Computers and Artificial Intelligence；Vol.3, No.3; 2026
E-ISSN: 3006 -3728 P-ISSN：3006-371X
DOI: https://doi.org/10.70267/cai.26v3n3.209216
Published by: Zeus Press

# Review of Applications of Machine Learning in Optoelectronic Information Science

**Dongchen Huang***

*School of Physics, Changchun University of Science and Technology, Changchun 130022, China*

*Corresponding author: Dongchen Huang.*

## Abstract

With the rapid advancement of artificial intelligence technology, machine learning—as one of its core branches—has demonstrated immense application value in the field of optoelectronic information science. This article systematically reviews the research progress on the practical applications of machine learning in several sub-fields of optoelectronic information science. It covers core areas such as the design and optimization of optoelectronic materials, photoelectric detection and imaging, optical communication systems, photonic device design, and photoelectron spectroscopy analysis. It also analyzes the major challenges faced in the current application process, such as data quality and model interpretability. Future research directions are outlined, providing a systematic reference basis for researchers in related fields.

## Keywords

machine learning, optoelectronic information science, optoelectronic materials, optical communication, photonic devices

## 1. Introduction

Optoelectronic information science is a cutting-edge technology field integrating optics, electronics and computer science. Relevant research results have been widely used in energy, communication, medical care, industrial testing and other industries, and have become an important driving force for the development of modern science and technology. With the increasing complexity of optoelectronic information systems, the traditional research method of relying on theoretical analysis and experimental trial and error gradually exposes the inherent limitations of low efficiency and high cost when dealing with high-dimensional nonlinear photoelectric problems.

Machine learning is the core branch of artificial intelligence. It can automatically explore the inherent laws of data through the learning and analysis of massive data, and complete the modeling, prediction and optimization of complex systems. It has a natural advantage in dealing with high-dimensional nonlinear problems and mining deep implicit information of data. It provides a new solution for the research of optoelectronic information science. In recent years, the cross-integration of machine learning and optoelectronic information science has become a hot topic in industry research. The continuous emergence of a large number of practical research results has injected new impetus into the technological innovation and development of the field of optoelectronic information.

This article systematically sorts out the progress of the practical application research of machine learning in various subdivisions of optoelectronic information science. Analyze the core challenges facing the current research. At the same time, the future development direction of this field is expected, aiming to provide comprehensive and systematic reference for researchers in related fields.

## 2. Applications of Machine Learning in the Design and Optimization of Optoelectronic Materials

Optoelectronic materials are the core foundation for the scientific development of optoelectronic information. Its performance parameters directly determine the functional realization and operation efficiency of optical components. The design and optimization of traditional optoelectronic materials mainly relies on theoretical calculation and experimental trial and error. The overall process is extremely time-consuming and the research and development cost is high. The introduction of machine learning technology provides a new technical path for the efficient design and performance optimization of optoelectronic materials.

### 2.1 Optimization of Process Parameters for Perovskite Solar Cells

With the advantages of high power conversion efficiency and low preparation cost, perovskite solar cells have become the core research direction in the field of optoelectronic energy conversion in recent years. However, the process parameters such as annealing temperature, anealing time, precursor concentration in the preparation process will directly affect the final performance of the battery. It is difficult to quickly lock the optimal parameter combination with the traditional trial and error method.

In the relevant research of *Optimization of fabrication process parameters for perovskite solar cells assisted by machine learning* [1], researchers built an efficiency prediction model for perovskite solar cells by integrating three types of algorithms: random forest, XGBoost, and AdaBoost algorithms. Through the study of massive process parameter data sets, accurate prediction of the efficiency of solar cells under different process parameters can be realized. The experimental results confirm that the model can effectively shorten the process optimization cycle and reduce the cost of research and development. The power conversion efficiency of the device prepared after optimization based on this model reaches 23.56%, which is greatly improved compared with the traditional optimization method.

In the special guest research *Efficient optoelectronic coupling simulation and machine learning for perovskite solar cells* [2], Based on an optoelectronic multiphysics coupling model, the researchers carried out a series of multi-physical field simulation calculations and obtained the full data on the optical and electrical performance of perovskite solar cells. A neural network model that correlates microphysical quantities and macroscopic optoelectronic response has been built. The model can accurately predict the optical and electrical properties of solar cells with an error range of less than 3%, and the computing speed far exceeds that of traditional simulation methods.

### 2.2 Screening of Organic Optoelectronic Materials

The traditional material screening method in the field of organic optoelectronic materials has a shortcoming of low efficiency, and it is difficult to quickly select the target materials with excellent comprehensive performance. In the relevant research of *Application of machine learning in the screening of organic optoelectronic materials* [3]. Researchers constructed a dataset with nearly 1,000 sets of organic polymer solar cell donor materials data. Focusing on the electronic characteristics of the material and the sequence characteristics of molecular fingerprints, Through the completion of training and learning through the random forest classification model, the rapid prediction of the performance of organic optoelectronic materials has been realized, which has significantly improved the work efficiency of material screening.

## 3. Applications of Machine Learning in Photoelectric Detection and Imaging

Since photoelectric detection and imaging technology have extremely important applications in the fields of industrial detection, biomedicine and security monitoring, Therefore, after the introduction of machine learning technology, the performance of the optoelectronic detection system has been greatly improved. The problems of noise interference and low defect recognition accuracy encountered in traditional testing are naturally solved.

### 3.1 Prediction of Photocathode Reflectivity

During the fabrication of photodetectors, the growth state of the photocathode directly determines the final performance of the detector. In the related research of *LSTM-based prediction model for the reflectivity of $K_2CsSb$ photocathodes* [4]. To address the industry challenge of predicting the growth state in advance during the fabrication of $K_2CsSb$ photocathodes, researchers have proposed a reflectivity prediction model based on a Long Short-Term Memory (LSTM) recurrent neural network. This model fully leverages the temporal correlations within reflectance data to perform feature extraction via a two-layer LSTM network, thereby achieving precise prediction of photocathode reflectance and providing a reliable technical means for the real-time monitoring of the photocathode fabrication process.

### 3.2 Surface Defect Identification of Optical Components

In high-power solid-state laser devices, surface defects in optical components can severely restrict the operating performance of the laser system. In the related research of *Small-sample surface defect identification of optical components based on deep learning* [5]. To address the need for precise detection of surface defects in optical components of high-power solid-state laser devices under small sample conditions, researchers have proposed the ICFNetV2 detection method, which integrates data augmentation and deep residual networks. This method effectively mitigates the gradient decay problem during the training of deep networks, significantly enhances the efficiency of cross-layer feature propagation, and ultimately achieves high-precision recognition of surface defects on optical components.

### 3.3 Photovoltaic Defect Detection

Defects in photovoltaic modules can severely reduce the power generation efficiency and lifespan of photovoltaic systems. In the related research of *Photovoltaic hot-spot detection integrating knowledge distillation and attention mechanism* [6]. In view of the technical difficulties of multi-scale target detection of hot spot faults of photovoltaic modules in complex environments. The researchers proposed a detection algorithm that integrates the mechanism of knowledge distillation and attention, and at the same time designed a functional module that integrates high-order space interaction and channel attention. It greatly enhances the expression ability of fault characteristic information and effectively improves the accuracy of photovoltaic hot spot detection.

### 3.4 Parameter Optimization of Optoelectronic Detection Systems

In photoelectric imaging systems such as underwater lidar and range-gated imaging, the imaging quality is often affected by environmental interference, equipment noise and other factors. The research team of Huazhong University of Science and Technology has carried out in-depth research on the performance improvement of this kind of system. In the study Online optimization of operating parameters for range-gated imaging systems based on Gaussian process regression [7]. The researchers adopted the machine learning method of Gaussian process regression to realize the online optimization of the working parameters of the range-gated imaging system. Effectively solve the problems of low debugging efficiency of traditional method parameters and poor adaptability in complex environments; At the same time, the team further combined the Bayesian estimation method and proposed the Bayesian reconstruction method for underwater three-dimensional distance selective imaging. The noise reduction enhancement of the imaging system has been realized, and the combination of the two studies has greatly improved the detection ability and imaging quality of the system.

## 4. Applications of Machine Learning in Optical Communication Systems

Since optical communication is the foundation of modern communication technology, and the data flow has grown rapidly in recent years, Therefore, the optical communication system is bound to face the problems of transmission capacity improvement, nonlinear effect compensation, and resource optimization and allocation. And machine learning technology provides a very natural new way of thinking for this.

### 4.1 Submarine Optical Cable Route Planning

The routing planning of submarine optical cables directly affects the stability and cost of communication. In the study *Research on machine learning-assisted multi-objective optimization for submarine optical cable route planning* [8]. The researchers proposed a machine learning-assisted multi-objective optimization submarine optical cable routing planning algorithm. Take advantage of reinforcement learning, learn independently, and optimize costs and risks synchronously. Considering the seabed topography, water depth and other parameters, it effectively improves the overall optimization ability of submarine optical cable routing planning and reduces cumulative costs and risks.

### 4.2 Optimization of Data Transmission in Optical Communication Networks

The faulty data in the optical communication network will affect the quality of data transmission. In the study of *Research on mathematical models of data transmission in optical communication networks based on deep learning* [9]. Researchers have designed a mathematical model of optical communication network data transmission based on deep learning. Using the convolutional neural network to extract the data transmission characteristics of the optical communication network, output data categories through the full connection layer and the Softmax classifier, and introduce the L1 regular penalty term and the optimization activation function to realize the sparseness of parameters, which effectively improves the quality and stability of optical communication network data transmission.

### 4.3 Prediction of Performance Degradation in OTN Networks

The performance degradation of the optical transmission network (OTN) will affect the quality of communication services. In the study of Prediction of performance degradation in OTN networks based on machine learning [10]. The researchers proposed an OTN network performance degradation prediction method based on machine learning. Through the study of network performance data, the advance prediction of OTN network performance degradation is realized, which provides strong support for network operation and maintenance and fault handling.

### 4.4 Synchronization in Chaotic Optical Communication

Chaotic optical communication has high security. However, the traditional chaotic synchronization method is less efficient and limited by the strict matching requirements of the hardware parameters of the transceiver, which makes it difficult to support high-speed and long-distance chaotic encrypted signal transmission. The research team of the Intelligent Optical Fiber System Laboratory of Shanghai Jiaotong University has carried out in-depth research on this issue. In the official research result *Coherent chaotic optical communication of 30 Gb/s over 340-km fiber transmission via deep learning* [11], The researchers proposed a chaotic confidential communication scheme based on coherent detection and neural network synchronization: Coherent detection is used to convert the chaotic signal to the digital domain, and the digital signal processing (DSP) algorithm is combined to compensate for the channel transmission damage of the chaotic encrypted signal. At the same time, chaotic synchronization in the digital domain is realized through neural networks, getting rid of the dependence of traditional light domain synchronization on hardware parameter matching. In the end, the team successfully realized the transmission experiment demonstration of 30 Gb/s phase chaotic encryption signal at 340 km optical fiber, which effectively improved the transmission distance and practical ability of chaotic optical communication, and ensured the physical layer security of communication.

## 5. Applications of Machine Learning in Photonic Device Design

Photonic device design needs to meet specific optical performance indicators such as transmission spectrum, reflection spectrum, field distribution, etc. The traditional positive design method needs to go through multiple rounds of iterative optimization, and there are shortcomings with long design cycles and low execution efficiency. The reverse design method based on machine learning technology can directly push back the device structure parameters according to the target optical characteristics, greatly improving the efficiency of device design.

### 5.1 Design of Photonic Crystal Waveguides

Photonic crystal waveguide is one of the core photonic devices, and its dispersion characteristics directly determine the overall performance of the device. In the study of *Design of broadband low-dispersion valley photonic crystal slow-light waveguides based on machine learning* [12]. The researchers used the two-dimensional plane wave expansion method to simulate the dispersion curve of the valley photonic crystal waveguide and analyzed the effect of the equilateral triangle air hole size near the interface on the broadband low dispersion slow light characteristics of the topological boundary state. The researchers built and trained a forward predictive neural network with a series classification regression structure, and finally completed the screening of valley photonic crystal waveguides with excellent slow light characteristics, which significantly improved the design efficiency of photonic crystal waveguides.

### 5.2 Optoelectronic Hybrid Neural Networks

The optoelectronic hybrid neural network combines the parallel linear computing ability of photon elements and the nonlinear processing advantages of electronic components, which has a high application potential in classification tasks. In the study of *On-chip training and noise immunity of optoelectronic hybrid optical convolutional neural networks* [13]. The researchers built a optoelectronic hybrid optical convolutional neural network architecture, The linear computing link of the architecture is realized by the optical computing layer based on the Mach-Zender interferometer. The pooled computing and model training process is completed in FPGA. At the same time, the team studied the FPGA-based on-chip training scheme, analyzed the impact of noise interference on the on-chip training effect, and finally effectively optimized the operation performance and anti-interference ability of the optoelectronic hybrid neural network.

### 5.3 Design of Carbon-based Photodetectors

Carbon-based photodetectors have excellent optical and electrical properties and have broad application prospects in the field of optoelectronic devices. In the study of *Research progress of waveguide-integrated carbon-based infrared photodetectors* [14]. The researchers summarized the relevant frontier explorations: At present, the academic community is assisting in the design of carbon-based heterojunction devices based on machine learning-driven structural optimization strategies. The relationship between the atomic structure of the carbon skeleton and the macroscopic performance is revealed through intelligent algorithms. At the same time, in combination with the van der Waals heterojunction project, the working wavelength of heterojunction devices has been successfully expanded to 1550 nm communication window, which provides a new theoretical framework for the design of carbon-based optical components.

## 6. Applications of Machine Learning in Photoelectron Spectroscopy Analysis

Photoelectron spectroscopy is a key technology for analyzing the electronic structure of materials. Its output data has the characteristics of high dimension and high complexity. Traditional data analysis methods cannot effectively mine the deep physical information contained in the data. The introduction of machine learning technology provides a new technical path for the processing and analysis of photoelectron energy spectrum data.

In the study of *Applications and prospects of machine learning in photoelectron spectroscopy* [15]. The researchers systematically sorted out various application scenarios of machine learning in the field of photoelectron spectroscopy, specifically covering spectral data noise reduction, electron structure analysis, chemical composition analysis and spectral prediction and other directions. The research results show that noise reduction processing of photoelectron energy spectrum data through machine learning models can effectively remove noise interference and improve the quality of spectral data. The machine learning model based on massive spectral data to complete training and learning can accurately analyze the electronic structure and chemical composition of materials, providing stable and reliable data support for material performance-related research.

## 7. Challenges and Prospects

Although machine learning has achieved fruitful application results in the field of optoelectronic information science, the existing research still faces many urgent challenges to be solved:

### 7.1 Data Quality and Annotation Issues

The performance of the machine learning model is highly dependent on the quality and scale of the training data set. High-quality experimental data in the field of optoelectronic information science is generally difficult to obtain and the collection cost is high. Data annotation work also needs strong professional knowledge support. The above factors restrict the performance optimization of machine learning models to a certain extent. In addition, there is an uneven distribution of samples in data sets in some subdivisions, which will directly lead to a decline in the identification and prediction performance of the model on a small number of sample categories.

### 7.2 Model Interpretability Issues

The vast majority of machine learning models, especially deep learning models, are classified as "black box" models, and their internal decision-making process lacks clear interpretability. Researchers in the field of optoelectronic information science need to clarify the decision-making basis of the model to verify the rationality and reliability of the output results of the model. Therefore, improving the interpretability of machine learning models is one of the core research directions in the future in this field.

### 7.3 Integration of Physical Mechanisms and Machine Learning

Most of the existing machine learning models adopt a pure data-driven construction mode, which has not achieved deep integration with the basic physical mechanism in the field of optoelectronic information science. Embedding the physical mechanism into the framework of the machine learning model can effectively improve the generalization ability and output reliability of the model, and reduce the dependence of the model on massive training data, which is the key research direction in the future in this field.

### 7.4 Limitations of Hardware and Computational Resources

The training and operation of some machine learning models, especially deep learning models, require a large amount of computing resources and time costs. This characteristic restricts its large-scale promotion in actual engineering scenarios to a certain extent. Therefore, developing lightweight machine learning models or relying on emerging technologies such as edge computing and optical computing to improve the operating efficiency of models is one of the important research directions in this field in the future.

In the future, with the continuous development of artificial intelligence technology, the application of machine learning in the field of optoelectronic information science will be more extensive and in-depth:

#### 7.4.1 Multi-model Fusion and Collaborative Optimization

By integrating different types of machine learning models, integrating the core advantages of various models, and comprehensively improving the comprehensive performance and generalization ability of the model. For example, the machine learning model that integrates physical information can be combined with the pure data-driven model to realize the complementarity of the advantages of different technical paths.

#### 7.4.2 Edge Computing and Intelligent Optoelectronic Systems

Deploying the trained machine learning model to the edge device can realize the real-time intelligent processing of the optoelectronic system and effectively improve the response speed and operation reliability of the system. For example, in optical communication systems, real-time processing of signals and dynamic optimization of system resources can be completed through edge computing technology.

#### 7.4.3 Automated Design and Optimization of Optoelectronic Systems

Develop a full-process automated machine learning design platform to realize the automated design and optimization iteration of optoelectronic materials, devices and systems. It can further improve the efficiency of research and development in the field and greatly reduce the cost of research and development.

### 7.4.4 Cross-disciplinary Integrated Applications

Further integrate machine learning with optoelectronic information science, quantum science, biomedicine and other fields, and open up new application scenarios, such as quantum optoelectronic systems, biomedical photoelectric detection, etc.

## 8. Conclusion

Since machine learning technology has brought new and powerful solutions to optoelectronic information science, it has excellent application prospects in the fields of optoelectronic material design and optimization, photoelectric detection and imaging, optical communication systems, photonic device design, and photoelectron spectroscopy analysis. Machine learning is conducive to improving the performance of optoelectronic systems, shortening the research and development cycle, and reducing the cost of research and development. Therefore, machine learning has become a powerful enabling technology for advancing research and innovation in optoelectronic information science.

Although there are still problems in data quality and model interpretability in the research on the integration of machine learning and optoelectronic information science. However, the integration of the two is deepening, so it is bound to bring new opportunities and breakthroughs to the development of optoelectronic information science. Therefore, in the future, relevant researchers should take the initiative to strengthen interdisciplinary cooperation, so that machine learning technology can effectively and fully serve the optoelectronic information science.

## References

[1] Gao, Y., Li, Y., Xia, C., & Chen, Q. (2025). Optimization of fabrication process parameters for perovskite solar cells assisted by machine learning. *Laser & Optoelectronics Progress, 62*(09), 315–321.

[2] Kong, R., Wei, Y., Chen, J., Ma, T., Zhan, Y., & Li, X. (2024). Efficient optoelectronic coupling simulation and machine learning for perovskite solar cells (Invited). *Laser & Optoelectronics Progress, 61*(01), 429–439.

[3] Zhou, Y., & Guo, P. (2022). Application of machine learning in the screening of organic optoelectronic materials. *Journal of Lanzhou Jiaotong University, 41*(04), 42–47.

[4] Wei, J., Qian, Y., & Cao, Y. (2022). LSTM-based prediction model for the reflectivity of K2CsSb photocathodes. *Applied Optics, 43*(06), 1037–1043.

[5] Shao, Y., Xin, C., & Chu, H. (2025). Small-sample surface defect identification of optical components based on deep learning. *High Power Laser and Particle Beams, 37*(12), 5–12.

[6] Hao, S., Wu, Y., Ma, X., Li, T., & Wang, H. (2023). Photovoltaic hot-spot detection integrating knowledge distillation and attention mechanism. *Optics and Precision Engineering, 31*(24), 3640–3650.

[7] Cheng, H., Li, W., Guo, W., & Yang, K. (2019). Online optimization of operating parameters for range-gated imaging systems based on Gaussian process regression. *Acta Optica Sinica, 39*(04), 198–205.

[8] Zhao, Z., Gao, G., Gan, W., Wang, H., Duan, M., & Kang, D. (2025). Research on machine learning-assisted multi-objective optimization for submarine optical cable route planning. *Optical Communication Technology, 49*(03), 10–15. https://doi.org/10.13921/j.cnki.issn1002-5561.2025.03.002

[9] Zhang, L., & Cheng, N. (2025). Research on mathematical models of data transmission in optical communication networks based on deep learning. *Laser Journal, 46*(01), 196–201. https://doi.org/10.14016/j.cnki.jgzz.2025.01.196

[10] Chen, L., Liao, L., Zhang, P., Zhu, D., Peng, Z., & Zhou, H. (2025). Prediction of performance degradation in OTN networks based on machine learning. *Study on Optical Communications, (02)*, 50–55. https://doi.org/10.13756/j.gtxyj.2025.240047

[11] Zhao, Y., Ke, J., Zhuge, Q., Hu, W., & Yi, L. (2022). Coherent chaotic optical communication of 30 Gb/s over 340-km fiber transmission via deep learning. *Optics Letters, 47*(11), 2650–2653. https://doi.org/10.1364/OL.453696

[12] Zhao, Y., Wang, Y., Hu, Z., Pu, Y., & Xi, X. (2025). Design of broadband low-dispersion valley photonic crystal slow-light waveguides based on machine learning. *Acta Optica Sinica, 45*(10), 145–152.

[13] Shao, X., Su, J., & Wang, J. (2025). On-chip training and noise immunity of optoelectronic hybrid optical convolutional neural networks. *Chinese Optics, 18*(05), 1124–1131.

[14] Lü, X., Zhang, J., Chen, J., Liu, Z., Zhao, W., Wu, Q., et al. (2025). Research progress of waveguide-integrated carbon-based infrared photodetectors. *Acta Physica Sinica, 74*(20), 341–355.

[15] Deng, X., Wu, L., Zhao, R., Wang, J., & Zhao, L. (2024). Applications and prospects of machine learning in photoelectron spectroscopy. *Acta Physica Sinica, 73*(21), 64–84.

## Funding

This research received no external funding.

## Conflicts of Interest

The authors declare no conflict of interest.

## Acknowledgment

This paper is an output of the science project.