Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.080623

REVIEW

Machine Learning-Driven Materials Design and Performance Prediction in Organic Solar Cells Emphasizing Ensemble Learning Models

Shafidah Shafian[1,*] **and Azlan Ismail**[2,3]

[1]Solar Energy Research Institute, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
[2]Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), Kompleks Al-Khawarizmi, Universiti Teknologi MARA (UiTM), Shah Alam, Selangor, Malaysia
[3]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM), Shah Alam, Selangor, Malaysia

*Corresponding Author: Shafidah Shafian. Email: norshafidah@ukm.edu.my

Received: 13 February 2026; Accepted: 29 April 2026; Published: 15 June 2026

ABSTRACT: Organic solar cells (OSCs) have progressed rapidly in recent years, driven by advances in donor polymers, non-fullerene acceptors, and increasingly complex binary and multicomponent blend architectures. Despite these achievements, device performance remains governed by strongly coupled molecular, morphological, and processing variables, making materials optimization inherently multidimensional and difficult to navigate using conventional trial-and-error approaches. The growing availability of experimental data and computational descriptors has therefore encouraged the integration of machine learning (ML) techniques into OSC research as a complementary strategy for accelerating materials discovery and device optimization. Among the available ML strategies, ensemble learning has proven particularly well suited to OSC systems, where datasets are often limited, heterogeneous, and derived from diverse experimental conditions. This review provides a focused and materials-oriented examination of ensemble learning applications in OSC research, spanning donor and acceptor screening, blend optimization, and stability-related prediction tasks. Bagging-based, boosting-based, and stacking-based approaches are discussed in relation to their roles in predicting key photovoltaic metrics, including power conversion efficiency (PCE), open-circuit voltage (Voc), short-circuit current density (Jsc), and related device parameters. The interplay between data sources, descriptor selection, and model performance is critically analyzed, with particular attention to model interpretability through feature-importance evaluation and other explainable learning techniques. Finally, current challenges are discussed, and ensemble learning is positioned as a practical and interpretable tool for accelerating rational OSC materials design.

KEYWORDS: Organic solar cells; performance prediction; machine learning; ensemble learning

1 Introduction

Organic solar cells (OSCs) have emerged as a promising class of photovoltaic technologies owing to their intrinsic advantages, including mechanical flexibility, lightweight construction, compatibility with low-temperature solution processing, and potential for low-cost, large-area fabrication [1–3]. Over the past decade, remarkable progress has been achieved in OSC performance, with laboratory-scale power conversion efficiencies (PCEs) increasing from below 1% in early demonstrations to values exceeding 20% in state-of-the-art devices. Recent reports of certified efficiencies approaching 19.4% in small-area devices and over 14% in large-area modules underscore the rapid evolution of OSCs toward practical applicability and commercialization [4–6]. These advances position OSCs as attractive candidates for emerging photovoltaic

applications such as building-integrated photovoltaics, portable electronics, and flexible energy-harvesting systems [7–10].

The rapid performance improvements in OSCs have been driven primarily by advances in organic semiconductor materials, particularly through the development of donor–acceptor architectures and non-fullerene acceptors (NFAs) with tailored electronic structures and optical properties [11–13]. Modern OSCs typically rely on finely engineered donor polymers or small molecules paired with NFAs that exhibit strong and tunable absorption in the visible and near-infrared regions, reduced nonradiative energy losses, and improved morphological compatibility. In parallel, progress in device architecture, interfacial engineering, and processing strategies has contributed to enhanced charge generation, transport, and extraction [14,15]. Despite these achievements, OSCs remain fundamentally distinct from inorganic photovoltaics in that their device performance is governed by a complex and highly coupled interplay among molecular structure, nanoscale morphology, and fabrication conditions [16–18].

From a materials science perspective, OSC optimization represents a uniquely challenging multivariable problem. Small changes in molecular backbone structure, side-chain chemistry, end-group functionalization, or blend composition can lead to substantial variations in energy-level alignment, exciton dynamics, phase separation behavior, and long-term stability [5,19]. This sensitivity is further amplified in modern device architectures, such as ternary and quaternary blends, where multiple components interact in nonlinear and system-specific ways [20–24]. As a result, the design space of OSC materials has expanded rapidly, encompassing an enormous combinatorial landscape of donor–acceptor pairs, molecular derivatives, and processing parameters.

Traditional trial-and-error experimental approaches, while essential for validating new concepts, are increasingly inadequate for efficiently navigating this expanding materials design space. Exhaustive experimental screening of all possible material combinations and fabrication conditions is impractical due to time, cost, and resource constraints. Physics-based computational methods, including quantum chemical calculations and mesoscale simulations, have provided valuable insights into structure–property relationships and charge-transport mechanisms. However, their direct application to high-throughput screening remains limited by computational expense and the difficulty of accurately capturing mesoscale morphology and processing-induced effects [25,26].

These challenges have motivated growing interest in data-driven approaches based on machine learning (ML) as complementary tools for accelerating OSC materials discovery and optimization [27–29]. By learning correlations between molecular descriptors, device parameters, and performance metrics from existing datasets, ML models provide a scalable framework for performance prediction, virtual screening, and trend analysis [30,31]. Over the past several years, ML has been increasingly applied to OSC research to predict key photovoltaic parameters, including PCE, open-circuit voltage (Voc), short-circuit current density (Jsc), and fill factor (FF). These studies demonstrate that ML models can extract meaningful patterns from heterogeneous experimental and computational data, enabling rapid evaluation of candidate materials and guiding experimental efforts toward promising regions of chemical space.

Despite this promise, the application of ML to OSC research faces several intrinsic challenges that distinguish it from other data-rich scientific domains. OSC datasets are typically modest in size and are often compiled from disparate literature sources with varying synthesis protocols, device architectures, and measurement conditions. Such heterogeneity introduces noise, bias, and inconsistencies that can limit the reliability and generalizability of purely data-hungry ML models. Furthermore, descriptor representations used to encode OSC materials may be high-dimensional, redundant, or only indirectly linked to the underlying physical processes governing device operation. These characteristics place important

constraints on model selection and necessitate ML strategies that balance predictive accuracy, robustness, and interpretability.

Among the available ML strategies, ensemble learning methods, which construct predictive models by combining multiple base learners, have attracted substantial attention in the OSC research due to their robustness, ability to mitigate overfitting, and enhanced interpretability [32–35]. These properties are particularly advantageous when working with complex materials datasets that are often small to medium in size and exhibit nonlinear structure–property relationships. While deep learning approaches, such as graph neural networks and transformer-based models, have demonstrated strong capabilities in learning representations directly from molecular structures, their effectiveness typically depends on the availability of large, high-quality datasets, which remain limited in OSC research. In contrast, ensemble learning methods are well suited to tabular and literature-derived datasets, making them a practical and reliable choice under realistic experimental conditions.

Ensemble learning approaches, including bagging-, boosting-, and stacking-based models, have demonstrated consistent performance across a wide range of OSC materials systems, including polymer donors, NFAs, and multi-component blends. Beyond improved predictive accuracy, ensemble models frequently provide feature-importance metrics and explainable outputs that enable researchers to connect data-driven results with established physical understanding. This capability is especially valuable in OSC research, where interpretability supports hypothesis generation and experimental validation. However, discussions of ensemble learning are often embedded within broader ML frameworks rather than examined as a distinct methodological category [36–38]. A focused analysis that highlights representative ensemble strategies, commonly used datasets and descriptors, and their role in guiding materials-level understanding can therefore provide additional clarity for researchers at the interface of materials science and data-driven modeling.

In this review, we provide a comprehensive overview of ensemble learning approaches applied to OSC research, with an emphasis on their role in materials design, performance prediction, and data-driven optimization. We first summarize the key classes of OSC materials and highlight the intrinsic complexity of their design space. We then discuss commonly used data sources and descriptor representations employed in OSC ML studies. Subsequently, representative ensemble learning strategies, including bagging-, boosting-, and stacking-based methods, are reviewed with attention to their predictive performance, interpretability, and practical limitations. Finally, current challenges and emerging opportunities are discussed, outlining future directions for integrating ensemble learning with high-throughput experimentation, computational chemistry, and physics-informed modeling to accelerate the rational design of next-generation OSC materials.

2 Organic Solar Cell Materials

Modern OSCs typically employ bulk heterojunction architectures composed of organic donor and acceptor materials that self-assemble into interpenetrating networks [39–43]. Efficient device operation requires the simultaneous optimization of light absorption, exciton generation and dissociation, charge transport, and charge extraction, all of which are highly sensitive to molecular structure and blend morphology. Small variations in backbone planarity, side-chain length, end-group functionalization, or processing conditions can significantly alter device performance. This sensitivity has driven rapid diversification of OSC materials, but it has also led to a substantial increase in the size and complexity of the design space.

From a materials discovery perspective, this expanding chemical space poses a fundamental bottleneck. The number of possible donor–acceptor combinations, molecular derivatives, blend compositions, and processing variables grows exponentially as new materials are introduced. Consequently, identifying optimal material systems through intuition-driven design or trial-and-error experimentation becomes increasingly

inefficient. In this context, ML has emerged as a powerful complementary tool, capable of learning structure–property–performance relationships directly from existing data and guiding materials selection and optimization in a data-driven manner. The following subsections review the major classes of OSC materials, highlighting their intrinsic design complexity and the specific opportunities they present for ML-assisted materials discovery.

2.1 Donor Materials for Organic Solar Cells

Donor materials play a central role in OSC operation by governing light absorption, exciton generation, hole transport, and energy-level alignment with acceptor materials [44]. Both polymer donors and small-molecule donors have been extensively explored, each offering distinct advantages and challenges from a materials design standpoint [45–48]. Polymer donors have historically dominated OSC research due to their strong absorption coefficients, mechanical flexibility, and compatibility with solution-based fabrication. Early systems such as poly (3-hexylthiophene) enabled foundational understanding of bulk heterojunction physics but were limited by narrow absorption windows and modest charge transport [49–51]. Subsequent generations of donor–acceptor copolymers introduced alternating electron-rich and electron-deficient units, allowing systematic tuning of optical bandgap, highest occupied molecular orbital (HOMO) energy levels, and backbone planarity. State-of-the-art polymer donors, including PM6-, D18-, and PBDB-T–derived systems, achieve high efficiencies through careful control of backbone rigidity, conjugation length, and side-chain architecture [52–55]. Small-molecule donors offer complementary benefits, including well-defined molecular structures, precise electronic properties, and improved batch-to-batch reproducibility [56]. However, their stronger crystallization tendencies often complicate morphology control, making performance highly sensitive to processing conditions. In both polymer and small-molecule donors, side-chain engineering plays a critical role by influencing solubility, molecular packing, phase separation, and mechanical stability.

From an ML perspective, donor materials present a high-dimensional design problem [57]. Molecular backbones, side chains, molecular weight, and electronic properties collectively define performance, yet their effects are often strongly correlated and nonlinear. Descriptor representations ranging from quantum-chemically derived frontier orbital energies to cheminformatics-based molecular fingerprints have therefore been widely employed to encode donor structures for ML models. Ensemble learning methods are particularly effective in this context, as they can handle descriptor redundancy and noisy experimental data while identifying key molecular features that dominate donor performance trends.

2.2 Acceptor Materials for Organic Solar Cells

Acceptor materials are equally critical to OSC performance, as they control electron transport, charge separation efficiency, and voltage losses. For many years, fullerene derivatives dominated OSC research due to their high electron mobility and favorable phase separation behavior [58,59]. However, their weak optical absorption, limited tunability of energy levels, and morphological instability motivated the development of alternative acceptor materials. The introduction of NFAs has fundamentally transformed OSC materials design [60–62]. NFAs based on fused-ring electron acceptors exhibit strong and tunable absorption in the visible and near-infrared regions, reduced nonradiative recombination losses, and improved compatibility with modern donor materials. Through systematic modification of core structures, end groups, and side chains, NFAs can be tailored to optimize lowest unoccupied molecular orbital (LUMO) levels, dipole moments, and intermolecular interactions.

At the same time, the chemical design space of NFAs is exceptionally large. Small structural modifications can lead to substantial changes in absorption spectra, energy-level alignment, and blend morphology.

This sensitivity makes comprehensive experimental exploration increasingly impractical. ML-based virtual screening has therefore become an attractive strategy for navigating NFA design space, enabling rapid evaluation of thousands to millions of hypothetical structures prior to synthesis [63]. Ensemble models have shown strong performance in capturing nonlinear relationships between NFA descriptors and device efficiency, making them well suited for acceptor optimization tasks.

2.3 Binary, Ternary, and Quaternary Blend Architectures

Most high-performance OSCs rely on bulk heterojunction blends in which donor and acceptor materials form nanoscale phase-separated networks [64]. Binary donor–acceptor systems have served as the foundation of OSC research and enabled systematic investigation of structure–property relationships [65]. However, to further enhance performance, ternary and quaternary blend architectures have been introduced, incorporating additional components to broaden absorption, optimize morphology, or reduce energy losses [24,66]. While multi-component blends have achieved record efficiencies, they dramatically increase materials complexity. The number of possible material combinations and composition ratios grows combinatorially, and interactions among components are often highly nonlinear and system-specific. As a result, identifying optimal blend compositions through conventional experimental approaches becomes increasingly inefficient.

ML provides a natural framework for addressing this challenge by learning performance trends across multidimensional composition spaces [67–69]. Ensemble learning models are particularly advantageous for ternary and quaternary systems, as they can robustly capture nonlinear interactions between multiple material components while remaining relatively insensitive to experimental noise. ML-assisted composition optimization has therefore emerged as a powerful strategy for guiding experimental design in complex blend systems.

2.4 Stability, Degradation, and Scalability Challenges

Despite significant improvements in efficiency, long-term stability and scalability remain critical barriers to the widespread deployment of OSCs [70,71]. Degradation mechanisms include photo-oxidation, thermal instability, morphological evolution, and chemical degradation at interfaces. These processes are influenced by intrinsic material properties as well as extrinsic factors such as device architecture and operating environment.

From a data perspective, stability studies present unique challenges. Degradation data are typically sparse, heterogeneous, and collected over long timescales, making conventional modeling approaches difficult. ML methods, particularly ensemble models that perform well with limited and noisy datasets, offer a promising avenue for predicting stability trends and identifying materials with improved lifetime characteristics. Integrating efficiency and stability prediction within unified ML frameworks represents an important opportunity for future OSC materials design [25,72,73].

2.5 Implications for Machine Learning–Assisted Materials Design

Across donor materials, acceptor materials, multi-component blends, and stability considerations, a common theme emerges which OSC materials design is characterized by high-dimensional, nonlinear, and interdependent variables that are difficult to optimize using intuition alone. ML provides a powerful complementary approach by enabling systematic analysis of existing data, rapid virtual screening of candidate materials, and identification of dominant structure–property–performance relationships.

Importantly, ML does not replace experimental or physics-based understanding but rather augments it by prioritizing promising design directions and reducing the cost of exploration. Ensemble learning methods, in particular, strike a practical balance between predictive accuracy and interpretability, making them well suited for guiding materials discovery and optimization in OSC research. This materials-centric perspective establishes a clear foundation for the data sources, descriptor representations, and ensemble learning strategies discussed in the following sections.

To provide a consolidated overview of the key materials design challenges in OSC and the corresponding roles that ML can play at different stages of optimization, Table 1 summarizes the major OSC material categories, dominant design variables, and representative ML opportunities and tasks.

Table 1: OSC materials design challenges and corresponding ML opportunities.

OSC Criteria	Key Design	Primary Challenge	ML Opportunity	Typical ML Task
Donor Materials	Conjugated backbone (building blocks), side-chain modifications, HOMO energy level, polymer molecular weight	Expansive molecular design space; performance depends on strongly correlated, nonlinear structure–property relationships	Identify critical molecular features and learn structure–performance correlations for high-efficiency donors	Performance prediction (PCE, Voc, Jsc) for new donor molecules; feature-importance analysis guiding donor design and screening
Acceptor materials	Fused-ring core architecture, electron-withdrawing end groups, side-chain substitutions, LUMO energy level, molecular dipole moment	Small structural tweaks (e.g., different end-groups) can cause large shifts in absorption, energy levels, and blend morphology—leading to an enormous candidate space	Evaluate thousands of NFA candidates and capture nonlinear structure–efficiency patterns	Efficiency prediction to rank and select promising acceptor molecules; classification models to flag high-performance vs. low-performance acceptors
Binary donor–acceptor blends	Donor–acceptor material pairing; blend ratio (D:A composition)	Interdependent electronic alignment and nanoscale morphology—small changes in material pair or ratio can dramatically affect performance	Data-driven pattern recognition across known D–A pairs to uncover which combinations yield high efficiency	Predicting device metrics (PCE, Jsc, Voc) for specific D-A pairs; screening and ranking donor–acceptor combinations to identify promising pairs

(Continued)

Table 1 (continued)

OSC Criteria	Key Design	Primary Challenge	ML Opportunity	Typical ML Task
Ternary/ quaternary blends	Multiple active components (additional donor(s) and/or acceptor(s)) and their composition ratios	Combinatorial explosion of possible multi-component formulations and mixing ratios; highly nonlinear, system-specific interactions among components	Capture interdependent effects and guide optimization; efficient search for optimal ternary/quaternary blend formulations	Performance forecasting for given ternary/quaternary compositions; recommending optimal material combinations and component ratios
Stability and degradation	Intrinsic material stability (resistance to photo-oxidation, thermal stress); morphology evolution during aging	Long-term performance data are sparse, noisy, and collected under varying conditions, hindering direct modeling	Learn degradation trends from limited data, enabling lifetime prediction and identification of stability-enhancing material features	Predicting efficiency loss over time (PCE vs. aging duration) for new devices; classifying stable vs. unstable materials; interpreting key degradation factors via model explainability

3 Data Sources and Descriptors in Machine Learning Studies of Organic Solar Cells

The application of ML in OSC research is intrinsically tied to the nature of the underlying materials design problem. Donor materials, non-fullerene acceptors, multi-component blends, and stability-critical device architectures each involve distinct variables, constraints, and sources of uncertainty. Consequently, the effectiveness of ML models depends not only on algorithm selection but also critically on the availability, quality, and representation of data used to describe OSC materials and devices.

In OSC research, ML models are typically trained using heterogeneous datasets that integrate experimental device performance metrics with molecular- and materials-level descriptors. For example, donor and acceptor optimization tasks rely primarily on structural and electronic-property descriptors, whereas blend composition and processing optimization require device-level and fabrication-related inputs. A clear understanding of data sources and descriptor representations is therefore essential for interpreting ML predictions and ensuring their relevance to practical materials design.

3.1 Data Sources

Data used in OSC ML studies can be broadly categorized into experimental performance data and computationally derived materials descriptors. Experimental datasets are most commonly compiled from the literature and include key photovoltaic metrics such as PCE, Voc, Jsc, and FF. However, literature-derived

datasets are typically limited in size and affected by variability in synthesis protocols, device architectures, and measurement conditions. Such heterogeneity introduces noise and bias that can obscure underlying structure–property relationships. This issue is particularly pronounced for complex design problems such as ternary and quaternary blends, where reported data may span a wide range of compositions and processing conditions. As a result, ML models applied to OSC datasets must be robust to inconsistencies and capable of generalizing from incomplete information.

Computational data, most commonly obtained from quantum chemical calculations, provides a complementary and more controlled source of information. Properties such as frontier molecular orbital energies, optical bandgaps, dipole moments, and reactivity descriptors are widely used to represent donor and acceptor materials in ML studies. These descriptors are especially relevant for materials design tasks including donor and acceptor screening and voltage optimization. Nevertheless, the generation of computational descriptors can be computationally expensive and sensitive to methodological choices, limiting scalability when applied to very large materials libraries.

3.2 Descriptor Representation

Descriptor representation serves as the critical interface between OSC materials and ML models, as it defines how materials and device characteristics are encoded for prediction tasks. In practice, three broad categories of descriptors are commonly employed: electronic structure descriptors, structure-based descriptors, and device- or process-level descriptors.

Electronic structure descriptors, such as HOMO/LUMO energy levels, bandgaps, and dipole moments, provide physically interpretable links to charge generation, separation, and recombination processes. These descriptors are particularly relevant for donor and acceptor materials design tasks and have been widely used in ML models targeting Voc, Jsc, and PCE prediction. Their relatively low dimensionality and clear physical meaning make them attractive for interpretable modeling, especially when datasets are limited. However, it is important to note that the values of these descriptors can depend on the computational protocols used to generate them. Differences in density functional theory (DFT) settings, including the choice of exchange–correlation functional and basis set, can introduce systematic variations in calculated properties such as HOMO/LUMO energy levels and bandgaps. As a result, descriptor values reported across different studies may not be directly comparable. When datasets are aggregated from multiple sources, these inconsistencies can introduce additional uncertainty that is unrelated to intrinsic material properties. Ensuring consistent computational settings or applying appropriate data harmonization strategies is therefore important for improving the reliability of ML models.

Structure-based descriptors, including molecular fingerprints derived from Simplified Molecular Input Line Entry System (SMILES) representations, enable scalable encoding of complex chemical structures without requiring explicit quantum chemical calculations [74]. These descriptors are particularly useful for large-scale virtual screening, where thousands to millions of hypothetical donor and acceptor materials must be evaluated efficiently. However, fingerprint-based representations are typically high-dimensional and may contain redundant or weakly relevant features, which can complicate model training and reduce interpretability if not properly managed.

Device- and process-level descriptors, such as blend composition ratios, film thickness, solvent choice, annealing temperature, and additive concentration, play a dominant role in modeling binary and multi-component OSC systems. These descriptors are essential for tasks such as composition optimization and process window identification, as they enable ML models to capture coupled materials–processing effects that cannot be described by molecular descriptors alone.

Fig. 1 illustrates a typical workflow in OSC ML studies, showing how experimental and computational data are transformed into descriptor sets and subsequently integrated into ensemble learning models for performance prediction and materials screening.

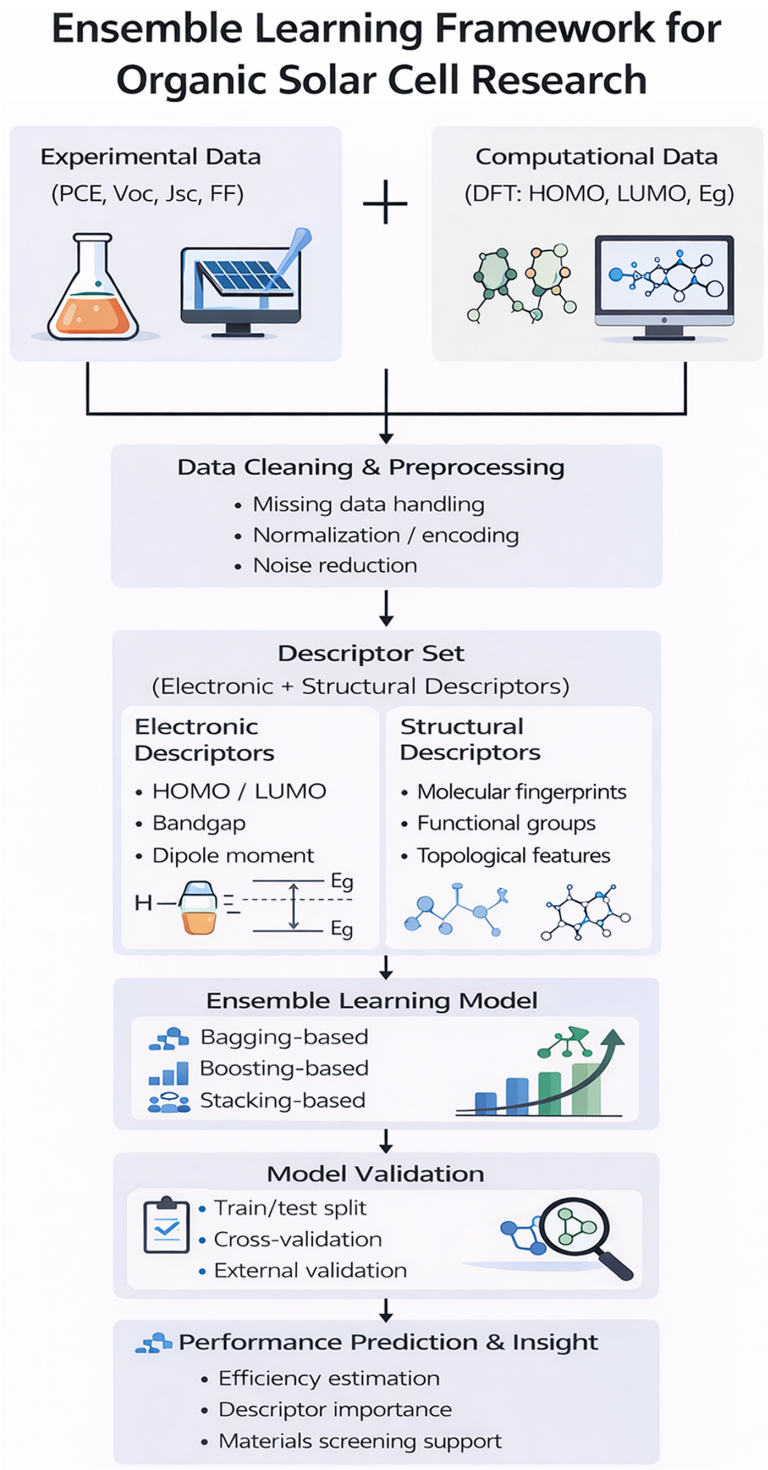

Figure 1: Workflow illustrates common data sources, descriptor generation, and their use in ensemble learning models for OSC research.

In addition to molecular and electronic descriptors, morphology and processing conditions play a critical role in determining OSC performance. However, incorporating these factors into ML models remains challenging. Experimental techniques such as grazing-incidence wide-angle X-ray scattering (GIWAXS) and atomic force microscopy (AFM) provide detailed information on nanoscale morphology, including crystallinity, domain size, and phase separation, but translating these measurements into standardized, quantitative descriptors suitable for ML is non-trivial. In practice, processing parameters such as solvent additives or annealing conditions, are often represented using simplified categorical variables. While this enables model implementation, it may not fully capture the continuous and interdependent nature of morphology evolution. Developing quantitative and standardized morphology descriptors therefore remains an important direction for improving descriptor representation in OSC ML studies. Ensemble learning methods can partially mitigate these limitations by integrating heterogeneous descriptor types and maintaining robustness to noisy or simplified representations. The implications of these limitations for model interpretability and reliability are further discussed in Section 5.2.

3.3 Data Preprocessing and Descriptor Selection

Effective ML modeling for OSCs critically depends on the quality and relevance of the input descriptors used to represent materials and device characteristics. Due to the inherently multiscale nature of OSC systems such as spanning molecular-level properties (e.g., HOMO/LUMO energies, dipole moments), morphological features (e.g., domain purity, crystallinity), and processing conditions (e.g., solvent choice, annealing temperature) descriptor sets are often high-dimensional and heterogeneous. Without careful preprocessing, such data can introduce noise, multicollinearity, and missing values, which degrade predictive performance and hinder model interpretability.

To address these challenges, common preprocessing steps include normalization or standardization of numerical variables, encoding of categorical features (such as processing conditions), and imputation of missing data. However, for literature-derived OSC datasets, these steps are not always straightforward due to inconsistencies in reporting formats, experimental protocols, and measurement conditions. As a result, preprocessing decisions can significantly influence model outcomes and should be regarded as a critical component of the ML workflow rather than a routine step.

Descriptor selection plays a central role in balancing model complexity and interpretability. High-dimensional descriptors, such as molecular fingerprints, enable efficient screening of large chemical spaces but may introduce redundancy and obscure physically meaningful relationships. In contrast, physically motivated descriptors, such as HOMO/LUMO energy levels or bandgaps, provide clearer interpretation but may not fully capture morphology- and processing-dependent effects. Feature selection and dimensionality reduction techniques, including principal component analysis, recursive feature elimination, and model-based selection using feature importance scores, are therefore essential for identifying relevant variables and reducing overfitting. In addition, post hoc interpretability methods such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) are increasingly used to examine how individual descriptors contribute to model predictions, particularly in complex ensemble learning models.

In addition to preprocessing and descriptor selection, model validation is a key methodological consideration that directly affects the reliability of ML predictions. In OSC studies, datasets are often limited in size and derived from heterogeneous sources, and simple random train–test splits may not adequately reflect real experimental variability. More robust validation strategies, such as cross-validation, external validation using independent datasets, or domain-aware data splitting, are therefore recommended to ensure that model performance is not overestimated and remains reproducible across different datasets. More robust validation strategies, such as k-fold cross-validation, external validation using independent datasets,

and domain-aware data splitting, are therefore recommended to ensure that model performance is not overestimated and remains reproducible across different datasets.

Following model validation, the selection of appropriate evaluation metrics is equally important for assessing model performance. These metrics depend on the specific learning task and are generally categorized into regression and classification metrics. Since these two categories serve different objectives, they should be interpreted separately to avoid misleading comparisons.

As summarized in Table 2, regression metrics evaluate continuous prediction accuracy (e.g., PCE or Voc), while classification metrics assess discrete decision outcomes such as material screening. These metrics serve different objectives depending on the learning task and should be interpreted within their specific modeling context rather than compared directly across studies.

Table 2: Evaluation metrics for classification and regression and tasks.

Types	Metrics	Purpose	Formula
Classification	Confusion Matrix	To summarize classification results by showing correct and incorrect predictions TP (True Positive), TN (True Negative, FP (False Positive), FN (False Negative)	$Confusion\ Matrix = \begin{matrix} TP & FN \\ FP & TN \end{matrix}$
	Accuracy	Measures overall correctness of predictions	$Accuracy = \frac{TP+TN}{TP+FN+FP+TN}$
	Precision	Measures how reliable positive predictions are	$Precision = \frac{TP}{TP+FP}$
	Recall	Measures ability to detect actual positives	$Recall = \frac{TP}{TP+FN}$
	F1 Score	Balances precision and recall	$F1 = \frac{2TP}{2TP+FP+FN}$
	Receiver Operating	Evaluates model performance across different thresholds	$Y: TPR = \frac{TP}{TP+FN}$
	Characteristic (ROC)	TPR (True Positive Rate), FPR (False Positive Rate)	$X: FPR = \frac{FP}{FP+TN}$
	Mean Squared Error (MSE)	Measures average squared prediction error	$MSE\,(y, \hat{y}) = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2$
	Root Mean Squared Error (RMSE)	Measures prediction error in original unit	$RMSE\,(y, \hat{y}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}$

(Continued)

Table 2 (continued)

Types	Metrics	Purpose	Formula
Regression	Mean Absolute Error (MAE)	Measures average absolute error	$MAE\left(y,\hat{y}\right)=\frac{1}{N}\sum_{i=1}^{N}\left\|\left(y_i-\hat{y}_i\right)\right\|$
	Mean Absolute Percentage Error (MAPE)	Measures error in percentage form	$MAPE\left(y,\hat{y}\right)=\frac{1}{N}\sum_{i=1}^{N}\frac{\left\|\left(y_i-\hat{y}_i\right)\right\|}{\left\|y_i\right\|}$
	Coefficient of Determination (R^2)	Measures how well model explains data variability	$R^2\left(y,\hat{y}\right)=1-\frac{\sum_{i=1}^{N}\left(y_i-\hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i-\hat{y}_i\right)^2}=\frac{1}{N}\sum_{i=1}^{N}\hat{y}_i$
	Pearson Correlation Coefficient (r)	Measures linear relationship between predicted and actual values	$r=\frac{\sum_{i=1}^{n}\left(y_i-\overline{y}\right)\left(\hat{y}_i-\overline{\hat{y}}\right)}{\sqrt{\sum_{i=1}^{n}\left(y_i-\overline{y}\right)^2}\sum_{i=1}^{n}\left(\hat{y}_i-\overline{\hat{y}}\right)^2}$

Ensemble learning methods, particularly tree-based models such as random forests (RF) and gradient boosting, offer advantages in handling high-dimensional and partially correlated descriptors and provide feature importance measures that support interpretability. However, it is important to recognize that such importance scores reflect statistical relationships rather than direct physical causality. Careful interpretation is therefore required when translating ML outputs into actionable materials design insights.

Overall, data preprocessing, descriptor selection, and model validation should be considered integral components of OSC ML modeling, as their impact on model performance can be comparable to, or even greater than, the choice of algorithm itself. These steps form the critical bridge between raw experimental or computational data and reliable, physically meaningful predictions.

3.4 Implications for Model Choice in OSC Research

The characteristics of OSC datasets which are limited size, noise, descriptor heterogeneity, and nonlinear interactions place important constraints on ML model selection. Models that require large, homogeneous datasets or highly curated descriptors may struggle to generalize across different materials systems.

In contrast, ensemble learning methods are well suited to the diverse ML tasks outlined in Table 1, as they can accommodate mixed descriptor types, capture nonlinear relationships, and remain robust under data-limited conditions. By explicitly aligning data sources and descriptor choices with specific OSC materials challenges, ML models can be more effectively designed, interpreted, and applied. This materials-oriented perspective provides a coherent bridge between the qualitative discussion of OSC design complexity in Section 2 and the detailed review of ensemble learning strategies presented in Section 4.

4 Ensemble Learning Methods for Performance Prediction and Materials Optimization in Organic Solar Cells

Ensemble learning methods have emerged as particularly effective ML strategies for addressing the intrinsic complexity of OSC systems. OSC datasets are typically heterogeneous, limited in size, and

influenced by coupled molecular, morphological, and processing variables. These characteristics favor ML approaches that are robust to noise, capable of capturing nonlinear relationships, and interpretable enough to provide materials-relevant insight. Ensemble learning meets these requirements by integrating multiple base learners to improve generalization performance while retaining the ability to identify dominant descriptors governing device behavior.

Predictive performance metrics reported across different studies are inherently influenced by variations in descriptor selection, dataset composition, and modeling protocols, which can lead to differences in reported results. Therefore, model performance is discussed in a context-dependent manner, focusing on the conditions under which specific ensemble approaches are effective.

Ensemble learning methods are commonly categorized into three major strategies: bagging-based, boosting-based, and stacking-based ensembles. Each strategy aligns naturally with different OSC materials design tasks, ranging from donor and acceptor screening to multicomponent blend optimization and stability prediction. Fig. 2 schematically summarizes these ensemble strategies, while representative applications are discussed in the following subsections [75].

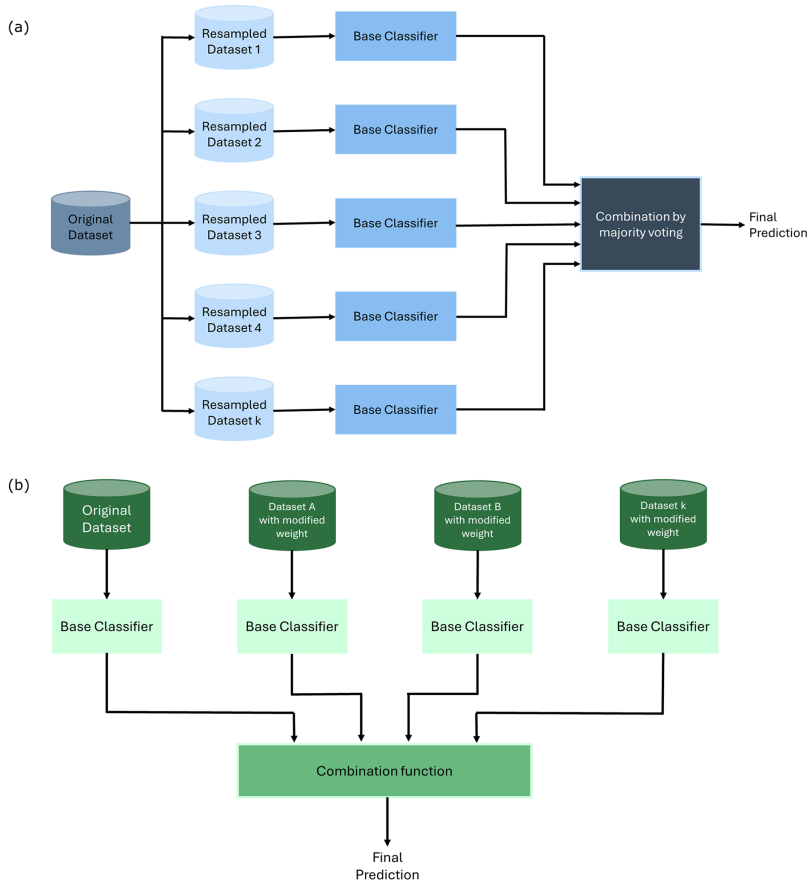

Figure 2: (Continued)

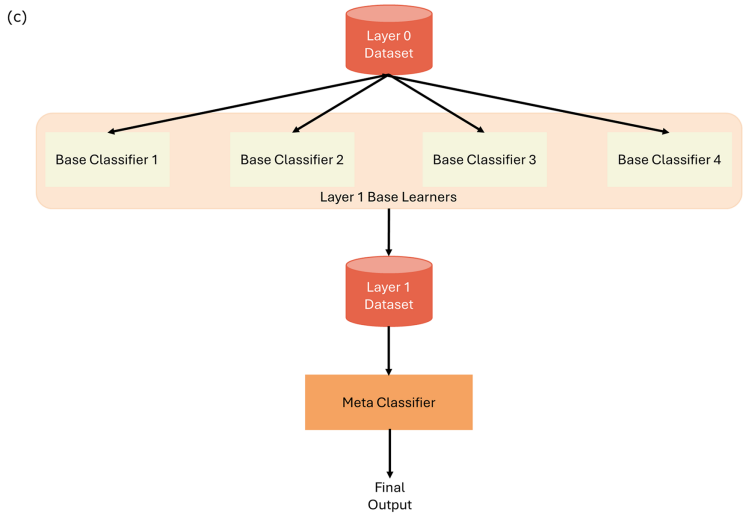

Figure 2: Overview of ensemble learning strategies including (**a**) bagging-, (**b**) boosting-, and (**c**) stacking-based methods. Recreated with permission from Ref. [75]. Copyright (2024). Springer Nature.

Fig. 2a illustrates the principle of bagging-based ensemble learning, in which multiple base learners are trained independently on different bootstrap samples drawn from the same dataset. Their predictions are subsequently aggregated through averaging or majority voting to improve prediction stability and robustness. Fig. 2b depicts the workflow of boosting-based ensemble learning, where base learners are trained sequentially and increasing emphasis is placed on data points that are poorly predicted by earlier models, with the final prediction obtained via a weighted combination of all learners. Fig. 2c presents stacking-based ensemble learning, in which multiple base learners generate individual predictions that serve as inputs to a meta-learner responsible for learning an optimal combination strategy; in practice, this requires careful validation design to avoid information leakage.

A summary of commonly used ensemble learning methods, their roles, and representative applications in OSC research is provided in Table 3.

Table 3: Ensemble learning methods commonly applied in OSC research.

Ensemble Strategy	Method	Base Learner	Key Characteristics	Representative OSC Studies
Bagging-based ensemble	Random Forest (RF)	Decision trees	Uses bootstrap sampling and random feature selection to build multiple trees; robust to noise	Sun et al. (2019) [76] Lee (2019) [77] Lee (2020) [78] Lee (2020) [79] Wu et al. (2020) [80] Hao et al. (2021) [69] Rodríguez-Martínez et al. (2021) [81] Suthar et al. (2023) [82] Zhao et al. (2024) [83] Li. et al. (2024) [68] Liu et al. (2025) [84]

(Continued)

Table 3 (continued)

Ensemble Strategy	Method	Base Learner	Key Characteristics	Representative OSC Studies
Boosting-based ensemble	Gradient Boosting Decision Tree (GBDT)/ Gradient Boosting Regressor (GBR)/ Gradient Boosting Regression Tree (GBRT)	Decision trees	Optimizes loss function through iterative tree construction	Sahu et al. (2019) [85] Suthar et al. (2023) [82] Lee (2024) [86] Das & Mondal (2024) [87] Siddique et al. (2025) [88] Tahir et al. (2025) [89]
	eXtreme Gradient Boosting (XGBoost)	Decision trees	Optimized gradient boosting with regularization and parallel training	Wang et al. (2023) [90] Zhao et al. (2023) [91] Lee (2023) [92] Lee (2023) [93] Lee (2024) [94] Lee et al. (2025) [95] Cui et al. (2026) [96]
Stacking-based ensemble	Stacking	Mixed models	Combines predictions from multiple base learners using a meta-model	Wang et al. (2023) [97] Akbar et al. (2025) [98] Valiente et al. (2026) [99]

4.1 Bagging-Based Ensemble Learning

Bagging-based ensemble learning constructs predictive models by training multiple base learners independently on different bootstrap samples drawn from the same dataset and aggregating their predictions. Among these methods, random forests (RF) have been the most widely adopted ensemble approach in OSC research due to its robustness against overfitting, tolerance to descriptor redundancy, and ability to handle noisy and heterogeneous datasets. From a materials design perspective, RF models are particularly well suited for OSC applications where experimental data are compiled from diverse literature sources and molecular descriptors are high dimensional, as is often the case for donor polymers and non-fullerene acceptors. RF models leverage random feature selection and ensemble averaging to reduce variance, making them reliable predictors even when datasets are modest in size.

Sun et al. (2019) [76] presented an ML–assisted framework for accelerating the discovery of high-performance OSC donor materials by establishing structure–property relationships directly from experimentally reported data. They constructed a comprehensive database of 1719 donor molecules collected from the literature and systematically evaluated different molecular representations (Fig. 3a), including images, SMILES strings, physicochemical descriptors, and multiple types of molecular fingerprints. Among the tested algorithms, RF emerged as the most reliable model for classifying donor materials into low- and high-performance categories based on PCE thresholds. RF demonstrated strong robustness when handling high-dimensional fingerprint representations, particularly those exceeding 1000 bits, outperforming neural network and support vector machine models under limited and noisy data conditions (Fig. 3b). Importantly, the RF models were not only used for retrospective prediction but also for prospective screening by designing and synthesizing ten new donor molecules based on RF predictions, achieving good agreement between predicted performance classes and experimental results (Fig. 3c). This work established RF-based ensemble learning as a practical prescreening tool that bridges data-driven modeling and experimental validation, highlighting its suitability for realistic OSC datasets characterized by structural complexity and experimental variability.

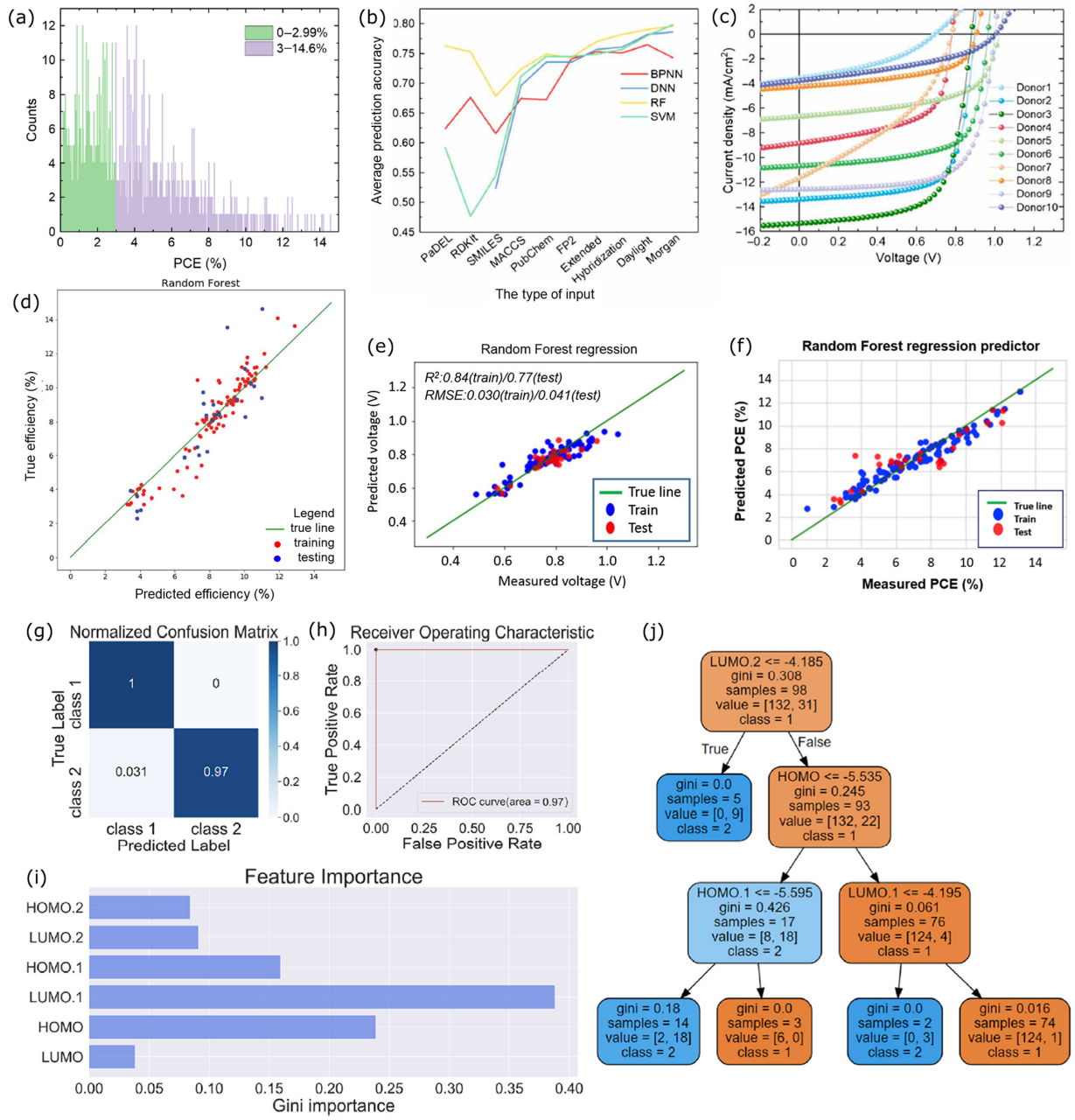

Figure 3: Summary of RF–based modeling applications in OSC research. (**a**) Distribution of PCE values of the 1719 molecules in compiled database, (**b**) performance comparison of four ML models, and (**c**) J-V curve of the OSC with the active layer using the predicted donor material. Reproduced with permission from Ref. [76]. Copyright (2019) The Authors (CC BY-NC). (**d**) Linear correlation between experimentally measured and RF-predicted PCE values (**e**) RF-predicted Voc as a function of measured energy loss, and (**f**) comparison between RF-predicted and experimentally measured PCE values. Reproduced with permission from Refs. [77–79]. Copyright (2019) WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim and Copyright (2020) Elsevier. (**g**) Normalized confusion matrix of the RF classification model, (**h**) receiver operating characteristic (ROC) curve with corresponding area under the curve (AUC), (**i**) Gini-based feature importance derived from the RF classifier, and (**j**) representative decision tree illustrating the classification logic based on true/false decision rules. Reproduced with permission from Ref. [69]. Copyright (2021) The Authors (CC-BY-NC-ND).

In a series of studies by Lee in 2019 [77] and 2020 [78,79], Lee demonstrated a consistent and progressively refined use of RF modeling to address different design challenges in OSCs. In the 2019 study focusing on fullerene-based ternary OSCs [77], RF was applied to a relatively small dataset of 124 devices collected from the literature, incorporating frontier molecular orbital energy levels and basic device structural information to predict PCE. The regression results, shown in Fig. 3d, indicate a reasonable agreement between predicted and experimental PCE values for both training and test datasets, with no clear evidence of overfitting. Beyond prediction accuracy, RF was further used to evaluate feature importance, highlighting the dominant role of donor frontier orbital energies. This analysis provided data-driven support for the importance of energy-level alignment in governing the performance of complex ternary blends. In a subsequent study aimed at understanding and improving the Voc of ternary OSCs [78], RF was employed more explicitly as a tool for extracting practical design rules. Regression models trained on frontier molecular orbital energies enabled accurate Voc prediction for both reported and previously unseen ternary systems. As illustrated in Fig. 3e, the predicted and measured Voc values exhibit a strong linear correlation, indicating reliable model performance. Importantly, the feature ranking obtained from the RF model allowed key electronic parameters responsible for voltage losses to be identified, translating statistical outcomes into physically meaningful guidance for selecting donor, acceptor, and third-component materials. In the third study, which extended the approach to non-fullerene OSCs [79], RF was applied to a broader class of materials using a limited set of easily accessible electronic descriptors. Despite the reduced descriptor complexity, the RF model achieved strong predictive performance for PCE, demonstrating that ensemble methods can remain effective under practical data constraints. Fig. 3f shows a clear linear relationship between predicted and measured PCE values, accompanied by a near-normal error distribution, suggesting good generalization to independent datasets. Together, these works illustrate how RF evolved into a tool not just for prediction but for gleaning design rules (through feature importance) across different OSC contexts.

Wu et al. (2020) [80] developed an ML framework to accelerate the discovery of high-performance donor–acceptor (D/A) pairs for non-fullerene OSCs, with RF serving as a key predictive and screening tool. Using a literature-derived dataset of 565 experimentally reported D/A combinations, they transformed donor and acceptor molecular structures into digitized fragment-based representations and trained RF models to predict PCE. When compared with other ML approaches, RF demonstrated strong predictive reliability and robustness to structural complexity and data variability. The trained RF model was subsequently applied to screen over 32 million hypothetical D/A combinations, from which several promising pairs were identified and experimentally validated, showing good agreement between predicted and measured efficiencies. This work highlights the effectiveness of RF in combining data-driven prediction with large-scale virtual screening to guide experimental efforts in non-fullerene OSC research.

Hao et al. (2021) [69] applied RF modeling to analyze and optimize energy-level alignment in non-fullerene ternary OSCs, with the aim of identifying high-efficiency material combinations from experimentally reported data. Using frontier molecular orbital energy levels of donors, major acceptors, and third components as descriptors, RF was implemented for both regression and classification tasks to predict PCE and to distinguish high-performance devices. Fig. 3g shows the normalized confusion matrix, indicating a high recall for efficient ternary devices despite the imbalanced dataset. Fig. 3h presents the receiver operating characteristic curve, with an area under the curve close to unity, confirming the strong discriminative ability of the RF model. Fig. 3i illustrates the feature-importance ranking, where the LUMO level of the major acceptor dominates device performance, followed by the donor HOMO, while the third-component energy levels provide a secondary tuning effect. Fig. 3j displays a representative decision tree from the RF ensemble, offering an intuitive view of how energy-level thresholds are used to classify device efficiency. Collectively, these results demonstrate how RF serves not only as a reliable

predictive model but also as an interpretable framework that links data-driven analysis with physically meaningful design guidance for non-fullerene ternary OSCs.

Rodríguez-Martínez et al. (2021) [81] employed RF modeling in combination with high-throughput experimentation to predict how the Jsc varies with donor–acceptor composition and active-layer thickness in binary OSCs. Thousands of experimental data points were generated using combinatorial thickness–composition libraries, and RF models were trained using a small set of intrinsic material descriptors, primarily frontier energy levels and charge carrier mobilities. Unlike purely descriptive approaches, RF was used as a predictive tool capable of extrapolating the photocurrent–composition phase space to material systems not included in the training set. This capability is illustrated in Fig. 4a, which presents a matrix of experimental Jsc–composition maps alongside RF-predicted curves for both training and unseen donor–acceptor pairs at different thicknesses. The close agreement between predicted and experimental trends demonstrates that RF can reliably capture nonlinear composition-dependent photocurrent behavior and identify optimal composition windows using minimal input information. This study highlights RF as a robust ensemble-learning approach for navigating complex experimental design spaces and guiding composition optimization in OSC devices.

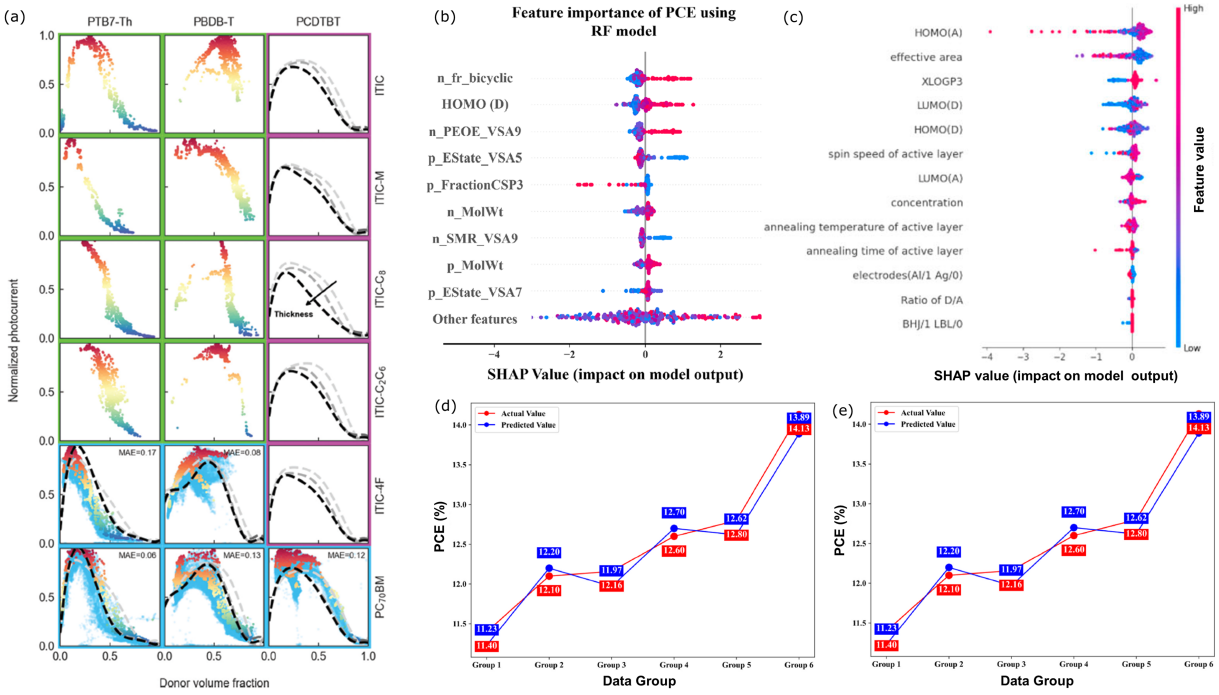

Figure 4: (Continued)

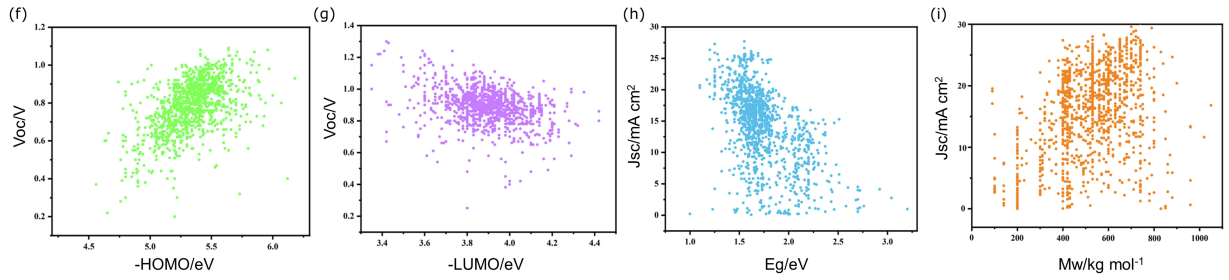

Figure 4: Summary of RF–based modeling applications in OSC research. (**a**) Combinatorial matrix of photocurrent phase diagrams for a set of high-performing polymer donor–acceptor system. Reproduced with permission from Ref. [81]. Copyright (2021) The Authors (CC BY 3.0) (**b**) Feature-importance ranking of frontier molecular orbital (FMO) and RDKit-derived descriptors for PCE, obtained from SHAP analysis of the RF model. Reproduced with permission from Ref. [82]. Copyright (2023). Royal Society of Chemistry. (**c**) SHAP value scatter plot for the RF model, highlighting the contribution and directionality of key descriptors influencing PCE prediction. Reproduced with permission from Ref. [83]. Copyright (2024) The Authors (CC BY). (**d**,**e**) Comparison between experimentally measured and RF-predicted PCE values for selected device groups in the database. Reproduced with permission from Ref. [68]. Copyright (2024) Elsevier. (**f**) Voc vs. HOMO. (**g**) Voc vs. LUMO. (**h**) Jsc vs. Eg. (**i**) Jsc vs. Mw, summarizing structure-property relationships derived from RF-assisted analysis. Reproduced with permission from Ref. [84]. Copyright (2025) The Authors (CC BY-NC-ND 4.0).

Suthar et al. (2023) [82] employed RF modeling to predict the photovoltaic performance of polymer–non-fullerene acceptor OSCs by learning relationships between molecular structure and device PCE from a large, literature-derived dataset. A database of 1242 experimentally reported donor–acceptor combinations was constructed, and RF models were trained using a combined set of frontier molecular orbital energies and chemically meaningful structural descriptors generated from molecular fingerprints. Among the evaluated models, RF showed the most reliable performance for predicting PCE, demonstrating strong agreement between predicted and experimental values while remaining robust to data heterogeneity. The role of RF in this study is further clarified in Fig. 4b, which presents the feature-importance ranking obtained from SHAPley Additive Explanation (SHAP) analysis applied to the RF model. This figure highlights that specific structural descriptors, such as bicyclic ring content and donor frontier orbital levels, contribute most strongly to efficiency predictions, providing clear insight into how molecular features influence device performance. Together, these results illustrate how RF serves both as a reliable predictive tool and as an interpretable framework for identifying key structure performance relationships in OSCs.

Zhou et al. (2024) [83] applied RF modeling to quantitatively evaluate how fabrication parameters influence the PCE of PM6:Y6 non-fullerene OSCs, addressing the challenge of optimizing multiple interdependent process variables simultaneously. A literature-derived dataset was constructed using experimentally reported preparation parameters, including effective device area, solvent and additive type, spin-coating speed, solution concentration, annealing temperature, and annealing time, and RF was selected as the most reliable model after comparison with several other ML algorithms. Beyond accurate efficiency prediction, RF was used as an interpretative tool through SHAP analysis to reveal how individual process parameters contribute positively or negatively to device performance. Fig. 4c clearly visualizes the SHAP value distribution for each fabrication parameter, directly showing each parameter's relative influence on efficiency. This subfigure demonstrates, for example, the dominant impact of effective area and solvent properties, while also highlighting optimal parameter ranges rather than single best values. Overall, this work illustrates how RF can be effectively used not only for predicting device efficiency, but also for providing practical, data-driven guidance for fabrication process optimization in OSCs.

Li et al. (2024) [68] applied RF modeling to systematically predict and screen the performance of donor–acceptor–acceptor ternary OSCs based on experimentally reported data. They constructed a curated dataset of 280 ternary devices using descriptors that reflect both materials selection and fabrication conditions, including donor–acceptor mass ratios, film morphology roughness, frontier molecular orbital energy levels, and molecular fingerprints. RF was benchmarked against several tree-based ensemble models and demonstrated superior performance in predicting PCE, maintaining good generalization on unseen data. The effectiveness of RF as a predictive model is illustrated in Fig. 4d,e, which shows the strong agreement between experimentally measured and RF-predicted efficiencies for representative ternary systems. Beyond prediction, the trained RF model was used to screen more than 429,000 unexplored ternary combinations, enabling the identification of high-efficiency candidates for future experimental investigation. This study demonstrates how RF can function as both a reliable predictor and a practical screening tool for navigating the complex compositional space of ternary OSCs.

Liu et al. (2025) [84] employed RF as the core learning model to establish structure–property relationships and guide the design of high-performance non-fullerene acceptor materials for OSCs. Using a curated dataset of 1343 experimentally reported acceptors, they introduced a polymer-unit fingerprint (PUFp) representation to describe macromolecular structures in a chemically meaningful manner, alongside conventional electronic descriptors such as HOMO, LUMO, bandgap, and molecular weight. Among several ML algorithms evaluated, RF exhibited the best predictive performance for PCE, while also enabling robust feature-importance and SHAP analyses. The practical role of RF in linking molecular properties to device behavior is clearly illustrated in Fig. 4f–i, which summarizes statistical relationships between key experimental parameters and photovoltaic metrics, including Voc vs. HOMO/LUMO levels and Jsc vs. bandgap and molecular weight. These plots visually reinforce the RF-derived insights by showing how energy-level alignment, bandgap narrowing, and molecular weight optimization jointly influence device performance, in agreement with established OSC physics. By combining accurate RF-based prediction with interpretable structure–property trends and subsequent virtual screening of new acceptor candidates, this work demonstrates how ensemble learning can move beyond prediction toward rational materials design in organic photovoltaics.

4.2 Boosting-Based Ensemble Learning

Boosting-based ensemble learning methods differ from bagging approaches by training base learners sequentially, with each new model emphasizing data points that are poorly predicted by earlier learners. This sequential error-correction mechanism enables boosting models to capture subtle nonlinear relationships between descriptors and target properties, making them especially valuable for fine-grained OSC performance optimization. Common boosting approaches applied in OSC research include gradient boosting methods (GBR, GBDT and GBRT). More advanced boosting frameworks, such as XGBoost, extend this concept by improving computational efficiency, handling complex feature interactions, and reducing overfitting. From a materials perspective, boosting-based models are particularly effective for capturing energy-level alignment effects, voltage losses, and charge-generation dynamics that are highly sensitive to small changes in molecular structure.

Sahu et al. (2019) [85] employed GBRT models to accelerate molecular discovery for OSC through large-scale virtual screening. In their work, experimentally reported efficiencies of small-molecule OSC were combined with quantum-chemically derived descriptors to train ML models capable of predicting PCE with significantly higher accuracy than conventional Scharber-type models. Among the tested algorithms, GBRT showed the strongest predictive performance, capturing nonlinear relationships between molecular structure, electronic properties, and device efficiency. The trained GBRT model was then applied to screen over

10,000 candidate donor molecules constructed from common building blocks, enabling the identification of promising chemical motifs and molecular arrangements associated with high efficiency. Beyond prediction, the gradient boosting framework was used to extract physically meaningful design rules, highlighting key electronic descriptors and structural features that govern device performance. This study demonstrates how ensemble boosting methods can serve not only as efficient screening tools but also as knowledge-extraction platforms for guiding rational molecular design in OSCs.

In a subsequent study, Suthar et al. (2023) [82] extended their ML framework by deliberately emphasizing gradient boosting to address performance parameters that were less accurately captured by RF models, particularly the Voc and Jsc. Using a large experimental dataset of 1242 polymer:NFA combinations, they compared multiple regression models and found that while RF remained most effective for predicting overall PCE, gradient boosting provided superior accuracy for Voc and Jsc, achieving stronger correlations with experimental values. This distinction reflects the ability of gradient boosting to sequentially correct prediction errors and capture subtle, non-linear dependencies associated with voltage losses and charge-generation processes. Through SHAP analysis, the study further showed that gradient boosting highlighted different descriptor sensitivities compared with RF, particularly emphasizing energy offsets, bandgap-related terms, and structure-dependent features influencing charge extraction. By positioning gradient boosting as a complementary model rather than a replacement, they demonstrated that different ensemble strategies can be selectively applied to target specific photovoltaic parameters, offering a more nuanced and physically informed pathway for optimizing OSC performance beyond efficiency alone.

Lee (2024) [86] employed GBDT models to predict key OSC performance parameters of polymer:NFA OSC using quantum-mechanically derived descriptors. A curated dataset of experimentally validated donor–acceptor systems was constructed, and gradient boosting was benchmarked against other ML approaches. Lee showed that gradient boosting achieved superior accuracy for predicting Jsc, Voc, and PCE, owing to its ability to sequentially refine errors and capture subtle nonlinear descriptor–performance relationships. Feature-importance and SHAP analyses further revealed that energy-level offsets, exciton-related parameters, and charge-transfer descriptors play dominant roles in governing device performance. This study highlights gradient boosting as a particularly effective tool for extracting physically meaningful structure–property trends in OSCs when descriptor interactions are complex and highly nonlinear.

Das and Mondal (2024) [87] applied GBR as part of a data-driven framework to predict multiple photovoltaic parameters, Jsc, Voc, fill factor, and PCE in polymer–non-fullerene OSCs. Using a carefully curated database of 300 experimentally reported devices and high-quality quantum mechanical descriptors, gradient boosting emerged as the best-performing model for Jsc, Voc, and PCE prediction, outperforming linear and bagging-based approaches. They emphasized gradient boosting's strength in handling limited yet information-rich datasets, where sequential learning enables improved sensitivity to descriptor interactions linked to charge generation and voltage losses. The trained model was subsequently used for large-scale virtual screening of thousands of new donor–acceptor combinations, demonstrating the suitability of gradient boosting for both accurate prediction and accelerated materials discovery in OSC research.

Siddique et al. (2025) [88] utilized GBR to guide the design of high-efficiency small-molecule donors for OSC. Starting from a literature-derived dataset of donor molecules described by a large set of chem-informatics descriptors, they systematically compared more than forty ML models and identified gradient boosting as the most accurate predictor of PCE. The emphasis on gradient boosting lies in its robustness against descriptor redundancy and its capacity to learn complex structure–efficiency relationships without overfitting. The optimized model was then applied to virtually screen and evaluate over 10,000 newly generated donor candidates, followed by chemical similarity and synthetic accessibility analyses. This work

demonstrates how gradient boosting can serve as a practical engine for rapid donor material screening and design in OSC development.

Tahir et al. (2025) [89] employed GBR as the primary predictive model to accelerate the discovery of polymer donors for OSC. Using Mordred-derived molecular descriptors for experimentally reported polymer donors, they compared several regression algorithms and showed that gradient boosting consistently delivered the highest predictive accuracy for PCE. The study emphasized gradient boosting's ability to balance bias and variance while capturing nonlinear dependencies between molecular structure and device efficiency. Beyond prediction, the model was used to generate and evaluate a virtual library of polymer donors, from which high-performance candidates were selected based on predicted efficiency, structural similarity, and synthetic feasibility. This work reinforces gradient boosting as a powerful and reliable approach for polymer donor design and high-throughput screening in OSC research.

Several studies highlight the power of XGBoost, an advanced boosting algorithm, for OSC predictions. Wang et al. (2023) [90] employed XGBoost to establish a quantitative link between material energy-level alignment and the Voc of binary OSC, addressing the long-standing challenge of rational Voc optimization. Using a dataset compiled from more than 400 experimentally reported devices, they trained and compared multiple ML models and identified XGBoost as the most accurate predictor of Voc, outperforming RF and other regression approaches. Beyond prediction accuracy, the XGBoost model was used as an interpretative tool through SHAP analysis to identify the dominant physical drivers of voltage generation. This role is clearly illustrated in Fig. 5a–c, where the predicted–measured Voc correlations are evaluated using different feature combinations. Fig. 5a shows that reliable Voc prediction can already be achieved when only the two most important descriptors, the donor HOMO and acceptor LUMO, are included, highlighting their central role in voltage control. In contrast, Fig. 5b demonstrates that using weakly relevant features alone leads to poor predictive performance, while Fig. 5c confirms that combining key and auxiliary features improves model robustness. Together, these results demonstrate how XGBoost serves not only as a high-accuracy regression model but also as a physically meaningful framework for extracting energy-level matching strategies to guide the design of high-Voc binary OSCs.

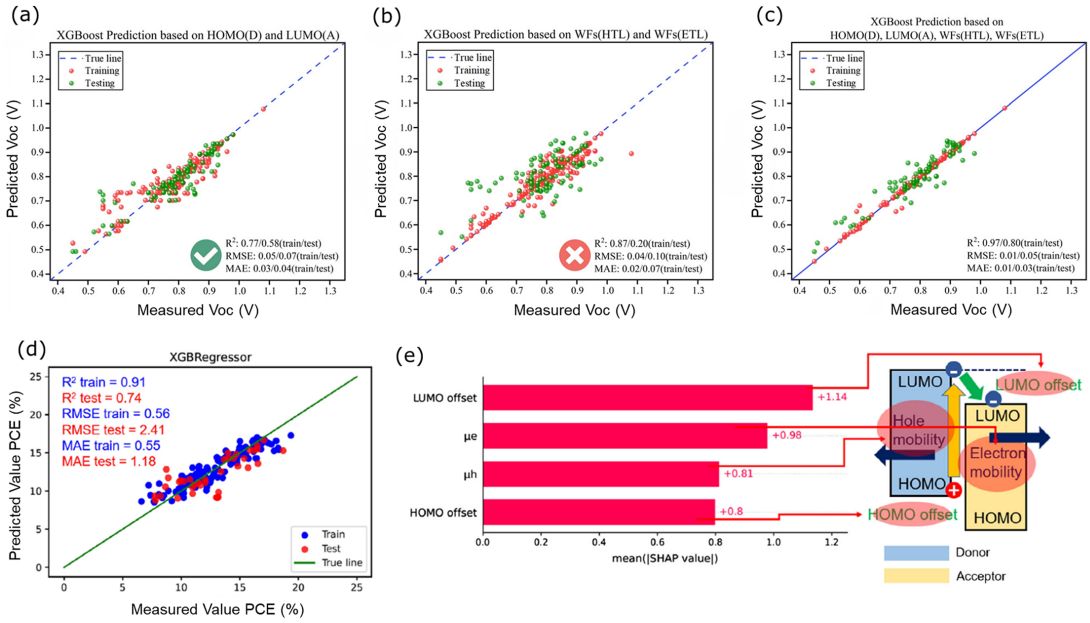

Figure 5: (Continued)

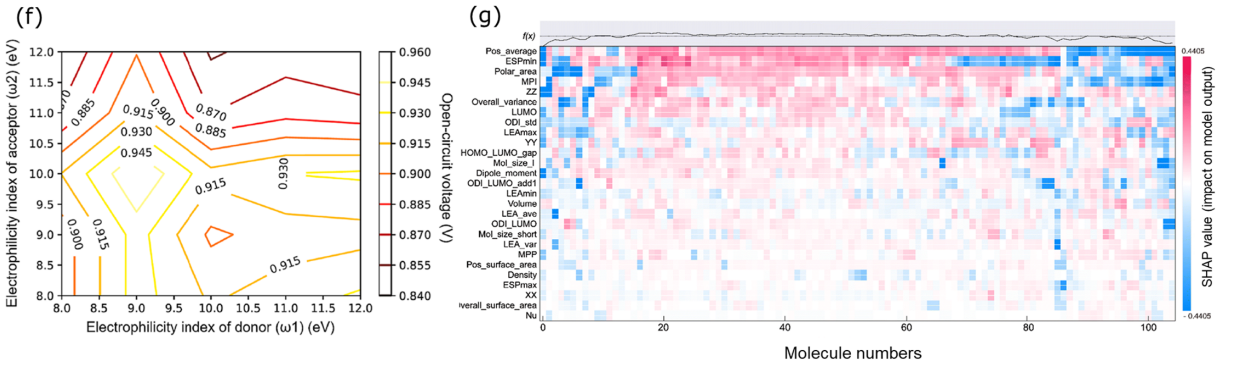

Figure 5: Performance and interpretability of XGBoost models in OSC studies. (**a**–**c**) Correlation between predicted and measured Voc using XGBoost based on (**a**) key energy-level descriptors (HOMO(D) and LUMO(A)), (**b**) less relevant descriptors (WF-HTL and WF-ETL), and (**c**) all four descriptors combined. Reproduced with permission from Ref. [90]. Copyright (2023) Royal Society of Chemistry. (**d**) Agreement between predicted and experimental PCE for training and testing datasets. (**e**) Feature importance derived from mean SHAP values in the XGBoost model using all descriptors. Reproduced with permission from Refs. [92,93]. Copyright (2023) Wiley-VCH GmbH. (**f**) XGBoost-predicted Voc contour map as a function of donor (ω_1) and acceptor (ω_2) electrophilicity indices. Reproduced with permission from Ref. [94]. Copyright (2024) Elsevier. (**g**) SHAP value heatmap showing the contribution of molecular descriptors to model predictions across 125 molecules, where color indicates the direction and magnitude of each feature's impact. Reproduced with permission from Ref. [96]. Copyright (2026) Royal Society of Chemistry.

Zhao et al. (2023) [91] applied XGBoost as a core predictive and screening tool to accelerate the design of high-efficiency all-small-molecule OSCs based on the Y6 acceptor. Using a literature-derived dataset of 85 experimentally reported Y6-based donor molecules, they constructed XGBoost regression models to predict PCE from a reduced set of carefully screened molecular structure descriptors. XGBoost was chosen for its strong performance on small datasets and its ability to capture nonlinear structure–performance relationships while controlling overfitting. Beyond prediction, the trained XGBoost model was integrated into a high-throughput virtual screening workflow, where nearly 10,000 new donor molecules were generated through scaffold–fragment recombination and ranked according to their predicted PCE. The highest-scoring candidates were further validated using DFT calculations, confirming favorable absorption spectra and energy-level alignment with Y6. This work demonstrates how XGBoost can function not only as an accurate regression model for OSC efficiency prediction but also as a practical engine for molecular screening and donor discovery in all-small-molecule OSCs.

Lee (2023) [92] applied XGBoost to accurately predict the PCE of NFA–based ternary OSC using a small set of experimentally accessible molecular descriptors. In this work, XGBoost was employed as a nonlinear regression model to capture complex relationships between device efficiency and effective molecular descriptors derived from frontier orbital energies, particularly electronegativity (χ) and hardness (η) of the donor, acceptor, and third component. The model achieved excellent predictive performance (training $R^2 \approx 0.98$, testing $R^2 \approx 0.71$), demonstrating that high accuracy can be obtained without relying on large numbers of structural or computationally expensive descriptors. As shown in Fig. 5d, the strong agreement between predicted and experimentally measured PCE values for both training and testing datasets highlights the robustness and generalization capability of the XGBoost model, validating its suitability for efficiency prediction in complex ternary systems. Overall, this study establishes XGBoost not only as a powerful predictive tool but also as a framework for extracting physically meaningful insights into energy-level matching and charge-transfer optimization in ternary OSC.

Lee (2023) [93] developed an interpretable ML framework using XGBoost to predict the Jsc of non-fullerene OSCs from experimentally accessible empirical descriptors. The XGBoost model, trained on a curated dataset of donor–acceptor systems, achieved high predictive accuracy ($R^2 > 0.8$) using only four physically meaningful features: donor–acceptor LUMO offset, HOMO offset, electron mobility, and hole mobility. Beyond prediction, XGBoost was coupled with SHAP to reveal the governing physical factors controlling photocurrent generation. As illustrated in Fig. 5e, the SHAP feature-importance analysis identifies the LUMO offset as the dominant contributor to Jsc, with charge carrier mobilities playing secondary roles and the HOMO offset having a comparatively weaker impact. This result quantitatively confirms that interfacial LUMO–LUMO energetic alignment is the primary driver of efficient photoinduced charge separation in non-fullerene OSC, providing a clear data-driven guideline for rational energy-level tuning in donor–acceptor material design.

Lee (2024) [94] employed an interpretable ML framework based on XGBoost to predict and analyze the Voc of non-fullerene acceptor–based ternary OSCs using both conventional electronic descriptors and chemically inspired global reactivity descriptors. In this work, XGBoost was specifically used as a nonlinear regression model to capture the complex, multivariate relationship between Voc and frontier molecular orbital energies as well as the electrophilicity index (ω) of the donor, acceptor, and third component, achieving robust predictive accuracy ($R^2 \approx 0.78$ for chemically inspired descriptors). Importantly, the integration of SHAP enabled physical interpretability, identifying the electrophilicity index of the acceptor (ω_2) as the dominant factor governing Voc enhancement in ternary blends. Fig. 5f visualizes this relationship through a ML-derived contour map of predicted Voc as a function of donor (ω_1) and acceptor (ω_2) electrophilicity indices, revealing an optimal descriptor region where Voc values approaching ~0.95 V can be achieved. This contour plot provides an intuitive design guideline, demonstrating that balanced electrophilicity between donor and acceptor materials is critical for maximizing voltage output in complex ternary OSC and highlighting the value of XGBoost-based models for rational energy-level engineering.

Lee et al. (2025) [95] employed an interpretable ML framework based on XGBoost to quantitatively predict the Jsc of NFA-based ternary OSC using a minimal set of physically meaningful energy-level descriptors. In this study, XGBoost was used as a nonlinear regression model to capture the complex relationships between Jsc and frontier molecular orbital energy offsets among donor, acceptor, and third-component materials, specifically ΔHOMO and ΔLUMO between different interfaces in the ternary blend. The model achieved reasonable predictive performance ($R^2 \approx 0.76$ for training and ≈0.65 for testing), demonstrating that experimentally accessible energy-offset descriptors are sufficient to describe photocurrent generation trends in ternary systems. By integrating SHAP, they further identified the HOMO offset between donor and acceptor (ΔHOMO(D–A)) as the most influential factor governing Jsc, highlighting the critical role of hole-transfer driving force in NFA-based ternary devices. Overall, this work establishes XGBoost as an effective and interpretable tool for linking energy-level alignment to photocurrent generation, providing a data-driven guideline for rational selection of third components and accelerated design of high-performance ternary OSCs.

Cui et al. (2026) [96] developed an intelligent data-driven framework that integrates large language model (LLM)–based literature mining, high-throughput quantum chemical calculations, and explainable ML to establish quantitative structure–property–performance relationships (QMSPRs) for Y-NFA in OSC. Within this framework, XGBoost was employed as a nonlinear regression model to predict device PCE from a carefully selected set of quantum-chemically derived molecular descriptors, while SHAP was used to ensure physical interpretability. The XGBoost–SHAP analysis revealed that electrostatic and polarity-related descriptors dominate PCE prediction, highlighting the critical role of molecular charge distribution in high-performance acceptors. Fig. 5g visualizes this interpretability at the molecule level through a SHAP

value heatmap, where each column represents an individual Y-series acceptor and each row corresponds to a molecular descriptor, with color intensity indicating the direction and magnitude of each descriptor's contribution to the predicted PCE. This figure demonstrates that XGBoost does not rely on a uniform feature hierarchy but instead captures molecule-specific combinations of descriptors, providing transparent insight into how distinct molecular designs translate into device performance.

4.3 Stacking-Based Ensemble Learning

Stacking-based ensemble learning combines predictions from multiple base models using a meta-learner, allowing different algorithms to contribute complementary strengths. This approach is particularly useful when datasets contain heterogeneous descriptors or when no single model consistently performs best across all tasks.

However, compared to bagging and boosting approaches, stacking methods require more careful validation design to avoid data leakage between base learners and the meta-learner. This added complexity, together with the relatively limited number of reported studies in OSC research, partly explains the less frequent application of stacking-based methods.

Wang et al. (2023) [97] employed a stacking-based ML strategy to accelerate materials discovery for OSC by integrating deep learning and ensemble learning into a unified screening framework. In their approach, a graph neural network was first used as a lower-level model to automatically extract structure-aware features directly from molecular graphs of donor and acceptor materials, capturing complex chemical information without relying on predefined descriptors or costly DFT calculations. These learned representations were then passed to a LightGBM gradient-boosted decision tree model acting as a higher-level ensemble learner, which predicted the PCE. This stacked architecture leverages the complementary strengths of deep learning (feature learning from molecular structure) and ensemble tree models (robust nonlinear regression and generalization), enabling fast, accurate, and scalable high-throughput screening of candidate organic photovoltaic materials. By validating predictions against reported devices and newly designed molecules, they demonstrated that stacking can effectively bridge molecular structure and device performance, providing a practical data-driven route for identifying high-efficiency NFA in OSC research.

Akbar et al. (2025) [98] applied a stacking-based ensemble ML approach to accelerate material discovery for OSC by targeting the identification of low-band-gap organic semiconductors that are favorable for high PCE. In their work, multiple base regression models were trained using molecular representations derived from RDKit fingerprints, each capturing different nonlinear relationships between chemical structure and electronic band gap. These individual predictions were then combined through a stacking ensemble regressor, which acted as a meta-learner to optimally weight and integrate the outputs of the base models, thereby reducing prediction bias and variance. The stacking strategy significantly outperformed all single models, achieving higher predictive accuracy and better generalization when evaluated on both test datasets and independent experimental OSC data from the literature. By reliably identifying low-band-gap donor and acceptor candidates, the study demonstrates that stacking ensembles provide a robust and scalable ML framework for screening organic photovoltaic materials, offering a practical route to guide molecular design and improve OSC performance beyond conventional single-model approaches.

Valiente et al. (2026) [99] employed a stacking-based ensemble strategy within an automated ML (AutoML) framework to model and predict the long-term performance degradation of polymer-based OSC, with a focus on P3HT:PCBM devices. Instead of relying on a single regression algorithm, their approach systematically benchmarked and combined multiple ML models with different inductive biases, allowing the AutoML pipeline to select and integrate the most informative learners into an optimized ensemble. This stacking-like methodology effectively captured the nonlinear and time-dependent relationships between

fabrication parameters, environmental conditions, and PCE decay over extended operational periods (>180 days). By aggregating information from diverse base models, the ensemble significantly improved prediction robustness and generalization, achieving very high accuracy ($R^2 > 0.9$) even for previously unseen devices. The study demonstrates that stacking-based ensemble learning is a powerful tool for OSC research, particularly for reliability and lifetime prediction, where complex degradation mechanisms cannot be adequately described by single-model approaches. Table 4 summarizes representative studies by categorizing the OSC system, prediction targets, dataset scale, descriptor representations, and ensemble learning models used.

Table 4: Representative datasets organized by ensemble learning strategy in OSC studies.

Year	Dataset Size	Materials Scope	Target	ML Model	Ref.
Bagging-based Ensemble					
2019	1700+ donors	Polymer donors	PCE	RF	[76]
2019	124 devices	Fullerene ternary	PCE	RF	[77]
2020	120+ ternary	Ternary	Voc	RF	[78]
2020	135 D/NFA	NFA	PCE	RF	[79]
2020	565 D/A	Binary NFA	PCE	RF	[80]
2021	157 NFA ternary	NFA ternary	PCE	RF	[69]
2021	1000 devices	Binary blends	Jsc	RF	[81]
2024	280 ternary	D:A1:A2 ternary	PCE	RF	[68]
2025	1343 acceptors	NFA	PCE	RF	[84]
Boosting-based Ensemble					
2023	1242 D/A	Polymer:NFA	Voc, Jsc, PCE	RF + GB	[82]
2023	400+ binary	Binary blends	Voc	XGBoost	[91]
2023	~100 ternary	NFA	PCE	XGBoost	[92]
2023	100+ binary	NFA	Jsc	XGBoost	[93]
2024	200+ D/A	Polymer:NFA	Voc, Jsc, PCE	GBDT	[86]
2024	300 devices	Polymer:NFA	Voc, Jsc, FF, PCE	GBR	[87]
2024	~120 ternary	NFA ternary	Voc	XGBoost	[94]
2025	1000+ donors	Small molecule donors	PCE	GBR	[88]
2025	800+ polymers	Polymer donors	PCE	GBR	[89]
2025	~150 ternary	NFA ternary blend	Jsc	XGBoost	[95]
2026	125 NFA	NFA	PCE	XGBoost	[96]
Stacking-Based Ensemble					
2023	1000+ molecules	Donor/NFA screening	PCE	GNN + LightGBM stacking	[97]

(Continued)

Table 4 (continued)

Year	Dataset Size	Materials Scope	Target	ML Model	Ref.
2025	1000+ candidates	Low bandgap materials	Bandgap	Multi-model stacking	[98]
2026	>500 time-series	P3HT:PCBM stability	PCE decay	AutoML stacking	[99]

4.4 Materials-Oriented Insights from Ensemble Learning

Ensemble learning models contribute not only to accurate prediction but also to interpretive analysis in OSC research. Their built-in mechanisms such as feature importance scoring, decision-tree structures, and model-agnostic tools like SHAP enable researchers to identify which molecular, morphological, or processing descriptors most strongly influence device performance [86,87,94,100]. This interpretability is particularly valuable in transforming complex data patterns into actionable design knowledge.

Rather than functioning as black-box regressors, ensemble models offer a transparent framework for evaluating the relationships between input variables and key metrics such as PCE, or stability. For example, SHAP-based analysis can highlight how variations in HOMO-LUMO offsets, dipole moments, or processing additives contribute to shifts in device output, while decision-tree structures help reveal threshold effects or nonlinear regimes that define performance plateaus.

Crucially, ensemble learning complements experimental and physics-based modeling by accelerating hypothesis generation and materials prioritization. It aids in narrowing the candidate space, suggesting optimal design windows, and flagging promising regions of the chemical and process parameter landscape for further investigation. This synergy is especially impactful under data-constrained conditions, where the cost or time associated with exhaustive physical testing remains prohibitive.

While this review focuses on ensemble learning methods, it is important to position these approaches within the broader ML landscape. In contrast to deep learning models, such as graph neural networks and transformer-based architectures that learn directly from raw molecular structures, ensemble methods are generally more effective for the small, heterogeneous, and descriptor-based datasets typical of OSC research. Deep learning approaches often require large, well-curated datasets and substantial computational resources, whereas ensemble models provide a practical balance between predictive performance, robustness, and interpretability under realistic data constraints. As such, these approaches should be viewed as complementary, with method selection depending on data availability, representation, and the specific objectives of OSC materials design.

To aid comparative understanding, Table 5 summarizes the core strengths, limitations, and typical applications of bagging-, boosting-, and stacking-based ensemble approaches in the OSC domain.

Table 5: Comparative overview of ensemble learning strategies in OSC research.

Ensemble Strategy	Key Strength	Limitations	Typical OSC Application
Bagging-based	• Reduces variance and overfitting • Performs well on small or noisy datasets • Offers feature importance for interpretability	• Less effective at capturing subtle nonlinear patterns • May require many base estimators for complex tasks	• Predicting PCE and related metrics • Screening donor/acceptor candidates • Interpreting key molecular or device descriptors
Boosting-based	• High predictive accuracy • Captures complex, nonlinear relationships • Effective for structured, tabular data	• Prone to overfitting without regularization • Requires careful tuning • Computationally more intensive	• Energy-level alignment modeling • Predicting PCE and related metrics • Predicting stability p • Process optimization and parameter tuning
Stacking-based	• Combines strengths of multiple models • Improves generalization, especially for complex tasks • Flexible across data types and targets	• Complex to implement and validate • Computationally demanding • Limited interpretability compared to single models	• Multi-target prediction (e.g., performance + stability) • Integrating material and process-level features

5 Challenges, Opportunities & Outlook

Despite the significant progress achieved through ML and ensemble learning approaches in OSC research, several fundamental challenges remain that limit their broader adoption and impact on practical materials design. Importantly, many of these challenges are not purely algorithmic in nature but arise from the intrinsic complexity of OSC materials systems, experimental variability, and the multifactorial trade-offs between efficiency, stability, and manufacturability. Addressing these challenges requires a shift from viewing ML solely as a predictive tool toward treating it as an integrated component of the OSC materials discovery pipeline.

5.1 Data Quality, Standardization, and Representativeness

One of the most persistent challenges in ML-assisted OSC research is the quality and consistency of available data. Most datasets are compiled from the literature, where variations in synthesis protocols, device architectures, and testing conditions introduce significant noise and bias. In particular, the prevalence of "champion bias," where high-efficiency devices are preferentially reported, leads to datasets that are not fully representative of the broader design space.

Even when nominally identical donor–acceptor systems are reported, differences in processing conditions, device stack configurations, and measurement protocols can lead to substantial performance variability. Such heterogeneity limits the reliability and generalizability of ensemble learning models. While these models are relatively robust to noise, they may still capture dataset-specific correlations rather than transferable physical relationships, leading to over-optimistic predictions that are difficult to reproduce experimentally.

Addressing these challenges requires improved data standardization, including consistent reporting of fabrication conditions, inclusion of negative or low-performance results, and the development of curated benchmark datasets. Recent advances in LLMs have introduced new opportunities for automated data extraction from scientific literature. In OSC research, LLM-based approaches can assist in parsing experimental reports, extracting device parameters, and structuring heterogeneous data into machine-readable formats [101]. Such tools have the potential to reduce manual data curation efforts and improve dataset scalability. However, challenges remain in ensuring extraction accuracy, consistency across reporting styles, and integration with existing ML pipelines. Without addressing these issues, the predictive capability of ML models will remain fundamentally constrained by data limitations.

5.2 Descriptor Limitations and Physical Interpretability

The choice of descriptors plays a central role in ML-based OSC research, as they define how materials and devices are represented in the model. In practice, two main types of descriptors are commonly used. Structure-based descriptors, such as molecular fingerprints derived from SMILES, enable efficient screening of large chemical spaces but often lack direct physical meaning. In contrast, physically motivated descriptors, such as HOMO/LUMO energy levels, bandgap, or dipole moment, are more interpretable but may not fully capture morphology and processing effects that also influence device performance.

Incorporating morphology and processing effects into ML models remains a significant challenge. Experimental techniques such as GIWAXS and AFM provide detailed information on nanoscale features, including crystallinity, domain size, and phase separation, but translating these measurements into standardized, quantitative descriptors suitable for ML is non-trivial. As a result, many studies simplify processing parameters, such as solvent additives or annealing conditions, into binary or categorical variables, which may not adequately capture their continuous and interdependent influence on device morphology and performance.

This leads to an inherent trade-off between scalability and physical relevance. While ensemble learning models can identify important features, these results should be interpreted with caution, as feature importance reflects statistical relationships within the dataset and does not necessarily indicate causality. Consequently, models may capture correlations that are valid only within a limited chemical and processing space, rather than uncovering fundamental physical mechanisms.

To improve interpretability, post hoc analysis techniques such as SHAP, LIME, permutation importance, and partial dependence plots are increasingly applied. SHAP is particularly useful because it quantifies both the magnitude and direction of each descriptor's contribution to individual predictions, while LIME provides

local, case-specific explanations of model behavior. In the context of OSC research, these methods can help identify key descriptors influencing device performance, including energy-level alignment, molecular structure, and processing conditions.

These interpretability approaches can provide practical guidance for materials design and process optimization by highlighting which descriptors most strongly influence predicted performance. For example, SHAP analysis may indicate that HOMO/LUMO alignment or specific solvent additives have a dominant impact on device efficiency, thereby suggesting directions for molecular tuning or processing optimization. Similarly, partial dependence plots can reveal trends between descriptors and target properties, offering additional insight into how performance responds to changes in key variables.

However, it is important to recognize that these methods primarily explain model behavior rather than underlying physical causation. As such, the derived insights should be treated as hypothesis-generating tools and validated through experimental or physics-based analysis. Their reliability is also strongly dependent on the quality, diversity, and representativeness of the underlying dataset.

To address these limitations, future work should focus on integrating multiple descriptor types, including electronic, structural, and processing-related features, as well as incorporating domain knowledge where possible. Such hybrid approaches can improve both predictive performance and physical interpretability, helping to bridge the gap between data-driven modeling and experimentally meaningful design strategies.

5.3 Model Validation, Generalization, and Applicability Domain

Reliable model validation is essential to ensure that ML predictions are meaningful and transferable. In many OSC studies, model performance is evaluated using simple train–test splits on relatively small datasets. While such approaches may report high accuracy, they can overestimate model performance when training and test data share similar characteristics.

A key concept in this context is the applicability domain, which defines the range of conditions for which a model can provide reliable predictions. Ensemble learning models generally perform well within the chemical and experimental space represented in the training data, but their reliability decreases when applied to new or structurally different materials.

To improve generalization, more rigorous validation strategies are required, including external validation using independent datasets and data-splitting methods that reflect realistic scenarios. In addition, uncertainty estimation can help identify predictions that are less reliable. Clearly defining the applicability domain is therefore essential to avoid overinterpretation and to ensure that ML predictions are used appropriately in guiding experimental research.

5.4 Stability, Degradation, and Practical Deployment

Although ML models have shown strong performance in predicting efficiency-related metrics such as PCE, long-term stability remains a major challenge in OSC research. Device degradation is influenced by multiple factors, including material properties, morphology evolution, and environmental conditions, making it difficult to model using conventional ML approaches.

One key limitation is the availability of data. Stability datasets are typically much smaller and less standardized than efficiency datasets, as they require long-term measurements under varying conditions. As a result, ML models trained on such data may have limited predictive reliability. While ensemble learning methods are relatively robust to small and noisy datasets, their effectiveness is still constrained by the quality and representativeness of the input data.

To address these challenges, future work should focus on integrating stability into ML frameworks by incorporating time-dependent data and developing standardized protocols for degradation studies. Combining performance and stability prediction within a unified framework is particularly important for identifying materials that are not only efficient but also suitable for real-world applications. Such efforts will help bridge the gap between laboratory-scale optimization and practical deployment of OSC technologies. These limitations highlight the need for stability-aware and manufacturability-oriented ML frameworks that better reflect real-world operating conditions.

5.5 Toward Closed-Loop and Physics-Informed Machine Learning Frameworks

Looking forward, one of the most transformative opportunities lies in the integration of ensemble learning models into closed-loop ML–experiment workflows, where predictions actively guide material synthesis, device fabrication, and subsequent data acquisition. Such iterative feedback loops have the potential to accelerate OSC materials discovery by focusing experimental efforts on the most informative regions of the design space, thereby reducing reliance on trial-and-error approaches.

In this context, active learning provides a practical strategy for implementing closed-loop discovery. Ensemble learning methods are particularly well suited for this paradigm, as the variance or disagreement among base learners can serve as a proxy for prediction uncertainty.

This enables the model to balance exploration of underrepresented regions with exploitation of high-performing candidates, which is especially important for OSC systems characterized by limited and heterogeneous datasets. At the same time, purely data-driven approaches remain constrained by data scarcity, dataset bias, and incomplete physical representation. Integrating ensemble learning with physics-based constraints and domain knowledge therefore represents a promising pathway toward more reliable and interpretable predictions. Embedding known relationships such as energy-level alignment, charge transport, and morphology–performance correlations can improve model robustness while reducing dependence on large datasets.

Despite these opportunities, several challenges remain, including the integration of heterogeneous experimental data, standardization of workflows, and the development of reliable uncertainty quantification. Addressing these issues will be essential for advancing toward practical closed-loop and active learning–driven OSC discovery systems.

6 Conclusions

This review has provided a materials-oriented synthesis of ML approaches applied to OSC research, with a particular focus on ensemble learning strategies for performance prediction and materials optimization. By framing ML as an enabling tool for OSC materials design rather than solely as a predictive technique, the review highlights how data-driven methods can address the intrinsic complexity arising from coupled molecular, morphological, and processing variables. Advances in donor polymers, non-fullerene acceptors (NFAs), and multicomponent blend architectures have rapidly expanded the OSC design space, making intuition-driven optimization increasingly insufficient. In this context, ML offers a scalable framework for extracting structure–property–performance relationships from heterogeneous experimental and computational data. The effectiveness of such approaches, however, depends strongly on the choice of data sources and descriptor representations, which must be aligned with specific materials design challenges to ensure meaningful and interpretable predictions.

The review demonstrates that ensemble learning methods including bagging-, boosting-, and stacking-based models are particularly well suited to OSC research due to their robustness under limited and

noisy data conditions. RF models enable reliable screening of diverse materials systems, boosting-based approaches capture subtle nonlinear effects associated with energy-level alignment and voltage losses, and stacking strategies facilitate the integration of heterogeneous descriptors. When applied in a materials-aware manner, these methods contribute not only to accurate performance prediction but also to practical guidance for materials selection and device optimization. Overall, ensemble learning has emerged as a practical and interpretable ML framework for advancing OSC materials research. Its impact lies in its ability to operate effectively within realistic experimental constraints while providing insight into the key parameters governing device performance. Continued integration of ensemble learning with careful data curation and materials-focused analysis will be essential for accelerating the rational development of next-generation OSCs.

Acknowledgement: Shafidah Shafian acknowledges the support provided by the Geran Translasi UKM (TR-UKM), grant number UKM-TR2024-09, funded by Universiti Kebangsaan Malaysia (UKM). The authors declare that no generative artificial intelligence (AI) or AI-assisted technologies were used to generate the scientific content, ideas, or analysis presented in this manuscript. AI tools were used exclusively for language polishing and improving readability, under full human oversight. The manuscript was carefully reviewed and revised by the authors to ensure its quality and consistency.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Shafidah Shafian; methodology, Shafidah Shafian; validation, Shafidah Shafian and Azlan Ismail; investigation, Shafidah Shafian and Azlan Ismail; data curation, Shafidah Shafian; writing—original draft preparation, Shafidah Shafian; writing—review and editing, Shafidah Shafian and Azlan Ismail; visualization, Shafidah Shafian; project administration, Shafidah Shafian; funding acquisition, Shafidah Shafian. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: This article does not involve data availability.

Ethics Approval: Not applicable.

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

References

1. Riede M, Spoltore D, Leo K. Organic solar cells—the path to commercial success. Adv Energy Mater. 2021;11(1):2002653. doi:10.1002/aenm.202002653.
2. Liu Y, Liu B, Ma C-Q, Huang F, Feng G, Chen H, et al. Recent progress in organic solar cells (Part I material science). Sci China Chem. 2022;65(2):224–68. doi:10.1007/s11426-021-1180-6.
3. Liu Y, Liu B, Ma C-Q, Huang F, Feng G, Chen H, et al. Recent progress in organic solar cells (Part II device engineering). Sci China Chem. 2022;65(8):1457–97. doi:10.1007/s11426-022-1256-8.
4. Research P. Best research-cell efficiency chart: national renewable energy laboratory (NREL). 2025 [cited 2026 Jan 1]. Available from: https://www.nrel.gov/pv/cell-efficiency.
5. Zhang Y, Deng W, Petoukhoff CE, Xia X, Lang Y, Xia H, et al. Achieving 19.4% organic solar cell via an in situ formation of pin structure with built-in interpenetrating network. Joule. 2024;8(2):509–26. doi:10.1016/j.joule.2023.12.009.
6. Basu R, Gumpert F, Lohbreier J, Morin P-O, Vohra V, Liu Y, et al. Large-area organic photovoltaic modules with 14.5% certified world record efficiency. Joule. 2024;8(4):970–8. doi:10.1016/j.joule.2024.02.016.
7. Xie C, Jiang X, Zhu Q, Wang D, Xiao C, Liu C, et al. Mechanical robust flexible single-component organic solar cells. Small Methods. 2021;5(9):2100481. doi:10.1002/smtd.202100481.
8. Qin F, Sun L, Chen H, Liu Y, Lu X, Wang W, et al. 54 cm^2 large-area flexible organic solar modules with efficiency above 13%. Adv Mater. 2021;33(39):2103017. doi:10.1002/adma.202103017.

9. Salehin FNM, Chelvanathan P, Goje AA, Ludin NA, Ibrahim MA, Shafian S. Design of blue, green and red colorful semitransparent films using Ag/SnO_2/Ag color filter for integrated into solar cells. Results Phys. 2025;70:108172. doi:10.1016/j.rinp.2025.108172.
10. Shafian S. Optical and structural analysis of layer thickness effects on transmission in semitransparent colorful organic photovoltaics. PaperASIA. 2025;41(4b):1–10. doi:10.59953/paperasia.v41i4b.569.
11. He K, Li X, Liu H, Zhang Z, Kumar P, Ngai JH, et al. D-a polymer with a donor backbone-acceptor-side-chain structure for organic solar cells. Asian J Org Chem. 2020;9(9):1301–8. doi:10.1002/ajoc.202000172.
12. Wan X, Li C, Zhang M, Chen Y. Acceptor–donor–acceptor type molecules for high performance organic photovoltaics–chemistry and mechanism. Chem Soc Rev. 2020;49(9):2828–42. doi:10.1039/d0cs00084a.
13. Zhao J, Yao C, Ali MU, Miao J, Meng H. Recent advances in high-performance organic solar cells enabled by acceptor–donor–acceptor–donor–acceptor (A-DA′ D-A) type acceptors. Mater Chem Front. 2020;4(12):3487–504. doi:10.1039/d0qm00305k.
14. Shin S, Shafian S, Ryu KY, Jeon YK, Kim WS, Kim K. Solution-processed TiO_2 nanoparticles functionalized with catechol derivatives as electron transporting layer materials for organic photovoltaics. Adv Mater Interfaces. 2022;9(14):2200118. doi:10.1002/admi.202200118.
15. Ryu KY, Shafian S, Shin J, Lee YJ, Lee M, Kim K. Linear polyurethane ionenes for stable interlayer of organic photovoltaics. J Power Sources. 2022;542:231772. doi:10.1016/j.jpowsour.2022.231772.
16. Hong M, Youn J, Ryu KY, Shafian S, Kim K. Improving the stability of non-fullerene-based organic photovoltaics through sequential deposition and utilization of a quasi-orthogonal solvent. ACS Appl Mater Interfaces. 2023;15(16):20151–8. doi:10.1021/acsami.3c02071.
17. Kim H, Heo Y, Na Y, Shafian S, Kim B, Kim K. Cross-linking-integrated sequential deposition: a method for efficient and reproducible bulk heterojunctions in organic solar cells. ACS Appl Mater Interfaces. 2024;16(41):55873–80.
18. Kim H, Kong Y-J, Kim W-S, Shafian S, Kim K. Enhancing reproducibility in organic solar cell fabrication via static sequential deposition with cross-linked polymer donor and nonfullerene acceptor. ACS Appl Polym Mater. 2024;6(10):5814–21. doi:10.1021/acsapm.4c00477.
19. Zhi HF, Jiang M, Zhang H, An Q, Bai HR, Jee MH, et al. Isomeric small molecule donor with terminal branching position directly attached to the backbone enables efficient all-small-molecule organic solar cells with excellent stability. Adv Funct Mater. 2023;33(25):2300878. doi:10.1002/adfm.202300878.
20. Günther M, Kazerouni N, Blätte D, Perea JD, Thompson BC, Ameri T. Models and mechanisms of ternary organic solar cells. Nat Rev Mater. 2023;8(7):456–71.
21. Xu X, Li Y, Peng Q. Ternary blend organic solar cells: understanding the morphology from recent progress. Adv Mater. 2022;34(46):2107476. doi:10.1002/adma.202107476.
22. Ma Q, Jia Z, Meng L, Zhang J, Zhang H, Huang W, et al. Promoting charge separation resulting in ternary organic solar cells efficiency over 17.5%. Nano Energy. 2020;78:105272. doi:10.1016/j.nanoen.2020.105272.
23. Zhang P, Zhang Z, Sun H, Li J, Chen Y, Wang J, et al. Reducing the voltage loss of Y-series acceptor based organic solar cells via ternary/quaternary strategies. Chin Chem Lett. 2024;35(2):108802. doi:10.1016/j.cclet.2023.108802.
24. Du M, Sun N, Cheng H, Liu X, Yi X, Guo Q, et al. Multiple-birth-acceptor: easily-synthesized mixture for easily-fabricated quaternary organic solar cells with beyond 20% efficiency. Angew Chem. 2025;137(49):e202515114.
25. Shafian S, Mohd Salehin FN, Lee S, Ismail A, Mohamed Shuhidan S, Xie L, et al. Development of organic semiconductor materials for organic solar cells via the integration of computational quantum chemistry and AI-powered machine learning. ACS Appl Energy Mater. 2025;8(2):699–722. doi:10.1021/acsaem.4c02937.
26. Shafian S, Husen MN, Xie L, Kim K. Predicting high-performance perovskite solar cells using AI-based machine learning models. Mater Today Sustain. 2025;31:101176.
27. Sahu H, Rao W, Troisi A, Ma H. Toward predicting efficiency of organic solar cells via machine learning and improved descriptors. Adv Energy Mater. 2018;8(24):1801032. doi:10.1002/aenm.201801032.
28. Padula D, Simpson JD, Troisi A. Combining electronic and structural features in machine learning models to predict organic solar cells properties. Mater Horiz. 2019;6(2):343–9. doi:10.1039/c8mh01135d.

29. Ahmed DR, Muhammadsharif FF. A review of machine learning in organic solar cells. Processes. 2025;13(2):393. doi:10.3390/pr13020393.
30. Mahmood A, Wang J-L. Machine learning for high performance organic solar cells: current scenario and future prospects. Energy Environ Sci. 2021;14(1):90–105. doi:10.1039/d0ee02838j.
31. Miyake Y, Saeki A. Machine learning-assisted development of organic solar cell materials: issues, analyses, and outlooks. J Phys Chem Lett. 2021;12(51):12391–401. doi:10.1021/acs.jpclett.1c03526.
32. Dong X, Yu Z, Cao W, Shi Y, Ma Q. A survey on ensemble learning. Front Comput Sci. 2020;14(2):241–58.
33. Mienye ID, Sun Y. A survey of ensemble learning: concepts, algorithms, applications, and prospects. IEEE Access. 2022;10:99129–49.
34. Zhou Z-H. Ensemble learning. In: Machine learning. Berlin/Heidelberg, Germany: Springer; 2021. p. 181–210.
35. Kunapuli G. Ensemble methods for machine learning. New York, NY, USA: Simon and Schuster; 2023. 352 p.
36. Feng J, Wang H, Ji Y, Li Y. Molecular design and performance improvement in organic solar cells guided by high-throughput screening and machine learning. Nano Sel. 2021;2(9):1629–41. doi:10.1002/nano.202100006.
37. Zhao ZW, Geng Y, Troisi A, Ma H. Performance prediction and experimental optimization assisted by machine learning for organic photovoltaics. Adv Intell Syst. 2022;4(6):2100261. doi:10.1002/aisy.202100261.
38. Jiang Y, Yao C, Yang Y, Wang J. Machine learning approaches for predicting power conversion efficiency in organic solar cells: a comprehensive review. Sol RRL. 2024;8(22):2400567. doi:10.1002/solr.202400567.
39. Shafian S, Hwang H, Kim K. Near infrared organic photodetector utilizing a double electron blocking layer. Opt Express. 2016;24(22):25308–16. doi:10.1364/oe.24.025308.
40. Hwang H, Lee H, Shafian S, Lee W, Seok J, Ryu KY, et al. Thermally stable bulk heterojunction prepared by sequential deposition of nanostructured polymer and fullerene. Polymers. 2017;9(9):456. doi:10.3390/polym9090456.
41. You Y-J, Saeed MA, Shafian S, Kim J, Kim SH, Kim SH, et al. Energy recycling under ambient illumination for internet-of-things using metal/oxide/metal-based colorful organic photovoltaics. Nanotechnology. 2021;32(46):465401. doi:10.1088/1361-6528/ac13e7.
42. Yuan J, Zhang C, Qiu B, Liu W, So SK, Mainville M, et al. Effects of energetic disorder in bulk heterojunction organic solar cells. Energy Environ Sci. 2022;15(7):2806–18. doi:10.1039/d2ee00271j.
43. He X, Liu ZX, Chen H, Li CZ. Selectively modulating componential morphologies of bulk heterojunction organic solar cells. Adv Mater. 2024;36(7):2306681. doi:10.1002/adma.202306681.
44. Yi J, Zhang G, Yu H, Yan H. Advantages, challenges and molecular design of different material types used in organic solar cells. Nat Rev Mater. 2024;9(1):46–62. doi:10.1038/s41578-023-00618-1.
45. Zhang T, An C, Cui Y, Zhang J, Bi P, Yang C, et al. A universal nonhalogenated polymer donor for high-performance organic photovoltaic cells. Adv Mater. 2022;34(2):2105803. doi:10.1002/adma.202105803.
46. Guo C, Fu Y, Li D, Wang L, Zhou B, Chen C, et al. A polycrystalline polymer donor as pre-aggregate toward ordered molecular aggregation for 19.3% efficiency binary organic solar cells. Adv Mater. 2023;35(41):2304921. doi:10.1002/adma.202304921.
47. Wang Y, Wang Y, Zhu L, Liu H, Fang J, Guo X, et al. A novel wide-bandgap small molecule donor for high efficiency all-small-molecule organic solar cells with small non-radiative energy losses. Energy Environ Sci. 2020;13(5):1309–17. doi:10.1039/c9ee04199k.
48. Gao H, Sun Y, Meng L, Han C, Wan X, Chen Y. Recent progress in all-small-molecule organic solar cells. Small. 2023;19(3):2205594. doi:10.1002/smll.202205594.
49. Shafian S, Jang Y, Kim K. Solution processed organic photodetector utilizing an interdiffused polymer/fullerene bilayer. Opt Express. 2015;23(15):A936–46. doi:10.1364/oe.23.00a936.
50. Chandrasekaran N, Kumar A, Thomsen L, Kabra D, McNeill CR. High performance as-cast P3HT: PCBM devices: understanding the role of molecular weight in high regioregularity P3HT. Mater Adv. 2021;2(6):2045–54.
51. Chatterjee S, Jinnai S, Ie Y. Nonfullerene acceptors for P3HT-based organic solar cells. J Mater Chem A. 2021;9(35):18857–86. doi:10.1039/d1ta03219d.
52. Seo S, Lee JW, Kim DJ, Lee D, Phan TNL, Park J, et al. Poly (dimethylsiloxane)-block-PM6 polymer donors for high-performance and mechanically robust polymer solar cells. Adv Mater. 2023;35(24):2300230. doi:10.1002/adma.202300230.

53. Wei N, Chen J, Cheng Y, Bian Z, Liu W, Song H, et al. Constructing multiscale fibrous morphology to achieve 20% efficiency organic solar cells by mixing high and low molecular weight D18. Adv Mater. 2024;36(41):2408934. doi:10.1002/adma.202408934.
54. Zhang Q, Yuan X, Feng Y, Larson BW, Su GM, Maung Maung Y, et al. Understanding the interplay of transport-morphology-performance in PBDB-T-based polymer solar cells. Sol RRL. 2020;4(4):1900524. doi:10.1002/solr.201900524.
55. Busireddy MR, Chen T-W, Huang S-C, Su Y-J, Wang Y-M, Chuang W-T, et al. PBDB-T-based binary-OSCs achieving over 15.83% efficiency via end-group functionalization and alkyl-chain engineering of quinoxaline-containing non-fullerene acceptors. ACS Appl Mater Interfaces. 2022;14(36):41264–74. doi:10.1021/acsami.2c09614.
56. Tang H, Yan C, Huang J, Kan Z, Xiao Z, Sun K, et al. Benzodithiophene-based small-molecule donors for next-generation all-small-molecule organic photovoltaics. Matter. 2020;3(5):1403–32. doi:10.1016/j.matt.2020.09.001.
57. Malhotra P, Khandelwal K, Biswas S, Chen F-C, Sharma GD. Opportunities and challenges for machine learning to select combination of donor and acceptor materials for efficient organic solar cells. J Mater Chem C. 2022;10(47):17781–811. doi:10.1039/d2tc03276g.
58. Mumyatov AV, Troshin PA. A review on fullerene derivatives with reduced electron affinity as acceptor materials for organic solar cells. Energies. 2023;16(4):1924. doi:10.3390/en16041924.
59. Li X, Kong X, Sun G, Li Y. Organic small molecule acceptor materials for organic solar cells. Escience. 2023;3(5):100171. doi:10.1016/j.esci.2023.100171.
60. Luo D, Jang W, Babu DD, Kim MS, Wang DH, Kyaw AKK. Recent progress in organic solar cells based on non-fullerene acceptors: materials to devices. J Mater Chem A. 2022;10(7):3255–95. doi:10.1039/d1ta10707k.
61. Wu Q, Ding S, Sun A, Xia Y. Recent progress on non-fullerene acceptor materials for organic solar cells. Mater Today Chem. 2024;41:102290. doi:10.1016/j.mtchem.2024.102290.
62. Armin A, Li W, Sandberg OJ, Xiao Z, Ding L, Nelson J, et al. A history and perspective of non-fullerene electron acceptors for organic solar cells. Adv Energy Mater. 2021;11(15):2003570. doi:10.1002/aenm.202003570.
63. Huang D, Wang K, Li Z, Zhou H, Zhao X, Peng X, et al. A machine learning prediction model for quantitative analyzing the influence of non-radiative voltage loss on non-fullerene organic solar cells. Chem Eng J. 2023;475:145958.
64. Xu X, Yu L, Meng H, Dai L, Yan H, Li R, et al. Polymer solar cells with 18.74% efficiency: from bulk heterojunction to interdigitated bulk heterojunction. Adv Funct Mater. 2022;32(4):2108797. doi:10.1002/adfm.202108797.
65. Wang J, Wang Y, Bi P, Chen Z, Qiao J, Li J, et al. Binary organic solar cells with 19.2% efficiency enabled by solid additive. Adv Mater. 2023;35(25):2301583.
66. Chen X, Wang D, Wang Z, Li Y, Zhu H, Lu X, et al. 18.02% Efficiency ternary organic solar cells with a small-molecular donor third component. Chem Eng J. 2021;424(9):130397. doi:10.1016/j.cej.2021.130397.
67. Cao J, Xu Z. Providing a photovoltaic performance enhancement relationship from binary to ternary polymer solar cells via machine learning. Polymers. 2024;16(11):1496. doi:10.3390/polym16111496.
68. Li J-H, Zhang C-R, Zhang M-L, Liu X-M, Gong J-J, Chen Y-H, et al. Machine learning study of D: A1: A2 ternary organic solar cells. Org Electron. 2024;125:106988. doi:10.1016/j.orgel.2023.106988.
69. Hao T, Leng S, Yang Y, Zhong W, Zhang M, Zhu L, et al. Capture the high-efficiency non-fullerene ternary organic solar cells formula by machine-learning-assisted energy-level alignment optimization. Patterns. 2021;2(9):1–12. doi:10.1016/j.patter.2021.100333.
70. Duan L, Uddin A. Progress in stability of organic solar cells. Adv Sci. 2020;7(11):1903259. doi:10.1002/advs.201903259.
71. Ding P, Yang D, Yang S, Ge Z. Stability of organic solar cells: toward commercial applications. Chem Soc Rev. 2024;53(5):2350–87. doi:10.1039/d3cs00492a.
72. Du X, Lüer L, Heumueller T, Classen A, Liu C, Berger C, et al. Revealing processing stability landscape of organic solar cells with automated research platforms and machine learning. InfoMat. 2024;6(7):e12554. doi:10.1002/inf2.12554.
73. David TW, Anizelli H, Jacobsson TJ, Gray C, Teahan W, Kettle J. Enhancing the stability of organic photovoltaics through machine learning. Nano Energy. 2020;78(3):105342. doi:10.1016/j.nanoen.2020.105342.

74. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. J Chem Inf Comput Sci. 1988;28(1):31–6. doi:10.1021/ci00057a005.
75. Sarmah U, Borah P, Bhattacharyya DK. Ensemble learning methods: an empirical study. SN Comput Sci. 2024;5(7):924.
76. Sun W, Zheng Y, Yang K, Zhang Q, Shah AA, Wu Z, et al. Machine learning–assisted molecular design and efficiency prediction for high-performance organic photovoltaic materials. Sci Adv. 2019;5(11):eaay4275. doi:10.1126/sciadv.aay4275.
77. Lee MH. Insights from machine learning techniques for predicting the efficiency of fullerene derivatives-based ternary organic solar cells at ternary blend design. Adv Energy Mater. 2019;9(26):1900891. doi:10.1002/aenm.201900891.
78. Lee M-H. A machine learning-based design rule for improved open-circuit voltage in ternary organic solar cells. Adv Intell Syst. 2020;2(1):1900108. doi:10.1002/aisy.201900108.
79. Lee M-H. Robust random forest based non-fullerene organic solar cells efficiency prediction. Org Electron. 2020;76(10):105465. doi:10.1016/j.orgel.2019.105465.
80. Wu Y, Guo J, Sun R, Min J. Machine learning for accelerating the discovery of high-performance donor/acceptor pairs in non-fullerene organic solar cells. npj Comput Mater. 2020;6(1):120. doi:10.1038/s41524-020-00388-2.
81. Rodríguez-Martínez X, Pascual-San-José E, Fei Z, Heeney M, Guimerà R, Campoy-Quiles M. Predicting the photocurrent–composition dependence in organic solar cells. Energy Environ Sci. 2021;14(2):986–94.
82. Suthar R, Abhijith T, Karak S. Machine-learning-guided prediction of photovoltaic performance of non-fullerene organic solar cells using novel molecular and structural descriptors. J Mater Chem A. 2023;11(41):22248–58. doi:10.1039/d3ta04603f.
83. Zhao X, Lei M, Wang K, Peng X, Li Z, Zhou H, et al. Exploring the impact of fabrication parameters in organic solar cells with PM6: Y6 using machine learning. AIP Adv. 2024;14(6):065325. doi:10.1063/5.0201580.
84. Liu X, Zhang X, Sheng Y, Zhang Z, Xiong P, Ju X, et al. Advancing organic photovoltaic materials by machine learning-driven design with polymer-unit fingerprints. npj Comput Mater. 2025;11(1):107. doi:10.1038/s41524-025-01608-3.
85. Sahu H, Yang F, Ye X, Ma J, Fang W, Ma H. Designing promising molecules for organic solar cells via machine learning assisted virtual screening. J Mater Chem A. 2019;7(29):17480–8. doi:10.1039/c9ta04097h.
86. Lee M-H. Predicting and analyzing the fill factor of non-fullerene organic solar cells based on material properties and interpretable machine-learning strategies. Sol Energy. 2024;267:112191. doi:10.1016/j.solener.2023.112191.
87. Das B, Mondal A. Predictive modeling and design of organic solar cells: a data-driven approach for material innovation. ACS Appl Energy Mater. 2024;7(20):9349–63.
88. Siddique B, Alomar TS, Tahir MH, AlMasoud N, El-Bahy ZM. Designing of small molecule donors with the help of machine learning for organic solar cells and performance prediction. J Photochem Photobiol A Chem. 2025;459:116026. doi:10.1016/j.jphotochem.2024.116026.
89. Tahir MH, Farrukh A, Alqahtany FZ, Badshah A, Shaaban IA, Assiri MA. Accelerated discovery of polymer donors for organic solar cells through machine learning: from library creation to performance forecasting. Spectrochim Acta Part A Mol Biomol Spectrosc. 2025;326:125298. doi:10.1016/j.saa.2024.125298.
90. Wang K, Guo C, Li Z, Zhang R, Feng Z, Fang G, et al. Machine learning assisted identification of the matched energy level of materials for high open circuit voltage in binary organic solar cells. Mol Syst Des Eng. 2023;8(6):799–809. doi:10.1039/D2ME00265E.
91. Zhao Q, Shan Y, Zhou H, Zhang G, Liu W. Machine learning-assisted performance prediction and molecular design of all-small-molecule organic solar cells based on the Y6 acceptor. Sol Energy. 2023;265:112115. doi:10.1016/j.solener.2023.112115.
92. Lee M-H. Interpretable machine learning model for the highly accurate prediction of efficiency of ternary organic solar cells based on nonfullerene acceptor using effective molecular descriptors. Sol RRL. 2023;7(14):2300307. doi:10.1002/solr.202300307.
93. Lee M-H. Frontier molecular orbital offset as an empirical descriptor for predicting short circuit current of nonfullerene organic solar cells. Sol RRL. 2023;7(18):2300533. doi:10.1002/solr.202300533.

94. Lee M-H. Investigation of the open-circuit voltage of non-fullerene acceptors-based ternary organic solar cells based on interpretable machine-learning approach and chemically inspired descriptors. J Photochem Photobiol A Chem. 2024;450:115430. doi:10.1016/j.jphotochem.2023.115430.
95. Lee MH, Chen YC, Chang YM, Hou B. Investigation of short-circuit current density in non-fullerene-based ternary organic solar cells by incorporating machine learning algorithms with effective descriptors. Sol RRL. 2025;9(10):2500167. doi:10.1002/solr.202500167.
96. Cui Y, Ma W, Yan H. Developing an intelligent data-driven framework for organic photovoltaic research. EES Solar. 2026;2(1):246–56. doi:10.1039/d5el00160a.
97. Wang H, Feng J, Dong Z, Jin L, Li M, Yuan J, et al. Efficient screening framework for organic solar cells with deep learning and ensemble learning. npj Comput Mater. 2023;9(1):200. doi:10.1038/s41524-023-01155-9.
98. Akbar B, Khan A, Tayara H. Optimized selection of low band gap semiconductors through integrated machine learning and experimental validation for enhanced organic solar cell performance. Soc Sci Res Netw. 2025 [cited 2026 Jan 1]. Available at SSRN: https://ssrn.com/abstract=5644270.
99. Valiente D, Rodríguez-Mas F, Alegre-Requena JV, Dalmau D, Flores M, Ferrer JC. General machine learning models for interpreting and predicting efficiency degradation in organic solar cells. Expert Syst Appl. 2026;296:128890. doi:10.2139/ssrn.5147681.
100. Song Z, Cao S, Yang H. An interpretable framework for modeling global solar radiation using tree-based ensemble machine learning and Shapley additive explanations methods. Appl Energy. 2024;364(11):123238. doi:10.1016/j.apenergy.2024.123238.
101. Sahu H, Mahmood A, Shafique LB, Ramprasad R. From corpus to innovation: advancing organic solar cell design with large language models. npj Comput Mater. 2025;12:27.