Advanced Powder Materials 5 (2026) 100431

Contents lists available at ScienceDirect

Advanced Powder Materials

journal homepage: www.keaipublishing.com/en/journals/advanced-powder-materials

Machine learning for carbon dots: Capabilities, limitations, and the path toward rational design

Shudie Shen[a], Leilei Zhang[a], Linquan Gan[a], Tao Li[a], Wenli Zhao[a], Yanfei Wang[a,c], Liang Zhu[a,b,*], Xiaoyu Zhao[a,b,c,**]

[a] Tianjin Key Laboratory of Brine Chemical Engineering and Resource Eco-utilization, College of Chemical Engineering and Materials Science, Tianjin University of Science and Technology, Tianjin, 300457, China
[b] State Key Laboratory of Bio-based Fiber Materials, Tianjin University of Science and Technology, Tianjin, 300457, China
[c] College of Energy Storage, Tianjin University of Science and Technology, Tianjin, 300457, China

ARTICLE INFO

Keywords:
Carbon dots
Machine learning
Materials design
Performance prediction
Materials application and development

ABSTRACT

Carbon dots (CDs), as significant luminescent carbon-based nanomaterials, exhibit broad application potential in sensing, biomedicine, and optoelectronic devices. However, their synthesis processes are highly nonlinear, and their structural heterogeneity is prominent, resulting in a long-standing lack of a decipherable framework for the synthesis-structure-property relationship, which has restricted on-demand design and controllable preparation. In recent years, machine learning (ML) has provided a data-driven paradigm to address this complexity, yet related research still faces challenges such as scattered data, unclear task boundaries, and insufficient model interpretability. This review systematically examines recent key advances in machine learning for CDs research, focusing on three core tasks: property prediction, synthesis and inverse design, and mechanism analysis. It critically analyzes the capabilities and limitations of various models in predicting emission wavelength, photoluminescence quantum yield, and phosphorescence lifetime. By comparing data sources, feature construction, and validation strategies, it points out that many current high-prediction accuracies primarily stem from statistical fitting rather than learning physical causality, particularly facing structural data bottlenecks in red or near-infrared emission and cross-system generalization. Furthermore, from the perspective of CDs application systems, the review systematically evaluates the practical enabling role of machine learning in CDs-related applications such as sensing, biomedicine, optoelectronics, and information encryption, clearly distinguishing between its role as a tool for performance optimization and its function as a key means for rational design of CDs materials. Finally, a future-oriented pathway for machine learning-driven CDs research is proposed: by constructing standardized, high-quality databases, introducing physically constrained and interpretable models, and combining active learning with closed-loop experimental validation, the field can advance from empirical trial-and-error toward a predictive, interpretable, and translatable paradigm of rational design.

1. Introduction

Carbon dots (Fig. 1), a class of nanomaterials with sizes typically below 10 nm [10–12], detailed information is presented in Table 1, possess distinct advantages such as tunable luminescence, excellent biocompatibility [13], high stability [14], and low toxicity [15]. They demonstrate broad application potential across multiple cutting-edge fields, including sensing [16], biomedicine [17], photocatalysis and energy applications [18–20], light-emitting diodes [21], and anti-counterfeiting encryption [22–24]. However, the synthesis–structure–property relationship of CDs remains complex and inadequately understood. The multivariate, nonlinear, and strongly coupled

* Corresponding author. Tianjin Key Laboratory of Brine Chemical Engineering and Resource Eco-utilization, College of Chemical Engineering and Materials Science, Tianjin University of Science and Technology, Tianjin 300457, China.
** Corresponding author. Tianjin Key Laboratory of Brine Chemical Engineering and Resource Eco-utilization, College of Chemical Engineering and Materials Science, Tianjin University of Science and Technology, Tianjin 300457, China.
E-mail addresses: shenshudie@mail.tust.edu.cn (S. Shen), 18822049585@163.com (L. Zhang), 2493213623@qq.com (L. Gan), litao@mail.tust.edu.cn (T. Li), zhaowenli@tust.edu.cn (W. Zhao), wangyanfei@tust.edu.cn (Y. Wang), zhuliang@tust.edu.cn (L. Zhu), xyz@tust.edu.cn (X. Zhao).
Peer review under the responsibility of Central South University.

https://doi.org/10.1016/j.apmate.2026.100431
Received 12 February 2026; Received in revised form 9 May 2026; Accepted 10 May 2026
Available online 22 May 2026
2772-834X/© 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

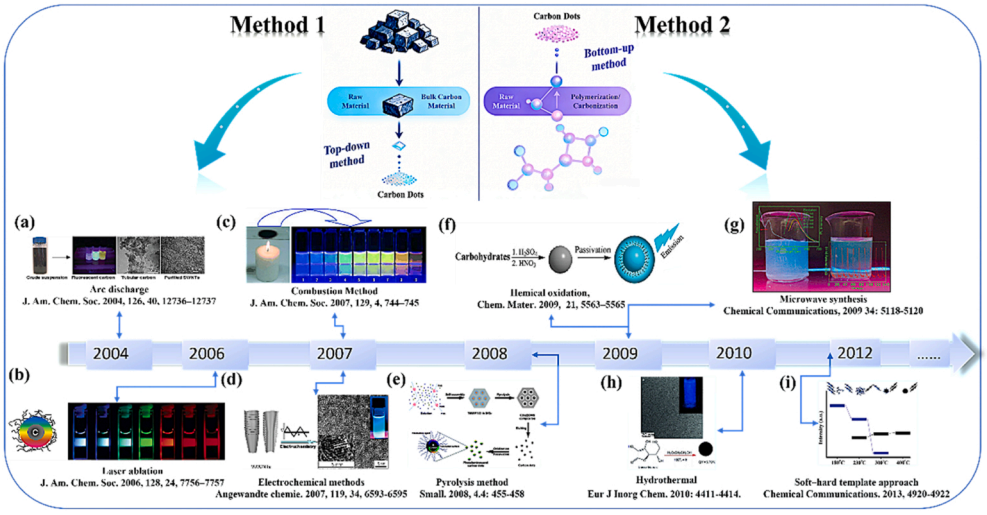


Fig. 1. Methods for CDs synthesis and timing. (a) Arc discharge methods [1]. Copyright 2004, with permission from American Chemical Society. (b) Laser ablation methods [2]. Copyright 2006, with permission from American Chemical Society. (c) Combustion methods [3]. Copyright 2007, with permission from John Wiley and Sons Ltd. (d) Electrochemical methods [4]. Copyright 2007, with permission from American Chemical Society. (e) Pyrolytic methods [5]. Copyright 2008, with permission from John Wiley and Sons Ltd. (f) Chemical oxidation methods [6]. Copyright 2009, with permission from American Chemical Society. (g) Microwave assisted heating [7]. Copyright 2009, with permission from the Royal Society of Chemistry. (h) Hydrothermal or solvothermal methods [8]. Copyright 2010, with permission from John Wiley and Sons Ltd. (i) Soft and hard templating methods [9]. Copyright 2013, with permission from the Royal Society of Chemistry.

Table 1
Classification, definition, and key characteristics of CDs.

Abbreviation	Full name	Definition	Structural features
CDs	Carbon dots	A broad umbrella term for zero-dimensional fluorescent carbon-based nanomaterials, typically with at least one dimension below 10 nm.	Generally composed of a carbonaceous core plus abundant surface functional groups, defects, and/or heteroatom dopants; may be crystalline, semicrystalline, or amorphous.
GQDs	Graphene quantum dots	A subclass of CDs with graphene-like crystalline domains, usually consisting of one or a few graphene layers.	Typically anisotropic, sheet-like or quasi-sheet-like nanostructures with highly conjugated sp^2 carbon frameworks; optical behavior strongly affected by edge states and π-domains.
CNDs	Carbon nanodots	A broad class of quasi-spherical carbon nanoparticles, often used as a general term for spherical CDs with amorphous or weakly crystalline cores.	Mostly spherical or quasi-spherical; typically contain carbon cores with lower graphitization than GQDs and rich surface functional groups.
CQDs	Carbon quantum dots	Generally refers to quasi-spherical carbon dots with more pronounced quantum-confinement-related behavior and a relatively higher degree of crystallinity than CNDs.	Spherical or quasi-spherical nanostructures with partially graphitic or crystalline domains; lattice fringes are often observed, indicating a more ordered carbon core.
CPDs	Carbonized polymer dots	Carbon dots derived from polymerization, crosslinking, and partial carbonization processes, typically retaining polymeric/organic characteristics in addition to carbonized domains.	Usually composed of disordered or weakly ordered carbon cores embedded in or connected with polymer-like/crosslinked subdomains and abundant surface groups.

nature of their synthesis processes renders traditional experience-based trial-and-error approaches inefficient, often making the discovery of high-performance CDs a matter of chance. In recent years, machine learning has emerged as an increasingly important data-driven tool in materials science, utilized for predicting material properties, screening novel compounds, and optimizing synthesis pathways [25–27]. Within this context, data-driven machine learning methods, leveraging their capacity to handle high-dimensional and nonlinear relationships, offer a promising opportunity to decipher the intricate correlations governing the structure-property relationships of CDs.

In the past period, substantial progress has been made in ML-assisted CDs research, with systematic contributions from multiple scholars laying a solid foundation for the advancement of this field [28], as shown in Table 2. Dananjaya et al. provided a comprehensive overview of the latest progress in the synthesis, functionalization, property modulation, and diverse applications of GQDs in polymer nanocomposites, including 3D printing technologies [49]. Their work also explored the emerging role of ML in optimizing GQD-based polymer composites. EI-Azazy et al. reviewed synthesis methods for CDs, ML-assisted optimization strategies, and their applications in water pollution control and electrocatalytic hydrogen evolution reactions [50]. More recently, Yan et al. systematically summarized the current state, model architectures, and technical pathways of ML in optimizing CDs synthesis and their application across various analytical detection fields [51]. Given the continuous advancement of ML-assisted research on carbon dots, it is important to classify the machine learning methods used in CDs research based on their functional roles and scientific contributions, as shown in Table 3. Additionally, systematically sorting out the application logic, core achievements and technical bottlenecks of ML in the field of carbon dots is of great significance for promoting sustainable innovation in this area. It highlights the core value of ML in the integrated and multi-domain application of CDs, offering a broader and more balanced perspective.

This review systematically outlines the commonly used ML algorithms in CDs research. By integrating recent relevant literature, it summarizes the general workflow and core paradigms for ML-assisted CDs studies, as illustrated in Fig. 2. Beginning at the fundamental research level, the review elaborates in detail on the key applications and technological breakthroughs of ML in the quantitative prediction of critical CDs properties, the optimization of synthesis processes, and the in-depth exploration of underlying mechanisms. A systematic evaluation of both the contributions and limitations of these approaches is provided. Subsequently, the focus shifts to the applied research level, where an in-depth analysis is presented on how ML has propelled the application of CDs in cutting-edge fields such as sensing and detection, photocatalysis and energy applications, biomedicine, light-emitting diodes, and information encryption and anti-counterfeiting. Typical case studies are incorporated to illustrate these advancements. Finally, the

Table 2
A summary table of common algorithms and applications in carbon dots application machine learning research.

Number	Prediction	Application	Algorithm	Data	Date	Ref
1	Contaminant	Sensor array	SVM, CNN, LSTM, CNN-LSTM, CNN-LSTM + Attention	1024	2025	[29]
2	Emerging Organic Pollutants (EOPs)	Sensor array	CNN, FCNN, RNN, DBN, TR, GRU, DNDF, CMAQ, SVR, GA-KELM, DBN-BP	137566	2025	[30]
3	NO_2^-	Sensor array	GPR	140	2026	[31]
4	Ion category and concentration	Sensor array	ANN	>240000	2025	[32]
5	Degrees of fermentation	Sensor array	ELM, SVM, LSSVM	210	2024	[33]
6	Biothiols	Sensor array	LDA, HCA, ANN, KNN, DT, SVM		2025	[34]
7	Heavy metal ions	Sensor array	NB, SVM, KNN, LDA, DT, ANN	90	2026	[35]
8	Quantum yield (QY)		DT, RF, GBDT, KNN, ET, XGBoost	480	2023	[36]
9	Methimazole monitoring	Sensor array	U-Net		2025	[37]
10	Temperature data		RNN		2025	[38]
11	Propazine concentration	Sensor array	SVR	400	2024	[39]
12	Amoxicillin	Sensors	PCA, LDA, Lasso, SVM	200	2025	[40]
13	Thiols	Sensor array	LDA, HCA		2022	[41]
14	Assess the risk of cardiovascular disease	Sensor array	Ridge-R, LSTM		2024	[42]
15	The AUC of the absorption curve in the 300–400 nm range.	Material discovery	ANN, BO		2025	[43]
16	Category and concentration	Sensor arrays	NDA, RF, SVM, EN	525	2025	[44]
17	Amino acids	Sensor arrays	ANN, NON-NEGATIVE-PARAFAC	600000	2021	[45]
18	Glucose	electrochemical sensing	LR, RF, LightGBM, ANN		2024	[46]
19	Cr (VI)	Sensor array	K-means, RR, XGB, SVR, Linear		2023	[47]
20	Differentiating nucleobases	Sensor array	PCA, HCA, KNN, SIMCA	35	2018	[48]

Table 3
Taxonomy of ML approaches in CDs research according to functional role and scientific contribution.

Category	Typical methods	Role of ML	Typical tasks in CDs research	Scientific contribution
Data reduction and pattern recognition	PCA, HCA, LDA, clustering	Extract patterns, classify responses, visualize data	Sensor-array classification, sample grouping, preliminary mechanism exploration	Improves data organization and discrimination, but usually does not provide design rules
Predictive modeling	MLR, SVR, SVM, RF, XGBoost, ANN	Learn input–output relationships	Predict emission wavelength, QY, phosphorescence lifetime, sensing signals	Accelerates screening and prediction, but often remains correlation-based
Optimization and decision-making	BO, GA, multi-objective optimization, active learning	Recommend conditions and improve experimental efficiency	Synthesis optimization, candidate selection, property maximization	Reduces trial-and-error and supports efficient exploration of design space
Mechanism-coupled and design-oriented ML	ML + DFT, interpretable ML, physically informed ML, closed-loop learning	Link descriptors to mechanism and support rational design	Mechanism analysis, precursor-informed design, transferable rule discovery	Most relevant to causality-aware rational design

review concentrates on the core bottlenecks faced by this interdisciplinary field, discusses potential solutions, and outlines promising future research directions. The overarching objective of this review is to provide a systematic, intelligent, and scalable data-driven methodology for CDs research, thereby fostering a fundamental transition in material development from an experience-oriented paradigm to an artificial intelligence-enhanced one. The overall framework of the review is presented in Scheme 1.

2. Application of ML to CDs

2.1. Prediction of CDs properties

Current research on the prediction of CDs properties has significantly enhanced the efficiency of experimental screening and has gradually shifted the field from traditional trial-and-error approaches to data-assisted decision-making. However, from a methodological perspective, most studies still mainly focus on performance prediction and parameter optimization, rather than truly entering the stage of mechanism-guided rational design. This feature is particularly evident in the three core properties of emission wavelength and color, fluorescence quantum yield, and afterglow characteristics of CDs. This is because the input variables in the models are empirical ones such as precursor type, reaction temperature, time or ratio, which mainly serve to compress the experimental search space and improve screening efficiency, but often still struggle to establish transferable structure-activity relationships, thus remaining a considerable distance from target-oriented reverse design. Therefore, to truly achieve mechanism-guided rational design, it is necessary to further strengthen the accumulation of high-quality data, enhance the physicochemical significance of descriptors, strengthen external validation and cross-system generalization, and deeply integrate machine learning with spectral characterization, theoretical calculation and electronic structure analysis, thereby establishing structure-property-function relationships that are both interpretable and transferable.

2.1.1. Emission wavelength and color

The photoluminescence (PL) mechanisms and origins of CDs remain incompletely understood, which limits the effective prediction of their emission colors. Currently, most syntheses yield CDs that primarily emit blue and green fluorescence, while achieving efficient and stable red and near-infrared (NIR) emission remains challenging [65]. The ability to precisely tune the emission wavelength and color of CDs is critical for expanding their applications in multicolor imaging [66], enhancing deep-tissue penetration [67,68], and reducing biological background interference [69]. Early strategies relied largely on adjusting macroscopic synthesis parameters to control luminescence, a process that remained essentially empirical and trial-and-error-based, resulting in low efficiency and poor controllability. In recent years, the research on the color regulation of CDs using machine learning can be roughly divided into three levels: (i) classification/regression of colors or wavelengths within the existing sample space; (ii) statistical mapping

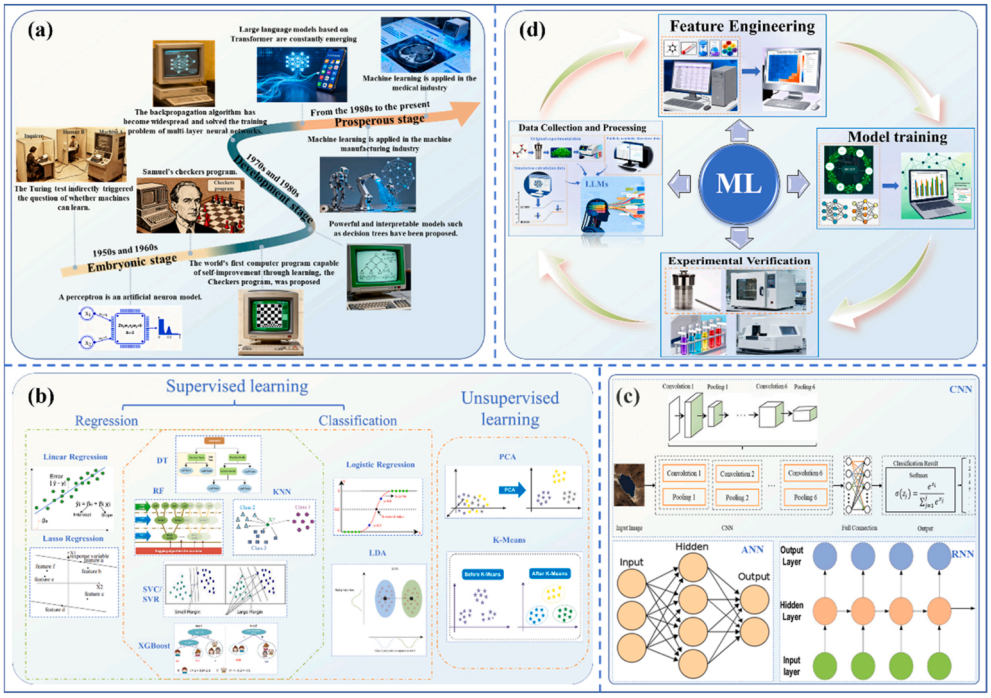
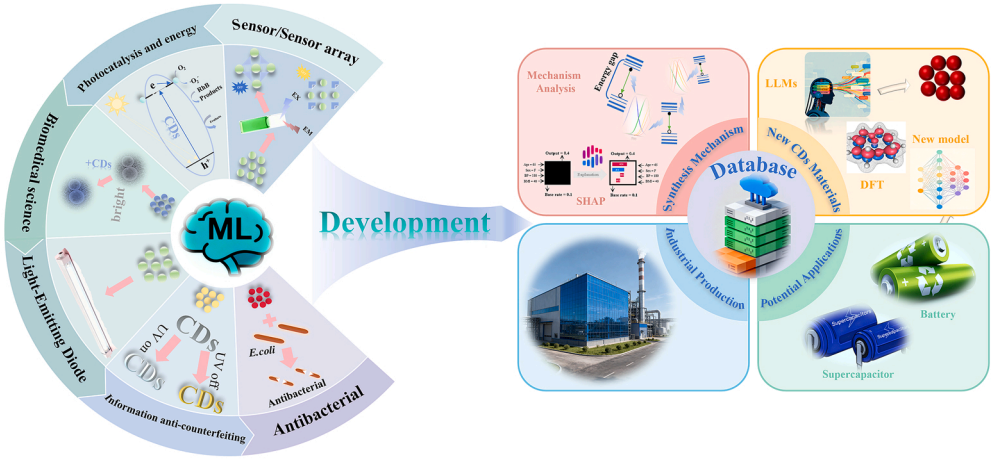


Fig. 2. (a) The development stage of artificial intelligence. (b-c) Commonly used algorithms of ML [52–61] and DL [62–64]. The above pictures have all been obtained with copyright permission. (d) The workflow of ML.



Scheme 1. Overview of ML-assisted CDs research, highlighting ML-driven CDs applications, and future research directions.

between synthetic variables and optical output; (iii) models based on precursors or structures, which are closer to true design.

For instance, Luo et al. constructed a classification model based on literature data that achieved high-accuracy identification of red versus non-red CDs, as shown in Fig. 3(a) [70]. Senanayak et al. further integrated literature and experimental data to establish a statistical correlation model between synthesis parameters and emission wavelength, illustrated in Fig. 3(b) [71]. It should be noted, however, that such studies essentially address the problem of color or wavelength identification under existing sample conditions, rather than the inverse design of precursors and synthesis conditions for a targeted emission wavelength. Especially in the red and near-infrared CDs systems, due to the significantly insufficient sample quantity and highly uneven distribution, the red CDs account for a minority in the training set. The model has difficulty capturing their key features, and the high accuracy of the model more reflects its fitting ability for blue-green emitting CDs. Its cross-system generalization and reverse design capabilities are still limited by the structural data bottleneck. This is because the model cannot directly represent the core microscopic structural factors that determine the performance of the CDs. It can only rely on the local statistical correlation between process parameters and performance, and therefore it is difficult to learn transferable structure-property rules. Tuchin et al. specifically addressed the core challenge of predicting the optical properties of red and NIR-emitting CDs, as depicted in Fig. 3(c) [72]. By integrating literature data to build ML models and employing both human-driven and machine-driven methods for data clustering, they constructed primarily linear-regression-based prediction models. While these studies have achieved accurate prediction of absorption and fluorescence peak positions, significant challenges persist in targeted design. Xing et al. proposed an ML prediction strategy that treats precursor combinations as variables [73]. By constructing a dataset of multiple precursor combinations and optimizing model parameters, they realized precise prediction of the PL performance of CDs under different excitation wavelengths, taking a step forward towards

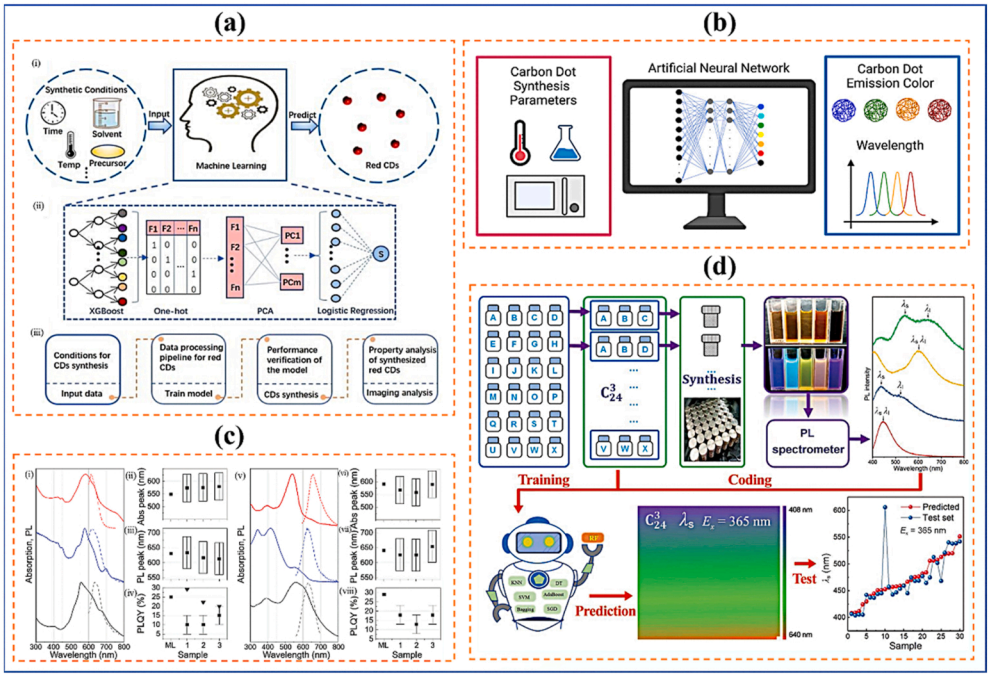


Fig. 3. (a) Semantic illustration of machine learning was introduced to assist researchers in synthesizing red fluorescent CDs [70]. Copyright 2022, with permission from the Royal Society of Chemistry. (b) Schematic illustration of machine learning for predicting the wavelength and color of carbon dots [71]. Copyright 2022, with permission from American Chemical Society. (c) Comparison of optical properties of CDs-1 [72]. Copyright 2024, with permission from John Wiley and Sons Ltd. (d) Schematic illustration of the CDs synthesis and machine learning process [73]. Copyright 2024, with permission from Springer Nature.

design-oriented modeling, as presented in Fig. 3(d). This strategy mainly operates at the level of precursor combination screening rather than being completely based on mechanism-based reverse design.

The optical properties of CDs serve as the critical bridge connecting synthesis conditions to their ultimate applications and form the central focus for prediction and optimization using ML. Although current ML methods have demonstrated progress in predicting CDs color and emission wavelength, they still face significant challenges including inadequate dataset quality, uneven sample distribution, and insufficient model generalizability. Furthermore, due to the lack of descriptors related to the mechanism, the above research did not deeply interpret the structure-activity relationship of carbon dots. The later model input should not be limited to process parameters alone; it should start from physical and chemical descriptors. With continued advances in data accumulation and algorithm development, the integration of ML with quantum chemical calculations and theoretical simulations at a deep level, and revealing the fundamental nature of luminescence from the electronic structure perspective, will emerge as a vital pathway toward achieving truly rational design of CDs.

2.1.2. *Fluorescent quantum yield*

Beyond absorption and emission wavelengths, photoluminescence QY is another core metric for evaluating the luminescent performance of CDs. It represents the efficiency of energy conversion in fluorescent materials and directly determines their sensitivity, brightness, and effectiveness in applications [74]. However, the solid-state photoluminescence QY of experimentally synthesized fluorescent CDs remains generally low, with poor controllability. Achieving precise regulation of QY still requires extensive experimental screening, which severely hinders the practical application of CDs materials. In this context, the introduction of ML offers precise guidance for the directional preparation of CDs with high QY, greatly accelerating the development process, reducing trial-and-error costs, and enabling a more rational design approach.

For instance, Han et al. applied ML to CDs synthesis, identifying key factors affecting the QY of CDs and successfully predicting the QY of green-emitting CDs, as illustrated in Fig. 4(a) [75]. Nevertheless, this work focused solely on single-objective optimization and was limited to only five process parameters as features. The model was forced to establish correlations between process parameters and performance to fit the training data, which may result in good performance on the training set but poor generalizability. When applied to new chemical systems or slightly different process conditions, predictions can deviate significantly. Guo et al. proposed an ML-guided multi-objective optimization framework that expanded the feature set to eight parameters and achieved precise synthesis of full-color CDs, as demonstrated in Fig. 4(b) [76]. Unfortunately, the input variables of the above research mainly consist of empirical process parameters, while the output is determined by the complex microstructure. The model is forced to fit the results using surface correlation, which may lead to good fitting results within the given system. However, when applied to new systems or different reaction paths, its predictive ability may still significantly decline, indicating that it is more of a local condition optimization tool rather than reaching the level of rational design driven by mechanisms.

To address the gap in elucidating precursor-related physical mechanisms, Chen et al. combined ML-based principal component analysis with group theory, encoding structural features such as precursor symmetry as inputs [77]. Their work revealed that molecular vibrational modes serve as a core mechanism influencing the QY of GQDs, as outlined in Fig. 4(c). Dolenkoa et al. focused on a specific precursor system and employed a multilayer perceptron (MLP) neural network to establish a correlation model, enabling precise regulation of the target QY across a range of 0-100%, as represented in Fig. 4(d) [78]. It is worth noting that this work is limited to a single precursor combination. The model input mainly consists of process parameters. Once the precursor is changed, the nucleation mode of CDs, the formation of local conjugated domains, the construction of surface states, and the way of doping with heteroatoms may all change, resulting in insufficient generalization ability of the model. Future research should expand the dataset to include a wide range of precursor chemicals, and the input features should also incorporate chemical meaning descriptors so that the model can learn cross-system patterns. More importantly, the input features should incorporate chemical meaning descriptors. Additionally, advanced modeling strategies, such as transfer learning and physics-based machine learning, can be used to enhance the model's robustness across different systems. Kannouma et al. further optimized the model architecture. Through the combination of MLP and genetic algorithm optimization, they first predicted the optimal precursor

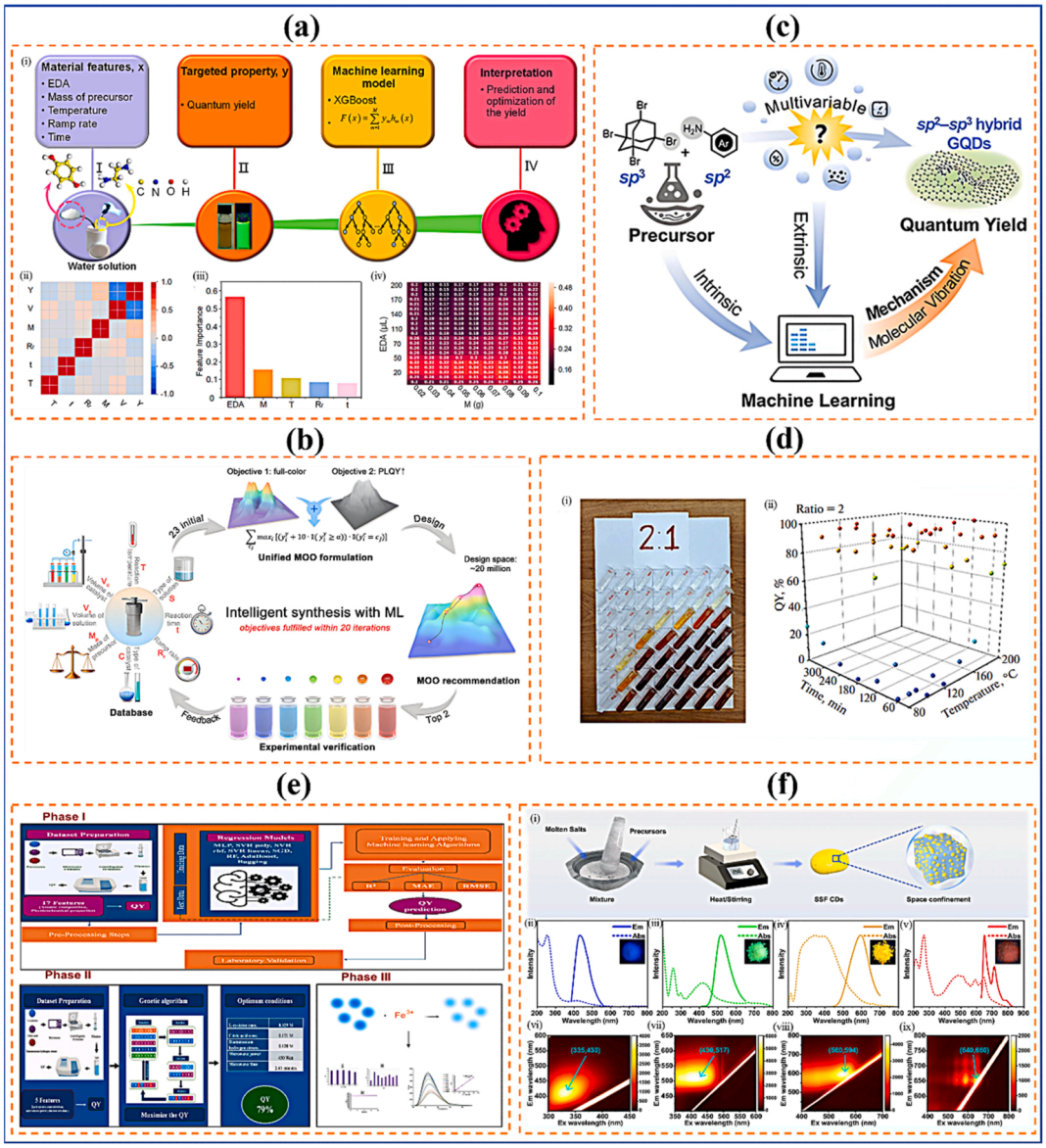


Fig. 4. (a) Application of ML for guided synthesis of CDs [75]. Copyright 2020, with permission from American Chemical Society. (b) Workflow of ML-guided synthesis of carbon quantum dots with superior optical properties [76]. Copyright 2024, with permission from Springer Nature. (c) Fluorescence quantum yield modulation in graphene quantum dots with the assistance of machine learning for symmetrically triggered precursors [77]. Copyright 2024, with permission from John Wiley and Sons Ltd. (d) Photograph of synthesized CDs samples and experimental QY values of the same dots [78]. Copyright 2025, with permission from Springer Nature. (e) Schematic diagram of the synthesis of high QY-CQDs using machine learning assistance and its application process [79]. Copyright 2025, with permission from Elsevier. (f) Synthesis and optical properties of full-color SSF CDs assisted by low-temperature molten salt [80]. Copyright 2025, with permission from Springer Nature.

combination based on 17 physicochemical features and subsequently optimized the synthesis parameters, as presented in Fig. 4(e) [79]. These investigations, however, remained at the laboratory-scale synthesis level and did not address the challenges of large-scale production or low solid-state luminescence efficiency.

Conventional synthesis methods often require high reaction temperatures, which can lead to excessive polymerization and carbonization of precursors, thereby reducing the solid-state fluorescence (SSF) QY. Moreover, CDs are prone to aggregation-induced fluorescence quenching in the solid state. The in-situ spatial confinement effect offered by low-temperature molten salts can suppress non-radiative transition pathways and significantly enhance solid-state QY. Building on this principle, Yu et al. developed a low-temperature molten-salt synthesis method. By utilizing an XGBoost model to optimize zinc-ion coordination and spatial confinement effects, they achieved kilogram-scale production of solid-state luminescent CDs with a remarkable QY of up to 99.86% and successfully applied them in Light-emitting diodes (LEDs), as demonstrated in Fig. 4(f) [80].

Current ML-based research on CDs QY has achieved breakthroughs in small sample prediction and multi-objective optimization. However, it is still constrained by issues such as limited dataset size and lack of external validation, risks of overfitting in the prediction and verification stages, and relatively traditional model architecture and insufficient industrialization adaptability. These limitations have restricted the future development of CDs. Future research still needs to break through in three core directions: data standardization, model interpretability, and application orientation. It should focus on the green industrialization system to promote the upgrade of machine learning from predicting the properties of CDs to becoming a core tool for guiding the design and industrial production of CDs materials.

2.1.3. CDs afterglow properties

The luminescence lifetime of fluorescent CDs is crucial for optimizing their instant optical performance and for developing fluorescent probes and sensors. However, as fields such as optical encryption and bioimaging continue to advance, reliance solely on instantaneous luminescence is no longer sufficient, creating an urgent need for new materials with controllable delayed-emission capabilities. This demand has shifted research focus toward room-temperature phosphorescence (RTP) materials [81]. Conventional RTP materials, typically composed of rare-earth elements, noble-metal complexes, or pure organic compounds, often involve complex processes, high costs, relatively short phosphorescence lifetimes, and potential biotoxicity [82]. Consequently, phosphorescent CDs have emerged as promising alternatives due to their facile preparation, low toxicity, and tunable luminescence. Nevertheless, phosphorescence properties are determined by a multitude of complex factors, making the discovery of high-performance materials highly serendipitous. Therefore, establishing predictive methods for phosphorescence lifetime is of significant value for enabling rational material design and accelerating application translation.

For example, Li et al. combined ML-based principal component analysis with group theory to reveal that precursor symmetry matching is critical for regulating the degree of sp^2-domain separation in nitrogen-doped GQDs [83]. This approach successfully led to the preparation of an ultralong-lifetime phosphorescent material with a lifetime of up to 28.5 s, as shown in Fig. 5(a). Furthermore, Yang et al. integrated active machine learning with density functional theory (DFT) and time-dependent DFT to accurately predict the excitation wavelength, emission wavelength, and afterglow lifetime of CDs, as illustrated in Fig. 5(b) [84]. However, these studies focused exclusively on solid-state systems, which limits the functional versatility of the materials and hinders their adaptation to the diverse requirements of complex application scenarios.

Although phosphorescent CDs can achieve efficient emission in the solid state through matrix-assisted stabilization, they still face the key challenge of triplet exciton quenching in liquid systems [86]. Guo et al. developed an interpretable ML model based on XGBoost, combined with SHAP analysis and quantum tunneling theory [85]. They clarified an inverse correlation between matrix thickness and the probability of triplet-electron leakage, and successfully prepared highly efficient phosphorescent CNDs in aqueous solution with a lifetime exceeding 10 s and a quantum yield of 10.18%, thus addressing the critical problem of phosphorescence quenching in liquid environments, as demonstrated in Fig. 5(c). While these studies have successfully used ML to predict the phosphorescence properties of CDs, the models rely heavily on highly controlled specific synthesis systems, and the conclusions drawn are primarily based on statistical correlations. As a result, the generalizability of these models and their capability for inverse material design remain limited.

Overall, these studies indicate that in the current research on phosphorescent carbon dots, the application of machine learning is evolving from empirical screening to mechanism-based optimization, but it has not yet reached the stage of fully standardized rational design. The performance of such models depends critically on the quantity and coverage of experimental data. Currently available datasets are still limited in size and often originate from single laboratories, which may introduce systematic bias. Furthermore, long-lifetime and high-quantum-yield samples are underrepresented in existing datasets, leading to insufficient predictive capability for high-performance CDs. Future efforts could benefit from the introduction of transfer learning and meta-learning techniques to reduce model dependence on specific systems while addressing the scarcity of high-performance samples through data augmentation. Additionally, emphasis should be placed on the transferability of models across different material systems, enabling their application to other types of carbon-based lumines-cent materials and facilitating the design of next-generation CDs with improved thermal stability and longer lifetimes.

2.2. ML-assisted CDs synthesis

The final properties of CDs are not determined by a single factor but result from the complex interplay of multiple variables. Traditional CDs synthesis relies on one-factor-at-a-time experimentation, requiring

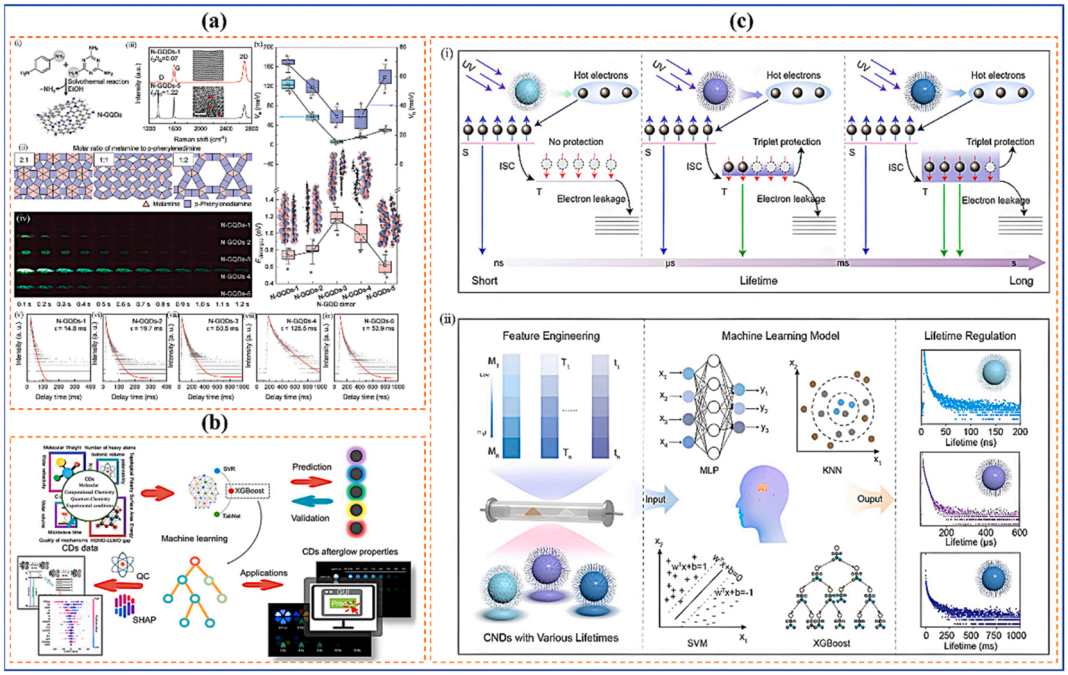


Fig. 5. (a) Modulation of the separability of sp^2 domains and phosphorescence lifetime of GQDs by changing the molar ratio of precursors [83]. Copyright 2024, with permission from John Wiley and Sons Ltd. (b) Schematic representation of carbon dots-based long afterglow material exploration and design utilizing active machine learning and quantum chemical simulations [84]. Copyright 2024, with permission from American Chemical Society. (c) The construction of the ML model based on the triplet excitons' behavior [85]. Copyright 2025, with permission from John Wiley and Sons Ltd.

extensive repetitive trials; optimizing the reaction conditions to obtain CDs with specific properties can thus become a time-consuming trial-and-error process [22]. ML offers the ability to uncover nonlinear synergistic or antagonistic interactions among synthesis parameters and to identify key influencing factors, thereby optimizing synthesis conditions and providing a theoretical foundation for the targeted design of CDs with tailored characteristics.

For instance, Pudza et al. employed response surface methodology and an artificial neural network (ANN) to optimize the synthesis of fluorescent CDs from cassava starch, as shown in Fig. 6(a) [87]. However, the entire model was trained, validated, and tested on only 30 data points, and the training data often failed to adequately cover the actual data distribution. The distribution of high-performance samples was uneven, resulting in the model being able to only learn local correlations and unable to be applied to cross-system patterns. On the other hand, the model only established a mathematical mapping between process parameters and performance, without exploring or revealing any underlying chemical or physical mechanisms, limiting the scientific value and extrapolation ability of the model. To address such limitations, Zhang et al. focused on optimizing the fluorescence activity of boron and nitrogen-co-doped GQDs [88]. They used a polynomial regression model followed by a greedy random-walk algorithm to globally search the parameter space and identify locally optimal parameters, as illustrated in Fig. 6(b). This approach, however, did not achieve a full global search of the entire synthesis parameter space; the optimization results remained confined to local optima, and extensive repetitive experiments were still required to verify parameter effects. Furthermore, without optimization, only the influence trends of individual parameters could be identified, while the interactions among multiple parameters remained unclear. To bridge the gap in global optimization, He et al. adopted random-forest model to accurately predict the inhibition efficiency of CDs-based corrosion inhibitors and integrated a genetic algorithm to intelligently optimize the synthesis pathway, as depicted in Fig. 6(c) [89]. This work represents a step beyond traditional trial-and-error methods and promotes the application of ML in material synthesis, particularly in the development of green functional materials.

The current limitations of machine learning research that is oriented towards synthesis are not only due to insufficient data size, but also because the models often learn the correlation between the process and performance, yet fail to clearly capture the chemical pathways of precursor transformation, carbonization, and structure formation. Future efforts should focus on constructing multi-source heterogeneous datasets, introducing transfer learning to enhance model generalization, establishing multi-objective optimization frameworks that integrate performance with cost and environmental considerations, and incorporating physicochemical mechanisms to strengthen model interpretability. These advancements will enable more accurate guidance for targeted design and optimization. Additionally, ML should be fully leveraged to develop CDs synthesis methods that are better suited for industrialization, thereby realizing scalable, standardized, and intelligent production of CDs materials.

2.3. Analysis of ML assistance mechanisms

The photoluminescence mechanisms of CDs are highly complex and not yet fully understood. The intricate formation processes complicate the regulation of CDs structure and properties as well as the establishment of correlations between synthesis parameters and CDs characteristics. Traditional analytical methods, often limited to fixed excitation wavelengths, cannot comprehensively extract spectral information under diverse conditions. In this context, ML provides a valuable tool for guiding the investigation of CDs luminescence mechanisms. For example, Dager et al. synthesized monodisperse carbon quantum dots from fennel seeds and employed ML methods to analyze their PL mechanism, recommending optimal excitation wavelengths for PL analysis as illustrated in Fig. 7(a) [90]. However, the application of ML in this study was confined to data classification and feature extraction rather than serving as a mechanism-interpretation tool; the interpretation of fluorescence origins still relied on conventional characterization techniques.

The detection mechanisms of CDs as sensors can be classified into three categories: fluorescence quenching, fluorescence enhancement,

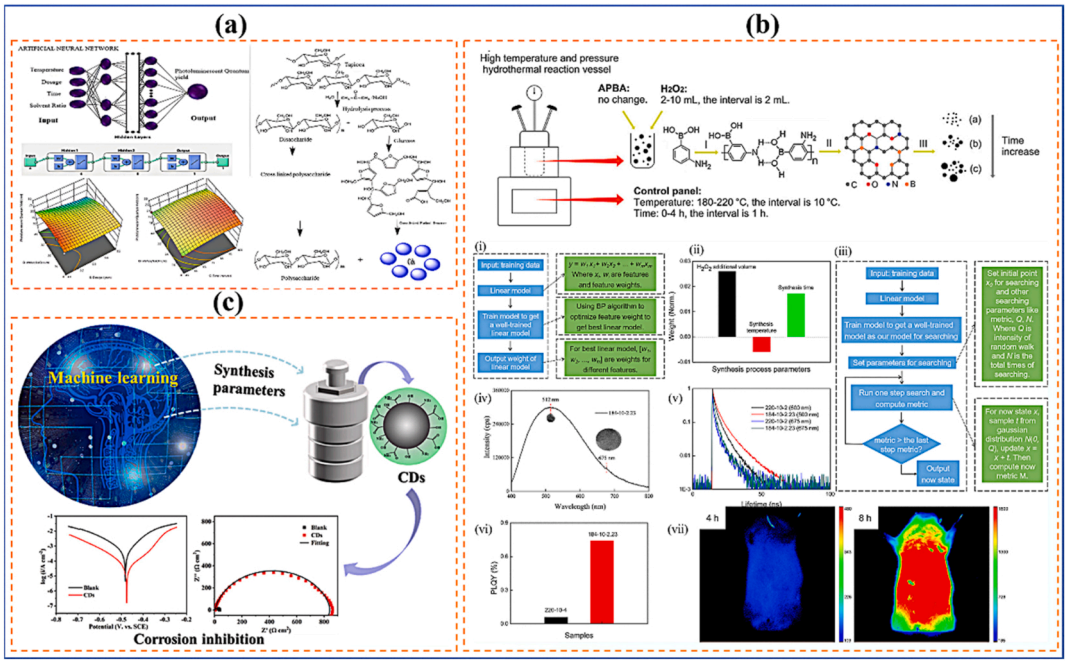


Fig. 6. (a) Schematic diagram of optimizing the process of synthesizing carbon dots from cassava flour using machine learning [87]. Copyright 2019, with permission from Multidisciplinary Digital Publishing Institute. (b) Schematic diagram of the synthesis process for optimizing B, N-GQDs using machine learning technology [88]. Copyright 2022, with permission from American Chemical Society. (c) Schematic diagram of the synthesis of carbon dots-based corrosion inhibitors controlled by machine learning [89]. Copyright 2023, with permission from Elsevier.

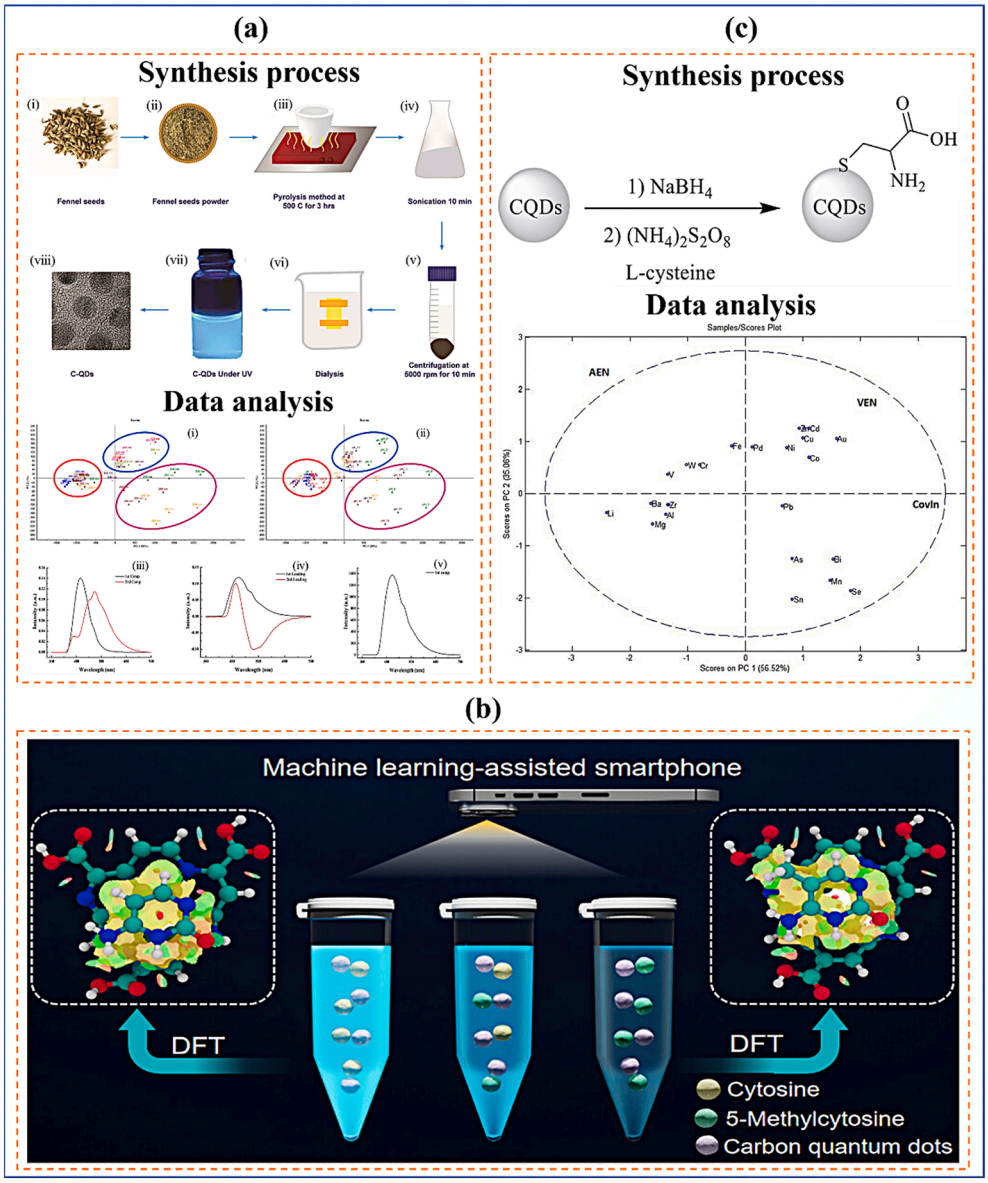


Fig. 7. (a) The schematic illustration shows the synthesis of CDs from fennel seed via pyrolysis method and analysis results of three models [90]. Copyright 2019, with permission from Springer Nature. (b) A schematic diagram of using ML to assist smartphone in interpreting dual-response CDs for simultaneous detection of cytosine and 5-methylcytosine [91]. Copyright 2023, with permission from American Chemical Society. (c) Thiol-ene conjugation of L-cysteine to the CQDs and the biplot of the first two PCs of PCA analysis [92]. Copyright 2021, with permission from Elsevier.

and fluorescence shift. Thonghlueng et al. constructed a dual-response fluorescent sensor capable of simultaneously detecting cytosine and 5-methylcytosine in urine. DFT was employed to elucidate the underlying sensing mechanism, and further integration of ML with a smartphone platform enabled accurate quantification of methylation levels, as shown in Fig. 7(b) [91]. Obviously, since the machine learning model is trained based on macroscopic image features rather than physical descriptors, it can only learn the empirical correlations for signal quantification, but cannot reveal the underlying electronic or chemical mechanisms that control the fluorescence behavior. Additionally, the methylation levels in both the training and testing data were not comprehensively covered, lacking intermediate gradients or extreme values, which may limit the prediction accuracy of the model for methylation levels outside the sampled range. Most importantly, this study offers a valuable perspective for future work: coupling DFT with ML could enable rapid identification of interaction mechanisms and verification of detection selectivity through a one-time DFT data accumulation followed by multiple model reuses. Salahinejad et al. applied ML to investigate the fluorescence-quenching mechanisms of CDs as fluorescent probes in heavy-metal detection [92]. By functionalizing CQDs with cysteine and combining quantitative structure-property relationship (QSPR) modeling with ML methods, intrinsic correlations between the physicochemical properties of heavy-metal ions and quenching efficiency were revealed. The ability of metal ions to accept electrons and their characteristics in the context of hard-soft acid-base theory were identified as key influencing factors, as depicted in Fig. 7(c). Nevertheless, although cross-validation was adopted during model training, an independent external validation set was not explicitly defined, and the actual predictive performance of the model requires further verification. Moreover, the physical significance of some descriptors needs to be further validated, as certain derived descriptors exhibit weak chemical interpretability.

Mechanistic prediction for CDs remains a challenging task within ML algorithms, especially in mastering precursor structural characteristics and improving prediction accuracy. Future efforts could benefit from coupling with DFT calculations to achieve a deeper, microscopic-level interpretation. Furthermore, ML could also be leveraged to explore mechanisms such as precursor carbonization and assembly, heteroatom doping and defect regulation, specific recognition between CDs and biomolecules, and drug loading and controlled-release mechanisms.

3. ML drives the development of CDs applications

Carbon dots have demonstrated significant application potential across various fields due to their excellent biocompatibility, low toxicity, and unique optical properties. However, their functional performance is highly dependent on multiple factors, including synthesis routes, structural characteristics, and environmental responses [93]. The traditional experience-driven research paradigm has struggled to systematically uncover the complex structure–property relationships, limiting the rational design and performance optimization of CDs materials. In this context, machine learning, with its powerful data-mining and pattern-recognition capabilities, offers a new paradigm for deciphering the intricate correlations between multi-dimensional features of CDs and their target performance. By constructing cross-domain databases that integrate synthesis parameters, physicochemical properties, and functional outputs, ML models can achieve accurate prediction and inverse guidance from material design to application performance. This approach promotes the intelligent development of CDs in areas such as sensing [94–96], photocatalysis and energy applications, biomedicine [97–99], light-emitting diodes, and anti-counterfeiting encryption [82, 100,101], thereby accelerating their advancement toward high performance and integrated functionality. Accordingly, this review will introduce the achievements and limitations of ML-driven CD development in the following aspects.

3.1. Sensing

Carbon dots possess surfaces rich in amino, carboxyl, hydroxyl, and other functional groups, which facilitate interactions with small molecules and drug molecules. They also exhibit high sensitivity to their surrounding environment, often accompanied by corresponding changes in fluorescence intensity or emission wavelength. Consequently, CDs are widely employed in constructing highly sensitive and selective fluorescent sensing platforms [102]. Fluorescent sensing based on CDs primarily achieves qualitative or quantitative detection by monitoring changes in fluorescence signals in the presence of target analytes. This approach offers strong flexibility and is applicable to a wide range of detection scenarios and analytes [103]. To date, CDs have achieved a series of advances in ion detection [104], the detailed information is listed in Table 4, temperature sensing [105], bacterial detection [106], the detailed information is listed in Table 5, pH monitoring [107], and drug detection, among others. However, as detection systems become increasingly complex and the variety of targets continues to expand, CDs-based sensing faces dual challenges. On one hand, the synthesis process of tailoring CDs with high selectivity for specific targets through trial-and-error methods remains highly uncertain. On the other hand, even when sensing materials are obtained, traditional data-processing methods struggle to rapidly and automatically interpret the resulting response signals. Within this dual context, the introduction of ML techniques, operating at either the material-design or signal-interpretation level, has become an inevitable trend for enhancing the intelligence and practicality of sensing systems [114].

Traditional ion-detection methods are often operationally complex and rely on large-scale instruments, hindering on-site application [115]. Several researchers have adopted a strategy combining CDs-based fluorescent sensing with ML algorithms to provide efficient, sensitive new solutions to this problem. For example, Tian et al. utilized the data-processing capability of ML to detect ions such as Cu^{2+} and Hg^{2+}, with a limit of detection (LOD) of 100 ppm [108]. To further meet practical requirements for rapid on-site detection, research has advanced toward portability and platform integration. Chen et al. developed near-infrared carbonized polymer dots, integrated them with an ML algorithm and 3D printing technology to construct a portable detection platform, and applied it to the ultrasensitive detection of trace Cu^{2+} in seawater and aquatic products, achieving an LOD as low as 0.24 nM [109]. For more complex gaseous pollutant detection scenarios, this technology has evolved further into integrated and functionalized sensing solutions. Qin et al. innovatively employed B, N-doped CDs combined with an AgCl/Ag composite fluorescent-film sensor, integrating a photo-oxidation-enrichment strategy with ML to realize the integrated detection of gaseous mercury (Hg^0) in coal-fired flue gas, with an LOD of 3.2×10^{-7} g m^{-3} [116]. This technological pathway, through material design, device integration, and algorithm optimization, systematically promotes the transition of ion detection from laboratory analysis toward field-deployable and intelligent applications.

Particularly for Fe^{3+} detection, existing fluorescent probes suffer from limitations such as weak anti-interference capability, relatively high LODs, and difficulty in achieving high-throughput analysis. Zheng et al. focused on optimizing the basic performance of Fe^{3+} fluorescent

Table 4
Summary of research on ML-assisted CDs ion detection.

Carbon dots	Carbon source	Synthesis method	Target analyte	Best model	R^2	LOD	Ref.
CQD	Citric acid, ethylenediamine	Hydrothermal	Cu^{2+}; Pb^{2+}; Fe^{3+}; Hg^{2+}; Cr^{6+}	KNN	0.9950	Cu^{2+}:100 p.m.; Pb^{2+}:10 nM; Fe^{3+}:100 pM; Hg^{2+}:10 nM; Cr^{6+}:10 nM	[101]
CPDs	o-phenylenediamine, 1,3,5-benzene tricarboxylic acid	One-step hydrothermal	Cu^{2+}	DT	0.9986	0.24 nM/0.47 nM	[102]
B,N-CDs	/	Hydrothermal	Hg^0	Linear-R	0.9823	3.2×10^{-7} g m^{-3}	[103]
NCDs	L-histidine, L-aspartic acid	One-step hydrothermal	Fe^{3+}	Linear-R	0.9586	0.716 mM	[104]
NN-uGQDs	Graphene foam	Electrochemical	Hg^{2+}; Fe^{3+}	Classical Linear-R	0.9920	Hg^{2+}: 0.001 mg L^{-1}; Fe^{3+}: 0.003 mg L^{-1}	[105]
CQDs	Cinnamyl alcohol	Solvothermal	Fe^{3+}	YOLOv8	0.9968	33 nM	[106]
Au@N-GQDs	/	/	Fe^{3+}	Stepwise Linear model	0.9800	1.0 nM	[107]
BU-CDs	Citric acid, cysteamine hydrochloride	Hydrothermal	Hg^{2+}; S^{2-}	Lasso-R	0.9910; 0.9939	Hg^{2+}: 0.834 μM; S^{2-}: 3.885 μM	[108]
T-CDs	β-Cyclodextrin, o-phenylenediamine	Hydrothermal	Cd^{2+}; Co^{2+}; Cu^{2+}; Fe^{3+}; Mn^{2+}; Zn^{2+}	XGBoost	0.9858 - 0.9977	Cd^{2+}: 1.63 μM; Co^{2+}: 3.14 μM; Cu^{2+}: 5.82 μM; Fe^{3+}: 6.52 μM; Mn^{2+}: 4.42 μM; Zn^{2+}: 6.78 μM	[109]

Table 5
Summary of research on ML-assisted CDs bacteria detection.

Carbon dots	Carbon source	Synthesis method	Target analyte	Best model	R^2	LOD	Ref.
CQDs	Citric acid, ethylenediamine	One-step hydrothermal	Bacterial species	RF	100% accuracy rate	10^2 CFU mL^{-1}	[110]
PM-CQDs; Amp-CQDs; Gen-CQDs;	Ammonium citrate	One-pot	Foodborne pathogens	Naïve Bayes	>0.9814	$<1 \times 10^3$ CFU mL^{-1}	[111]
NS-GQDs; S-NS-GQDs; K-NS-GQDs; B-NS-GQDs	/	/	Bacteria	KNN	>0.9803	O157:H7: 84.0 CFU mL^{-1}; *S. aureus*: 109.5 CFU mL^{-1}; S. typhimurium: 41.6 CFU mL^{-1}; *P. aeruginosa*: 48.9 CFU mL^{-1}; *L. monocytogenes*: 168.8 CFU mL^{-1}	[112]
Ag-CQDs; Cu-CQDs; Zn-CQDs	L-tryptophan and citric acid; 4-aminobenzene and ethylenediamine; o-phenylenediamine	Hydrothermal	Bacteria	KNN	0.9820	$<1 \times 10^3$ CFU mL^{-1}	[113]

probes through material design and simple ML models, but the LOD remained at the μM level, and the issues of interference and efficiency in multi-ion coexistence were not addressed [110]. To move toward practical application, Llaver et al. aimed to achieve simultaneous multi-target detection and cope with complex matrices [111]. By combining functionalized graphene quantum dots with an ML algorithm, they realized the simultaneous detection and quantification of Hg^{2+} and Fe^{3+} in real water samples, with an LOD for Fe^{3+} as low as 53.7 nM. However, this work was still confined to liquid-phase detection and lacked high-throughput analytical capability. Li et al. focused on constructing an integrated, high-throughput, and even solid-state intelligent detection platform [112]. They developed a fluorescent detection system combining hydrophobic deep eutectic solvents with carbon quantum dots, paired with the YOLOv8 ML algorithm to build a high-throughput point-of-care testing platform. This system achieved rapid and intelligent detection of trace Fe^{3+} in serum and drinking water, with an LOD as low as 33 nM. Das et al. constructed an Au@N-GQDs heterostructure and performed in-depth optimization using over 20 supervised learning models, the Stepwise Linear model was selected, with an R^2 value of 0.98 [113]. Since the sensing response is mainly dominated by a small number of variables with clear physical meanings, and the input feature dimension is low and the sample size is limited, the stepwise linear model can achieve excellent predictive performance while maintaining high interpretability through effective variable selection. Ultimately, in a solid-state sensing mode, they achieved highly sensitive detection of Fe^{3+} in river water with excellent anti-interference performance and an LOD of only 1 nM, marking a dual advance toward practicality and ultimate sensitivity in this field.

The studies described above demonstrate the feasibility of ML-assisted CDs sensing. However, most approaches remain single-channel, suffering from insufficient anti-interference capability and multi-target recognition ability in complex environments. The development of sensor arrays based on multicolor or multi-channel signals has become a key technological transition for advancing ML-assisted CDs sensing toward practical applications [117,118]. Research strategies have deepened along two complementary directions: one involves constructing integrated multicolor sensors for specific application scenarios. For instance, Lu et al. proposed an ML system based on a multicolor fluorescence-colorimetric paper sensor, enabling the visual detection of Hg^{2+}, S^{2-}, and the food spoilage marker H_2S, providing a direct solution for on-site food-quality monitoring, as shown in Fig. 8(a) [119]. The other direction focuses on developing multi-channel sensor arrays with general-purpose recognition capability to address broader complex detection needs. Tang et al. constructed a six-channel sensor array based on triple-emission CDs and combined it with ML to achieve simultaneous identification and quantitative detection of six metal ions, demonstrating high stability and anti-interference performance in real water samples [126]. This offers a data-driven innovative solution for rapid multi-metal-ion detection.

Beyond ion sensing, CDs can also precisely monitor temperature via changes in fluorescence intensity. Compared with traditional contact-type sensors, non-contact sensing based on fluorescence-intensity changes offers advantages such as high spatial resolution and rapid response, making it particularly suitable for microscopic or non-invasive temperature measurement in scenarios such as living organisms and microfluidic chips. However, broad emission spectra and reliance on a single fluorescent parameter make the measurements susceptible to environmental interference, limiting their accuracy. Döring et al. systematically compared multiple sensing strategies and optimized the processing and modeling of multi-dimensional fluorescence data using ML algorithms, significantly improving the temperature-measurement accuracy of different CDs to 0.54 K, 1.04 K, representing an improvement of 23%-53% over the best single-variable fitting accuracy, as illustrated in Fig. 8(b) [120]. Nevertheless, this study also did not delve into the correlation between features and the physicochemical mechanisms of CDs, leaving the model essentially a black box.

The integration of carbon-based fluorescent sensor arrays with ML technology offers an innovative path for bacterial detection that is rapid, specific, and does not require complex instrumentation. Wang et al. constructed a paper-based sensor array by covalently linking antibiotics to carbon quantum dots and combined it with basic ML algorithms to achieve accurate discrimination of five bacterial species, within a detection range spanning from 1.0×10^3 CFU mL^{-1} to 1.0×10^7 CFU mL^{-1} [127]. Xiao et al. optimized data-feature extraction and algorithm selection within the same CQD-modified system and employed a Naïve Bayes algorithm to accomplish both bacterial species identification and concentration differentiation [128]. Although the model's prediction performance was excellent, its accuracy heavily depended on the representativeness and distribution of the training data, leading to a significant drop in accuracy for unknown samples or new environments. Zhang et al. upgraded the sensing material, utilized multi-dimensional fluorescence signals to train the model, and optimized algorithm parameters to enhance the stability and sensitivity of bacterial identification, as shown in Fig. 8(c) [121]. However, these studies focused primarily on bacterial species classification and identification; although concentration detection was mentioned, no systematic concentration-regression model or quantitative-prediction framework was established. Zhu et al. adopted a metal-doping strategy to construct an integrated recognition-inactivation platform [129]. ML not only achieved 100% rapid identification and quantification of bacteria but also further established a predictive model linking detection signals to bactericidal efficacy. Currently, ML models in this area often

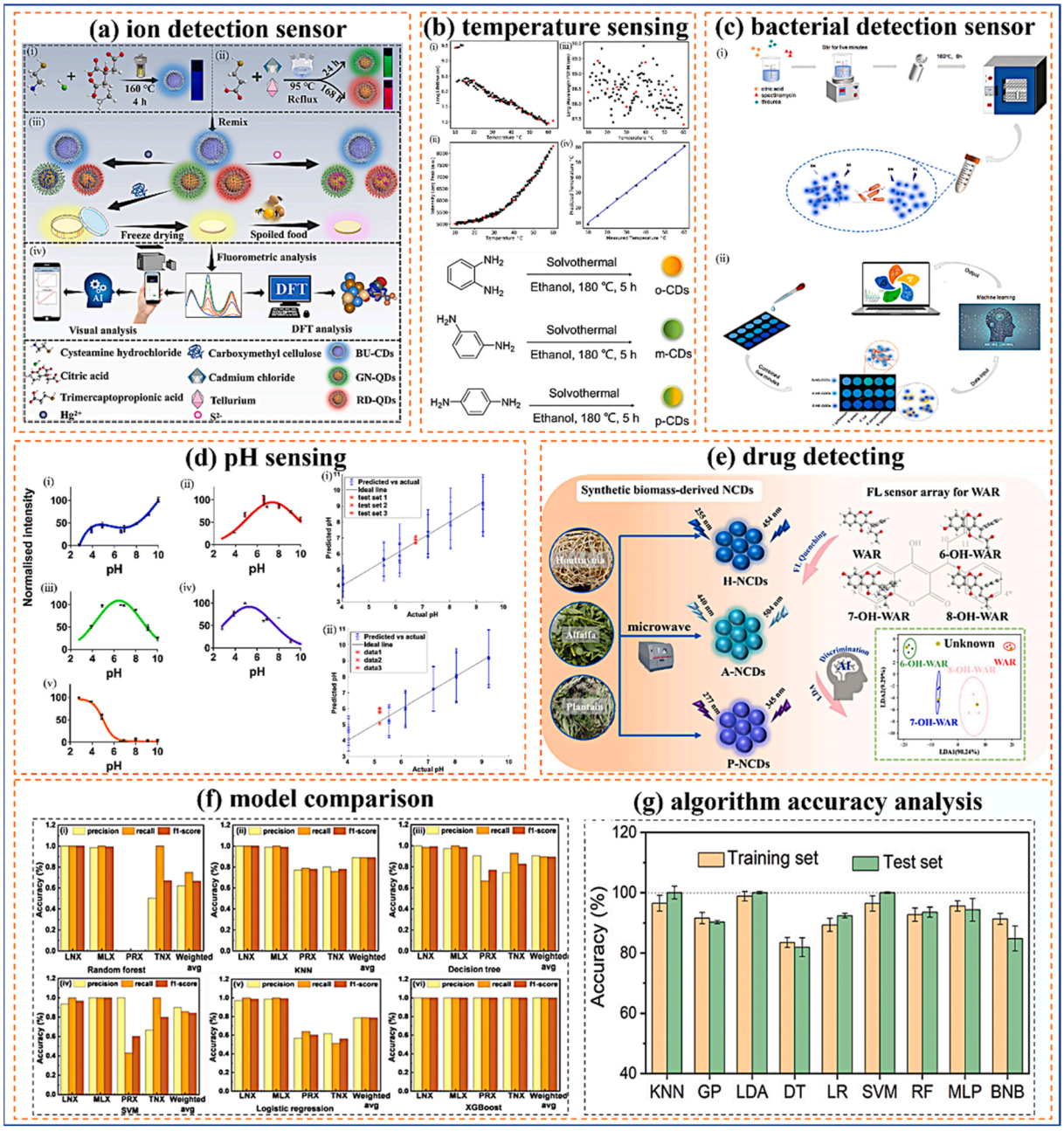


Fig. 8. (a) Schematic diagram of the ML system monitoring Hg^{2+} and sulfide using a polychromatic fluorescent colorimetric paper sensor [119]. Copyright 2023, with permission from American Chemical Society. (b) Augmentation and linear regression were applied to temperature sensing using p-CDs. Synthesis of o-CDs, m-CDs, and p-CDs, and color coding of the CDs spheres illustrates their emission color [120]. Copyright 2024, with permission from American Chemical Society. (c) Schematic illustration of the NS-GQDs synthesis method and ML-assisted the fluorescence sensor array to distinguish the five types of bacteria [121]. Copyright 2025, with permission from American Chemical Society. (d) Synthesized CDs with variable pH response patterns and GPR prediction by using two CDs array built by CD43 and CD69 [122]. Copyright 2024, with permission from the Royal Society of Chemistry. (e) Schematic illustration of the construction of three biomass-derived NCDs and application to precise discrimination of WAR and its metabolites [123]. Copyright 2025, with permission from American Chemical Society. (f) Comparison of precision, recall, and f1-score across six models for LNX, MLX, PRX, and TNX discrimination [124]. Copyright 2025, with permission from Springer Nature. (g) Accuracy analysis of training set and test set by 9 machine learning algorithms [125]. Copyright 2025, with permission from Elsevier.

rely on single-dimensional fluorescence data and exhibit poor adaptability to mixed-bacterial samples, commonly remaining trapped in small-data, demonstration-level applications. Future breakthroughs are expected to come from constructing large, high-quality databases and developing robust algorithms capable of handling complex real-world scenarios.

As one of the most frequently monitored chemical parameters, the high-precision measurement of pH distribution and dynamic changes is crucial in numerous chemical and life-science fields. However, developing nanoscale sensors that combine high spatial and temporal resolution and can map pH in real time within complex systems remains a key challenge in sensing technology [130]. Guo et al. regulated the surface functional groups of CDs to construct a multifunctional pH-sensor array and combined it with LDA and GPR to achieve accurate classification and quantitative detection within the pH range of 3-10, as shown in Fig. 8(d) [122]. However, in complex samples, certain substances may non-specifically absorb fluorescence or produce background signals; single-dimensional intensity data cannot distinguish

target pH response from interference signals, leading to detection errors. Moreover, single-dimensional data are susceptible to external influences, making it difficult to establish stable linear relationships, which increases the mean-square error of predictions and fails to meet high-precision detection requirements. Wu et al. constructed an anti-interference triple-emission ratiometric fluorescent sensor, integrated the YOLO v3/v5x/v8 series of deep-learning algorithms, and developed a smart detection mini-program [131]. This study converted fluorescence signals into RGB three-channel color data collectable by a smartphone, used algorithms to automatically identify and analyze the linear relationship between the R/G ratio and pH, and achieved quantitative detection, thereby realizing portable operation.

Currently, drug detection is at a critical juncture of transitioning from precise but cumbersome laboratory analysis toward intelligent, low-cost, highly selective, and convenient point-of-care diagnostics. Li et al. constructed a nitrogen-doped CDs fluorescent sensor array and employed an LDA algorithm to achieve discriminative classification of warfarin and its metabolites, as illustrated in Fig. 8(e) [123]. However, the study focused on classification and identification and did not utilize ML to construct concentration-prediction models, limiting the depth of application in quantitative analysis. Moreover, the study did not systematically compare multiple algorithms, raising questions about algorithm suitability. Going a step further, Zhang et al. employed a fluorine and nitrogen-co-doping strategy to successfully develop a dual-excitation F,N-CDs fluorescent sensor. They systematically compared six algorithms for evaluating six model prediction indicators, such as LNX, MLX, PRX, and TNX, as shown in Fig. 8(f). Finally, they achieved detection in a wide concentration range of 0.097 to 25 μM for different drugs of the sibutramine class, with a detection limit of 97 [124]. The model-training data in this work were based on single-drug concentration gradients; training samples covering mixed-drug proportions were not included, preventing the model from deducing the specific content of each drug in mixed systems from fluorescence signals. Zhang et al. used molecularly imprinted-modified CDs to construct a four-channel fluorescent sensor array, systematically evaluated nine ML algorithms, and found that the optimal model was LDA, as indicated in Fig. 8(g). This indicates that the current output features of the array are still fundamentally more suitable for linear discrimination rather than deeper mechanism learning or complex feature mining [125]. By leveraging specific recognition sites formed by different template molecules, distinct fluorescence-quenching fingerprint patterns were generated for different sulfonamide drugs in mixed system. Combined with the LDA algorithm, the simultaneous screening and quantitative analysis of 12 types of sulfonamide drugs were ultimately.

In most studies related to sensing, machine learning has significantly enhanced the interpretation, classification, and quantitative analysis of signals. However, the issue of the unclear relationship between the structure and application of carbon dots remains unsolved, which hinders the further development of CDs. Furthermore, the unclear structure–property–application relationships of CDs also hinder further advancement of ML algorithms. Therefore, in addition to establishing large-scale databases, subsequent research needs to thoroughly explore the intrinsic connections between synthesis parameters and the structure-application principles of CDs, integrating them with experimental results. Moreover, future sensors will no longer be passive measurement tools but intelligent systems capable of interacting with the environment and making autonomous decisions. Integrating CDs sensors with microfluidics and wearable devices, models can learn and compensate for environmental interference in real time, enabling continuous, non-invasive, real-time monitoring. By analyzing dynamic data trends, impending abnormal events can be predicted, and feedback signals can be sent to actuators, constructing a perception-decision-action closed loop. This evolution will ultimately shift the paradigm from static detection to dynamic feedback.

3.2. Photocatalysis and energy applications

Carbon dots exhibit unique and critical application value in the fields of photocatalysis and energy conversion, primarily owing to their excellent light-absorption capability, tunable band structure, efficient charge separation and migration characteristics, and favorable chemical and photostability [132]. In photocatalysis, CDs can not only serve as independent photocatalysts, directly driving pollutant degradation and water purification reactions via their surface active sites [133], but also, more importantly, act as efficient co-catalysts when combined with semiconductor materials. This effectively broadens the light-response range of the composite system, suppresses the recombination of photogenerated electron-hole pairs, and enhances interfacial charge transfer, thereby significantly improving the efficiency and selectivity of processes such as photocatalytic water splitting for hydrogen production, CO_2 reduction to fuels, and organic synthesis. However, the optimization of CDs functionality involves the complex synergy of multiple factors including elemental composition, size distribution, surface states, doping type and concentration, etc. Relying solely on experimental exploration to elucidate structure–property relationships is inefficient and costly. Against this backdrop, the deep integration of machine learning with CDs research has notably accelerated the development of novel CDs for photocatalysis and energy applications.

The oxidation of cyclohexane to adipic acid is an industrially important reaction, but existing technologies (e.g., nitric-acid oxidation) suffer from low efficiency and heavy pollution. Wang et al. synthesized polymer carbon dots and employed an XGBoost algorithm to predict the optimal catalytic conditions for C–H bond oxidation of cyclohexane, constructing a continuous quasi-homogeneous catalytic system as shown in Fig. 9(a) [134]. Although the continuous quasi-homogeneous catalytic process developed in this work offers an innovative paradigm for greening and intelligentizing chemical processes, the uneven distribution of training samples raises questions about its adaptability.

Conventional industrial methods for hydrogen peroxide (H_2O_2) production are energy-intensive and polluting, while photocatalytic approaches represent an ideal and environmentally benign alternative. CDs, owing to their unique optoelectronic properties, tunable interfacial chemistry, excellent stability, and green sustainable characteristics, demonstrate multiple advantages in photocatalytic H_2O_2 generation. ML further provides a data-driven solution for designing green catalytic systems. For instance, Liu et al. utilized ML combined with transient photovoltage technology to guide the design and synthesis of photocatalysts from biomass for photocatalytic H_2O_2 production, as illustrated in Fig. 9(b) [135]. Xu et al., on the basis of optimizing the synthesis of CDs-chitosan composite catalysts via ML, further investigated interfacial charge-transfer dynamics using techniques such as transient potential scanning and transient photocurrent measurements [139]. However, the ML models in these studies were optimized only based on initial synthesis conditions, without considering possible structural evolution, surface reconstruction, or deactivation mechanisms of the catalysts during photocatalytic processes.

Solar energy is indispensable on a global scale. However, current photovoltaic technologies are constrained by their cost-effectiveness and energy-conversion efficiency, which limits the deep development of a solar economy [140]. The development of novel photovoltaic materials that are low-cost, highly efficient, and environmentally friendly is therefore required. Mona et al. adopted a combined approach involving experimental measurements, numerical simulations, and ML to explore the forward-scattering mechanism introduced by CQDs in titanium dioxide and analyzed their potential in photovoltaic applications, as shown in Fig. 9(c) [136]. This research aimed to enhance the light-harvesting efficiency of photovoltaic devices, providing technical support for the design of light-management layers in solar cells. Nevertheless, the ML model is more akin to conducting statistical fitting of the concentration-morphology parameter-attenuation performance

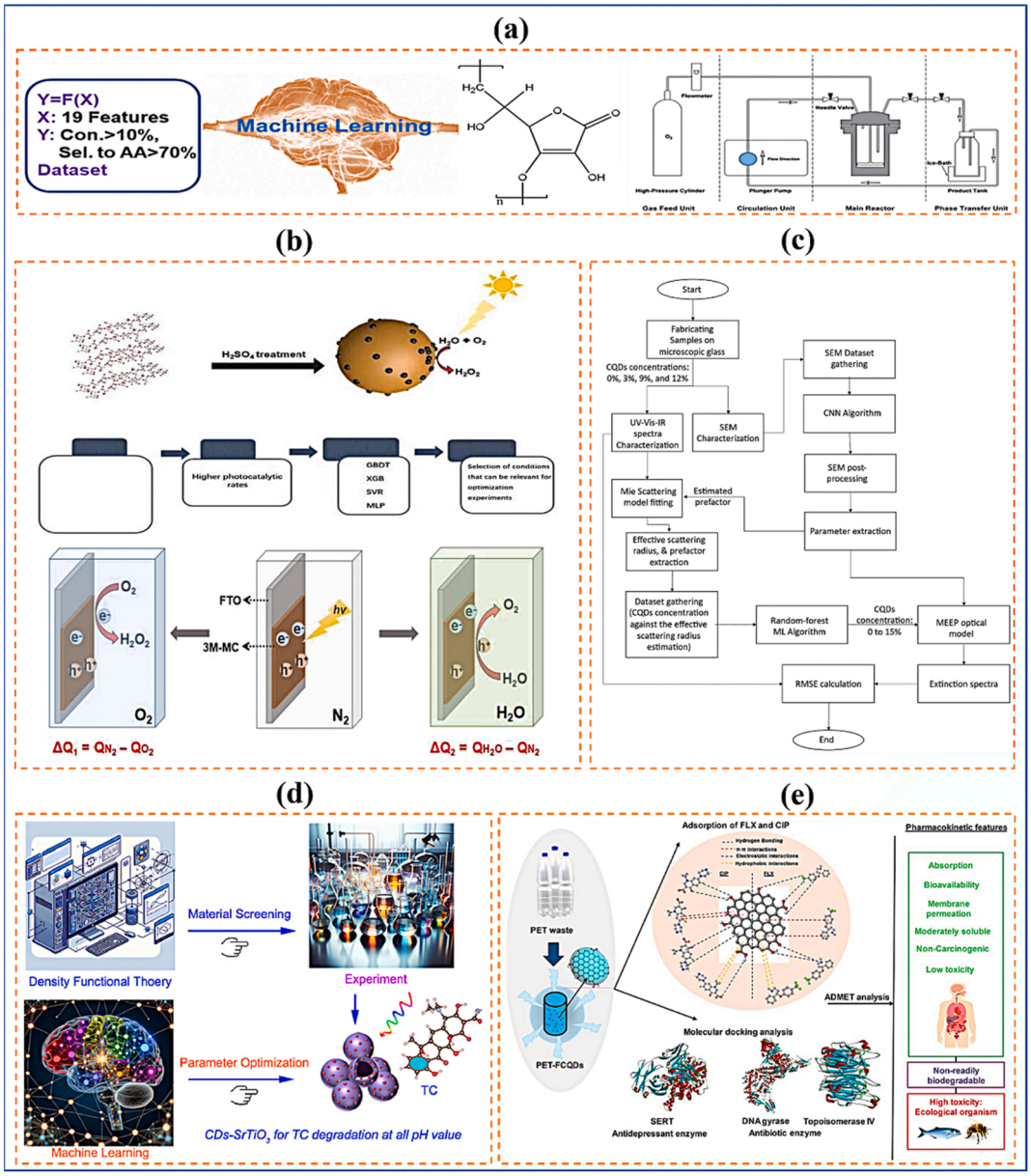


Fig. 9. (a) Schematic diagram of the catalytic oxidation design of cyclohexane the designed continuous flow reactor with four functional parts [134]. Copyright 2022, with permission from American Chemical Society. (b) Preliminary exploration through machine learning [135]. Copyright 2022, with permission from Springer Nature. (c) Overall experimental, numerical and ML-based optimization process flow diagrams of TiO_2 films containing CQDs at different concentrations [136]. Copyright 2024, with permission from Elsevier. (d) Schematic diagram of machine learning and DFT dual-guidance for carbon dots implanted $SrTiO_3$ hollow nanosphere for efficient all-pH-value photocatalysis [137]. Copyright 2025, with permission from Elsevier. (e) Schematic diagram of Synthesis, adsorption mechanism and application [138]. Copyright 2025, with permission from Elsevier.

relationship within a limited experimental space, rather than directly revealing the underlying physical mechanism of forward scattering, leaving the model's interpretability still relatively weak.

The degradation of organic pollutants and water-body remediation have become major challenges in environmental science [141,142]. Wang et al. optimized the CDs-$SrTiO_3$ photocatalytic system by combining DFT with ML. However, the machine learning part essentially mainly established an empirical mapping between experimental parameters and degradation performance, without explicitly incorporating the microscopic structural information such as Ti–C bond formation, surface oxygen functional groups, interface electron coupling, and PDS activation that determine the catalytic activity into the model, as depicted in Fig. 9(d) [137]. Due to the limited size of the training data and the lack of strict external validation, the model is more likely to learn the local statistical rules within the current experimental space rather than the structural-performance causal relationship with cross-system transfer ability. Therefore, the ML in this work is more suitable to be regarded as a parameter optimization tool rather than a mechanism-driven rational design framework. Li et al. synthesized CDs through different nitrogen-doping strategies, constructed a multi-dimensional response dataset, and employed various supervised-learning algorithms to achieve accurate identification of 16 types of dihydroxybenzene isomers and their mixtures [143]. Although the ANN performed best, its decision-making process remained a black box, making it difficult to interpret the classification logic, which hinders mechanistic understanding or further optimization of sensor design. Hussauni et al. investigated the performance of a zeolitic imidazolate framework-graphene quantum dot composite adsorbent for removing

Congo red dye and employed ensemble models to improve prediction accuracy [144]. Concurrently, SHAP analysis was incorporated to clarify key influencing factors, compensating for the lack of model interpretability.

PET plastic pollution permeates the entire lifecycle of energy production, storage, transmission, and waste treatment. Therefore, accurately monitoring PET-plastic content and adopting efficient removal/resource-recovery methods can not only directly reduce pollution load but also lower environmental costs through resource cycling. Enyoh et al. synthesized fluorescent carbon quantum dots by pyrolyzing PET plastic waste and, using Box-Behnken design combined with ML optimization, achieved efficient removal of fluoxetine and ciprofloxacin from aqueous solutions, as shown in Fig. 9(e) [138]. In this study, ML was only focused on predicting adsorption efficiency, without analyzing the correlation between feature variables and the adsorption mechanism. Priyanto et al., based on CDs photoluminescence imaging and ML, combined image-feature analysis with deep learning to realize highly sensitive quantitative detection of PET microplastics, and favorable detection limits ($LOD = 0.771 \pm 0.030\ mg\ L^{-1}$; $LOQ = 2.336 \pm 0.099\ mg\ L^{-1}$). SHAP analysis was used to clarify the physical meaning of key features [145]. Although this work addressed the model-optimization problem for single-pollutant detection, it likewise did not involve the treatment of multi-component complex systems, and its practical applicability requires further verification.

The integration of ML with CDs in photocatalysis and energy fields is driving a transition toward a digital-intelligent-manufacturing paradigm. However, its development is still constrained by challenges such as limited dataset sizes and the disconnect between experimental validation and real-world application scenarios. Future efforts should focus on constructing high-quality shared databases and promoting high-throughput automated experimental closed-loops to achieve an efficient leap from rational design to industrial application.

3.3. Biomedical

Carbon dots possess tunable luminescent properties, good water solubility, strong membrane-penetration ability, as well as advantages such as low toxicity and high biocompatibility, making them promising for a wide range of biomedical applications [146,147]. They are suitable not only for in vitro detection and imaging but also for in vivo disease diagnosis and therapy, with specific directions including bioimaging [148], drug or gene delivery [149], antibacterial and anticancer therapy [150], and biomarker detection [151], the detailed information is listed in Table 6. However, the fragmented research paradigm separating synthesis from application severely limits the development efficiency of high-performance CDs for diagnostic and therapeutic platforms. To address this challenge, integrating ML to connect rational design, controllable synthesis, and performance prediction of CDs has become an inevitable direction. During the synthesis stage, ML can guide the rational design and efficient synthesis of CDs with tailored properties by establishing predictive models linking synthesis conditions to structural features. At the application stage, ML can integrate multi-omics data and bioimaging information to achieve accurate prediction of therapeutic outcomes. Deep integration of ML with CDs biomedical research is expected to drive the field toward an intelligent research paradigm of prediction-design-validation, thereby enabling more efficient and precise development of next-generation CDs-based theranostic platforms.

Biomarkers refer to detectable indicators that objectively reflect physiological states, pathological changes, or external environmental influences in organisms, covering various types such as proteins, nucleic acids, and disease markers. As emerging nanomaterials, CDs exhibit significant potential for highly sensitive detection of diagnostic biomarkers due to their excellent photophysical properties and good biocompatibility. For example, Saren et al. constructed an OPCA model combined with a multicolor quantum-dot-based immunosensor to achieve quantitative detection of gastrointestinal tumor biomarkers, the concentrations of the corresponding tumor markers were all below 8.5–51.5 $ng\ mL^{-1}$, as shown in Fig. 10(a) [152]. However, ML in this study was used only for gene-feature screening and did not address intelligent analysis of detection signals or modeling of multi-marker interference. Yu et al. developed a DNA xerogel "probe" based on N-doped carbon quantum dots and utilized an SVM algorithm for rapid detection of the kidney-injury biomarker neutrophil gelatinase-associated lipocalin [158]. This approach addressed the deficiency in combining signals with ML, improving diagnostic accuracy, but still fell short in high-throughput analysis. Wu et al. developed a WeChat mini-program, constructing an intelligent sensing platform based on a triple-emission ratiometric fluorescent probe and deep learning to achieve high-throughput detection of spermine, thereby overcoming the issue of low sample-processing efficiency, though the system lacked the ability to distinguish interference in complex matrices [159]. Further advancing this direction, Wang et al. employed B, N-co-doped CDs, used K-Means for feature expansion, and applied RF model to realize ratiometric fluorescence and dual-mode detection of dopamine [160]. This study not only enhanced model interpretability but also identified key influencing factors through feature-importance analysis.

Conventional diagnostic methods for oral squamous cell carcinoma (OSCC) suffer from long processing times, complex operations, and dependence on expert experience, failing to meet the need for rapid intraoperative diagnosis. Xiao et al. developed a multi-channel nanosensor based on a core-shell CD/gold nanoparticle composite material [153]. By combining various ML algorithms, they achieved subtype classification and malignancy grading of OSCC, reducing the traditional pathological diagnosis time from 48 h to 15 min, as illustrated in Fig. 10 (b). However, the dataset was not fused with multi-source clinical data and relied solely on fluorescence signals, leaving the model's generalizability across different samples and centers unverified.

Due to their nanoscale size, multifunctional groups, and ability to penetrate cell membranes, CDs can serve as carriers for drug delivery, enabling targeted administration and controlled release of anticancer drugs. Moreover, conjugation of CDs with drugs can enhance the water solubility of poorly soluble drugs and reduce toxic side effects on normal cells. Ahmad et al. studied a nickel-carbon nanodot/glycerol composite

Table 6
Summary of research on ML-assisted CDs drug detection.

Carbon dots	Carbon source	Synthesis method	Target analyte	Best model	R^2	LOD	Ref.
N-CQDs	Carbamide, aniline, ethylenediamine	Hydrothermal	Neutrophil gelatinase-associated lipocalin	SVM	0.9970	0.0035 $ng\ mL^{-1}$	[137]
D-CNPs	Urea, anhydrous citric acid	Solvothermal	Spermine	YOLOv5	0.9937	0.18 μM	[143]
B, N-CDs	3-aminobenzeneboronic acid, proline	One-pot hydrothermal	Dopamine	D_8-RFR	0.9642	0.22 μM	[144]
FCQDs	4-(trifluoromethyl)benzene-1,2-diamine, citric acid	Pyrolysis	Fluoroquinolone antibiotics	FLD, HCA	>0.9900	4.20 nM	[148]
B-CDs; G-CDs	Urea and citric acid; Phenylenediamine, ascorbic acid	One-step hydrothermal	Antibiotics	LDA	0.9900	MNZ: 0.098 μM; NOR: 0.11 μM	[149]

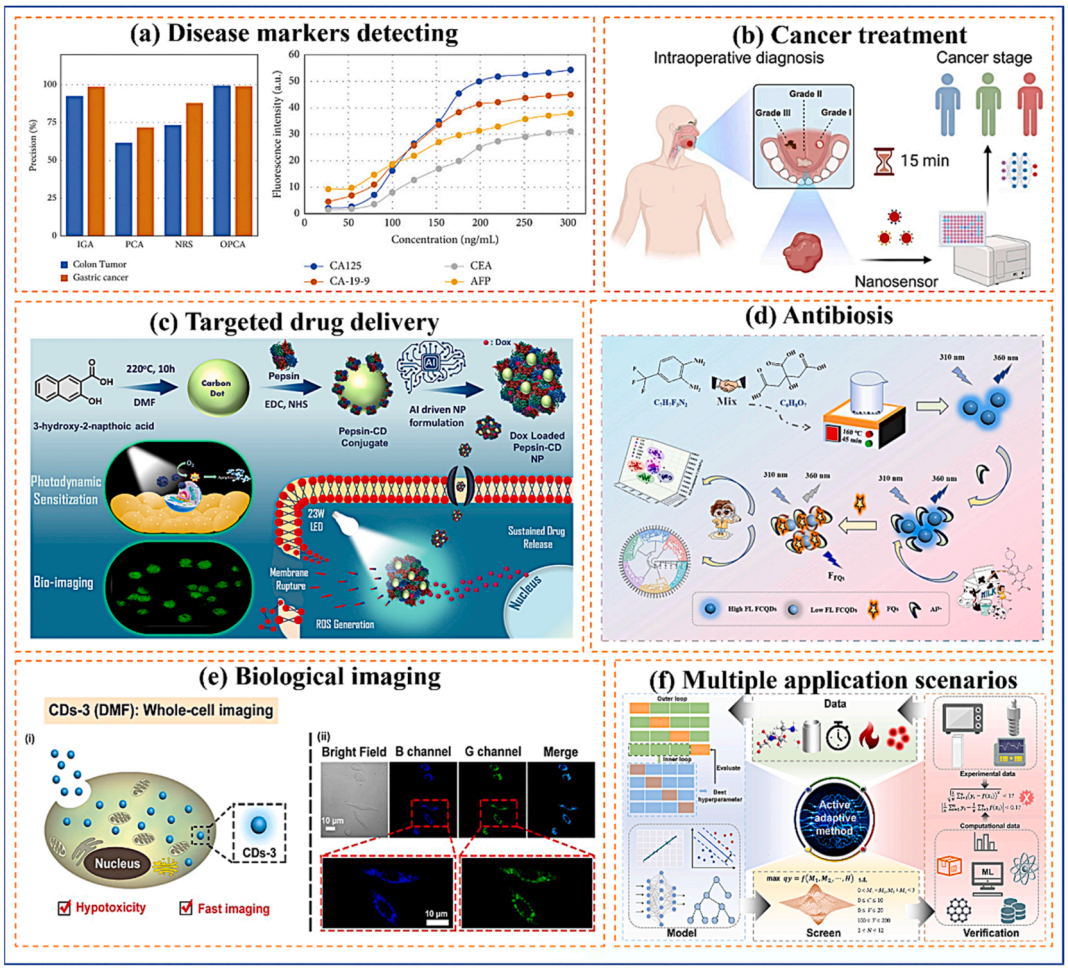


Fig. 10. Biomedical Applications of CDs. (a) Detection of disease markers [152]. Copyright 2022, with permission from John Wiley and Sons Ltd. (b) Cancer treatment [153]. Copyright 2025, with permission from Springer Nature. (c) Targeted drug delivery [154]. Copyright 2025, with permission from the Royal Society of Chemistry. (d) Antibiosis [155]. Copyright 2025, with permission from Elsevier. (e) Biological imaging [156]. Copyright 2012, with permission from American Chemical Society. (f) Multiple application scenarios [157]. Copyright 2024, with permission from Springer Nature.

nanofluid, employing a convolutional neural network (CNN) model combined with the finite-difference method to efficiently predict flow and thermal properties, providing theoretical support for applications of nanofluids in drug delivery and energy storage [161]. However, ML in this work was used only for optimizing and validating numerical results; no model was established to correlate nanoparticle concentration or structural features with thermal performance, and the prediction logic of the model was not interpreted, resulting in weak explainability. In contrast, Banerjee et al. directly developed a multifunctional nano-theranostic platform through covalent conjugation of CDs with pepsin. By using ML to optimize nanoparticle preparation conditions, they achieved synergistic functions including enhanced bioimaging, drug delivery, and photodynamic therapy, as shown in Fig. 10(c) [154]. Concurrently, SHAP analysis was applied to clarify key influencing factors, compensating for the lack of interpretability in the aforementioned study. Moreover, ML was extended from performance prediction to preparation optimization, accomplishing a transition from data fitting to mechanism-assisted interpretation.

Currently, antibiotics remain the most widely used antibacterial approach, primarily acting by directly killing or inhibiting bacterial growth [162]. Therefore, rapid and accurate detection technologies are needed to provide technical support for on-site screening and batch testing. Yin et al. used fluorine-doped carbon quantum dots as probes and combined LDA and HCA algorithms to achieve accurate classification of six structurally similar fluoroquinolone antibiotics and quantitative detection of ciprofloxacin, the LOD was 4.20 nM [155]. Nevertheless, ML in this study was employed only for signal clustering and verification of action mechanisms; it did not address the need for simultaneous detection of multiple antibiotic classes or rapid on-site analysis, and the model was not adapted for precise identification of multi-component interference in complex matrices, as illustrated in Fig. 10(d). These problems can be solved by introducing multi-label regression and blind source separation algorithms for the detection of multi-component coexistence. A portable rapid detection platform can be constructed by combining smart phones with paper-based chips. Moreover, transfer learning and computational chemistry can be utilized to enhance the model's robustness in complex matrices and the interpretability of the molecular mechanism. Wang et al. constructed a three-channel fluorescent array based on multicolor CDs and quantum dots, integrated LDA with a self-developed RTMDet deep-learning app, and achieved 100% accurate classification of eight antibiotics [163]. On-site visual analysis was accomplished via a smartphone app and portable device. Moreover, their model could quantify the concentration ratios of mixed antibiotics, further expanding the practical value of ML in antibiotic detection.

CDs can enter cells, interact with organelles such as the nucleus and mitochondria, and enable organelle-specific imaging. For instance, Hong et al. used an XGBoost model to achieve precise prediction of fluorescence intensity and emission centers of CDs at room temperature, successfully preparing CDs-3 with customized fluorescent properties and

applying them to HeLa-cell imaging, as shown in Fig. 10(e) [156]. The fluorescence signal of single-emission CDs is susceptible to environmental factors, whereas dual-emission CDs can use their two emission peaks as mutual internal references. By monitoring changes in the intensity ratio of the dual peaks, systematic errors caused by external interference can be compensated, making them particularly suitable for complex samples. Nurjis et al. employed Maillard-reaction by-products to construct a tunable dual-emission optical platform [164]. By using the Levenberg-Marquardt algorithm to fit the nonlinear decay kinetics of fluorescence intensity, the optical properties of two CQDs were accurately validated. Furthermore, both CQDs were applied to cell imaging and onion-epidermal-cell staining, demonstrating their potential as stable and practical bio-fluorescent probes. In this study, ML was utilized for post-processing of the experimental data.

Machine learning has demonstrated strong potential in materials science; however, most studies lack iterative feedback and a closed-loop experimental validation process. For instance, Tang et al. proposed an active-adaptive methodology (AAM) that integrates ML models to achieve intelligent optimization of CDs synthesis and a significant enhancement in quantum yield [157]. The optimized CDs were successfully applied to Co^{2+}/Fe^{3+} ion detection. Furthermore, the application scenarios of CDs were expanded, enabling their use in early caries diagnosis and tooth-remineralization therapy, as illustrated in Fig. 10(f). Nevertheless, the data source was limited, relying solely on the authors' own 200 experimental datasets without incorporating literature data or theoretical calculations. Consequently, the generalization capability of the model across different precursors and synthesis methods remains to be verified.

Currently, most ML-assisted CDs studies in the biomedical field are based on small-sample datasets, and the data often originate from single centers or experimental batches. The lack of external validation using multi-center, large-sample, and multi-source data leads to models that are prone to overfitting and exhibit insufficient generalizability, thereby limiting the universality and reliability of the conclusions. Additionally, systematic validation with animal models or clinical samples is often absent. To address these limitations, multi-center collaborations should be established to expand the scale of datasets and enhance the generalization capability of ML models across different precursors and synthesis processes. Concurrently, the application of CDs should be extended to advanced scenarios such as mult-target detection and in vivo imaging, further strengthening the synergistic innovation between algorithms and experiments.

3.4. Light-emitting diode

Light-emitting diodes are modern efficient solid-state devices for electro-optical conversion, have become core components of contemporary lighting and display technologies due to their low energy consumption, long lifetime, and environmental friendliness [140]. Compared with conventional rare-earth phosphors and potentially toxic metal-based semiconductor quantum dots, carbon dots offer advantages such as tunable emission color, high photoluminescence quantum yield, strong resistance to photobleaching, excellent photostability, and low cost, making them ideal candidates for next-generation fluorescent conversion materials in LEDs. Currently, CDs-based LEDs can be classified into two types: phosphor-converted LEDs (PC-LEDs) and electroluminescent LEDs (EL-LEDs). In PC-LEDs, CDs are typically employed as phosphors and excited by a blue or near-ultraviolet (NUV) chip serving as the light source [165]. In EL-LEDs, CDs function as the emissive layer that emits light upon electrical excitation [21]. However, the performance optimization of CD-LEDs involves the complex coupling of multiple factors, including the luminescent properties of CDs, charge-carrier transport, interface compatibility, and device architecture design. Relying solely on experimental trial-and-error for material screening and device-fabrication strategies is inefficient. To address these challenges, ML techniques have emerged. ML can be used to construct cross-scale correlation models, enabling intelligent material design, simulation and optimization of device performance, and prediction of lifetime and stability.

For example, Wang et al. applied ML to regulate the correlated color temperature (CCT) of white-light LEDs based on CDs [166]. Using XGBoost and MLP models, precise adjustment of white-light LEDs within the CCT range of 3093-11018 K was achieved, as shown in Fig. 11(a). However, ML in this study was only employed for performance-parameter matching and did not address the directional correlation between CDs synthesis parameters and optoelectronic properties. Furthermore, compared with CDs emitting at shorter wavelengths, red-emitting CDs possess lower energy, greater potential for deep-tissue fluorescence imaging, and more stable luminescence lifetimes, making them more suitable for imaging and optoelectronic applications. In light of this, Cao et al. developed highly efficient red-emitting nitrogen and sulfur-co-doped CDs. By implementing a MATLAB-based program for rapid latent-fingerprint recognition, the CDs were successfully applied in plant-growth LEDs, as illustrated in Fig. 11(b) [167].

The aforementioned studies required the separate preparation of red-emitting CDs to adapt to plant-illumination applications, indicating that functional extension relied on the independent development of specific materials. Zhang et al. prepared nitrogen-doped CDs and employed a CNN algorithm to construct a smartphone-based sensing platform for Pd^{2+} detection [168]. Subsequently, they reused the yellow-emission characteristic of the N-CDs to fabricate warm-white-light LEDs, thereby achieving multi-scenario application of CDs without the need for additional development of new CDs materials, as demonstrated in Fig. 11(c).

The electrical performance of diodes is significantly influenced by temperature. Conventional experimental measurement of current-voltage (I-V) characteristics across the full temperature range is time-consuming, labor-intensive, and costly. Existing applications of ML in diode-parameter prediction have mostly focused on single models, lacking multi-algorithm comparisons and in-depth adaptation to temperature-dependent properties. For instance, Ersöz et al. prepared a nanocomposite diode using lanthanum-doped polyethyleneimine-functionalized graphene quantum dots as the core component [169]. By employing a gradient-boosting (GB) model, they achieved accurate prediction of temperature-dependent I-V characteristics, thereby extending the application of ML to electrical-performance modeling of carbon-based electronic devices, as shown in Fig. 11(d).

In ML-assisted CDs-LED research, limitations such as single-source data, model interpretability confined to feature ranking, and a lack of deep mechanistic analysis have not yet been fully overcome. Future efforts should focus on multi-dimensional data fusion and mechanism interpretation as core objectives to facilitate the industrial application of LEDs. Moreover, most current CDs-based LEDs operate in photoluminescence mode; the development of efficient and stable electroluminescent CDs-LEDs using ML represents an important future direction. Equally important, leveraging ML to assist the inverse design of CDs for light-emitting devices that meet the requirements of Micro-LEDs, wearable electronics, and epidermal electronics is an emerging research frontier.

3.5. Information anti-counterfeiting encryption

In recent years, with continuous advances in technology, traditional anti-counterfeiting materials have increasingly struggled to meet the growing demand for security and functionality, creating an urgent need to develop novel and efficient luminescent anti-counterfeiting materials. Carbon dots, owing to their facile preparation, low cost, high stability, excellent water solubility, and combined fluorescence and afterglow luminescence, have emerged as ideal candidates for constructing advanced anti-counterfeiting inks. Moreover, the abundant hydroxyl groups on filter-paper surfaces provide ample hydrogen-bonding sites

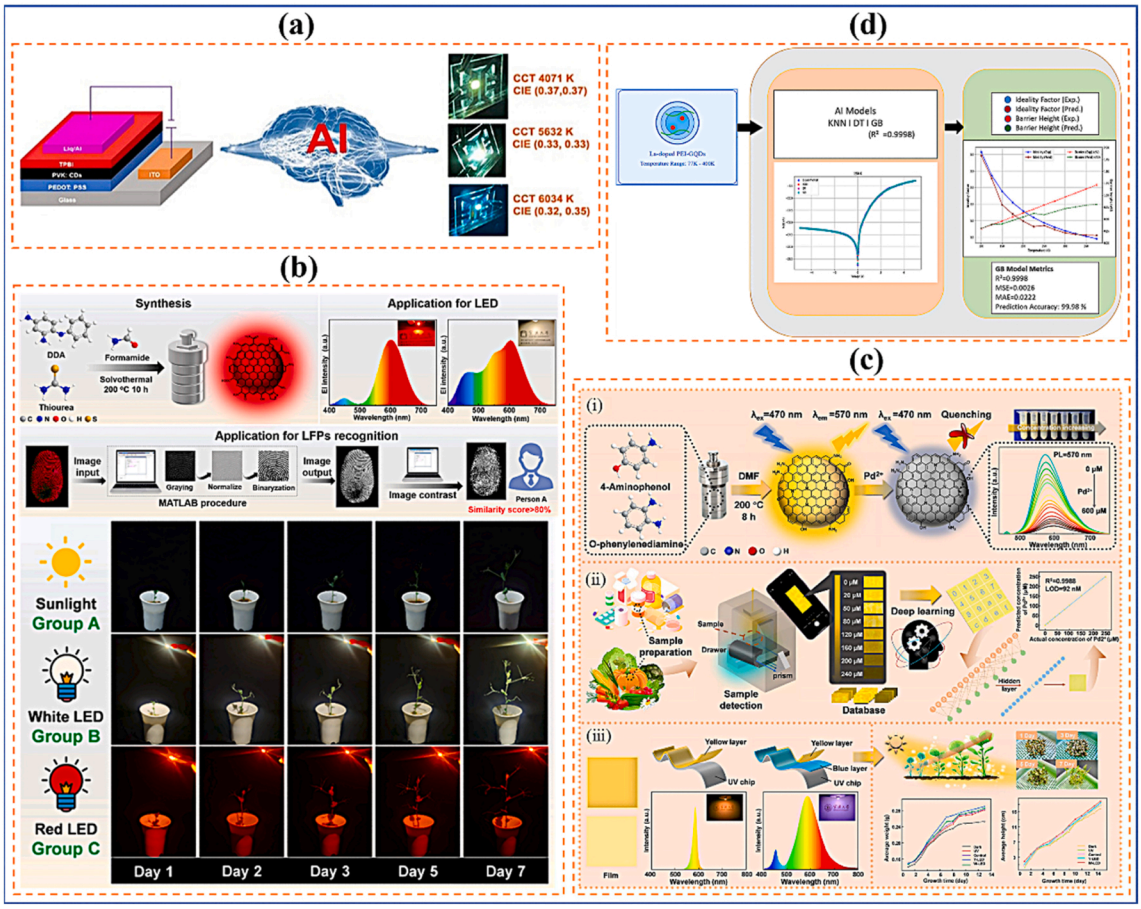


Fig. 11. (a) Schematic diagram of a CD WLEDs with adjustable correlation color temperature based on ML [166]. Copyright 2021, with permission from John Wiley and Sons Ltd. (b) The progress in synthesis and applications of the prepared R-CDs, as well as panoramic images of pease growth for 7 days under the condition of sunlight, WLED and RLED [167]. Copyright 2024, with permission from Elsevier. (c) Diagram of the synthesis process of N-CDs and multiple applications of N-CDs [168]. Copyright 2025, with permission from Springer Nature. (d) Schematic diagram of artificial intelligence driving data generation based on diode temperature-dependent current-voltage characteristics [169]. Copyright 2025, with permission from Elsevier.

for CDs, promoting strong adhesion and stable luminescence, making them promising for next-generation security inks.

The luminescent properties of CDs are largely governed by the synergistic regulation of multiple factors, including precursor selection, reaction conditions, and surface modification [101]. This complex interdependence makes the optimization of CDs-based anti-counterfeiting materials via traditional experience-driven approaches inefficient, costly, and poorly reproducible. Within this context, ML offers a new pathway. For instance, Chen et al. systematically developed an ML-assisted method for the controlled synthesis of multicolor CDs, enabling rational design and optimization of their optical properties and application in high-capacity information encryption, as illustrated in Fig. 12(a) [170]. Since the solvent itself is not a simple numerical variable suitable for direct input into the model, the authors incorporated the physical and chemical nature of the solvent into the model, moving from simple empirical optimization to a data-driven understanding of the synthesis-luminescence relationship. This provides new ideas for more rational design and improvement of the performance of CDs in the future. If the model is provided with the microscopic structural information of carbon dots, such as local conjugated domains, surface defect states, and the distribution of heteroatoms, it will reveal the deep structural-property causal relationship.

Compared with fluorescent CDs, phosphorescent CDs possess long-lifetime afterglow emission, which allows information encryption and delayed reading in the time dimension, effectively countering instant counterfeiting. Furthermore, phosphorescent emission does not require continuous excitation and can persist after the excitation source is turned off, not only improving the concealment of anti-counterfeiting features but also enabling the construction of more complex dynamic encryption patterns through multi-color phosphorescence design, greatly enhancing the security level and difficulty of imitation. Muyassiroh et al. used ML-guided synthesis to prepare room-temperature phosphorescent CDs, achieved optimization of phosphorescence lifetime, and constructed a time-dependent information-encryption system, as shown in Fig. 12(b) [171]. In this process, ML was employed for parameter optimization, but it did not address the precise customization of multi-color phosphorescence. More importantly, CDs relying solely on single-fluorescence emission still face limitations such as auto-fluorescence background interference and limited lifetime in practical applications [173]. Yan et al. proposed a method that combines ML to simultaneously predict multi-dimensional properties covering both fluorescence and phosphorescence, achieving multi-color customization and advanced encryption, as depicted in Fig. 12(c) [82]. In the studies mentioned above, ML served as a predictive tool to assist material synthesis. In contrast, Wang et al. employed ML as a decryption tool. By leveraging the time-resolved phosphorescence characteristics and spatial arrangement of carbon nitride quantum dots, they constructed a CNN-driven multi-dimensional encryption platform [172]. Through matrix encoding and Rubik's-cube-like transformations, multiple layers of information confusion were achieved, significantly increasing encryption complexity and resistance to cracking, as demonstrated in Fig. 12(d).

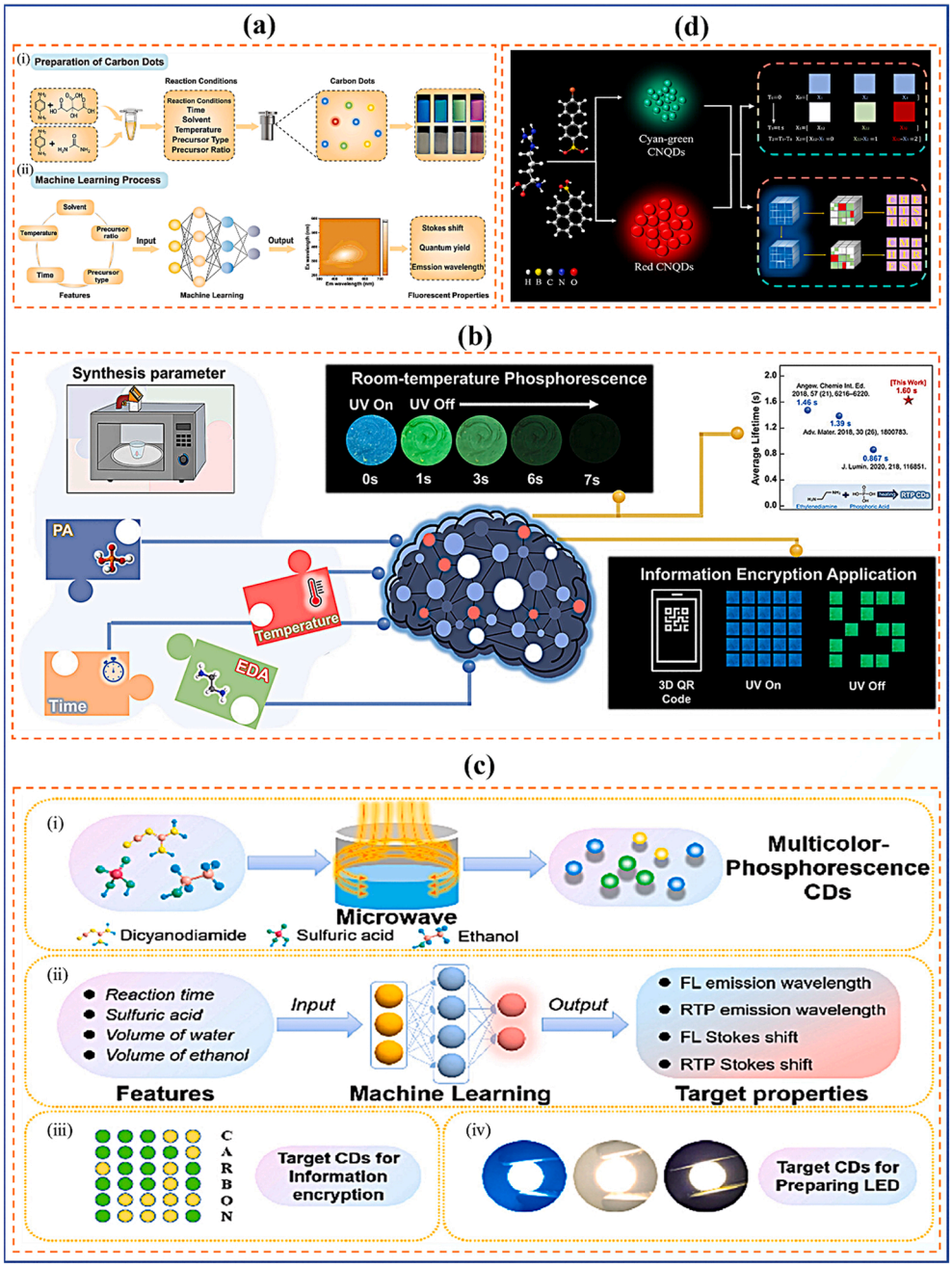


Fig. 12. (a) ML-assisted multicolor CDs synthesis workflow and application [170]. Copyright 2023, with permission from John Wiley and Sons Ltd. (b) Schematic diagram of the process for synthesizing RTP CDs guided by ML to enhance phosphorescent lifetime and information encryption [171]. Copyright 2024, with permission from American Chemical Society. (c)Workflow and application of ML-assisted multicolor phosphorescent CDs preparation and applications [82]. Copyright 2025, with permission from Elsevier. (d) Schematic Diagram of CNN for research and application of Long-life and multi-color RTP- CQDs [172]. Copyright 2025, with permission from American Chemical Society.

The main bottleneck in current research utilizing ML lies in the insufficient integration of multi-dimensional data and the disconnect between mechanisms and applications. Future efforts should focus on multi-source data integration, multi-objective algorithm optimization, and the construction of intelligent systems, aiming to realize the on-demand design and intelligent manufacturing of CDs-based optical-information materials, as well as their high-reliability applications in smart-security fields.

3.6. Others

As research deepens, ML-assisted CDs studies are being extended to more diverse applications. For instance, Bian et al. addressed the public health threat posed by multidrug-resistant bacteria (MRB) by focusing on the efficient development of antibacterial CDs [174]. They proposed an ML-assisted strategy that explores the intrinsic relationships between CDs synthesis conditions, physicochemical properties, and antibacterial performance, thereby enabling rapid screening and performance prediction of antibacterial CDs, as shown in Fig. 13(a). Momina et al. synthesized CDs from sugarcane bagasse as an adsorbent for rapid removal of cationic and anionic dyes from wastewater [175]. By employing ML to accurately capture the nonlinear relationships in the adsorption process, a low-cost and sustainable solution for wastewater treatment was provided, as illustrated in Fig. 13(b). Solomon et al. prepared nitrogen- and sulfur-co-doped CDs (N,S-CDs) that serve as efficient corrosion inhibitors with both environmental friendliness and multi-metal protection capability [176]. Through a combination of experiments, DFT, and ML, the underlying mechanism of action was clarified, offering an environmentally benign solution for multi-metal corrosion protection, as illustrated in Fig. 13(c).

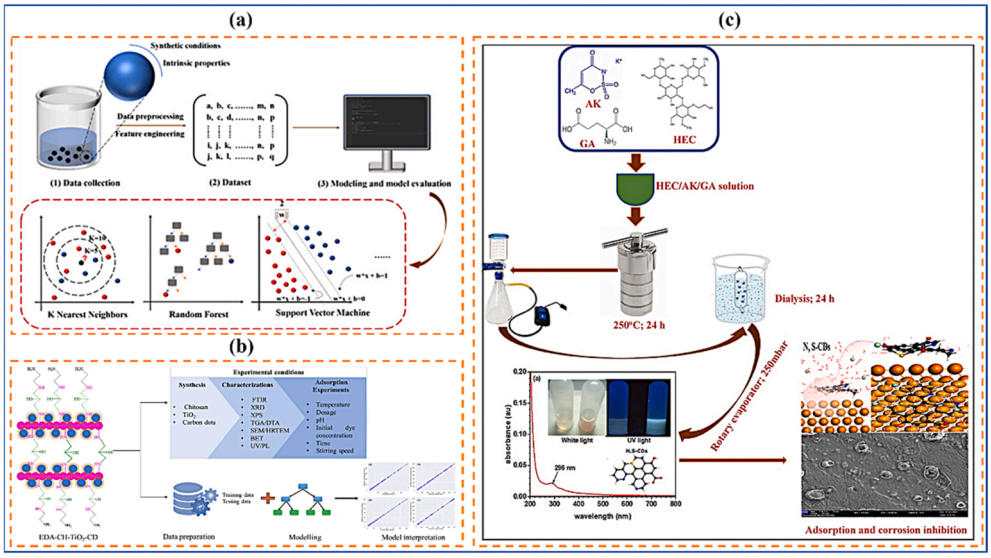


Fig. 13. ML-assisted CDs application. (a) Antibacterial [174]. Copyright 2024, with permission from Dove Medical Press. (b) Adsorption [175]. Copyright 2024, with permission from Elsevier. (c) Corrosion inhibition [176]. Copyright 2025, with permission from Elsevier.

4. Challenges and outlook

The integration of machine learning with CDs research is gradually shifting the field from experience-driven exploration toward a data-driven research paradigm. However, the development of ML-assisted CDs research still faces numerous challenges: such as insufficient data quality and standardization, limited mechanism explanation, slow progress from prediction to actual material design, and the lack of a complete path from laboratory-scale optimization to practical deployment. Therefore, future progress not only depends on improving model performance, but also requires strengthening the connections between data infrastructure, physical understanding, material design, and scalable applications.

4.1. LLM-assisted construction and updating of CDs databases

The lack of high-quality, standardized, and comparable data remains a primary bottleneck limiting the robustness and generalization of ML models in CDs research. To alleviate this constraint, LLMs can be employed to assist in semi-automated literature data extraction, normalization, and curation, enabling the construction of an open and structured CDs-specific database, avoiding the risks of small sample sizes and overfitting. During this process, it is necessary to pay attention to the instability of entity recognition, the incompleteness and inconsistency of synthetic condition reports, the variability of characterization protocols, the strong context dependence of application-related performance indicators, and the frequent distribution of key information across charts, tables, and supplementary materials.

4.2. Integrating physical interpretability with data-driven mechanistic analysis

The second major challenge is that the relationship between the structure and properties of CDs has not been fully elucidated. Although many research reports have shown encouraging accuracy in property prediction, optimization, or application layer analysis, these results are often closer to empirical fitting rather than true learning of the underlying physicochemical causal relationships. Future efforts should therefore integrate physical constraints and chemical prior knowledge, for example through multiscale simulations such as density functional theory and molecular dynamics. In parallel, interpretability tools including SHAP and LIME can be employed to quantify the contributions of synthesis parameters and structural descriptors, thereby enhancing the mechanistic relevance of ML-based predictions.

4.3. Precursor-based intelligent design for accelerated materials discovery

Another core challenge is that currently, most applications of machine learning in CDs research still mainly involve prediction and optimization within known systems rather than specifically designing new materials, resulting in poor generalization ability of the models. In the future, it is necessary to shift from process-centered optimization to structure and precursor-oriented intelligent design. This requires more incorporation of precursor molecular structures, functional groups, local conjugation features, heteroatom chemistry and quantum chemical descriptors in the models, in order to closely link the structure and properties of CDs, expand the model application, and then transition from empirical optimization to rational material design.

4.4. Towards intelligent, scalable and sustainable manufacturing of CDs

At present, the research on CDs mainly remains at the stage of laboratory-scale batch synthesis, lacking systematic studies for production on a kilogram-scale. In the future, it is necessary to combine continuous-flow synthesis platforms with real-time process monitoring and machine learning-driven control strategies to build an intelligent feedback optimization system, thereby achieving dynamic regulation and stable scaling of the synthesis process. On this basis, by integrating green chemistry principles, developing sustainable manufacturing paths based on biomass precursors, green solvent systems and low-energy consumption processes will help reduce environmental burdens and enhance the practical application feasibility of materials. Through the coordinated advancement of intelligence and greenness, the production of CDs is expected to achieve repeatable, scalable and energy-saving industrial manufacturing, while meeting the dual requirements of high performance and environmental compatibility.

5. Conclusions

Machine learning has significantly improved the efficiency of performance screening and application realization in CDs research, yet it has not fundamentally transformed the knowledge-generation paradigm for CDs material design. Within established precursor systems and synthesis pathways, ML models are capable of effectively narrowing the

experimental search space through multi-parameter correlation learning and demonstrate clear ad-vantages in obtaining key optical properties such as emission wavelength, photoluminescence quantum yield, and afterglow lifetime. However, these advantages are primarily manifested in accelerating the exploration of existing chemical space rather than providing reliable extrapolation to unknown structure–property relationships. From an application perspective, beyond substituting for material design per se, the core value of ML in CDs-based sensing, information encryption, and related optical applications lies in its efficient decoding capability for complex, multi-channel optical signals, thereby significantly enhancing the recognizability and information capacity of the system in real-world scenarios. In contrast, in engineering-oriented systems such as CDs-based light-emitting devices, which place greater emphasis on structural controllability and mechanistic closed-loop understanding, current ML research remains largely confined to performance correlation and parameter optimization. Its leading role in design is constrained by factors such as limited data scale, absence of physical constraints, and sample homogeneity. Overall, ML-assisted CDs research still faces a series of core challenges spanning the entire data-mechanism-synthesis-application chain, mainly reflected in data foundation, mechanistic understanding, controllable fabrication, and practical application. Addressing these challenges requires coordinated innovation in database construction, model development, synthesis strategies, and industrial advancement to promote the continued development and practical implementation of this interdisciplinary field.

Ultimately, whether machine learning evolves from an efficient accelerator into a genuine design engine for carbon dots will be determined by its ability to couple data quality, physical interpretability, synthesis controllability, and application-driven validation within a unified research workflow.

CRediT authorship contribution statement

Shudie Shen: Data curation, Investigation, Writing – original draft. **Leilei Zhang:** Visualization. **Linquan Gan:** Formal analysis. **Tao Li:** Software. **Wenli Zhao:** Software. **Yanfei Wang:** Project administration. **Liang Zhu:** Conceptualization, Supervision, Writing – review & editing. **Xiaoyu Zhao:** Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China [grant numbers NNSFC 22108204 and NNSFC 22178272], the Innovative Research Team of the Tianjin Municipal Education Commission [grant numbers TD13-5008], and the Yangtze Scholars and Innovative Research Team in Chinese University [grant numbers IRT-17R81]

References

[1] X. Xu, R. Ray, Y. Gu, H.J. Ploehn, L. Gearheart, K. Raker, W.A. Scrivens, Electrophoretic analysis and purification of fluorescent single-walled carbon nanotube fragments, J. Am. Chem. Soc. 126 (2004) 12736–12737.
[2] Y.-P. Sun, B. Zhou, Y. Lin, W. Wang, K.A.S. Fernando, P. Pathak, M.J. Meziani, B. A. Harruff, X. Wang, H. Wang, P.G. Luo, H. Yang, M.E. Kose, B. Chen, L.M. Veca, S.-Y. Xie, Quantum-sized carbon dots for bright and colorful photoluminescence, J. Am. Chem. Soc. 128 (2006) 7756–7757.
[3] H. Liu, T. Ye, C. Mao, Fluorescent carbon nanoparticles derived from candle soot, Angew. Chem. 119 (2007) 6593–6595.
[4] J. Zhou, C. Booker, R. Li, X. Zhou, T.-K. Sham, X. Sun, Z. Ding, An electrochemical avenue to blue luminescent nanocrystals from multiwalled carbon nanotubes (MWCNTs), J. Am. Chem. Soc. 129 (2007) 744–745.
[5] A.B. Bourlinos, A. Stassinopoulos, D. Anglos, R. Zboril, M. Karakassides, E. P. Giannelis, Surface functionalized carbogenic quantum dots, Small 4 (2008) 455–458.
[6] H. Peng, J. Travas-Sejdic, Simple aqueous solution route to luminescent carbogenic dots from carbohydrates, Chem. Mater. 21 (2009) 5563–5565.
[7] H. Zhu, X. Wang, Y. Li, Z. Wang, F. Yang, X. Yang, Microwave synthesis of fluorescent carbon nanoparticles with electrochemiluminescence properties, chem, Commun. Now. 34 (2009) 5118–5120.
[8] B. Zhang, C. Liu, Y. Liu, A novel one-step approach to synthesize fluorescent carbon nanoparticles, Eur. J. Inorg. Chem. 2010 (2010) 4411–4414.
[9] Y. Yang, D. Wu, S. Han, P. Hu, R. Liu, Bottom-up fabrication of photoluminescent carbon dots with uniform morphology via a soft–hard template approach, Chem. Commun. 49 (2013) 4920–4922.
[10] X. Huang, L. Bao, R. Bao, L. Liu, J. Tao, J. Wang, Z. Zhang, Z. Ge, S. Tan, J. Yi, F. Meng, Reinforced copper matrix composites with highly dispersed nano size TiC in-situ generated from the carbon polymer dots, Adv. Powder Mater. 2 (2023) 100090.
[11] Y. Wang, S. Wu, High-performance green fluorescent carbon dots: synthesis, properties, and applications in temperature sensing, and information anti-counterfeiting, Mater. Sci. Eng. B 319 (2025) 118355.
[12] S. Ghanbarlou, D. Kahforoushan, H. Abdollahi, P. Zarrintaj, A. Alomar, C. Villanueva, S.M. Davachi, Advances in quantum dot-based fluorescence sensors for environmental and biomedical detection, Talanta 294 (2025) 128176.
[13] W. Fawaz, J. Hasian, I. Alghoraibi, Synthesis and physicochemical characterization of carbon quantum dots produced from folic acid, Sci. Rep. 13 (2023) 18641.
[14] A. Barati, M. Shamsipur, E. Arkan, L. Hosseinzadeh, H. Abdollahi, Synthesis of biocompatible and highly photoluminescent nitrogen doped carbon dots from lime: analytical applications and optimization using response surface methodology, Mater. Sci. Eng. C 47 (2015) 325–332.
[15] C. Xia, S. Zhu, T. Feng, M. Yang, B. Yang, Evolution and synthesis of carbon dots: from carbon dots to carbonized polymer dots, Adv. Sci. 6 (2019) 1901316.
[16] S. Sahana, A. Gautam, R. Singh, S. Chandel, A recent update on development, synthesis methods, properties and application of natural products derived carbon dots, Nat. Prod. Bioprospecting. 13 (2023) 51.
[17] R. Lamba, Y. Yukta, J. Mondal, R. Kumar, B. Pani, B. Singh, Carbon dots: synthesis, characterizations, and recent advancements in biomedical, optoelectronics, sensing, and catalysis applications, ACS Appl. Bio Mater. 7 (2024) 2086–2127.
[18] M. Singh, I. Malik, F. Hazzazi, A. Kumar, Advances in carbon-dots based nanomaterials for electrocatalytic water splitting: fundamental, challenges, and future prospects, Int. J. Hydrog. Energy 179 (2025) 151623.
[19] H. Ren, A. Labidi, Z. Miao, J. Liang, X. Feng, M. Li, Y. Zhao, C. Wang, Emerging of carbon quantum dots-based powder materials for photocatalytic environmental remediation and chemical production, Adv. Powder Mater. 5 (2026) 100372.
[20] Q. Ni, M. Zhang, B. Tang, W. Hou, K. Wang, H. Guo, J. Zhang, T. Han, M. Wu, L. Wang, Rapid synthesis of carbon quantum dot-integrated metal–organic framework nanosheets via electron beam irradiation for selective 5-hydroxymethylfurfural electrooxidation, Adv. Powder Mater. 4 (2025) 100267.
[21] R. Chen, Z. Wang, S. Zheng, N. Jiang, K. Lin, L. Zeng, R. Zhang, F. Huang, D. Chen, Carbon dots: electroluminescent devices for lighting and display applications, 2025 100517 eScience.
[22] M. Alafeef, I. Srivastava, T. Aditya, D. Pan, Carbon dots: from synthesis to unraveling the fluorescence mechanism, Small 20 (2024) 2303937.
[23] R. Mohammadzadeh Kakhki, M. Mohammadpoor, Machine learning-driven approaches for synthesizing carbon dots and their applications in photoelectrochemical sensors, Inorg. Chem. Commun. 159 (2024) 111859.
[24] J. Yu, X. Yong, Z. Tang, B. Yang, S. Lu, Theoretical understanding of structure–property relationships in luminescence of carbon dots, J. Phys. Chem. Lett. 12 (2021) 7671–7687.
[25] X. Zhang, W. Zhou, X. Li, T. Xu, Y. Yu, L. Zheng, G. Jin, S. Zhang, Machine learning-driven insights into the microstructure and properties of high-entropy alloys, Adv. Powder Mater. 4 (2025) 100331.
[26] J. Tu, Z. Wang, W. Li, L. Tan, L. Huang, F. Liu, SAMM: a general-purpose segmentation model for material micrographs based on the segment anything model 2, Adv. Powder Mater. 5 (2026) 100404.
[27] Z. Hu, C. Huang, L. Xie, L. Hua, Y. Yuan, L. Zhang, Machine learning assisted quality control in metal additive manufacturing: a review, Adv. Powder Mater. 4 (2025) 100342.
[28] Z. Liang, Y. Deng, Z. Shi, X. Liao, H. Zong, L. Ren, X. Li, X. Zeng, P. Hu, W. Ke, B. Wu, K. Wang, J. Qian, W. Xu, F. Chen, AI-driven design of powder-based nanomaterials for smart textiles: from data intelligence to system integration, Adv. Powder Mater. 5 (2026) 100356.
[29] T. Jena, S. Shinde, S.E. Roselin, C. Sujana, V.S. Rao, K. Immanuvel Arokia James, Deep learning-optimized carbon quantum dot biosensors for emerging contaminant monitoring, Microchem. J. 218 (2025) 115389.
[30] S. Srinivasulu, G. Nagarajan, R.I. Minu, Quantum deep learning assisted carbon quantum dot sensors for classification of emerging organic pollutants in air quality monitoring, Microchem. J. 221 (2026) 116703.
[31] M. Guan, G. Wang, K. Li, X. Liu, Z. Guo, M. Li, Controlled synthesis of multicolor fluorescent carbon quantum dots based on phenylenediamine isomers and their application in nitrite smart sensing and fingerprint trace detection systems, Dyes Pigments 246 (2026) 113400.

[32] B.U.M. Enriquez, M. Rangel-Ayala, Y. Kumar, J.E. Garcia, J.F. Gomez-Aguilar, S. Khandual, V. Agarwal, Multifunctional arthrospira platensis biomass derived carbon dots: sensing/removal of heavy metal ions, high-power light-emitting devices, and some machine learning assisted approaches for solid state sensor, J. Environ. Chem. Eng. 13 (2025) 117827.
[33] S. Cao, S. Dong, Y. Chen, J. Shen, Q. Cui, C. Sun, J. Ning, L. Li, Rapid fluorescence detection of black tea fermentation degree based on cobalt ion mediated carbon quantum dots, Food Control 165 (2024) 110610.
[34] M. Wei, M. Yang, H. Leng, Y. Shu, A novel intelligent sensing strategy: integration of metal-doped carbon dots nanozymes and machine learning for rapid screening of biothiols in disease, Sens. Actuators B Chem. 444 (2025) 138304.
[35] X. Zhang, W. Zhu, S. Zhang, F. Wang, Machine learning-driven graphene quantum dot fluorescence array for rapid heavy metal ions detection and identification, Environ. Res. 290 (2026) 123489.
[36] J. Chen, M. Zhang, Z. Xu, R. Ma, Q. Shi, Machine-learning analysis to predict the fluorescence quantum yield of carbon quantum dots in biochar, Sci. Total Environ. 8896 (2023) 165136.
[37] S. Bao, S. Zhi, X. Yang, Y. Wang, Z. Yang, J. Liu, P. Pan, Y. Qi, J. He, Deep learning-based GQDs/AgNPs composite enhanced electrochemiluminescence portable biosensor for ultrasensitive point-of-care methimazole monitoring, Microchem. J. 217 (2025) 114910.
[38] S. Ahmad, H.A. Alsalmah, Exploring the thermal attributes of nano-composition (GQDs+BiSe+Ag) suspended in therminol VP-1: an artificial intelligence based approach, Case Stud. Therm. Eng. 72 (2025) 106231.
[39] H. Liu, J. You, C. Liu, Z. Zhang, A. Sun, G. Hao, X. Shi, Machine learning-assisted wide-gamut fluorescence visual test paper for propazine determination in fish and seawater samples, Sens. Actuators B Chem. 413 (2024) 135843.
[40] D. Chang, H. Shi, J. Feng, Y. Liu, Y. Yan, L. Shi, Y. Yang, S. Shuang, C. Dong, Coupling machine learning with smartphone-assisted sensing platform: sample-to-answer visual detection of amoxicillin using ratiometric fluorescence carbon dots, Microchem. J. 216 (2025) 114505.
[41] X. Zhu, T. Li, X. Hai, S. Bi, A nanozyme-based colorimetric sensor array as electronic tongue for thiols discrimination and disease identification, Biosens. Bioelectron. 213 (2022) 114438.
[42] V. Gurjar, N. Nazeer, A. Bhargava, N. Soni, P. Kaur, R. Tiwari, P.K. Mishra, Deep learning-enabled nanophotonic test leveraging poly-L lysine-tethered carbon quantum dots to assess the risk of cardiovascular disease, Microchem. J. 207 (2024) 112164.
[43] Ş. Özönder, H.K. Küçükkartal, Rapid discovery of graphene nanoflakes with desired absorption spectra using DFT and bayesian optimization with neural network kernel, J. Phys. Chem. A 129 (2025) 4591–4600.
[44] T. Zhai, W.T. Gu, M. Yu, Y.D. Shen, J.M. Liu, S. Wang, Machine learning-assisted multicolor fluorescence assay for visual data acquisition and intelligent inspection of multiple food hazards regardless of matrix interference, ACS Sens. 10 (2025) 4725–4732.
[45] N. Tuccitto, L. Fichera, R. Ruffino, V. Cantaro, G. Sfuncia, G. Nicotra, G. T. Sfrazzetto, G. Li-Destri, A. Valenti, A. Licciardello, A. Torrisi, Carbon quantum dots as fluorescence nanochemosensors for selective detection of amino acids, ACS Appl. Nano Mater. 4 (2021) 6250–6256.
[46] N. Ullah Khan, B. Prasad Sharma, S. Hussain Tumrani, M. Zahoor, R. Ali Soomro, T. Küçükdeniz, S. Karakuş, E. Ramadan Elsharkawy, J. Lu, S.M. El-Bahy, Z.M. El-Bahy, Enhanced detection of glucose with carbon quantum dot-modified copper oxide: computational insight and machine learning modeling of electrochemical sensing, Microchem. J. 204 (2024) 110936.
[47] M. Zhang, H. He, Y. Huang, R. Huang, Z. Wu, X. Liu, H. Deng, Machine learning integrated high quantum yield blue light carbon dots for real-time and on-site detection of Cr(VI) in groundwater and drinking water, Sci. Total Environ. 904 (2023) 166822.
[48] Q. Zhuang, W. Cao, Y. Ni, Y. Wang, Synthesis-identification integration: one-pot hydrothermal preparation of fluorescent nitrogen-doped carbon nanodots for differentiating nucleobases with the aid of multivariate chemometrics analysis, Talanta 185 (2018) 491–498.
[49] V. Dananjaya, S. Marimuthu, R. (Chunhui) Yang, A.N. Grace, C. Abeykoon, Synthesis, properties, applications, 3D printing and machine learning of graphene quantum dots in polymer nanocomposites, Prog. Mater. Sci. 144 (2024) 101282.
[50] M. El-Azazy, A.I. Osman, M. Nasr, Y. Ibrahim, N. Al-Hashimi, K. Al-Saad, M.A. Al-Ghouti, M.F. Shibl, A.H. Al-Muhtaseb, D.W. Rooney, A.S. El-Shafie, The interface of machine learning and carbon quantum dots: from coordinated innovative synthesis to practical application in water control and electrochemistry, Coord. Chem. Rev. 517 (2024) 215976.
[51] F. Yan, R. Bai, J. Huang, X. Bian, Y. Fu, Machine learning-assisted carbon dots synthesis and analysis: state of the art and future directions, TrAC Trends Anal. Chem. 184 (2025) 118141.
[52] M.J. Volk, I. Lourentzou, S. Mishra, L.T. Vo, C. Zhai, H. Zhao, Biosystems design by machine learning, ACS Synth. Biol. 9 (2020) 1514–1533.
[53] F. Li, Y. Yang, E. Xing, From lasso regression to feature vector machine, Adv Neural Inform Process Syst 18 (2005).
[54] B. Charbuty, A. Abdulazeez, Classification based on decision tree algorithm for machine learning, J. Appl. Sci. Technol. Trends 2 (2021) 20–28.
[55] H.A. Salman, A. Kalakech, A. Steiti, Random forest algorithm overview, Babylon. J. Mach. Learn. 2024 (2024) 69–79.
[56] E.Y. Boateng, J. Otoo, D.A. Abaye, Basic tenets of classification algorithms k-nearest-neighbor, support vector machine, random forest and neural network: a review, J. Data Anal. Inf. Process. 8 (2020) 341–357.
[57] D. Mustafa Abdullah, A. Mohsin Abdulazeez, Machine learning applications based on SVM classification a review, Qubahan Acad. J. 1 (2021) 81–90.
[58] A. Arista, Comparison decision tree and logistic regression machine learning classification algorithms to determine covid-19, Sinkron 7 (2022) 59–65.
[59] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., ACM, San Francisco California USA, 2016, pp. 785–794.
[60] M. Nour, K. Polat, Automatic classification of hypertension types based on personal features by machine learning algorithms, Math. Probl Eng. 2020 (2020) 2742781.
[61] S. Naeem, A. Ali, S. Anam, M.M. Ahmed, An unsupervised machine learning algorithms: comprehensive review, Int. J. Com. Dig. Sys. 13 (2023) 911–921.
[62] M. Sahu, R. Dash, A survey on deep learning: convolution neural network (CNN), in: D. Mishra, R. Buyya, P. Mohapatra, S. Patnaik (Eds.), Intell. Cloud Comput., Springer Singapore, Singapore, 2021, pp. 317–325.
[63] I.C. Peace, A.O. Uzoma, S.A. Ita, Effect of learning rate on artificial neural network in machine learning, Int. J. Eng. Res. 4 (2015) 359–363.
[64] L. Alzubaidi, J. Zhang, A.J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M.A. Fadhel, M. Al-Amidie, L. Farhan, Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, J. Big Data 8 (2021) 53.
[65] Q. Li, D. Cheng, H. Gu, D. Yang, Y. Li, S. Meng, Y. Zhao, Z. Tang, Y. Zhang, J. Tan, S. Qu, Aggregation-induced color fine-tunable carbon dot phosphorescence covering from green to near-infrared for advanced information encryption, Chem. Eng. J. 462 (2023) 142339.
[66] J. Bian, Z. Wang, J. Ge, W. Gong, K. Nan, Z. Liu, J. Guo, Q. Zhao, X. Wang, L. Wan, Recent advances of fluorescent carbon dots for biological imaging and biosensing, Chem. Eng. J. 519 (2025) 165135.
[67] R. Lawrance, P. Chowdhury, H.C. Lin, Y.H. Chan, The luminous frontier: transformative NIR-IIa fluorescent polymer dots for deep-tissue imaging, RSC Appl. Polym. 2 (2024) 749–774.
[68] P. Lesani, I. Zare, M. Khetarpaul, Z. Lu, A. Ghaee, H. Zreiqat, Red to near-infrared carbon dots: synthesis, cellular interactions, drug loading, and therapeutic applications, Adv. Drug Deliv. Rev. 228 (2026) 115742.
[69] X. Zhang, D. Ren, Y. Wei, H. Ma, Red light, green chemistry: on-site nitrite detection in pre-processed vegetables using diazotization-active carbon dots and smartphone-enabled dual-mode fluorometric/colorimetric readout, Anal. Chim. Acta 1377 (2025) 344643.
[70] J.B. Luo, J. Chen, H. Liu, C.Z. Huang, J. Zhou, High-efficiency synthesis of red carbon dots using machine learning, Chem. Commun. 58 (2022) 9014–9017.
[71] R.D. Senanayake, X. Yao, C.E. Froehlich, M.S. Cahill, T.R. Sheldon, M. McIntire, C.L. Haynes, R. Hernandez, Machine learning-assisted carbon dot synthesis: prediction of emission color and wavelength, J. Chem. Inf. Model. 62 (2022) 5918–5928.
[72] V.S. Tuchin, E.A. Stepanidenko, A.A. Vedernikova, S.A. Cherevkov, D. Li, L. Li, A. Döring, M. Otyepka, E.V. Ushakova, A.L. Rogach, Optical properties prediction for red and near-infrared emitting carbon dots using machine learning, Small 20 (2024) 2310402.
[73] C. Xing, G. Chen, X. Zhu, J. An, J. Bao, X. Wang, X. Zhou, X. Du, X. Xu, Synthesis of carbon dots with predictable photoluminescence by the aid of machine learning, Nano Res. 17 (2024) 1984–1989.
[74] H. Chen, B. Wu, S. Ji, L. Dang, H. Wei, Z. Wang, Nitrogen-doped carbon dots with high quantum yield for the detection of Fe^{3+} and antibacterial property, Mater. Today Commun. 47 (2025) 113021.
[75] Y. Han, B. Tang, L. Wang, Y. Lu, C. Guan, L. Zhang, M. Le, Z. Liu, M. Wu, Machine learning driven synthesis of carbon dots with enhanced quantum yields, ACS Nano 14 (2020) 14761–14768.
[76] H. Guo, Y. Lu, Z. Lei, H. Bao, M. Zhang, Z. Wang, C. Guan, B. Tang, Z. Liu, L. Wang, Machine learning-guided realization of full-color high-quantum-yield carbon quantum dots, Nat. Commun. 15 (2024) 4843.
[77] L. Chen, S. Yang, Y. Li, Z. Liu, H. Wang, Y. Zhang, K. Qi, G. Wang, P. He, G. Ding, Precursor symmetry triggered modulation of fluorescence quantum yield in graphene quantum dots, Adv. Funct. Mater. 34 (2024) 2401246.
[78] S.A. Dolenko, K.A. Laptinskiy, A.A. Korepanova, S.A. Burikov, T.A. Dolenko, Intelligent control of the synthesis of luminescent carbon dots with the desired photoluminescence quantum yield using machine learning, Opt. Mem. Neural Netw. 34 (2025) 18–29.
[79] R.E. Kannouma, A.M. Gab Allah, A.H. Kamal, M.A. Hammad, F.R. Mansour, Machine learning and genetic algorithm prediction of high quantum yield carbon quantum dots for chemical analysis, Microchem. J. 208 (2025) 112499.
[80] Y. Lan, G.S. Zheng, R.W. Song, J.N. Hao, J.L. Liu, C.L. Shen, J.Y. Zhu, S. Cao, J. L. Zhao, Q. Lou, C.X. Shan, Low-temperature molten-salt enabled synthesis of highly-efficient solid-state emitting carbon dots optimized using machine learning, Nat. Commun. 16 (2025) 8167.
[81] J. Tan, D. Yang, C. Lu, Y. Li, X. Gu, P. Sha, S. Qu, X. Wang, L. Fu, Q. Li, Achieving efficient phosphorescence of carbon dots via a synergistic enhancement strategy for bright field anti-counterfeiting, Chem. Eng. J. 512 (2025) 162102.
[82] Z. Yan, J. Li, S. Zhou, X. Yang, Machine learning-assisted controllable preparation of multicolor phosphorescent carbon dots, Carbon 241 (2025) 120394.
[83] Y. Li, L. Chen, S. Yang, G. Wei, X. Ren, A. Xu, H. Wang, P. He, H. Dong, G. Wang, C. Ye, G. Ding, Symmetry-triggered tunable phosphorescence lifetime of graphene quantum dots in a solid state, Adv. Mater. 36 (2024) 2313639.
[84] H. Yang, Z. Ran, Y. Luo, S. Liu, W. Xu, J. Liu, J. Cui, B. Lei, C. Hu, J. Zhuang, Y. Liu, Y. Xiao, Exploration and design of carbon dot-based long afterglow materials using active machine learning and quantum chemical simulations, ACS Nano 18 (2024) 29203–29213.
[85] R. Guo, S.Y. Song, Q. Cao, W.B. Zhao, X.Y. Li, H.F. Zhang, F.K. Li, Y.C. Liang, C. X. Shan, K.K. Liu, Machine learning-driven achieving efficient phosphorescent

carbon nanodots in aqueous solution by suppressing triplet electron leakage, Adv. Mater. 37 (2025) 2505925.
[86] J. Liao, X. Chen, J. Li, X. Gong, Engineering strategy for achieving transparent aqueous multicolor long lifetime room temperature phosphorescent carbon dots, Nano Energy 146 (2025) 111532.
[87] M. Yahaya Pudza, Z. Zainal Abidin, S. Abdul Rashid, F. Md Yasin, A.S.M. Noor, M. A. Issa, Sustainable synthesis processes for carbon dots through response surface methodology and artificial neural network, Processes 7 (2019) 704.
[88] Q. Zhang, Y. Tao, B. Tang, J. Yang, H. Liang, B. Wang, J. Wang, X. Jiang, L. Ji, S. Li, Graphene quantum dots with improved fluorescence activity via machine learning: implications for fluorescence monitoring, ACS Appl. Nano Mater. 5 (2022) 2728–2737.
[89] H. He, S. E, L. Ai, X. Wang, J. Yao, C. He, B. Cheng, Exploiting machine learning for controlled synthesis of carbon dots-based corrosion inhibitors, J. Clean. Prod. 419 (2023) 138210.
[90] A. Dager, T. Uchida, T. Maekawa, M. Tachibana, Synthesis and characterization of mono-disperse carbon quantum dots from fennel seeds: photoluminescence analysis using machine learning, Sci. Rep. 9 (2019) 14004.
[91] J. Thonghlueng, S. Ngernpimai, A. Chuaephon, W. Phanchai, T. Wiwasuku, Y. Wanna, K. Wiratchawa, T. Intharah, R. Thanan, C. Sakonsinsiri, T. Puangmali, Dual-responsive carbon quantum dots for the simultaneous detection of cytosine and 5-methylcytosine interpreted by a machine learning-assisted smartphone, ACS Appl. Mater. Interfaces 15 (2023) 40141–40152.
[92] M. Salahinejad, S. Sadjadi, M. Abdouss, Investigating fluorescence quenching of cysteine-functionalized carbon quantum dots by heavy metal ions: experimental and QSPR studies, J. Mol. Liq. 334 (2021) 116067.
[93] X. Li, H. Luo, D. Zhai, Y. Wan, G. He, D. Hu, H. Hou, D. Zhang, S. Zhang, Enhanced capacitive energy storage of polyetherimide at high temperatures by integration of electrical insulation and thermal conductivity, Adv. Powder Mater. 4 (2025) 100286.
[94] Q. Fu, J. Zhang, K. Zhang, S. Sun, Z. Dong, Multicolor solid-state fluorescence and multicolor afterglow carbon dots: preparation, luminescence regulation, and applications, J. Mater. Chem. A 13 (2025) 36–72.
[95] S. Pandit, T. Banerjee, I. Srivastava, S. Nie, D. Pan, Machine learning-assisted array-based biomolecular sensing using surface-functionalized carbon dots, ACS Sens. 4 (2019) 2730–2737.
[96] Z. Chen, Y. Liu, Z. Kang, Diversity and tailorability of photoelectrochemical properties of carbon dots, Acc. Chem. Res. 55 (2022) 3110–3124.
[97] A. Döring, A.L. Rogach, Utilizing deep learning to enhance optical sensing of ethanol content based on luminescent carbon dots, ACS Appl. Nano Mater. 5 (2022) 11208–11218.
[98] X. He, Y. Zheng, C. Hu, B. Lei, X. Zhang, Y. Liu, J. Zhuang, The afterglow of carbon dots shining in inorganic matrices, Mater. Horiz. 11 (2024) 113–133.
[99] X. Wang, H. Wang, W. Zhou, T. Zhang, H. Huang, Y. Song, Y. Li, Y. Liu, Z. Kang, Carbon dots with tunable third-order nonlinear coefficient instructed by machine learning, J. Photochem. Photobiol. Chem. 426 (2022) 113729.
[100] H. Zhang, L. Sun, X. Guo, J. Xu, X. Zhao, Y. Xia, Multicolor fluorescent/room temperature phosphorescent carbon dot composites for information encryption and anti-counterfeiting, Appl. Surf. Sci. 613 (2023) 155945.
[101] J. Hu, M. Lei, L. Yan, L. Chen, Y. Yang, J. Zheng, X. Liu, B. Xu, Carbon dot-based afterglow composite with tunable room temperature phosphorescence and thermally activated delayed fluorescence and their anti-counterfeiting and encryption application, Chem. Eng. J. 489 (2024) 151245.
[102] M. Aggarwal, P. Sahoo, S. Saha, P. Das, Machine learning-mediated ultrasensitive detection of citrinin and associated mycotoxins in real food samples discerned from a photoluminescent carbon dot barcode array, J. Agric. Food Chem. 71 (2023) 12849–12858.
[103] A. Bocan, R.S. Moakhar, C. del R. Mata, M. Petkun, T.D. Iure-Grimmel, S. G. Yedire, H. Shieh, A.K. Jahromi, S.S. Mahshid, S. Mahshid, Machine-learning-aided advanced electrochemical biosensors, Adv. Mater. 37 (2025) 2417520.
[104] M. Li, M. Han, X. Song, Y. Chen, Z. Wu, L. Guan, F. Wang, D. Wang, L. Lv, X. Li, A design strategy of ratiometric optical probe for Fe^{3+} ions and ascorbic acid detection based on europium-doped carbon dots, J. Rare Earths (2025) S1002072125003837.
[105] Z. He, Z. Wang, L. Ga, J. Ai, P. Liu, Sulfur-nitrogen doped carbon dots for Bi^{3+} ion detection and temperature sensing, Talanta Open 12 (2025) 100482.
[106] B.M. Vijaybhai, A. Singh, G. Singh, A.K. Tangra, M.S. Sankhla, Advances in carbon quantum dot-based nanocomposites for targeted detection of food toxins: a focus on safety applications, J. Food Compos. Anal. 148 (2025) 108321.
[107] Z. Riahi, A. Khan, J.W. Rhim, G.H. Shin, J.T. Kim, Carbon dot-based pH-responsive indicators for intelligent food packaging and food safety control, Trends Food Sci. Technol. 163 (2025) 105200.
[108] C. Tian, Y. Lee, Y. Song, M.R. Elmasry, M. Yoon, D.H. Kim, S.Y. Cho, Machine-learning-enhanced fluorescent nanosensor based on carbon quantum dots for heavy metal detection, ACS Appl. Nano Mater. 7 (2024) 5576–5586.
[109] R. Chen, J. You, Y. Chen, Z. Zhang, A. Sun, H. Liu, X. Shi, Broad application prospects near-infrared carbonized polymer dots combined with machine learning for the detection of Cu^{2+} in seawater and aquatic products, Sens. Actuators B Chem. 404 (2024) 135245.
[110] N. Zheng, M. Zhang, B. Liu, J. Zhang, F. Zhang, Z. Ma, Z. Nie, J. Tang, G. Chen, A machine learning approach to prediction of the highly sensitive and selective model for specific detection of Fe^{3+} ions using nitrogen-doped carbon dots as fluorescent probe, J. Photochem. Photobiol. Chem. 472 (2026) 116856.
[111] M. Llaver, S.D. Barrionuevo, J.M. Núñez, A.L. Chapana, R.G. Wuilloud, M. H. Aguirre, F.J. Ibañez, Fluorescent graphene quantum dots-enhanced machine learning for the accurate detection and quantification of Hg^{2+} and Fe^{3+} in real water samples, Environ. Sci. Nano 11 (2024) 2703–2715.
[112] H. Li, H. Yue, H. Li, M. Zhu, X. He, M. Liu, X. Li, F. Qiu, High-throughput point-of-care serum iron testing utilizing machine learning-assisted deep eutectic solvent fluorescence detection platform, J. Colloid Interface Sci. 680 (2025) 389–404.
[113] R. Das, A. Paria, P.K. Giri, Machine learning framework for selective and sensitive metal ion sensing with nitrogen-doped graphene quantum dots heterostructure, Carbon 232 (2025) 119800.
[114] L. Kang, J. Wang, X. Lin, J. Feng, N. Duan, Z. Wang, S. Wu, A full-color carbon quantum dots fluorescence sensing array combined with machine learning for rapid bacterial detection and classification, Adv. Healthc. Mater. 15 (2025) e02916.
[115] S. Thakur, A. Ojha, S.K. Kansal, N.K. Gupta, H.C. Swart, J. Cho, A. Kuznetsov, S. Sun, J. Prakash, Advances in powder nano-photocatalysts as pollutant removal and as emerging contaminants in water: analysis of pros and cons on health and environment, Adv. Powder Mater. 3 (2024) 100233.
[116] Y. Qin, F. Zhang, R. Tang, C. Yuan, C. Cui, C. Yan, T.D. James, L. Wang, M. Li, W. Zhu, Boosting gaseous mercury detection via photooxidation-enrichment fluorescent membrane with machine learning, Small 22 (2025) e13585.
[117] X. Wang, J. Huo, H. Gu, H. Jing, Y. Rao, Z. Zhao, X. Li, Y. Liu, H. Ma, Machine learning-assisted construction of a lignin carbon dots sensor array for detecting food colorants, Food Chem. 496 (2025) 146948.
[118] Z. Cai, C. Yang, Z. Zhang, W. Xing, S. Yang, Ratiometric fluorescence sensor based on carbon quantum dots for hydrocortisone detection in food samples, Food Chem. 495 (2025) 146636.
[119] Z. Lu, M. Chen, T. Liu, C. Wu, M. Sun, G. Su, X. Wang, Y. Wang, H. Yin, X. Zhou, J. Ye, Y. Shen, H. Rao, Machine learning system to monitor Hg^{2+} and sulfide using a polychromatic fluorescence-colorimetric paper sensor, ACS Appl. Mater. Interfaces 15 (2023) 9800–9812.
[120] A. Döring, Y. Qiu, A.L. Rogach, Improving the accuracy of carbon dot temperature sensing using multi-dimensional machine learning, ACS Appl. Nano Mater. 7 (2024) 2258–2269.
[121] X. Zhang, W. Zhu, L. Mei, S. Zhang, J. Liu, F. Wang, Machine learning-enhanced bacteria detection using a fluorescent sensor array with functionalized graphene quantum dots, ACS Appl. Mater. Interfaces 17 (2025) 3084–3096.
[122] H. Guo, P. Lesani, H. Zreiqat, E.J. New, A fluorescent sensor array based on carbon dots for the accurate determination of pH, Sens. Diagn 3 (2024) 1923–1934.
[123] J. Li, S. Wu, X. Shi, Y. Cao, H. Hao, J. Wang, Q. Han, Machine learning-assisted biomass-derived carbon dots as fluorescent sensor array for discrimination of warfarin and its metabolites, Langmuir 41 (2025) 1694–1702.
[124] Y. Zhang, Q. Ma, S. Gao, X. Liu, H. Xing, H. Hu, L. Wang, W. Li, T. Zhang, Y. Hou, D. Chen, Ratiometric determination and discrimination of oxicams via dual-excitation carbon dots assisted by machine learning, Anal. Chem. 97 (2025) 21428–21437.
[125] Z. Zhang, Y. Yuan, Y. Tian, Y. Zhao, T. Chen, J. Gao, L. Chen, Machine learning-assisted carbon quantum dots-molecularly imprinted polymer array for synchronous screening of multiple sulfonamides, Sens. Actuators B Chem. 444 (2025) 138374.
[126] Y. Tang, P. Zhu, Q. Xu, J. Wang, Machine learning assists the sensor array constructed by the tri-emission carbon dots to detect multiple metal ions, Microchem. J. 201 (2024) 110536.
[127] F. Wang, M. Xiao, J. Qi, L. Zhu, Paper-based fluorescence sensor array with functionalized carbon quantum dots for bacterial discrimination using a machine learning algorithm, Anal. Bioanal. Chem. 416 (2024) 3139–3148.
[128] M. Xiao, L. Mei, J. Qi, L. Zhu, F. Wang, Functionalized carbon quantum dots fluorescent sensor array assisted by a machine learning algorithm for rapid foodborne pathogens identification, Microchem. J. 201 (2024) 110701.
[129] L. Zhu, L. Mei, Y. Xuan, F. Wang, Machine learning assisted paper-based fluorescent sensor array with metal-doped multicolor carbon quantum dots for identification and inactivation of bacteria, Talanta 293 (2025) 128035.
[130] S. Kalytchuk, T. Malina, F. Mravec, K. Poláková, L. Zdražil, Š. Kment, A.L. Rogach, M. Otyepka, R. Zbořil, pH sensing and imaging in living cells based on fluorescence lifetime of carbon dot nanosensors, Biosens. Bioelectron. 291 (2026) 118022.
[131] C. Wu, H. Chang, X. Chen, S. Yang, Y. Dai, P. Tan, Y. Chen, C. Shen, Z. Lu, M. Sun, G. Su, Y. Wang, Y. Zou, H.M.D. Wang, H. Rao, T. Liu, Deep learning-assisted rapid assessment of food freshness using an anti-interfering triple-emission ratiometric fluorescent sensor, ACS Sustain. Chem. Eng. 12 (2024) 2465–2475.
[132] Z. Haoge, C. Shuo, W. Snehal, K. Kimleng, J.S. Cross, Microalgae-derived carbon quantum dots enhance photosynthesis and accumulation of energy storage compounds in microalgal communities, J. Environ. Chem. Eng. 13 (2025) 119751.
[133] B. Zhang, W. Wang, L. Zhu, N. Li, X. Chen, J. Tian, X. Zhang, Simultaneously enhanced permeability and anti-fouling performance of polyethersulfone ultrafiltration membranes by structural control and mixed carbon quantum dots, J. Membr. Sci. 641 (2022) 119931.

[134] X. Wang, S. Chen, Y. Ma, T. Zhang, Y. Zhao, T. He, H. Huang, S. Zhang, J. Rong, C. Shi, K. Tang, Y. Liu, Z. Kang, Continuous homogeneous catalytic oxidation of C–H bonds by metal-free carbon dots with a poly (ascorbic acid) structure, ACS Appl. Mater. Interfaces 14 (2022) 26682–26689.
[135] Y. Liu, X. Wang, Y. Zhao, Q. Wu, H. Nie, H. Si, H. Huang, Y. Liu, M. Shao, Z. Kang, Highly efficient metal-free catalyst from cellulose for hydrogen peroxide photoproduction instructed by machine learning and transient photovoltage technology, Nano Res. 15 (2022) 4000–4007.
[136] M. Samir, Z. Ismail, A.M. Agour, H. Nageh, S.O. Abdellatif, Exploring forward scattering mechanisms in TiO_2 with carbon quantum dots: insights into photovoltaic applications, Opt. Mater. 148 (2024) 114922.
[137] L. Wang, T. Yang, M. Wei, R. Guan, W. Wei, J. Jiang, Machine learning and DFT dual-guidance of carbon dots implanted $SrTiO_3$ hollow nanosphere for efficient all-pH-value photocatalysis, J. Mater. Sci. Technol. 217 (2025) 169–181.
[138] C. Ebere Enyoh, Q. Wang, Box–behnken design and machine learning optimization of PET fluorescent carbon quantum dots for removing fluoxetine and ciprofloxacin with molecular dynamics and docking studies as potential antidepressant and antibiotic, Sep. Purif. Technol. 362 (2025) 131975.
[139] H. Xu, J. Li, L. Li, Z. Li, J. Wang, F. Liao, H. Huang, Y. Liu, Z. Kang, Unveiling charge transfer mechanisms in machine learning-optimized carbon dots-chitosan catalysts for H2O2 photogeneration, Mater. Res. Bull. 195 (2026) 113877.
[140] A.K. Astakala, S.Y. Lee, J. Gautam, K.B. Thapa, I. In, S.J. Lee, S.J. Park, Engineering inorganic perovskite solar cells: overcoming efficiency and stability barriers for next-generation photovoltaics, Adv. Powder Mater. 5 (2026) 100354.
[141] B. Fang, Z. Xing, D. Sun, Z. Li, W. Zhou, Hollow semiconductor photocatalysts for solar energy conversion, Adv. Powder Mater. 1 (2022) 100021.
[142] S. Li, M. Cai, Y. Liu, C. Wang, R. Yan, X. Chen, Constructing $Cd_{0.5}Zn_{0.5}S/Bi_2WO_6$ S-scheme heterojunction for boosted photocatalytic antibiotic oxidation and Cr (VI) reduction, Adv. Powder Mater. 2 (2023) 100073.
[143] Y. Li, Y. Liang, X. He, Y. Yin, X. Sun, C. Li, B. Yang, J. Wang, Fluorescence-based sensor array using nitrogen-doped carbon dots: machine learning-assisted discrimination of dihydroxybenzene isomers and their ternary blends, Microchem. J. 220 (2026) 116498.
[144] M. Hussaini, M.S. Vohra, S.A. Onaizi, Investigating the removal of Congo red dye using ZIF-8 and GQD composite: characterization, kinetics, isotherm, thermodynamics, optimization, and machine learning studies, Colloids Surf. A Physicochem. Eng. Asp. 726 (2025) 137915.
[145] A. Priyanto, D.A. Hapidin, D. Maulina, M.P. Aji, K. Khairurrijal, Novel simple accurate detection of microplastics based on image of photoluminescent nanoparticle carbon dots via machine learning and deep feature embedding, J. Environ. Manage. 398 (2026) 128545.
[146] A.K. Singh, P. Itkor, Z.A.N. Hanani, M. Lee, K.K. Gaikwad, Y.S. Lee, Facile and green synthesis of multifunctional nitrogen-doped carbon quantum dots from ginkgo biloba leaves for antibacterial, antioxidant, and bioactive coating applications, food packag, Shelf Life 51 (2025) 101581.
[147] S. Rodríguez-Varillas, J. Espina-Casado, R. Badía Laíño, A. Fernández-González, T. Fontanil López, Carbon dots in oncology: multifunctional nanoplatforms for diagnosis, targeted therapy, and drug discovery, Drug Discov. Today 30 (2025) 104470.
[148] A.S. Rasal, T.M. Subrahmanya, S. Kizhepat, G. Getachew, A.V. Ghule, R.S. Devan, W.S. Hung, M.Z. Fahmi, A. Wibrianto, J.Y. Chang, Carbon quantum dots: classification-structure-property-application relationship for biomedical and environment remediation, Coord. Chem. Rev. 533 (2025) 216510.
[149] J. Li, Q. Li, Q. Yang, Q. Tang, X. Hu, Q. Liu, L. Zhang, A highly permeable nanoplatform based on functionalized carbon dots for synergistic reactive oxygen/nitrogen species tumor therapy, Mater. Today Nano 28 (2024) 100532.
[150] J. Xu, W. Ji, K. Zhu, M. Liu, H. Gao, M. Yan, Q. Zeng, Application of fluorescence resonance energy transfer in homogeneous immunoassay using quantum dots, Coord. Chem. Rev. 542 (2025) 216872.
[151] Y. Guo, Y. Liu, J. Yang, Y. Wu, H. Lian, X. Tang, C. Li, W. Cui, Z. Guo, Carbon dots meet artificial intelligence: applications in biomedical engineering, J. Mater. Chem. B 13 (2025) 13905–13927.
[152] G. Saren, L. Zhu, Y. Han, Quantitative detection of gastrointestinal tumor markers using a machine learning algorithm and multicolor quantum dot biosensor, Comput. Intell. Neurosci. 2022 (2022) 9022821.
[153] Y. Xiao, B. Peng, X. Zhu, Y. Long, F. Liu, M. Liu, H. Li, Y. Zhang, S. Yao, Machine learning-assisted multichannel nanosensor based on core-shell carbon dot/gold nanoparticle composites for rapid and precise intraoperative oral cancer malignancy diagnosis, Sci. China Chem. 68 (2025) 5273–5283.
[154] S. Banerjee, S. Mandal, S. Singh, A. Sen, P. Das, Machine learning-guided fabrication of carbon dot–pepsin nano-conjugates for enhanced bioimaging, synergistic drug delivery, and visible light-induced photosensitization, Nanoscale 17 (2025) 14233–14247.
[155] X. Yin, S. Yang, G. Du, L. Zhang, Y. Xiang, N. Ye, Exploration of the fluorine−fluorine interaction mechanism in fluoroquinolone antibiotics recognition and ciprofloxacin detection on the basis of fluorine−doped carbon quantum dots and machine learning, Food Chem. 492 (2025) 145560.
[156] Q. Hong, X.Y. Wang, Y.T. Gao, J. Lv, B.B. Chen, D.W. Li, R.C. Qian, Customized carbon dots with predictable optical properties synthesized at room temperature guided by machine learning, Chem. Mater. 34 (2022) 998–1009.
[157] Y. Tang, Q. Xu, X. Zhang, R. Zhu, N. Zhao, J. Wang, Expediting carbon dots synthesis by the active adaptive method with machine learning and applications in dental diagnosis and treatment, Nano Res. 17 (2024) 10109–10118.
[158] M. Yu, R. Ye, T. Zeng, L. Tan, Z. Zhao, W. Gao, X. Chen, Z. Lian, Y. Ma, A. Li, J. Hu, Constructing an ultra-rapid nanoconfinement-enhanced fluorescence clinical detection platform by using machine learning and tunable DNA xerogel "probe,", Anal. Chem. 95 (2023) 15690–15699.
[159] C. Wu, P. Tan, X. Chen, H. Chang, Y. Chen, G. Su, T. Liu, Z. Lu, M. Sun, Y. Wang, Y. Zou, J. Wang, H. Rao, Machine learning-assisted high-throughput strategy for real-time detection of spermine using a triple-emission ratiometric probe, ACS Appl. Mater. Interfaces 15 (2023) 48506–48518.
[160] X. Wang, L. Wang, M. Wu, Y. Zheng, R. Wang, T. Shao, S. Zeng, A. Li, R. Li, Q. Yue, Portable and intelligent ratio fluorometry and colorimetry for dual-mode detection of dopamine based on B, N-codoped carbon dots and machine learning, Talanta 294 (2025) 128288.
[161] S. Ahmad, H.A. Alsalmah, Thermal optimization of nickel-carbon nanodots using neural network and numerical simulation, Case Stud. Therm. Eng. 74 (2025) 106768.
[162] C.H. Wang, Y.Y. Tang, Y.J. Li, F.L. Xiao, L.D. Zhou, Q.H. Zhang, C.S. Yuan, Synergistic highly sensitive electrochemical detection of enrofloxacin using eichhornia crassipes-derived carbon dots and molecularly imprinted polymer networks, J. Environ. Chem. Eng. 13 (2025) 118648.
[163] Y. Wang, S. Wang, Z. Zhang, Y. Ning, W. Wang, C. Liang, Y. Jia, Rapid visual discrimination and quantification of multiple antibiotics based on machine learning-assisted fluorescence sensor array, Food Control 178 (2025) 111521.
[164] F. Nurjis, R. Ali, H. Ali, Harnessing maillard reaction byproducts for dual emissive carbon quantum dots: a tunable optical platform, RSC Adv. 15 (2025) 30711–30726.
[165] S. Rahupathy, M. Sivanandhan, A. Parasuraman, Designing pH-functionalized carbon dots as white light emitting phosphor for cool white LED application, Carbon 245 (2025) 120826.
[166] X. Wang, B. Wang, H. Wang, T. Zhang, H. Qi, Z. Wu, Y. Ma, H. Huang, M. Shao, Y. Liu, Y. Li, Z. Kang, Carbon-dot-based white-light-emitting diodes with adjustable correlated color temperature guided by machine learning, Angew. Chem. 133 (2021) 12693–12698.
[167] J. Cao, R. Chen, L. Wang, H. Xing, H. Hu, X. Yang, C. Gu, S. Tang, D. Chen, Meticulously designed red emission carbon dots as efficient and stable luminescent species for rapid latent fingerprint recognition and light emitting diodes, Chem. Eng. J. 491 (2024) 152121.
[168] H. Zhang, J. Hu, J. Chen, J. Pan, J. Wang, X. Wang, D. Chen, Machine learning-assisted N-doped carbon dots for sensitive Pd^{2+} detection and high-performance LED applications in plant cultivation, Microchim. Acta 192 (2025) 731.
[169] B. Ersöz, A. Öter, Z. Berktaş, H.İ. Bülbül, A. Di Bartolomeo, Ş. Sağıroğlu, E. Orhan, Artificial intelligence-driven data generation for temperature-dependent current–voltage characteristics of diodes, FlatChem 51 (2025) 100847.
[170] J. Chen, J.B. Luo, M.Y. Hu, J. Zhou, C.Z. Huang, H. Liu, Controlled synthesis of multicolor carbon dots assisted by machine learning, Adv. Funct. Mater. 33 (2023) 2210095.
[171] D.A.M. Muyassiroh, F.A. Permatasari, T. Hirano, T. Ogi, F. Iskandar, Machine learning-guided synthesis of room-temperature phosphorescent carbon dots for enhanced phosphorescence lifetime and information encryption, ACS Appl. Nano Mater. 7 (2024) 5465–5475.
[172] W. Wang, K. Song, S. Zou, Q. Xia, X. Li, K. Sun, H. Ma, Long-lifetime and multicolor room-temperature phosphorescence of carbon quantum dots derived from biomass: applications on convolutional neural network, ACS Appl. Nano Mater. 8 (2025) 10641–10650.
[173] X. Wu, Y. Chen, S. Fu, L. Xu, L. Ye, Multiple atoms doped carbon dots with room temperature phosphorescence for anticounterfeiting and information encryption, ACS Appl. Nano Mater. 7 (2024) 28348–28357.
[174] Z. Bian, T. Bao, X. Sun, N. Wang, Q. Mu, T. Jiang, Z. Yu, J. Ding, T. Wang, Q. Zhou, Machine learning tools to assist the synthesis of antibacterial carbon dots, Int. J. Nanomedicine 19 (2024) 5213–5226.
[175] M. Momina, Q. Qurtulen, H. Salimi Shahraki, A. Ahmad, Z. Zaheer, Machine learning approaches to predict adsorption performance of sugarcane derived-carbon dot −based composite in the removal of dyes, Sep. Purif. Technol. 351 (2024) 127937.
[176] M.M. Solomon, U.O. Amune, X. He, H. Wang, J. He, D. Hu, F.B. Tesema, H. Do, A. A. Mansour, R. Salghi, S.A. Umoren, Combined experimental, theoretical, & machine learning studies of anticorrosion properties of hydroxyethyl cellulose, L-glutamic acid, & potassium acesulfame-derived carbon dots, Carbon 245 (2025) 120772.

Abbreviation Glossary

AE: Autoencoder
AI: Artificial Intelligence
ANN: Artificial Neural Network
BO: Bayesian Optimization
CDs: Carbon Dots
CNN: Convolutional Neural Network
DFT: Density Functional Theory

DNDF: Deep Neural Decision Forest
DT: Decision Tree
ELM: Extreme Learning Machine
EN: Elastic Net
ET: Extra Trees
FCNN: Fully Convolutional Neural Network
GA-KELM: Genetic Algorithm Kernel Extreme Learning Machine
GBM: Gradient Boosting Machine
GBDT: Gradient Boosting Decision Tree
GDST: Gradient Descent with Stochastic Term
GPR: Gaussian Process Regression
HCA: Hierarchical Clustering Analysis
KNN: K-Nearest Neighbor
LDA: Linear Discriminant Analysis
LSSVM: Least Squares Support Vector Machine
LSTM: Long Short-Term Memory
ML: Machine Learning
MP: McCulloch-Pitts Neuron
NDA: Normal Discriminant Analysis
PLS-DA: Partial Least Squares Discriminant Analysis
RNN: Recurrent Neural Network
RR: Ridge Regression
SVM: Support Vector Machine
TR: Transformer Reinforcement Learning

Shen Shudie (1998-) is a doctoral candidate whose primary research focus is on machine learning-assisted synthesis of carbon dots and the luminescence mechanisms. Email: shenshudie@mai.tust.edu.cn.

Zhu Liang (1980-) is a Professor and PhD supervisor, mainly engaged in the research of development and mechanism analysis of industrial crystallization and nanomaterials. Corresponding author. E-mail: zhuliang@tust.edu.cn.

Zhao Xiaoyu (1986-) is a Professor and PhD supervisor working on electrochemical processes and crystallization mechanisms in functional nanomaterials. Corresponding author. xyz@tust.edu.cn.