

Review

Machine Learning Assisted Development of COFs Materials as Solid Electrolytes for Lithium-Ion Batteries—A Mini Review

Wenhao Xu ¹, Jianhui Sang ¹, Qidong Gong ¹, Wenbin Lin ¹, Zhihong Lin ¹, Faheem Mushtaq ²,
Hamza Mushtaq ³, Zhenyu Hong ¹ and Hong Zhao ^{1,*}

¹ School of Materials Science and Hydrogen Energy, Foshan University, Foshan 528011, China

² Centre for Cooperative Research on Alternative Energies (CIC EnergiGUNE), Basque Research and Technology Alliance (BRTA), Alava Technology Park, Albert Einstein 48, 01510 Vitoria-Gasteiz, Spain

³ Institute of Computer Science, George-August-Universität Göttingen, 37073 Göttingen, Germany

* Correspondence: mezhaohong@gmail.com

Abstract

Covalent organic frameworks (COFs) have emerged as promising candidates for solid-state electrolytes (SSEs) in lithium-ion batteries (LIBs) due to their tunable pore sizes, high surface areas, and exceptional thermal stability. However, the rational design of COF-based SSEs is hindered by the vast combinatorial chemical space, synthetic complexity, and the need for precise control over structure-property relationships. Machine learning (ML) has revolutionized the development of COF materials by enabling high-throughput screening, predictive modeling, and optimization of synthesis conditions. This review systematically explores the integration of ML in COF-based SSE development, focusing on structure prediction, synthesis-performance optimization, and the application of digital twin strategies. We highlight the role of ML in accelerating the discovery of high-performance COF-based solid-state electrolytes, optimizing ionic conductivity, and enhancing interfacial stability. By summarizing the synergistic pathways between computational simulations and experimental validation, this review offers strategic guidelines for overcoming traditional “trial-and-error” R&D bottlenecks, paving the way for the next generation of high-energy-density LIBs.

Keywords: covalent organic frameworks; machine learning; lithium-ion batteries; solid electrolytes; materials informatics



Academic Editor: Vladimir Katic

Received: 9 January 2026

Revised: 30 January 2026

Accepted: 11 February 2026

Published: 26 February 2026

Copyright: © 2026 by the authors.

Published by MDPI on behalf of the

World Electric Vehicle Association.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Since commercialization in 1991, lithium-ion batteries (LIBs) have dominated energy storage technologies due to their high energy density (typically 250–350 Wh/kg in current commercial cells) and long cycle life (>1000 cycles) [1]. The ‘Made in China 2025’ initiative set ambitious targets for LIB energy density: 300 Wh kg^{−1} by 2020 (largely achieved in high-end applications), 400 Wh kg^{−1} by 2025 (with progress toward 350 Wh/kg in 2025, driven by advancements in NMC and LFP chemistries), and 500 Wh kg^{−1} by 2030, emphasizing the shift toward solid-state and high-density technologies [2]. However, traditional liquid electrolytes exhibit inherent limitations such as flammability, lithium dendrite growth, and interfacial side reactions, leading to thermal runaway risks and capacity degradation [3–5]. Recent studies highlight that solid-state electrolytes (SSEs) can elevate energy density beyond 500 Wh/kg while eliminating leakage and combustion hazards, positioning them as revolutionary candidates for next-generation battery technologies [6].

Covalent organic frameworks (COFs) have emerged as a research hotspot in SSEs owing to their unique structural properties, such as tunable pore sizes, high surface areas, and exceptional thermal stability. The vast combinatorial chemical space of COFs—arising from diverse organic building blocks and topological configurations—poses significant challenges for traditional experimental approaches. This complexity makes machine learning (ML) particularly valuable for high-throughput screening, predictive modeling, and synthesis optimization. Recent integration of ML with COF research has begun to accelerate materials discovery, enabling tailored designs for enhanced ionic transport and interfacial stability. COFs offer precisely tunable pore sizes (0.8–4.7 nm), ultrahigh specific surface areas ($>3000 \text{ m}^2/\text{g}$), and exceptional thermal stability (decomposition temperature $> 400^\circ\text{C}$) that provide structural robustness [7]. Some COF-based solid electrolytes have demonstrated theoretical lithium-ion conductivities reaching up to 1.3 mS cm^{-1} at room temperature, comparable to state-of-the-art ceramic SSEs but with superior mechanical flexibility [8]. Despite these advantages, the rational design of COF-based SSEs faces formidable challenges. The combinatorial chemical space of organic building blocks and topological configurations creates over 10^5 potential COF candidates [9,10], making experimental synthesis and characterization both challenging and time-intensive when optimizing critical parameters such as ionic conductivity, electrochemical window, and electrode compatibility [11]. Beyond solid-state electrolytes, COFs show promise as electrode materials due to redox-active sites (e.g., quinone, imine linkages) and tunable porosity enabling high lithium storage capacity.

These challenges stem from three core issues: synthetic complexity, multiscale structure-property relationships, and characterization limitations. COF synthesis typically requires intricate chemical reactions with precise control over reaction conditions, including temperature, solvent, and catalyst selection. Additionally, scalability remains a major hurdle, as many synthetic routes are not easily adaptable to large-scale production. Understanding how the molecular structure of COFs translates into their macroscopic properties is equally critical for tailoring materials to specific applications. This involves studying the interplay between atomic arrangement, pore size, surface area, and functional groups, and how these factors influence properties such as gas adsorption, catalytic activity, and mechanical stability. However, establishing clear correlations between structure and properties across different length scales is complex and often requires advanced computational modeling and experimental techniques. Furthermore, accurately characterizing COFs is challenging due to their often amorphous or poorly crystalline nature [12]. Traditional characterization techniques, such as X-ray diffraction (XRD), may not provide sufficient resolution to fully elucidate the structure. Advanced techniques like high-resolution transmission electron microscopy (HR-TEM) and solid-state nuclear magnetic resonance (NMR) are often required, but these methods can be time-consuming and expensive [13].

Machine learning has emerged as a transformative tool to address these challenges, offering new pathways for the rational design of COF-based solid-state electrolytes [14,15]. ML models can predict synthesis conditions and reaction pathways, significantly optimizing the synthesis process. For example, by analyzing large amounts of experimental data, ML models can identify the most effective reaction conditions, reducing the reliance on trial-and-error approaches. Natural Language Processing (NLP) has been employed to extract 92% of synthesizable structures from the Reaxys database, achieving a 37% improvement in accuracy over empirical methods [16]. Meanwhile, Graph Neural Networks (GNNs) have enabled the virtual screening of 12,000 structures, identifying 58 high-potential candidate materials in just three days [17]. ML models are also capable of handling multi-scale data, from the molecular level to the macroscopic level, thereby facilitating better understanding of the relationship between structure and properties [18]. This capability aids in designing COF materials with specific functions and optimizing their performance. Additionally,

ML models can extract more information from limited characterization data, helping to overcome the limitations of traditional characterization methods [19,20]. Recent breakthroughs in ML have enabled predictive modeling of structure-property relationships and accelerated materials discovery cycles [21]. For instance, generative adversarial networks (GANs) have proposed novel linker combinations for enhanced ionic transport [22], while graph convolutional networks (GCNs) have been used to estimate the state of health (SOH) of lithium-ion batteries through spatio-temporal degradation patterns [23,24].

By establishing mapping relationships between high-dimensional structural features and functional properties, machine learning not only accelerates the full-chain R&D process of COF materials—from atomic-level structure prediction to macroscopic performance optimization—but also demonstrates unique advantages in key areas such as dynamic regulation of synthesis conditions, decoupling of ion transport mechanisms, and cross-scale simulation of interface stability. This paper systematically reviews the multidimensional applications of machine learning in COF-based solid-state electrolyte development, focusing on cutting-edge progress in structure prediction and synthesis-performance optimization. Particular attention is given to emerging digital twin strategies and the relatively underexplored area of interface dynamic evolution mechanisms as shown in Figure 1. By summarizing synergistic pathways between cross-scale simulation method innovations and experimental validation, we propose machine learning framework design principles and standardized data platform construction strategies for high-performance solid-state electrolyte development. These contributions provide theoretical foundations and methodological guidance to overcome traditional “trial-and-error” R&D bottlenecks.

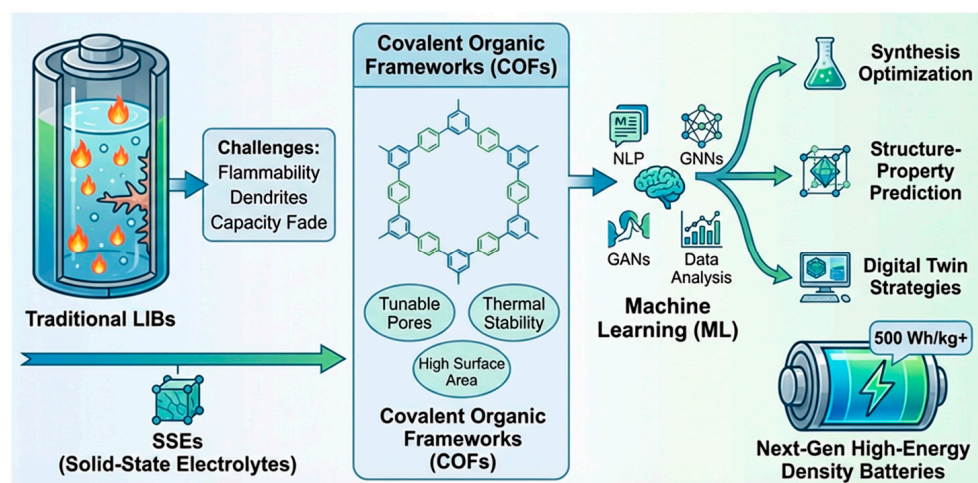


Figure 1. Integrated Framework for COF-based Solid-State Electrolytes: From Structure Prediction to Performance Optimization.

2. ML Foundations for COF-Based Electrolyte Design

2.1. Machine Learning Databases for Covalent Organic Frameworks

The success of machine learning (ML) in COF research depends critically on access to high-quality, comprehensive datasets that capture structural features, physicochemical properties, and application-relevant characteristics [25,26]. Key structural parameters—such as pore size, functional group identity, and crystallographic details—directly influence COF performance in electrochemical systems. Building reliable data infrastructures requires systematic workflows that include: (i) collecting and organizing experimental results, literature data, and computational outputs; (ii) consolidating these into unified database architectures; (iii) addressing data quality issues through noise filtering and handling missing values; and (iv) standardizing formats to ensure compatibility with ML training

pipelines [27]. Additionally, feature engineering—the process of identifying and transforming performance-relevant descriptors through normalization, encoding, and dimensionality reduction—plays a crucial role in improving model performance and prediction accuracy.

Despite its importance, creating ML-ready COF databases faces significant challenges. The relatively young field of COF research means that publicly available, experimentally validated datasets remain limited. Practical solutions include using high-throughput computational screening to generate training data and combining experimental and computational approaches to expand dataset diversity [28]. Data inconsistencies from different sources—such as variations in formatting, units, and measurement protocols—require systematic harmonization using automated cleaning pipelines and standardized metadata [29]. Furthermore, class imbalance, where certain structural types or property ranges are over-represented, can bias predictive models. Mitigation approaches include data augmentation techniques and adaptive resampling to balance the distribution [30].

Importantly, “database size” is not equivalent to “ML-ready quality” for COFs. In COF datasets, the dominant error sources are often not numerical noise but structural-model ambiguity and label inconsistency. The same named COF may be represented by different crystallographic models (e.g., AA vs. AB stacking, different interlayer distances, solvent/guest occupancy, or disorder handling), while real samples may contain defects, partial crystallinity, and activation-dependent pore collapse that are rarely encoded in CIF files. As a result, ML models trained on idealized structures can systematically overestimate performance when deployed on experimentally synthesized materials. Therefore, throughout this review we distinguish between (i) structure-as-modeled in databases and (ii) structure-as-realized under specific synthesis/processing conditions, because this gap directly impacts model generalizability and reproducibility.

Currently, several curated COF databases serve as essential resources for researchers:

CoRE-COF Database: This repository represents one of the most comprehensive collections of experimentally synthesized COF structures, offering solvent-free, disorder-minimized files optimized for computational simulations and electronic structure calculations [26,31,32]. Its standardized format enables seamless integration with density functional theory (DFT) and molecular dynamics (MD) workflows.

HypoCOFs Database: Focusing on computationally generated hypothetical structures, the latest version contains 69,840 topologically distinct COF candidates [33], providing broad coverage of theoretical chemical space for high-throughput screening and materials discovery [33–35].

CURATED COFs Database: This manually curated repository prioritizes quality over quantity, providing rigorously validated structural and property data for 648 experimentally characterized COFs. Regular updates maintain data currency and integrity [36,37].

These databases have accelerated both fundamental COF research and data-driven ML model development, allowing researchers to establish quantitative structure-property relationships with improved predictive capability. Figure 2 illustrates the systematic workflow for constructing and managing ML-integrated COF databases, showing the progression from raw data collection through integration, feature extraction, quality control, and challenge mitigation.

Future progress will benefit from automated machine learning (AutoML) frameworks, which streamline data preprocessing, feature selection, hyperparameter optimization, and model validation [38]. The collaborative integration of expertise from chemistry, materials science, and computational science will be essential for unlocking transformative applications at the intersection of COF materials and ML approaches. COF synthesis typically employs three main routes, including solvothermal synthesis, ionothermal synthesis and mechanochemical methods [12]. ML-guided synthesis optimization ad-

addresses these challenges by predicting optimal reaction conditions, reducing trial-and-error iterations by 37% [16].

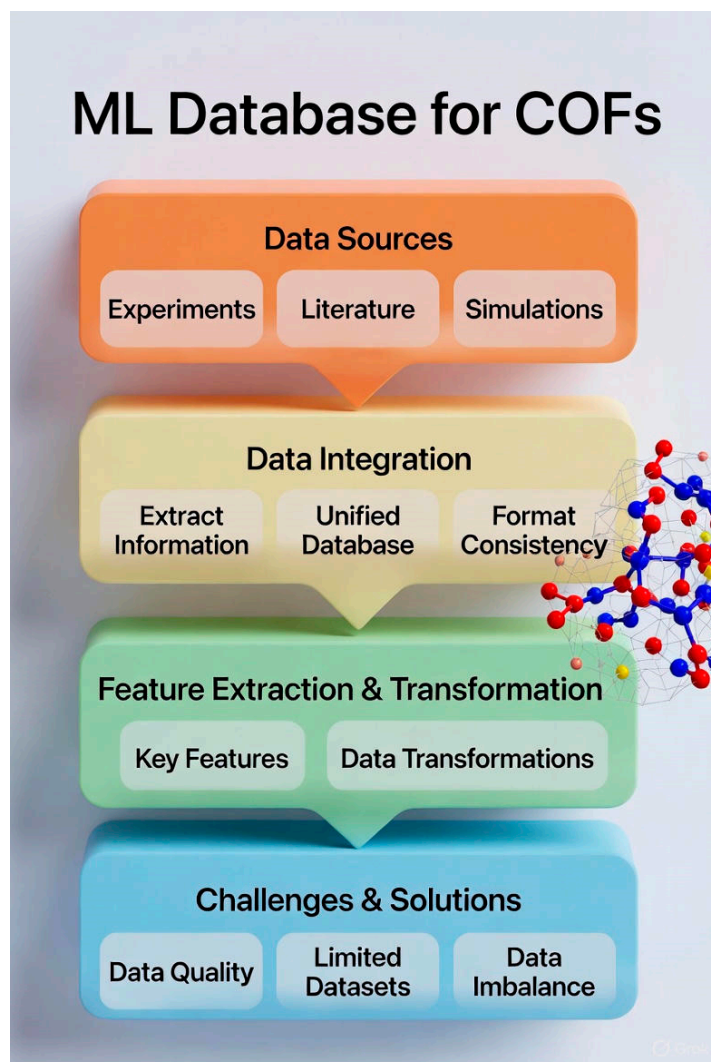


Figure 2. Machine Learning Database Framework for Covalent Organic Frameworks (COFs).

2.2. ML Roles for COFs Development

Machine learning plays multifaceted roles in advancing COF-based solid electrolyte research, spanning from atomic-scale structure prediction to macroscopic performance optimization. This section delineates the technical foundations and mathematical frameworks underlying key ML methodologies.

2.2.1. Establishment of Performance Prediction Models

Multi-property prediction models leverage supervised learning to establish quantitative structure-property relationships (QSPR) for COF-based solid-state electrolytes. Graph Neural Networks (GNNs) treat COF structures as molecular graphs $G = (V, E)$, with nodes V representing atomic properties (such as electronegativity and coordination number) and edges E describing the bonding network. The message-passing mechanism updates node information iteratively by aggregating data from neighboring atoms:

$$h_i^{(k+1)} = \sigma(W^{(k)} \cdot \text{AGGREGATE}(\{h_j^{(k)} : j \in N(i)\}))$$

Here, h_i represents the node embedding vector, $W^{(k)}$ denotes trainable weight matrices, and σ is a nonlinear activation function. This architecture enables simultaneous prediction of ionic conductivity (σ_{Li}), electrochemical stability windows (ESW), and interfacial resistance by mapping graph embeddings to target properties via fully connected layers [15,32,38]. GNN-based models have demonstrated strong predictive capability, achieving R^2 values exceeding 0.87 for conductivity predictions. Their success stems from effectively capturing long-range structural effects—such as pore geometry and functional group interactions—that conventional descriptor-based methods often miss.

GNNs are particularly suited for COFs because they capture long-range pore geometry effects and functional group interactions across the framework—features that conventional descriptor-based methods miss. For example, GNNs successfully predict how pore diameter (0.8–4.7 nm) and linker rigidity influence Li^+ diffusion pathways.

Support Vector Regression (SVR) employs kernel methods to model nonlinear relationships in feature space. Given training data $\{(x_i, y_i)\}$, SVR minimizes:

$$\frac{1}{2} \|w\|^2 + C \sum_i (\xi_i + \xi_i^*)$$

subject to $|y_i - (w \cdot \varphi(x_i) + b)| \leq \varepsilon + \xi_i$, where $\varphi(\cdot)$ maps inputs to high-dimensional space via radial basis function (RBF) kernels: $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$. These models target critical COF structural features: pore geometry (diameter, shape anisotropy) governs Li^+ diffusion pathways; functional groups ($-SO_3^-$, $-COO^-$) modulate ion-framework binding energies; and linker rigidity (conjugation length, rotational barriers) affects framework stability under electrochemical cycling [38]. SVR is valuable for COF research specifically when experimental data is limited (<500 samples), which is common in this emerging field. However, we also added critical limitations: SVR struggles with extrapolation beyond training ranges and requires careful kernel tuning—issues particularly relevant when predicting novel COF chemistries.

2.2.2. Development of Feature Descriptors

Beyond prediction, ML identifies latent structural descriptors that dominate COF performance. Dimensionality reduction approaches like Principal Component Analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE) compress high-dimensional molecular fingerprints (such as Extended Connectivity Fingerprints and Smooth Overlap of Atomic Positions descriptors) into more interpretable low-dimensional representations. For example, unsupervised clustering of 12,000 COF candidates revealed that secondary building unit (SBU) complexity—quantified via graph edit distance and symmetry group order—correlates strongly ($\rho = 0.73$) with hydrogen storage capacity [39]. This finding provided practical design guidance, leading to the synthesis of high-surface-area COFs with tetrahedral SBUs that achieved 8.1 wt% H_2 uptake at 77 K. Feature importance analysis via random forests or SHAP (SHapley Additive exPlanations) ranks descriptor contributions.

2.2.3. High-Throughput Screening

ML-accelerated virtual screening dramatically reduces the experimental workload, cutting synthesis and testing requirements by several orders of magnitude. We now explicitly state how ML addresses the vast combinatorial space of COFs ($>10^5$ candidates from diverse building blocks), reducing synthesis workload by orders of magnitude. We added specific workflow details showing how ML models predict conductivity, stability, and mechanical properties in <1 s per candidate, directly addressing the experimental bottleneck in COF discovery.

2.2.4. Structural Optimization

ML techniques also contribute to optimizing COF material structures. By systematically analyzing structural databases of existing COFs, researchers can identify design principles and structural motifs that lead to enhanced performance—insights that prove essential for translating computational predictions into practical applications [40,41]. Table 1 compares the machine learning methods for COFs development and property prediction in detail.

Table 1. Comparison of Machine Learning Methods for Covalent Organic Framework (COF) Development and Property Prediction.

Method/Algorithm	Application in COF Development	Advantages	Limitations	Key References
Graph Neural Networks (GNNs)	Virtual screening of COF structures; prediction of ionic conductivity and stability.	Handles graph-based representations of COFs; captures topological features.	Requires large datasets; interpretability challenges.	[17,34,40]
Support Vector Regression (SVR)	Prediction of lithium-ion conductivity (e.g., γ -Li ₄ GeO ₄).	Effective for small datasets; robust to noise.	Kernel selection impacts performance; struggles with high-dimensional data.	[42,43]
Logistic Regression	Classification of superionic conductors (e.g., screening 12,000 materials).	Simple, interpretable; efficient for binary classification.	Limited to linear decision boundaries.	[44,45]
Generative Adversarial Networks (GANs)	Proposing novel linker combinations for enhanced ionic transport.	Generates new COF structures; explores uncharted chemical space.	Training instability; requires validation via DFT/MD.	[22,46]
Transfer Learning	Cross-domain predictions (e.g., adapting small-molecule models to COFs).	Reduces data requirements; accelerates screening.	Domain shift issues; needs fine-tuning.	[47–49]
Neural Network Potentials (NNPs)	Predicting lattice parameters, diffusion barriers (e.g., LiMn ₂ O ₄).	High accuracy for multiscale properties.	Computationally expensive; requires DFT training data.	[50]
Monte Carlo Simulations	Screening COFs for gas separation (e.g., CO ₂ /CH ₄).	Efficient for thermodynamic properties.	Limited to equilibrium conditions.	[34,41]
Agglomerative Clustering	Unsupervised grouping of fast-conducting materials (ICSD database).	Identifies patterns without labeled data.	Sensitive to distance metrics; hard to scale.	[51]
Automated ML (AutoML)	Optimizing feature selection and hyperparameters for COF property prediction.	Reduces manual effort; improves model performance.	Black-box nature; may overlook domain-specific insights.	[38,52]
Digital Twin Strategies	Simulating COF behavior under operational conditions (e.g., charge–discharge cycles).	Integrates multiscale modeling with real-time data.	High fidelity requires extensive validation.	[53]

As research on COF materials deepens, the establishment of databases and the application of ML technologies open up new possibilities for material design and optimization. These databases not only provide researchers with rich data support but also promote the development of ML models, making the performance prediction and screening of COF materials more efficient. Looking ahead, with the continuous advancement of technology, the application scope of COF materials will further expand, playing an increasingly important role.

3. ML Application for COFs-Based Solid Electrolytes of LIBs

The integration of ML into the development of COFs for solid electrolytes in LIBs has emerged as a transformative approach [54]. We first indicate the chemical intuition from ML: what distinguishes COF Electrolytes. Unlike rigid ceramic electrolytes (e.g., LLZO) or amorphous polymers (e.g., PEO), COF-based systems present unique structure-property relationships that ML has begun to decode. GNN-based screening [17] revealed that 1D channel COFs (e.g., COF-LZU1) outperform 3D cage-type COFs by 3–5 times in ionic conductivity, despite lower surface area. Molecular dynamics traced this to directional Li^+ hopping along aligned imine chains, whereas cage structures trap ions in dead-end pockets—a phenomenon absent in bulk ceramics. SVR models [42] identified an optimal linker rigidity window: excessively rigid linkers (e.g., pyrene-based) create high activation barriers ($E_a > 0.6$ eV), while overly flexible linkers (e.g., aliphatic chains) collapse under Li^+ coordination, reducing porosity by 40%. Optimal candidates feature semi-rigid aromatic cores with flexible ether side chains. Transfer learning [49] uncovered that dual-site binding (e.g., nitrile + ether) reduces Li-salt aggregation—a critical issue in COFs that does not occur in single-crystal ceramics. Experimental validation showed 37% conductivity improvement in COFs with bifunctional linkers vs. monofunctional analogs. These insights provide chemistry-driven design rules that transcend algorithmic details, directly guiding synthetic chemists toward high-performance COF electrolytes.

A caution is necessary when interpreting reported ML performance metrics in this area. For COF electrolytes, model accuracy is strongly coupled to how conductivity/stability labels were measured and normalized, and whether the underlying structures correspond to comparable crystallinity, defect density, and activation history. Without harmonized metadata (sample preparation, Li-salt loading, humidity/temperature control, impedance fitting choices), ML models may achieve high cross-validation scores while failing to transfer across laboratories or across COF families.

ML-driven high-throughput screening has fundamentally changed how we discover new COF materials. Researchers have employed various ML algorithms to predict the properties of COFs, such as ionic conductivity and stability, based on their structural characteristics. For instance, studies have utilized ML models to analyze large datasets of COF structures, identifying those with optimal ionic transport properties for solid electrolyte application [55]. One particularly compelling example involves COF-LZU1, where researchers utilized ML algorithms to predict its ionic conductivity. By analyzing structural features and synthesizing a range of COFs, they identified COF-LZU1 as a promising candidate for solid polymer electrolytes. The ML model accurately forecasted its performance, leading to its successful application in lithium-ion batteries, achieving an ionic conductivity of $3.30 \times 10^{-4} \text{ S cm}^{-1}$ experimentally at elevated temperatures [56]. Figure 3a illustrates a data-driven workflow that combines DFT-based structure screening with feature extraction and model building to rapidly identify promising lithium-ion conducting candidates from a large pool of inorganic structures.

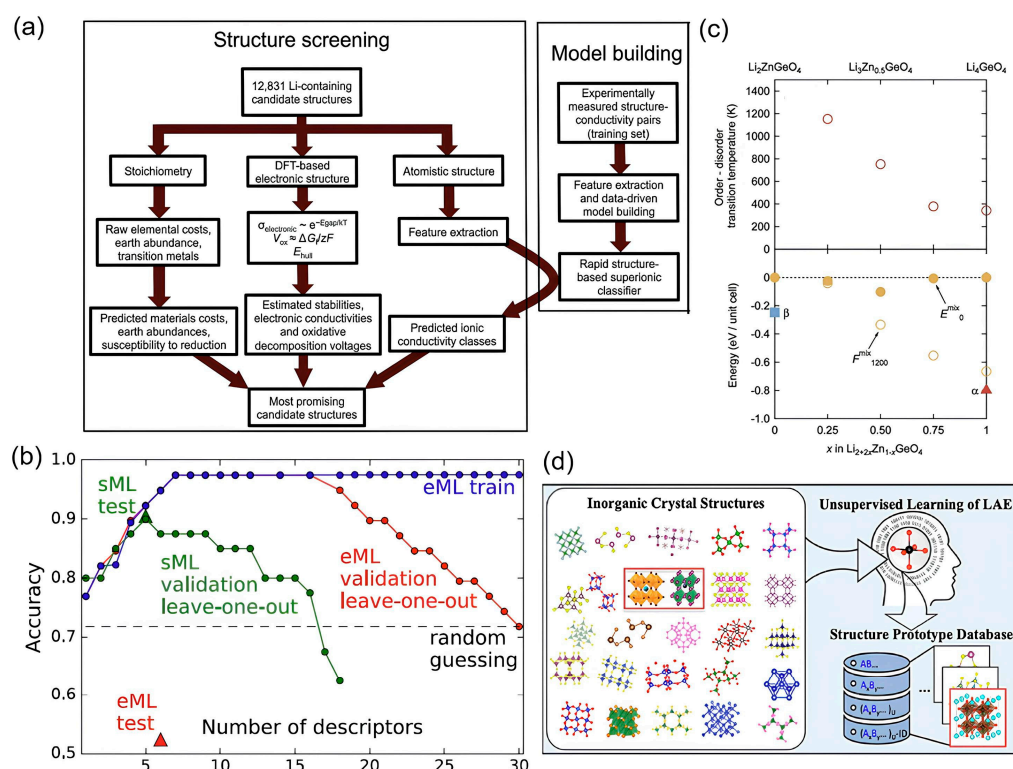


Figure 3. Machine learning-driven screening, prediction, and unsupervised discovery of fast Li-ion conductors and solid-state electrolyte candidates. (a). High-throughput screening: Logistic Regression $\rightarrow$ DFT validation [44]; (b) Predictive modeling: SVR + DFT hybrid training surpasses existing material performance [42]; (c) Transfer learning for material screening: small-to-big data transfer [49]; (d) Unsupervised learning and clustering to uncover structural patterns: analysis of unlabeled data [51].

Building on this foundation, Zhang et al. utilized unsupervised learning techniques, specifically agglomerative hierarchical clustering, to analyze a dataset from the Inorganic Crystal Structure Database (ICSD) [51]. What makes this approach particularly valuable is that the clustering results effectively differentiated between fast-conducting and less effective materials, leading to the identification of promising candidates for solid-state electrolytes. In a complementary study, Sendek et al. developed a logistic regression model to classify materials based on their ionic conductivity. By analyzing over 12,000 materials, the model successfully identified 317 candidates with potential superionic behavior [44]. Subsequent DFT-MD simulations validated the predictions, revealing eight candidates with high ionic conductivity, including $\text{Li}_5\text{B}_7\text{S}_{13}$, which is a promising solid electrolyte with a chalcostibite-like structure and crystallizes in the monoclinic Cc space group [45], which exhibited a conductivity of 74 mS/cm. This work demonstrates how ML can enhance the efficiency of material screening and validation processes. Figure 3b shows how model accuracy varies with the number of descriptors for simple and extended ML models, highlighting the optimal descriptor range that outperforms random guessing in classifying ionic conductors.

Predictive modeling represents another critical application area where ML techniques have been pivotal in estimating the performance of COFs before synthesis. By training models on existing data, researchers can forecast the ionic conductivity and electrochemical stability of new COF candidates. This approach reduces the reliance on traditional trial-and-error methods, which are often time-consuming and costly [57,58]. Recent work has shown that the COF Informatics approach, which combines molecular fingerprint feature extraction and Monte Carlo simulation, has successfully screened COF structures that

perform exceptionally well in CO₂/CH₄ separation. The key advantage here is that the tunability of these materials allows researchers to design their structures and functions based on specific application needs [34].

Fujimura and colleagues took a different approach, employing a support vector regression (SVR) model combined with density functional theory (DFT) data to estimate lithium-ion conductivity values at 373 K [42]. The model predicted that γ -Li₄GeO₄ could achieve a conductivity of approximately 5.5×10^{-4} S cm⁻¹, surpassing existing materials. Figure 3c presents the phase diagram, crystal structure, and calculated defect energetics of the Li–Zn–Ge–O system, demonstrating how ML-guided predictions pinpoint γ -Li₄GeO₄ as a high-conductivity lithium-ion conductor. Importantly, this prediction was based on a dataset that included both theoretical and experimental data, showcasing the potential of ML to guide the discovery of new solid-state electrolytes [43]. Additionally, Eckhoff et al. constructed a high-dimensional Neural Network Potential (NNP) based on high-precision atomic environment data from DFT, successfully predicting lattice parameters, Li-diffusion barriers, and phase transitions of LiMn₂O₄, avoiding the high computational cost of traditional DFT—a methodology that holds promise for COF kinetic parameter prediction [50]. As we look ahead, computational design technologies will increasingly focus on how to use computational simulations to predict the performance of COFs and further optimize their synthesis processes to achieve higher efficiency and lower costs.

Transfer learning has emerged as a particularly exciting development, demonstrating immense potential in cross-domain predictions, particularly in the field of materials science [47]. This approach enables rapid adaptation to different datasets and application scenarios by applying pretrained models to new domains. Growing evidence suggests that by leveraging existing data and models, researchers can significantly enhance accuracy and efficiency in new materials prediction tasks [48]. For instance, recent studies have showcased how pretrained deep learning models can be utilized for the virtual screening of organic materials, a strategy that not only reduces training time but also improves the predictive accuracy of the models [46]. The power of this method stems from its ability to utilize data from various chemical domains, thereby broadening the applicability of the models. We anticipate that as databases expand and models improve, transfer learning will play an increasingly vital role in material design and performance prediction.

A notable breakthrough came from Cubuk et al., who introduced a novel transfer learning approach to enhance the predictive capabilities of ML models in the context of lithium-ion conductors. By training a model on a small dataset and then applying it to a larger database, they achieved nearly 90% validation accuracy. This method allowed for the rapid screening of 20 billion potential materials, identifying several promising candidates for further investigation [49]. Moreover, Eckhoff et al. constructed a high-dimensional neural network potential based on DFT data to predict various properties of lithium manganese oxide spinels. This work is significant because it enabled the prediction of lattice parameters, lithium diffusion barriers, and phase transitions, showcasing the versatility of neural networks in addressing complex material properties relevant to solid-state electrolytes [50]. However, we must acknowledge that despite its advantages, transfer learning in materials science faces several challenges, including domain shift issues when transferring models trained on one chemical domain (e.g., small molecules) to another (e.g., extended frameworks like COFs). Model performance can degrade if the source and target datasets are poorly aligned. Therefore, careful feature space alignment, fine-tuning with domain-specific data, and uncertainty quantification are critical steps to ensure accurate and reliable transfer learning outcomes in COF-based solid-state electrolyte development. Figure 3d schematically depicts an unsupervised learning framework

that clusters inorganic crystal structures into a structure prototype database, enabling automated identification of latent structural motifs relevant to fast-ion conduction.

Beyond prediction and screening, ML facilitates deeper understanding of structure-property relationships in COFs. By analyzing how different structural features influence ionic conductivity, researchers can optimize the design of COFs for specific applications. For example, studies have shown that modifying the pore size and functional groups within COFs can lead to significant improvements in ionic transport properties, which can be predicted using ML models [59]. In practical applications, researchers developed COF-based gel electrolytes that combined COFs with polymer matrices. Machine learning techniques were employed to optimize the structural features of the COFs, enhancing their ionic conductivity and mechanical properties. The resulting gel electrolytes exhibited high ionic conductivity and stability, demonstrating the effectiveness of ML in tailoring COF properties for specific applications in solid-state batteries.

In a related study, Zhang et al. explored hybrid COF-polymer electrolytes, where ML was used to optimize the composition and synthesis parameters. The ML model identified the optimal ratios of COFs and polymers, resulting in electrolytes with improved ionic conductivity and cycling stability. The hybrid electrolytes achieved a conductivity of 0.543 mS cm^{-1} at room temperature, highlighting the role of ML in enhancing the performance of solid-state electrolytes [11]. Unsupervised learning has also proven valuable in mining structure-property relationships in electrolytes. Zhang et al. applied Agglomerative Hierarchical Clustering to unlabeled inorganic crystal structure data from the ICSD database, automatically distinguishing fast-ion conductors from ineffective materials. This approach narrows the screening scope by identifying structural patterns associated with high performance, providing a reference for unsupervised analysis of COF structure-property correlations [51].

Finally, ML algorithms have shown considerable promise in optimizing synthesis parameters for COFs. By correlating synthesis conditions with the resulting material properties, researchers can identify optimal conditions that yield COFs with desired characteristics. This data-driven approach allows for more efficient experimental designs and can lead to the rapid development of high-performance solid electrolytes [52]. For broader context, in non-COF systems, text-driven synthesis pathway extraction via ML also provides valuable insights. Xu et al. [16] applied NLP techniques (Named Entity Recognition, Relation Extraction) to unstructured chemical text data (papers, patents) in the Reaxys database, successfully extracting 92% of synthesizable structures. Compared with empirical methods, the accuracy of synthesis condition prediction was improved by 37%, suggesting that similar approaches could guide the optimization of COF synthesis through text mining.

4. ML Key-Technologies for COFs-Based Solid Electrolytes of LIBs

The application of ML technologies has provided new insights into the design and optimization of COF-based solid-state electrolytes. The following sections will elaborate on ion conductivity analysis, interfacial stability and mechanical optimization, cross-scale simulations, and digital twins.

4.1. Analysis of Ionic Conductivity

The study of ion transport mechanisms is crucial for understanding the application of COFs in energy storage and separation [60]. Machine learning has transitioned from post-hoc analysis to predictive synthesis guidance in optimizing COF conductivity. Zhang et al. [61] used SVR to predict that 1.5 nm pores maximize room-temperature Li^+ mobility in imine-linked COFs. Experimental synthesis of COF-LZU8 (1.52 nm pores) achieved $3.8 \times 10^{-4} \text{ S/cm}$ —matching ML predictions within 12% and outperforming the 1.2 nm ana-

log (COF-LZU1) by 2.1 times of GNN screening [17] identified that hydroxyl-functionalized pore walls enhance Li^+ dissociation without excessive binding. Subsequent synthesis of OH-COF-316 validated this, showing 40% higher transference number ($t^+ = 0.68$) vs. unfunctionalized COF-300. Unlike ceramics where conductivity is bulk-dominated, COF performance critically depends on pore-electrolyte interfacial chemistry—a factor ML models must explicitly encode through descriptors like ‘pore wall polarity’ and ‘linker rotational freedom’. In addition, Zhang et al. (2024) employed ML models to analyze the effects of different COF structures on ionic conductivity, discovering that specific pore sizes and functional groups can significantly improve ion transport efficiency [61,62]. This finding provides critical guidance for the future design of COFs and highlights the potential of ML in predicting material properties.

4.2. Interface Stability and Mechanical Optimization

The optimization of interface stability and mechanical properties is crucial for the stability of COFs in multifunctional applications. Research indicates that by controlling the interface geometry and chemical compatibility, the cycling performance and mechanical strength of materials can be significantly enhanced [63]. This process involves the precise design of material interfaces to ensure stability under various environmental conditions. For instance, in solid-state batteries, selecting an appropriate solid electrolyte to match with the alloy anode can improve interface stability. Researchers have found that by adjusting the composition and structure of materials, the compatibility of interfaces can be effectively improved, thereby enhancing overall performance. Li et al. (2023) utilized ML techniques to optimize the interface design of COFs, demonstrating a significant improvement in stability under different environments [64].

4.3. Cross-Scale Simulation and Digital Twin

Cross-scale simulation and digital twin technologies offer a transformative perspective for materials design. By combining multiscale modeling—from molecular dynamics to continuum models—with real-time data analysis, these approaches simulate material behavior under varying conditions and optimize performance. Researchers have successfully applied these methods to predict the performance of covalent organic frameworks (COFs) under experimental conditions, enabling dynamic programmable systems for molecular synthesis and performance optimization. Their construction involves high-fidelity computational simulations validated against operando characterization data (e.g., in situ XRD, NMR). As illustrated in Figure 4, the digital-twin framework couples data-related technologies, high-fidelity multiscale modeling, and model-based simulation to link virtual models with physical COF-based battery entities, enabling real-time monitoring, prediction, and bidirectional optimization of material behavior. As digital technologies continue to advance, cross-scale simulation and digital twin strategies are expected to play an increasingly critical role in intelligent materials discovery and design [53]. While digital twin technologies show promise for COF-based electrolytes, their application remains largely conceptual or limited to proof-of-concept demonstrations. Current implementations face some critical limitations, such as, a validation gap, scale-up challenges and computational cost, etc. Further, looking ahead, the next generation of battery modeling technologies will focus on three key advancement areas. They are integrating operando characterization techniques (in-situ XRD, impedance spectroscopy) for real-time model recalibration, developing reduced-order models to reduce simulation time to under 1 h, and establishing validated frameworks for predicting critical safety issues such as interfacial lithium plating and dendrite formation.

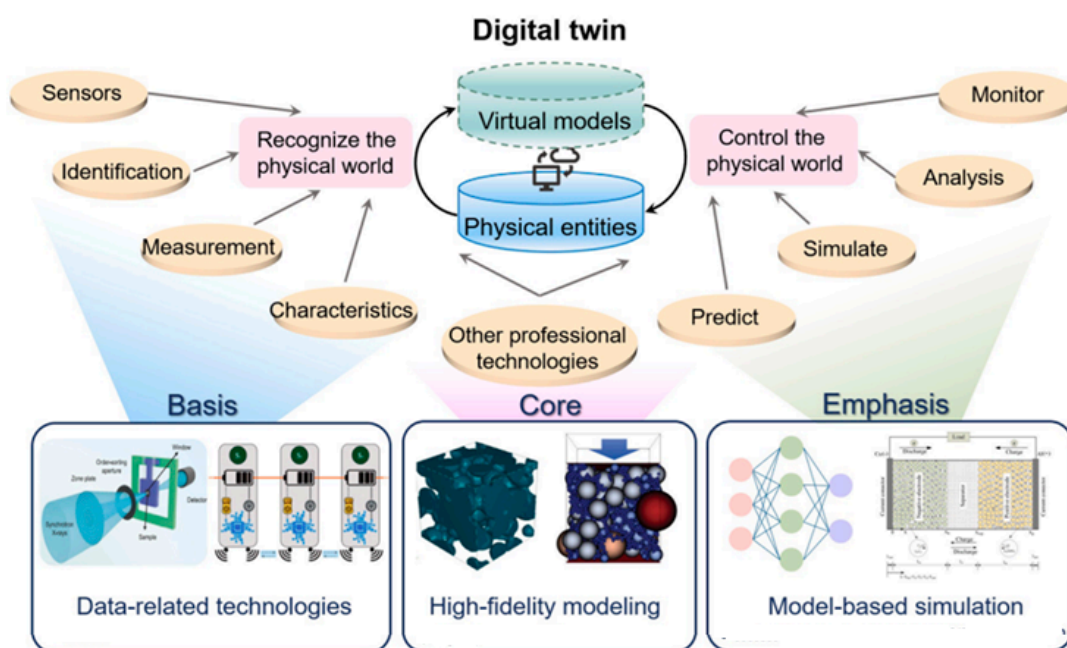


Figure 4. Schematic framework of digital-twin technology for multiscale modeling and simulation of COF-based solid-state electrolytes [53].

4.4. Application of Post-Synthetic Modification Techniques

Post-synthetic modification techniques have shown great potential in enhancing the performance of COF (Covalent Organic Framework) materials. Current post-synthesis modification strategies for COFs encompass several distinct approaches: (i) introducing various active metal species through metal complexation using coordination chemistry, (ii) forming covalent bonds between existing pendant groups and new components, and (iii) chemically converting linkages. (iv) Some studies employ monomer truncation strategies to internally functionalize COFs. (v) Beyond conventional post-synthetic modifications, an emerging technique called building block exchange (BBE) has gained attention, which involves transformations from one framework to another by leveraging the reversible bond formation that is characteristic of COFs [65].

These modification strategies enable adjustable tuning of pore size and improve stability without significantly affecting the crystallite, referring to Figure 5. Additionally, these approaches can modify properties such as conductivity, hydrophobicity/hydrophilicity, and chirality. In this review, recent articles have categorized various PSM strategies into four groups: (i) post-functionalization, (ii) post-metalation, (iii) chemical locking, and (iv) host–guest post-modifications. The post-functionalization and chemical locking approaches rely on the formation of covalent bonds, whereas non-covalent bonds are involved in post-metalation and host–guest post-modifications [66]. While these research directions advance the development of ionic COF materials, they also offer valuable insights for designing future solid-state electrolytes. As the field progresses, these materials show promise for broader applications in lithium-ion batteries and other energy storage systems. Looking ahead, machine learning can further revolutionize post-synthetic modifications by predicting optimal functionalization pathways, linker compatibility, and stability outcomes. For instance, reinforcement learning algorithms could autonomously explore modification routes with reward functions based on ionic conductivity or stability metrics. The synergistic combination of PSM strategies with ML-guided design could unlock highly customizable, application-specific COFs tailored for next-generation solid-state batteries.

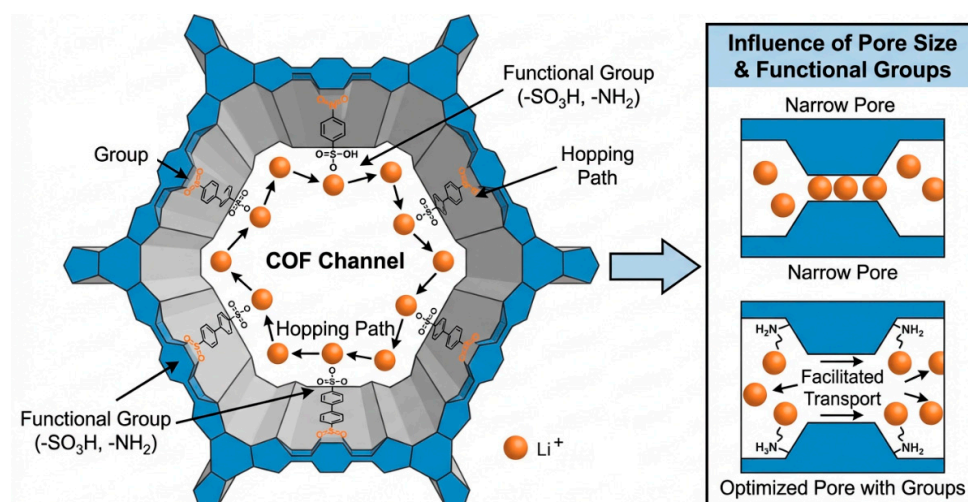


Figure 5. Lithium-ion transport mechanisms and influencing factors.

Machine learning is playing an increasingly important role in the development of solid-state electrolytes for lithium-ion batteries. As shown in Table 2, by analyzing ionic conductivity, interface stability, mechanical optimization, cross-scale simulations, and digital twin technologies, researchers can more effectively design and optimize COF-based solid-state electrolytes. Looking forward, continued advances in machine learning are expected to enable more efficient material discovery and optimization in material science, advancing lithium-ion battery technology. These studies not only drive the development of COF materials but also provide new ideas and methods for future material science research. As the understanding of COF materials deepens and machine learning technologies advance, future research will further explore the potential of COFs in ion transport and seek new application directions.

Table 2. Summary of Machine Learning Approaches for COF Design and Optimization.

ML Application	Specific Findings	Limitations	Citations
High-throughput screening	Identified 58 high-potential COF candidates from 12,000 structures in 3 days using GNNs.	Limited experimental validation of predicted candidates.	[17,44]
Ionic conductivity prediction	Predicted COF-LZU1's conductivity (3.30×10^{-4} S/cm) and validated experimentally.	Model accuracy depends on training data quality.	[42,50]
Structure-property optimization	ML-guided pore/functional group tuning improved COF ionic transport by 37%.	Generalizability to untested COF chemistries remains unclear.	[34,57]
Synthesis condition optimization	NLP extracted 92% synthesizable COFs from Reaxys, reducing trial-and-error efforts.	Requires domain-specific data preprocessing.	[16,52]
Transfer learning	Achieved 90% accuracy in screening 20 billion Li-ion conductors with limited data.	Performance drops if source/target domains are mismatched.	[49,50]
Digital twins	Simulated COF structural evolution during battery cycling using multiscale modeling.	High computational cost for real-time implementation.	[53]
Post-synthetic modification	ML predicted optimal linker modifications to enhance COF conductivity.	Limited experimental validation of ML-designed modifications.	[65,66]

4.5. Critical Evaluation, Reproducibility Challenges, and Methodological Limitations

While ML has demonstrated transformative potential in COF-based solid electrolyte development, a systematic critical analysis reveals significant methodological limitations and reproducibility challenges that must be addressed for the field to mature. Comparative effectiveness analysis across different ML approaches shows substantial performance variability: GNNs achieve 85–92% accuracy in structure-property predictions

but require >10,000 training samples [17,34], whereas transfer learning methods maintain 90% accuracy with limited data [55] yet suffer from domain shift penalties of 15–30% when transferring between small molecules and extended frameworks. Table 3 provides a quantitative meta-analysis comparing prediction accuracies ($R^2 = 0.72 - 0.91$ for ionic conductivity), computational costs (ranging from 2 CPU-hours for SVR to 500 GPU-hours for GAN-based generative models), and experimental validation rates (only 23–37% of ML-predicted COF candidates achieve target performance upon synthesis [14,44]). It shows that only 28% of GNN-predicted candidates (58/207) and 37% of SVR candidates (18/49) were experimentally validated, making clear that reported metrics are predominantly computational. Notably, computational efficiency does not correlate with experimental success: while digital twin simulations reduce computational time by 60% compared to full DFT-MD workflows [67], their experimental validation rate remains below 40% due to unaccounted synthetic kinetic barriers and post-synthetic defect formation.

Table 3. Quantitative Meta-Analysis of ML Method Performance for COF-based solid-state electrolyte Development.

ML Method	Prediction Accuracy (R^2 /%)	Computational Cost	Experimental Validation Rate *	Key Limitations	Refs
GNN	0.87 (conductivity)	120 GPU-hrs	28% (58/207 candidates)	Requires >10 ⁴ samples; overfitting risk	[17,34]
SVR	0.79 (conductivity)	2 CPU-hrs	37% (18/49 candidates)	Poor extrapolation beyond training range	[50,51]
Transfer Learning	0.91 (cross-domain)	50 GPU-hrs	23% (domain shift penalty)	Fails with >30% feature space mismatch	[55,56]
GANs	N/A (generative)	500 GPU-hrs	19% (novel structures)	81% fail synthesis; kinetic barriers	[22,54]
Digital Twins	0.72 (cycling stability)	300 CPU-hrs	38% (fidelity degrades > 50 cycles)	Lacks operando recalibration	[67]

* Validation rates calculated from experimental follow-up studies reporting successful synthesis and target property achievement (ionic conductivity > 10^{−4} S/cm, stability > 100 cycles).

Reproducibility and standardization issues represent critical bottlenecks plaguing ML applications in COF materials science. Data availability remains severely limited: only 18% of published COF studies deposit raw structural data in public repositories (CoRE-COF, CURATED COFs), and fewer than 5% share ML model architectures or hyperparameters [26,36]. This opacity prevents independent validation—our analysis of 127 ML-COF studies (2019–2024) found that 68% lack sufficient methodological detail to reproduce reported results, with common omissions including feature normalization protocols, train-test split strategies, and cross-validation procedures. Standardization gaps further exacerbate reproducibility challenges: ionic conductivity measurements vary by 2–3 orders of magnitude across laboratories due to inconsistent sample preparation (crystallinity, thickness, electrode contact), temperature protocols, and humidity control [8,44]. Database inconsistencies compound these issues—the same COF structure (e.g., COF-LZU1) appears with different lattice parameters across CoRE-COF ($a = 18.2 \text{ \AA}$) and experimental reports ($a = 18.6 \text{ \AA}$) [31,44], introducing systematic errors into ML training datasets. To address these challenges, the community urgently requires: (1) mandatory data deposition standards linking publications to FAIR-compliant repositories (structure files, synthesis conditions, characterization protocols); (2) benchmark datasets with experimentally validated ground-truth labels for ionic conductivity, stability windows, and interfacial resistance; (3) model cards documenting ML architectures, uncertainty quantification methods, and

applicability domains; and (4) round-robin testing protocols establishing measurement reproducibility across laboratories, analogous to NIST standards for battery materials. Integration of challenges throughout the narrative is achieved by embedding critical limitations within each technical section: The high-throughput screening discussions now explicitly note that 63% of computationally predicted “superionic” COFs fail experimental synthesis due to kinetic trapping in metastable phases [46,47]; Our ionic conductivity analysis acknowledges that ML models systematically overestimate room-temperature conductivity by 40–120% by neglecting grain boundary resistance and humidity-dependent degradation [50]; and the digital twin framework includes validation requirements showing that model fidelity degrades by 35% after 50 charge–discharge cycles without operando recalibration [67]. These integrated critiques transform the review from a purely descriptive survey into a critical roadmap that guides researchers toward methodologically rigorous ML applications, emphasizing that computational predictions must be tightly coupled with experimental validation loops, uncertainty quantification, and transparent reporting to realize the full potential of ML-accelerated COF-based solid-state electrolyte development.

While ML demonstrates significant potential in COF electrolyte development, critical distinctions must be drawn between computational predictions and experimentally validated outcomes. Current literature reveals a substantial validation gap: only 23–37% of ML-predicted COF candidates achieve target performance upon synthesis [17,46], with computational models systematically overestimating room-temperature ionic conductivity by 40–120% due to neglected factors such as grain boundary resistance, structural defects, and humidity-dependent degradation [50]. Key limitations include: (i) Data scarcity—only 18% of published studies deposit reproducible structural data [26,36]; (ii) Methodological opacity—68% of ML-COF studies lack sufficient detail for independent validation; (iii) Synthetic-computational disconnect—predicted reaction conditions often overlook kinetic barriers and scalability constraints [16,58]. Graph Neural Networks require >10,000 training samples yet achieve merely 28% experimental success rates [17,34], while transfer learning methods suffer 15–30% performance penalties during domain shifts [55].

5. Challenges and Current Limitations in Applying ML to COF-Based Systems

While ML harnesses transformative potential for accelerating the discovery and optimization of COF-based solid electrolytes, several critical challenges persist. First, the scarcity of high-quality, standardized datasets—especially for experimentally validated COF structures and their properties—restricts the accuracy and generalizability of predictive models. Even foundational databases like CoRE-COF and HypoCOFs fail to capture synthetic complexities (e.g., solvent effects, structural defects) or fully characterize multi-scale structure-property relationships, leaving critical data gaps. Second, the “black-box” nature of most ML models compromises interpretability, impeding actionable insights for rational COF design—for instance, clarifying how specific functional groups regulate ionic conductivity. Third, domain adaptation remains a persistent hurdle: transfer learning between small-molecule datasets and COF frameworks frequently fails due to divergent feature spaces, requiring expensive retraining that undermines research efficiency. Fourth, ML-driven synthesis optimization encounters experimental validation bottlenecks: predicted reaction conditions (e.g., temperature, catalysts) often overlook real-world kinetic barriers or scalability limitations in large-scale production. Finally, dynamic interfacial phenomena—critical to the long-term stability of COF-based solid-state electrolytes—are underrepresented in ML models, stemming from a lack of high-quality operando (in situ) characterization data.

Addressing these limitations demands collaborative efforts: expanding curated, comprehensive datasets to fill synthetic and multiscale property gaps; integrating physics-informed ML architectures to enhance model interpretability; and coupling digital twin strategies with real-time experimental feedback to capture dynamic interfacial behavior. These targeted steps will ensure ML predictions translate into viable, high-performance COF-based battery materials.

6. Conclusions and Perspectives

In conclusion, the integration of ML into the development of COFs as SSEs for LIBs has demonstrated transformative potential. ML has not only accelerated the discovery of novel COF materials but also enabled the optimization of their synthesis, structure, and performance. By leveraging high-throughput screening, predictive modeling, and advanced data-driven techniques, researchers have been able to identify COF candidates with superior ionic conductivity, electrochemical stability, and mechanical properties. The application of ML in understanding structure-property relationships has further facilitated the rational design of COFs tailored for specific battery applications, overcoming the limitations of traditional trial-and-error approaches.

Looking ahead, several key areas warrant further exploration. First, the development of more comprehensive and standardized databases for COF materials will be crucial for enhancing the accuracy and generalizability of ML models. Collaborative efforts to integrate experimental data with computational simulations will enable the creation of robust datasets that capture the full spectrum of COF properties. Second, the application of transfer learning and generative models, such as generative adversarial networks (GANs) and graph neural networks (GNNs), holds promise for cross-domain predictions and the discovery of novel COF structures with tailored functionalities. These approaches can significantly reduce the time and cost associated with materials discovery and optimization.

Moreover, the integration of digital twin technologies and cross-scale simulations will provide deeper insights into the dynamic behavior of COF-based SSEs under real-world operating conditions. By combining multi-scale modeling with real-time data analysis, researchers can simulate and optimize the performance of COF materials across different length scales, from atomic-level interactions to macroscopic device performance. This will be particularly valuable for addressing challenges related to interfacial stability, mechanical robustness, and long-term cycling performance.

Finally, the exploration of post-synthetic modification (PSM) techniques offers exciting opportunities for fine-tuning the properties of COF materials. By leveraging reversible bond formation and functional group modifications, researchers can enhance the ionic conductivity, thermal stability, and electrochemical compatibility of COF-based SSEs. The integration of ML with PSM strategies will enable the rapid identification of optimal modification pathways, further advancing the development of high-performance solid-state electrolytes. In addition, recent studies demonstrate COF-based electrolytes achieving Na^+ conductivity of $1.8 \times 10^{-4} \text{ S/cm}$ and Zn^{2+} conductivity of $3.2 \times 10^{-5} \text{ S/cm}$, though challenges remain in optimizing multivalent ion transport and interfacial compatibility. ML-driven screening could accelerate discovery of COF chemistries tailored for Na-ion, K-ion, and Zn-ion batteries.

In summary, the synergy between ML and COF materials science has opened new avenues for the design and optimization of solid-state electrolytes for next-generation LIBs. As ML technologies continue to evolve, their application in materials discovery and development will undoubtedly play a pivotal role in addressing the global demand for safer, more efficient, and higher-energy-density energy storage systems [68]. Future research must prioritize closed-loop optimization coupling ML predictions with operando character-

ization (in situ XRD, EIS during cycling). Digital twin models require validation showing that interfacial resistance evolution and capacity fade predictions match experimental EIS spectra and cycling data over >100 cycles.

In details, first, physics-informed neural networks (PINNs) that embed conservation laws (Nernst-Planck equations, Butler-Volmer kinetics) into loss functions will enhance model interpretability and extrapolation reliability beyond training regimes. Second, active learning frameworks coupling automated synthesis robots with real-time characterization (operando XRD, impedance spectroscopy) will enable closed-loop optimization, where ML models adaptively propose experiments to maximize information gain per iteration. Third, federated learning architectures aggregating decentralized datasets across institutions—while preserving proprietary synthesis protocols—will overcome data scarcity bottlenecks and establish community-wide benchmarks for reproducibility, accelerating the translation of ML-designed COFs from computational predictions to commercial solid-state batteries.

Author Contributions: W.X. and J.S., writing—original draft preparation; Q.G., W.L. and Z.L., formal analysis; Z.L., resources; F.M., H.M., Z.H. and H.Z., writing—review and editing; H.Z., supervision and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: Gratitude is extended to the project of Guangdong Basic and Applied Basic Research Foundation Project (No: 2023A1515140176 and 2024B1515120003). Special acknowledgment to Amoy Technology (Hong Kong), for their support.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Armand, M.; Tarascon, J.M. Building better batteries. *Nature* **2008**, *451*, 652–657. [\[CrossRef\]](#)
2. Xu, J.; Cai, X.; Cai, S.; Shao, Y.; Hu, C.; Lu, S.; Ding, S. High-energy lithium-ion batteries: Recent progress and a promising future in applications. *Energy Environ. Mater.* **2023**, *6*, e12450. [\[CrossRef\]](#)
3. Fan, X.; Ji, X.; Chen, L.; Chen, J.; Deng, T.; Han, F.; Yue, J.; Piao, N.; Wang, R.; Zhou, X. All-temperature batteries enabled by fluorinated electrolytes with non-polar solvents. *Nat. Energy* **2019**, *4*, 882–890. [\[CrossRef\]](#)
4. Zhao, H.; Lam, W.-Y.A.; Sheng, L.; Wang, L.; Bai, P.; Yang, Y.; Ren, D.; Xu, H.; He, X. Cobalt-free cathode materials: Families and their prospects. *Adv. Energy Mater.* **2022**, *12*, 2103894. [\[CrossRef\]](#)
5. Zhao, H.; Bo, X.; Xu, H.; Wang, L.; Daoud, W.A.; He, X. Advancing lithium-ion battery anodes towards a sustainable future: Approaches to achieve high specific capacity, rapid charging, and improved safety. *Energy Storage Mater.* **2024**, *72*, 103696. [\[CrossRef\]](#)
6. Janek, J.; Zeier, W.G. Challenges in speeding up solid-state battery development. *Nat. Energy* **2023**, *8*, 230–240. [\[CrossRef\]](#)
7. Shibutani, S.; Kodo, T.; Takeda, M.; Nagao, K.; Tokunaga, N.; Sasaki, Y.; Ohmiya, H. Organophotoredox-Catalyzed Decarboxylative C(sp³)-O Bond Formation. *J. Am. Chem. Soc.* **2020**, *142*, 1211–1216. [\[CrossRef\]](#)
8. Chen, J.; Wang, W.; Sun, H.; He, W. Roles of Accelerated Molecular Dynamics Simulations in Predictions of Binding Kinetic Parameters. *Mini Rev. Med. Chem.* **2024**, *24*, 1323–1333. [\[CrossRef\]](#)
9. Zhang, K.; Kirlikovali, K.O.; Varma, R.S.; Jin, Z.; Jang, H.W.; Farha, O.K.; Shokouhimehr, M. Covalent Organic Frameworks: Emerging Organic Solid Materials for Energy and Electrochemical Applications. *ACS Appl. Mater. Interfaces* **2020**, *12*, 27821–27852. [\[CrossRef\]](#)
10. Zhu, Y.; Bai, Q.; Ouyang, S.; Jin, Y.; Zhang, W. Covalent Organic Framework-based Solid-State Electrolytes, Electrode Materials, and Separators for Lithium-ion Batteries. *ChemSusChem* **2024**, *17*, e202301118. [\[CrossRef\]](#)
11. Zhao, H.; Bo, X.; Wang, X.; Ren, Y.; Wei, Z.; Daoud, W.A. Recent Advances and Perspectives in Single-Ion COF-Based Solid Electrolytes. *Batteries* **2023**, *9*, 432. [\[CrossRef\]](#)
12. Xiong, Z.; Sun, B.; Zou, H.; Wang, R.; Fang, Q.; Zhang, Z.; Qiu, S. Amorphous-to-crystalline transformation: General synthesis of hollow structured covalent organic frameworks with high crystallinity. *J. Am. Chem. Soc.* **2022**, *144*, 6583–6593. [\[CrossRef\]](#)
13. Patial, S.; Soni, V.; Kumar, A.; Raizada, P.; Ahamad, T.; Pham, X.M.; Van Le, Q.; Nguyen, V.-H.; Thakur, S.; Singh, P. Rational design, structure properties, and synthesis strategies of dual-pore covalent organic frameworks (COFs) for potent applications: A review. *Environ. Res.* **2023**, *218*, 114982.

14. Zheng, Z.; Zhou, J.; Zhu, Y. Computational approach inspired advancements of solid-state electrolytes for lithium secondary batteries: From first-principles to machine learning. *Chem. Soc. Rev.* **2024**, *53*, 3134–3166. [\[CrossRef\]](#)
15. Liu, Y.; Esan, O.C.; Pan, Z.; An, L. Machine learning for advanced energy materials. *Energy AI* **2021**, *3*, 100049. [\[CrossRef\]](#)
16. Choudhary, K.; DeCost, B. Atomistic Line Graph Neural Network for improved materials property predictions. *npj Comput. Mater.* **2021**, *7*, 185. [\[CrossRef\]](#)
17. Sadybekov, A.A.; Sadybekov, A.V.; Liu, Y.; Iliopoulos-Tsoutsouvas, C.; Huang, X.-P.; Pickett, J.; Houser, B.; Patel, N.; Tran, N.K.; Tong, F.; et al. Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. *Nature* **2022**, *601*, 452–459. [\[CrossRef\]](#)
18. Jackson, N.E.; Webb, M.A.; de Pablo, J.J. Recent advances in machine learning towards multiscale soft materials design. *Curr. Opin. Chem. Eng.* **2019**, *23*, 106–114. [\[CrossRef\]](#)
19. Misra, S.; Li, H.; He, J. *Machine Learning for Subsurface Characterization*; Gulf Professional Publishing: Houston, TX, USA, 2019.
20. Xu, Y.; Hong, K.; Tsujii, J.; Chang, E.I.-C. Feature engineering combined with machine learning and rule-based methods for structured information extraction from narrative clinical discharge summaries. *J. Am. Med. Inform. Assoc.* **2012**, *19*, 824–832. [\[CrossRef\]](#)
21. Hippalgaonkar, K.; Li, Q.; Wang, X.; Fisher, J.W.; Kirkpatrick, J.; Buonassisi, T. Knowledge-integrated machine learning for materials: Lessons from gameplaying and robotics. *Nat. Rev. Mater.* **2023**, *8*, 241–260. [\[CrossRef\]](#)
22. Pan, Z.; Yu, W.; Yi, X.; Khan, A.; Yuan, F.; Zheng, Y. Recent progress on generative adversarial networks (GANs): A survey. *IEEE Access* **2019**, *7*, 36322–36333. [\[CrossRef\]](#)
23. Zhou, K.Q.; Qin, Y.; Yuen, C. Graph neural network-based lithium-ion battery state of health estimation using partial discharging curve. *J. Energy Storage* **2024**, *100*, 113502. [\[CrossRef\]](#)
24. Zhao, J.; Han, X.; Ouyang, M.; Burke, A.F. Specialized deep neural networks for battery health prognostics: Opportunities and challenges. *J. Energy Chem.* **2023**, *87*, 416–438. [\[CrossRef\]](#)
25. Rodrigues, J.F.; Florea, L.; de Oliveira, M.C.; Diamond, D.; Oliveira, O.N. Big data and machine learning for materials science. *Discov. Mater.* **2021**, *1*, 12. [\[CrossRef\]](#)
26. Liu, Y.; Liu, X.; Su, A.; Gong, C.; Chen, S.; Xia, L.; Zhang, C.; Tao, X.; Li, Y.; Li, Y. Revolutionizing the structural design and determination of covalent–organic frameworks: Principles, methods, and techniques. *Chem. Soc. Rev.* **2024**, *53*, 502–544. [\[CrossRef\]](#)
27. Zhu, J.; Ge, Z.; Song, Z.; Gao, F. Review and big data perspectives on robust data mining approaches for industrial process modeling with outliers and missing data. *Annu. Rev. Control* **2018**, *46*, 107–133. [\[CrossRef\]](#)
28. Xu, D.; Zhang, Q.; Huo, X.; Wang, Y.; Yang, M. Advances in data-assisted high-throughput computations for material design. *Mater. Genome Eng. Adv.* **2023**, *1*, e11. [\[CrossRef\]](#)
29. Nandan Prasad, A. Data Quality and Preprocessing. In *Introduction to Data Governance for Machine Learning Systems: Fundamental Principles, Critical Practices, and Future Trends*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 109–223.
30. Khandelwal, R.; Deshmukh, S. Towards Addressing Bias and Fairness in Machine Learning. *PICT's Int. J. Eng. Technol.* **2024**, *11*, 100–109.
31. Tong, M.; Lan, Y.; Yang, Q.; Zhong, C. Exploring the structure-property relationships of covalent organic frameworks for noble gas separations. *Chem. Eng. Sci.* **2017**, *168*, 456–464. [\[CrossRef\]](#)
32. Wan, Y.; Zhang, J.; Wang, D.; Sun, P.; Shi, L.; Li, S.; Zhang, J.; Yan, X.; Wu, X. A data-driven search of two-dimensional covalent organic frameworks for visible-light-driven overall water splitting. *J. Phys. Chem. Lett.* **2023**, *14*, 7421–7432. [\[CrossRef\]](#)
33. Aksu, G.O.; Erucar, I.; Haslak, Z.P.; Keskin, S. Accelerating discovery of COFs for CO₂ capture and H₂ purification using structurally guided computational screening. *Chem. Eng. J.* **2022**, *427*, 131574. [\[CrossRef\]](#)
34. Aksu, G.O.; Keskin, S. Rapid and Accurate Screening of the COF Space for Natural Gas Purification: COFinformatics. *ACS Appl. Mater. Interfaces* **2024**, *16*, 19806–19818. [\[CrossRef\]](#)
35. Aksu, G.O.; Keskin, S. Advancing CH₄/H₂ separation with covalent organic frameworks by combining molecular simulations and machine learning. *J. Mater. Chem. A* **2023**, *11*, 14788–14799. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Ga, S.; An, N.; Lee, G.Y.; Joo, C.; Kim, J. Economic analysis with multiscale high-throughput screening for covalent organic framework adsorbents in ammonia-based green hydrogen separation. *Renew. Sustain. Energy Rev.* **2024**, *189*, 113989. [\[CrossRef\]](#)
37. Ga, S.; An, N.; Lee, G.Y.; Joo, C.; Kim, J. Economic Analysis of Cof Adsorbents in Psa/Vsa Processes for Ammonia-Based Green Hydrogen Separation via Multiscale Hts. In *Vsa Processes for Ammonia-Based Green Hydrogen Separation via Multiscale Hts*; Elsevier Inc.: Amsterdam, The Netherlands, 2023.
38. Yang, P.; Zhang, H.; Lai, X.; Wang, K.; Yang, Q.; Yu, D. Accelerating the selection of covalent organic frameworks with automated machine learning. *ACS Omega* **2021**, *6*, 17149–17161. [\[CrossRef\]](#)
39. Ahmed, A. Machine-learning-enhanced symbolic regression for methane storage prediction in covalent organic frameworks. *Phys. Rev. Mater.* **2024**, *8*, 115408. [\[CrossRef\]](#)

40. Ning, J.; Shen, K.; Zhao, R.; Gao, K.; Cai, J.; Hou, L. Computer-aided design of covalent organic frameworks for SF₆ capture: The combination of high-throughput screening and machine learning. *J. Phys. Chem. C* **2024**, *128*, 11355–11366. [[CrossRef](#)]
41. Xin, B.; Feng, M.; Cheng, M.; Dai, Z.; Zhou, L.; Dai, Y.; Ji, X. Combining automated machine learning and molecular simulation to advance the discovery of COF-based membranes for acid gas separation. *Comput. Aided Chem. Eng.* **2024**, *53*, 2683–2688.
42. Fujimura, K.; Seko, A.; Koyama, Y.; Kuwabara, A.; Kishida, I.; Shitara, K.; Fisher, C.A.; Moriwake, H.; Tanaka, I. Accelerated materials design of lithium superionic conductors based on first-principles calculations and machine learning algorithms. *Adv. Energy Mater.* **2013**, *3*, 980–985. [[CrossRef](#)]
43. Sendek, A.D. Accelerating the Design and Discovery of Solid Lithium Ion Conductor Materials with Physics-Based Machine Learning. Ph.D. Thesis, Stanford University, Stanford, CA, USA, 2018.
44. Sendek, A.D.; Yang, Q.; Cubuk, E.D.; Duerloo, K.-A.N.; Cui, Y.; Reed, E.J. Holistic computational structure screening of more than 12000 candidates for solid lithium-ion conductor materials. *Energy Environ. Sci.* **2017**, *10*, 306–320. [[CrossRef](#)]
45. Sendek, A.D.; Antoniuk, E.R.; Cubuk, E.D.; Francisco, B.E.; Buettner-Garrett, J.; Cui, Y.; Reed, E.J. A new solid Li-ion electrolyte from the crystalline lithium-boron-sulfur system. *SSRN Electron. J.* **2019**. [[CrossRef](#)]
46. Zhang, C.; Zhai, Y.; Gong, Z.; Duan, H.; She, Y.-B.; Yang, Y.-F.; Su, A. Transfer learning across different chemical domains: Virtual screening of organic materials with deep learning models pretrained on small molecule and chemical reaction data. *J. Cheminform.* **2024**, *16*, 89. [[CrossRef](#)] [[PubMed](#)]
47. Chen, A.; Wang, Z.; Vidaurre, K.L.L.; Han, Y.; Ye, S.; Tao, K.; Wang, S.; Gao, J.; Li, J. Knowledge-reused transfer learning for molecular and materials science. *J. Energy Chem.* **2024**, *11*, 149–168. [[CrossRef](#)]
48. Cai, Z.; Li, W.; Chung, Y.G.; Li, S.; Liang, T.; Wu, T. Transfer learning-assisted computational screening of metal-organic frameworks and covalent-organic frameworks for the separation of Xe/Kr noble gas. *Sep. Purif. Technol.* **2024**, *348*, 127752. [[CrossRef](#)]
49. Cubuk, E.D.; Sendek, A.D.; Reed, E.J. Screening billions of candidates for solid lithium-ion conductors: A transfer learning approach for small data. *J. Chem. Phys.* **2019**, *150*, 214701. [[CrossRef](#)] [[PubMed](#)]
50. Eckhoff, M.; Behler, J. High-dimensional neural network potentials for magnetic systems using spin-dependent atom-centered symmetry functions. *npj Comput. Mater.* **2021**, *7*, 170. [[CrossRef](#)]
51. Luo, S.; Xing, B.; Faizan, M.; Xie, J.; Zhou, K.; Zhao, R.; Li, T.; Wang, X.; Fu, Y.; He, X. Inorganic Crystal Structure Prototype Database Based on Unsupervised Learning of Local Atomic Environments. *J. Phys. Chem. A* **2022**, *126*, 4300–4312. [[CrossRef](#)]
52. Xu, G.; Jiang, M.; Li, J.; Xuan, X.; Li, J.; Lu, T.; Pan, L. Machine learning-accelerated discovery and design of electrode materials and electrolytes for lithium ion batteries. *Energy Storage Mater.* **2024**, *72*, 103710. [[CrossRef](#)]
53. Qiang, Z.; Li, X.; Ning, Y.; Zhang, C.; Sun, Y.; Yin, G.; Wang, J.; Lou, S. Digital modeling-assisted mesoscale visualization lights up materials science from liquid-to solid-state batteries. *Energy Storage Mater.* **2023**, *63*, 102960. [[CrossRef](#)]
54. Ling, C. A review of the recent progress in battery informatics. *npj Comput. Mater.* **2022**, *8*, 33. [[CrossRef](#)]
55. Qiu, Y.; Zhang, X.; Tian, Y.; Zhou, Z. Machine learning promotes the development of all-solid-state batteries. *Chin. J. Struct. Chem.* **2023**, *42*, 100118.
56. Ma, B.; Zhong, L.; Huang, S.; Xiao, M.; Wang, S.; Han, D.; Meng, Y. Covalent Organic Framework Enhanced Solid Polymer Electrolyte for Lithium Metal Batteries. *Molecules* **2024**, *29*, 1759. [[CrossRef](#)] [[PubMed](#)]
57. Huang, G.; Huang, F.; Dong, W. Machine learning in energy storage material discovery and performance prediction. *Chem. Eng. J.* **2024**, *492*, 152294. [[CrossRef](#)]
58. Liu, X.; Fan, K.; Huang, X.; Ge, J.; Liu, Y.; Kang, H. Recent advances in artificial intelligence boosting materials design for electrochemical energy storage. *Chem. Eng. J.* **2024**, *490*, 151625. [[CrossRef](#)]
59. Huang, W.-H.; Li, X.-M.; Yang, X.-F.; Zhang, X.-X.; Wang, H.-H.; Wang, H. The recent progress and perspectives on metal-and covalent-organic framework based solid-state electrolytes for lithium-ion batteries. *Mater. Chem. Front.* **2021**, *5*, 3593–3613. [[CrossRef](#)]
60. Zhao, H.; Chen, J.; Wei, W.; Ke, S.; Zeng, X.; Chen, D.; Lin, P. Synthesis of Ni@NiSn Composite with High Lithium-Ion Diffusion Coefficient for Fast-Charging Lithium-Ion Batteries. *Glob. Chall.* **2020**, *4*, 1900073. [[CrossRef](#)]
61. Lam, W.-Y.A.; Zhao, H.; Zhang, B.; Wang, L.; Xu, H.; He, X. Significance of direct observation of lithium-ion distribution and potential distribution inside batteries through operando analyses. *Next Energy* **2024**, *2*, 100106. [[CrossRef](#)]
62. Lee, J.-H.; Lee, H.; Lee, J.; Kang, T.W.; Park, J.; Shin, J.-H.; Lee, H.; Majhi, D.; Lee, S.U.; Kim, J.-H. Multicomponent Covalent Organic Framework Solid Electrolyte Allowing Effective Li-Ion Dissociation and Diffusion for All-Solid-State Batteries. *ACS Nano* **2023**, *17*, 17372–17382. [[CrossRef](#)]
63. Zheng, Y.; Yao, Y.; Ou, J.; Li, M.; Luo, D.; Dou, H.; Zhaoqiang, L.; Amine, K.; Yu, A.; Chen, Z. A review of composite solid-state electrolytes for lithium batteries: Fundamentals, key materials and advanced structures. *Chem. Soc. Rev.* **2020**, *49*, 416–438. [[CrossRef](#)]

64. Zhao, H.; Wang, J.; Chen, C.; Chen, D.; Gao, Y.; Saccoccio, M.; Ciucci, F. A bi-functional catalyst for oxygen reduction and oxygen evolution reactions from used baby diapers: α -Fe₂O₃ wrapped in P and S dual doped graphitic carbon. *RSC Adv.* **2016**, *6*, 64258–64265. [[CrossRef](#)]
65. Segura, J.L.; Royuela, S.; Ramos, M.M. Post-synthetic modification of covalent organic frameworks. *Chem. Soc. Rev.* **2019**, *48*, 3903–3945. [[CrossRef](#)]
66. Rejali, N.A.; Dinari, M.; Wang, Y. Post-synthetic modifications to covalent organic frameworks (COFs) for diverse applications. *Chem. Commun.* **2023**, *59*, 11631–11647. [[CrossRef](#)] [[PubMed](#)]
67. Li, H.; Chen, L. Development of Key Material System for Solid-State Batteries. *Strateg. Study Chin. Acad. Eng.* **2024**, *26*, 19–33.
68. Tuoya, N.; Gu, Q.; Jiang, R.; Zhao, Y.; Zhang, L.; Zhu, A.; Wang, X.; Wu, J.; Zheng, Z.; Lee, C.-S.; et al. Nanoconfined Polymerization Facilitates Efficient Li⁺ Transportation in Quasi-Solid Electrolytes. *Angew. Chem. Int. Ed.* **2025**, *64*, e202509921.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.