Progress and Challenges of Reverse Design of AI-enabled Functional Materials

Haoyan Li *

College of Materials and Chemical Engineering, South China Agricultural University, Guangzhou, China

* Corresponding Author Email: 1968195440@qq.com

Abstract. Against the backdrop of global energy structure transformation and high-end manufacturing industry upgrading, the development of new functional materials has become the key to scientific and technological competition. This paper systematically reviews how artificial intelligence (AI) drives the profound change of the functional material research and development paradigm. Through in-depth analysis of the atomic component evolution of AI in energy catalysis, the geometric topology construction in the field of optoelectronic and electromagnetic, and the multi-scale optimization in structural mechanical materials, this paper shows the great potential of deep coupling between generative models and physical laws. At the same time, this paper also faces technical bottlenecks, such as the lack of experimental data, the poor synthesis of design structure, and the lack of transparency of the model. The research conclusion points out that with the intervention of cutting-edge technologies such as physical information assistance, automatic laboratory and large language models, the material research and development will move towards the automatic closed-loop direction of deep integration of data, experiment and theory. This review not only reveals the theoretical value of AI in complex performance optimization but also has important academic and practical significance in breaking the bottleneck of the material development cycle and accelerating the large-scale customization of high-performance materials.

Keywords: Functional Materials, Inverse Design, Artificial Intelligence, Generative Models.

1. Introduction

Against the backdrop of current energy transformation and rapid change of semiconductor technology, the innovation of functional materials has become the core driving force to promote the development of the industry. Whether it is to improve the efficiency of catalytic reaction, optimize the performance of photoelectric chips, or construct metamaterials with special physical properties, it puts forward high requirements for the composition design and structure control of materials. These materials are not only the basis of clean energy technology, but also the physical carrier to achieve a new generation of information technology breakthroughs. Their R&D level is directly related to the competitiveness of high-end manufacturing.

In the past few decades, the discovery of functional materials mainly depended on the intuitive experience of researchers and the experimental exploration based on the "trial and error method". Although in recent years, high-throughput screening combined with density functional theory (DFT) and other calculation methods has significantly improved the speed of discovery of new materials, the forward design mode of "modeling before prediction" still has obvious limitations. In the face of almost unlimited chemical combination space, the traditional screening methods are often like "looking for a needle in a haystack", which not only consumes huge computational resources, but also is difficult to directly give the optimal structural scheme under complex performance requirements. This lag in the research and development paradigm has become the main bottleneck restricting the large-scale application of new functional materials [1-3].

In order to meet the above challenges, artificial intelligence (AI) - enabled reverse design is emerging as a new research and development paradigm. Unlike traditional methods, reverse design emphasizes "starting from the end", that is, directly generating the optimal chemical composition or geometric topology according to the preset target performance [4]. This paper aims to systematically review how AI technology promotes functional materials from simple "attribute prediction" to deep

"structure automatic generation". From a theoretical point of view, this method can mine the complex nonlinear mapping behind the data and deepen human understanding of the physical nature of materials. From a practical point of view, it greatly reduces the R&D cycle and provides an efficient and accurate solution for the customized design of high-performance materials.

2. AI-oriented design progress of typical functional materials

2.1. Energy and Catalysis

In the research of energy conversion and storage, the core goal of reverse design is to break through the limitation of chemical reaction kinetics by accurately regulating the electronic structure of the material surface. For the hydrogen evolution and oxygen evolution reaction (HER/OER), researchers used a graph neural network (GNN) to capture the deep connection between crystal topology and intermediate adsorption energy. Zheng et al. Proposed the ASB-GCNN model based on CGCNN to solve the problem of difficult prediction of complex systems, and decomposed the material structure into active layer, surface layer, and bulk layer. Through the combination of classification model and shap value analysis, 600 kinds of Ma α Z ₄-based monatomic catalysts were screened. It was found that the hydrogen adsorption free energy of $V_1/HfSn_2N_4(S)$was as low as 0.06 EV, which was close to the performance of platinum-based catalysts [5]. Traditional experimental methods can only verify known materials, while generative models (such as VAE) show a powerful "from nothing" design ability: they can directly conduct reverse search in a wide space of chemical components and accurately locate the atomic ratio at the peak of activity in the "volcanic map". In addition, in the research and development of new battery electrolytes, AI combines molecular fingerprint coding with active learning to quickly identify target molecules with high ionic conductivity and a wide electrochemical window in a massive organic molecular library. In early 2024, Microsoft and Pacific Northwest National Laboratory used Azure Quantum Elements to apply an AI model to about 32million inorganic chemical formulas and screened out about 500000 candidate materials for predicting thermodynamic stability. About 18 candidate materials were selected for the synthesis test, and finally a solid electrolyte containing lithium ion and sodium ion was found, which reduced the lithium content by about 70% while maintaining the ionic conductivity [6]. This method has changed the previous "blind screening" mode and significantly shortened the R&D cycle from theoretical prediction to laboratory verification.

2.2. Optoelectronic and Electromagnetic Fields

The design logic of optoelectronic and electromagnetic materials is undergoing a paradigm shift from a single "band gap adjustment" to a complex "space morphology Engineering". In the design of photovoltaic materials, reverse engineering mainly focuses on the precise control of energy band structure. By learning the intrinsic physical relationship between lattice distortion and band evolution, an AI model can deduce the optimal doping concentration and lattice parameters according to the preset spectral absorption characteristics. Alhashmi, A., By combining first principles calculation and machine learning algorithm, they learned the internal relationship between lattice distortion (structural change caused by doping) and band evolution through AI model, accurately identified the optimal doping strategy (CA doped Ba site or Ti doped Zr site) of $BaZrS_3$ perovskite, successfully reduced the band gap from~1.75 EV to~1.26 EV, and achieved the reverse design of the optimal doping concentration and lattice parameters based on the target spectral absorption characteristics (Shockley Queisser limit~1.3 EV) [7].

In the research and development of metasurfaces, the design dimension is further upgraded to the topology optimization of sub wavelength scale. With the help of the generation of confrontation network (GAN) or diffusion model, designers only need to define the required optical phase or amplitude response, and AI can automatically generate non-intuitive and asymmetric nano antenna structures. This "design on demand" method completely breaks the dependence of human experience

on conventional geometry (such as cylinder or block) and realizes the unprecedented precise control of electromagnetic wave [8].

2.3. Structure and Mechanics

The challenge of structural material design for a long time has been how to balance the contradictory variables of strength and ductility, and AI provides a new way to solve this problem. In the development of high entropy alloys (HEAS), in the face of the "combinatorial explosion" problem caused by multi-element superposition, AI introduces physical information neural network (pinn) and thermodynamic phase diagram constraints, which can accurately predict the phase stability at different temperatures. This allows researchers to reverse eliminate the areas prone to brittle phase and select the best element ratio that can maintain the single-phase solid solution. In the field of lightweight and high-strength composites, the design space of AI extends from micro components to macro topology. Through a reinforcement learning algorithm, AI can simulate the non-uniform growth mechanism of biological bones and reverse construct the lattice structure with high energy absorption efficiency. This multi-scale optimization design ensures that the material can still show excellent mechanical bearing capacity at very low density [9][10].

2.4. Summary.

Table 1. Summary and Comparison.

Field	Core Inverse Design Targets	Mainstream AI Models	Key Physical Constraints
Energy Catalysis	Adsorption energy, redox potential	GNN, VAE, Active Learning	Electron cloud density, coordination number
Optoelectronics & Electromagnetics	Absorption/reflection spectrum, band gap	GAN, Diffusion Models, CNN	Maxwell's equations, periodic boundary conditions
Structural Mechanics	Specific strength, fracture toughness, phase stability	Regression Trees, Reinforcement Learning, PINN	Crystal symmetry, thermodynamic phase diagram

Through the systematic review of various research directions (see Table 1), researchers can find that the technical path and physical constraints of AI in the design of enabling functional materials show obvious characteristics of domain differentiation.

In the field of energy and catalysis, the focus of reverse design mainly focuses on the regulation of active sites at the atomic scale. The models (such as GNN and VAE) need to deal with the highly sensitive electron cloud density and coordination environment and optimize the reaction kinetics by accurately adjusting the adsorption energy. In contrast, the design logic of optoelectronic and electromagnetic materials is more embodied in the construction of spatial geometric topology. Using Gan or diffusion model, researchers can achieve customized control of spectrum and wavefront characteristics by changing the morphology of sub wavelength scale under the physical framework of Maxwell equations.

In addition, for structural and mechanical materials, the task of AI evolved to find the optimal balance of performance in the multidimensional parameter space. With the help of pinn and reinforcement learning, the design process should not only meet the constraints of crystal symmetry and thermodynamic phase diagram, but also co optimize multiple mutually exclusive indicators such as specific strength and ductility, so as to break the long-standing "performance trade-off" dilemma in materials science. This comparison shows that for materials with different physical mechanisms, the key premise to achieve efficient reverse design is to choose the algorithm architecture and representation strategy that fit their essential characteristics.

3. Current Challenges

Although AI has shown great potential in the field of material design, the following core obstacles still need to be overcome in order to realize the comprehensive transformation from theoretical prediction to industrial application.

3.1. Data Level "Small Sample" Dilemma

The lack of high-quality and standardized experimental data is the primary bottleneck restricting the performance of AI models. Different from the field of computer vision with massive open source data, the experimental results of functional materials are often scattered in different research groups and literatures. Due to the lack of uniform characterization criteria, and researchers tend to publish only "successful" experimental results (i.e., positive data), the model faces serious data sparsity and sample bias during training. This "small sample" situation greatly limits the generalization ability of the deep learning model, making it prone to inaccurate prediction or overfitting when exploring the unknown chemical space.

3.2. Disconnection Between "Digital Model" and "Physical Entity"

Another severe challenge of AI reverse design is how to ensure the experimental synthesizability of the design scheme. The generation algorithm can often give the structure scheme with the best performance from the mathematical level, but in the real physical world, these structures may not be prepared due to thermodynamic instability or kinetic path obstruction. If AI ignores the realistic constraints such as material growth mechanism, precursor activity and process cost, the "ideal material" produced by AI can only stay in the computer simulation stage. Therefore, how to integrate complex physical synthesis constraints into the generation strategy is the key to open up the closed loop of reverse design.

3.3. Balance Decision of Multiple Performance Objectives

In practical industrial applications, functional materials often need to take into account a variety of mutually exclusive performance indicators. For example, high-efficiency optoelectronic materials may be accompanied by poor chemical stability, or high-performance structural alloys are often expensive. This multi-objective trade-off requires AI to find "Pareto optimal solution" in multidimensional space. However, with the increase of constraints, the difficulty of optimization increases exponentially. The model not only needs to have strong search efficiency but also needs to have high sensitivity in dealing with high-dimensional nonlinear conflicts, so as to ensure that the final scheme achieves global balance in comprehensive performance.

4. Future Outlook

4.1. Physical Information Driven

Future research will no longer rely solely on the pure data-driven model but will focus on building a physical information-driven AI model (Physics-informed AI). By implementing the Schrodinger equation, thermodynamic laws, and other basic physical criteria as constraints into the neural network, researchers can guide the model to generate a design scheme that conforms to both statistical laws and physical knowledge. This method can effectively make up for the lack of experimental data, improve the prediction accuracy of the model under extreme conditions, and make AI truly have the ability to leap from "observing correlation" to "understanding causality".

4.2. Automatic Laboratory

The experimental paradigm of material research and development is gradually evolving in self-driving labs. In this system, AI is not only responsible for the initial reverse design, but also automatically performs material synthesis and performance testing through the robot system, and

feeds back the results to the design end in real time. This unmanned closed-loop system can realize all-weather autonomous iteration and greatly eliminate the uncertainty of human operation. By shortening the R&D cycle from several years to several weeks, the automatic laboratory will become the key infrastructure for the large-scale discovery of functional materials in the future.

4.3. Knowledge Driven Decision Support

The introduction of large language models (LLMS) will provide unprecedented knowledge support for reverse design. Using advanced natural language processing technology, LLM can automatically extract key formulas, synthesis paths, and failure cases from millions of scientific literatures, providing a deep experience background for the design model. In the research and development process, the large language model can not only act as a "literature expert" to assist researchers in scheme evaluation but also optimize the design strategy through logical reasoning. This cooperation mode, which combines human prior knowledge and machine generation ability, will build a new knowledge-driven R&D framework.

5. Conclusion

This paper systematically reviews the paradigm change of artificial intelligence in the research and development of functional materials, focusing on its customized applications in the fields of energy, optoelectronics and structural mechanics. By comparing the design paths under different physical mechanisms, this paper deeply discusses the practical bottlenecks such as data sparsity, experimental composability, and model interpretability. Finally, this paper looks forward to the broad prospects of cutting-edge technologies such as physical information-driven, automatic laboratory and large language models in building a full closed loop of intelligent material research and development. With the improvement of computing power and the evolution of algorithm architecture, the role of AI in reverse design of functional materials is changing from "auxiliary tool" to "decision core". In the future, the development in this field will mainly focus on the deep integration of physical logic, the automation of experimental processes, and the comprehensive utilization of multidimensional knowledge.

References

[1] Sanchez-Lengeling B, Aspuru-Guzik A. Inverse molecular design using machine learning: Generative models for matter engineering. Science, 2018, 361(6400): 360-365.

[2] Kim C, Chandrasekaran A, Huan T D, et al. Polymer genome: A data-powered polymer informatics platform for property predictions. The Journal of Physical Chemistry C, 2018, 122(29): 17575-17585.

[3] Gomez-Bombarelli R, Wei J N, Duvenaud D, et al. Automatic chemical design using a data-driven continuous representation of molecules. ACS Central Science, 2018, 4(2): 268-276.

[4] National Academies of Sciences, Engineering, and Medicine. The Materials Genome Initiative: 10 years on. Washington, D.C.: National Academies Press, 2023.

[5] Zheng J, Wang S, Deng S, et al. Accelerating the screening of modified MA_2Z_4 catalysts for hydrogen evolution reaction by deep learning-based local geometric analysis. Energy & Environmental Materials, 2024, 7(6): e12743.

[6] Chen X, et al. Accelerating computational materials discovery with artificial intelligence and cloud high-performance computing: From large-scale screening to experimental validation. arXiv preprint, 2024.

[7] Alhashmi A, Kanoun M B, Goumri-Said S. Machine learning for halide perovskite materials ABX_3 (B = Pb, X = I, Br, Cl): Assessment of structural properties and band gap engineering for solar energy. Materials, 2023, 16(7): 2657.

[8] So S, Rho J. Designing nanophotonic structures using conditional deep convolutional generative adversarial networks. Nanophotonics, 2019, 8(7): 1255-1261.

[9] Qi X, Sun Z, Li Y, et al. Machine learning-guided design of single-phase high-entropy alloys with superior mechanical properties. npj Computational Materials, 2023, 9: 191.

[10] Ma Z, Liu H, Liu Y, et al. Design of new energy-absorbing lattice cell configuration by dynamic topology optimization. Metals, 2024, 14(12): 1348.