification techniques, IOS Press: Amsterdam 160 (2007) 3–24.
[28] D.W. Marquardt, R.D. Snee, Ridge regression in practice, Am. Stat. 29 (1) (1975) 3–20.
[29] E. Schulz, M. Speekenbrink, A. Krause, A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions, J. Math. Psychol. 85 (2018) 1–16.
[30] J. Durbin, Estimation of parameters in time-series regression models, J. r. Stat. Soc. Ser. B Methodol. 22 (1) (1960) 139–153.
[31] J.C. Stoltzfus, Logistic regression: a brief primer, Acad. Emerg. Med. 18 (10) (2011) 1099–1104.
[32] V. Svetnik, A. Liaw, C. Tong, J.C. Culberson, R.P. Sheridan, B.P. Feuston, Random forest: a classification and regression tool for compound classification and QSAR modeling, J. Chem. Info. Comp. Sci. 43 (6) (2003) 1947–1958.
[33] M.A. Hearst, S.T. Dumais, E. Osuna, J. Platt, B. Scholkopf, Support vector machines, IEEE Intell. Syst. 13 (4) (1998) 18–28.
[34] K.P. Murphy, Naive bayes classifiers, University of British Columbia, 2006.
[35] T.M. Kodinariya, P.R. Makwana, Review on determining number of Cluster in K-Means Clustering, Inter. J. 1 (6) (2013) 90–95.
[36] D.A. Reynolds, Gaussian mixture models, Encyclopedia of Biometrics 741 (2009) 659–663.
[37] T. Kohonen, Essentials of the self-organizing map, Neural Netw. 37 (2013) 52–65.
[38] M.E. Celebi, K. Aydin, Unsupervised Learning Algorithms, Vol. 9, Springer:, New York, 2016.

[39] G.R. Naik, D.K. Kumar, An overview of independent component analysis and its applications, Informatica 35 (1) (2011).
[40] Y. Wang, B. Seo, B. Wang, N. Zamel, K. Jiao, X.C. Adroher, Fundamentals, materials, and machine learning of polymer electrolyte membrane fuel cell technology, Energy and AI 1 (2020) 100014.
[41] C. Chen, Y. Zuo, W. Ye, X. Li, Z. Deng, S.P. Ong, A critical review of machine learning of energy materials, Adv. Energy Mater. 10 (8) (2020) 1903242.
[42] S.H. Samadi, B. Ghobadian, M. Nosrati, Prediction of higher heating value of biomass materials based on proximate analysis using gradient boosted regression trees method, Energy Sources, Part A: Recov., Utiliz., [Environ. Effects] 43 (6) (2021) 672–681.
[43] Y. Okamoto, Y. Kubo, Ab initio calculations of the redox potentials of additives for lithium-ion batteries and their prediction through machine learning, ACS Omega 3 (7) (2018) 7868–7874.
[44] M.A. Shandiz, R. Gauvin, Application of machine learning methods for the prediction of crystal system of cathode materials in lithium-ion batteries, Comput. Mater. Sci. 117 (2016) 270–278.
[45] G. Pilania, P.V. Balachandran, C. Kim, T. Lookman, Finding new perovskite halides via machine learning, Front. Mater. 3 (2016) 19.
[46] S.I. Gallant, Perceptron-based learning algorithms, IEEE Trans. Neural Netw. 1 (2) (1990) 179–191.
[47] S. Sharma, S. Sharma, A.J.T.D.S. Athaiya, Activation Functions in Neural Networks. 6 (12) (2017) 310–316.
[48] M.W. Gardner, S. Dorling, Artificial Neural Networks (the Multilayer Perceptron)—a Review of Applications in the Atmospheric Sciences. 32 (14–15) (1998) 2627–2636.
[49] F. Murtagh, Multilayer perceptrons for classification and regression, Neurocomputing 2 (5–6) (1991) 183–197.
[50] Albawi, S.; Mohammed, T. A.; Al-Zawi, S. In Understanding of a convolutional neural network, 2017 international conference on engineering and technology (ICET), Ieee: 2017; pp 1-6.
[51] Krizhevsky, A.; Sutskever, I.; Hinton, G. E. J. A. i. n. i. p. s., Imagenet classification with deep convolutional neural networks. Adv. Neural Infor. Processing Systems 2012, 25.
[52] Simonyan, K.; Zisserman, A., Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:.00234 2014.
[53] K. He, X. Zhang, S. Ren, J. Sun, In *Deep residual learning for image recognition*, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[54] Jain, A.; Zamir, A. R.; Savarese, S.; Saxena, A. In Structural-rnn: Deep learning on spatio-temporal graphs, Proceedings of the ieee conference on computer vision and pattern recognition, 2016; pp 5308-5317.
[55] Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y., Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:.00234 2014.
[56] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.
[57] I. Sutskever, O. Vinyals, Le Q V. J (2014) 27.
[58] M. Rupp, A. Tkatchenko, K.-R. Müller, O.A. Von Lilienfeld, Fast and accurate modeling of molecular atomization energies with machine learning, Phys. Rev. Lett. 108 (5) (2012) 058301.
[59] Jalem, R.; Kimura, M.; Nakayama, M.; Kasuga, T. J. J. o. c. i.; modeling, Informatics-aided density functional theory study on the Li ion transport of Tavorite-type LiMTO4F (M3+–T5+, M2+–T6+). J. Chem. Inf. Model. 2015, 55 (6), 1158-1168.
[60] M. Ziatdinov, O. Dyck, A. Maksov, X. Li, X. Sang, K. Xiao, R.R. Unocic, R. Vasudevan, S. Jesse, S.V. Kalinin, Deep learning of atomically resolved scanning transmission electron microscopy images: chemical identification and tracking local transformations, ACS Nano 11 (12) (2017) 12742–12752.
[61] K.T. Schütt, F. Arbabzadah, S. Chmiela, K.R. Müller, A. Tkatchenko, Quantum-chemical insights from deep tensor neural networks, Nat. Comm. 8 (1) (2017) 13890.
[62] L.P. Kaelbling, M.L. Littman, A.W. Moore, Reinforcement learning: A survey, J. Artif. Intell. Res. 4 (1996) 237–285.
[63] M.A. Wiering, M. Van Otterlo, Reinforcement learning, In Adaptation, Learning, Optimization, Springer: New York 12 (2012) 729.
[64] C.J. Watkins, P. Dayan, Q-learning, In Machine Learning, Springer: New York 8 (1992) 279–292.
[65] K. Arulkumaran, M.P. Deisenroth, M. Brundage, A.A. Bharath, Deep reinforcement learning: A brief survey, IEEE Signal Process Mag. 34 (6) (2017) 26–38.
[66] V. François-Lavet, P. Henderson, R. Islam, M.G. Bellemare, J.J.F. Pineau, An introduction to deep reinforcement learning. *Foundations Trends®*, Mach. Learn. 11 (3–4) (2018) 219–354.
[67] J. Fan, Z. Wang, Y. Xie, Z. Yang, In *A theoretical analysis of deep Q-learning*, PMLR, Learning for dynamics and control, 2020, pp. 486–489.
[68] S. Gu, T. Lillicrap, I. Sutskever, S. Levine, In *Continuous deep q-learning with model-based acceleration*, PMLR, International conference on machine learning, 2016, pp. 2829–2838.
[69] J. Liang, V. Makoviychuk, A. Handa, N. Chentanez, M. Macklin, D. Fox, In Gpu-accelerated robotic simulation for distributed reinforcement learning, Conference on Robot Learning, PMLR (2018) 270–282.
[70] C.-Y. Lee, B.-J. Chou, C.-F. Huang, Data science and reinforcement learning for price forecasting and raw material procurement in petrochemical industry, Adv. Eng. Inform. 51 (2022) 101443.
[71] P. Rajak, A. Krishnamoorthy, A. Mishra, R. Kalia, A. Nakano, P. Vashishta, Autonomous reinforcement learning agent for chemical vapor deposition synthesis of quantum materials, NPJ Comput. Mater. 7 (1) (2021) 108.
[72] F. Sui, R. Guo, Z. Zhang, G.X. Gu, L. Lin, Deep reinforcement learning for digital materials design, ACS Mater. Lett. 3 (10) (2021) 1433–1439.
[73] R.K. Vasudevan, E. Orozco, S.V. Kalinin, Discovering mechanisms for materials microstructure optimization via reinforcement learning of a generative model, Mach. Learn.: Sci. Technol. 3 (4) (2022) 04LT03.
[74] E. Nikolaev, N. Zakharova, V. Zakharov, In *Smart manufacturing control system based on deep reinforcement learning*, IOP Conference Series: Materials Science and Engineering, IOP Publishing (2021) 012061.
[75] Croitoru, F.-A.; Hondru, V.; Ionescu, R. T.; Shah, M., Diffusion models in vision: A survey. *IEEE Transactions on Pattern Analysis*.
[76] W. Feller, Diffusion processes in one dimension, Trans. Am. Math. Soc. 77 (1) (1954) 1–31.
[77] A. Hyvärinen, P. Dayan, Estimation of non-normalized statistical models by score matching, J. Mach. Learn. Res. 6 (4) (2005).
[78] L. Yang, Z. Zhang, Y. Song, S. Hong, R. Xu, Y. Zhao, W. Zhang, B. Cui, M.-H. Yang, Diffusion Models: A Comprehensive Survey of Methods and Applications. 56 (4) (2023) 1–39.
[79] J. Ho, A. Jain, P. Abbeel, Denoising diffusion probabilistic models, Adv. Neural Inf. Proces. Syst. 33 (2020) 6840–6851.
[80] Y. Song, S. Ermon, Generative modeling by estimating gradients of the data distribution, Adv. Neural Inf. Proces. Syst. 32 (2019).
[81] Y. Song, S. Ermon, Improved techniques for training score-based generative models, Adv. Neural Inf. Proces. Syst. 33 (2020) 12438–12448.
[82] Y. Song, C. Durkan, I. Murray, S. Ermon, Maximum likelihood training of score-based diffusion models, Adv. Neural Inf. Proces. Syst. 34 (2021) 1415–1428.
[83] P. Lyngby, K.S. Thygesen, Data-driven discovery of 2D materials by deep generative models, NPJ Comput. Mater. 8 (1) (2022) 232.
[84] M. Alverson, S. Baird, R. Murdock, J. Johnson, T.D. Sparks, Generative adversarial networks and diffusion models in material discovery, Digital Discovery (2024).
[85] K.-H. Lee, G.J. Yun, Microstructure reconstruction using diffusion-based generative models, Mech. Adv. Mater. Struct. (2023) 1–19.
[86] C. Duan, Y. Du, H. Jia, H.J. Kulik, Accurate transition state generation with an object-aware equivariant elementary reaction diffusion model, Nat. Comput. Sci. 3 (12) (2023) 1045–1055.
[87] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, Adv. Neural Inf. Proces. Syst. 30 (2017).
[88] OpenAI, Gpt-4 technical report. 2023, 2, 3.
[89] Kenton, J. D. M.-W. C.; Toutanova, L. K. In Bert: Pre-training of deep bidirectional transformers for language understanding, Proceedings of naacL-HLT, 2019; p 2.
[90] Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; Stoyanov, V. J. a. p. a., Roberta: A robustly optimized bert pretraining approach. **2019**.
[91] Dong, Q.; Li, L.; Dai, D.; Zheng, C.; Wu, Z.; Chang, B.; Sun, X.; Xu, J.; Sui, Z., A survey for in-context learning. arXiv preprint arXiv:.00234 2022.
[92] K.M. Jablonka, Q. Ai, A. Al-Feghali, S. Badhwar, J.D. Bocarsly, A.M. Bran, S. Bringuier, L.C. Brinson, K. Choudhary, D. Circi, 14 examples of how LLMs can transform materials science and chemistry: a reflection on a large language model hackathon, Digital Discovery 2 (5) (2023) 1233–1250.
[93] V. Korolev, P. Protsenko, Accurate, interpretable predictions of materials properties within transformer language models, Patterns 4 (10) (2023).
[94] E.A. Olivetti, J.M. Cole, E. Kim, O. Kononova, G. Ceder, T.-Y.-J. Han, A. M. Hiszpanski, Data-driven materials research enabled by natural language processing and information extraction, Appl. Phys. Rev. 7 (4) (2020).
[95] D. Flam-Shepherd, K. Zhu, A. Aspuru-Guzik, Language models can learn complex molecular distributions, Nat. Comm. 13 (1) (2022) 3293.
[96] X.L. Wang, R.J. Xiao, H. Li, L.Q. Chen, Quantitative structure-property relationship study of cathode volume changes in lithium ion batteries using ab-initio and partial least squares analysis, J Materiomics 3 (3) (2017) 178–183.
[97] R. Jalem, M. Nakayama, T. Kasuga, An efficient rule-based screening approach for discovering fast lithium ion conductors using density functional theory and artificial neural networks, J Mater Chem A 2 (3) (2014) 720–734.
[98] W. Li, J.E. Zhu, Y. Xia, M.B. Gorji, T. Wierzbicki, Data-Driven Safety Envelope of Lithium-Ion Batteries for Electric Vehicles, Joule 3 (11) (2019) 2703–2715.
[99] P. Fermín-Cueto, E. McTurk, M. Allerhand, E. Medina-Lopez, M.F. Anjos, J. Sylvester, G. dos Reis, Identification and machine learning prediction of knee-point and knee-onset in capacity degradation curves of lithium-ion cells, Energy and AI 1 (2020).
[100] K. Min, B. Choi, K. Park, E. Cho, Machine learning assisted optimization of electrochemical properties for Ni-rich cathode materials, Sci Rep-Uk 8 (2018).
[101] N. Kireeva, V.S. Pervov, Materials Informatics Screening of Li-Rich Layered Oxide Cathode Materials with Enhanced Characteristics Using Synthesis Data, Batteries Supercaps 3 (5) (2020) 427–438.
[102] R.P. Joshi, J. Eickholt, L. Li, M. Fornari, V. Barone, J.E. Peralta, Machine Learning the Voltage of Electrode Materials in Metal-Ion Batteries, ACS Appl Mater Interfaces 11 (20) (2019) 18494–18503.
[103] B.W. Deng, P.C. Zhong, K. Jun, J. Riebesell, K. Han, C.J. Bartel, G. Ceder, CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling, Nat Mach Intell 5 (9) (2023) 1031–1041.
[104] O. Allam, B.W. Cho, K.C. Kim, S.S. Jang, Application of DFT-based machine learning for developing molecular electrode materials in Li-ion batteries, Rsc Adv 8 (69) (2018) 39414–39420.

[105] S.V.S. Sowndarya, J.N. Law, C.E. Tripp, D. Duplyakin, E. Skordilis, D. Biagioni, R. S. Paton, P. St John, C., Multi-objective goal-directed optimization of de novo stable organic radicals for aqueous redox flow batteries, Nat Mach Intell 4 (8) (2022), 720-+.

[106] R. Jalem, M. Kimura, M. Nakayama, T. Kasuga, Informatics-Aided Density Functional Theory Study on the Li Ion Transport of Tavorite-Type LiMTO4F (M(3+)-T(5+), M(2+)-T(6+)), J Chem Inf Model 55 (6) (2015) 1158–1168.

[107] K. Fujimura, A. Seko, Y. Koyama, A. Kuwabara, I. Kishida, K. Shitara, C.A. J. Fisher, H. Moriwake, I. Tanaka, Accelerated Materials Design of Lithium Superionic Conductors Based on First-Principles Calculations and Machine Learning Algorithms, Adv. Energy Mater. 3 (8) (2013) 980–985.

[108] Y. Zhang, X.F. He, Z.Q. Chen, Q. Bai, A.M. Nolan, C.A. Roberts, D. Banerjee, T. Matsunaga, Y.F. Mo, C. Ling, Unsupervised discovery of solid-state lithium ion conductors, Nat Commun 10 (2019).

[109] A.D. Sendek, Q. Yang, E.D. Cubuk, K.A.N. Duerloo, Y. Cui, E.J. Reed, Holistic computational structure screening of more than 12 000 candidates for solid lithium-ion conductor materials, Energ Environ Sci 10 (1) (2017) 306–320.

[110] L. Zhang, Y. Liu, Y. You, A. Vinu, L. Mai, NASICONs-type solid-state electrolytes: The history, physicochemical properties, and challenges, Interdiscip Mater. 2 (1) (2023) 91–110.

[111] G. Xi, M. Xiao, S. Wang, D. Han, Y. Li, Y. Meng, Polymer-based solid electrolytes: material selection, design, and application, Adv. Funct. Mater. 31 (9) (2021) 2007598.

[112] Z. Xiao, T. Long, L. Song, Y. Zheng, C. Wang, Research progress of polymer-inorganic filler solid composite electrolyte for lithium-ion batteries, Ionics 28 (1) (2022) 15–26.

[113] N. Lv, Q. Zhang, Y. Xu, H. Li, Z. Wei, Z. Tao, Y. Wang, H. Tang, PEO-based composite solid electrolyte for lithium battery with enhanced interface structure, J. Alloys Compd. 938 (2023) 168675.

[114] T. Nakayama, Y. Igarashi, K. Sodeyama, M. Okada, Material search for Li-ion battery electrolytes through an exhaustive search with a Gaussian process, Chem Phys Lett 731 (2019).

[115] E. Heid, M. Fleck, P. Chatterjee, C. Schröder, A.D. MacKerell, Toward Prediction of Electrostatic Parameters for Force Fields That Explicitly Treat Electronic Polarization, J Chem Theory Comput 15 (4) (2019) 2460–2469.

[116] D. Bedrov, J.P. Piquemal, O. Borodin, A.D. MacKerell, B. Roux, C. Schröder, Molecular Dynamics Simulations of Ionic Liquids and Electrolytes Using Polarizable Force Fields, Chem Rev 119 (13) (2019) 7940–7995.

[117] T. Xie, A. France-Lanord, Y.M. Wang, Y. Shao-Horn, J.C. Grossman, Graph dynamical networks for unsupervised learning of atomic scale dynamics in materials, Nat Commun 10 (2019).

[118] K.T. Schütt, H.E. Sauceda, P.J. Kindermans, A. Tkatchenko, K.R. Müller, SchNet - A deep learning architecture for molecules and materials, J Chem Phys 148 (24) (2018).

[119] L.D. Ellis, S. Buteau, S.G. Hames, L.M. Thompson, D.S. Hall, J.R. Dahn, A New Method for Determining the Concentration of Electrolyte Components in Lithium-Ion Cells, Using Fourier Transform Infrared Spectroscopy and Machine Learning, J Electrochem Soc 165 (2) (2018) A256–A262.

[120] Y.T. Chen, M. Duquesnoy, D.H.S. Tan, J.M. Doux, H.D. Yang, G. Deysher, P. Ridley, A.A. Franco, Y.S. Meng, Z. Chen, Fabrication of High-Quality Thin Solid-State Electrolyte Films Assisted by Machine Learning, ACS Energy Lett 6 (4) (2021) 1639–1648.

[121] R.R. Richardson, C.R. Birkl, M.A. Osborne, D.A. Howey, Gaussian Process Regression for Capacity Estimation of Lithium-Ion Batteries, Ieee T Ind Inform 15 (1) (2019) 127–138.

[122] H. El-Bousiydy, T. Lombardo, E.N. Primo, M. Duquesnoy, M. Morcrette, P. Johansson, P. Simon, A. Grimaud, A.A. Franco, What Can Text Mining Tell Us About Lithium-Ion Battery Researchers' Habits? Batteries Supercaps 4 (5) (2021) 689.

[123] A. Van der Ven, Z. Deng, S. Banerjee, S.P. Ong, Rechargeable Alkali-Ion Battery Materials: Theory and Computation, Chem Rev 120 (14) (2020) 6977–7019.

[124] V.L. Deringer, Modelling and understanding battery materials with machine-learning-driven atomistic simulations, J Phys-Energy 2 (4) (2020).

[125] W.W. Li, Y. Ando, E. Minamitani, S. Watanabe, Study of Li atom diffusion in amorphous Li3PO4 with neural network potential, J Chem Phys 147 (21) (2017).

[126] Z. Deng, C. Chen, X.G. Li, S.P. Ong, An electrostatic spectral neighbor analysis potential for lithium nitride, Npj Comput Mater 5 (2019).

[127] Y.Q. Shao, M. Hellström, A. Yllö, J. Mindemark, K. Hermansson, J. Behler, C. Zhang, Temperature effects on the ionic conductivity in concentrated alkaline electrolyte solutions, Phys Chem Chem Phys 22 (19) (2020) 10426–10430.

[128] P.B. Jorgensen, A. Bhowmik, Equivariant graph neural networks for fast electron density estimation of molecules, liquids, and solids, Npj Comput Mater 8 (1) (2022).

[129] Bhowmik, P. B. J. a. A., DeepDFT: Neural Message Passing Network for Accurate Charge Density Prediction. *arXiv* **2020**.

[130] J.H. Chang, D. Kleiven, M. Melander, J. Akola, J.M. Garcia-Lastra, T. Vegge, CLEASE: a versatile and user-friendly implementation of cluster expansion method, J Phys-Condens Mat 31 (32) (2019).

[131] P.M. Tagade, S.P. Adiga, S. Pandian, M.S. Park, K.S. Hariharan, S.M. Kolake, Attribute driven inverse materials design using deep learning Bayesian framework, Npj Comput Mater 5 (2019).

[132] C.H. Wang, K. Aoyagi, P. Wisesa, T. Mueller, Lithium Ion Conduction in Cathode Coating Materials from On-the Fly Machine Learning, Chem Mater 32 (9) (2020) 3741–3752.

[133] Z. Ahmad, T. Xie, C. Maheshwari, J.C. Grossman, V. Viswanathan, Machine Learning Enabled Computational Screening of Inorganic Solid Electrolytes for Suppression of Dendrite Formation in Lithium Metal Anodes, ACS Cent Sci 4 (8) (2018) 996–1006.

[134] K. Hatakeyama-Sato, T. Tezuka, M. Umeki, K. Oyaizu, AI-Assisted Exploration of Superionic Glass-Type Li(+) Conductors with Aromatic Structures, J Am Chem Soc 142 (7) (2020) 3301–3305.

[135] K. Miwa, R. Asahi, Machine learning-enhanced molecular dynamics simulations for Nb-doped lithium garnet-type oxide Li7-xLa3(Zr2-xNbx)O12, Phys Rev Mater 2 (10) (2018).

[136] M.R. Johan, S.M.M. Yasin, S. Ibrahim, Bayesian Neural Networks Model for Ionic Conductivity of Nanocomposite Solid Polymer Electrolyte System (PEO - LiCF3SO3 - DBP - ZrO2), Int J Electrochem Sc 7 (1) (2012) 222–233.

[137] S. Ibrahim, M.R. Johan, Conductivity, Thermal and Neural Network Model Nanocomposite Solid Polymer Electrolyte System (PEO-LiPF-EC-CNT), Int J Electrochem Sc 6 (11) (2011) 5565–5587.

[138] Z.C. Xu, J. Wang, Q. Fan, P.D. Lund, J. Hong, Improving the state of charge estimation of reused lithium-ion batteries by abating hysteresis using machine learning technique, J Energy Storage 32 (2020).

[139] A.R. Natarajan, A. Van der Ven, Machine-learning the configurational energy of multicomponent crystalline solids, Npj Comput Mater 4 (2018).

[140] M. Soleymanibrojeni, C.R. Caldeira Rego, M. Esmaeilpour, W. Wenzel, An active learning approach to model solid-electrolyte interphase formation in Li-ion batteries, J. Mater. Chem. A 12 (4) (2024) 2249–2266.

[141] H. Jia, M. Horton, Y. Wang, S. Zhang, K.A. Persson, S. Meng, M. Liu, Persona of Transition Metal Ions in Solids: A Statistical Learning on Local Structures of Transition Metal Oxides, Adv Sci (weinh) 9 (27) (2022) e2202756.

[142] M. Roberts, J. Owen, High-Throughput Method to Study the Effect of Precursors and Temperature, Applied to the Synthesis of LiNiCoMnO for Lithium Batteries, ACS Comb Sci 13 (2) (2011) 126–134.

[143] M.D. Fleischauer, J.M. Topple, J.R. Dahna, Combinatorial investigations of Si-M (M = Cr plus Ni, Fe, Mn) thin film negative electrode materials, Electrochem Solid St 8 (2) (2005) A137–A140.

[144] M.S. Beal, B.E. Hayden, T. Le Gall, C.E. Lee, X.J. Lu, M. Mirsaneh, C. Mormiche, D. Pasero, D.C.A. Smith, A. Weld, C. Yada, S. Yokoishi, High Throughput Methodology for Synthesis, Screening, and Optimization of Solid State Lithium Ion Electrolytes, Acs Comb Sci 13 (4) (2011) 375–381.

[145] B. Aoun, C. Yu, L.L. Fan, Z.H. Chen, K. Amine, Y. Ren, A generalized method for high throughput in-situ experiment data analysis: An example of battery materials exploration, J Power Sources 279 (2015) 246–251.

[146] B.M. Kalejahi, M. Bahram, A. Naseri, S. Bahari, M. Hasani, Multivariate curve resolution-alternating least squares (MCR-ALS) and central composite experimental design for monitoring and optimization of simultaneous removal of some organic dyes, J Iran Chem Soc 11 (1) (2014) 241–248.

[147] D. Kirikkaleli, T.S. Adebayo, Do renewable energy consumption and financial development matter for environmental sustainability? New Global Evidence. *Sustainable Development* 29 (4) (2021) 583–594.

[148] S.C. Kishore, S. Perumal, R. Atchudan, M. Alagan, A.K. Sundramoorthy, Y.R. Lee, A Critical Review on Artificial Intelligence for Fuel Cell Diagnosis, Catalysts 12 (7) (2022) 743.

[149] X. Ren, Y. Wang, A. Liu, Z. Zhang, Q. Lv, B. Liu, Current progress and performance improvement of Pt/C catalysts for fuel cells, J Mater Chem A 8 (46) (2020) 24284–24306.

[150] T. Zhou, H. Shan, H. Yu, C.a. Zhong, J. Ge, N. Zhang, W. Chu, W. Yan, Q. Xu, H. a. Wu, C. Wu, Y. Xie, Nanopore Confinement of Electrocatalysts Optimizing Triple Transport for an Ultrahigh-Power-Density Zinc-Air Fuel Cell with Robust Stability, Adv. Mater. 32 (47) (2020) 2003251.

[151] T. Nguyen, G. Lin, H. Ohn, D. Hussey, D. Jacobson, M. Arif, Measurements of two-phase flow properties of the porous media used in PEM fuel cells, ECS Trans. 3 (1) (2006) 415.

[152] N. Zhang, T. Zhou, J. Ge, Y. Lin, Z. Du, W. Wang, Q. Jiao, R. Yuan, Y. Tian, W. Chu, High-density planar-like Fe2N6 structure catalyzes efficient oxygen reduction, Matter 3 (2) (2020) 509–521.

[153] Z. Cheng'an, Z. Tianpei, Z. Nan, C. Minglong, X. Youxue, Y. Wensheng, C. Wangsheng, Z. Xusheng, X. Qian, G. Jiankai, W. Changzheng, Central electron-enriched NO-FeN₄ sites as superior acidic oxygen reduction reaction electrocatalysts for proton exchange membrane fuel cells, JUSTC 53 (3) (2023) 0304-1–0304-10.

[154] S. Chugh, C. Chaudhari, K. Sonkar, A. Sharma, G. Kapur, S. Ramakumar, Experimental and modelling studies of low temperature PEMFC performance, Int. J. Hydrogen Energy 45 (15) (2020) 8866–8874.

[155] T. Elmer, M. Worall, S. Wu, S.B. Riffat, Fuel cell technology for domestic built environment applications: State of-the-art review, Renew. Sustain. Energy Rev. 42 (2015) 913–931.

[156] S. Srinivasan, R. Mosdale, P. Stevens, C. Yang, Fuel cells: reaching the era of clean and efficient power generation in the twenty-first century, Annu. Rev. Energy Env. 24 (1) (1999) 281–328.

[157] E.H. Yu, X. Wang, U. Krewer, L. Li, K. Scott, Direct oxidation alkaline fuel cells: from materials to systems, Energ Environ Sci 5 (2) (2012) 5668–5680.

[158] M.E. Scofield, H. Liu, S.S. Wong, A concise guide to sustainable PEMFCs: recent advances in improving both oxygen reduction catalysts and proton exchange membranes, Chem. Soc. Rev. 44 (16) (2015) 5836–5860.

[159] T. Fuller, K. Gallagher, Phosphoric acid fuel cells, Elsevier, In Materials for fuel cells, 2008, pp. 209–247.

[160] A.C. Rady, S. Giddey, S.P. Badwal, B.P. Ladewig, S. Bhattacharya, Review of fuels for direct carbon fuel cells, Energy Fuel 26 (3) (2012) 1471–1488.

[161] M. Fallah Vostakola, B. Amini Horri, Progress in material development for low-temperature solid oxide fuel cells: A review, Energies 14 (5) (2021) 1280.

[162] T. Jahnke, G. Futter, A. Latz, T. Malkow, G. Papakonstantinou, G. Tsotridis, P. Schott, M. Gérard, M. Quinaud, M. Quiroga, Performance and degradation of Proton Exchange Membrane Fuel Cells: State of the art in modeling from atomistic to system scale, J Power Sources 304 (2016) 207–233.
[163] M.H. Robson, Non-platinum group metal oxgyen reduction catalysts and their mechanism in both acid and alkaline media: The effect of the catalyst precursor and the ionomer on oxygen reduction, The University of New Mexico, 2013.
[164] T. Sutharssan, D. Montalvao, Y.K. Chen, W.-C. Wang, C. Pisac, H. Elemara, A review on prognostics and health monitoring of proton exchange membrane fuel cell, Renew. Sustain. Energy Rev. 75 (2017) 440–450.
[165] A. Demuren, R.L. Edwards, Modeling proton exchange membrane fuel cells—a review. *50 Years of CFD in Engineering Sciences: A Commemorative Volume in Memory of D*, Brian Spalding (2020) 513–547.
[166] G.S. Fanourgakis, K. Gkagkas, E. Tylianakis, G.E. Froudakis, A universal machine learning algorithm for large-scale screening of materials, JACS 142 (8) (2020) 3814–3822.
[167] B. Wang, B. Xie, J. Xuan, K. Jiao, AI-based optimization of PEM fuel cell catalyst layers for maximum power density via data-driven surrogate modeling, Energ. Conver. Manage. 205 (2020) 112460.
[168] R. Ding, W. Yin, G. Cheng, Y. Chen, J. Wang, R. Wang, Z. Rui, J. Li, J. Liu, Boosting the optimization of membrane electrode assembly in proton exchange membrane fuel cells guided by explainable artificial intelligence, Energy and AI 5 (2021) 100098.
[169] D.E. Keyes, L.C. McInnes, C. Woodward, W. Gropp, E. Myra, M. Pernice, J. Bell, J. Brown, A. Clo, J. Connors, E. Constantinescu, D. Estep, K. Evans, C. Farhat, A. Hakim, G. Hammond, G. Hansen, J. Hill, T. Isaac, X. Jiao, K. Jordan, D. Kaushik, E. Kaxiras, A. Koniges, K. Lee, A. Lott, Q. Lu, J. Magerlein, R. Maxwell, M. McCourt, M. Mehl, R. Pawlowski, A.P. Randles, D. Reynolds, B. Rivière, U. Rüde, T. Scheibe, J. Shadid, B. Sheehan, M. Shephard, A. Siegel, B. Smith, X. Tang, C. Wilson, B. Wohlmuth, Multiphysics simulations, Int. J. High Perform. Comput. App. 27 (1) (2013) 4–83.
[170] Y. Su, C. Yin, S. Hua, R. Wang, H. Tang, Study of cell voltage uniformity of proton exchange membrane fuel cell stack with an optimized artificial neural network model, Int. J. Hydrogen Energy 47 (67) (2022) 29037–29052.
[171] W. Li, X. Wang, L. Wang, L. Jia, R. Song, Z. Fu, W. Xu, An LSTM and ANN Fusion Dynamic Model of a Proton Exchange Membrane Fuel Cell, IEEE T Ind Inform 19 (4) (2023) 5743–5751.
[172] E.T. Sayed, H. Rezk, M.A. Abdelkareem, A.G. Olabi, Artificial neural network based modelling and optimization of microalgae microbial fuel cell, Int. J. Hydrogen Energy 52 (2023) 1015–1025.
[173] T.H. Le, D. Thakur, P.K.T. Nguyen, Modeling and optimization of direct urea-hydrogen peroxide fuel cell using the integration of artificial neural network and bio-inspired algorithms, J. Electroanal. Chem. 922 (2022) 116783.
[174] D.D. Nguyen, Q.D.T. Pham, M. Tanveer, H. Khan, J.W. Park, C.W. Park, G.M. Kim, Deep learning–based optimization of a microfluidic membraneless fuel cell for maximum power density via data-driven three-dimensional multiphysics simulation, Bioresour. Technol. 348 (2022) 126794.
[175] J. Peng, J. Huang, X.-L. Wu, Y.-W. Xu, H. Chen, X. Li, Solid oxide fuel cell (SOFC) performance evaluation, fault diagnosis and health control: A review, J Power Sources 505 (2021) 230058.
[176] Shen, Fuzzy Model Based Control for Energy Management and Optimization in Fuel Cell Vehicles | IEEE Journals & Magazine | IEEE Xplore.
[177] M. Bressel, M. Hilairet, D. Hissel, B. Ould Bouamama, Model-based aging tolerant control with power loss prediction of Proton Exchange Membrane Fuel Cell, Int. J. Hydrogen Energy 45 (19) (2020) 11242–11254.
[178] V.M. Nagulapati, S.S. Kumar, V. Annadurai, H. Lim, Machine learning based fault detection and state of health estimation of proton exchange membrane fuel cells, Energy and AI 12 (2023) 100237.
[179] J. Zuo, H. Lv, D. Zhou, Q. Xue, L. Jin, W. Zhou, D. Yang, C. Zhang, Deep learning based prognostic framework towards proton exchange membrane fuel cell for automotive application, Appl. Energy 281 (2021) 115937.
[180] A. Perera, P. Kamalaruban, Applications of reinforcement learning in energy systems, Renew. Sustain. Energy Rev. 137 (2021) 110618.
[181] Y. Zhang, C. Zhang, R. Fan, C. Deng, S. Wan, H. Chaoui, Energy management strategy for fuel cell vehicles via soft actor-critic-based deep reinforcement learning considering powertrain thermal and durability characteristics, Energ. Conver. Manage. 283 (2023) 116921.
[182] W. Chen, J. Peng, J. Chen, J. Zhou, Z. Wei, C. Ma, Health-considered energy management strategy for fuel cell hybrid electric vehicle based on improved soft actor critic algorithm adopted with Beta policy, Energ. Conver. Manage. 292 (2023) 117362.
[183] R. Song, X. Liu, Z. Wei, F. Pan, Y. Wang, H. He, Safety and Longevity-Enhanced Energy Management of Fuel Cell Hybrid Electric Vehicle with Machine Learning Approach, IEEE Trans. Transp. Electrif. (2023) 1.
[184] C. Jia, H. He, J. Zhou, J. Li, Z. Wei, K. Li, A novel health-aware deep reinforcement learning energy management for fuel cell bus incorporating offline high-quality experience, Energy 282 (2023) 128928.
[185] Y. Zheng, X.-L. Wu, D. Zhao, Y.-W. Xu, B. Wang, Y. Zu, D. Li, J. Jiang, C. Jiang, X. Fu, X. Li, Data-driven fault diagnosis method for the safe and stable operation of solid oxide fuel cells system, J Power Sources 490 (2021) 229561.
[186] Y. Tian, Q. Zou, Z. Jin, Z. Lin, Data-driven diagnosis of the high-pressure hydrogen leakage in fuel cell vehicles based on relevance vector machine, Int. J. Hydrogen Energy 47 (24) (2022) 12281–12292.
[187] Y. Xing, B. Wang, Z. Gong, Z. Hou, F. Xi, G. Mou, Q. Du, F. Gao, K. Jiao, Data-Driven Fault Diagnosis for PEM Fuel Cell System Using Sensor Pre-Selection Method and Artificial Neural Network Model, IEEE Trans. Energy Convers. 37 (3) (2022) 1589–1599.
[188] Z. Zhang, S. Li, Y. Xiao, Y. Yang, Intelligent simultaneous fault diagnosis for solid oxide fuel cell system based on deep learning, Appl. Energy 233–234 (2019) 930–942.
[189] B. Zuo, Z. Zhang, J. Cheng, W. Huo, Z. Zhong, M. Wang, Data-driven flooding fault diagnosis method for proton-exchange membrane fuel cells using deep learning technologies, Energ. Conver. Manage. 251 (2022) 115004.
[190] X. Zhang, X. Guo, Fault diagnosis of proton exchange membrane fuel cell system of tram based on information fusion and deep learning, Int. J. Hydrogen Energy 46 (60) (2021) 30828–30840.
[191] H. Yuan, D. Tan, X. Wei, H. Dai, Fault Diagnosis of Fuel Cells by a Hybrid Deep Learning Network Fusing Characteristic Impedance, IEEE Trans. Transp. Electrif. (2023) 1.
[192] R. Kötz, M. Carlen, Principles and applications of electrochemical capacitors, Electrochim. Acta 45 (15–16) (2000) 2483–2498.
[193] E.R.M. Goikolea, Nanotechnology in Electrochemical Capacitors, in: L. M. Rodriguez-Martinez, N. Omar (Eds.), Emerging Nanotechnologies in Rechargeable Energy Storage Systems, Elsevier, Boston, 2017.
[194] K.T. Butler, D.W. Davies, H. Cartwright, O. Isayev, A. Walsh, Machine learning for molecular and materials science, Nature 559 (7715) (2018) 547–555.
[195] E.O. Pyzer-Knapp, J.W. Pitera, P.W. Staar, S. Takeda, T. Laino, D.P. Sanders, J. Sexton, J.R. Smith, A. Curioni, Accelerating materials discovery using artificial intelligence, high performance computing and robotics, NPJ Comput. Mater. 8 (1) (2022) 84.
[196] M.L. Green, B. Maruyama, J. Schrier, Autonomous (AI-driven) materials science, Appl. Phys. Rev. 9 (3) (2022).
[197] P. Sharma, T. Bhatti, A review on electrochemical double-layer capacitors, Energy Conversion Management 51 (12) (2010) 2901–2912.
[198] R. Burt, G. Birkett, X. Zhao, A review of molecular modelling of electric double layer capacitors, Phys. Chem. Chem. Phys. 16 (14) (2014) 6519–6538.
[199] X. Hu, S. Li, H. Peng, A comparative study of equivalent circuit models for Li-ion batteries, J. Power. Sources 198 (2012) 359–367.
[200] L. Zhang, X. Hu, Z. Wang, F. Sun, D.G. Dorrell, A review of supercapacitor modeling, estimation, and applications: A control/management perspective, Renewable Sustainable Energy Reviews 81 (2018) 1868–1878.
[201] S. Aderyani, P. Flouda, S. Shah, M. Green, J. Lutkenhaus, H. Ardebili, Simulation of cyclic voltammetry in structural supercapacitors with pseudocapacitance behavior, Electrochim. Acta 390 (2021) 138822.
[202] Y. Liu, B. Guo, X. Zou, Y. Li, S. Shi, Machine learning assisted materials design and discovery for rechargeable batteries, Energy Storage Mater. 31 (2020) 434–450.
[203] M.-F. Ng, J. Zhao, Q. Yan, G.J. Conduit, Z.W. Seh, Predicting the state of charge and health of batteries using data-driven machine learning, Nat. Mach. Intell. 2 (3) (2020) 161–170.
[204] S. Nanda, S. Ghosh, T. Thomas, Machine learning aided cyclic stability prediction for supercapacitors, J. Power. Sources 546 (2022) 231975.
[205] H. Su, C. Lian, J. Liu, H.J.C.E.S. Liu, Machine learning models for solvent effects on electric double layer capacitance, Chem. Eng. Sci. 202 (2019) 186–193.
[206] H.v. Helmholtz, Studien über electrische Grenzschichten, Ann. Phys. 243 (7) (1879) 337–382.
[207] D. Li, S. Li, S. Zhang, J. Sun, L. Wang, K. Wang, Aging state prediction for supercapacitors based on heuristic kalman filter optimization extreme learning machine, Energy 250 (2022) 123773.
[208] Y. Zhou, Y. Wang, K. Wang, L. Kang, F. Peng, L. Wang, J. Pang, Hybrid genetic algorithm method for efficient and robust evaluation of remaining useful life of supercapacitors, Appl. Energy 260 (2020) 114169.
[209] A.Y. Mehandzhiyski, X. Wang, C. Anquetil-Deck, D. Chen, B.A. Grimes, Microscopic Insight to Nonlinear Voltage Dependence of Charge in Carbon-Ionic Liquid Supercapacitors, Energy Mater Adv. (2021).
[210] S. Jha, S. Bandyopadhyay, S. Mehta, M. Yen, T. Chagouri, E. Palmer, H. Liang, Data-Driven Predictive Electrochemical Behavior of Lignin-Based Supercapacitors via Machine Learning, Energy Fuels 36 (2) (2022) 1052–1062.
[211] M. Shariq, S. Marimuthu, A.R. Dixit, S. Chattopadhyaya, S. Pandiaraj, M. Muthuramamoorthy, A.N. Alodhyab, M. Khaja Nazeeruddin, A.N. Grace, Machine learning models for prediction of electrochemical properties in supercapacitor electrodes using MXene and graphene nanoplatelets, Chem. Eng. J. 484 (2024) 149502.
[212] P. Liu, Y. Wen, L. Huang, X. Zhu, R. Wu, S. Ai, T. Xue, Y. Ge, An emerging machine learning strategy for the assisted-design of high-performance supercapacitor materials by mining the relationship between capacitance and structural features of porous carbon, J. Electroanal. Chem. 899 (2021) 115684.
[213] Q. Dou, H.S. Park, Perspective on high-energy carbon-based supercapacitors, Energy Environ. Mater. 3 (3) (2020) 286–305.
[214] Y. Zhu, S. Murali, M.D. Stoller, K.J. Ganesh, W. Cai, P.J. Ferreira, A. Pirkle, R. M. Wallace, K.A. Cychosz, M.J.S. Thommes, Carbon-Based Supercapacitors Produced by Activation of Graphene 332 (6037) (2011) 1537–1541.
[215] H. Jiang, P.S. Lee, C. Li, 3D carbon based nanostructures for advanced supercapacitors, Energy Environ. Sci. 6 (1) (2013) 41–53.
[216] X. Chen, R. Paul, L. Dai, Carbon-based supercapacitors for efficient energy storage, Nat. Sci. Rev. 4 (3) (2017) 453–489.
[217] S. Zhu, J. Li, L. Ma, C. He, E. Liu, F. He, C. Shi, N. Zhao, Artificial neural network enabled capacitance prediction for carbon-based supercapacitors, Mater. Lett. 233 (2018) 294–297.
[218] H. Su, S. Lin, S. Deng, C. Lian, Y. Shang, H. Liu, Predicting the capacitance of carbon-based electric double layer capacitors by machine learning, Nano. Adv. 1 (6) (2019) 2162–2166.

[219] J. Abdi, T. Pirhoushyaran, F. Hadavimoghaddam, S.A. Madani, A. Hemmati-Sarapardeh, S.H. Esmaeili-Faraj, Modeling of capacitance for carbon-based supercapacitors using Super Learner algorithm, J. Energy Storage 66 (2023) 107376.
[220] R. Kushwaha, M.K. Singh, S. Krishnan, D.K. Rai, Machine learning enabled property prediction of carbon-based electrodes for supercapacitors, J. Mater. Sci. 1–11 (2023).
[221] S.I. Ahmed, S. Radhakrishnan, B.B. Nair, R. Thiruvengadathan, Efficient hyperparameter-tuned machine learning approach for estimation of supercapacitor performance attributes, J. Phys. Comm. 5 (11) (2021) 115011.
[222] W.Z. Tawfik, S.N. Mohammad, K.H. Rahouma, E. Tammam, G.M. Salama, An artificial neural network model for capacitance prediction of porous carbon-based supercapacitor electrodes, J Energy Storage 73 (2023) 108830.
[223] T. Wang, R. Pan, M.L. Martins, J. Cui, Z. Huang, B.P. Thapaliya, C.-L. Do-Thanh, M. Zhou, J. Fan, Z. Yang, Machine-learning-assisted material discovery of oxygen-rich highly porous carbon active materials for aqueous supercapacitors, Nat. Commun 14 (1) (2023) 4607.
[224] M. Rahimi, M.H. Abbaspour-Fard, A. Rohani, Synergetic effect of N/O functional groups and microstructures of activated carbon on supercapacitor performance by machine learning, J. Power. Sources 521 (2022) 230968.
[225] S. Ghosh, S. Barg, S.M. Jeong, K. Ostrikov, Heteroatom-doped and oxygen-functionalized nanocarbons for high-performance supercapacitors, Adv. Energy Mater. 10 (32) (2020) 2001239.
[226] S. Mishra, R. Srivastava, A. Muhammad, A. Amit, E. Chiavazzo, M. Fasano, P. Asinari, The impact of physicochemical features of carbon electrodes on the capacitive performance of supercapacitors: a machine learning approach, Sci. Rep. 13 (1) (2023) 6494.
[227] B. Reddy, P. Narayana, A. Maurya, U.M.R. Paturi, J. Sung, H.-J. Ahn, K. Cho, N. Reddy, Modeling capacitance of carbon-based supercapacitors by artificial neural networks, J. Energy Storage 72 (2023) 108537.
[228] A.G. Saad, A. Emad-Eldeen, W.Z. Tawfik, A.G. El-Deen, Data-driven machine learning approach for predicting the capacitance of graphene-based supercapacitor electrodes, J. Energy Storage 55 (2022) 105411.
[229] A. Chenwittayakhachon, K. Jitapunkul, B. Nakpalad, P. Worrayotkovit, S. Namuangruk, P. Sirisinudomkit, P. Iamprasertkun, Machine Learning Approach to Understanding the 'synergistic'pseudocapacitive Effects of Heteroatom Doped Graphene, 2D Mater 10 (2) (2023) 025003.
[230] Y. Wang, J. Sha, S. Zhu, L. Ma, C. He, C. Zhong, W. Hu, N. Zhao, Data-driven design of carbon-based materials for high-performance flexible energy storage devices, J. Power. Sources 556 (2023) 232522.
[231] A. Jamaluddin, D. Harjunowibowo, S. Budiawanti, N.A. Kurdhi, D.T. Lai, S. Ramesh, Applying machine learning to understand the properties of biomass carbon materials in supercapacitors, Energy Rep. 10 (2023) 3125–3132.
[232] X. Lu, P. Liu, K. Bisetty, Y. Cai, X. Duan, Y. Wen, Y. Zhu, L. Rao, Q. Xu, J. Xu, An emerging machine learning strategy for electrochemical sensor and supercapacitor using carbonized metal–organic framework, J. Electroanalytical Chem. 920 (2022) 116634.
[233] K. Lu, D. Chang, X. Ji, M. Li, W. Lu, Machine learning aided discovery of the layered double hydroxides with the largest basal spacing for super-capacitors, Int. J. Electrochemical Sci. 16 (11) (2021) 211146.
[234] S. Ghosh, G.R. Rao, T.J.E.S.M. Thomas, Machine learning-based prediction of supercapacitor performance for a novel electrode material: Cerium oxynitride, Energy Storage Mater. 40 (2021) 426–438.
[235] S. Jha, S. Bandyopadhyay, S. Mehta, M. Yen, T. Chagouri, E. Palmer, H. Liang, Data-driven predictive electrochemical behavior of lignin-based supercapacitors via machine learning, Energy Fuels 36 (2) (2021) 1052–1062.
[236] S. Yu, Q. Wang, J. Wang, C. Fang, Y. Li, J. Ge, B. Fang, Computational Modeling Guided Design of Metal-Organic Frameworks for Photocatalysis—A Mini Review, Catal. Sci. Technol. 13 (2023) 6583–6603.
[237] T. Zhou, N. Zhang, M. Chen, Y. Xie, W. Yan, W. Chu, X. Zheng, Q. Xu, J. Ge, C. Wu, Central electron-enriched NO-FeN 4 sites as superior acidic oxygen reduction reaction electrocatalysts for proton exchange membrane fuel cells, JUSTC 53 (3) (2023) 0304-1.
[238] Y. Xu, J. Ge, C.W. Ju, Machine Learning in Energy Chemistry: Introduction, Challenge and Perspective, Energy Adv. 2 (2023) 896–921.
[239] M.T.M. Koper, A.Z. Weber, K. Chan, J. Cheng, Introduction: computational electrochemistry, Chem. Rev. 122 (12) (2022) 10579–10580.
[240] B.C. Melot, Design and preparation of materials for advanced electrochemical storage, Accounts Chem. Res. 46 (5) (2013) 1226–1238.
[241] J. Deb, L. Saikia, K.D. Dihingia, G.N. Sastry, ChatGPT in the Material Design: Selected Case Studies to Assess the Potential of ChatGPT, J. Chem. Inform. Model. (2024).
[242] H. Wang, L. Zhang, J. Han, DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics, Comput. Phys. Commun. 228 (2018) 178–184.