Journal of Electronic Research and Application

Editor-in-Chief

Joselito A. Dolot

Lyceum of the Philippines University-Batangas, Philippines

BIO-BYWORD SCIENTIFIC PUBLISHING PTY LTD
(619 649 400)
Level 10
50 Clarence Street
SYDNEY NSW 2000

ISSN (ONLINE): 2208-3510
ISSN (PRINT): 2208-3502

Journal of Electronic Research and Application

Focus and Scope

Journal of Electronic Research and Application is an international, peer-reviewed and open access journal which publishes original articles, reviews, short communications, case studies and letters in the field of electronic research and application.

Topics covered but not limited to:

- Automation
- Circuit Analysis and Application
- Electric and Electronic Measurement Systems
- Electrical Engineering
- Electronic Materials
- Electronics and Communications Engineering
- Power Systems and Power Electronics
- Signal Processing
- Telecommunications Engineering
- Wireless and Mobile Communication

About Publisher

Bio-Byword Scientific Publishing is a fast-growing, peer-reviewed and open access journal publisher, which is located in Sydney, Australia. As a dependable and credible corporation, it promotes and serves a broad range of subject areas for the benefit of humanity. By informing and educating a global community of scholars, practitioners, researchers and students, it ndeavors to be the world's leading independent academic and professional publisher. To realize it, it keeps creative and innovative to meet the range of the authors' needs and publish the best of their work.

By cooperating with University of Sydney, University of New South Wales and other world-famous universities, Bio-Byword Scientific Publishing has established a huge publishing system based on hundreds of academic programs, and with a variety of journals in the subjects of medicine, construction, education and electronics.

Publisher Headquarter

BIO-BYWORD SCIENTIFIC PUBLISHING PTYLTD
Level 10
50 Clarence Street
Sydney NSW 2000
Website: www.bbwpublisher.com
Email: info@bbwpublisher.com

2025, Volume 9, Issue 2

Table of Contents

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Enhancing Indoor Object Detection with xLSTM Attention-Driven YOLOv9 for Improved 2D-Driven 3D Object Detection

Yu He[1]*, Chengpeng Jin[2], Xuesong Zhang[1]

[1]School of Railway Intelligent Engineering, Dalian Jiaotong University, Dalian 116028, Liaoning, China
[2]School of Faculty of Land and Resources Engineering, Kunming University of Science and Technology, Kunming 650031, Yunnan, China

Corresponding author: Yu He, heyu3517@gmail.com

Abstract: Three-dimensional (3D) object detection is crucial for applications such as robotic control and autonomous driving. While high-precision sensors like LiDAR are expensive, RGB-D sensors (e.g., Kinect) offer a cost-effective alternative, especially for indoor environments. However, RGB-D sensors still face limitations in accuracy and depth perception. This paper proposes an enhanced method that integrates attention-driven YOLOv9 with xLSTM into the F-ConvNet framework. By improving the precision of 2D bounding boxes generated for 3D object detection, this method addresses issues in indoor environments with complex structures and occlusions. The proposed approach enhances detection accuracy and robustness by combining RGB images and depth data, offering improved indoor 3D object detection performance.

Keywords: Deep learning; Object detection; Attention mechanism

Online publication: February 24, 2025

1. Introduction

Currently, 3D object detection algorithms are widely applied in environmental perception systems, particularly in fields such as robotic control, intelligent surveillance, and autonomous driving. Recent 3D object detection algorithms typically rely on 3D sensors to capture spatial data [1-3]. Although traditional high-precision LiDAR and high-precision camera devices provide relatively accurate spatial information, their high costs and complex deployment requirements make them unsuitable for consumer-grade devices. In contrast, RGB-D sensors, exemplified by Kinect, have become an ideal choice for the consumer market due to their low cost and ease of deployment, making them well-suited to meet the demands of indoor environmental perception.

However, despite the significant cost advantage of RGB-D sensors over LiDAR, they still exhibit certain limitations in accuracy, especially in terms of precision and depth perception in complex indoor environments. Nonetheless, RGB-D sensors provide real-time RGB images and corresponding 3D depth information, making it possible to optimize 3D object detection algorithms using RGB images. In indoor object detection tasks, RGB images not only provide rich texture and color information but also help enhance detection robustness, particularly in low-light or occluded environments. Therefore, the integration of RGB images with 3D depth data holds significant research value and practical importance, particularly for the application of consumer-grade RGB-D sensors in indoor 3D object detection.

The F-ConvNet algorithm offers a promising approach by using 2D bounding boxes (2D BBoxes) generated by a 2D detector to slice the 3D space, thereby assisting in 3D object detection [4]. However, algorithms, such as F-ConvNet that rely on 2D detectors to generate 2D bounding boxes, have some shortcomings. Specifically, the generated 2D bounding boxes lack specificity and fail to fully account for the complexity and occlusions inherent in indoor environments. As a result, these methods struggle to precisely identify and localize indoor objects, particularly in narrow and complex spaces.

To address this issue, we propose an enhanced method that integrates the attention-driven YOLOv9 algorithm with xLSTM, incorporating it into F-ConvNet as the 2D detection module [5,6]. By introducing the attention mechanism of xLSTM, we can more accurately focus on and capture features relevant to 3D objects during the generation of 2D bounding boxes, especially in indoor environments. The temporal characteristics and long-short term memory capabilities of xLSTM enable it to better handle the dynamic changes and complex structures of indoor objects, thereby generating more precise and targeted 2D bounding boxes. This improvement significantly enhances the performance of F-ConvNet in indoor object detection, allowing the 3D object detection algorithm to more effectively combine RGB images and depth information, achieving higher detection accuracy and robustness.

In this paper, our main contributions can be summarized as follows:

(1) We integrate xLSTM into YOLOv9 to enhance its focus on indoor object detection, enabling it to generate more precise 2D bounding boxes (2D BBoxes). This improvement allows for better identification and localization of objects in indoor environments.

(2) We incorporate the xLSTM-enhanced YOLOv9 into the F-ConvNet framework, improving the accuracy of the 2D bounding boxes used for 3D object detection. This integration helps to better handle the complexities of indoor environments, providing more accurate and reliable detection results.

2. The proposed method

2.1. Vision-LSTM (ViL) + YOLOv9

Vision-LSTM (ViL) is a visual general-purpose network backbone constructed using xLSTM with residual blocks, as shown in Figure 1 [6]. The process is as follows: First, the input image is split into several patches, and then these patches are linearly projected. Each patch is augmented with a learnable positional information vector. Next, these patches are processed by alternating mLSTM blocks. For even-numbered blocks, the sequence is first flipped and then processed by the mLSTM layer. The processed sequence is normalized and finally passed through a linear projection for classification. The ViL encoder processes this sequence of patches, averaging the outputs of the first and last patch, and then the final classification result is output through a linear classification head.

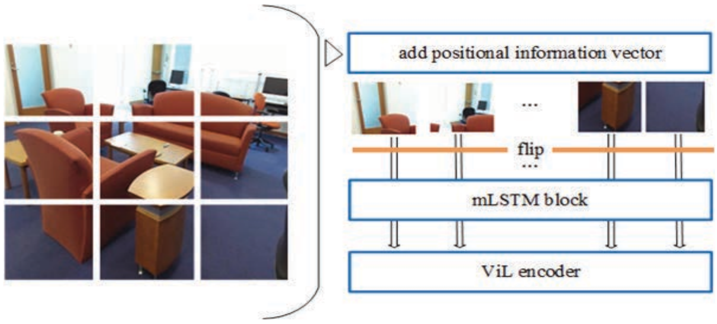

Figure 1. Vision-LSTM (ViL) architecture diagram

Vision-LSTM (ViL) utilizes mLSTM blocks to process image patches in a manner similar to time-series processing, introducing the concept of sequence processing into image tasks, akin to handling time-series data. We integrate Vision-LSTM (ViL) into YOLOv9 by introducing the EfficientLSTM module to enhance the temporal processing capability of image features. The ViLBlock and SequenceTraversal modules contained in EfficientLSTM are responsible for performing temporal processing of image features through mLSTM (multi-layer long short-term memory) networks. Specifically, the core component of ViL-mLSTM blocks is embedded into the feature extraction part of YOLOv9, helping the network capture dynamic changes and spatial relationships within the image.

2.2. xLSTM_YOLOv9 + F-ConvNet

In the process of integrating xLSTM_YOLOv9 into F-ConvNet, we begin by separating the RGB-D data into RGB and Depth components. The RGB data is then passed through xLSTM_YOLOv9, where the 2D detector generates 2D bounding boxes (2DBBoxes). These 2DBBoxes are subsequently mapped into the frustum generation structure of F-ConvNet to assist in the creation of the frustum. The entire workflow is illustrated in **Figure 2**.

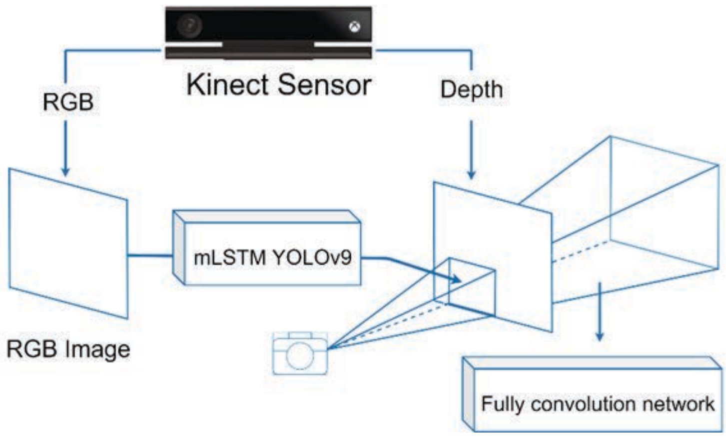

Figure 2. Integrating xLSTM_YOLOv9 into F-ConvNet

This approach leverages the strengths of both 2D and 3D detection, enabling more accurate object localization. By utilizing the RGB data for 2D bounding box generation through xLSTM_YOLOv9, we can refine the frustum generation process, ensuring it focuses on the most relevant areas of the scene. This integration not only improves detection accuracy in complex environments but also enhances the robustness of the system by efficiently combining the spatial and depth information. As a result, this method optimizes the handling of occlusions and varying object scales, making it especially effective in real-world 3D object detection tasks.

3. The proposed method

We integrated the xLSTM structure into YOLOv9 to enhance the performance of indoor object detection algorithms. To validate our proposed approach, we utilized the SUN RGB-D dataset. In this dataset, we selected 5050 scenes for validation and 5285 scenes for training. For the 2D experiments, we used the 2D data provided by the SUN RGB-D dataset [7]. The experimental results are presented in **Table 1**.

From **Table 1**, it is evident that after incorporating the xLSTM structure, the performance of the YOLOv9 algorithm in indoor object detection has been significantly improved, especially when dealing with objects that are prone to occlusion. The introduction of xLSTM allows the model to better capture temporal information, thereby enhancing its ability to recognize objects in dynamic scenes and complex environments. Compared to the traditional YOLOv9, the model with xLSTM exhibits stronger robustness in occluded object detection, small object recognition, and background noise suppression. This improvement is particularly crucial for real-world applications of object detection in indoor environments, where in complex scenarios with multiple objects interwoven, the xLSTM structure can effectively alleviate the recognition difficulties posed by occlusions and challenging backgrounds.

Table 1. Shortcut keys for the template

Method	bed	table	sofa	chair	toilet	desk	dresser	night stand	bookshelf	bathtub	mAP0.25
YOLOv9	0.81	0.46	0.61	0.71	0.86	0.33	0.48	0.69	0.54	0.60	0.61
xLSTM YOLOv9	0.85	0.55	0.69	0.76	0.87	0.35	0.49	0.69	0.54	0.52	0.64

In **Table 2**, we compare several typical RGB-D-based algorithms and also compare them with the F-ConvNet baseline algorithm, validating the effectiveness of our proposed approach. The evaluation metric used in **Table 2** is mAP@0.25 proposed by Song *et al.* [7].

Next, we integrated xLSTM_YOLOv9 into F-ConvNet, and the results are shown in **Figure 3**. In **Figure 3**, the first row displays the 2D RGB images from the SUN RGB-D dataset, the second row shows the detection results of the RGB scene using xLSTM_YOLOv9, and the third row presents the F-ConvNet detection results in the 3D point cloud scene, generated using the detection results from xLSTM_YOLOv9 and the frustum.

Table 2. Comparison of methods on different object categories

Method	bathtub	bed	bookshelf	chair	desk	dresser	night stand	soft	table	toilet	mean
DSS [8]	44.2	78.8	11.9	61.2	20.5	6.4	15.4	53.5	50.3	78.9	42.1
COG [9]	58.26	63.67	31.80	62.17	45.19	15.47	27.36	51.02	51.29	70.07	47.63
2Ddriven3D [10]	43.45	64.48	31.40	48.27	27.93	25.92	41.92	50.39	37.02	80.40	45.12
PointFusion [11]	37.26	68.57	37.69	55.09	17.16	23.95	32.33	53.83	31.03	83.80	45.38
Ren *et al.* [12]	76.2	73.2	32.9	60.5	34.5	13.5	30.4	60.4	55.4	73.7	51.0
F-PointNet [13]	43.3	81.1	33.3	64.2	24.7	32.0	58.1	61.1	51.1	90.9	54.0
F-ConvNet [4]	61.32	83.19	36.46	64.4	29.67	35.1	58.42	66.61	53.34	86.99	57.55
Ours	63.95	84.69	32.74	77.84	24.63	34.0	61.4	66.32	50.65	88.59	58.32

Figure 3. The detection result diagram of the F-ConvNet algorithm after integrating xLSTM_YOLOv9

Disclosure statement

The authors declare no conflict of interest.

Author contributions

Conceptualization: Yu He, Xuesong Zhang
Investigation: Yu He, Chengpeng Jin
Formal analysis and writing: Yu He, Chengpeng Jin

References

[1] Hu Y, Yang J, Chen L, et al., 2023, "Planning-Oriented Autonomous Driving". Proceedings of the IEEE/CVF Conference on CVPR, 2023: 17853–17862.

[2] Yang H, Zhang S, Huang D, et al., 2024, Unipad: A Universal Pre-Training Paradigm for Autonomous Driving. Proceedings of the IEEE/CVF Conference on CVPR, 2024: 15238–15250.

[3] Min C, Zhao D, Xiao L, et al., 2024, Driveworld: 4D Pre-Trained Scene Understanding via World Models for Autonomous Driving. Proceedings of the IEEE/CVF Conference on CVPR, 2024: 15522–15533.

[4] Wang Z, Jia K, 2019, Frustum ConvNet: Sliding Frustums to Aggregate Local Point-Wise Features for Amodal 3D Object Detection. 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2019: 1742–1749.

[5] Wang CY, Yeh IH, Liao HYM, 2024, YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information. Computer–ECCV2024, 15089: 1–21.

[6] Beck M, Poppel K, Spanring M, et al., 2024, xLSTM: Extended Long Short-Term Memory. arXiv:2405.04517v1, 2024: 1–56.

[7] Song S, Lichtenberg SP, Xiao J, 2015, SUN RGB-D: A RGB-D Scene Understanding Benchmark Suite. Proceedings of the IEEE Conference on CVPR. Boston, Massachusetts, 567–576.

[8] Song S, Xiao J, 2016, Deep Sliding Shapes for Amodal 3D Object Detection in RGB-D Images. Proceedings of the

IEEE Conference on CVPR, Las Vegas, Nevada, 808–816.

[9] Ren Z, Sudderth EB, 2016, Three-Dimensional Object Detection and Layout Prediction Using Clouds of Oriented Gradients. Proceedings of the IEEE Conference on CVPR, Las Vegas, Nevada, 1525–1533.

[10] Lahoud J, Ghanem B, 2017, 2D-Driven 3D Object Detection in RGB-D Images. Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 4622–4630.

[11] Xu D, Anguelov D, Jain A, 2018, Pointfusion: Deep Sensor Fusion for 3D Bounding Box Estimation. Proceedings of the IEEE Conference on CVPR, Salt Lake City, Utah, 244–253.

[12] Ren Z, Sudderth EB, 2018, 3D Object Detection with Latent Support Surfaces. Proceedings of the IEEE Conference on CVPR, Salt Lake City, Utah, 937–946.

[13] Qi CR, Liu W, Wu C, et al., 2018, Frustum Pointnets for 3D Object Detection from RGB-D Data. Proceedings of the IEEE Conference on CVPR, Salt Lake City, Utah, 918–927.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Application of Resource Orchestration Theory in the Research of AI and HI to Achieve Value Creation

Ni Ran*

Guizhou University of Finance and Economics, Guiyang 550025, Guizhou, China

Corresponding author: Ni Ran, 728332610@qq.com

Abstract: This paper grasps the research theme of artificial intelligence (AI) and human intelligence (HI) synergy to create value, and analyzes the development status of AI and HI in the current context of digital intelligence, as well as the significance of their synergy to empower value creation. At the same time, the theory of resource arrangement is introduced, and the connotation and composition of the theory are summarized, as well as the development in the field of research and application. This paper focuses on revealing the intrinsic relationship between resource orchestration theory and AI and HI collaborative work, aiming to fully explore the potential of resource orchestration theory in the collaborative innovation of AI and HI, and put forward practical suggestions based on this.

Keywords: Resource arrangement theory; Artificial intelligence (AI) and human intelligence (HI) synergy; Value creation.

Online publication: February 24, 2025

1. Introduction

In the era of digital intelligence, artificial intelligence (AI) as a disruptive emerging technology, has become the core driving force for the vigorous development of the digital economy [1]. The rapid development of AI technology not only provides unprecedented opportunities for China's economic transformation and relay innovation but also profoundly changes the logic of enterprise value creation [2]. With the continuous improvement of the "human-like" nature of AI and its own algorithm capabilities, coupled with the rapid expansion of AI's commercial application territory, the possibility and necessity of coexistence and collaboration between AI and humans are constantly increasing. The combined effect of human-machine collaboration can realize the two-way reinforcement of the intelligence of both parties, amplify the capabilities of both parties in their respective fields of expertise, and create new work value [3].

As a methodology for managing resources, resource allocation theory clarifies how enterprises allocate resources and establish competitive advantages to obtain maximum value from a dynamic perspective, which provides novel ideas for enterprises to obtain and transform resources and create value [4]. In the collaboration between AI and HI, it is crucial to strategically allocate resources such as personnel, finances, materials, and information. Resource orchestration theory plays a key role in identifying which resources are essential and effective, as well as in determining how to integrate and utilize them for maximum efficiency. This approach ensures robust support for the collaborative efforts of AI and HI. In addition, the resource allocation theory provides a complete and reasonable explanatory framework for analyzing how enterprises allocate effective resources and develop the ability system in production and operation activities, to build competitive advantages and maximize overall benefits.

Based on this, this paper analyzes the basic principles and theoretical framework of resource orchestration theory, discusses the research progress of resource orchestration theory in the field of enterprise management and innovation, analyzes the internal relationship between resource orchestration theory and AI and HI synergy, clarifies the role of resource orchestration theory in helping to realize the synergistic value creation of AI and HI, and puts forward relevant practical suggestions.

2. AI and HI collaboration to achieve value creation

2.1. AI and HI synergy

AI technologies, especially machine learning and deep learning, are gradually merging with HI to form a new model of human-machine collaboration [5]. This collaborative model not only improves work efficiency and accuracy but also promotes further innovation and development of technology. With the trend of deep integration of AI and HI collaboration, its application scenarios are also deepening and expanding. For example, in sales training, AI can assist sales managers in improving sales agents with top sales performance rankings [6]; The use of emotion recognition software in call service centers can effectively regulate employees' performance in interpersonal emotions and make them obtain higher emotional satisfaction [7]. In the field of knowledge production, such as news writing, artistic creation, education and teaching, the high-density information search and data processing capabilities of AI combined with the advantages of human beings in knowledge, experience, logic, ethics, etc., can play a huge advantage and significantly improve production efficiency [8].

2.2. AI and HI synergistically empower value creation

Through the deployment and application of AI technology, enterprises can reconstruct business value and empower value creation. First, AI and HI working together can significantly improve productivity. AI can take over repetitive tasks, freeing up manpower for more complex and higher-value work. At the same time, the synergy between AI and HI can also optimize workflows, reduce human error, and improve overall work efficiency. Second, AI and HI can work together to drive innovation. By combining human creativity with the computing power of machines, new possibilities can be explored and new products and services can be developed. Finally, AI working in tandem with HI can enhance the adaptability of organizations. In a rapidly changing market environment, AI can quickly respond to market changes and adjust strategies, while HI can judge market trends based on experience and make corresponding decisions, which can enable organizations to be more agile and agile in responding to market challenges. Therefore, enterprises can create new value through the application of AI technology and human intelligence as well as promote the excellent growth and sustainable development of enterprises.

3. Overview of resource allocation theory

Resource allocation theory improves the concept of resource base based on the perspective of action, and explains the process of acquiring, integrating, and utilizing resources to establish competitive advantage and obtain maximum value from the dynamic dimension. Sirmon proposed the resource orchestration theory by combining resource management and asset orchestration. He argued that, guided by strategy, the process through which managers search for and select key resources—along with structuring, bundling, and utilizing them—can be transformed into a competitive advantage for enterprises [9].

3.1. Theoretical composition of resource arrangement

According to existing research, resource allocation theory includes three sub-processes: resource structuring, resource capacity (resource bundling), and resource leverage (resource utilization) [10]. Resource structuring is the basis of enterprise resource management, which refers to the process of acquiring, accumulating, and stripping resources to form a resource portfolio. Resource capability emphasizes the process of forming unique and difficult-to-imitate capabilities through integration and allocation, which are an important source of competitive advantage for enterprises in market activities. Resource leverage refers to making full use of the company's capabilities, as well as specific market opportunities, through the optimal allocation and use of resources, enterprises can continue to improve their competitiveness and market position. In general, the three processes of resource orchestration theory are interrelated and interdependent. Together, these three processes form the core framework of resource orchestration theory, which provides important guidance for enterprises to achieve competitive advantage and sustainable development.

3.2. Research field of resource allocation theory

With the deepening of research, the research field of resource allocation theory has been expanding. Initially, the theory was mainly applied in the field of business management and project management to help organizations better manage and utilize resources, and improve productivity and project success. However, in recent years, the research on resource allocation theory has been expanded to a wider range of fields, such as digital transformation, innovation and entrepreneurship, capacity cultivation, and value creation [11–15]. The research in these fields not only enriches the content of resource orchestration theory but also provides new application scenarios and challenges. In practical applications, such as the green transformation of manufacturing, engineering project management, breakthroughs in key core technologies of enterprises, and digital transformation of education, the practical application effect of resource allocation theory has also been widely recognized. These successful cases provide strong support and reference for the further development and application of resource allocation theory.

4. The role of resource orchestration theory in the synergy between AI and HI to achieve value creation

As mentioned above, the core of resource orchestration theory is to achieve the best benefits and maximize goals through the rational allocation and integration of resources. The essence of AI and HI collaboration is to improve business efficiency, meet the individual needs of users, and promote enterprise innovation and development. From this point of view, there is an intrinsic relationship between resource orchestration theory and AI and HI working together to create value. In the collaborative work of AI and HI, enterprises will give full play to the complementarity and synergy between the two according to business needs, form a resource combination by considering resource allocation and integration, and give full play to the advantages of the combination to achieve the optimal resource utilization effect. Only by continuously optimizing the allocation of resources and improving the efficiency of resource utilization can AI and HI give full play to their capabilities and maximize complementarity and synergy to promote enterprises to create new business value.

Resource orchestration theory also provides system support and guidance for the collaborative work of AI and HI. Before integrating AI and HI, enterprises need to clarify their requirements and goals as well as analyze and sort out existing problems for business processes. In this case, enterprises can apply resource orchestration theory to conduct detailed analysis and evaluation of various resources (including AI systems, human resources, data resources, etc.), and fully consider factors such as resource availability and utilization efficiency to determine the optimal allocation and allocation mode of various resources. Based on the results of resource analysis, enterprises can formulate a collaborative work strategy between AI and HI as well as consider factors such as the function, performance, compatibility, and ease of use of the AI system. After the AI and HI collaborative work system is built, it is necessary to conduct sufficient testing and optimization, find and solve existing problems in a timely manner, continuously optimize the system, and enhance the effect of collaborative work. In the end, it can help enterprises maximize the use of resources and the successful implementation of projects, create new business value and customer value for enterprises, reshape the status of the industry, and accelerate the iterative upgrading of enterprises.

5. Conclusions and prospects

Resource orchestration theory plays an indispensable role in the synergy between AI and HI to achieve value creation. In the AI and HI collaboration scenario, the application of resource orchestration theory can scientifically identify which resources are most critical to value creation and optimize allocation based on business needs. In addition, in the practice of AI and HI collaboration, enterprises can promote the birth and development of new technologies, new products, and new models through reasonable resource orchestration, accelerate the pace of innovation of enterprises and industries, form stronger comprehensive capabilities, and create greater value.

With the continuous development of AI technology, the application of resource orchestration theory in the collaboration between AI and HI will be more in-depth in the future. Enterprises need to continuously explore the deep integration mode of technology and management to achieve the optimal allocation of resources and the maximization of collaborative efficiency. At the same time, enterprises need to constantly explore new collaboration models and value creation methods to adapt to the changing market environment and customer needs.

Disclosure statement

The author declares no conflict of interest.

References

[1] Cheng Y, Cheng Y, He Q, 2022, Research on Enterprise Value Creation Model and Mechanism of Artificial Intelligence Technology Application—Based on the Perspective of Knowledge Base View. Science and Technology Management Research, 42(15): 173–187.

[2] Li S, Luo J, Zhang Z, 2023, How AI Capabilities Promote Enterprises to Achieve Value Co-Creation: An Exploratory Case Study Based on the Interaction between Enterprises and Customers. China Industrial Economics, 2023(05): 174–192.

[3] Xu Y, Li W, 2020, AI Reshaping the Work World and the Adaptation of Vocational Education Informatization. China E-Education, 2020(01): 93–98.

[4]Zhao H, Chen H, Jiang P, et al., 2022, The Impact of Platform Ecological Embedding and Data Empowerment on the Innovation Flexibility of Small and Medium-Sized Manufacturing Enterprises: Based on the Perspective of Resource Orchestration. Research and Development Management, 34(5): 1–15.

[5] Jiang Y, Dong H, 2023, On the Two-Way Interaction between Artificial Intelligence and Human Intelligence. Dialectics of Nature Communications, 45(11): 14–25.

[6] Kim JH, Kim M, Kwak DW, et al., 2022, Home-Tutoring Services Assisted with Technology: Investigating the Role of Artificial Intelligence Using a Randomized Field Experiment. Journal of Marketing Research, 59(1): 79–96.

[7] Henkel AP, Bromuri S, Iren D, et al., 2020, Half Human, Half Machine–Augmenting Service Employees with AI for Interpersonal Emotion Regulation. Journal of Service Management, 31(2): 247–265.

[8] Chen Q, Zhang L, 2023, Ethical Reflection on Artificial Intelligence in Education: Phenomenon Analysis and Vision Construction—An Analytical Perspective Based on “Human-Computer Collaboration.” Journal of Distance Education, 41(03): 104–112.

[9] Sirmon DG, Hitt MA, Ireland RD, et al., 2011, Resource Orchestration to Create Competitive Advantage: Breadth, Depth, and Life Cycle Effects. Journal of Management, 37(5): 1390–1412.

[10] Baert C, Meuleman M, Debruyne M, et al., 2016, Portfolio Entrepreneurship and Resource Orchestration. Strategic Entrepreneurship Journal, 10(4): 346–370.

[11] Wang P, Peng L, Jin H, et al., 2024, Research on the Influencing Factors and the Path of Digital Transformation in Colleges and Universities—Based on the Theory of Dynamic Capability and Resource Arrangement. Journal of Distance Education, 42(01): 41–49.

[12] Liu L, Long J, Liu R, et al., 2024, How Big Data Capability Drives Enterprise Innovation: Based on the Perspective of Resource Orchestration. Science and Technology Progress and Countermeasures, 41(07): 72–80.

[13] Tian Z, Chen H, 2023, How Can Manufacturing Companies Build Digital Capabilities? A Case Study Based on Resource Arrangement Theory. Management Case Studies and Reviews, 16(04): 489–509.

[14] Jia J, Chen H, Xu W, et al., 2023, How Can Digital Startups Achieve Value Creation through Resource Orchestration? Research & Development Management, 35(05): 131–149.

[15] Zhang Y, Sun X, Qian Y, 2022, Value Creation and Evolution in the Digital Transformation of Traditional Manufacturing Enterprises: A Longitudinal Single Case Study from the Perspective of Resource Orchestration. Economic Management, 44(04): 116–133.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research on π/4QPSK Modulation Communication Transmission System and FPGA Implementation

Yuhui Chu, Yunjie Yuan*, Xiaolei Cai, Yanyan Wang, Peijie Yin

Xi'an Electronic Engineering Research Institute, Xi'an 710100, Shaanxi, China

Corresponding author: Yunjie Yuan, yinpeijie789@hotmail.com

Abstract: This paper examines the π/4QPSK modulation communication transmission system, analyzing the performance advantages and disadvantages of π/4QPSK in comparison to QPSK. It also presents a comprehensive FPGA implementation scheme for a modulation communication transmission system, integrating RS channel coding, framing, frequency conversion, and other modules. This design is based on practical research and development requirements. The Xilinx Spartan6 chip board was used for board-level verification. The π/4QPSK modulated signal was transmitted via D/A conversion and radio frequency, with the transmitted waveform was looped back for reception. After A/D processing, the correctness of the designed modulation transmission scheme was verified.

Keywords: π/4QPSK modulation; FPGA

Online publication: February 24, 2025

1. Introduction

In communication systems, the transmitter primarily consists of modules such as modulation, encoding, and framing. The design of the transmitter's signal directly affects the performance of the entire communication system. Digital modulation primarily involves mapping signals to constellation points, with different types of constellation diagrams directly affecting the system's noise resistance and the bandwidth resources it consumes. $\pi/4$ Quadrature Phase Shift Keying ($\pi/4$QPSK) combines the advantages of Quadrature Phase Shift Keying (QPSK) and Offset Quadrature Phase Shift Keying (OQPSK), making it a form of quadrature phase shift keying modulation. The constellation diagram of QPSK modulation can experience phase jumps of up to π, which can lead to spectral expansion and larger side lobes, as well as being affected by the nonlinearity of power devices. In contrast, OQPSK limits phase jumps to $\pi/2$, resulting in less bandwidth usage and better anti-interference capabilities, but with higher implementation complexity. $\pi/4$QPSK limits phase jumps to $\pi/4$ and $3\pi/4$, with minimal phase changes, making the spectrum more compact after filtering and limiting compared to the previous two. This results in better bandwidth utilization under small-scale fading conditions. Single-carrier systems have a lower peak-to-average power ratio compared to multi-carrier systems, reducing the requirements for RF devices. Therefore, in communication systems with low throughput, such as messaging systems, or in systems with high real-time requirements, such as voice systems, single-carrier systems provide significant cost and overhead advantages. Thus, research on $\pi/4$QPSK communication systems is necessary.

$\pi/4$QPSK employs differential encoding, allowing for both coherent and non-coherent demodulation (differential demodulation). Non-coherent demodulation does not require the recovery of a local carrier, significantly reducing the impact of carrier synchronization inaccuracies on the system [1,2]. Additionally, non-coherent demodulation enables rapid synchronization, reducing the need for synchronization codes in the frame structure compared to coherent demodulation, thereby increasing the proportion of effective data and improving spectral efficiency. In fast-fading channels, differential demodulation can achieve lower bit error rates than coherent demodulation. However, $\pi/4$QPSK also has drawbacks, such as the need for linear power amplifiers and the potential loss of paired symbols in the event of reception errors [2]. Considering the trade-offs between the advantages and disadvantages of $\pi/4$QPSK modulation, it is suitable for burst wireless communication systems. In a study conducted by Zhou, $\pi/4$QPSK modulation was applied to ship communication systems to improve bandwidth efficiency and system flexibility [3]. As for Liu *et al.*, a lookup table method was used to synthesize $\pi/4$QPSK waveforms through multiple square-root raised cosine filters, achieving simpler implementation with minimal performance loss [4]. Zhou *et al.* proposed a cross-product frequency discriminator demodulation method to address frequency offset issues in burst transmissions of $\pi/4$QPSK [5]. The literature mentioned above primarily analyzes $\pi/4$QPSK from an algorithmic perspective, whereas this paper validates the system's transmission scheme through the FPGA implementation of the $\pi/4$QPSK transmitter.

2. $\pi/4$QPSK modulation transmitter implementation scheme

As shown in **Figure 1**, the transmitter of the $\pi/4$QPSK modulation system mainly consists of RS encoding, framing, serial-to-parallel conversion, differential encoding, and up-conversion modules. The transmitted signal is finally output through a DAC, resulting in the RF output of the $\pi/4$QPSK modulated waveform.

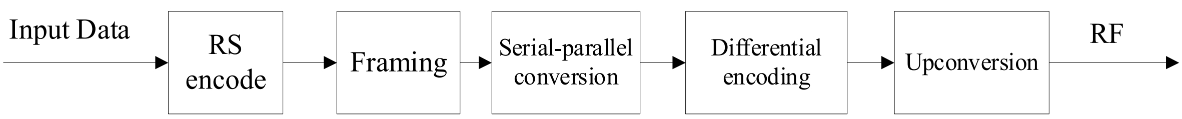

Figure 1. $\pi/4$QPSK modulation system

Reed-Solomon (RS) coding is a powerful forward error correction code suitable for correcting burst and random errors. It treats the original data symbols as polynomial coefficients and attaches a generator polynomial to the original data to form the encoded codeword. Therefore, RS coding has low overhead and implementation complexity, making it suitable for $\pi/4$QPSK systems. As shown in **Figure 2**, the RS encoder is implemented using a linear feedback shift register structure, with a division-based structure. The input consists of m original data symbols representing polynomial coefficients, and the generator polynomial is defined as $g(x)=(x-\alpha)(x-\alpha^2)...(x-\alpha^{2t})$, where α is the primitive element of the finite field $GF(2^m)$. After passing through this linear feedback shift register system, the output is the codeword.

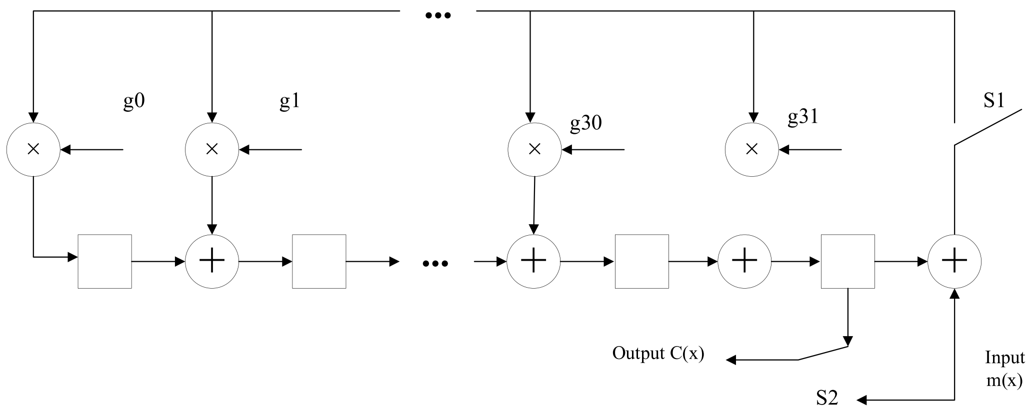

Figure 2. RS encoder structure

Since synchronization and other processing are required at the receiver, preamble symbols such as frame headers and synchronization frames (used by the receiver to detect the signal) are added to the output of the RS encoder. The final framed output remains a serial bitstream.

The serial-to-parallel conversion module maps the bitstream into I and Q signals, which are then fed into the π/4QPSK module. π/4QPSK modulation uses eight different carrier phases to represent digital information, as shown in the signal space diagram in **Figure 3**. The eight phase states are $\{0,\ \pi/4,\ \pi/2,\ 3\pi/4,\ \pi,\ -3\pi/4,\ -\pi/2,\ -\pi/4\}$. These phase states can be viewed as two sets of QPSK signal phases, with the black solid points forming one set $\{0,\ \pi/2,\ \pi,\ -\pi/2\}$ and the white hollow points forming the other set $\{\pi/4,\ 3\pi/4,\ -3\pi/4,\ -\pi/4\}$. The phase state transitions in π/4QPSK can only occur between these two QPSK phase groups, meaning the maximum phase jump in π/4QPSK is $3\pi/4$. The differentially encoded data is then up-converted and mixed before being sent to the DAC, resulting in the final transmitted π/4QPSK modulated signal.

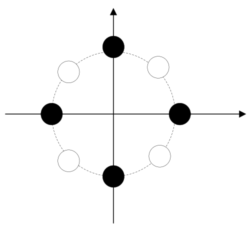

Figure 3. π/4QPSK constellation diagram

3. FPGA implementation of the transmitter

This paper uses the Xilinx Spartan6 FPGA hardware platform to implement the π/4QPSK modulation transmitter. The FPGA platform, equipped with hardware description languages, implements signal processing algorithms by building digital circuits, offering real-time processing, parallelism, and flexibility, making it suitable for communication systems. The FPGA platform can support multiple modulation schemes, data rates, and flexible waveform configurations. The π/4QPSK modulation transmitter scheme designed in this paper was verified through simulation on the Vivado platform and waveform debugging using Chipscope. The simulation and measured waveform analyses are as follows:

As shown in **Figure 4**, the simulation results of the RS encoder's input and output are presented. The input data (datain) and output data (dataout) are shown, with the encoded results sequentially appended to the input data. The encoded output is available in the next clock cycle, verifying the low encoding delay. The input data length is 16 bytes, and the encoded output length is 32 bytes, achieving RS(32,16) encoding.

As shown in **Figure 5**, the simulation results of the framing module are presented. The output (data_bit) is the framed bitstream, combining the synchronization frame, frame header, and encoded data into a serial bitstream.

As shown in **Figure 6**, the simulation results of the modulation module are presented. The I and Q signals are the differentially encoded outputs, and the I and Q signals after shaping filtering and up-conversion are shown. The final synthesized I and Q signals are output as the modulated signal. The bitstream enters the modulation module, is converted into I and Q signals through serial-to-parallel conversion, and then differentially encoded. The differentially encoded I and Q signals undergo shaping filtering. A DDS generates a 25 MHz carrier, which is multiplied with the shaped I and Q signals. The resulting I and Q signals are summed to produce the final modulated output signal.

The final modulated $\pi/4$QPSK signal is output through a DAC and looped back into the hardware board. The waveform after A/D processing is shown in **Figure 7**, which represents the observable measured data waveform. The waveform is clear with minimal glitches, demonstrating that the designed transmitter hardware implementation meets practical application requirements.

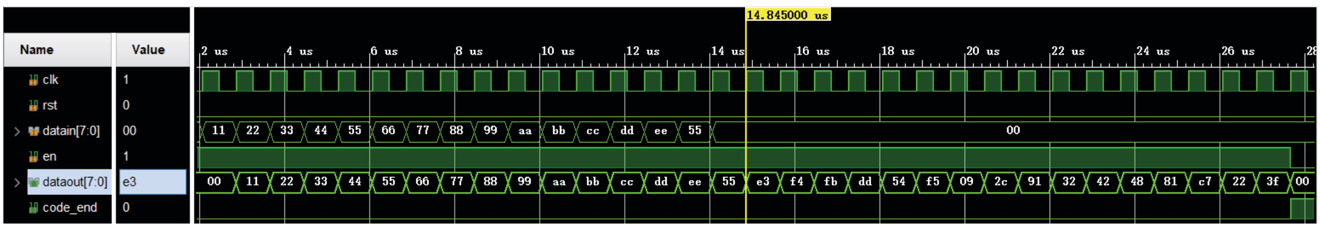

Figure 4. RS encoding simulation results

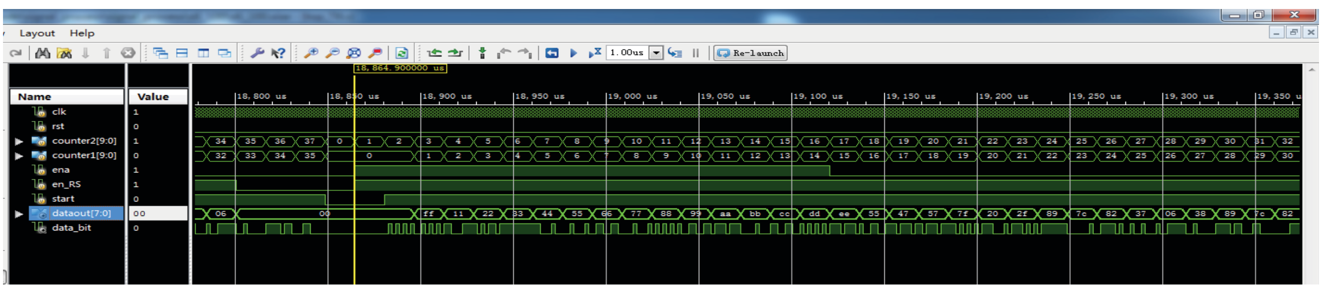

Figure 5. Framing module output simulation results

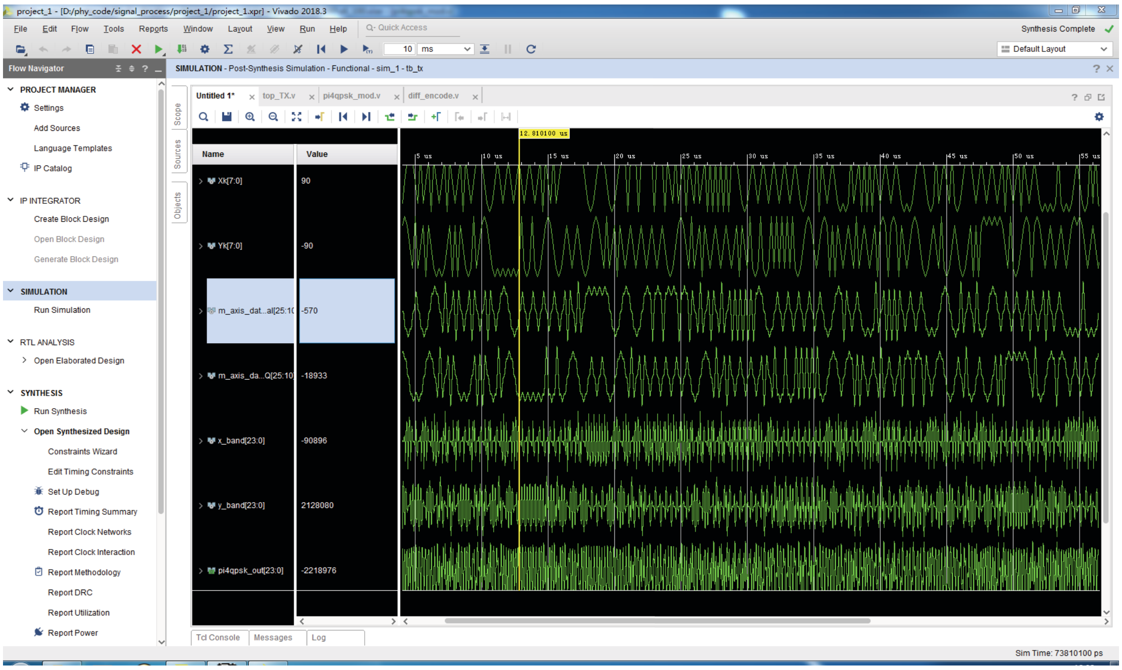

Figure 6. Modulation output simulation results

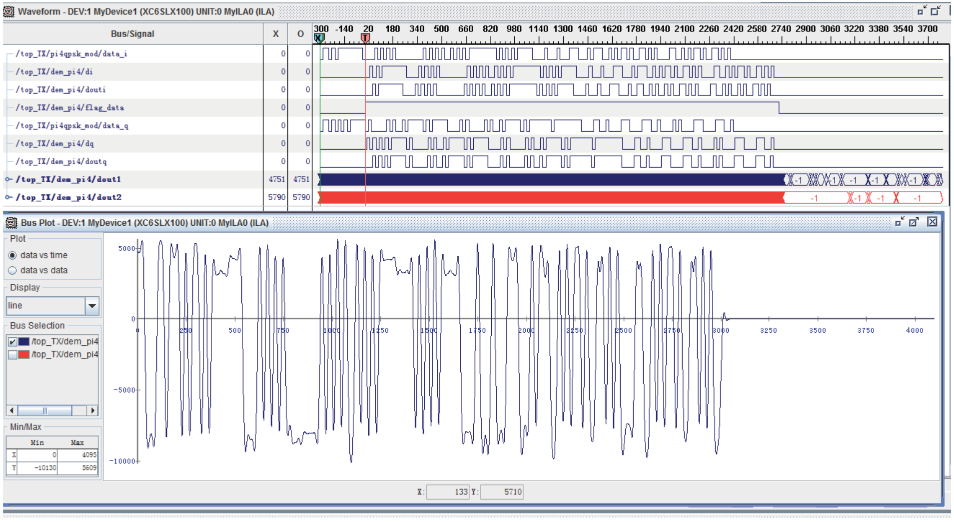

Figure 7. Measured waveform after D/A output and A/D input

4. Conclusion

This paper primarily introduces the simulation and FPGA implementation of the π/4QPSK modulation system, specifically including the implementation of RS encoding, framing, serial-to-parallel conversion, differential encoding, and up-conversion. The specific implementation methods for each module are discussed and the proposed methods are realized on the FPGA platform. By utilizing debugging software to observe the output signals and the outputs of each module, satisfactory results have been achieved.

Disclosure statement

The authors declare no conflict of interest.

Author contributions

Conceptualization: Yunjie Yuan
Investigation: Yuhui Chu, Xiaolei Cai
Formal analysis: Yanyan Wang, Peijie Yin
Writing: Yuhui Chu

References

[1] Yuan Y, 2012, Software Radio-Based π/4-DQPSK Modulation and Demodulation System, thesis, Huaqiao University.

[2] Jiang N, 2005, Research on Digital π/4-DQPSK Modulation and Demodulation and FPGA Implementation, thesis, University of Electronic Science and Technology of China.

[3] Zhou G, 2019, Design of NC-OFDM System Based on π/4QPSK for Ship Communication Systems. Electronic Measurement Technology, 42(23): 128–131.

[4] Liu Z, Sun Y, Liu Y, et al., 2016, Design and Implementation of π/4QPSK Based on Multi-Channel Synthesis. Journal of China Academy of Electronics and Information Technology, 11(06): 642–648.

[5] Zhou F, Zheng L, Tao W, 2017, Digital Demodulation of π/4-QPSK Signals Based on Cross-Product Frequency Discriminator. Telemetry and Remote Control, 38(02): 49–55.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Electrochemically Driven Nickel-Catalyzed Phenol Synthesis via Sustainable Oxygen Atom Transfer from Nitrous Oxide

Qiyu Zhang[1], Wenbin Xie[2]*, Yasser M. A. Mohamed[3], Hossam A. El Nazer[3], Ahmed A. Elnazer[4], Yinghui Han[1]*

[1]College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 101408, China
[2]College of Materials Science and Opto-Electronic Technology, University of Chinese Academy of Sciences, Beijing 101408, China
[3]Photochemistry Department, National Research Centre, Dokki, Giza, P. O. 12622, Egypt
[4]Geological Sciences Department, National Research Centre, El Buhouth St., Dokki, Giza, P. O. 12622, Egypt

Corresponding author: Yinghui Han, hanyinghui@ucas.ac.cn; Wenbin Xie, xiewenbin@ucas.ac.cn

Abstract: The valorization of nitrous oxide (N_2O) as an oxygen atom donor presents an attractive opportunity for green chemistry applications, leveraging both its industrial abundance and thermodynamically favorable oxidation potential. However, practical implementation has been constrained by the inherent kinetic inertness and poor coordinating ability of N_2O. While prior studies achieved N_2O-mediated conversion of aryl halides to phenols, such transformations necessitated stoichiometric chemical reductants and elevated pressure (2 atm), posing challenges in operational safety and process scalability. This study focuses on an electrochemical strategy that enables efficient oxygen atom transfer under ambient pressure through controlled current application. This methodology facilitates the selective transformation of aryl iodides to phenols without external reducing agents, establishing an environmentally benign synthetic pathway. By replacing traditional chemical reductants with electrons as the sole reducing equivalent, our approach addresses critical sustainability challenges in aromatic oxygenation chemistry while maintaining operational simplicity under mild conditions.

Keywords: Nitrous oxide; Electrochemical synthesis; Aryl iodides; Revalorization; Nickel-catalysis

Online publication: March 21, 2025

1. Introduction

The development of efficient valorization strategies for nitrous oxide (N_2O) is of great significance due to its potential as an O-atom donor. N_2O possesses three key attributes in this regard: (i) thermodynamically favorability as an oxidant, (ii) a benign oxygen transfer capability, and (iii) abundance as an industrial byproduct. However, its practical implementation remains hindered by intrinsic limitations including kinetic inertness, minimal dipole moment, and weak coordination to metal catalysts, which collectively impede activation under mild conditions, as shown in **Figure 1a** [1]. While previous advances by Cornella *et al.* demonstrated N_2O-mediated hydroxylation of aryl halides under mild thermal conditions, these protocols relied on stoichiometric chemical reductants and required pressurized N_2O (2 atm), introducing significant safety considerations and scalability constraints (**Figure 1b**) [2–4]. To address these limitations, we propose a novel electrochemical strategy employing constant current as the sole electron source to drive the catalytic cycle in this work. This methodology enables selective conversion of aryl iodides to phenolic products under ambient pressure conditions, effectively bypassing the need of exogenous reducing agents (**Figure 1c**).

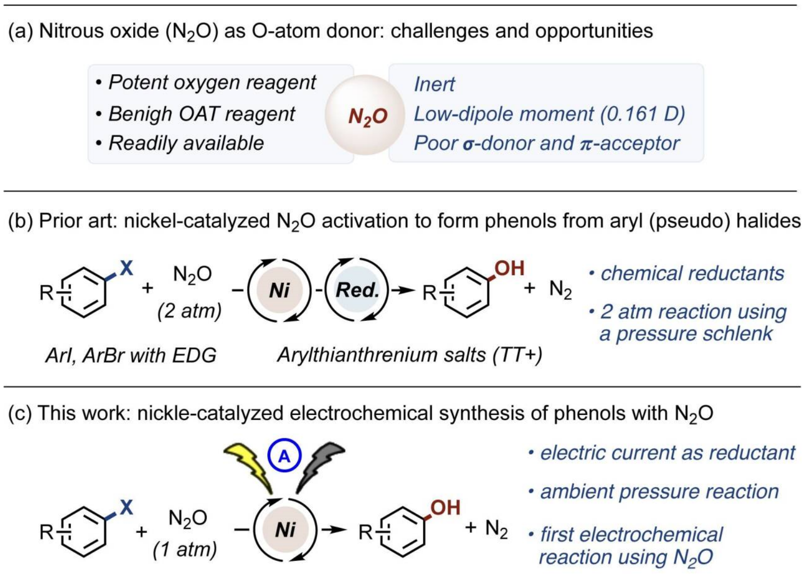

Figure 1. Nickel-catalyzed electrochemical synthesis of phenols with nitrous oxide. (a) Challenges and opportunities of using nitrous oxide as an O-atom donor. (b) Prior work. (c) This work.

2. Experimental methods

The electrocatalytic reaction setup, which is shown in **Figure 2**, is assembled [5].

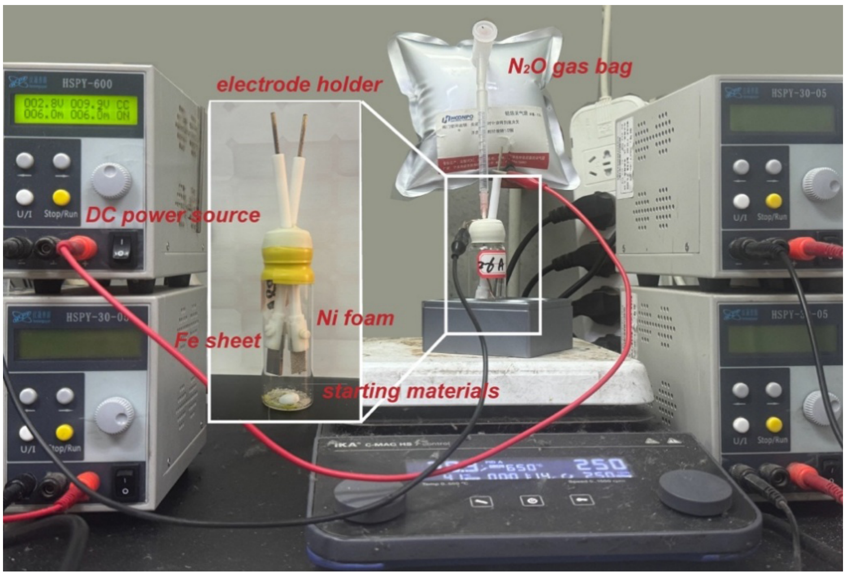

Figure 2. Easily hand-made electrochemical cell.

A commercially available nickel foam electrode is cut into pieces about 1.5 cm by 3 cm and a custom-made iron sheet of the same size is fixed with electrode clips. In a glovebox, the reaction substrate, catalyst, ligand, and additives are weighed into a 30 mL glass vial. After piercing the rubber septum with the two electrode holders, the vial is removed from the glovebox. The reaction vial is then connected to a N_2O gas cylinder and a Schlenk line via a T-shaped connector. The vial is evacuated and filled with N_2O for three times, followed by injecting 6 mL of DMA solvent that has been pre-saturated with N_2O for 30 minutes. The reaction system is stirred at room temperature for 15 minutes before connecting the circuit. After the reaction is complete, the reaction mixture is subjected to GC analysis.

3. Results and discussion

After a series of experiment, it is found that under a N_2O atmosphere, **1a,** undergoes activation and C-O bond formation at room temperature in the presence of a Ni(II) pre-catalyst bearing ligand **L2** and a constant reduction current in 0.1 M DMA bubbled with N_2O for 30 min, affording the desired product in moderate yield (**Table 1**, Entry 1). The reaction is shown in **Figure 3**. Notably, this reaction employs N_2O at 1 atmosphere, thereby avoiding the safety risks associated with 2 atm of pressure.

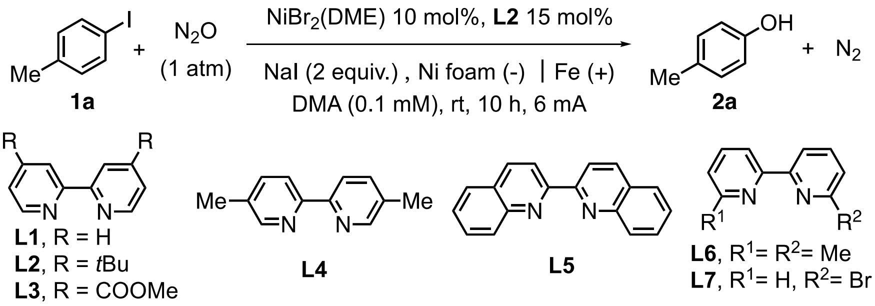

Figure 3. Reaction equation and screened ligands

Table 1. Reaction optimization with substrate 1a*

Entry	Deviation from standard conditions	Yield (%)[b]
1	None	47
2	Fe (+) | Ni (-)	N.D.
3	$NiCl_2$	6
4	L1	21
5	L3	N.D.
6	L4	9
7	L5	12
8	L6	7
9	L7	18
10	$ZnBr_2$	27

* Reaction conditions: 1a. (0.6 mmol), $NiBr_2$(DME) (10.0 mol%), **L2** (15 mol%), NaI (2.0 equiv.) and DMA (6 ml) in an undivided cell with a nickel foam electrode as cathode and an iron sheet as a sacrificial anode; b. The yields were determined by GC-MS with anthracene as the internal standard.

As expected, the choice of electrodes is critical to the success of the reaction. Various electrode combinations (Entry 2) were screened and ultimately, a nickel foam cathode and an iron sheet anode were selected as the optimal combination. Next, different nickel catalysts and ligands (Entry 3-9) were evaluated. It is found that the highest yield was achieved when the substituent at the 4-position was electron-donating and sterically hindered. Finally, different additives (Entry 10) were screened and sodium iodide was identified as the optimal additive, which is consistent with its reported unique role in reductive coupling catalysis [2].

Based on Cornella *et al.*'s experimental studies and Baldinelli *et al.*'s calculation studies, a plausible mechanism for this Ni-catalyzed synthesis of phenols with N_2O via electrochemical reduction is proposed in **Figure 4** [2, 6]. Initially, oxidative addition of Ni(0) to the 1a generates a Ni(II) complex. Next, the cathode donates two electrons to the pyridine-based ligand coordinated to the nickel center, significantly lowering the activation energy barrier for the insertion process, thereby making the entire catalytic cycle kinetically feasible. Then, O insertion and N_2 extrusion tend to occur in a concerted manner. Finally, the catalytic cycle is closed through cathodic reduction, regenerating the Ni(0) species.

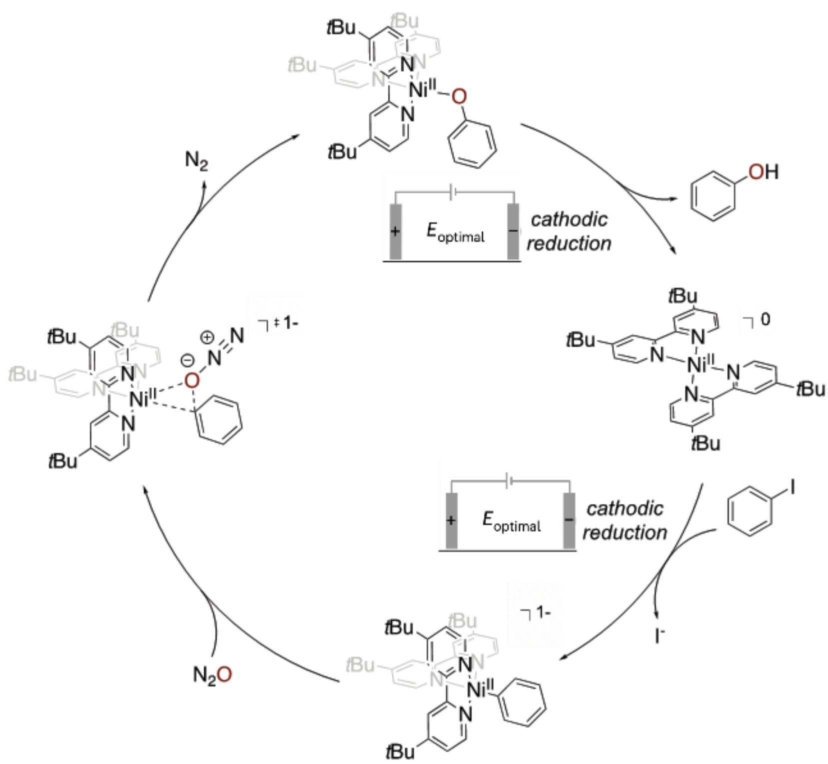

Figure 4. Proposed catalytic cycle.

4. Conclusions

In summary, from this study, the electrochemical Ni(0)-catalyzed construction of C–O bonds from aryl iodides was achieved, marking the first example of this type of transformation. Significantly, this catalytic paradigm harnesses electrons to drive the reaction forward rather than conventional stoichiometric reagents， fundamentally reconfiguring the activation pathway for N_2O utilization. Our findings establish a novel conceptual framework for implementing this environmentally problematic gas in value-added synthetic transformations. Ongoing efforts are focused on delving deeper into the underlying mechanism and broadening the range of compatible substrates. This work ultimately bridges the gap between sustainable electrochemistry and greenhouse gas valorization in synthetic organic chemistry.

Funding

National Natural Science Foundation of China (Project No.: 52320105003, Project No.: 52303019), the CAS-ANSO Co-funding Research Project (Project No.: CAS-ANSO-CF-2024), the National Key R&D Program of China (Project No.: 2023YFC3707201), the Fundamental Research Funds for the Central Universities (Project No.: E3ET1803), and China Postdoctoral Science Foundation (Project No: 2024T170904)

Disclosure statement

The authors declare no conflict of interest.

Author contributions

Conceptualization and funding: Yinghui Han

Investigation, data analysis, and writing: Qiyu Zhang

Guidance: Wenbin Xie

Review and editing: Yasser M. A. Mohamed, Hossam A. El Nazer, Ahmed A. Elnazer

References

[1] William BT, 2010, Binding and Activation of N2O at Transition-Metal Centers: Recent Mechanistic Insights. Angew Chem Int Ed, 49: 1018-1024. http://doi.org/10.1002/anie.200905364.

[2] Vaillant FL, Calbet AM, Gonzalez-Pelayo S, et al., 2022. Catalytic Synthesis of Phenols with Nitrous Oxide. Nature, 604: 677-683. http://doi.org/10.1038/s41586-022-04516-4.

[3] Ni SY, Cornella J, 2023, Catalytic Hydroxylation of Arylthianthrenium Salts with Nitrous Oxide. Tetrahedron, 145: 133602. http://doi.org/10.1016/j.tet.2023.133602.

[4] Zhang HH, Rodrigalvarez J, and Martin R, 2023, C(sp^2)-H hydroxylation via catalytic 1,4-Ni migration with N_2O. J Am Chem Soc, 145: 17564-17569. http://doi.org10.1021/jacs.3c07018.

[5] Jiao KJ, Li ZM, Xu XT, et al., 2018, Palladium-Catalyzed Reductive Electrocarboxylation of Allyl Esters with Carbon Dioxide. Org Chem Front, 5: 2244-2248. http://doi.org10.1039/c8qo00507a.

[6] Baldinelli L, Belanzoni P, Bistoni G, et al., 2024, Mechanism of Nitrous Oxide Activation in C(sp2)-O Bond Formation Reactions Catalyzed by Nickel Complexes. J Am Chem Soc, 146: 6016-6024. http://doi.org/10.1021/jacs.3c12922.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Variable Speed Limit Plate Signs Based on Depth Recognition Control

Xiaoxi Liu[1]*, Libin Zhang[1], Xinrui Wang[2], Hanqi Tang[2], Shuming Zhang[1]

[1]School of Civil Engineering and Transportation, Northeast Forestry University, Harbin 150040, China
[2]College of Computer and Control Engineering, Northeast Forestry University, Harbin 150040, China

Corresponding author: Xiaoxi Liu, lxx2335691039@163.com

Abstract: Predictive control (PC) is an advanced control algorithm, which is widely used in industrial process control. Among them, model-based predictive control (MPC) is an important branch of predictive control. Its basic principle is to use the system model to predict future behavior and determine the current control action by optimizing the objective function. Based on the algorithm combined with three different sections using deep learning technology to identify vehicles and output the optimal speed limit, to achieve the effect of traffic flow optimization.

Keywords: Traffic safety; Variable speed limit; MPC algorithm

Online publication: April 2, 2025

1. Research background

As a core part of the transportation infrastructure, the highway network has a perfect facility system and an efficient, fast, and convenient transportation service capacity. In daily travel, the highway system plays an important role and becomes a key indicator to assess the level of economic development. According to the latest data, as of July 2022, China's total highway mileage was about 461,000 km, and by the end of June 2023, the total number of vehicles had increased to about 426 million vehicles. Through the comparative analysis of historical data, it can be seen that the development of highways in China shows a steady growth trend. However, compared with the growing traffic demand, the speed of highway construction and expansion is relatively backward. This lag leads to the increasing contradiction between traffic demand and road supply, which then becomes the root cause of traffic congestion.

The traffic congestion phenomenon has obvious temporal and spatial characteristics, especially in peak hours and specific road sections. The study shows that when the traffic demand of the main highway exceeds the designed capacity, the bottleneck section will appear and the capacity will decrease, leading to vehicles queuing in the upstream section. As traffic demand continues to grow, congestion on bottleneck sections may spread upstream, further exacerbating traffic congestion. This not only reduces the capacity and service level of roads but also may cause serious problems such as traffic accidents, environmental pollution, and waste of resources.

To solve the problem of traffic congestion, the traditional solution strategies mainly include increasing the construction of traffic lines and related facilities. However, these measures often require significant funding input and high late-stage maintenance costs. In contrast, the global optimization regulation strategy is more advantageous. This strategy does not require a large amount of capital investment, and the later maintenance cost is low, and can fundamentally solve the problem of traffic congestion. Therefore, it is particularly important to find an effective regulation method to maximize the use of highway line resources, reduce the loss, and ensure smooth traffic.

In the field of global optimization control, advanced methods such as the model prediction control (MPC) method and urban traffic signal timing optimization have been widely used. The main advantage of the MPC approach is its ability to adapt to changing traffic flows and consider future traffic conditions, thus providing a more flexible and precise basis for traffic management decisions.

Based on the above analysis, this paper proposes a variable control plate optimization design scheme based on real-time road condition change. The scheme aims to realize the optimal regulation of highway traffic flow by changing the organization of traffic flow in specific sections and peak hours, to improve travel efficiency and smoothness. This research can not only help to alleviate the traffic congestion problem but also provide the scientific basis and technical support for future highway traffic management.

2. Local and international research

Hoogen *et al.* demonstrated the effectiveness of variable speed limit control in reducing speed dispersion and reducing the frequency of traffic waves [1]. Zackor *et al.*, through the comparative analysis of the German expressway, indicated that the variable speed limit control can improve the road capacity by about 5% [2]. Ulfarsson *et al.*, based on U.S. highway data, found that variable speed limit control had significant improvement when speed dispersion was high [3].

Bertini *et al.* compared the data before and after implementing variable speed limit control on an Australian expressway, which confirmed the positive role of this control strategy in reducing travel time [4]. Harbord *et al.* evaluated the implementation effect of variable speed limit control on U.K. highways and found that it not only improved the road capacity but also enhanced the comfort of passengers [5]. A study by Kwon *et al.* on highway construction areas in the United States showed that variable speed limit control reduced speed dispersion by 25% during peak traffic demand hours [6]. Papageoriou *et al.* made an in-depth analysis of the data of variable speed limit control on a European expressway and found that this method realized the stability of traffic flow and improved traffic efficiency by increasing the critical density of road sections [6]. Jonkers *et al.* test on the Dutch highway showed that the application of variable speed limit control can effectively reduce the occurrence of traffic waves [7].

Rivey found that the variable speed limit control implemented on French highways significantly reduced the incidence of traffic accidents [8]. Hoogendoorn *et al.* analyzed the traffic conditions after the implementation of variable speed limit control on Dutch expressways and found that the control measures improved the efficiency of traffic circulation in the control section by about 4% [9,10]. In terms of model research, Hegyi *et al.* improved the expected speed calculation equation of the METANET model and optimized the operation efficiency of traffic flow by adjusting the speed limit value in the upstream speed limit area of the bottleneck section [11,12].

In China, the variable speed limit control technology was first applied to real roads in 1990. Based on the coil data, the speed limit value of 60 km/h, 80 km/h, and 100 km/h is displayed through seven variable speed limit sign boards, corresponding to the traffic state of congestion, and free flow respectively. In recent years, the practical application of variable speed limit control has been conducted on some expressways in China. In 2010, a variable speed limit control system was installed in the 10 km section of Hang Expressway. By analyzing the data collected within half a year, it was found that the variable speed limit control could significantly reduce the dispersion degree of the speed [13,14].

3. Design principle

3.1. Design ideas

The idea is to explore the research background and importance of variable rate-limiting methods and to evaluate the strengths and disadvantages of existing research models. Based on the existing literature, we construct a predictive control model and plan the technical paths and methods of the study. The operation characteristics of expressways are deeply analyzed, especially the traffic flow characteristics during peak hours and special sections. The reduced traffic operating capacity and the variable speed limit control are explained and discussed in detail to assess its potential impact on traffic flow. Combined with the predictive control model (MPC) algorithm and constraints, a predictive control model is proposed to provide theoretical support for the subsequent experiments to achieve the goals of this study. Define a cost function: The cost function is the sum of the squares of the differences between the speed limit and the target speed. Specifically, the goal is to minimize the square sum of vehicle density and speed differences, aiming to optimize traffic flow control and ensure that the vehicle travels under specified speed limits and as close to the target speed as possible to improve the fluency and efficiency of traffic flow. By minimizing the objective function, the optimal control strategy can be determined to optimize the traffic flow. The model parameters were adjusted, and the optimal rate-limit value was determined.

3.2. Study methods

3.2.1. Methods to be adopted

(1) Literature: Through consulting relevant papers, books, pictures, and other materials, understand the current development of deep learning identification technology and OpenCV visual database processing and analysis information, and learn the advanced technology used in related research projects.

(2) Investigation and analysis: To investigate the application status of the existing variable speed limit plates locally and internationally, and to analyze their advantages and disadvantages.

(3) Experimental method: Using the existing conditions, make a prediction model based on deep learning recognition technology, carry out feasibility verification, eliminate other interference factors, and further improve the determined overall research scheme.

(4) Expert consultation: Consult and communicate with professional teachers and professors about the technical problems arising in the research and development process, to gain more experience and inspiration, and speed up the research and development process of group projects.

(5) Summary method: Summarize the problems encountered in MPC and the development and innovation of the prediction and control system through text, photos, data, and other forms, and make the corresponding feasibility proof, to get a complete set of variable speed limit system based on deep learning recognition technology.

3.2.2. Technical route

Firstly, through an extensive review of relevant literature, this research focuses on the application of deep learning recognition technology and OpenCV visual library in information processing and analysis. This helps the research team to quickly develop a research plan that is both reasonable and feasible. Secondly, a variable rate-limiting control system is designed based on deep learning and OpenCV image processing technology. The functions intended to be realized by the system include the identification of the vehicle position and the measurement of the movement speed. Thirdly, this study also established the model and preliminarily completed the design of the overall experimental structure. Finally, the experiment is conducted through the simulation scene to identify the shortcomings of the system in the current stage, and the system is further optimized and improved according to the experimental data, to develop a relatively mature variable rate-limiting control system driven by deep learning.

4. Innovative features

4.1. Vehicle identification and positioning based on deep learning

Design and train deep learning models suitable for vehicle recognition, such as convolutional neural networks (CNN) or pre-trained models (such as ResNet, YOLO, etc.), which are trained through a large-scale vehicle image data set to improve the accuracy and robustness of vehicle recognition.

4.2. Traffic flow analysis and variable speed limit strategy

Accurately judge the number and density of vehicles in each lane, and realize the real-time monitoring and analysis of the traffic flow through the analysis of the traffic flow situation and road conditions. According to the results of the traffic flow analysis and combined with the actual situation of the road, the variable speed limit strategy is designed and implemented, and the goal of traffic fluency and road safety is achieved by adjusting the driving speed of the vehicle.

4.3. Make the models and perform the experiments

A traffic signal light model based on depth identification technology is made to simulate the operation of vehicles under variable speed limits and test and verify the feasibility of the scheme.

5. Application prospects

The technology uses a high-resolution camera, which can accurately capture road images, providing a rich data basis for subsequent traffic monitoring. At the same time, the technology also combines deep learning models, such as convolutional neural networks (CNN), the advanced algorithms that can efficiently process and analyze images to accurately identify vehicles on the road and monitor traffic flow in real time.

To meet the processing needs of large amounts of data, the technology also integrates advanced data processing units, which have strong computing power and can achieve fast data processing and decision-making ability, ensuring that the system can make accurate judgments of traffic conditions in a short time.

Additionally, the technology is equipped with variable information signs (VMS), which can display speed limits in real-time and provide timely road condition information to the driver. Through the wireless communication module, the technology can communicate with the central traffic management system in real-time, and realize information sharing and collaborative processing [15].

The technology has many advantages. Firstly, it is highly flexible and can dynamically adjust the speed limit value according to real-time traffic conditions, weather conditions, or special events to ensure road safety and smooth flow. Secondly, the technology helps to improve road safety and reduce the occurrence of traffic accidents. Thirdly, it can also optimize the traffic flow, reduce the congestion phenomenon, and improve the efficiency of road use [16]. Fourthly, the technology can also help to reduce vehicle emissions, reduce environmental pollution, and achieve green travel. Finally, the system has a high degree of automation, low maintenance cost, and convenient for daily management and maintenance.

6. Conclusion

While the initial investment in the technology may be high, the long-term benefit is significant. By improving traffic efficiency and safety, and reducing the costs caused by accidents and congestion, the technology can bring considerable economic benefits to enterprises and society. In addition, the government's support policies for traffic safety and environmental protection may also provide additional subsidies or incentives for the promotion of the technology to further promote the adoption and application of the technology.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Hoogen EVD, Smulders S, 1994, Control by Variable Speed Signs: Results of the Dutch Experiment. Seventh International Conference on "Road Traffic Monitoring and Control", 1994: 145–149.

[2] Zackor H, 1979, Self-sufficient control of speed on freeways. Proceedings of the International Symposium on Traffic Control, 2: 226–249.

[3] Ulfarsson GF, Shankar VN, Vu P, 2005, The Effect of Variable Message and Speed Limit Signs on Mean Speeds and Speed Deviations. International Journal of Vehicle Information and Communication Systems, 1(1–2): 69–87.

[4] Bertini RL, Boice S, Bogenberger K, 2005, Using ITS Data Fusion to Examine Traffic Dynamics on a Freeway with Variable Speed Limits. 2005 IEEE Intelligent Transportation Systems, 2005: 1006–1011.

[5] Zeng Y, 2018, Research on Intelligent Vehicle Target Recognition and Learning Control Method based on Deep Neural Network, thesis, National University of Defense Technology.

[6] Papageorgiou M, Kosmatopoulos E, Papamichail I, 2008, Effects of Variable Speed Limits on Motorway Traffic Flow. Transportation Research Record, 2047(1): 37–48.

[7] Jonkers, E, Wilmink IR, Stoelhorst H, et al., 2011, Results of Field Trials with Dynamic Speed Limits in the Netherlands: Improving Throughput and Safety on the A12 Freeway. 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2011: 2168–2173.

[8] Rivey F, 2010, Evaluation of the Dynamic Speed Limit System on the A13 motorway in France. Presented at Easy-Way Annual Forum, Lisboa.

[9] Hoogendoorn, Edara P, Sun C, 2013, Operational Analysis of Freeway Variable Speed Limit System: Case Study of Deployment in Missouri. Journal of Intelligent Transportation Systems Technology Planning and Operations, 19(4): 1–14.

[10] Saha P, Young R, 2007, Weather-Based Safety Analysis for the Effectiveness of Rural VSL Corridors. Transportation Research Board 93rd Annual Meeting, 2014: 2293.

[11] Hegyi A, Schutter BD, Hellendoorn J, 2005, Optimal Coordination of Variable Speed Limits to Suppress Shock Waves. IEEE Transactions on Intelligent Transportation Systems, 6(1): 102–112.

[12] Allaby P, Hellinga B, Bullock M, 2007, Variable Speed Limits: Safety and Operational Impacts of a Candidate Control Strategy for Freeway Applications. IEEE Transactions on Intelligent Transportation Systems, 8(4): 671–680.

[13] Li J, 2019, Research on the Robot Control Algorithm Based on Visual Feedback, thesis, Hangzhou Dianzi University.

[14] Li Z, 2014, Variable Speed Limit Control Technology for Fast Roads, thesis, Nanjing Southeast University.

[15] Papageorgiou M, Blosseville J, Hadj-Salem H, 1990, Modelling and Real-time Control of Traffic Flow on the Southern Part of Boulevard Peripherique in Paris: Part I: Modelling. Transportation Research Part A: General, 24(5): 345–359.

[16] Zhang K, 2016, Network Application Identification and Traffic Control Based on Deep Packet Detection, thesis, Guangdong University of Technology.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Application Strategies of Artificial Intelligence and Big Data Technology in Computer Monitoring and Control

Yumin Yuan, Xiyuan Li*

Chongqing University of Science and Technology, Chongqing 401331, China.

***Corresponding author:** Xiyuan Li, 3171005983@qq.com

Abstract: This article focuses on the current computer monitoring and control as the research direction, studying the application strategies of artificial intelligence and big data technology in this field. It includes an introduction to artificial intelligence and big data technology, the application strategies of artificial intelligence and big data technology in computer hardware, software, and network monitoring, as well as the application strategies of artificial intelligence and big data technology in computer process, access, and network control. This analysis aims to serve as a reference for the application of artificial intelligence and big data technology in computer monitoring and control, ultimately enhancing the security of computer systems.

Keywords: Computer monitoring and control; Artificial intelligence technology; Big data technology; Hardware and software; Network security

Online publication: March 28, 2025

1. Introduction to artificial intelligence and big data technology

With the comprehensive application and popularization of modern computer network technology, computer monitoring and control have become crucial to ensuring the security of network information. In specific computer monitoring and control tasks, artificial intelligence (AI) and big data technology exhibit significant advantages [1]. Based on this, researchers first need to fully understand AI and big data technology. Then, combining current computer monitoring and control requirements, they can explore the application of AI and big data technology to fully leverage their advantages.

Artificial intelligence (AI) technology is an important branch of modern computer science. It is a cutting-edge technology with comprehensive characteristics and its basic application goal is to equip computer systems with human-like wisdom and capabilities, including perception, learning, reasoning, decision-making, and task execution abilities. Currently, key components of AI technology include machine learning, deep learning, natural language processing, computer vision, and knowledge graph techniques [2]. With advantages such as high efficiency and intelligence, this technology has been widely applied in many current fields.

Meanwhile, big data technology is an emerging technology in the modern computer information field. Its basic application purpose is to quickly acquire valuable information from massive amounts of diverse data. Presently, the main components of big data technology consist of data collection, data processing, data storage, data analysis, and data visualization techniques [3]. Thanks to its high efficiency, precision, and visualization capabilities, big data technology has been extensively used in various current domains and has exhibited significant advantages supported by computer technology.

In the current field of computer monitoring and control, the integrated advantages of artificial intelligence and big data technology are mainly manifested in the following aspects. Firstly, the effective integration of the two can further enhance the computer's data processing capabilities, enabling timely and effective collection, integration, storage, management, and analysis of massive data in monitoring and control [4]. Secondly, effective integration can further improve the computer's data analysis and decision-making abilities, providing support for real-time data analysis and intelligent decision-making in monitoring and control. Thirdly, effective integration can enhance the computer's fault diagnosis and recovery capabilities, enabling timely and reasonable fault diagnosis and recovery during monitoring and control.

2. Main application strategies of artificial intelligence and big data technology in computer monitoring

For current computer monitoring, there are three main application directions of artificial intelligence and big data technology: hardware monitoring, software monitoring, and network monitoring. The following is an analysis of the main application strategies of artificial intelligence and big data technology in this monitoring work.

2.1. Application of artificial intelligence and big data technology in hardware monitoring

Hardware monitoring is a key aspect of computer security monitoring. Through reasonable monitoring of computer hardware such as power supply, hard disk, memory, and CPU, corresponding hardware failures or abnormalities can be detected in a timely manner and handled promptly to prevent hardware damage from adversely affecting the stable and safe operation of the computer. During specific monitoring, artificial intelligence technology is mainly used for real-time monitoring and analysis of various hardware states, and making scientific predictions about their subsequent operation [5]. For example, for hardware abnormalities obtained via sensors, technicians can use machine learning technology to establish a corresponding hardware failure prediction model, train and analyze its state data, and make timely and accurate predictions about upcoming failures. Big data technology is mainly used to quickly detect, analyze, and determine corresponding hardware failures, ensuring they are handled promptly to significantly reduce downtime due to failures. For instance, in hardware monitoring, technicians can use big data mining technology to promptly identify abnormal hardware operating parameters and analyze the causes of failures, subsequent development trends, and impact levels with the help of big data analysis technology [6]. This allows for timely detection and handling of hardware failures or abnormalities during computer operation, preventing further development of hardware failures that may lead to downtime and ensuring the safe and stable operation of the computer.

2.2. Application of artificial intelligence and big data technology in software monitoring

For current computer security monitoring, software monitoring is also an important aspect. In this aspect, artificial intelligence and big data technology can be applied in the following areas. The first is computer software performance monitoring and optimization. In this process, technicians can use big data technology to collect and analyze various performance index parameters during the actual operation of the software system to detect abnormal data in a timely manner.

For abnormal data, clustering algorithms in machine learning technology can be used to perform cluster analysis on similar events to quickly determine the type of software problem. The second is computer software fault diagnosis and repair. In this process, technicians can use big data technology to collect real-time data such as alarms and logs generated during software operation for analysis and learning by relevant algorithm models under artificial intelligence technology. This allows for timely detection and determination of data that deviates significantly from normal data, enabling rapid diagnosis of computer software faults.

For diagnosed faults, technicians can quickly develop reasonable repair plans by combining artificial intelligence and big data technology to achieve timely and effective repair of software problems [7]. The third is computer software security monitoring and protection. In this process, technicians can use big data technology to collect real-time security information, such as operating data and code information in the computer software system. With the support of machine learning technology, based on the characteristics of known viruses or vulnerabilities, they can analyze whether new operating data and code segments are consistent or similar to them, achieving timely and effective monitoring of software security. This allows for effective identification and prevention of various software viruses and intrusion detections.

2.3. Application of artificial intelligence and big data technology in network monitoring

In computer security monitoring, reasonable network monitoring is crucial. Based on this, technicians need to utilize current advanced artificial intelligence and big data technology to conduct real-time and effective monitoring of computer networks. For current computer network security monitoring, artificial intelligence and big data technology are primarily implemented through corresponding support software. The first is network traffic monitoring, including information such as its source, port, protocol, and destination. During specific monitoring, tools like Tcpdump and Wireshark can be used as primary support software. With the help of these tools, malicious and abnormal network traffic can be detected in a timely manner.

The second is network protocol monitoring, including protocol function, structure, and security. During monitoring, WinDump and the TCP/IP toolkit can serve as the main support software. With these tools, protocol vulnerabilities, security risks, and specific protocol attack behaviors in the network can be judged promptly. Finally, there is network attack monitoring, which includes various computer network and system attack behaviors. During specific monitoring, Suricata and Snort can be used as the main support software. With the help of these tools, various network attack behaviors can be detected and intercepted in a timely manner to prevent illegal attacks on the computer network.

3. Main application strategies of artificial intelligence and big data technology in computer control

From the perspective of current computer control work, there are also three main application directions of artificial intelligence and big data technology: process control, access control, and network control. During specific control operations, technicians need to combine actual situations and reasonably introduce and apply artificial intelligence and big data technology to ensure computer control effectiveness and enhance its security. The following are the main application strategies of artificial intelligence and big data technology in computer control.

3.1. Application of artificial intelligence and big data technology in process control

In the process of computer security control based on artificial intelligence and big data technology, process control is the primary application direction of these technologies. Process control refers to managing programs that are in the running stage in the computer system to ensure the security and stability of the entire computer system. Currently, state monitoring, resource management, and priority scheduling are the main components of computer process control. In the specific process of process control, Process Explorer and Process Monitor are the primary monitoring software supported by artificial intelligence and big data technology. With the help of these monitoring software, various program process states running in the computer system can be monitored and controlled in real-time, including monitoring their running time, memory usage, and CPU occupancy, as well as controlling the suspension and termination of processes.

During this period, the main application strategy of artificial intelligence and big data technology is to collect, analyze, and process massive amounts of computer process data to achieve timely and effective prediction, prevention, and control of system failures. For example, machine learning algorithms can be used to train historical process data of various programs in the computer and establish a corresponding failure prediction model based on this. Then, based on big data technology, various process state data can be collected and acquired in real-time. With the support of the established failure prediction model, timely and accurate predictions can be made about upcoming failures during program operation, and effective measures can be taken to repair and maintain abnormal programs. This allows reasonable control of various program processes in the computer, effectively reducing the failure rate of the computer system caused by program failures, making the overall computer system operation safer and reliable.

3.2. Application of artificial intelligence and big data technology in access control

Access control is also a major application direction of artificial intelligence and big data technology in the field of computer security control. Currently, the main goal of computer access control is to manage the resources accessed by various processes or users within the computer system, restrict access to critical resources and sensitive data, and avoid unauthorized process or user access. This ensures the secure operation of the computer system. There are two types of computer access control technologies based on artificial intelligence and big data technology.

The first is identity verification technology, including facial recognition, fingerprint recognition, username and password recognition, etc. In this technology, big data technology can provide support for multi-source data collection (including basic user information, behavior data, environmental data, etc.), data integration and preprocessing, data analysis, and feature extraction. At the same time, with the help of machine learning algorithms, deep learning algorithms, artificial intelligence model training, and optimization in artificial intelligence technology, the accuracy of identity verification can be further improved, providing support for computer access control [8].

The second is access control technology based on user permissions, including different levels of permissions such as reading, writing, modifying, and executing network information by users. When applied specifically, intelligent technologies such as rule engines, machine learning, and deep learning can achieve reasonable control over simple rules, permission grants, and resource access to prevent unauthorized users from accessing important data or sensitive information. The rational application of big data technologies such as data collection, data integration, data cleaning, and preprocessing can provide support and assistance to existing artificial intelligence technologies, improving the efficiency and accuracy of their analysis and control to achieve good access control effects.

3.3. Application of artificial intelligence and big data technology in network control

For modern computer security control, network control is one of the most critical control elements. Its basic goal is to monitor and control computer networks to maintain their safe and stable operation. Based on this, technicians should fully recognize the importance of network control and reasonably introduce artificial intelligence and big data technologies to provide more advanced technical support for this work [9]. In terms of current network control, the main application methods of artificial intelligence and big data technologies are to analyze and process massive network data in a reasonable way to achieve prediction and prevention of various network security issues. In this process, technicians can use machine learning technology to train historical network data, build a computer network attack prediction model based on this, and use this model to conduct real-time monitoring of all network traffic data to achieve accurate prediction of various network attacks [10]. Then, based on the prediction results and supported by artificial intelligence and big data technologies, in-depth research on prevention and response measures can be conducted to timely develop reasonable prevention and control plans so that computer network attacks can be effectively controlled.

4. Conclusion

In summary, artificial intelligence and big data technologies are key technologies in the modern network information age. These technologies not only provide strong support for the intelligent, efficient, and mass data operation of computer information systems but are also key technologies to ensure computer operation safety. Based on this, researchers and technicians need to strengthen related technology research, combine the current actual development of computers and their future development trends, reasonably introduce and apply artificial intelligence and big data technologies to give full play to their technical advantages, to ensure the safety of computer hardware, software, and its network environment and prevent unnecessary risks. This provides technical support for the safe and stable operation of computers.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Xue L, Xin T, 2024, Application of Big Data and Intelligent Control Technology in Computer Network Systems. Integrated Circuit Applications, 41(4): 214–215.

[2] Yang Y, 2024, Research on Computer Network Security Defense System Based on Artificial Intelligence Technology. Cyberspace Security, 15(4): 306–309.

[3] Wu Z, Yu Y, Liu S, 2024, Application of Condition Monitoring and Auxiliary Control System in Smart Substation.

Information Technology, 2024(1): 170–175 + 188.
[4] Yao Y, 2024, Computer Network Security Risks and Control Strategies in the Era of Big Data. China Public Security, 2024(1): 151–153.
[5] Wang L, 2024, Application of Big Data and Intelligent Control Technology in Computer Network Information Security Systems. Integrated Circuit Applications, 41(4): 292–293.
[6] Cong H, Wang H, Tao S, et al., 2023, Analysis of Computer Network Security and Risk Control Strategies. Integrated Circuit Applications, 40(10): 284–285.
[7] Yin K, 2022, Research on Computer Network Security Risk Control in the Context of Artificial Intelligence. Telecommunications Power Technology, 39(21): 177–179.
[8] Wang Q, 2023, Research on Security Control of Confidential Information Systems in Computer Networks. Information and Computers, 35(23): 193–195.
[9] Huang J, Cao M, Lin S, 2024, Design and Simulation of a Warehouse Monitoring and Warning System Based on Unity3D. Industrial Control Computer, 37(3): 70–72.
[10] Guo Q, 2024, Greenhouse Intelligent Control System Based on IoT Big Data. Computer Application Digest, 40(22): 66–68.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Motion Analysis of a New Final Coal Sorter Based on Multi-Body Dynamics

Qin-ci Xu*, Fu-hao Zhang, Fang-fang Zhang

Branch of Yankuang Shandong Research Institute of Coal Science and Technology, Zoucheng 273516, Shandong, China

Corresponding author: Qin-ci Xu, 17863817554@163.com

Abstract: Research has been conducted on dry coal selection technology to achieve efficient and clean utilization of coal and reduce water resource waste. To investigate the effective separation of 6–1 mm fine coal using this approach, the vibrating cascade sorter was examined, with particular attention given to its dynamic characteristics and motion behavior. Dry coal vibrating cascade sorter motion simulation experiments were carried out and the mechanical system motion characteristics of the vibrating cascade sorter were simulated and analyzed using multi-body dynamics software ADAMS, including the motion curve and spatial trajectory of the sorter body. Theoretical calculations of the sorter body's motion characteristics, including displacement amplitude, velocity amplitude, and acceleration amplitude, were compared with simulation results derived from multi-body dynamics. The comparison revealed a strong agreement between the computational model and analytical predictions, with a maximum deviation of less than 3.75%. The dynamic behavior of the vibrating cascade sorter at various rotational speeds was evaluated and contrasted against predictive models, with the highest discrepancy between the observed and predicted outcomes being less than 7.6%.

Keywords: Vibrating cascade sorter; Dry coal selection technology; ADAMS; Simulation; Motion characteristics

Online publication: April 2, 2025

1. Introduction

In China, over 60% of coal resources are distributed in the inland western regions with scarce water resources. Currently, wet coal washing technology is the main method, but it requires as much as 3 tons of water for each ton of coal processed. To reduce water waste, it is necessary to break through the technical difficulties of dry coal washing technology. Domestic and foreign scholars have carried out a large amount of research work in the field of dry coal washing desulfurization technology, including heavy medium separation, air separation, electrostatic separation and magnetic separation, etc.

Building on the technical architecture of the Clean Coal Center's research information management platform, this study integrates industrial software tools such as ANSYS, EDEM, and MATLAB to perform coupled screening simulation experiments. These simulations systematically examine the effects of various screening parameters on the looseness and stratification of coal seams while exploring the separation mechanism of the dry fine coal vibrating separation device. By combining vibration cascade separation theory with advanced numerical simulation methods, the study provides a comprehensive analysis aimed at improving the efficiency and effectiveness of fine coal dry separation processes.

2. Kinematics theory

The kinematic model of the vibrating cascade sorter is shown in **Figure 1**. Based on the kinematic model, displacement, velocity, acceleration, and other vibration parameters can be obtained. These parameters are natural properties of the vibrating cascade sorter and provide an important basis for evaluating its rationality and reliability [1]. During the operation of the vibrating cascade sorter, the heart shaft performs uniform circular motion while the sorting bed surface performs reciprocating motion similar to a straight line. The displacement, velocity, and acceleration of the sorting bed surface are constantly changing [2]. Therefore, it is of great significance to study the motion characteristics of the sorting bed surface. To study the motion characteristics of the bed surface, an absolute coordinate system OXY was established with the rotating center of the heart shaft as the origin O [3]. The X-axis is parallel to the bed surface and points in the direction of the material flow, while the Y-axis is perpendicular to the bed surface. According to the working principle of the vibrating cascade sorter, simplification and equivalent treatment were carried out on the heart shaft, eccentric sleeve, and spring plate. Meanwhile, e represents the eccentricity and ω represents the rotational speed of the heart shaft.

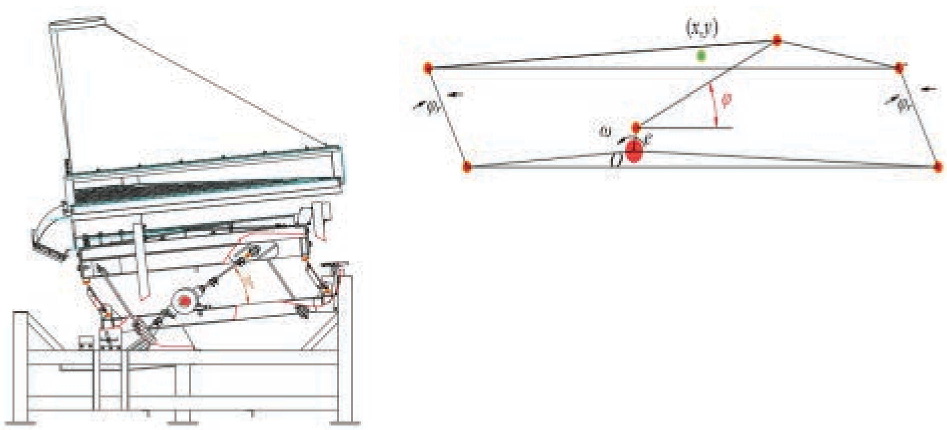

Figure 1. Schematic diagram of motion of vibrating cascade sorter

The vibrating screen of the sorting machine transmits the rotation of the eccentric wheel to the surface of the bed and the rising airflow causes the mixed material to move up and down and flip on the bed surface, causing the material on the bed surface to move towards the discharge end at a constant speed [4]. The material on the bed surface is given a certain amount of kinetic energy while being screened, so the difference in its mass results in different amounts of kinetic energy [5]. The mixed coal with a larger mass has greater kinetic energy, allowing it to overcome the frictional force given by the bed surface and the frictional force between particles to move to the right, while the coal powder with a smaller mass has less kinetic energy, causing them to move to the far left due to the interaction of forces [6]. As a result, the mixed coal is sorted by particle size and discharged from left to right at the outlet, ultimately completing the screening process. The main idea behind the design of the sorting machine is to use the different particle sizes of the mixed coal to discharge the suitable material from the far left into the selection process, while the gangue and large impurities are discharged from the right side.

When the eccentric shaft rotates, there is displacement in the *X* and *Y* direction of the sorting bed surface.

$$x = e\cos\varphi\cos\omega t \tag{1}$$

$$y = e\sin\varphi\cos\omega t \tag{2}$$

In the formula, e is the eccentricity of the heart axis rotation, m; φ is the angle between the plate spring and the sorting bed surface, °; ω is the rotational speed of the heart axis, r/min; t is the time, s. The expressions of velocity and acceleration on the sorting bed can be obtained by calculating the first and second derivatives of equations (1) and (2) with respect to time t respectively.

$$\dot{x} = -e\omega\cos\varphi\sin\omega t \tag{3}$$

$$\ddot{x} = -e\omega^2\cos\varphi\cos\omega t \tag{4}$$

$$\dot{y} = -e\omega\sin\varphi\sin\omega t \tag{5}$$

$$\ddot{y} = -e\omega^2\sin\varphi\cos\omega t \tag{6}$$

The formula below can be used to calculate the amplitude in *X* and *Y* directions of the sorting bed surface.

$$A_x = e\cos\varphi \tag{7}$$

$$A_y = e\sin\varphi \tag{8}$$

3. Establishment of multi-body dynamics simulation model

Under the interface of ADAMS software, a simplified simulation model of a vibration cascade sorter was established, as shown in **Figure 2** and **Figure 3**. The eccentricity distance e is 8 mm, and the inclination angle φ of the plate spring is 30 degrees. The spring stiffness is set according to theoretical calculations. To overcome the influence of gravity, the spring is preloaded before the simulation. The gravity acceleration is set to 9.81 m/s^2. The material selected is steel, with a corresponding density of 7.82×10^3 kg/m^3 [7]. The constraint relationships among the heart shaft and the eccentric sleeve, the eccentric sleeve and the plate spring, the plate spring and the sorting bed surface, and the sorting bed surface and the support frame are all defined as joint rotational pairs. A motion rotational drive is applied to the joint rotational pair between the heart shaft and the eccentric sleeve, with the rotational speed function specified as STEP(time, 0, 0, t0, ω0), indicating that the rotational speed increases smoothly from 0 to ω0 within the first t0 seconds. After this period, the heart shaft rotates at a constant angular velocity. In this study, the total simulation time, acceleration time (t0), target angular velocity (ω0), and simulation step size are set to 20 s, 10 s, 38.7 $rad{\cdot}s^{-1}$ (equivalent to 370 $r{\cdot}min^{-1}$), and 0.002 s, respectively, with experimental data recorded at every 0.002-second interval.

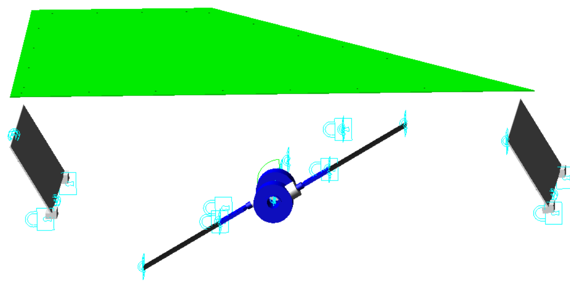

Figure 2. Multi-body dynamic simulation model of vibration cascade sorter

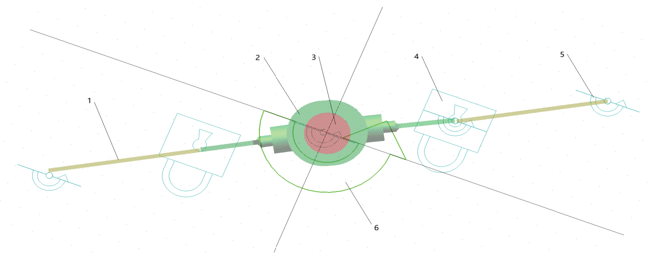

Figure 3. Simulation model driven settings; (1) Plate spring, (2) Eccentric bushing, (3) Manshaft, (4) Fixed pair, (5) Rotating pair, (6) Drive the vice

4. Multi-body dynamics simulation analysis

The time-domain response curve of the sorting bed displacement is shown in **Figure 4**. As illustrated, from 0 to 1.4 seconds, the vibrating cascade sorter operates in the startup phase, during which the vibration frequency of the bed surface gradually increases. After 1.4 seconds, the system reaches stable operating conditions, and the time-domain response curve reflects the displacement behaviour of the bed surface in both the *X* and *Y* directions. Under these stable conditions, the displacement of the sorting bed surface exhibits smooth and periodic variations, with a consistent phase difference of zero between the *X* and *Y* directions. The displacement amplitude of the sorting bed surface is 7.01 mm in the *X* direction and 3.85 mm in the *Y* direction.

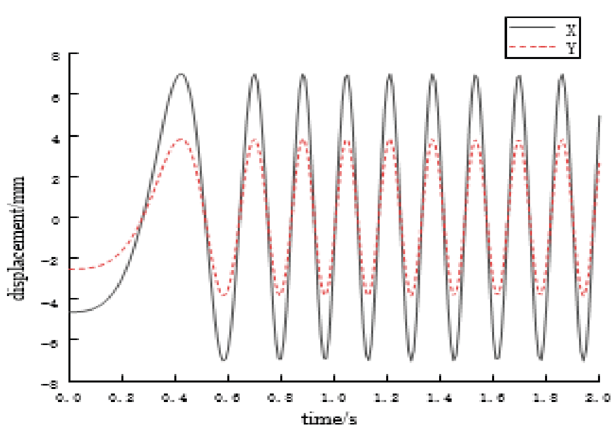

Figure 4. Based on the time domain response curve of the displacement characteristics of the sorting bed surface of MBD vibration cascade sorter

Figure 5 presents the spatiotemporal characteristic curve and the displacement Lissajous figure corresponding to the 0–20 second displacement-time domain response. The spatiotemporal characteristic curve reveals that the spatial displacement of the sorting bed is initially irregular and chaotic but gradually transitions to a stable state. Once stable operating conditions are reached, the Lissajous figure shows that the displacement trajectory of the sorting bed forms a tilted straight line, indicating a consistent and periodic motion pattern.

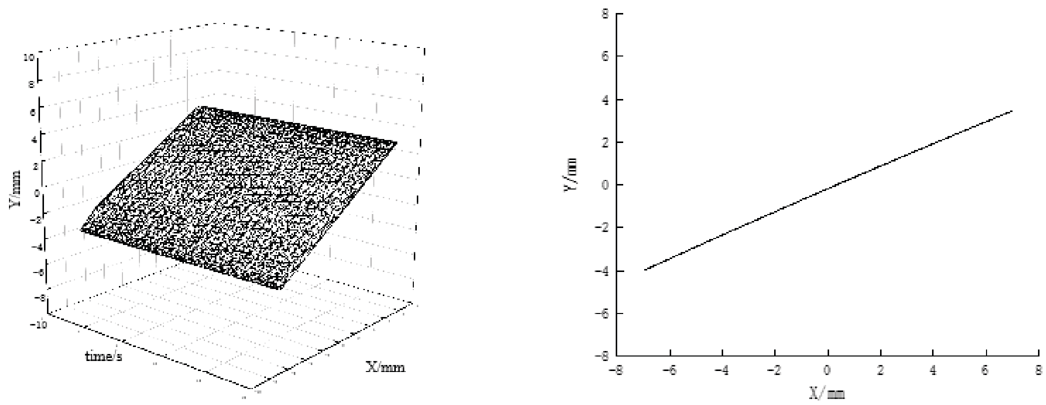

Figure 5. Space-time characteristic curve and displacement Lissajous of the sorting bed surface based on MBD vibration cascade sorting machine

Figure 6 illustrates the time-domain response curve of the sorting bed surface velocity. As shown in the figure, during the start-up phase (0–1.4 s), the velocity amplitude of the sorting bed surface gradually increases in both the *X* and *Y* directions. After 1.4 seconds, the velocity amplitude stabilizes, indicating the transition to steady-state operation. Under stable conditions, the velocity amplitude of the sorting bed surface reaches 271.7 mm·s^{-1} in the *X* direction and 149.3 mm·s^{-1} in the *Y* direction. Throughout this stable phase, both the magnitude and direction of the velocity remain consistent.

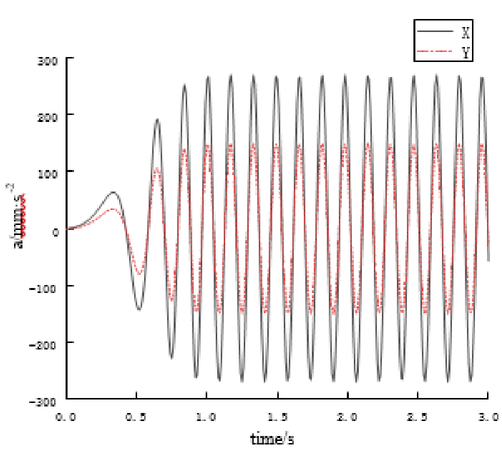

Figure 6. Time domain response of velocity characteristics based on MBD vibration cascade sorter

Figure 7 shows the time-domain response curve of the sorting bed surface acceleration. From the figure, it can be seen that during the start-up phase (0–1.4 s), the amplitude of the acceleration of the sorting bed surface along the X and Y directions gradually increases. After 1.4 s, the amplitude reaches the maximum and begins to operate stably. In the stable state, the amplitude of the sorting bed surface acceleration along the X direction is 10.6 mm·s^{-2}, and the amplitude of the acceleration along the Y direction is 6.0 mm·s^{-2}. Under stable operating conditions, the acceleration direction of the sorting bed surface along the X direction remains the same [9].

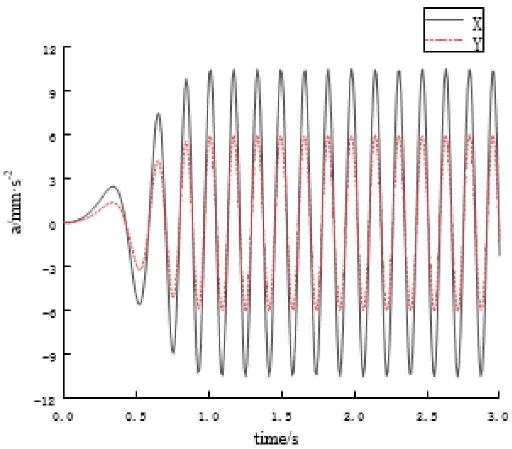

Figure 7. Time-domain response curve of acceleration characteristics based on MBD vibration cascade sorter

The theoretical calculation results of the motion characteristics, such as displacement amplitude, velocity amplitude, and acceleration amplitude of the vibrating cascade sorter bed surface based on the kinematic model are compared with the simulation results based on multi-body dynamics (MBD), as shown in **Table 1**. From the data presented, it is evident that the computational outcomes align closely with the predicted values. The highest discrepancy stands at just -3.75%, demonstrating the high reliability and precision of the theoretical predictions [10].

Table 1. Comparison between theoretical results and simulation results

Parameter		Theoretical result	Simulation result	Relative error
Sort the bed surface	$\lvert x\rvert$/mm	6.93	7	1.01
	$\lvert\dot{x}\rvert$ /mm s^{-1}	268.1	271.7	1.34
	$\lvert\ddot{x}\rvert$ /mm s^{-2}	10.4	10.6	1.92
	$\lvert y\rvert$ /mm	4.0	3.85	-3.75
	$\lvert\dot{y}\rvert$ /mm s^{-1}	154.69	149.3	-3.49
	$\lvert\ddot{y}\rvert$ /mm s^{-2}	5.98	6.0	0.33

5. Conclusion

This article first analyzed the motion theory of the vibrating cascade sorter and obtained the motion characteristic parameters of it. The motion characteristics of the vibrating cascade sorter were simulated and analyzed using the multibody dynamics software ADAMS. The correctness and feasibility of the theoretical model calculation and analysis method were verified. The motion characteristics of the sorter body, including displacement amplitude, velocity amplitude, and acceleration amplitude, show a clear dependence on eccentricity and rotational speed. As the eccentricity increases, all three amplitudes increase correspondingly. However, when the rotational speed of the heart shaft increases, the displacement amplitude of the sorter body remains constant, while both the velocity and acceleration amplitudes gradually increase.

Using principles of complex system mechanics, a digital model of the vibrating cascade sorter was created. Simulations were conducted to derive the movement patterns and three-dimensional paths of the sorter's body. The predicted values for key dynamic properties, including displacement range, velocity range, and acceleration range, were compared against the computational outcomes. The findings from the simulations closely matched the analytical predictions, showing a highest discrepancy of under 3.75%.

The parametric study showed that the motion characteristics of the sorter body, including displacement amplitude, velocity amplitude, and acceleration range, all rise with greater eccentricity. When the rotational speed of the central shaft increases, the displacement range of the sorter body stays constant, while the velocity and acceleration ranges progressively grow. The motion characteristics of the vibrating cascade sorter are optimal for the separation device load screening efficiency when the separation bed inclination is 30°, the amplitude is 8mm, and the heart axis rotation.

The article analyzes the dynamic characteristics and motion laws of the vibrating cascade sorter as the research object. To lay the foundation for the further design of large-scale coal sorting equipment and explore the working mechanism of the vibrating cascade sorter, the next step will focus on improving the screening efficiency and accuracy. This will be achieved through optimizing the structural design of the vibrating sorter, as well as the shape and material of the screen, among other aspects. For example, using high-strength and wear-resistant materials to make the screen can prolong the service life of the vibrating sorter and improve screening efficiency.

In summary, the next research directions for the vibrating sorter include improving screening efficiency and accuracy, reducing energy consumption and noise, expanding application fields, and achieving automation control.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Hou SU, Ma LQ, Huang G, 2016, Mechanism of the Effect of Fine Coal with Different Densities on Coarse Coal Flotation. Journal of China Coal Society, 41(7): 1813–1819.

[2] Zhang S, Zhang XX, 2011, Experimental Study on Vibration Countercurrent Dry Separation Machine. Coal Preparation Technology, 2011(1): 1–3 + 75.

[3] Yang YF, Chen QR, Li Q, 2002, Current State of Coal Dry Beneficiation with Air-Dense Medium Fluidized Bed. Journal of Huainan Institute of Technology, 2002(1): 36–39.

[4] Zhao YM, Yang XL, Zhao PF, 2015, The Invention Relates to a Vibrating Fluidized Bed Drying Separation Body Machine: CN103196281B.

[5] Song SL, Zhao YM, Luo ZF, 2012, Fine Coal Particles Separation by Air-Solid Magnetic Fluidization. Journal of China Coal Society, 37(9): 1586–1590.

[6] Li LJ, Hu WY, Gao ZC, et al., 2016, The Crank and Rocker are Broken by Vibration Kinematics Analysis and Experiment of Shell Camellia Fruit Roughing Machine. Transactions of the Chinese Society of Agricultural

Engineering, 32(7): 28–35.
[7] Qi GQ, Lv JB, 2007, Variance Analysis on Sinusoid Frequency Estimators Based on the Argument of the Sample Autocorrelation Function. Journal of Dalian Maritime University, 2007(4): 5–9.
[8] Xu WF, 2019, Study on Composite Dry Separation Characteristics of Fine-Grained Coal Less Than 6 mm, thesis, Huazhong University of Science and Technology.
[9] Dong L, Zhang YD, Zhao YM, 2020, Fluidization Characteristics and Fine Coal Separation in a Gas-Solid Vibrated Fluidized Bed. Journal of China Coal Society, 2020(6): 29–34.
[10] Sun YY, Liu MX, Wang LJ, 2010, Vibration Modal Analysis System Based on LabVIEW. Coal Mine Machinery, 31(10): 149–151.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Synergy Between Resilient Networks and Random Forests in Online Fraud Detection

Junxi Wang[1], Ningtao Sun[2], Yuhan Lv[3], Jiayi Zhou[4], Yue Xiao[5]*

[1]School of Intelligent Manufacturing and Aviation, Zhuhai College of Science and Technology, Zhuhai 519041, Guangdong, China
[2]School of Computer Science, Zhuhai College of Science and Technology, Zhuhai 519041, Guangdong, China
[3]School of Liberal Arts, Zhuhai College of Science and Technology, Zhuhai 519041, Guangdong, China
[4]School of Logistics Management and Engineering, Zhuhai College of Science and Technology, Zhuhai 519041, Guangdong, China
[5]School of Data Science, Zhuhai College of Science and Technology, Zhuhai 519041, Guangdong, China

***Corresponding author:** Yue Xiao, xiaoyue810@zcst.edu.cn

Abstract: This paper explores the synergistic effect of a model combining Elastic Net and Random Forest in online fraud detection. The study selects a public network dataset containing 1781 data records, divides the dataset by 70% for training and 30% for validation, and analyses the correlation between features using a correlation matrix. The experimental results show that the Elastic Net feature selection method generally outperforms PCA in all models, especially when combined with the Random Forest and XGBoost models, and the ElasticNet + Random Forest model achieves the highest accuracy of 0.968 and AUC value of 0.983, while the Kappa and MCC also reached 0.839 and 0.844 respectively, showing extremely high consistency and correlation. This indicates that combining Elastic Net feature selection and Random Forest model has significant performance advantages in online fraud detection.

Keywords: Fraudulent websites; Machine learning; Elastic Net; Random forests

Online publication: March 28, 2025

1. Introduction

With the rapid development of Internet technology, the Internet has become an indispensable part of daily life. It provides a convenient platform for information acquisition, communication, and business transactions. However, this convenience also brings a series of security challenges, especially the increasing number of network fraud. Network fraud not only harms the economic interests of users, but also threatens the security and trust of the entire network ecosystem. Therefore, how to effectively identify and prevent network fraud has become an important topic in the field of network security.

Among the many cybersecurity issues, the detection of fraudulent websites is particularly critical. Fraudulent websites usually disguise themselves as legitimate websites and induce users to disclose sensitive information, such as login credentials and credit card information, which brings direct financial losses and privacy risks to users. Traditional rule-based detection methods appear to be incompetent in the face of ever-changing fraudulent tactics, so researchers have begun to seek more advanced techniques to improve the accuracy and efficiency of detection [1].

In recent years, machine learning techniques have received a lot of attention for their power in data analysis and pattern recognition. The ability of machine learning algorithms to learn and recognize complex patterns from large amounts of data has led to great potential in the field of fraud detection. In recent years Dou *et al.* proposed a new model called CAmouflage-REsistant GNN (CARE-GNN), which aims to strengthen the aggregation process of the GNN through three neural modules to prevent the fraudster's camouflage behaviors and demonstrated the effectiveness of the method on two real fraud datasets through experimental results [2]. Kim *et al.* proposed a fraud detection method based on graph neural network (GNN) to solve the fraud detection problem through a dynamic relational attention aggregation mechanism [3]. The method learns the node representation of each relation and uses a learnable attention function to aggregate the node representations, assigning different attention coefficients to each relation. Xu *et al.* proposed SEC-GFD, a semi-supervised GNN fraud detector, which improves detection through hybrid filtering and local environment constraints and experimentally proves that it outperforms other methods [4].

Based on this research, this paper will explore the application of models combining Elastic Net and Random Forest in fraudulent website detection. Elastic Net is an integrated regression technique that incorporates L1 and L2 regularization to enhance the generalization ability of the model by promoting sparsity in feature selection and enhancing model stability. Meanwhile, Random Forest, an integrated learning method, effectively improves the accuracy and robustness of the model by constructing multiple decision trees and integrating their predictions. This combination aims to optimize the performance of fraud detection to identify and prevent online fraud more effectively.

2. Model construction

2.1. Analysis of data sets

A publicly available web dataset was selected for this study, which contains 1781 data records covering 14 different feature dimensions. For model training and validation, the dataset was divided in a 7:3 ratio, where 70% of the data was used to construct a training set for model learning and the remaining 30% was used as a validation set to evaluate the model's performance and generalization ability. This division aims to ensure that the model can make full use of the data during the training process, and at the same time, the validation set is used to monitor the overfitting situation of the model, to achieve the optimization and refinement of the model. To assess the correlation between different features in the dataset, this study adopts the correlation matrix technique for the analysis. By constructing the correlation matrix, as shown in **Figure 1**, it is possible to visualize the correlation coefficients between the features.

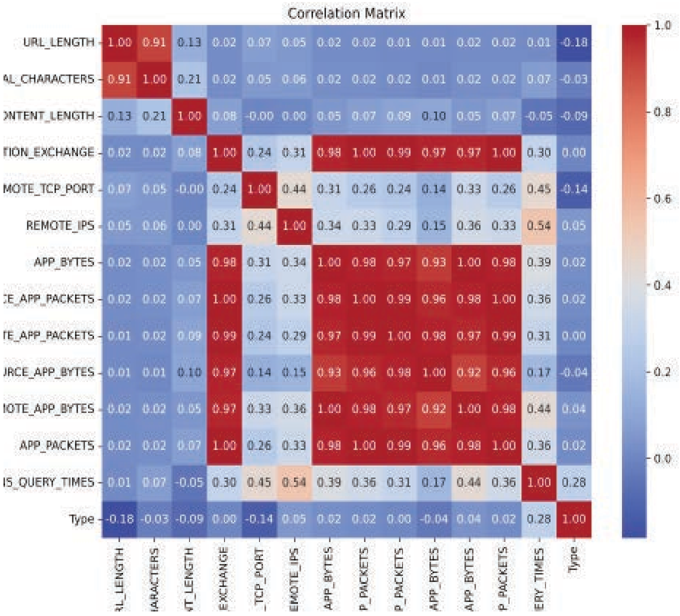

Figure 1. Correlation matrix diagram

2.2. Elastic networks

As shown in **Figure 1**, it can be observed that there is a high degree of positive correlation between features such as "APP_BYTES" and "APP_PACKETS", which may lead to the problem of multicollinearity during the modelling process, thus affecting the stability and prediction accuracy of the model [5, 6]. To solve this problem, this study adopts the Elastic Net regularization technique. The Elastic Net regularization technique combines the advantages of Lasso regression (L1 regularization) and Ridge regression (L2 regularization). By introducing L1 and L2 regularization at the same time, it is able to balance the feature selection and model smoothing, which effectively reduces the risk of overfitting and improves the prediction performance. The objective function of elastic network regression can be expressed as shown in Equation 1 below:

$$\mathcal{L}(\omega) = \frac{1}{2n}\|y - X_{\omega}\|_2^2 + \alpha\left(\rho\|w\|_1 + \frac{1-\rho}{2}\|w\|_2^2\right) \quad (1)$$

In this equation, X denotes the feature matrix, y is the target variable, the vector of regression coefficients to be estimated, n is the number of samples, is the number of features, and is the regularization strength parameter, which controls the degree of regularization of the model, and is the weighted ratio of L1 regularization (Lasso) to L2 regularization (Ridge). Specifically, when $\rho = 1$, the elasticity network degenerates into a Lasso regression model; when = 0, it degenerates into a Ridge regression model; and when $0 < \rho < 1$, the elasticity network combines the advantages of L1 and L2 regularization, which is capable of feature selection as well as avoiding the problem of covariance between features.

In the objective function of the elastic network, the error term is $\frac{1}{2n}\|y - X_{\omega}\|_2^2$. The measure is the error of fit of the model, the regularization term is $\alpha\left(\rho\|w\|_1 + \frac{1-\rho}{2}\|w\|_2^2\right)$. It is used to control the complexity of the model and to prevent overfitting. The L1 regularization term (Lasso term) enables feature selection by summing the absolute values of the weights so that certain weight coefficients converge to zero, while the L2 regularization term (Ridge term) smoothes the model parameters and reduces instability due to covariance by summing the squares of the weight coefficients.

Optimization of the model can be achieved by adjusting the parameters α and ρ, which govern the strength and balance of the regularization terms. The parameter α controls the overall regularization strength; larger values of α increase the penalty, leading to a more simplified model that can prevent overfitting. Meanwhile, ρ determines the relative contributions of L1 and L2 regularizations within the elastic net framework, allowing the model to flexibly select relevant features while suppressing irrelevant ones. Through the appropriate tuning of these parameters, the model's generalization ability is enhanced, ultimately improving its detection accuracy.

2.3. Random forest

In recent years, significant progress has been made in Random forest techniques, making them a prominent choice for ransomware detection due to their robustness and ability to handle high-dimensional data [7]. Similarly, Pathak *et al.* used various machine learning models for phishing website detection and after comparison found that random forest model was able to show better performance in this task [8]. The valid data used for detecting fraudulent websites in this study contains multiple dimensions and the data in each dimension is discrete, so the random forest technique can also be used for fraudulent website detection. Random forest, as an integrated learning model proposed based on bagging strategy (Bagging), consists of multiple decision trees. In random forest, each decision tree is an independent and uncorrelated classifier and each tree will only learn on a certain subset of the whole data. Because of this, the random forest can capture the nonlinear structure in the data through different decision trees and comprehensively acquire important features in the data while effectively reducing the risk of overfitting during the training process. The principle of the training and testing steps of the random forest model in this study is shown in **Figure 2**.

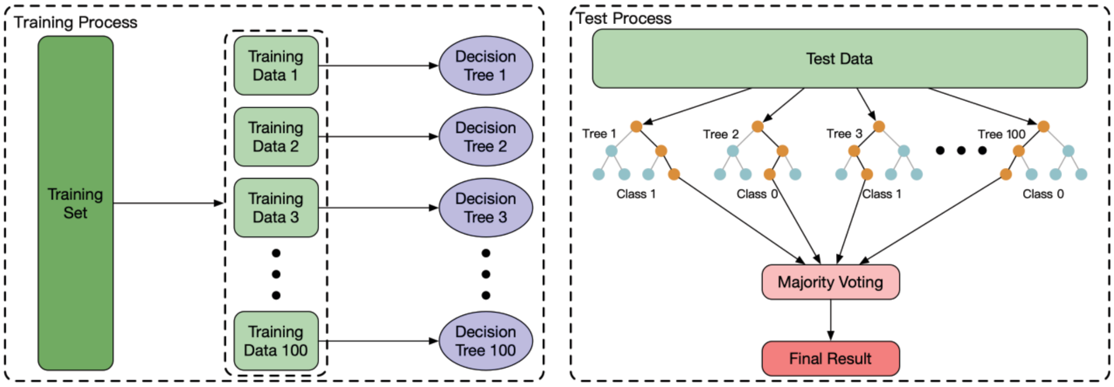

Figure 2. Random forest structure diagram

To ensure that the learning model used in this study has sufficient predictive power, the random forest model designed in this study consists of 100 decision trees. By using the fraud detection website dataset, the data was divided into 100 parts to train each decision tree, and the majority voting result of multiple decision trees was selected as the final result in the test. Ultimately, the depth of each decision tree in the random forest model trained in this study was moderate, with an average depth of 14.04. The training data indicated that the model was trained in a way that ensured that the model was able to capture the data features effectively while avoiding the risk of overfitting. With the bagging strategy feature, the random forest model applied to fraudulent website detection uses the majority voting method to predict the results of all decision trees with the prediction formula shown in Equation 2:

$$\hat{y} = arg \max_{C\in\{0,1\}} \sum_{t=1}^{T} 1(h_t(x) = C) \tag{2}$$

3. Results of the experiment

3.1. Model evaluation indicators

(1) Accuracy: Accuracy is one of the most intuitive metrics to evaluate in machine learning, measuring the number of samples the model predicts correctly as a percentage of the total number of samples.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

(2) Specificity: Specificity refers to the rate of true negative cases, i.e., the proportion of cases correctly predicted as negative by the model, and it reflects the model's ability to classify negative cases.

$$Specificity = \frac{TN}{TN+FP}$$

(3) Log Loss: In binary classification problems, log-loss is a metric for assessing the performance of a classification model, which measures the difference between the model's predicted probability and the actual occurrence of a label.

$$Log\ Loss = -\frac{1}{N}\sum_{i=1}^{N} [y_i \log (p_i) + (1 - y_i) \log (1 - p_i)]$$

(4) AUC: AUC indicates the ability of the model to identify positive classes; the higher the AUC value, the better the classification ability of the model.

(5) Kappa: The Kappa coefficient is a measure of classification accuracy that evaluates the agreement between two evaluators (or classifiers) by taking into account both observed agreement and chance agreement.

$$Kappa = \frac{P_A-P_e}{1-P_e}$$

(6) MCC: MCC is a metric for assessing the performance of classification models that takes into account the values of all four confusion matrices and is not affected by category imbalance.

$$MCC = \frac{TP\times TN-FP\times FN}{\sqrt{(TP+FP)\times(TP+FN)\times(TN+FP)\times(TN+FN)}}$$

3.2. Model results

Table 1 shows that two feature selection methods, PCA and ElasticNet, were used to compare model performance in online fraud detection.

Table 1. Table of evaluation indicators for different outcome model analyses

Method	Model	Accuracy	Specificity	Log Loss	AUC	Kappa	MCC
PCA	Logistic Regression	0.867	0985	0.308	0.809	0.0007	0.001
PCA	SVM	0.878	1.000	0.349	0.739	0.0000	0.000
PCA	XGBoost	0.933	0.970	0.237	0.893	0.667	0.669
PCA	Random Forest	0.943	0.989	0.288	0.920	0.697	0.711
ElasticNet	Logistic Regression	0.908	0.983	0.227	0.918	0.450	0.485
ElasticNet	SVM	0.886	1.000	0.239	0.933	0.103	0.233
ElasticNet	XGBoost	0.959	0.974	0.094	0.989	0.809	0.810
ElasticNet	Random Forest	0.968	0.993	0.102	0.983	0.839	0.844

Results indicate that ElasticNet outperformed PCA across all models, especially in Random Forest and XGBoost. The ElasticNet + Random Forest model achieved the highest accuracy (0.968), AUC (0.983), Kappa (0.839), and MCC (0.844). ElasticNet + XGBoost also performed well with an accuracy of 0.959 and AUC of 0.989. Under PCA, Random Forest had the best accuracy (0.943) and AUC (0.920). SVM achieved perfect specificity (1.000) under both methods, while Logistic Regression had lower specificity (0.983 under ElasticNet). In terms of log loss, ElasticNet + XGBoost and ElasticNet + Random Forest had the lowest values (0.094 and 0.102, respectively). Overall, ElasticNet + Random Forest provided the best performance for web fraud detection.

The performance of ElasticNet regularization combined with four machine learning models (Logistic Regression, SVM, XGBoost, and Random Forest) for classification is shown in **Figure 3** and **Figure 4**. The confusion matrices indicate that ElasticNet + Random Forest had the best performance for positive classes (only 3 misclassifications), while ElasticNet + Logistic Regression was best for negative classes (462 correct predictions, 8 misclassifications). The scatterplots show that ElasticNet + SVM had good linear separability, while ElasticNet + XGBoost and ElasticNet + Random Forest had some misclassifications but still performed well overall. Overall, ElasticNet + Random Forest appears to be the best model for this classification task.

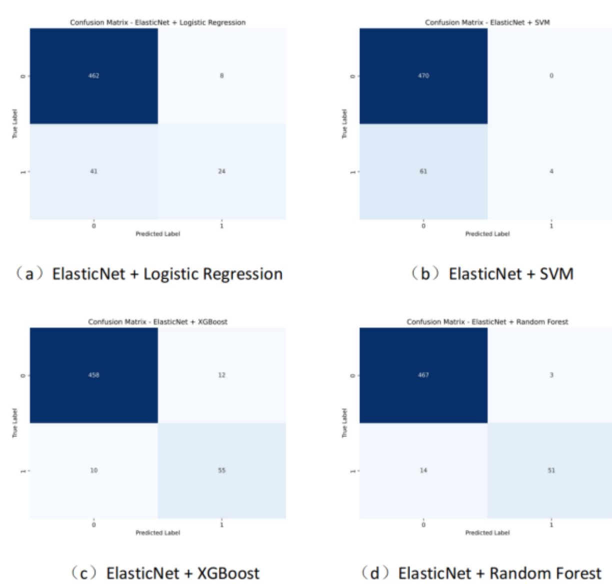

Figure 3. Confusion matrix for each combined model

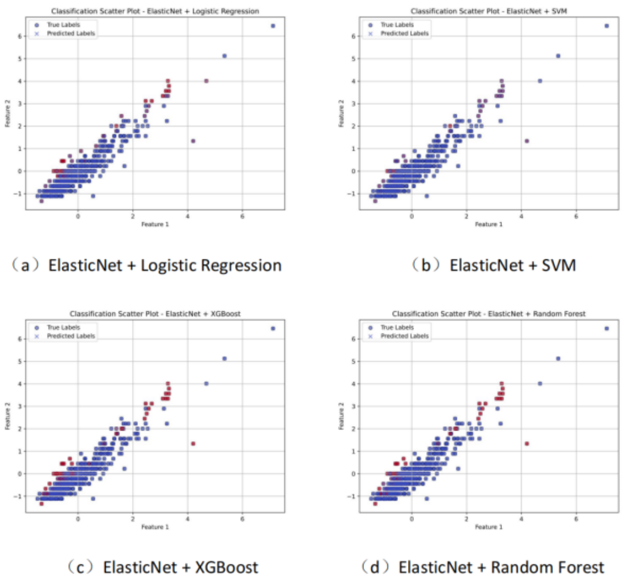

Figure 4. Calibration curves for each combination

4. Conclusion

In this study, we validate the effectiveness of the model by combining ElasticNet with Random Forest and in online fraud detection. By comparing different feature selection methods and machine learning models, it is found that the combination of ElasticNet and Random Forest performs the best in terms of evaluation metrics such as accuracy, AUC value, Kappa, and MCC. The ElasticNet regularization technique not only reduces the feature complexity and the risk of overfitting but also achieves the feature selection, while the Random Forest effectively captures the nonlinear structure and important features in the data through the majority voting method. This combination not only improves the generalization ability of the model but also enhances the robustness and accuracy of the model, providing an efficient technical solution in the field of online fraud detection.

Funding

Guangdong Innovation and Entrepreneurship Training Programme for Undergraduates "Automatic Classification and Identification of Fraudulent Websites Based on Machine Learning" (Project No.: DC2023125)

Disclosure statement

The authors declare no conflict of interest.

Author contributions

Model construction and data analysis: Junxi Wang, Ningtao Sun
Writing: Yuhan Lv, Jiayi Zhou
Project guidance: Yue Xiao

References

[1] Pourhabibi T, Ong KL, Kam BH, et al., 2020, Fraud Detection: A Systematic Literature Review of Graph-Based Anomaly Detection Approaches. Decision Support Systems, 133: 113303.

[2] Dou Y, Liu Z, Sun L, et al., 2020, Enhancing Graph Neural Network-Based Fraud Detectors Against Camouflaged Fraudsters. Proceedings of the 29th ACM International Conference on Information & Knowledge Management, 2020: 315–324.

[3] Kim H, Choi J, Whang JJ, 2023, Dynamic Relation-Attentive Graph Neural Networks for Fraud Detection. 2023 IEEE International Conference on Data Mining Workshops (ICDMW), 2023: 1092–1096.

[4] Xu F, Wang N, Wu H, et al., 2024, Revisiting Graph-Based Fraud Detection in Sight of Heterophily and Spectrum. Proceedings of the AAAI Conference on Artificial Intelligence, 38(8): 9214–9222.

[5] Zhang Y, Ye W, Zhang J, 2017, A Generalized Elastic Net Regularization with Smoothed ℓq Penalty for Sparse Vector Recovery. Computational Optimization and Applications, 2017, 68: 437–454.

[6] Zou H, Hastie T, 2005, Regularization and Variable Selection via the Elastic Net. Journal of the Royal Statistical Society Series B: Statistical Methodology, 67(2): 301–320.

[7] Rafapa J, Konokix A, 2024, Ransomware Detection Using Aggregated Random Forest Technique with Recent Variants. Authorea, 2024: 1–8.

[8] Pathak P, Shrivas AK, 2024, Development of Proposed Model Using Random Forest with Optimization Technique for Classification of Phishing Website. SN Computer Science, 5: 1059.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Exploration of the Application of Artificial Intelligence Technology in the Transformation of Old Objects

Tonghuan Zhang, Xinyu Yang, Ying Chen*, Qiufan Xie*

Chongqing Institute of Engineering, Chongqing 400056, China

***Corresponding author:** Ying Chen, cying88@163.com; Qiufan Xie, CherFen@foxmail.com

Abstract: With the rapid development of technology, artificial intelligence (AI) is increasingly being applied in various fields. In today's context of resource scarcity, pursuit of sustainable development and resource reuse, the transformation of old objects is particularly important. This article analyzes the current status of old object transformation and the opportunities brought by the internet to old objects and delves into the application of artificial intelligence in old object transformation. The focus is on five aspects: intelligent identification and classification, intelligent evaluation and prediction, automation integration, intelligent design and optimization, and integration of 3D printing technology. Finally, the process of "redesigning an old furniture, such as a wooden desk, through AI technology" is described, including the recycling, identification, detection, design, transformation, and final user feedback of the old wooden desk. This illustrates the unlimited potential of the "AI + old object transformation" approach, advocates for people to strengthen green environmental protection, and drives sustainable development.

Keywords: Artificial Intelligence (AI); Old object transformation; Environmental protection

Online publication: April 2, 2025

1. Introduction

With the advancement of technology, artificial intelligence (AI) has gradually become a focal point of attention in various industries and integrated into our daily lives. It has not only achieved remarkable application results in fields like smart manufacturing, smart healthcare, and smart education, but also demonstrated its strong potential and value in areas such as artistic creation, bringing more possibilities to artistic expression. Meanwhile, the pursuit of sustainable development and resource reuse has always been a hot topic, attracting widespread attention. In 2021, the National Development and Reform Commission issued the "14th Five-Year Plan for Circular Economy Development," stating that "developing a circular economy is a major strategic direction for China's economic and social development" [1]. The circular economy, with efficient resource utilization and recycling as its core, follows the principles of reduction, reuse, and recycling. Characterized by low consumption, low emissions, and high efficiency, it is essentially a green, low-carbon, and sustainable economic development model [2]. As an important component of resource recycling, the recycling of second-hand products helps reduce resource waste, lower environmental pollution, and brings new opportunities for economic development [3]. Upcycling, largely sourced from recycled second-hand products, is a crucial part of the circular economy. Both are closely related, aiming to achieve effective resource utilization and promote sustainable development by reducing resource consumption and waste generation. The application of AI technology further facilitates this process, accelerating resource recycling and the green transformation of the economy by optimizing resource management and promoting changes in consumption patterns. In this context, the integration of upcycling, an environmentally friendly and economical concept, with AI technology, opens up new possibilities for our lives.

2. Concept and significance of upcycling

Explaining the meaning of upcycling from an environmental perspective refers to the process of transforming discarded, seemingly valueless items into something new and useful through creative redesign. This not only avoids waste but also realizes the repeated use of resources, conserving them and contributing to environmental protection [4]. This concept carries multiple significances. Firstly, it turns waste into treasure, preventing the earth from crisis due to resource depletion, emphasizing everyone's responsibility for environmental protection and green development. Secondly, it stimulates creativity and imagination. Every discarded item carries a story and the upcycler transforms it into a new artwork or utility item through inspiration. This process tests the creator's hands-on ability and innovative thinking, sparking boundless imagination and enriching life with creativity.

3. Current situation and challenges of upcycling

In modern times, with the rise of environmental awareness, upcycling has gained increasing popularity and shows broad development prospects. However, it also faces a series of challenges.

3.1. Instability of the upcycling supply chain

The raw materials for upcycling rely on the supply of second-hand items, and the instability of this supply chain can affect market operations. The main reasons include the scattered sources of waste items, diverse varieties, difficulties in fine classification, transportation challenges, low recycling value, limited reuse applications, poor market environment, and immature artistic processing of some old items.

3.2. Incomplete recycling system and model

In 2023, China's total recycled volume of ten types of renewable resources, including scrap steel, scrap non-ferrous metals, scrap plastics, waste paper, scrap tires, discarded electrical and electronic products, scrapped motor vehicles, waste textiles, waste glass, and waste batteries, reached approximately 376 million tons, showing an increase of 1.5% compared to 2022 [5]. Despite the rising total recycling volume and the expanding scale of the waste material recycling industry, it faces issues such as the low recycling rate of low value-added renewable resources (such as waste textiles, waste glass, and packaging materials). There is an urgent need to establish a stable and reliable recycling system and model to promote effective resource recycling.

3.3. Low brand recognition

Currently, there are various upcycled products in the market, such as using empty glass bottles to decorate various food items, repurposing idle buckets as planters, and transforming unusable canvas bags into refreshed goodies with paintbrushes and inspiration. However, these are common upcycling practices in daily life, where people recreate the value of an item at low cost through DIY. Most of these activities are spontaneously carried out by individuals during their free time or by art lovers, without establishing brand recognition. Although there are some online platforms dedicated to upcycling, their brand awareness is relatively low, requiring resources for market promotion.

3.4. Reliance on manual labor and low efficiency

Traditional upcycling heavily relies on manual labor and the upcycler must possess craftsmanship and ideas. Transforming an item can be time-consuming and inefficient, and sometimes technical limitations make it difficult to meet the growing market demand. Additionally, handling complex discarded items like old appliances and furniture requires high-level professional knowledge and skills. Therefore, improving the efficiency and quality of upcycling has become an urgent issue to address.

4. Opportunities brought by the Internet to the transformation of old items

(1) Initiate the "Internet + Recycling" model for transforming old items. This model will be beneficial to expanding and optimizing the supply chain for transforming old items, as well as increasing the influence of such transformations, calling on people to pay attention to environmental protection and low-carbon concepts.

(2) Establish an old item recycling and collection system, set up a hierarchical model of "online and offline stores - old item processing factories", and build an efficient collection network for waste items to improve collection efficiency.

(3) Initiate the "Internet + Promotion" model for transforming old items, providing one-stop services. Brand promotion can be done through e-commerce, enhancing publicity efforts and delivering brand messages. Apart from increasing brand awareness, we should also improve the credibility and service quality of the old item transformation brand, track user experience, and promote "trade-in" programs where the "new" refers to valuable new products made from transformed old items.

(4) Initiate the "AI" model for transforming old items to improve efficiency. AI can be used to quickly generate ideas, providing designers with more inspiration and helping them improve work efficiency. AI technology lowers the threshold for creativity, allowing more and more people to experience becoming a "creator". Anyone can easily put their desired old item transformation results on paper, significantly saving on design costs.

5. The application of artificial intelligence in the transformation of old items

AI technology plays a crucial role in the transformation of old items. The application of AI in this field mainly focuses on the following five aspects.

5.1. Intelligent identification and classification

AI technology can be used for the identification and classification of old items. Through image recognition and machine learning algorithms, different types of old items can be quickly and accurately identified. This helps divert old items from landfills or recycling stations, providing raw materials for subsequent transformations. This reduces reliance on human resources and greatly improves the efficiency and accuracy of old item transformations. For example, JD Ambati, the founder and CEO of EverestLabs, stated in an interview, "Our AI can provide precise data on the shape, size, weight, material, packaging type, commodity value, and even brand information for each recyclable item that flows through the factory".

EverestLabs, a company with an independent industrial 3D vision system, has established a proprietary dataset containing over 5 billion recyclable objects to train its recognition algorithm. Their data and robotics platform, RecycleOS, can classify objects with over 95% accuracy. Additionally, the development of AI has facilitated the implementation of smart recycling machines in multiple cities. These machines can identify four major categories of items: paper, metal, plastic, and fabric, including specific items like newspapers, faucets, drink bottles, and plush toys [6]. After users deposit waste items, the camera inside the recycling machine takes a photo, and an AI risk control model determines if the item meets deposit requirements. The machine quickly weighs and verifies the item. Upon successful verification, users receive corresponding rewards. When the recycling machine is full, built-in sensors send information to the logistics platform via the Internet of Things. The system then automatically allocates transport vehicles and plans the best route to ensure timely clearance.

Furthermore, integrating AI technology into the transformation of old furniture allows for intelligent sorting and classification. Old furniture can be transferred to the "old item processing factory" where AI performs image recognition again, quickly identifying information such as color, material, texture, and style of the old furniture, providing a basis for subsequent transformations.

5.2. Intelligent evaluation and prediction

AI technology can evaluate the value of old items, helping users understand their condition, remaining lifespan, and predict their potential and market value after transformation. This aids people in better selecting and transforming old items to achieve their maximum economic value.

5.3. Automation Integration

During the transformation of old items, sometimes interpolation and combination of old items are required. For complex disassembly and assembly tasks, AI automation technology, such as robotics, can be integrated to assist in completing these complex tasks and improve transformation efficiency.

Additionally, AI can optimize the transformation process through algorithms, reducing labor costs. For example, a certain mobile phone manufacturer announced its first smart recycling robots for mobile phones, named "Taz" and "Dave". These robots are primarily used for dismantling and recycling metal materials from mobile phones, such as copper from the mainboard, aluminum and titanium from the back cover, gold from the camera, and rare earth elements from acoustic components.

The Australian furniture brand Eva transformed 16,000 discarded mobile phone cases into a sofa, which was displayed at CASETiFY STUDiO in Sydney's CBD. Combining innovation and environmental protection, Re/CASETiFY™ poured contemporary creativity into the sofa design named Re/Treat, hoping to provide a space to relax the mind and emotions before continuing the journey towards sustainable development. Chengdu artist Liwei Yang loves to transform ordinary daily items into impressive artworks. He used 2,200 discarded mobile phone cases to create a colorful parasol, evoking childhood memories and continuing the second life of

Re/CASETiFY™ products.

5.4. Intelligent design and optimization

Artificial intelligence art can create excellent artworks by learning from existing data and artistic paradigms [7]. AI technology can provide intelligent design suggestions and optimization solutions for the transformation of old items based on user needs and preferences within artistic paradigms. This not only improves the transformation effect but also saves creation time and cost, providing creators with ideas. For example, in clothing transformation, AI can intelligently recommend suitable transformation plans based on the user's body shape, preferences, and temperament. For instance, for an outdated shirt, AI can analyze its color, pattern, and cutting, and propose new design concepts, such as adding fashionable decorative accessories, changing the collar, cuffs, and length variations, to give the old clothing a new look.

5.5. AI + 3D printing + transformation of old items

AI can quickly analyze design drawings, optimize production processes, and even perform real-time monitoring and adjustments during production. 3D printing, on the other hand, can rapidly turn AI-optimized designs into physical objects. From personalization, multifunctionality, and sustainability to cost reduction, 3D printing has already proven to be a great helper for artists in the creative world. Combining AI with 3D printing to give new life to old items is also an exploration of creativity [8].

For example, in November 2023, the Japanese design group Honoka exhibited an innovative 3D-printed furniture series at the Designart Tokyo exhibition. Named the "Tatami Transformation Project", the series uses a unique material made from a mixture of straw from discarded tatami mats and biodegradable resin. Manufactured using the ExtraBold large-scale 3D printer, the furniture incorporates elements of traditional Japanese craftsmanship, such as weaving, lattice, or knitting effects. The products include lampshades, lighting fixtures, tables, stools, etc., in a variety of colors, with a translucent characteristic that gives them a unique texture.

6. Case study analysis

Taking a typical case of transforming an old item, such as a wooden desk, through AI technology to redesign it in a way that meets both modern aesthetic and practical needs. The transformation process of this case includes recycling, identification, detection, design, transformation, and finally, user feedback.

6.1. Recycling phase of old wooden bookshelf

Wood can be sorted and recycled through scrap stations or smart recycling machines.

6.2. Old item detection phase

The AI model first performs a deep detection on the "old item - wooden desk". This includes detection of various aspects such as color, material analysis, structural analysis, and the extent of external damage. Through machine vision and deep learning technology, AI accurately evaluates the condition of the old item and generates a data report to provide data support for subsequent transformation designs.

6.3. Design phase

Based on artistic paradigm standards, AI design software automatically generates multiple transformation plans according to the detection results of the "old item - wooden bookshelf," combining modern aesthetics and practical needs. These plans include not only exterior design but also structural optimization and functional enhancements for customers to choose from, with the customer selecting a final optimized plan.

6.4. Manufacturing phase

In the manufacturing process, the AI model combines 3D printing and robotic manufacturing technology to implement the design plan. Through a precise manufacturing process, it ensures that the transformed item meets the design requirements and has high durability, achieving the ideal effect for the customer.

6.5. Evaluation phase

The evaluation of the AI model in the case of transforming old items mainly focuses on the following aspects:

(1) Accuracy evaluation: By comparing the AI model's design plan with manually designed plans and the actual effect of the final product, the accuracy of the AI model in detecting and designing old items can be evaluated. If the design plan and final product are highly consistent with the expected goals, it indicates high accuracy of the AI model.

(2) Efficiency evaluation: The application of the AI model can greatly improve the efficiency of transforming old items. Through automated design, automated manufacturing, and other processes, it reduces manual operation steps and time costs. Therefore, efficiency is one of the important indicators for evaluating the application of AI models in transforming old items.

(3) User feedback evaluation: The success of the final product also needs to consider user feedback. By collecting user feedback on the product, the advantages, disadvantages, and improvement directions of the product can be understood. If users are highly satisfied with the product, it indicates that the application of the AI model is successful.

(4) Environmental protection evaluation: Applying AI technology in transforming old items can not only improve resource utilization but also reduce waste generation. Therefore, environmental protection is also an important indicator for evaluating AI model applications. By comparing and analyzing resource consumption and waste generation before and after the transformation, the contribution of the AI model to environmental protection can be evaluated.

7. Conclusion and Future Prospects

The application of AI in transforming old items brings new possibilities and convenience to our lives. AI technology not only more efficiently identifies, evaluates, designs, and transforms old items, achieving their maximum economic and environmental value, but also integrates artistic creation with technology. In the future, it is hoped that the application of AI technology in the field of transforming old items will be more widespread and in-depth, realizing the intelligent and efficient transformation of old items. Let creativity change lives, and with the concept of sustainable development deeply rooted in people's hearts, transforming old items will become a more common way of life. It can penetrate smart cities, smart communities, revitalize villages, and contribute to saving resources and protecting the ecological environment for the earth.

Funding

2023 College Student Innovation and Entrepreneurship Training Program - Provincial and Ministerial Level (Chongqing City): Jiangjiang - A DIY Old Object Transformation Platform Integrating AI Technology (Project No.: S202312608036)

Disclosure statement

The authors declare no conflict of interest.

References

[1] Pan Y, 2021, Interpretation of the "14th Five-Year Plan" for Circular Economy Development—Accelerating the Construction of a Recycling System for Waste Materials and Building a New Pattern of Circular Economy Development. Renewable Resources and Circular Economy, 14(07): 23 + 44.

[2] Chen Y, 2023, The Circular Economy Is Increasingly Recognized. Financial Times, China.

[3] Song H, 2024, Developing a Circular Economy to Promote Sustainable Development. China Business Times, China.

[4] Ni X, 2017, The Value of Upcycling in Environmental Education. Low Carbon World, 2017(02): 24–25.

[5] China Material Recycling Association, 2024, 2024 China Recycling Industry Development Report. Resource Recycling, (07): 27–38.

[6] Du P, 2024, Technology 'Assists' in Giving New Life to Old Things. Science and Technology Daily, China.

[7] Sheng W, 2025, Looking at the Reality and Future of AI Painting from Three Perspectives. Chinese Literary Criticism, (01): 18–26 + 125.

[8] Yin S, 2023, 3D Printing: Walking Between Art and Technology. China Art News, China.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Design Discussion of a Wireless Fire Alarm System Based on Data Fusion Technology

Qun Wu*, Jinyang Wu

Chongqing Institute of Engineering, Chongqing 400056, China.

Corresponding author: Qun Wu, 18399668970@163.com

Abstract: This article explores the design of a wireless fire alarm system supported by advanced data fusion technology. It includes discussions on the basic design ideas of the wireless fire alarm system, hardware design analysis, software design analysis, and simulation analysis, all supported by data fusion technology. Hopefully, this analysis can provide some reference for the rational application of data fusion technology to meet the actual design and application requirements of the system.

Keywords: Data fusion technology; Fire alarm system; Wireless alarm; Hardware design; Software design

Online publication: March 27, 2025

1. Introduction

Wireless fire alarm is a major application direction of modern science and technology. Especially, data fusion technology plays an indispensable role and advantage [1]. Based on this, researchers need to combine the practical application requirements of the system to reasonably determine the main application ideas of data fusion technology in it. Based on this, with the help of data fusion technology, the hardware and software parts are reasonably designed to make it more flexible and intelligent. Only in this way can more advanced technical support be provided for modern fire alarms to achieve timely discovery and rescue of fire situations [2].

2. Basic design idea of wireless fire alarm system supported by data fusion technology

Under normal circumstances, the occurrence of a fire can be divided into four stages: early stage, smoldering stage, flame heat release stage, and attenuation stage. In the early stage, the flame, smoke, gas, and heat released by combustion are usually relatively small. In the smoldering stage, a large amount of smoke will be generated, but the flame heat will not be very high [3]. In the flame exothermic stage, a large amount of heat will be released and radiated, and the temperature of the fire site will rise rapidly under open-flame conditions. In the attenuation stage, as the combustible material gradually burns out, the flame will gradually become smaller, and the released smoke and heat will gradually decrease. Based on this, when a fire occurs, the wireless alarm system should try to detect the fire situation before the flame exothermic stage and issue an alarm in a timely manner [4].

To achieve this goal, in this system design, researchers mainly use wireless sensors to collect various fire data and use a wireless network composed of sensor nodes and coordinator nodes to transmit fire data to the host computer in a timely manner. With the help of data fusion technology in the host computer, the fire data is fused and processed, and an image recognition function is added to meet the timely perception of the fire situation and timely alarm requirements [5].

3. Analysis of hardware design of wireless fire alarm system

3.1. Hardware design supported by sensors and wireless transmission

In this design of the wireless fire alarm system, the designer mainly designs the wireless sensor network hardware of the system based on the CC2530 chip of the ZigBee communication protocol, which integrates A/D conversion, MCU, and RF transceiver modules, etc., making it small enough and low power consumption, suitable for all kinds of wireless sensor network nodes [6]. The sensor and wireless sensor network unit are set as sensor nodes, and three analog numerical sensors are set on each node, including CO sensor, smoke sensor, and temperature sensor. The sensor data is converted by an A/D converter and then sent to the host computer through the wireless RF module composed of CC2530 chip.

3.2. Hardware design supported by coordinator and wireless transmission

In the wireless sensor network structure of the system, the designer only sets up a coordinator, whose main role is to be responsible for the establishment, maintenance, and control of wireless sensor network nodes and to collect monitoring data from various sensors [7]. Its wireless transmission support hardware mainly includes RS232 serial port, MUC, and wireless transceiver unit. The hardware module is connected to the host computer in a wired manner to meet its data transmission needs and the specific data format definition needs to be realized under the support of the corresponding program. Compared with the sensor node, the biggest difference of the hardware is the addition of a serial port module. **Figure 1** shows the basic schematic diagram of the serial circuit of the coordinator of the wireless fire alarm system:

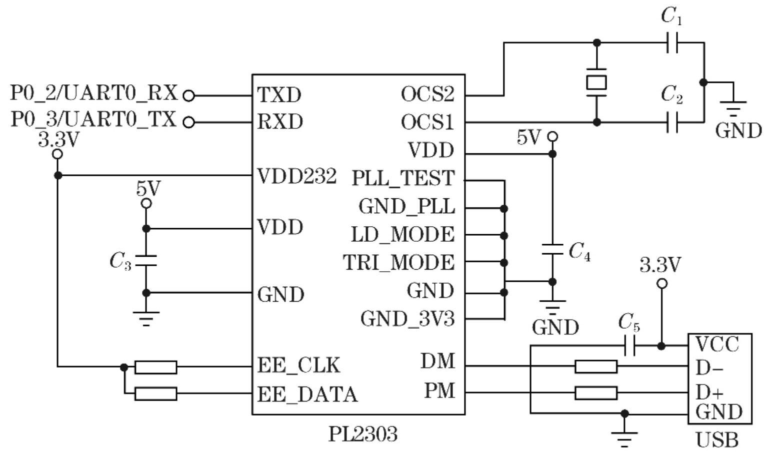

Figure 1. Schematic diagram of the basic principle of the serial port circuit of the wireless fire alarm system coordinator

4. Analysis of software design for wireless fire alarm system supported by data fusion technology

4.1. Algorithm design supported by data fusion technology

In terms of current data fusion technology for fire monitoring, the neural network algorithm is the most suitable supporting algorithm. This algorithm has a very strong nonlinear mapping capability, excellent adaptive and self-learning abilities, and good fault tolerance. This is a typical parallel processing algorithm. In practical applications, researchers can utilize this algorithm to build a perception-based classifier. Within a wireless fire alarm system, the classifier functions as a binary classifier that produces two possible outcomes: one representing valid alarm information and the other representing invalid alarm information. Based on this classifier, researchers can introduce a neural network algorithm supported by data fusion technology, using the monitoring information obtained from sensors as a basis to calculate fire alarm results in a timely and accurate manner through learning and training, thereby verifying the validity of fire alarm information [8].

There are also many forms of current neural network algorithms, among which the most representative algorithm model is the feedforward neural network algorithm. Based on this, in this system design, researchers have introduced a feedforward neural network algorithm with a hidden layer. It is a multilayer perceptron structure that can be used for nonlinear discrimination and its training algorithm is the particle swarm optimization algorithm. Considering that different types of sensors produce data within different ranges, and to prevent discrepancies in data size during actual computations, this algorithm standardizes sensor data with unknown minimum and maximum values using a Gaussian normalization algorithm [9]. The following is the algorithm formula for its normalization process:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

In this equation, z represents the standardized processing result of sensor monitoring data; μ represents the mean value of original sensor monitoring data; σ represents the variance of original sensor monitoring data. The standardized sensor monitoring data is used as input data for the neural network algorithm, ensuring that each connection within the connected hidden layers corresponds to a weight value (represented as here). The total value of each node in the hidden layer (represented as here) is used as the weighted sum of input values. The calculation formula is:

$$S_i = \sum_{j=1}^{n} w_j X_j + w_0 \quad (2)$$

In this equation, represents the total number of nodes; j represents the j-th node; represents the j-th input value; represents the initial weight value. Using the Sigmoid function as the activation function in neuron model calculation, with the support of this activation function, the neural network algorithm calculates the value in the hidden layer according to the following formula:

$$S_j = sigmoid(S_i) \quad (3)$$

For the parameters transmitted from the hidden layer to the output layer after processing, the neural network algorithm will continue to calculate them through a weighted sum approach. The calculation formula is as follows:

$$Y_j = \sum_{j=1}^{n} v_j S_j + v_0 \quad (4)$$

Based on this equation, represents the weighted sum of the output values; represents the total number of output parameters; *j* represents the jth parameter output in the output layer; represents the jth parameter value output; represents the starting output value. For the output layer, the neural networks algorithm mainly implements classification processing through the Softmax function, so that the sum of the output points after adding the values is 1.0, and the result of each output point is considered to be the basis for its content probability judgment. In this calculation mode, the output value of each output point can be expressed as:

$$Y_j = Soft\max(Y_j) \tag{5}$$

Based on the parameter values output by each output point, the neural network algorithm model can accurately determine the fire monitoring results and the probability of fire occurrence from various sensors on site. Typically, if the probability value is judged to be high, the model will use it as the fire alarm judgment result. In the practical application of data fusion technology based on neural network algorithms, the most critical process is its training. Therefore, in the specific design of the system, researchers need to make a reasonable choice of training algorithms. In the current wireless fire alarm system, the Particle Swarm Optimization (PSO) algorithm is a typical algorithm under global random optimization conditions for data fusion technology based on neural networks. Its basic principle is to use various individual components in the particle swarm to perform repeated iterative processing within the overall data space to find the optimal solution. In the process of calculation through the PSO algorithm, a series of virtual particles will be fully recorded and used as the optimal solution to the problem [10]. The neural network algorithm corresponding to each optimal solution represents a series of weights. With these weights, the neural network algorithm can make the calculated output value closest to the output value in the known training data, thus achieving scientific determination of the best weights. In each iteration, the neural network algorithm updates the speed and position of individual particles according to the following formulas:

$$v(t+1) = w \cdot v(t) + c_1 r_1 [p(t) - x(t)] + c_2 r_2 [g(t) - x(t)] \tag{6}$$

$$x(t+1) = x(t) + v(t+1) \tag{7}$$

In these equations, *t* represents the number of iterations; *w* represents the coefficient of inertia; and represent the acceleration coefficient; and represent random numbers randomly distributed in the range of [0,1]; *g (t)* represents the optimal position of all particles at the current time; *v* and *x* represent the potential solution corresponding to the desired problem. In this algorithm, each particle can record its own optimal position (represented here by *p (t)*), and *g (t)* represents the optimal position of all particles at the current time. Based on the above algorithm process, after repeated iterations, the optimal weight under the conditions of the neural networks algorithm can be accurately calculated. According to the specific calculation results, it can be timely and accurately judged and decided whether to immediately fire alarm. The closer the calculated optimal weight is to 1, the greater the probability of actual fire occurrence, and the system can combine the display results in the image recognition module to issue a fire alarm in time.

4.2. Design of the system software image recognition module

In this designed system, the image recognition program is only used to provide assistance for the fire judgment of the wireless sensor network system. In the specific design, the researchers use the color recognition model based on the RGB (red, green, and blue) flame color to obtain the color component and combine the color saturation in the HIS (hue, brightness, saturation) model to make a reasonable setting of the R component threshold. Under normal circumstances, when the R component is between 115 and 135 and the S component is between 55 and 65, if the recognized color meets the following formula, the recognized color is considered to be the flame color:

$$R > R_T \tag{8}$$

$$R \geq G > B \tag{9}$$

$$S \geq (255 - R) \cdot S_T / R_T \tag{10}$$

From these equations, represents the red component threshold identified by the system, and represents the saturation threshold identified by the system. By designing the image recognition module in the system reasonably based on the above conditions, timely and accurate identification of flame features can be achieved, and statistics and display of recognition results can be properly done, thus providing support and assistance for fire warning judgments based on neural network algorithms.

4.3. System software design for the host computer program

In this system, the main function of the host computer is to receive monitoring data from various sensors, store the data in a database, and utilize data fusion algorithms and image recognition programs to make timely and accurate judgments on detected fire situations, thereby ensuring the prompt transmission of fire alarm information. To achieve this goal, in this design, researchers mainly designed the following program for the host computer through C# programming language. After starting, the neural network algorithm module and image recognition module will start simultaneously.

After the neural network algorithm module starts, the system will detect the serial port and turn on the serial port receiving function to determine if the serial port data is correct. If it is correct, the next step will be executed; if it is incorrect, the serial port reception will continue until the serial port data is confirmed to be correct. The correct serial port data needs to be processed by the data fusion algorithm and based on the calculation results, it will be judged whether it meets the alarm criteria. If it meets the criteria, the system will immediately send an alarm; if not, it will return to the data fusion algorithm for further calculation.

After the image recognition module starts, the system will judge whether an image has been received. If an image is received, the next processing step will continue; if no image is received, the image sensor will continue to be turned on until an image is received. For the received image, after processing and recognition by the module, it will judge whether the flame color is detected. If the flame color is detected, the system will immediately send an alarm; if no flame color is detected, it will return to the image processing step.

5. Simulation analysis of the wireless fire alarm system supported by data fusion technology

For the wireless fire alarm system designed based on data fusion technology in this study, to verify its application effect, researchers simulated it with fire data through virtual simulation. During the specific simulation, researchers introduced 100 sets of normalized data and performed fusion calculations on them using the neural network algorithm supported by the system's data fusion technology. Analysis of the calculation results showed that the error could be controlled to about 3%, indicating that the accuracy of the data fusion calculation results introduced into the system is very high. **Table 1** shows the simulation results of the wireless fire alarm system supported by data fusion technology in this study:

Table 1. Simulation results of the wireless fire alarm system supported by data fusion technology in this study

Serial number	CO	Temperature	Smoke	Calculation results	Data results
1	0.750	0.900	0.220	0.943	0.960
2	0.930	0.880	0.200	0.920	0.930
3	0.780	0.380	0.310	0.765	0.780
4	0.680	0.320	0.510	0.653	0.680
5	0.350	0.210	0.300	0.360	0.350

At the same time, considering that there is a direct correlation between the accuracy of neural network algorithm calculation results and the quantity and precision of its training data, researchers need to introduce more high-precision training data to train the neural network algorithm model in subsequent practical applications. This can further enhance the accuracy of the neural network algorithm model's calculation results, fully leverage the advantages of data fusion technology, and make more timely and precise judgments on fire situations. Thus, it provides strong support for the timely detection, alarm, and suppression of fires.

6. Conclusion

In summary, data fusion technology is a critical technology in the design of modern wireless fire alarm systems, assisting in determining fire situations and making timely alarm decisions. Based on this, researchers should fully recognize the application advantages of data fusion technology and clarify its application methods. By considering the requirements of fire monitoring and fire alarm in the practical application of wireless fire alarm systems, data fusion technology should be reasonably introduced to establish corresponding algorithm models, enabling reasonable analysis and processing of fire monitoring data and providing scientific decision-making for whether to immediately execute the fire alarm procedure. Simultaneously, with the support of image recognition technology, further assistance should be provided for fire judgment and fire alarm in the system to ensure the timeliness and accuracy of fire alarms.

Funding

Chongqing Engineering University Undergraduate Innovation and Entrepreneurship Training Program Project: Wireless Fire Automatic Alarm System (Project No.: CXCY2024017); Chongqing Municipal Education Commission Science and Technology Research Project: Development and Research of Chongqing Wireless Fire Automatic Alarm System (Project No.: KJQN202401906)

Disclosure statement

The author declares no conflict of interest.

References

[1] Li Q, 2025, Automatic Fire Alarm Method for Large Public Buildings Based on Multi-Source Data Fusion. Henan Science and Technology, 2:34–37.

[2] Lei Z, 2024, Research and Design of Fire Alarm System, thesis, Xi'an University of Technology.

[3] Wang X, 2023, Research and Design of Fire Alarm System Based on Data Fusion Technology, thesis, Jilin Jianzhu University.

[4] Zhao Q, 2023, Research on Indoor Fire Warning System Based on Data Fusion, thesis, Shandong Jianzhu University.

[5] Hu B, Wang Q, Ren K, 2023, Design of Intelligent Building Fire Alarm System Based on Multi-Source Information Fusion. Journal of Shangluo University, 37(2):53–61.

[6] Tao P, Jiang Q, Song W, Liu X, Li J, 2024, Design and Application of Dynamic Simulation Technology in Fire Scene Simulation Training for Firefighters. Fire Protection Today, 9(8):35–37.

[7] Zhu P, Zhang S, Li J, 2023, Tunnel Intelligent Inspection System Based on Multi-Sensor Fusion. Science Technology and Engineering, 23(2):648–655.

[8] Huang G, Hou L, Mei S, et al., 2023, Precise Rescue System for Building Fires Based on Bluetooth and UWB Fusion Positioning. Modern Information Technology, 7(11):93–96.

[9] Xiao X, Zhang W, Zhu W, 2023, Research on a "Smart" Residential Fire Self-Rescue Guidance System Based on 5G. Computer Programming Skills and Maintenance, 11:72–75.

[10] Zhang J, 2024, Automatic Electrical Fire Alarm System Based on Video Recognition Technology. Automation Technology and Applications, 43(7):163–167.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

3D Computational Modeling and Stability Analysis of Highway Slope: A Case Study from the X104 Section in Ganxian County

Fujie Dai*, Yiwen Jin, Yongliang Wang, Jiajun Li

School of Resources and Environmental Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, China

***Corresponding author:** Fujie Dai, 2737864087@qq.com

Abstract: Highway planning requires geological surveys and stability analysis of the surrounding area. In the early stage of the survey, the modeling and stability analysis of the survey area can be carried out by using GIS software to intuitively understand the topography of the study area. The use of DEM to extract terrain factors can be used for simple stability analysis and the source data is easy to obtain, simple to operate, fast to analyze, and reliable analysis results. In this paper, taking the X104 road section in Ganxian County as an example, the ArcGIS platform is used to carry out 3D modeling visualization and stability analysis, and the stability evaluation map of the study area is obtained.

Keywords: 3D modeling; Stability; GIS; Highway planning

Online publication: March 28, 2025

1. Introduction

Three-dimensional geological modeling involves constructing geological information within a three-dimensional environment on a computer and creating models for spatial information management, geological interpretation, spatial analysis and prediction, geological statistics, entity content analysis, and graphical visualization. It is a key technology widely used in geological exploration, digital geology, geophysics, mine surveying, mining geology, and GIS [1–3]. It is an emerging discipline formed by the intersection of graphic, image, and scientific computing visualization [4, 5]. Compared with the traditional two-dimensional method, the dynamic display effect of three-dimensional geological modeling can make people understand the complex geological spatial relationship more intuitively. It can provide verification and interpretation for survey and experimental work, and the geological analysis function is powerful [6, 7]. Volume visualization technology uses 3D voxels to describe objects, including all the information inside and outside the object, which provides a strong theoretical basis and visualization approach for 3D geological modeling, but the relevant algorithms are still insufficient and need further research and improvement [8].

2. ArcGIS spatial analysis method

It consists of server components, mobile and desktop applications, and developer tools that interact with each other through the ArcGIS REST API and common file formats.

Spatial analysis is the analysis and operation of spatial graph data or other data in the spatial database, as well as the mixture between the two. Spatial analysis is the most unique function of GIS, and it is also the most essential difference between GIS and other information systems. The terrain factor analysis algorithm based on DEM is now very mature and the extraction and analysis of terrain factors can be quickly carried out by GIS software [9, 10].

3. 3D visualization of the surface model

The combination of 3D visualization technology and GIS technology has produced 3D GIS, which can query and manage spatial information and display 3D models. Whether 3D visualization can be realized is also an important difference between 3D GIS and 2D GIS [11–13]. ArcScene is the 3D visualization and analysis module of the ArcGIS Desktop product family, which can more intuitively understand the undulating form of regional terrain or interpret the location of slopes and valleys in 3D scenes than floor plans [14, 15]. Based on the DEM data of the study area, a three-dimensional geological model of the study area is established as shown in **Figure 1**.

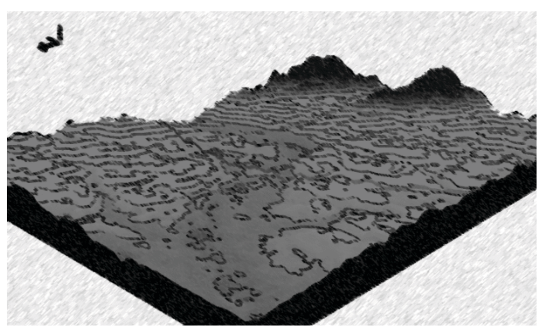

Figure 1. Three-dimensional geological model of the study area.

4. Discussion and Conclusions

Because the units and dimensions of the various terrain factor layers are different, they cannot be calculated directly. Therefore, it is necessary to reclassify and assign values to these layers before applying weights. The scale of each factor should be determined based on the actual situation to ensure consistency and accuracy in the analysis [16]. As can be seen from the **Figure 2**, in the study area, the red area is more stable and the green area is less safe. Most of the areas passed by the study section are stable areas, which are suitable for the route selection of highways, which verifies the conclusion of the stability analysis in this paper.

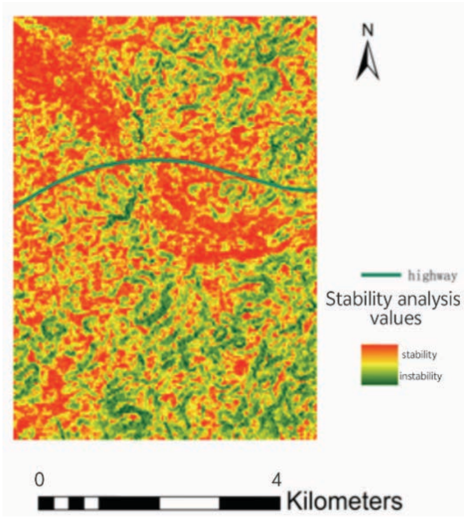

Figure 2. Stability analysis results of the study area.

In this paper, GIS spatial analysis and 3D visualization methods are used, and the main results are as follows:

(1) ArcScene has powerful 3D model modeling and display capabilities, as long as DEM data is provided, it can quickly generate a visual 3D surface model of the area, which can intuitively reflect the topography of the area and facilitate the communication between designers and construction personnel.

(2) Use the method of weighted superposition analysis to evaluate the stability of the regional slope, although the conclusion can be quickly drawn, but directly divide each factor into four grades for evaluation and there are many areas between the two grades in the results obtained, such an analysis is fuzzy, so whether the data can be processed more accurately when reclassifying and assigning each terrain factor will directly affect the reliability of the results.

(3) The use of DEM to extract topographic factors for regional stability analysis, the data source is easy to obtain, fast, and the working cycle is short, but the factors such as regional geological lithology, faults, weak structural planes, groundwater, earthquakes and rainfall are not considered. Regional slope stability evaluation is more to find the possible unstable areas in the study area, to prepare for the subsequent higher degree of investigation and evaluation work.

(4) The actual route selection of this section is basically located in the stable section, which confirms the stability analysis results of this paper.

Funding

National Undergraduate Training Program for Innovation and Entrepreneurship (Project No.: 202310407006)

Disclosure statement

The authors declare no conflict of interest.

References

[1] Wang B, Shi B, Song Z, 2009, A Simple Approach to 3D Geological Modelling and Visualization. Bull Eng Geol Environ, 68: 559–565.

[2] Wu Z, Guo F, Li J, 2019, The 3D Modelling Techniques of Digital Geological Mapping. Arabian Journal of Geosciences, 12: 467.

[3] Wu X, Liu G, Weng Z, Tian Y, Zhang Z, Li Y, Chen G, 2021, Constructing 3D Geological Models Based on Large-Scale Geological Maps. Open Geosciences, 13(1): 851–866.

[4] Deng Y, Zhang J, Sun Y, Tian Y, Chen Q, Qiu B, 2024, Key Technology for Section Analysis of 3D Geological Models Based on the Online Virtual Earth Environment. Earth Science Informatics, 17: 441–453.

[5] Wang X, Guo J, Fu S, et al, 2024, Towards Automatic and Rapid 3D Geological Modelling of Urban Sedimentary Strata from a Large Amount of Borehole Data Using a Parallel Solution of Implicit Equations. Earth Science Informatics, 17: 421–440.

[6] Zhu L, Li M, Li C, et al, 2013, Coupled Modeling Between Geological Structure Fields and Property Parameter Fields in 3D Engineering Geological Space. Engineering Geology, 167: 105–116.

[7] Shang H, Shen Y, Li S, et al, 2023, An Automated Mapping Method of 3D Geological Cross-Sections Using 2D Geological Cross-Sections and a DEM. ISPRS International Journal of Geo-Information, 12(4): 147.

[8] Fang H, Liu Y, Shi B, Wu Z, Wang B, 2002, Three-Dimensional Geological Modeling and Its Engineering Application. Hydrogeology & Engineering Geology, 2002(03): 52–55.

[9] Wang Y, 2019, Extraction and Analysis of Topographic Factors on the Loess Plateau in Northern Shaanxi Province Based on DEM. Western Development (Land Development Engineering Research), 4(12): 28–36.

[10] Zhu H, 2003, Research on Spatial Data Mining of Digital Elevation Model (DEM), thesis, Northwestern University.

[11] Sirakov N, Granado I, Muge F, 2002, Interpolation Approach for 3D Smooth Reconstruction of Subsurface Objects. Computers & Geosciences, 28: 877–885.

[12] Rabinowicz M, Boulegue J, Genthon P, 1998, Two- and Three-Dimensional Modeling of Hydrothermal Convection in the Sedimented Middle Valley Segment, Juan de Fuca Ridge. Journal of Geophysical Research, 103: 24045–24065.

[13] Du F, 2010, Research and Application of 3D Visualization of Cities Based on ArcScene, thesis, Xi'an University of Science and Technology.

[14] Calcagno P, Chiles J, Courrioux G, et al, 2008, Geological Modelling from Field Data and Geological Knowledge: Part I. Modelling Method Coupling 3D Potential-Field Interpolation and Geological Rules. Physics of the Earth and Planetary Interiors, 171: 147–157.

[15] Shen D, Wong D, Camelli F, 2013, An ArcScene Plug-in for Volumetric Data Conversion, Modeling and Spatial Analysis. Computers & Geosciences, 61: 104–115.

[16] Huang D, 2016, Slope Stability Analysis Based on GIS and Limit Equilibrium Principle, thesis, North China Electric Power University (Beijing).

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research and Design of Intelligent Inspection System for Thermal Power Plants

Wei Zhang*, Tingfeng Zhang

College of Electrical Engineering, Liaoning University of Technology, Jinzhou 121001, China

***Corresponding author:** Wei Zhang, 1326001933@163.com

Abstract: To meet the demand for intelligent and unmanned development in thermal power plants, an intelligent inspection system has been designed. This system efficiently performs inspection tasks and monitors the operational parameters of key equipment in real-time. The collected data is uploaded to the monitoring center, allowing operation and maintenance personnel to access equipment information promptly. Data analysis is used to provide fault warning and diagnosis for critical equipment. The system employs the Pure Pursuit algorithm, which effectively avoids obstacles and ensures path continuity and stability. Simulation results show that the Pure Pursuit algorithm significantly improves the navigation accuracy and task efficiency of the inspection robot, ensuring the reliability of thermal power plant inspections.

Keywords: Thermal power plants; Intelligent inspection; Parameter acquisition; Path planning

Online publication: March 28, 2025

1. Introduction

Currently, the reliability of primary equipment in thermal power plants is relatively low. It cannot fully meet the requirements for unmanned operations. When issues such as equipment failure, gas leaks, or fires occur, they cannot be detected and addressed in time. This poses serious risks to both personnel and equipment safety and may even result in casualties [1]. Traditional inspection methods rely on manual operations. Personnel must regularly visit each inspection point with handheld instruments to check electrical equipment and collect a large amount of operational data. This approach is physically demanding, inefficient, and unreliable. Additionally, the complex working environment increases the risk of accidents and injuries during inspections [2].To address the need for intelligent and unmanned operations in thermal power plants, it is necessary to develop and implement intelligent inspection systems. These systems can collect and monitor equipment data in real time, detect safety hazards promptly, and eliminate them before they escalate [3].

The research and application of intelligent inspection systems are gaining increasing attention in industrial production. These systems significantly enhance the predictability of equipment performance and provide valuable references for equipment condition monitoring and evaluation [4–8]. Moreover, intelligent inspection systems can replace manual inspections. They reduce labor costs, improve inspection efficiency, minimize safety risks, and enhance overall productivity [9–15]. This paper introduces the design of an intelligent inspection system for thermal power plants. The system efficiently performs inspection tasks, monitors key equipment parameters in real time, and uploads the data to a central monitoring system. It enables operators and maintenance personnel to access up-to-date equipment information and, through data analysis, achieve fault prediction and diagnosis for critical equipment. The system uses the Pure Pursuit algorithm, which effectively avoids obstacles and ensures path continuity and stability.

2. Design of the intelligent inspection system for thermal power plants

The intelligent inspection system is built on a mecanum wheel chassis, with the STM32 microcontroller as the MCU (Microcontroller Unit). The system integrates multiple functional modules, including a temperature detection module, gas detection module, navigation module, video detection and image processing module, OLED display module (for real-time display of data such as temperature, harmful or flammable gas presence, and smoke concentration), motor drive module, audio-visual alarm module, and communication module. The overall design of the system is shown in **Figure 1**.

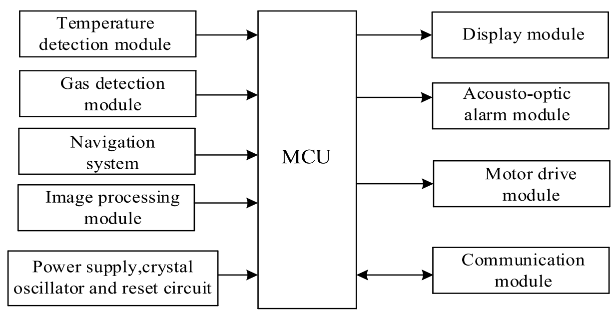

Figure 1. Overall system design

3. Design of key functional modules in the intelligent inspection system

3.1. Positioning and navigation

The intelligent inspection system uses laser navigation technology. It determines the robot's position in the global coordinate system through a laser sensor. The system adopts artificial landmarks for positioning. By using a rotating laser sensor to detect the landmarks in the environment, the position and orientation of the sensor in the global coordinate system are calculated using trigonometric geometry.

3.2. Gas detection and infrared temperature measurement

The gas detection module in the system uses an MQ-2 smoke and gas sensor module. The MQ-2 sensor is suitable for detecting various non-combustible smoke, natural gas, and liquefied petroleum gas. It has high sensitivity, especially for detecting non-combustible smoke containing carbon dioxide or hydrogen gas, and offers strong anti-interference performance. The temperature detection module uses infrared thermal imaging technology for online monitoring. It performs real-time temperature measurement of key components such as pipelines, valves, and other related equipment in the power plant.

3.3. Image processing module

The image processing module is designed to recognize parameters displayed on analog or digital meters in the thermal power plant, such as thermometers and liquid level gauges. The process of on-site instrument parameter recognition is shown in **Figure 2**. These parameters are converted into digital signals and uploaded to the monitoring center in real time. The module uses the Open MV image processing module. Open MV is a fully independent and programmable vision processing module with strong extensibility. It is designed for video and machine vision processing in embedded systems. The module is reprogrammed using C language, enabling advanced data structures to quickly handle complex video images and textual outputs. This makes it a highly efficient vision sensor for embedded applications.

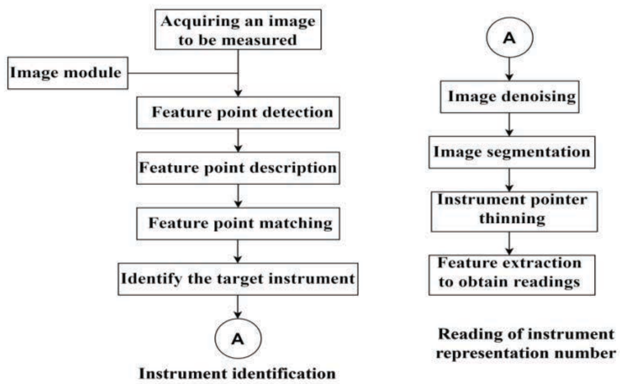

Figure 2. Flowchart of on-site instrument parameter recognition.

3.4. Vibration and audio signal analysis and collection

The vibration and audio signal collection module is designed to collect real-time vibration and audio signals from on-site equipment such as motors, fans, and generators. It analyzes the collected data to monitor whether the equipment's vibration and noise levels are normal. Using fault diagnosis algorithms, the system can detect abnormalities in real time. It provides early warnings of potential issues, allowing preventive measures to be taken in advance and avoiding operational accidents.

4. Path tracking algorithm of the intelligent inspection system

4.1. Grid map construction

Grid maps are commonly used in autonomous navigation for inspection robots. These maps are typically created using a combination of Radar sensors and SLAM algorithms, particularly in indoor environments with unknown conditions. However, during the actual map-building process, Radar sensors are affected by noise and motion distortion. This can result in significant discrepancies between the generated grid map and the real environment, which hinders effective path planning for the inspection robot. To address this issue, this paper combines manual measurements with the grid map generation algorithm in MATLAB to create a 2D grid map. The specific steps are as follows:

(1)Manual measurements are used to determine the overall dimensions of the pump room, the exact size of drainage equipment and the locations of the equipment within the pump room. Based on this information, a 2D floor plan of the pump room is drawn, as shown in **Figure 3**.

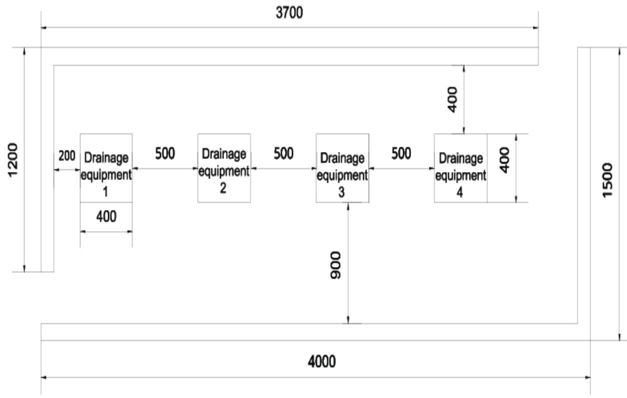

Figure 3. Geographical layout of the pump room

(2) Based on the 2D floor plan of the pump room, the obstacles and navigable areas are digitized into a 15 × 40 matrix. In this matrix, the value "0" represents navigable areas, while "1" represents obstacle areas. The digitized 2D layout is shown in **Figure 4**.

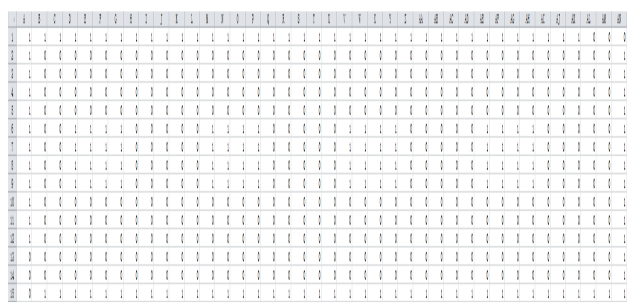

Figure 4. Digitized 2D layout

(3) The positions of the equipment in the pump room are known and not changed frequently. Therefore, combining manual measurements with MATLAB's grid map generation algorithm is a feasible approach to build the 2D grid map of the pump room. The resulting 2D grid map is shown in **Figure 5**.

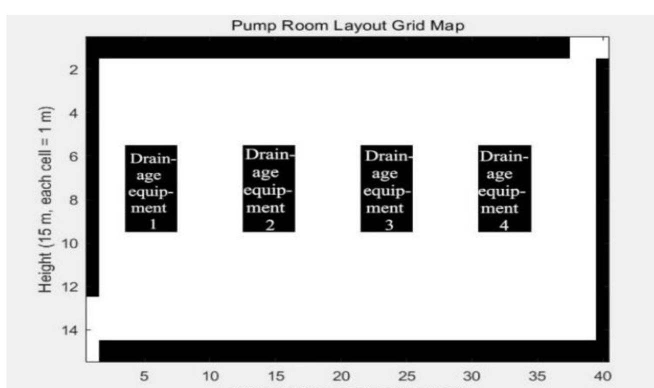

Figure 5. Equipment grid map

4.2. Path planning based on the Pure Pursuit algorithm

4.2.1. Principle of the Pure Pursuit algorithm

The Pure Pursuit algorithm is used to implement path planning for the intelligent inspection system. It ensures that the inspection system moves smoothly along the planned path. **Figure 6** shows the principle of the Pure Pursuit algorithm.

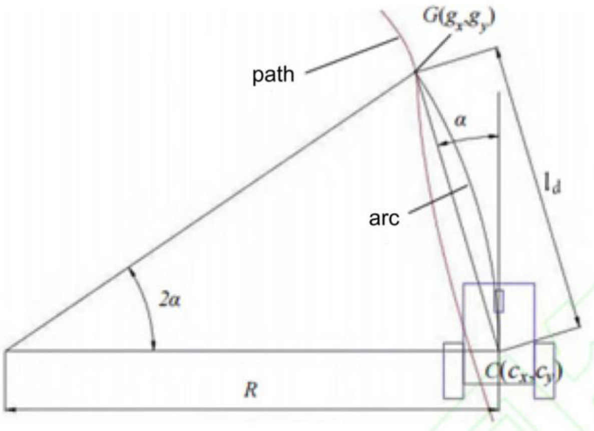

Figure 6. Principle of the Pure Pursuit algorithm

In **Figure 6**, $G(g_x, g_y)$ represents the next target point to be tracked. This point lies on the planned global path. The current position of the reference point C is given as (c_x, cy). The distance from the current position of the inspection system to the target point is denoted as l_d(m). The angle α(°) represents the angle between the vehicle's heading and the direction of the target point. Using the distance formula between two points, ldl_d can be calculated as follows:

$$l_d = \sqrt{(g_x - c_x)^2 + (g_y - c_y)^2} \tag{1}$$

$$v(t) = \frac{V_R + V_L}{2} = k_v l_d \tag{2}$$

In Equation 2, k_v represents the relationship between v_t (the velocity) and l_d (the distance to the target point). Using the law of sines, the following equation can be derived:

$$\frac{l_d}{sin(2\alpha)} = \frac{R}{sin\left(\frac{\pi}{2} - \alpha\right)} \tag{3}$$

$$\frac{l_d}{2\,sin(\alpha)\,cos(\alpha)} = \frac{R}{cos(\alpha)}$$

$$\frac{l_d}{sin(\alpha)} = 2R$$

Thus, the required turning angle for the inspection system can be calculated as:

$$\alpha = arcsin\left(\frac{l_d}{2R}\right) \tag{4}$$

4.2.2. Application of the Pure Pursuit algorithm

Figure 7 shows the path planning results for the pump room using the Pure Pursuit algorithm. The algorithm calculates the distance and angle between the reference point and the target point in real time. It dynamically adjusts the robot's direction to ensure smooth path progression. Combined with the laser navigation module, the Pure Pursuit algorithm effectively avoids static obstacles and maintains path continuity and stability. Simulation results show that the Pure Pursuit algorithm significantly improves the navigation accuracy and task efficiency of the inspection robot, ensuring the reliability of thermal power plant inspections.

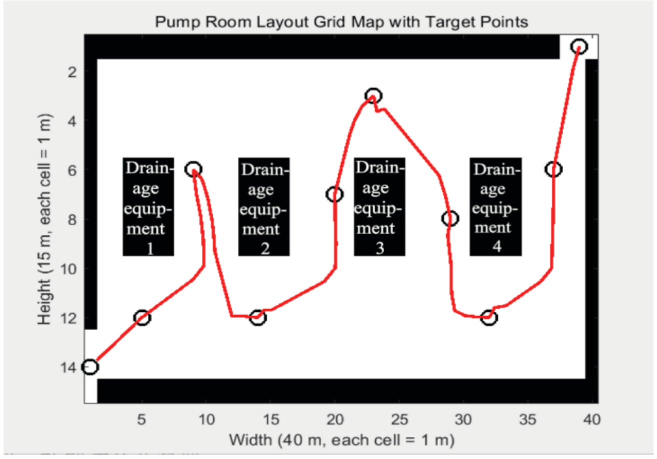

Figure 7. Path planning with the Pure Pursuit algorithm

5. Conclusion

The intelligent inspection system designed in this paper has autonomous path planning and navigation capabilities. It can collect and monitor thermal power plant equipment information in real time, accurately detecting the operational parameters of key equipment. The data is uploaded to the monitoring center, ensuring that operation and maintenance personnel are promptly informed of the equipment status. Through data analysis, the system achieves fault warning and diagnosis, effectively identifying and eliminating safety risks. The system uses the Pure Pursuit algorithm, which effectively avoids obstacles and ensures path continuity and stability. Simulation results show that the Pure Pursuit algorithm significantly improves the navigation accuracy and task efficiency of the inspection robot, ensuring the reliability of thermal power plant inspections. Additionally, the application of the intelligent inspection system reduces production costs, minimizes safety hazards, and increases production efficiency. It meets the demand for intelligent and unmanned development in thermal power plants.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Zhou L, Zhang Y, Sun Y, et al., 2011, Research and Application of Intelligent Substation Inspection Robots. Automation of Electric Power Systems, 35(19): 85–88 + 96.

[2] Delobel L, Aufrere R, Debain C, et al, 2019, A Real-Time Map Refinement Method Using a Multi-Sensor

Localization Framework. IEEE Transactions on Intelligent Transportation Systems, 20(5): 1644–1658.

[3] Keqiang J, Zhizhou S, Yongcheng L, et al, 2017, Development and Application of the Rail-Type Inspection Robot Used in Substation Rooms. MATEC Web of Conferences, 139(5): 00210.

[4] Wu X, Li D, Liu T, et al, 2020, Research on Anti-Fall Algorithms of Substation Intelligent Inspection Robots. Advances in Intelligent Systems and Interactive Applications.

[5] Wang C, Yin L, Zhao Q, et al, 2020, An Intelligent Robot for Indoor Substation Inspection. Industrial Robot, 47(5): 705–712.

[6] Chen X, Ren G, 2020, Key Technologies and Development Trends of Intelligent Manufacturing and Robot Application. IOP Conference Series: Earth and Environmental Science, 461(1): 012049.

[7] Xiao P, Zhang C, Feng H, et al., 2010, Research on GPS Navigation of Substation Inspection Robots. Sensors and Microsystems, 29(08): 23–25+28.

[8] Ren A, Sun Y, Ma T, 2009, Research on DGPS/DR Combined Navigation Technology for Mobile Robots. Machinery and Electronics, 2009(03): 58–61.

[9] Wei P, Zhang Z, Zhang C, et al., 2009, Research on Navigation System of Unattended Substation Inspection Robots. Automation and Instrumentation, 24(12): 5–8+25.

[10] Zhao J, Feng S, Sun T, et al, 2020, Functional Design and Application of Intelligent Robot Technology in Coal-Fired Smart Power Plants. Energy Technology, 2020(4): 35–42.

[11] Hu X, Wang F, 2021, Research on Path Planning and Optimization Technology of Intelligent Robots. Robot Technology and Applications, 43(2): 15–22.

[12] Delobel L, Aufrere R, Debain C, et al, 2019, A Real-Time Map Refinement Method Using a Multi-Sensor Localization Framework. IEEE Transactions on Intelligent Transportation Systems, 20(5): 1644–1658.

[13] Lin H, Xiang D, Ouyang J, et al, 2021, Research Review on Path Planning Algorithms for Mobile Robots. Computer Engineering and Applications, 57(18): 38–48.

[14] Patle B, Babu L, Pandey A, et al, 2019, A Review: On Path Planning Strategies for Navigation of Mobile Robots. Defence Technology, 15(4): 582–606.

[15] Song X, Chen C, Fan Q, et al, 2019, Application Research on New Inspection Robots Based on Intelligent Sensors. Henan Science and Technology, (10): 19–21.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research on the Prevention and Control Measures of Soil and Water Conservation in Water Conservancy Projects

Zhiyong Chang[1], **Yicui Yang**[2]*

[1]Guangxi Vocational College of Water Resources and Electric Power, Nanning 530023, Guangxi, China
[2]Guangxi Communications Design Group Co., Ltd, Nanning 530029, Guangxi, China

Corresponding author: Yicui Yang, yangyicui168@163.com

Abstract: With the development of our country's social economy, the construction scale of water conservancy project has had an obvious expansion. In the construction of water conservancy projects, certain impacts on the surrounding water and soil conditions are inevitable. These impacts may lead to problems such as soil erosion, which can directly affect local production, livelihoods, and the natural ecological environment on which people depend. In severe cases, such issues may even hinder the progress and quality of the water conservancy project itself. Therefore, in the construction of water conservancy projects, soil and water conservation work is extremely important. Based on this, this paper mainly aimed at the prevention and control of water and soil conservation of water conservancy projects launched the relevant analysis and research.

Keywords: Water conservancy project; Water and soil conservation; Prevention and control

Online publication: March 28, 2025

1. Introduction

With the rapid development of water conservancy engineering construction in China, soil and water conservation has gradually attracted wide attention [1]. While water conservancy engineering provides important support for economic development and social progress, it also inevitably brings environmental problems such as soil erosion. These problems will not only cause a certain negative impact on the ecological balance, but also affect the long-term stability and security of water conservancy projects. Therefore, it is of great significance to strengthen the exploration and research on the prevention and control measures of water and soil conservation in water conservancy projects.

2. The harm of soil and water loss

Soil and water loss refers to the process of removing soil from its original place in the form of water flow under the influence of natural erosion, human activities, and other factors, leading to land degradation and ecological environment deterioration [2]. This kind of loss not only affects the productivity of the land but also causes serious damage to the environment. Specifically, the harm of soil and water loss is mainly manifested as follows:

First, it will reduce agricultural productivity. Soil erosion will directly weaken the soil quality of cultivated land and greatly reduce the fertility of soil, which will have a certain impact on the growth of crops and lead to a decline in agricultural output, which is not conducive to the sustainable development of the agricultural industry. Second, it will aggravate the problem of water shortage. Soil erosion will lead to a large amount of sediment is washed into the river, the increase of sediment makes the transparency of the water body decreased, will not only affect the water quality, but also reduce the storage capacity of the reservoir and the available water resources, which will pose a threat to the safety of water supply, and even affect people's health.

Furthermore, it will exacerbate flooding. One of the most immediate consequences of soil erosion is the destruction of regional vegetation. Once vegetation is damaged, during the rainy season, the limited ground cover combined with steep slopes makes it difficult to effectively intercept and slow down surface runoff. As a result, the soil's ability to retain water is significantly reduced, increasing the risk of severe flooding [3]. However, if the precipitation intensity exceeds the permeability of the soil surface, it will lead to the formation of runoff on the ground, which may even cause serious flooding disasters and ultimately pose a greater threat to people's lives and property safety [4].

Additionally, it will destroy the surrounding ecological environment. Soil erosion is usually accompanied by changes in the structure of the ecosystem. For example, sediment accumulation can lead to the degradation of wetlands and water environments, resulting in reduced biodiversity. Additionally, the long-distance transport of sediment may damage ecosystems far from the erosion source, such as obstructing waterways and disrupting the habitats of aquatic organisms. These are direct manifestations of ecological and environmental degradation.

Lastly, it will greatly reduce the economic and social benefits of water conservancy projects. Once the sediment and plant residues produced by soil erosion enter the downstream river or reservoir with the water flow, it will reduce the flood discharge capacity of the river and the storage capacity of the reservoir. In serious cases, there may even be problems such as dam collapse, which will have a certain impact on the economic and social benefits of the water conservancy project.

3. Causes of soil and water loss in water conservancy projects

In the construction of water conservancy projects, the main reasons for soil erosion are as follows:

(1) During the construction period, operators generally need to frequently dig soil or move rocks; these operations are often accompanied by the risk of soil erosion [5]. In addition, if the operators do not pay enough attention to the disposal of waste soil and rock, it usually leads to the risk of soil erosion.

(2) In the construction process of water conservancy projects, it is usually necessary for operators to use large rolling equipment to compact the road. However, this approach can easily cause significant changes to the construction site's topography, landforms, surface vegetation, and other natural features. In severe cases, it may even disrupt surrounding underground runoff, greatly reducing the stability of the soil or mountain structure [6]. In this way, when encountering rainfall or prolonged gravitational effects, the ground is prone

to severe soil erosion problems.

(3) During construction, the ground and vegetation are usually damaged to varying degrees, which will not only upset the original balance between the two, but also reduce the erosion resistance of the soil structure. As a result, in the event of heavy rainfall or continuous erosion by daily water flow, it will directly lead to serious soil erosion.

4. Water and soil conservation work management technology

4.1. Structural control technology

In water conservancy projects, structural control technology is an important means of soil and water conservation work, mainly including engineering measures and ecological measures, aimed at effectively preventing and controlling soil erosion. The commonly used structural prevention and control means mainly include dam body, revetment, and so on [7]. Specifically, the dam body is usually set on the slope and river bed, and the relevant staff needs to be able to design its slope and height according to the terrain characteristics and hydrological conditions. Under normal circumstances, the height of the dam body is controlled at 1–3 meters to effectively intercept watershed runoff and reduce sediment loss.

Revetment technology is primarily a preventive and control measure aimed at mitigating soil loss along both sides of a river. Common methods include gabions, concrete blocks, and vegetation slope protection. Among these, gabion revetments can effectively enhance the stability of riverbanks, resist wave impact and river erosion. Typically, groups of gabions are installed at intervals of 50 meters, and trees are planted every 900 meters to form green belts, thereby reducing soil erosion [8]. Vegetation slope protection involves planting drought-resistant and deep-rooted plants to consolidate the soil, which not only helps beautify the environment but also effectively enhances the soil's resistance to erosion.

In addition, soil reinforcement and soil consolidation technology is also an important part of structural control technology. This technology mainly utilizes materials such as geotextiles, geomesh, and reinforced earth walls to enhance soil stability, reduce the risks of rainwater infiltration and soil liquefaction, and achieve more effective soil consolidation.

4.2. Vegetation restoration and control technology

Vegetation restoration and control technology is an effective technical means for soil and water conservation. In water conservancy projects, relevant workers can reduce soil erosion and improve soil quality by using plant roots to consolidate soil and conserve water sources. Under normal circumstances, we can choose some adaptive and developed roots of native plants for vegetation restoration, such as ash, Chinese ash, neem, etc., in order to effectively enhance the ability of soil erosion resistance. As for the planting density of these plants, it can usually be set to 300 to 500 plants per mu, so that the biomass per unit area can be effectively increased while ensuring the coverage [9]. When implementing the plant configuration of soil and water conservation functions, relevant staff need to fully take into account the growth habits and ecological niches of different plants, and rationally mix deep-rooted and shallow rooted plants to form a multi-level root network, such as strip, layered, or ribbon configuration. In the face of steep slopes and areas prone to landslides, plants with strong drought and wind resistance should be selected as far as possible, such as wild rhododendron and honeysuckle, to enhance slope stability.

4.3. Comprehensive treatment technology

In the comprehensive management of water and soil conservation in water conservancy projects, it can be achieved mainly through the following measures:

(1) Source control

Source control mainly involves the reasonable planning and management of land use, actively promoting and publicizing crop rotation, fallow systems, and relevant laws and policies, while continuously enhancing public environmental awareness. These measures aim to improve soil structure, strengthen its ability to conserve water and soil, and ultimately reduce soil erosion.

(2) Process control

Process control refers to measures taken during specific stages to mitigate soil erosion. For example, during flood discharge, it is essential to carefully consider factors such as the volume of water released, the impact on flood discharge capacity, waste accumulation, and riverbed exposure. In flood prevention efforts, the construction of flood-resistant structures should prioritize vegetation restoration in idle areas and ensure the protection of flood-prone sections. Additionally, increasing vegetation coverage through afforestation and grassland restoration can effectively reduce surface runoff and enhance the soil's water storage capacity, thereby improving soil and water conservation during the process.

(3) Terminal management

Terminal management focuses on strengthening the monitoring and evaluation of soil and water resources to achieve dynamic and continuous management. In the process of soil and water conservation, a comprehensive monitoring network should be established to track changes in soil and water loss. By leveraging advanced technologies such as remote sensing, GIS, and other technical tools, real-time monitoring of soil erosion and the impact of water conservancy project construction can be achieved. Based on monitoring data, targeted regional control measures should be formulated according to the specific conditions of different areas, to ensure that water conservancy projects effectively fulfill their role in soil and water conservation [10].

4.4. Engineering protection technology

In the process of water conservancy project construction, soil and water conservation control can also be achieved by using engineering protection technology. For example, during construction, cut-off ditches can be set up to discharge wastewater and prevent its uncontrolled flow. In addition, when implementing soil and water conservation measures in material collection yards, materials can be gathered along designated lines or areas to minimize disturbance and reduce the unnecessary waste of water resources [11]. At the same time, the operators should reasonably formulate the mining mode according to the actual situation and establish a certain drainage channel, as far as possible to reduce the waste of resources and prevent the loss of water resources. This is the best way to better prepare for the later greening and flood discharge during the flood season.

5. Water conservancy project soil and water conservation prevention and treatment measures

5.1. Forestry measures

Forestry measures are an important measure of soil and water conservation prevention and control in water conservancy projects, which can effectively reduce soil and water loss and maintain ecological balance. First of all, the existing forest cover should be protected to ensure its integrity and health, which is an effective measure to reduce the risk of soil exposure to wind and water erosion. For degraded forests, forest coverage and stability can be continuously improved through forest restoration and vegetation reconstruction, to better maintain soil fertility and water resources as well as achieve the purpose of gradually restoring soil and water conservation ability [12]. Among them, during this process, relevant personnel should reasonably manage forest cover, including optimizing and adjusting forest structure and density, as well as selecting and configuring appropriate tree species and vegetation. These measures can effectively reduce the impact of wind on soil erosion and slow down water flow velocity, thereby enhancing the soil's ability to resist erosion and improving overall ecological stability.

Secondly, forest fire prevention should be properly carried out by establishing firebreaks, regularly clearing combustible materials, and strengthening monitoring and management. These measures help reduce the occurrence of forest fires, thereby avoiding soil erosion and ecological damage caused by forest burning. And for the forest land that has been felled, exploited, or destroyed by other ways, the effect of forestry prevention and control should be effectively improved through land reclamation, soil restoration, afforestation, and other ways. Finally, we can also use the way of forest tending and management to strengthen the regular management and effective management of forest thinning, renewal, and tending activities, to further strengthen the effect of forestry prevention and management measures.

5.2. Agricultural measures

Agricultural measures play an important role in the prevention and control of water and soil conservation in water conservancy projects. From the present point of view, the common agricultural measures mainly include vegetation cover, intercropping, as well as soil and water conservation tillage technology. For example, by maintaining vegetation cover and intercropping, the risk of soil exposure to rain and water erosion can be greatly reduced. Furthermore, by introducing integrated forest and grassland agricultural management systems, such as farmland shelterbelt networks and gully forest networks, the soil environment can be effectively stabilized, thereby reducing soil erosion and enhancing ecological resilience. In addition, farmers in Shangri-La and other areas will take advantage of the local terrain characteristics to build terraces, which is also an important measure to solve the problem of soil erosion [13].

5.3. Engineering measures

Engineering measures play a key role in the prevention and control of water and soil conservation in water conservancy projects. At present, the common engineering measures mainly include gully control, slope protection engineering, construction of drainage system, construction of shelterbelt and so on [14]. For example, by building fixed structures or taking vegetation protection measures, soil slope erosion can be effectively avoided; Through the construction of drainage system, the water can be rationally discharged, which is conducive to avoiding soil liquefaction or landslide. IIn addition, by constructing earth and rock dams in the upper reaches of channels and rivers, rainfall runoff can be effectively intercepted, the flow velocity reduced, and the risk of soil erosion and scouring mitigated. This approach not only enhances soil moisture and provides a continuous water source but also contributes to stabilizing the soil environment and promoting vegetation restoration.

5.4. Ecological measures

At present, common ecological measures mainly include wetland protection and restoration, biodiversity protection, forest ecological restoration, and so on. For example, by protecting wetlands and restoring damaged wetlands, soil hydrologic regulation and water purification functions can be effectively improved, and the risk of soil erosion can be greatly reduced. A series of biodiversity protection measures such as the establishment of protected areas and the protection of endangered species can well maintain the ecological balance and improve the stability of the ecosystem, to achieve the purpose of reducing soil erosion [15].

However, in this process, attention should not only be paid to the fertilization and irrigation of plants, but also to the prevention and control of diseases and pests. On the one hand, during the process of plant fertilization and irrigation, it is essential to ensure the proper application of fertilizers and rational use of water to provide sufficient soil nutrients and moisture for plants in the early stages of growth, thus preventing plant death caused by drought. Techniques such as drip irrigation or sprinkler irrigation can be adopted to improve water resource utilization efficiency and reduce unnecessary water waste.

On the other hand, in the control of plant diseases and pests, biological control and physical control can be used, such as the introduction of natural enemies insects to control pests. By setting traps to catch adult insects, the frequency of use of chemical pesticides can be effectively reduced, which is conducive to further improving the stability of the overall ecosystem. In addition, in the process of water conservancy engineering construction and vegetation conservation, attention should also be paid to the protection of soil structure. Heavy machinery should avoid directly working on the slope, to reduce soil compaction and improve soil permeability as well as water retention ability. At the same time, in the management of vegetation restoration, a regular monitoring mechanism is set up to evaluate plant growth and soil quality, and timely adjustment of conservation measures according to the monitoring results.

6. Conclusion

In short, in the view of green ecology, the development of the country's social economy must not be based on the destruction of ecological environment. Therefore, in the construction of water conservancy projects, the relevant operators should pay full attention to the development of soil and water conservation work as well as actively take scientific and effective prevention and control measures to protect the surrounding ecological environment. Specifically, the use of structural control technology, vegetation restoration and control technology, comprehensive management technology, engineering protection technology, engineering protection technology and other technical means should be used to strengthen forestry measures, agricultural measures, engineering measures, ecological measures of water conservancy engineering soil and water conservation effect, to promote our water conservancy to achieve high quality and sustainable development.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Liu J, 2024, Influence of Soil and Water Loss in Hydraulic Engineering Construction and Its Prevention Strategies. Water Safety, 2024(24): 166–168.

[2] Zhang Q, 2024, Analysis of Soil and Water Conservation and Sustainable Development in Hydraulic Engineering Construction. China Water Transport, 2024(24): 100–101+121.

[3] Gu Z, 2024, Application of Soil and Water Conservation in Hydraulic Engineering Construction. Value Engineering, 43(26): 158–161.

[4] Li F, 2024, Application of Soil and Water Conservation Technology in Hydraulic Engineering Construction. Science and Technology of Shanxi Soil and Water Conservation, 2024(1): 19–21.
[5] Luo R, 2023, Analysis of Soil and Water Conservation and Sustainable Development in Water Conservancy Project Construction. Agricultural Development and Equipment, 2023(10): 111–113.
[6] Wang C, 2023, Research on Prevention and Treatment of Water and Soil Conservation in Hydraulic Engineering. Water Safety, 2023(10): 76–78.
[7] Mao M, 2023, Analysis of Soil and Water Conservation and Sustainable Development in Water Conservancy Project Construction. Shaanxi Water Resources, 2023(4): 98–100.
[8] Chen A, 2023, Existing Problems and Countermeasures of Soil and Water Conservation in Farmland Water Conservancy Construction. Control of Huai River, 2023(3): 64–65 + 68.
[9] Hu S, 2023, Analysis on the Application of Soil and Water Conservation Technology in Water Conservancy Project Construction. Control of Huai River, 2023(1): 63–64.
[10] Li S, 2022, Case Study on Soil and Water Conservation Measures in Hydraulic Engineering. Value Engineering, 41(28): 8–10.
[11] Li X, 2022, Discussion on Comprehensive Control Measures of Soil and Water Loss in Water Conservancy Project Protection. Agricultural Development and Equipment, 2022(8): 122–124.
[12] Zhang G, 2022, Discussion on Soil and Water Conservation Measures in Water Conservancy Project Construction Area. Southern Agriculture, 16(12): 211–213.
[13] Liu A, 2021, Analysis of Soil and Water Conservation Management During Water Conservancy Project Construction Period. China High-Tech, 2021(23): 139–140.
[14] Lin X, Deng F, 2021, Exploration of Soil and Water Conservation Design in Water Conservancy Project Construction. Shaanxi Water Conservancy, 2021(11): 134–135 + 138.
[15] Fan M, 2021, Comprehensive Control Measures of Soil and Water Loss in Water Conservancy Project Protection. China Science and Technology Information, 2021(22): 57–58.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research on the Design Scheme of "Boundless Vision" Social App for Visually Impaired People

Jiazhen Zhang[1], **Dan Ni**[1,2]*, **Meihui Song**[1], **Liuyan Chen**[1]

[1]City Institute, Dalian University of Technology, Dalian 116600, Liaoning, China
[2] Batangas State University, Batangas 4200, Calabarzon, Philippines

***Corresponding author:** Dan Ni, nidandan1989@gmail.com

Abstract: The "Boundless Vision" project focuses on creating novel user interfaces to improve the experience of visually impaired individuals when using smart devices. By incorporating features such as high-contrast color schemes, enlarged text, and voice control, an intuitive and accessible interface is developed. Moreover, the project leverages online volunteer services to offer telephone and video guidance, aiding visually impaired users in navigating urban environments and accessing information. Through a comprehensive process involving needs assessment, UI design, system development, and service evaluation, the project aims to boost the independence of visually impaired individuals and foster greater societal awareness of visual impairment challenges. Ultimately, the project will deliver a suite of user interface designs that can serve as a model for advancing inclusive technology across society.

Keywords: Visually impaired people; UI design; User interface; Navigation service; Volunteer platform

Online publication: April 2, 2025

1. Introduction

With the rapid development of technology, smart devices have gradually become indispensable in modern life. However, traditional interfaces often fail to meet the needs of visually impaired people. Many assistive tools on the market are still complex and lack clear feedback, causing difficulties for visually impaired users. This limits their social participation and autonomy. Therefore, developing suitable interfaces is crucial for improving their quality of life.

The project aims to provide convenient and efficient smart device experiences for visually impaired individuals through innovative UI design and an online volunteer service platform [1]. It focuses on understanding their needs, optimizing interface design with multi-sensory feedback, and offering navigation services to enhance independence and self-confidence. The platform also utilizes social resources to ensure high-quality volunteer services, promoting inclusiveness and fairness in the digital age.

2. Requirement analysis and user research

2.1. Investigation on the needs of visually impaired individuals

Visually impaired people face many challenges when using smart devices. A demand survey found that their main needs are simplifying operations, enhancing interface readability, and adding voice interaction [1]. Existing interfaces are often too complex, with small fonts and insufficient contrast, making them hard to use and read. Traditional touchscreen operations lack tactile feedback, making it difficult to locate and operate interface elements accurately [1]. They rely more on voice recognition, but existing systems often have low recognition rates and slow responses.

2.2. Major challenges faced by visually impaired individuals when using smart devices

2.2.1. Issues concerning interface design

When visually impaired people use smart devices, interface design issues are particularly prominent. Traditional interfaces often have small fonts and low contrast, making it hard for users to distinguish content [2]. Button and icon designs lack operability optimization, with insufficient voice prompts or tactile feedback, leading to confusion. Existing UI designs do not fully consider the cognitive characteristics and habits of visually impaired users, resulting in complex menus and cumbersome operations. To solve these problems, designers need to use high-contrast colors, large fonts, voice control, and simple layouts.

2.2.2. Challenges of navigation services

For visually impaired individuals, navigation services are crucial for enhancing independence and security [2]. Existing services face challenges in real-time accuracy and detailed guidance. Voice navigation often fails to update promptly or provide sufficient road information in complex urban environments. Navigation is usually limited to single feedback modes, which may not meet diverse needs in different settings [2]. For example, video navigation can be more effective indoors by providing spatial details. Current systems lack features like dynamic obstacle alerts and facility guidance. Diversified navigation services need optimization based on user habits and feedback to improve safety and convenience.

2.2.3. The core issues reflected by users' feedback

Based on feedback from visually impaired users, core issues lie in device operability, service real-time nature, and personalized user experience. Many users report that existing devices lack simplicity, with complex interfaces and inaccurate voice recognition leading to frequent misoperations. Voice navigation provides some assistance but often lacks timely feedback, especially in noisy environments where voice commands are hard to recognize. Personalized settings are insufficient, as different users have varying needs such as adjusting voice speed, volume, font size, or contrast. Collecting this feedback provides valuable references for future design improvements.

3. Analysis of characteristics and demands of visually impaired individuals

3.1. Physiological characteristics

The physiological characteristics of visually impaired individuals are mainly reflected in the absence or limitation of visual perception, which leads to the inability to obtain visual information normally. The degree of visual impairment varies, ranging from complete blindness, low vision or visual field loss. Completely blind individuals cannot perceive light or images, while those with low vision can only partially perceive light but cannot clearly identify objects. Individuals with visual field loss can only see a local area. Visually impaired individuals rely on hearing, touch, and smell to make up for the deficiency in vision and are particularly sensitive to sounds and identify object features through touch. When designing interfaces for visually impaired users, services that are in line with their perception methods, such as large fonts, high-contrast color schemes, and voice prompts, should be provided.

3.2. Psychological Characteristics

The psychological characteristics of visually impaired individuals are closely related to their life experiences and environmental adaptability. They have strong auditory, tactile, and spatial perception abilities, allowing unique environmental perception. They exhibit high independence and self-management awareness, forming suitable life patterns and coping strategies. They often have strong resilience but may also experience depression or anxiety. Technological products should consider their emotional needs, providing functions that enhance self-confidence and reduce anxiety.

3.3. Behavioral Characteristics

The behavioral characteristics of visually impaired individuals reflect their adaptive habits. They rely on other senses for strong environmental perception and spatial orientation, using hearing to judge object positions. When using devices, they prefer voice input and auditory feedback over visual displays. In daily activities, they adopt a cautious gait and rely heavily on tactile perception for object information. Designing intelligent products should focus on optimizing tactile feedback and simplifying operations to reduce errors.

4. UI design and function development

4.1. Principles and concepts of UI design

For UI design targeting visually impaired individuals, the first principle is simplicity and intuitiveness [3]. The interface should avoid complex elements and information overload, allowing quick understanding and operation. High contrast and large fonts are essential to ensure clear perception of content. Accessibility and operability must be considered, ensuring smooth interaction through voice recognition or touchscreen operations. Voice and tactile feedback enhance user-friendliness, helping users confirm operations or obtain information. Design should be optimized based on user feedback to better meet actual needs.

4.2. Design process

The design process is crucial for effective UI design [4]. It starts with analyzing requirements to understand the needs and habits of visually impaired users through surveys and interviews. Based on this analysis, designers establish UI principles like simplicity, color contrast, font size, and voice interaction. Next, they create prototypes and collect user feedback for adjustments. Finally, designers work with developers to implement and refine the design based on feedback, ensuring the final product meets user needs. The process emphasizes user participation and multiple iterations to maximize user experience.

4.3. Interface design and user experience

In the interface design of this project, we focused on ensuring accurate and timely information transmission for visually impaired users. We adopted a high-contrast color scheme and large fonts to enhance visibility.

Buttons and function items have voice description functions, providing clear feedback after the user clicks. Tactile feedback, such as vibration prompts, helps confirm operation completion. The design emphasizes user-centeredness, simplicity, and efficiency, ensuring a pleasant and smooth user experience through repeated testing and optimization.

4.4. Design and development of navigation system

In this project, the navigation system aims to provide precise and real-time route guidance for visually impaired individuals [4]. It integrates telephone and video navigation: telephone navigation uses voice commands for clear location and path guidance, while video navigation combines real-time images and environment recognition for intuitive guidance in complex environments [5]. The system employs AI for accurate path planning and dynamic route adjustments. It supports personalized settings like voice speed, volume, and language type [6]. Enhanced voice recognition ensures timely and clear feedback, improving travel convenience and independence for visually impaired users.

4.5. Design case presentation

As shown in **Figure 1**, this design interface adopts a style that is highly concise and has a high contrast ratio. The high-contrast approach enables visually impaired people to perceive color differences more easily and perform simple operations.

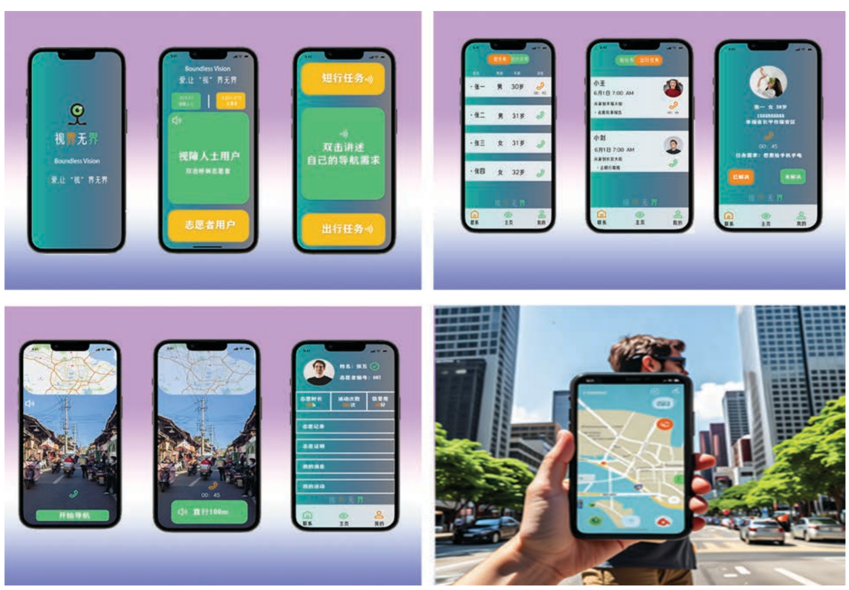

Figure 1. The "Boundless Vision" Social App UI Design

5. Service testing and optimization

5.1. User testing and feedback collection

During the project development process, user testing is crucial to ensure service feasibility and determine optimization directions. Visually impaired individuals conducted functional tests, providing feedback on UI interface and navigation system accuracy and usability [6]. Users encountered issues with voice recognition accuracy, navigation clarity, and interface layout [7]. Feedback was collected through interviews, phone followups, and online surveys. Regular user discussions and experience activities provide valuable data for design adjustments.

5.2. Service optimization and adjustment

Based on user feedback, we have made multiple optimizations to the service. We optimized the voice input module and adopted advanced voice recognition technology to improve accuracy and efficiency. We added real-time location updates and optimized voice navigation to ensure smoother guidance in complex environments. Interface design was simplified, improving button layout, font size, and contrast for better readability and convenience. A regular update and maintenance mechanism was added to continuously optimize the system.

6. Summary

Through user testing and feedback, we have successfully created a UI design and navigation system for visually impaired people. Continuous feedback collection and real-time optimization ensure the design's rationality and convenience. We will maintain user interaction and continuously improve the product. This project enhances autonomous mobility for the visually impaired and provides new solutions for visual impairment assistance. We expect this service to be promoted to more cities, promoting accessible environment construction and inclusive social development.

Funding

The 2024 Class A Project of the University-level College Students' Innovation and Entrepreneurship Training Program in Liaoning Province, titled "Unlimited Vision - UI Design for Visually Impaired People"(Project No.: X202413198012).

Disclosure statement

The author declares no conflict of interest.

References

[1] Liu C, Wei R, 2025, Research on Optimization Design of Human-Computer Interaction Interface Based on Visual Communication Technology. Electronic Design Engineering, 2025: 3–4.

[2] Wang H, Zhang X, Li S, et al., 2025, Design of Assisting Glasses for the Blind Based on Machine Vision and TGAM Module. Information Recording Materials, 2025: 1–2.

[3] Lin Y, 2024, Design Strategies of Mobile UI Interface Based on Visual Communication Design. Footwear Technology and Design, 2024: 1–2.

[4] Wang L, 2024, Application of Visual Communication Design in Mobile UI Interface Design. Da Gu, 2024: 1–2.

[5] Wang X, 2024, Voice Navigation for Braille Maps Facilitates Blind People's Travel. Shenzhen Special Zone Daily, 2024: 1.

[6] Lu N, Luo K, Wang Q, et al., 2024, Research on Design of Intelligent Navigation System for Blind People's Travel. Internet of Things Technology, 2024: 2–3.

[7] Liang D, 2020, Application of Visual Communication Design in Mobile UI Interface. Computer Products and Circulation, 2020: 1–2.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

The Exploration of the Application of Electronic Circuit Simulation Technology in Integrated Circuit Design

Shengjie He*

Tongji University, Shanghai 200092, China

***Corresponding author:** Shengjie He, 2250323@tongji.edu.cn

Abstract: With the rapid development of Internet technology, the application of electronic circuit simulation technology is more and more extensive, and now it has been applied to integrated circuit design. Because the electronic circuit simulation technology has high efficiency, flexible and simple application, as well as stable performance, it has shown more and more good application prospects in integrated circuit design. Based on the strong development trend of electronic circuit simulation technology, it will be more and more widely used in daily life in the future, so the research on electronic circuit simulation technology is more and more in-depth. In this paper, the application of electronic circuit technology in integrated circuit design is studied, hoping that the technology can provide a more concise and efficient research and development way for electronic applications.

Keywords: Electronic circuit; Simulation technology; Integrated circuit

Online publication: March 28, 2025

1. Introduction

In the era of information technology, electronic products have become deeply integrated into all aspects of people's production and daily life. As a result, electronic circuit simulation technology has undergone continuous advancements, now providing strong support for the efficient design and development of integrated circuits. Integrated circuit is an important part that affects the function and operation of electronic products. To meet the needs of rapid upgrading of electronic products, integrated circuit design must keep up with the pace of The Times and ensure the development of electronic products. The classification of integrated circuits is mainly based on their density and size. The application of electronic circuit simulation technology in integrated circuit design offers new approaches for advancing electronic products. Therefore, electronic product designers should focus on integrated circuits during the design process, flexibly apply chip-level system concepts to product development, and continuously integrate and optimize circuit functions to achieve higher performance and greater miniaturization.

It is integrated within the chip, which shortens the development cycle of electronic products and provides strong support for their research and development. In this paper, the concept, function, and characteristics of electronic circuit simulation technology are analyzed, and the specific application strategy of this technology in integrated circuit development, project development, and circuit function is discussed. Finally, the future development trend of this technology is analyzed to provide some help for related research.

2. Overview of electronic circuit simulation technology

Electronic circuit simulation technology, based on virtual reality, extends system simulation by relying on advanced computer system technology and requiring robust hardware and software platforms. It reflects the functional role of integrated circuit analysis, utilizing digital modeling and numerical simulation methods to replicate and analyze the operating states of electronic circuits. It can simulate the operation of the integrated circuit and then support the circuit to improve the accuracy. At present, the development of electronic equipment requires not only equipment manufacturing and commissioning but also the identification and timely resolution of problems during the development process. Due to the complexity of this manufacturing and commissioning workflow, the error rate is relatively high [1]. At present, the application of electronic circuit simulation technology in the development of electronic products reduces the workload of test work and improves the accuracy of test results.

At this stage, the integrated circuit performance test mainly relies on the tester to carry out. Compared with the manual test, the tester detection not only has high reliability, but also the test is more standardized and safer. The tester can simulate various software application scenarios, identify and correct problems in a timely manner, thereby improving the development efficiency of integrated circuits. This process not only promotes the continuous innovation of integrated circuit design concepts but also ensures the operational stability and functional diversity of integrated circuits, laying a solid foundation for the advancement of electronic components and the design of electronic products. In addition, electronic circuit simulation technology can accurately reflect the actual structure and operational characteristics of electronic circuits, enabling the integration, construction, and simulation of complete circuit systems. This provides strong support for R&D personnel in optimizing integrated circuit design, enhancing design efficiency and accuracy [2].

3. Function and feature analysis of electronic circuit simulation software

3.1. The function of electronic circuit simulation software

There are two kinds of electronic circuit simulation software, one is OrCAD PSpice Designer software, the system is mainly composed of circuit simulation, component processing, and schematic processing of three parts which is based on Spice general language written component-level simulation software to carry out work that is widely used in electronic information retrieval. The software builds a circuit scheme and virtual circuit components, which can accurately describe the circuit components and present the circuit details. The software system can support circuit structure analysis, circuit operation simulation, and provide circuit debugging suggestions [3]. However, its analysis and calculation time is long, and the system simulation convergence is poor.

The second kind of electronic software is Saber software. It features advanced simulation technology and adaptable operating conditions, making it not only widely applied in the field of power electronics but also demonstrating excellent performance in areas such as machinery and optoelectronics. The software can be compatible with hybrid simulation and analyze problems from multiple dimensions. The software can support integrated circuit operation simulation, data visualization analysis, etc., but its shortcomings lie in the low simulation success rate and complex operation [4].

3.2. The characteristics of electronic circuit simulation technology

3.2.1. High efficiency

In the past, the testing of electronic products relied on manual operation, but the test method has more security risks and rough details, so the testing of electronic products under manual operation is more difficult. However, with the application of electronic circuit simulation technology, the testing of electronic products has become more comprehensive and precise [5]. Because electronic circuit simulation technology can simulate various scenarios in the operation and application of electronic products, it can also correct the shortcomings of electronic products in time, so it shortens the development cycle of electronic products and provides strong support for the research and development of electronic products.

3.2.2. Diversity

The development of electronic products is more complex, debugging, production, and modification of the links of work may affect the quality of products. Electronic circuit simulation technology can be applied to all aspects of electronic product and integrated circuit development, supporting automatic debugging and providing more accurate modification schemes to promote the design and optimization of integrated circuits and electronic products. At the same time, due to the diverse application scenarios of integrated circuits and electronic products, electronic circuit simulation technology can simulate different application scenarios, provide more effective optimization suggestions for circuit simulation, ensure the quality of electronic products, and extend the life of electronic products [6].

3.2.3. Stability

With the development of information technology, the functions of electronic products are becoming more and more complex and the density of integrated circuits is increasing. Researchers apply chip-level system design technology to the development of electronic products, continuously enhancing the functional complexity of electronic products while significantly improving their operational stability. In the future, electronic circuit simulation technology will play an increasingly vital role in the development of electronic products.

4. Electronic circuit simulation technology for the embodiment of electronic application value

4.1. Development of integrated circuits

Electronic circuit simulation technology has brought revolutionary changes to integrated circuit design and promoted the rapid development of integrated circuit research. With the continuous development of information technology, electronic products have put forward higher requirements for integrated circuit design. As a key component of electronic products, integrated circuits should be able to support the operation of electronic products in complex scenarios. The electronic circuit simulation technology can ensure that the integrated circuit can be used in various working scenarios and meet the actual needs by simulating various working scenarios and debugging the integrated circuit repeatedly. Electronic circuit simulation technology updates the research and development of IC and ensures the function and efficiency of IC [7].

4.2. Project development and circuit functions

After the development of an electronic product, it will not be directly mass-produced and put into use but must undergo feasibility testing to ensure that its safety and operational performance meet relevant standards. Only after passing these tests can the product enter the mass production stage and be officially applied. The electronic circuit simulation technology replaces the manual operation of the electronic product feasibility test so that the detection work efficiency is higher and the accuracy is higher. At the same time, the electronic circuit simulation technology can simulate the operation of the product through virtual simulation, so the detection is more comprehensive and efficient. The adjustment of parameters of high temperature, high voltage, and high current will have an important impact on the function of integrated circuits. By setting parameters in the electronic circuit simulation system, the system will simulate the operation of products in special environments and test the limit value of the circuit, which can improve the service ability of the circuit [8]. It is evident that applying circuit simulation technology to the testing of electronic products can significantly enhance the scientific rigor and accuracy of the testing process while also shortening the research and development cycle. It can be seen that electronic circuit simulation technology has many advantages in the development of integrated circuits.

4.3. Effective verification of circuit functions

Designing an integrated circuit scheme is the first and most crucial step in the entire research and development process of integrated circuits. The design must be scientific and functionally complete, and it must undergo various feasibility verifications to ensure it meets fundamental application requirements. Once the precision of components and circuits in electronic products changes, it will have a certain impact on the function and performance of the product. Therefore, it is important to carry out feasibility tests for integrated circuits. The application of electronic circuit simulation technology can effectively enhance the feasibility verification of electronic applications. By combining electronic circuit simulation technology with traditional circuit feasibility analysis, developers can conduct comprehensive testing of electronic products, promptly identify functional errors or issues that deviate from actual conditions, thereby reducing the likelihood of circuit design failures. This approach also enhances the functional stability of electronic products to a certain extent [9].

4.4. Circuit design optimization

Temperature has a great impact on the stability of electronic equipment and the change of temperature can easily cause the change of the internal components of the chip, thus affecting its performance. The application of electronic circuit simulation technology can effectively avoid the above problems. By using simulation technology to replicate different temperature environments, developers can observe circuit changes, analyze circuit characteristics, and make targeted improvements. This process ultimately leads to the development of high-performance electronic devices. Electronic circuit simulation technology can simulate various environments or conditions that are difficult to observe in actual operation, helping designers analyze circuit behavior under extreme conditions and optimize the circuit design accordingly. The reliability of the circuit in practical application is ensured by conducting virtual tests on the circuit [10].

5. The specific application of electronic circuit simulation technology in integrated circuit design

Based on the electronic circuit simulation technology, users can set different parameters in the system according to the work needs, simulate various operating scenarios, and provide ideas for integrated circuit debugging and design. The integrated circuit operation is simulated and its structure is described to support R&D personnel to optimize the technology. At the same time, through the simulation system, it can also design and process its frequency response, output signal, etc., to study and analyze in the future and optimize the research results. Compared with traditional integrated circuit design methods, the application of electronic circuit simulation technology significantly reduces the time and cost of circuit design and development, while greatly improving product quality and performance. The following is an analysis of the application of electronic circuit simulation technology in integrated circuit design.

5.1. PSPICE simulation software

In practice, Cadence's PSPICE simulation software is widely used in large-scale integrated circuit design. The software can support repeated digital analog mixed simulation of the circuit and provide the corresponding correction suggestions so that the performance of the entire circuit is more perfect. The main function of PSPICE software is to simulate the circuit or text file so that researchers can understand various performance indicators of the circuit according to the results of simulation processing [11].

5.2. The specific functions of PSPICE simulation software

PSPICE simulation software has two basic functions: the basic analysis module (referred to as PSPICE AD) and the advanced analysis module (referred to as PSPICE AA). The basic analysis module primarily includes DC analysis, AC analysis, time domain analysis, and other functions. The advanced analysis module mainly includes several functions such as parameter scanning, temperature analysis, and worst-case analysis. Its simulation function is mainly the following modes:

(1) Noise analysis simulation refers to the process of analyzing the equivalent output noise and equivalent input noise of a given input signal within a specified frequency range and along a defined calculation path.

(2) DC bias simulation, in the presence of inductance and capacitor disconnect two cases in the line, is used to determine the steady-state operating point of the line. In the transient process of small interference and transient analysis, it will automatically find the DC balance point so as to provide the necessary initial conditions for the transient analysis, and the model parameters are nonlinearized.

(3) DC scan analysis refers to analyzing and calculating the DC output characteristics of a circuit as a specific parameter varies within a defined range.

(4) AC sweep analysis is used to calculate the amplitude-frequency and phase-frequency characteristics of small signals, as well as the input and output impedance of the circuit.

(5) Parameter scanning analysis involves analyzing the circuit's performance as the parameters of specific components vary according to a defined pattern or rule.

The basic process of simulation using PSPICE software includes the following steps: drawing the circuit schematic, setting simulation parameters, running the simulation, observing and analyzing the simulation results, adjusting the circuit, and modifying the simulation parameters accordingly.

5.3. The application of electronic circuit simulation software in integrated circuit design

5.3.1. Modeling classification and simulation process

PSpice simulation software is commonly used in the design of large integrated circuits. During the circuit design process, PSpice enables mixed-signal simulation testing, allowing designers to repeatedly modify the circuit based on test results. This iterative process helps to fully ensure that the performance indicators of the designed circuit meet the required design standards. The main function of PSpice simulation software is to simulate circuits and text files, etc., and allow relevant personnel to apply the final simulation results to various parameter data and performance of product circuit design [12].

5.3.2. PSpice simulation technology function

(1) Electronic circuit simulation function design for electronic products

In the product-oriented LSI design, the PSpice simulation software of Cadence is used to simulate and analyze all kinds of complex circuit performance indicators, to obtain the corresponding circuit data and realize real-time visualization processing. In this way, it is convenient for the circuit designer to observe the simulation results during simulation and analysis. At the same time, the system can also evaluate the circuit design and various performance indicators of the whole system efficiently in a more intuitive way. In short, the simulation software mainly includes two parts: Basic Analysis Module (AD) and Advanced Analysis Module (AD). On the basis of this basic analysis module, the simulation and analysis of the performance of the DC circuit and AC circuit are completed, and the time domain of the entire circuit is comprehensively analyzed [13].

In practical application, the method of noise simulation analysis is used to set the frequency, output equivalent, and input noise level of the test circuit. In addition, PSpice can analyze the bias point of DC circuits and evaluate the short-circuit factor, enabling the calculation of the circuit's operating state under short-circuit conditions or capacitor failure, thus accurately determining the overall condition and reliability of the circuit. On this basis, the transient characteristics of the system are further verified. From the perspective of high-level analysis module, this system mainly includes three modules: worst simulation analysis, temperature analysis, and parameter scanning. Finally, using the scanning function, the DC output characteristics of the product can be displayed. It can also be summed up and judged from the change of electrical parameters in the circuit. The simulation system has powerful functions such as numerical analysis, matrix calculation, scientific data visualization, modeling and simulation of nonlinear dynamic system, etc.

The R&D personnel can apply electronic simulation software to verify the calculation results and apply the system to automatically obtain the results of multivariate equations, differential equations of dynamic circuits, and equations of phases, etc. Electronic simulation experiments can be used to carry out online simulation experiments. Saving time in preparing experimental materials can significantly improve experimental efficiency [14]. For example, the application of Multisim to carry out the design of digital circuit board, directly input the graphics of circuit schematic diagram and circuit hardware description language into the system, and carry out online simulation, which greatly improves the work efficiency.

(2) Circuit design simulation process

The testing process of PSpice simulation software generally follows these steps: First, the simulation circuit of the electronic product is described, ensuring that all circuit nodes are clearly marked within the system, and simulation parameters are properly set. Second, the output waveform generated by the electronic simulation is recorded, and the simulation results are carefully observed and analyzed to determine whether further debugging of the product circuit is necessary. Finally, through iterative simulation debugging, a complete and optimized circuit design scheme is formed.

Taking a sinusoidal oscillation circuit as an example, the application of electronic circuit simulation technology in its design can be described as follows: under specific excitation conditions, the transient response of the circuit is calculated, and the resulting transient data is saved in a file with a ".dat" extension. By using the analysis functions within the simulation window, the signal waveform obtained from the simulation can be visually analyzed. During the circuit design simulation experiment, the proposed integrated circuit design scheme is input into the electronic circuit simulation system. Through the circuit input module (Schematics) of PSpice, the operation of the electronic circuit is simulated, and a circuit diagram of the multi-harmonic oscillator is generated. This multi-harmonic oscillator, typically implemented using a 555 timer chip, provides output through its pin 3 (OUT), effectively demonstrating the circuit's oscillation behavior.

(3) DC operation mode

The PSpice software system will label and analyze the DC operating point of the integrated circuit design scheme. During this operation, the staff can select the menu bar of the software, click "Open", and open the function of the DC operating point of the circuit in the pop-up dialog box. Then, the system can realize the automatic analysis of the parameters of the DC operating point of the circuit. The system generates the analysis results and saves them in the output file to support the researchers to query.

(4) Fourier analysis

Based on the Fourier analysis module in the simulation system, researchers can analyze the DC waveform and harmonic components of the circuit. In the Fourier analysis module, researchers input the fundamental frequency of the integrated circuit, after which the system automatically performs harmonic calculations and saves the analysis results in a file. In this way, when the first spike is the output signal frequency, its digital parameter is 4 KHz, and the second spike's second harmonic component is 8 KHz, so that the fourth spike is the fourth harmonic component, its value is 16 KHz. Finally, through comparative analysis, it is found that the frequency shown is consistent with the reality.

6. The future development trend of the application of electronic circuit simulation technology

The application of electronic circuit simulation function design technology has opened a new road for electronic applications. It can significantly improve the development efficiency of electronic products, shorten the research and development cycle and reduce costs while ensuring the quality of electronic products through simulation technology. Although the electronic circuit simulation function design technology in the development of computers only accounts for a small part, it is also a part that can not be ignored. Its use on the development and change of electronic applications has played a role in promoting, and this technology is constantly improving and perfecting. The current electrical simulation technology can only be used for the simulation of circuits and computer hardware systems, but not for the digital system of computer CPU. However, in the near future, the simulation technology of electronic circuits will be further improved and perfected to make the algorithm more accurate and to simulate the program of the CPU [15].

CPU program simulation is a form of electronic system simulation. With the continuous advancement and upgrading of circuit simulation technology, the application of simulation technology in CPU program development is expected to be realized in the future, enabling full simulation of program operation states. This will bring transformative changes to the research and development of electronic products. Based on the current development trend of electronic circuit simulation technology, it can be predicted that the future should be along the direction of shortening the research and development time of electronic products and improving the research and development efficiency. It is believed that in the future, with the continuous improvement and refinement of electronic circuit simulation function design technology, the development methods of electronic products will become increasingly advanced, and development tools will become more efficient, greatly shortening the research and development cycle of electronic products.

The application of electronic circuit simulation function design technology ensures that electronic product research and development personnel can quickly discover the development direction and improvement methods, so the application of this technology has an important role in promoting the development of electronic applications. From the perspective of electronic circuit simulation function design technology, the technology still has a great space for development, after continuous improvement and improvement, the technology will play a more important role in the development of electronic applications, therefore, R&D personnel should continue to explore in the actual work, the technology to further improve, so that electronic circuit simulation technology better serve the People's Daily life.

7. Conclusion

To sum up, as the structure and functions of electronic products become increasingly complex, the functional requirements for integrated circuits are also rising. At the same time, with the rapid replacement and upgrading of electronic products, the research and development of integrated circuits must continuously improve efficiency to keep pace with the demands of modern society. Based on electronic circuit simulation technology, integrated circuit (IC) research and development can simulate circuit operation under various application scenarios, providing strong support for optimizing circuit design concepts and improving R&D efficiency.

As an information tool, electronic circuit simulation technology overcomes the limitations of traditional manual testing, making system testing more standardized and safer. It greatly shortens the research and development cycle of integrated circuits, reduces development costs, and provides strong support for the advancement of electronic product development. It can be said that the application of electronic circuit simulation technology opens a new chapter in China's electronics industry, and promotes the research and development of electronic products to take a big step forward.

In the future, leveraging the advantages of efficient operation, flexible application, and stable performance of electronic circuit simulation technology, the development of electronic products is certain to move toward a faster and more efficient direction. However, electronic circuit simulation technology still has considerable room for development, both in terms of technical improvement and optimization. Its application scope also needs to be further expanded. Continuous support and innovation from relevant research and development personnel are essential to advance this technology, drive the sustained growth of China's electronics industry, and better serve the improvement of people's well-being and quality of life.

Disclosure statement

The author declares no conflict of interest.

References

[1] Zhao W, Wei P, 2012, Application of PSpice Simulation Technology in Electronic Circuit Design. Journal of Shaoxing University of Arts and Sciences (Natural Science), 32(09): 25–29.

[2] Mao J, Shen H, Gu J, et al., 2016, Application of LabVIEW FPGA Real-Time Simulation in Modern Power Electronic Technology Experiment Teaching. Laboratory Research and Exploration, 35(09): 108–113.

[3] Ma G, Wang J, Pan B, 2018, Application Research of Virtual Simulation Technology in Electronic Technology Course Teaching. Compilation of Scientific Research Results of Research on Teacher Education Capacity Building, College of Electronic Countermeasures, National University of Defense Technology.

[4] Wang X, Zhang X, Yin J, et al., 2019, Instrument Availability Evaluation for Nuclear Power Plant Based on Circuit Simulation Technology. Progress Report on Nuclear Science and Technology in China. Progress Report on Nuclear Science and Technology in China: Proceedings of the 2019 Annual Conference of the Chinese Nuclear Society, School of Nuclear Science and Engineering, North China Electric Power University.

[5] Sun D, Zhang X, Zhao T, et al., 2020, Exploration of Electronic Circuit Experiment Teaching Based on Virtual Simulation Technology. Journal of Higher Education, 2020(35): 108–111.

[6] Zhang J, 2020, Exploration on Application of Electronic Circuit Simulation Technology in Integrated Circuit Design. Electronic Components and Information Technology, 4(10): 94–96.

[7] Su J, Liu J, Jin Y, et al., 2020, Research on the Value of Electronic Circuit Simulation Technology in Electronic Application Development. Paper Making Equipment & Materials, 49(04): 118–119.

[8] Liu M, 2024, Analysis on the Function of Electronic Circuit Simulation Technology in Electronic Application Development. Electronic Components and Information Technology, 8(01): 76–78 + 82.

[9] Mo Z, Deng R, Yu H, et al., 2023, Application of Online Virtual Simulation Technology in Electronic Circuit Experiment. Digital Technology and Application, 41(6): 68–70.

[10] Cao T, 2023, Analysis on the Role of Electronic Circuit Simulation Technology in Electronic Application Development. Electronic Components and Information Technology, 7(3): 29–31 + 40.

[11] Ma C, Zhao M, Meng X, et al., 2019, Research on the Role of Electronic Circuit Simulation Technology in Electronic Application Development. Science and Technology Communication, 11(11): 120–121.

[12] Fu N, 2019, Application of Circuit Simulation Software in Electronic Technology Teaching Practice. Science and Technology Information, 17(15): 194–195.

[13] Hu G, 2019, The Role of Electronic Circuit Simulation Technology in the Development of Electronic Applications. Electronic Testing, 2019(8): 137–138.

[14] Luo H, 2019, Application of Virtual Simulation Technology in Electronic Technology Teaching in Secondary Vocational Schools. Wireless Internet Technology, 16(4): 151–152+166.

[15] Zhao Y, Li B, 2019, Application of EveryCircuit Circuit Simulation Software in Electronic Technology Teaching. Computer Knowledge and Technology, 15(6): 112–113.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Elevators in the Context of Green Manufacturing

Chenshi Zhong*

Guangdong Engineering Polytechnic, Guangzhou 510000, Guangdong, China

***Corresponding author:** Chenshi Zhong, zhongchenshi@gdep.edu.cn

Abstract: With the continuous development and wide application of modern technology, more enterprises and institutions in our country have begun to strengthen the research and manufacturing of intelligent elevators, aiming to build a new mode of modern manufacturing with the concept of green development as the core. Therefore, this paper takes green manufacturing as the background, analyzes the design significance of intelligent elevator lightweight, expounds its design ideas, and in the end introduces its specific design practice path.

Keywords: Green manufacturing; Intelligent elevator; Lightweight

Online publication: March 28, 2025

1. Introduction

Intelligent elevator lightweight refers to reducing the elevator's volume, weight, cost, and other factors while maintaining its essential functions. This is achieved by using new materials, optimizing structural design, and enhancing technology, ultimately improving the elevator's efficiency and convenience. It can be seen that strengthening the lightweight design of intelligent elevators is one of the important means to promote the realization of green manufacturing goals in China.

2. The significance of lightweight design of intelligent elevator under the background of green manufacturing

In the context of green manufacturing, the lightweight design of intelligent elevator can first effectively reduce the overall weight of the elevator. The structure of the elevator system is relatively complex and its optimization design usually requires comprehensive consideration of a variety of factors, which is difficult. The advanced structural optimization design scheme and the application of new materials can be organically combined to reduce the weight of the elevator while ensuring the performance and safety of the elevator [1]. Secondly, the lightweight design of intelligent elevator can effectively reduce the energy consumption of elevator operation and reduce the cost of manufacturing and maintenance.

Moreover, while the elevator is lightweight, it can also take into account the comfort, safety, and functionality of the elevator, which is the concrete embodiment of achieving the goal of green manufacturing. Finally, the intelligent elevator lightweight design, through the integration of intelligent simulation technology, enables precise simulation and accurate prediction of elevator structure and material usage. This not only enhances the efficiency and accuracy of intelligent elevator green design but also contributes to the development of a more refined elevator lightweight design and manufacturing system, ultimately promoting technological progress within the industry [2].

3. The light weight design idea of intelligent elevator under the background of green manufacturing

3.1. Analyze the market demand and technical background of elevator lightweight, and determine the direction of lightweight research

With the continuous acceleration of urbanization and the increasing number of densely populated areas, the use of elevators is increasing [3]. In this situation, the advantages of lightweight elevators are gradually highlighted. Compared with ordinary elevators, lightweight elevators have lower energy consumption, higher energy efficiency, and lower maintenance costs, which can better meet the development needs of the social market under the background of green manufacturing. Moreover, with the continuous development of material science, structural design, manufacturing process, and other fields, people's research and application level of lightweight materials has been significantly improved, which provides more technical support for elevator lightweight [4]. For example, carbon fiber composite materials have the characteristics of high strength, high stiffness and light weight, and have been widely used in elevator lightweight design [5]. The development and application of new manufacturing technologies such as 3D printing have also, to a certain extent, provided some new ideas and methods for the design and manufacture of lightweight elevator parts [6]. In addition, coupled with the policies and systems related to energy conservation and environmental protection, green building, and sustainable development issued by the national government, it also provides support and guarantee for the development of lightweight elevator [7]. It can be seen that the current market demand for elevator lightweight is relatively large, and the technical support is relatively rich, which lays a solid foundation for us to determine the direction of elevator lightweight research.

3.2. Learn from successful cases and advanced technologies to study the status quo and development trend of elevator lightweight at home and abroad

According to the published information, the Ultra Rope developed by KONE Company is an ultra-light, ultra-wear-resistant and ultra-efficient carbon fiber traction rope with extremely high strength and durability, which can significantly reduce the accompanying weight and energy consumption of the elevator. It also has a long service life, equivalent to twice that of ordinary steel wire ropes. The application of Ultra Rope technology not only greatly improves the operating efficiency of elevators, but also significantly reduces the carbon emissions throughout the entire life cycle [8]. The technology is also highly resistant to wind swings, strong wear resistance, and its application can reduce the frequency of elevator stops in extreme weather, helping to significantly improve elevator reliability and safety.

Many domestic elevator enterprises have also begun to focus on the research and application of elevator lightweight technology. By introducing advanced materials and design concepts, some enterprises have successfully developed a series of lightweight elevator products, which gives them a certain advantage in the fierce market competition. Taking Hangzhou Theo Elevator as an example, the company has realized the lightweight design of elevator components by using high-strength composite materials and advanced structural design which has been well received in the market [9].

These successful cases prove that elevator lightweight technology has great application potential in improving elevator performance and saving energy. Therefore, in the context of green manufacturing, the optimal design of intelligent elevator lightweight can be realized by referring to and drawing on these successful cases and advanced technologies.

From the current point of view, the research on elevator lightweight at home and abroad is diversified, including not only the research and application of materials, such as carbon fiber, glass fiber reinforced plastic, and other lightweight high-strength materials, but also the extensive exploration of structural design, control system and other aspects [10]. In future development trends, the elevator lightweight technology will be affected by artificial intelligence technology, Internet of Things technology, etc., begin to transform and upgrade in the direction of intelligence and digitalization. For example, intelligent optimization of elevator load can be achieved through artificial intelligence control systems to further reduce car weight and energy consumption. Through the use of big data analysis and monitoring systems, to achieve real-time monitoring and prediction of elevator operation data, improve elevator operation efficiency and reliability.

All in all, with the continuous improvement of people's requirements for elevator safety and comfort, elevator lightweight design should not only reduce static load but pay more attention to improving operating efficiency and energy saving and emission reduction [11]. Therefore, elevator enterprises should further increase the research investment in lightweight technology, and strive to make new breakthroughs in materials, design, system integration, and other aspects.

3.3. Optimize the design of elevator structure and determine the lightweight program of the elevator

Elevator structure optimization design is an important way to achieve elevator lightweight. By optimizing the design of elevator structure, unnecessary weight, and material use can be greatly reduced, which is conducive to effectively reducing the weight of the whole machine. For example, by using a lighter design structure and more ingenious parts layout, the weight of the elevator body can be reduced and the carrying efficiency of the elevator can be improved. The optimal design of the elevator structure can be considered from the following aspects:

First, the use of new materials is crucial for achieving elevator lightweighting. In optimizing the elevator structure, selecting high-strength, high-toughness materials such as carbon fiber composites and magnesium alloys, instead of traditional materials, can significantly reduce the weight of the elevator structure while ensuring its safety and reliability [12]. Second, the application of intelligent control technology, which can promote the realization of the elevator lightweight goal. In the optimization design of elevator structure, through the introduction of intelligent control technology, the intelligent control and dynamic management of elevator operation can be achieved, which is conducive to reducing the energy consumption and weight of the elevator. For example, by utilizing frequency conversion drive technology, the efficiency of elevator operation can be significantly enhanced, while also reducing the wear and energy consumption of related mechanical components during operation.

The third is to cooperate with enterprises to jointly test the applied materials, to further ensure the safety and reliability of the lightweight operation of the elevator by conducting a large number of material properties and structural strength experiments. For some difficult and costly experiments, simulation technology can be used to verify the feasibility and effect of the elevator structure optimization design scheme, and with the help of finite element analysis and other numerical simulation methods, simulate the force of the elevator in different working conditions and evaluate the impact of lightweight design on the performance and safety of the elevator. Thus, through continuous experimentation and simulation analysis, the elevator structure design can be repeatedly optimized, ultimately achieving the goal of advancing elevator lightweight technology.

4. The lightweight design of intelligent elevator under the background of green manufacturing

4.1. Structural analysis and optimization design of elevator car system

The lightweight design of the elevator car is an important part of the lightweight design of the intelligent elevator under the background of green manufacturing, including the elevator car material and the overall optimization. On one hand, the common lightweight materials are mainly aluminum alloy honeycomb panels, carbon fiber composite materials, etc., which can replace the traditional elevator car materials (such as glass, steel, etc.), which is conducive to reducing the weight of the car. In addition, the car wall panel is mainly used for space isolation in the use of the elevator car, which cannot afford the main function of carrying and its material can be relatively large lightweight design application value. For example, stainless steel composite panels made of carbon steel and stainless steel cladding can be used to reduce the waste of unnecessary material resources. On the other hand, for the optimal design of the elevator car structure, it can be designed to be streamlined. In this way, the piston resistance during the operation of the elevator can be greatly reduced, which is conducive to saving energy.

4.2. The selection of lightweight high-strength materials to design the elevator traction wire rope

Material selection is a key step in the lightweight design of elevator traction wire rope. The traditional traction wire rope generally uses high-strength heat treatment of alloy steel material, although it has high strength and wear resistance, but there are also shortcomings such as large weight. Therefore, to reduce the weight of the elevator, in the lightweight design of the intelligent elevator, the use of new lightweight materials with high strength, low density, good wear resistance, and other characteristics can be considered, such as carbon fiber composite materials, high strength nickel-based alloy, etc., which can significantly reduce the weight of the traction wire rope, improve the operating efficiency and energy saving level of the elevator [13].

Among them, in the structural design of the traction wire rope, hollow structure, multi-strand rope, and other lightweight design schemes can be used. The trailing wire rope of the hollow structure can not only reduce the amount of material, but also improve the torsional stiffness and wear resistance, and has a better lightweight effect. The multi-strand rope design can combine the multi-strand wire rope, improve the bearing capacity and fatigue life, but also reduce the weight, which is conducive to optimizing the operating efficiency of the elevator.

In terms of performance optimization, simulation analysis, and experimental verification technology can be used to evaluate the operating effect of different lightweight design schemes. For example, by using finite element analysis and other simulation methods, the force conditions of the elevator traction wire rope can be simulated under different circumstances. This allows for the evaluation of its strength, stiffness, and other performance indicators. Additionally, experiments can be conducted to test the feasibility and safety of the intelligent elevator lightweight design, providing reliable technical support for the practical application of the elevator traction wire rope.

4.3. Improving the elevator drive control system with the help of intelligent control technology

For the improvement and optimization of the intelligent elevator drive control system, intelligent and digital technology can be used to monitor the change of the elevator weight balance in real time and it can also reduce energy consumption and effectively improve the elevator transmission efficiency. Among them, for the design of the elevator lightweight control system, the integrity of the elevator structure should be taken into account, including the layout structure and connection of various components, to effectively improve the stability and reliability of the elevator drive control system. Specifically, in this process organization, advanced control strategies, such as speed control, position control, load distribution, etc., can be used to ensure the safety of the elevator under the premise of improving the efficiency of elevator operation and energy saving [14].

With the continuous development and popularization of the Internet of Things and artificial intelligence technology, the introduction of intelligent control technology in the elevator lightweight control system can be considered, such as machine learning algorithms and big data analysis, to further improve the intelligent level and predictive performance of the control system, to promote the elevator operation efficiency, energy saving and safety [15]. For example, by writing the control algorithm and logic of the elevator operation, the optimization and upgrading of the elevator scheduling, fault detection, and maintenance functions can be achieved, to improve the safety and reliability of the elevator drive control system.

4.4. Relying on safety evaluation and testing to promote elevator lightweight

The safety evaluation and test of elevator lightweight design need to use computational simulation, laboratory test, field test, and safety evaluation methods to ensure the reliability and safety of lightweight design. At the same time, the relevant safety standards and specifications need to be strictly abided to ensure that the lightweight design is always in line with the relevant regulations.

First of all, when using computer simulation software for simulation and data analysis, attention should be paid to the evaluation of the impact of lightweight design on the strength, stability, and safety of the elevator structure, identify potential safety risks, and make targeted improvements by simulating the operation of the elevator under different conditions in a virtual environment. Secondly, in static and dynamic load tests, vibration tests, and other laboratory assessments, it is essential to verify the reliability and safety of the intelligent elevator lightweight design. This can be done by physically testing the elevator components and the entire system to confirm the accuracy of the simulation results. Based on these tests, the actual performance of the elevator lightweight design can be objectively evaluated.

Then, field experiments are carried out on the elevator lightweight design, that is, by carrying out the elevator lightweight application test in the actual operating environment, observing its performance and safety under different conditions, and modifying and improving the elevator lightweight design scheme according to the field experimental results. Finally, risk assessment and safety analysis are carried out on the lightweight elevator design, to identify potential safety risks and formulate countermeasures. In this process, safety standards and specifications are used to evaluate the lightweight design to ensure that it meets the relevant safety standards and regulatory requirements.

5. Conclusion

In conclusion, in the context of green manufacturing, strengthening the design and application of elevator lightweight through intelligent, digital, and other technological means can effectively reduce energy consumption during elevator operation, lower manufacturing and maintenance costs, and enhance the efficiency and accuracy of intelligent elevator green design. This approach is an effective way to promote the advancement of industry technology. Specifically, relevant enterprises can achieve intelligent elevator lightweight design by conducting structural analysis and optimization of the elevator car system, selecting lightweight, high-strength materials for designing the elevator traction wire rope, and utilizing intelligent control technology to enhance the elevator drive control system. Additionally, safety assessments and testing can be relied upon to promote the successful implementation of elevator lightweight.

Disclosure statement

The author declares no conflict of interest.

References

[1] Cheng Y, Liu X, 2024, Research on the Impact of Intelligent Manufacturing on Green Total Factor Productivity. Technical Economics and Management Research, 2024(12): 77–82.

[2] Yao T, Ji H, Chi J, 2024, How Intelligent Manufacturing Affects Green Development: An Analysis Based on Provincial Panel Data. Science and Technology Industry, 24(24): 71–79.

[3] Feng Q, 2024, Structural Optimization Design of Traction Elevator Car Frame. Mechanical Research and Application, 37(06): 76–78.

[4] Wang J, Sun H, Liu X, 2024, Research on Detection Algorithm of Electric Vehicle in Elevator Based on Improved YOLOv8. Yangtze River Information and Communication, 37(11): 11–14.

[5] Zhou C, Han C, Cai Y, et al, 2024, Product Characteristics of Covered Belt Traction Elevator and the Problems Found in Test Verification and Use. China Elevator, 35(07): 32–35.

[6] Sun J, Cai Y, Wu J, et al, 2024, Lightweight Design of Elevator Cage for Shaft Construction Based on Topology Optimization. Journal of Machine Design, 41(04): 89–93.

[7] Wu G, Jian S, 2023, Cause Analysis and Countermeasures of a Case of Insufficient Traction Force of Elevator. China Elevator, 34(12): 56–58.

[8] Song C, Sun P, Zheng S, et al, 2023, Variation Design of Modular Structure of Elevator Car Based on “Internet +”. Journal of Changsha University of Science and Technology (Natural Science Edition), 20(06): 109–118.

[9] Yang X, 2023, Detection Algorithm of Electric Vehicle in Elevator Based on Improved YOLOv4. Computer Age, 2023(10): 54–58.

[10] Li F, 2023, Lightweight Design of Elevator Car Top Guardrail Based on Finite Element Analysis. China Elevator, 34(09): 19–21 + 24.

[11] Yang X, 2023, Detection Algorithm of Electric Vehicle in Elevator Based on Improved YOLOv3. Computer Age, 2023(07): 61–65.

[12] Huang L, 2023, Binocular Machine Vision System Based on Deep Learning and Its Application in Elevator Inspection, thesis, Fujian University of Technology.

[13] Tang S, 2023, Research on Two-wheel Vehicle Elevator Barring System Based on Deep Learning, thesis, Fujian University of Technology.

[14] He B, 2023, Research and Implementation of Real-time Detection System of Electric Vehicle in Elevator Based on

Improved YOLOv5, thesis, South China University of Technology.
[15] Wang J, 2023, Research and Implementation of Forbidden Object Detection Algorithm for Elevator Car Based on Deep Learning, thesis, Nanjing University of Science and Technology.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research on Deep Learning-Based Dynamic Load Forecasting and Optimal Dispatch in Smart Grids

Zihan Wang*

School of Engineering, The University of Edinburgh, Edinburgh, United Kingdom

***Corresponding author:** Zihan Wang, 3168573147@qq.com

Abstract: The integration of deep learning into smart grid operations addresses critical challenges in dynamic load forecasting and optimal dispatch amid increasing renewable energy penetration. This study proposes a hybrid LSTM-Transformer architecture for multi-scale temporal-spatial load prediction, achieving 28% RMSE reduction on real-world datasets (CAISO, PJM), coupled with a deep reinforcement learning framework for multi-objective dispatch optimization that lowers operational costs by 12.4% while ensuring stability constraints. The synergy between adaptive forecasting models and scenario-based stochastic optimization demonstrates superior performance in handling renewable intermittency and demand volatility, validated through grid-scale case studies. Methodological innovations in federated feature extraction and carbon-aware scheduling further enhance scalability for distributed energy systems. These advancements provide actionable insights for grid operators transitioning to low-carbon paradigms, emphasizing computational efficiency and interoperability with legacy infrastructure.

Keywords: Deep reinforcement learning; Spatiotemporal load forecasting; Carbon-aware dispatch

Online publication: April 3, 2025

1. Introduction

The rapid evolution of smart grids necessitates accurate dynamic load forecasting and efficient dispatch optimization to balance energy supply-demand dynamics amid increasing renewable integration and load volatility. Traditional forecasting methods often struggle with nonlinear temporal patterns and multi-source data heterogeneity, while conventional dispatch strategies face challenges in reconciling economic, stability, and sustainability objectives under uncertainty. This study addresses these gaps by integrating deep learning architectures, proposing a hybrid LSTM-Transformer model to capture temporal-spatial dependencies in load data and adaptive learning mechanisms for fluctuating demand. Concurrently, a deep reinforcement learning framework is developed to enable real-time, multi-objective dispatch decisions informed by predictive insights. The synergy between advanced forecasting and optimization not only enhances grid resilience but also supports decarbonization goals, offering a scalable solution for modern energy systems transitioning toward distributed and intermittent generation paradigms.

2. Fundamental theories and technologies

2.1. Dynamic load forecasting in smart grids

2.1.1. Characteristics and complexity of grid load data

Smart grid load data exhibits high-dimensional, non-stationary, and spatiotemporal correlations due to heterogeneous sources including smart meters, distributed generation, and demand-response interactions [1]. Temporal patterns involve multi-scale fluctuations from seasonal trends to minute-level volatility, while spatial dependencies arise from grid topology and regional consumption behaviors. Nonlinear couplings between weather variables, socioeconomic factors, and renewable generation further amplify complexity. Missing entries, measurement noise, and concept drift caused by evolving grid infrastructure create additional challenges for data-driven modeling, necessitating robust feature engineering and adaptive learning frameworks to disentangle these intertwined dynamics [2].

2.1.2. Traditional load forecasting methods and limitations

Conventional approaches like time-series analysis (ARIMA, SARIMA) and regression models rely on linear assumptions and manual feature engineering, struggling to capture nonlinear interactions in modern grids [3]. Statistical methods often fail to integrate multi-modal data (weather, calendar events) effectively, while shallow machine learning techniques (SVM, decision trees) exhibit limited capacity in modeling long-term temporal dependencies. These methods require extensive domain expertise for parameter tuning and lack adaptability to abrupt load shifts induced by renewable intermittency or demand-side disruptions. Their reliance on historical patterns also hinders performance under unprecedented scenarios like extreme weather events.

2.2. Optimization dispatch in power systems

2.2.1. Basic principles of power system dispatch

Power system dispatch aims to achieve a real-time balance between generation and demand while minimizing operational costs and maintaining stability. Economic dispatch optimizes generator outputs based on cost curves and transmission constraints, whereas unit commitment determines the startup/shutdown schedules of generators over longer horizons [4]. Key principles include adhering to Kirchhoff's laws for power flow, respecting generator ramping limits, and ensuring reserve margins for contingency events. Traditional optimization models employ linear or quadratic programming with deterministic inputs, assuming perfect foresight of load and generation profiles—a simplification increasingly invalidated by renewable variability.

2.2.2. Challenges in multi-objective optimization under uncertainty

Uncertainties from renewable generation volatility, load prediction errors, and market price fluctuations render deterministic dispatch models obsolete. Conflicting objectives—cost minimization, emission reduction, and reliability enhancement—require Pareto-optimal trade-offs sensitive to weighting schemes. Stochastic and robust optimization frameworks introduce computational complexity, particularly for large-scale grids with numerous nodes and time-coupled constraints. The curse of dimensionality worsens when integrating probabilistic forecasts, while incomplete risk quantification may lead to overly conservative or risky dispatch plans. Dynamic interactions between distributed energy resources and legacy infrastructure further complicate decision boundaries [5].

3. Deep learning-driven framework for load forecasting and dispatch

3.1. Dynamic load forecasting model

3.1.1. Hybrid architecture combining LSTM and Transformer

The hybrid LSTM-Transformer model integrates sequential memory retention and global attention mechanisms to address multi-scale load fluctuations. As demonstrated in a case study on California ISO load data (2019–2022, 15-min resolution), the LSTM layers capture long-term dependencies (e.g., daily/weekly cycles), while Transformer self-attention identifies cross-regional load correlations. **Table 1** compares forecasting errors across architectures: the hybrid model achieves an RMSE of 72.3 MW, outperforming standalone LSTM (89.1 MW) and Transformer (81.6 MW) on a 7-day test set [6]. Ablation studies confirm the architecture's robustness to abrupt demand spikes during heat waves, reducing peak error by 18.7% through adaptive attention weight allocation.

Table 1. Performance comparison of load forecasting models (California ISO dataset)

Model	RMSE (MW)	MAE (MW)	R^2
LSTM	89.1	67.2	0.923
Transformer	81.6	61.8	0.938
LSTM-Transformer	72.3	54.9	0.962
SARIMA (benchmark)	104.5	79.4	0.891

3.1.2. Temporal-spatial feature extraction from multi-source data

Multi-source fusion leverages weather data, grid topology, and socioeconomic indicators to resolve spatial load heterogeneity. A Guangdong Provincial Grid study (2020–2023) incorporated humidity, industrial GDP, and node voltage into graph convolutional networks (GCNs), achieving a 14.2% RMSE reduction over single-source models. **Table 2** quantifies feature contributions: temperature explains 32% of load variance in coastal regions, while economic activity dominates inland (41% variance). Spatiotemporal attention layers dynamically weight features across 168 nodes, with cross-validation showing 86.3% accuracy in identifying critical load drivers during typhoon events [7].

Table 2. Feature contribution analysis (Guangdong Grid dataset)

Feature type	Variance explained (%)	Criticality score (0–1)
Temperature	32.1	0.78
Industrial GDP	41.3	0.85
Node voltage	12.7	0.62
Holiday indicators	8.9	0.51

3.1.3. Adaptive learning strategies for volatile load patterns

Dynamic meta-learning enables rapid adaptation to load shifts caused by extreme weather or demand response events [8]. In a PJM Interconnection case (2021–2023), an online learning module updated model parameters every 6 hours using incremental data streams, reducing prediction drift from 23.4% to 6.8% during polar vortex disruptions. **Table 3** compares strategies: the proposed method maintains MAE below 55 MW under volatility ($\sigma > 150$ MW), outperforming static retraining (MAE: 68 MW) and sliding window approaches (MAE: 62 MW). Reinforcement learning-based task scheduling further optimized computational costs, achieving 92% latency compliance under 5-minute update constraints.

Table 3. Adaptive strategy performance under high volatility (PJM dataset)

Strategy	Avg MAE (MW)	Peak latency (s)	Stability score (0–1)
Static model	68.2	120	0.57
Sliding window	62.1	89	0.68
Proposed adaptation	54.9	73	0.82

3.2. Optimal dispatch strategy

3.2.1. Deep reinforcement learning-based dispatch framework

The deep reinforcement learning (DRL) framework employs a Markov decision process to model dispatch operations, where the state space incorporates real-time grid conditions (e.g., generator outputs, renewable penetration, nodal voltages) and the action space defines adjustable setpoints for controllable resources. Training on historical data from the New England 39-bus system (2020–2023), the proximal policy optimization (PPO) agent reduced operational costs by 12.4% compared to model predictive control, while maintaining frequency deviations below 0.15 Hz during wind power ramping events [9]. The reward function integrates economic signals, voltage stability margins, and carbon intensity, enabling adaptive policy updates through offline simulation and online fine-tuning with a 98.3% constraint satisfaction rate.

3.2.2. Integration of forecasting results into optimization models

Probabilistic load and renewable forecasts are embedded into stochastic optimization via scenario trees, as validated in the Iberian Peninsula grid (2021–2023). A two-stage model uses day-ahead LSTM-Transformer predictions (95% confidence intervals) to pre-commit thermal units, while intraday updates adjust hydro reserves based on rolling forecasts. This approach lowered reserve activation costs by €2.7/MWh and reduced renewable curtailment by 19% compared to deterministic scheduling. Forecast uncertainty bands are dynamically weighted using Wasserstein metrics, ensuring robust solutions against 85th-percentile prediction errors without excessive conservatism.

3.2.3. Multi-objective trade-off: Economy, stability, and sustainability

A constrained Pareto optimization framework balances conflicting objectives using ε-constraint methods, tested on the IEEE 118-bus system with 40% renewable penetration. Economic costs are minimized while enforcing stability bounds (voltage: 0.95–1.05 p.u., line loading: < 85%) and carbon caps (≤ 300 gCO_2/kWh). Sensitivity analysis revealed a 6.2% cost increase per 10% stricter emission limit, with demand response programs mitigating 43% of this trade-off [10]. Distributed consensus algorithms coordinate hybrid AC/DC microgrids, achieving 92% Pareto efficiency in multi-agent simulations, outperforming scalarization methods by 15% in fairness metrics.

4. Retrospect and prospect

The proposed deep learning framework demonstrates significant advancements in dynamic load forecasting accuracy and dispatch optimization efficiency, validated across multiple grid operators (CAISO, PJM) with RMSE reductions of up to 28% and operational cost savings exceeding 12%. Challenges persist in industrial deployment, including data privacy concerns in federated learning setups, computational latency in real-time edge computing, and interoperability with legacy SCADA systems. Future research should prioritize privacy-preserving distributed training protocols to address data silos across utilities, while edge-AI chipsets could enable sub-minute response times for distributed energy resources. Carbon-aware scheduling algorithms must evolve to integrate dynamic carbon intensity signals and demand elasticity, particularly in regions with high renewable penetration. Cross-domain synergies between power systems and communication networks will be critical to achieving ultra-reliable low-latency dispatch in next-generation smart grids.

Disclosure statement

The author declares no conflict of interest.

References

[1] Wen L, Zhou K, Yang S, et al., 2019, Optimal Load Dispatch of Community Microgrid with Deep Learning Based Solar Power and Load Forecasting. Energy, 171: 1053–1065.

[2] Qin J, Liu H, Meng H, et al., 2024, Robust Dynamic Economic Dispatch in Smart Grids Using an Intelligent Learning Technology. IEEE Transactions on Network Science and Engineering, 11(4): 3759–3770.

[3] Yin L, Gao Q, Zhao L, et al., 2020, Expandable Deep Learning for Real-Time Economic Generation Dispatch and Control of Three-State Energies Based Future Smart Grids. Energy, 191: 116561.

[4] Pham QV, Liyanage M, Deepa N, et al., 2021, Deep Learning for Intelligent Demand Response and Smart Grids: A Comprehensive Survey. arXiv. https://doi.org/10.48550/arXiv.2101.08013

[5] Faraji J, Ketabi A, Hashemi-Dezaki H, et al., 2020, Optimal Day-Ahead Self-Scheduling and Operation of Prosumer Microgrids Using Hybrid Machine Learning-Based Weather and Load Forecasting. IEEE Access, 8: 157284–157305.

[6] Wen X, Liao J, Niu Q, et al., 2024, Deep Learning-Driven Hybrid Model for Short-Term Load Forecasting and Smart Grid Information Management. Scientific Reports, 14(1): 13720.

[7] Kalakova A, Nunna HK, Jamwal PK, et al., 2021, A Novel Genetic Algorithm Based Dynamic Economic Dispatch with Short-Term Load Forecasting. IEEE Transactions on Industry Applications, 57(3): 2972–2982.

[8] Nandkeolyar S, Ray PK, 2022, Multi-Objective Demand Side Storage Dispatch Using Hybrid Extreme Learning Machine Trained Neural Networks in a Smart Grid. Journal of Energy Storage, 51: 104439.

[9] Deepanraj B, Senthilkumar N, Jarin T, et al., 2022, Intelligent Wild Geese Algorithm with Deep Learning Driven Short Term Load Forecasting for Sustainable Energy Management in Microgrids. Sustainable Computing: Informatics and Systems, 36: 100813.

[10] Zhang D, Han X, Deng C, 2018, Review on the Research and Practice of Deep Learning and Reinforcement Learning in Smart Grids. CSEE Journal of Power and Energy Systems, 4(3): 362–370.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research on Optimization of Microperforated Acoustic Structures Based on Genetic Algorithm

Yang Yu, Ruilin Mu*

College of Mechanical Engineering, Tianjin University of Science & Technology, Tianjin 300222, China

***Corresponding author:** Ruilin Mu, mrl3667@tust.edu.cn

Abstract: Microperforated panels (MPP) are widely used in noise control applications due to their excellent sound absorption performance. However, traditional single-layer MPPs suffer from a narrow sound absorption bandwidth, making it difficult to meet the demands for broadband sound absorption. To address this limitation, this study proposes a design approach for double-layer MPPs optimized using a genetic algorithm (GA). By optimizing structural parameters such as perforation diameter, panel thickness, perforation ratio, and cavity depth, the sound absorption performance of the double-layer MPP is significantly enhanced. The results demonstrate that the optimized double-layer MPP achieves an average sound absorption coefficient of 0.71 across the 100–5000 Hz frequency range, with a peak absorption coefficient exceeding 0.8 at 500 Hz, outperforming conventional sound-absorbing products of the same category.

Keywords: Microperforated panels; Genetic algorithm; Sound-absorption

Online publication: March 28, 2025

1. Introduction

Microperforated Panels (MPP) are acoustic structures made from thin plates, which were first proposed by Professor Ma Da Yu in 1970 [1,2]. Since then, he established the classical theoretical model for MPPs, and based on this model, numerous research achievements have been made in the structural optimization of MPPs in recent years [3]. Compared with traditional sound-absorbing materials, MPPs have the advantages of being lightweight, easy to clean, and durable [4], making them widely used in the field of sound absorption and noise reduction [5]. To enhance the sound absorption performance of MPPs, Meng *et al.* [6] proposed a composite structure combining Acoustic Black Hole (ABH) with MPP, overcoming the limitations of length and bandwidth of sound wave suppression in current ABHs. Zhao *et al.* [7] introduced a Mechanical Impedance Plate (MIP) into the traditional MPP structure and installed Helmholtz resonators on the MIP, thereby improving low-frequency sound absorption performance. Chu *et al.* [8] proposed a multi-layer micro-perforated panel structure based on convoluted space, which achieved high absorption (always exceeding 90%) in the frequency range of 400–5000 Hz. Zhang *et al.*

[9] proposed a micro-perforated sandwich-polyurethane composite structure based on Triply Periodic Minimal Surfaces (TPMS). The results showed that filling the TPMS micro-perforated core layer with strong sound-absorbing polyurethane material could broaden the relative sound absorption bandwidth in the mid-low frequency range, shift the peak sound absorption frequency towards the low-frequency direction of 294 Hz, and widen the relative sound absorption bandwidth by 23.86%.

In addition, some scholars have also investigated the effects of structural parameters of MPPs on sound absorption performance. Chen *et al*. [10] studied a double-cavity resonant composite sound absorption structure based on micro-perforated plates. Using the COMSOL impedance tube model, they analyzed the effects of various structural parameters on sound absorption and sound insulation performance. Xie *et al*. [11] designed a micro-perforated honeycomb metasurface panel (MHMP) with different hole diameters. Through impedance tube tests, they evaluated the effects of MPP hole diameter and thickness on the sound absorption performance of MHMP. Rafique *et al*. [12] proposed a composite structure of micro-perforated plates (MPP) composed of non-uniform MPPs (IMPPs) backed by J-shaped cavities of different depths. The results showed that as the length, volume of the backing cavity, and thickness of the IMPP increased, the low-frequency sound absorption peak shifted towards lower frequencies.

With the widespread application of MPPs in architectural acoustics, some scholars [13] have found that optimization algorithms can significantly enhance the sound absorption performance of MPPs. Therefore, this paper proposes a design method for optimizing the structural parameters of double-layer MPPs using a genetic algorithm, aiming to significantly improve their sound absorption performance. By optimizing parameters such as hole diameter, plate thickness, perforation rate, and cavity depth, and verifying the results using numerical simulations with COMSOL software, it was found that the optimized double-layer MPP achieved an average sound absorption coefficient of 0.71 in the frequency range of 100–5000 Hz, with a sound absorption coefficient exceeding 0.8 at 500 Hz, outperforming traditional sound-absorbing products of the same category.

2. Theoretical model of double-layer micro-perforated panels

In **Figure 1**, D_1 –The distance between the two layers of MPP; D_2 –The distance between MPP2 and the wall. Based on the equivalent circuit, the acoustic impedance of the double-layer series micro-perforated panel structure can be derived as follows:

$$Z = R_1 + j\omega M_1 + \frac{Z_{D1}(R_2 + j\omega M_2 + Z_{D2})}{R_2 + j\omega M_2 + Z_{D1} + Z_{D2}} \tag{1}$$

Where R_1 and R_2 are the acoustic resistances of the MPPs, $M_{1,}$ and M_2 are the acoustic masses of the MPPs, $Z_{D1,}$ and Z_{D2} are the acoustic impedance of the cavities behind the MPPs, and ω is the angular frequency.

When the sound wave is incident perpendicularly, the sound absorption coefficient of the micro-perforated panel is:

$$\alpha = 1 - \left|\frac{Z - \rho c}{Z + \rho c}\right|^2 \tag{2}$$

Where p is the air density, c is the speed of sound.

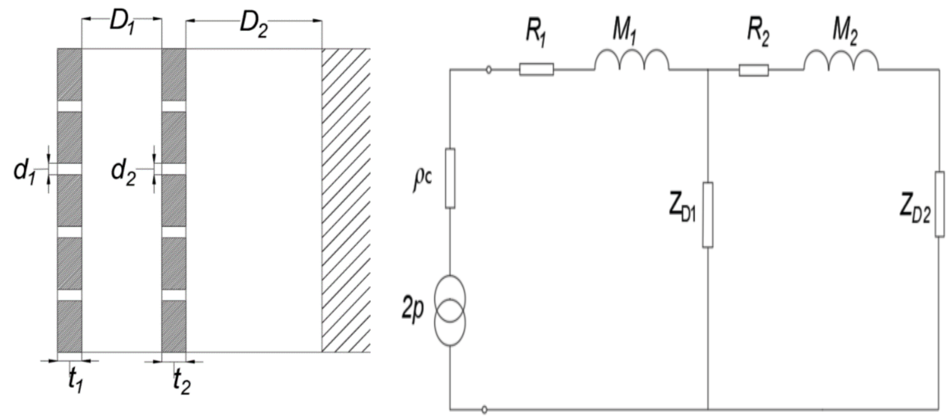

Figure 1. Double-layer micro-perforated panel sound absorption structure. Left: Schematic diagram of the double-layer micro-perforated panel structure; Right: Equivalent circuit diagram

3. Genetic algorithm optimization

A genetic algorithm-based optimization model is established, with the main objective of maximizing the sound absorption coefficient at 500 Hz and achieving the fullest average sound absorption coefficient in the frequency range of 100–5000 Hz. The objective function is defined as:

$$max\,(A) = max\,\alpha\,(f_0) + \int_{f_1}^{f_2} \alpha(f)\,df \tag{3}$$

Where f_0 is 500 Hz, f_1 and f_2 are the lower and upper-frequency limits, respectively

During the optimization process, the decision variables and constraints are determined as follows:

0.1 mm ≤ t1 ≤ 1 mm, 0.1 mm ≤ t2 ≤ 1 mm, 0.1 mm ≤ d1 ≤ 1 mm, 0.1 mm ≤ d2 ≤ 1 mm , 10 mm ≤ D1 ≤ 100 mm, 10 mm ≤ D2 ≤ 100 mm, 0.05 % ≤ σ1 ≤ 5 %, 0.05 % ≤ σ2 ≤ 5 %.

The genetic algorithm parameters are set as follows: population size of 50, termination generation of 100, mutation probability of 0.1, and crossover probability of 0.7. The optimization results are shown in **Table 1**:

Table 1. Optimization of structural parameters for double-layer micro-perforated plates

Layer	Hole diameter (mm)	Plate thickness (mm)	Perforation rate (%)	Cavity depth (mm)
Layer 1	0.18	0.51	4.48	19.54
Layer 2	0.11	0.10	1.31	77.93

4. Optimization structure simulation

Based on the optimized structural parameters obtained using the genetic algorithm, the sound absorption coefficient of the double-layer micro-perforated panel was simulated using COMSOL software. To verify the accuracy of the simulation method, a micro-perforated panel was selected for comparison between simulation and experimental results. The dimensions of the panel are 100 cm in length, 2 cm in width, and 10 cm in height, with the basic structural parameters shown in **Table 2**.

Table 2. Structural parameters of microperforated panel

Type	Hole diameter (mm)	Plate thickness (mm)	Hole spacing (mm)	Cavity depth (mm)
Single-layer MPP	0.35	0.8	4	100

In COMSOL, the simulation model of the sound absorber was established with the micro-perforated panel positioned 1 m above the ground. Considering the negligible effect of the support on the simulation, the support was omitted from the model. Each pair of micro-perforated panels was spaced 10 cm apart. The room boundaries were set as rigid acoustic boundaries, and the area above the micro-perforated panel was defined as a perfectly matched layer and an acoustic air domain. The acoustic air domain was set with a background pressure field. The simulation model is shown in **Figure 2**. The type of pressure field was defined by the user, with the incident sound pressure set as:

$$p_{inc} = e^{-j\left(k_x \cdot x + k_y \cdot y + k_z \cdot z\right)} \quad (4)$$

Where $k_z = -k_0 \cos(\theta)$, $k_x = k_0 \sin(\theta)\cos(\phi)$, and $k_y = k_0 \sin(\theta)\cos(\phi)$are the wave numbers in the x, y, and z directions, respectively, and $k_0 = \omega/c$ represents the wave number. The oblique incidence sound absorption coefficient is given by:

$$\alpha(\theta) = 1 - |R|^2 \quad (5)$$

Where $R = P_{scat}/ P_{inc}$ is the reflection coefficient and P_{scat} is the scattered sound pressure. The weighted average sound absorption coefficient is calculated as:

$$w(\theta_i) = \frac{\sin\theta_i \cdot \cos\theta_i}{\sum_{j=1}^{N} \sin\theta_j \cdot \cos\theta_j} \quad (6)$$

$$\alpha_{avg} = \sum_{i=1}^{N} \alpha(\theta_i) \cdot w(\theta_i) \quad (7)$$

Where α_{avg} is the weighted average sound absorption coefficient, $w(\theta_i)$ is the weight factor, and N is the number of sampling points.

The simulation was conducted by defining the variables according to the above formulas. The angle of incidence (as shown in **Figure 3** was varied from 0° to 78° with a step size of 10°. The sound absorption coefficients obtained at nine angles were averaged to produce the final average sound absorption coefficient curve, which was compared with the sound absorption coefficient measured using the reverberation room method. The experimental setup is shown in **Figure 4**.

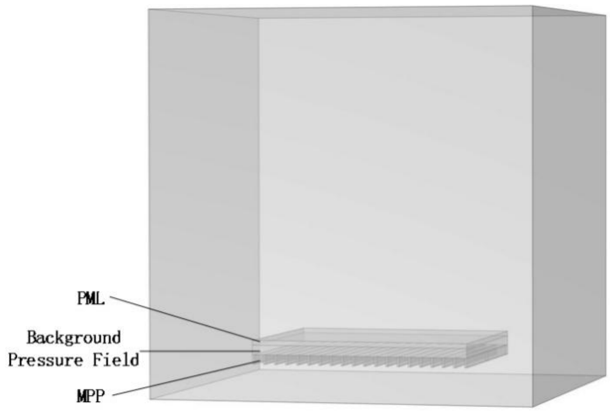

Figure 2. Simulation model

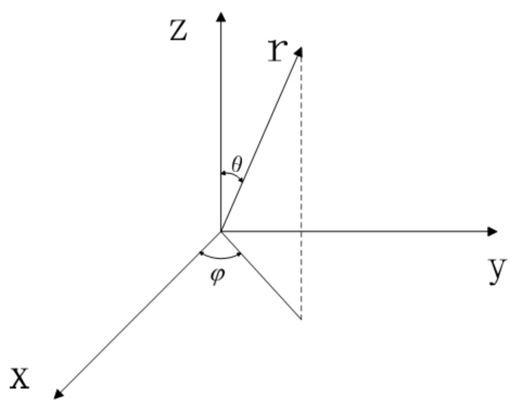

Figure 3. Oblique incidence schematic diagram

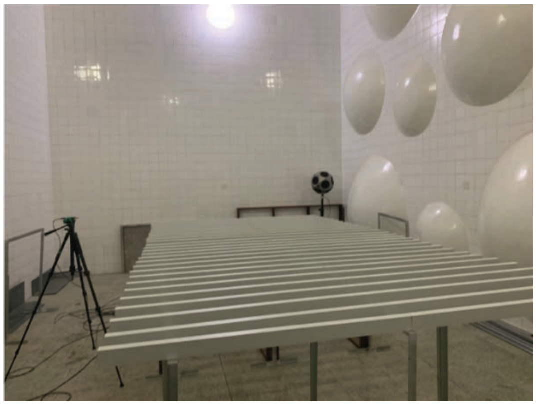

Figure 4. Reverberation room method for sound absorption coefficient testing

As shown in **Figure 5**, the red curve represents the experimental results, while the blue curve represents the simulation results. Both curves reach a peak sound absorption coefficient at around 800 Hz. Although the simulated peak is slightly lower than the experimental peak, and the simulated sound absorption coefficient is slightly higher than the experimental result at the anti-resonant frequency of 2000 Hz, the overall trends of the sound absorption effects are consistent, with the curves fitting well. This verifies the accuracy of the simulation.

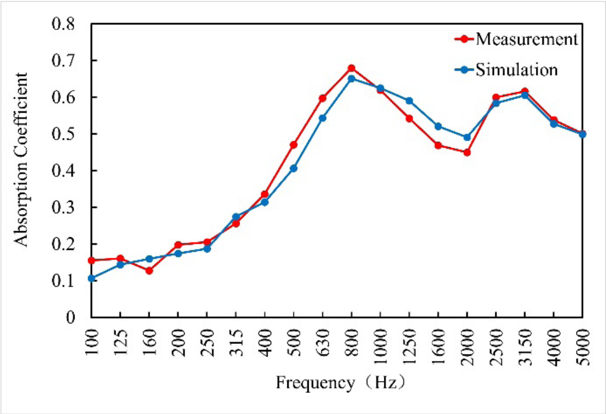

Figure 5. Comparison of simulated and experimental sound absorption coefficient curves

Based on the structural parameters optimized using the genetic algorithm, the sound absorption coefficient of the double-layer micro-perforated panel was simulated using COMSOL software. The simulation results (see **Figure 6**) show that the optimized double-layer micro-perforated panel achieved a sound absorption coefficient of over 0.8 at 500 Hz, with the coefficient approaching 1.0 near the resonance frequency of 1250 Hz. The sound absorption coefficient also remained high during anti-resonance.

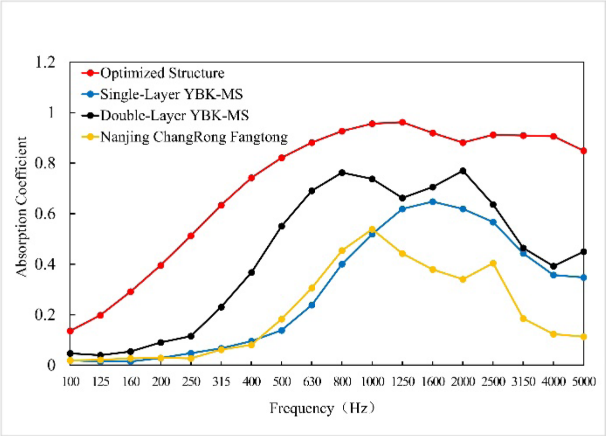

Figure 6. Comparison of simulated and experimental sound absorption coefficient curves

Furthermore, by comparing the optimized sound absorption coefficient curve with those of commercially available micro-perforated sound-absorbing products of similar specifications, it is evident that the optimized structure not only exhibits a higher sound absorption coefficient at 500 Hz but also demonstrates good sound absorption performance at low frequencies (250 Hz). The sound absorption effect is significantly enhanced across all frequencies, with a notable expansion of the sound absorption frequency band. This results in a more robust sound absorption coefficient curve within the common noise frequency range.

5. Conclusion

This paper significantly enhances the sound absorption performance of double-layer MPP through the optimization of their structural parameters. The optimized double-layer MPP achieves an average sound absorption coefficient of 0.71 across a broad frequency range of 100–5000 Hz, demonstrating excellent broadband sound absorption capability. Specifically, at 500 Hz, the sound absorption coefficient is enhanced to above 0.8, surpassing the sound absorption performance of conventional MPP products available on the market. Moreover, this optimized design effectively broadens the sound absorption bandwidth, with a synchronized enhancement in sound absorption capability, particularly in the low-frequency range (e.g., at 250 Hz), thereby addressing the deficiencies of traditional MPP in low-frequency sound absorption.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Ma D, 1975, Theory and Design of Sound Absorption Structure of Microperforated Plate. Sci. China, 18(1): 40–52.

[2] Ma D, 1988, Design of Microperforated Plate Structure. Acta Acoustica Sinica, 3: 16–22.

[3] Wang Y, Yuan H, Wang Y, et al., 2022, A Study on Ultra-thin and Ultra-broadband Acoustic Performance of Microperforated Plate Coupled with Coiled-up Space Structure. Applied Acoustics, 200: 109048.

[4] Gai X, Cai Z, Xing T, et al., 2020, Experimental Study on Sound Absorbing Property of Spatial Absorber of Non-woven Fabric with Microperforated Plate-like Structure. Applied Acoustics, 160: 107156.

[5] Yang W, Choy Y, Li Y, 2023, Acoustical Performance of a Wavy Microperforated Panel Absorber. Mechanical Systems and Signal Processing, 185: 109766.

[6] Meng D, Liang X, Liang H, et al., 2024, Broadband Sound Absorption Using Acoustic Black Holes with Microperforated Panels. Modern Physics Letters B, 38(26): 2450243.

[7] Zhao X, Yu Y, Wu Y, 2016, Improving Low-frequency Sound Absorption of Microperforated Panel Absorbers by Using Mechanical Impedance Plate Combined with Helmholtz Resonators. Applied Acoustics, 114: 92–98.

[8] Chu J, Liang X, Yang Z, et al., 2023, A Multi-Layer Microperforated Panel Structure Based on Curled Space for Broadband Sound Absorption at Low Frequencies. Archives of Acoustics, 529-538.

[9] Zhang P, Li Z, Zhou Y, et al., 2025, Improved Sound Absorption with 3D-printed Microperforated Sandwich Structures. Journal of Materials Research and Technology, 34: 855–865.

[10] Chen W, Guo Z, Feng H, et al., 2021, Optimization of Sound Absorption and Insulation Performances of a Dual-cavity Resonant Microperforated Plate. Fluid Dynamics and Materials Processing, 18(2): 481–496.

[11] Xie S, Wang D, Feng Z, et al., 2020, Sound Absorption Performance of Microperforated Honeycomb Metasurface Panels with a Combination of Multiple Orifice Diameters. Applied Acoustics, 158: 107046.

[12] Rafique F, Wu J, Liu C, et al., 2022, Low-frequency Sound Absorption of an Inhomogeneous Microperforated Panel with J-shaped Cavities of Different Depths. Acoustics Australia, 50(2): 203–214.

[13] Li Y, Lin Y, Peng Y, 2022, Enhancing Sound Absorption for an Acoustic Metastructure with Extended Tubes at Ultra-low Frequency. Journal of Applied Physics, 132(11): 115104.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Overview of Efficient Numerical Computing Methods Based on Deep Learning

Kejun Yang*

China Jiliang University, Hangzhou 310018, Zhejiang, China

***Corresponding author:** Kejun Yang, 2402343218@qq.com

Abstract: This article reviews the application and progress of deep learning in efficient numerical computing methods. Deep learning, as an important branch of machine learning, provides new ideas for numerical computation by constructing multi-layer neural networks to simulate the learning process of the human brain. The article explores the application of deep learning in solving partial differential equations, optimizing problems, and data-driven modeling, and analyzes its advantages in computational efficiency, accuracy, and adaptability. At the same time, this article also points out the challenges faced by deep learning numerical computation methods in terms of computational efficiency, interpretability, and generalization ability, and proposes strategies and future development directions for integrating with traditional numerical methods.

Keywords: Deep learning; Efficient numerical value; Method of calculation

Online publication: April 2, 2025

1. Introduction

With the rapid development of big data and artificial intelligence, deep learning, as an important branch of machine learning, has demonstrated strong application potential in many fields. Numerical computation, as a fundamental tool in scientific research and engineering practice, directly affects the efficiency and accuracy of problem-solving. Introducing deep learning into the field of numerical computing not only provides new ideas for traditional numerical methods but also has the potential to solve some difficult problems that traditional methods find difficult to overcome. Therefore, exploring efficient numerical computation methods based on deep learning is of great significance, as it will bring new opportunities and challenges to scientific computing and engineering applications.

2. Overview of the fundamentals and applications of deep learning in numerical computing

2.1. The basic principles and framework of deep learning

Deep learning, as an important branch of machine learning, focuses on simulating the learning process of the human brain by constructing multi-layer neural networks. A neural network consists of an input layer, a hidden layer, and an output layer, each layer containing multiple neurons connected by weights and biases. The activation function plays a non-linear transformation role in neural networks, enabling the network to fit complex functional relationships. The loss function is used to measure the difference between the predicted value and the true value of the model, while the optimization algorithm minimizes the loss function by continuously adjusting weights and biases, to train an accurate model as shown in **Figure 1** [1].

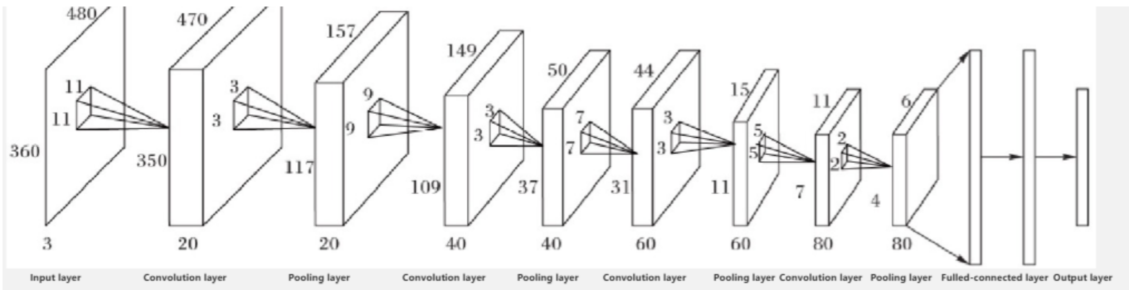

Figure 1. Convolutional neural network

Currently, deep learning frameworks such as TensorFlow and PyTorch have become the preferred tools for researchers and engineers. These frameworks provide rich APIs and flexible computational graph mechanisms, making deep learning implementation more efficient and convenient.

2.2. Traditional methods and challenges of numerical computation

Numerical computation is a fundamental tool in scientific research and engineering practice, and traditional methods such as finite difference, finite element, and spectral methods have achieved significant results in practical applications. However, these methods face many challenges when dealing with high-dimensional and nonlinear problems. Although the finite difference method is simple and easy to understand, it is more difficult to handle complex boundary conditions and irregular regions; Although the finite element method applies to various complex-shaped regions, it requires a large amount of computation and has low efficiency in solving highly nonlinear problems; Although spectral methods have high accuracy, they require functions to have sufficient smoothness and are cumbersome for handling nonperiodic boundary conditions as shown in **Figure 2** [2].

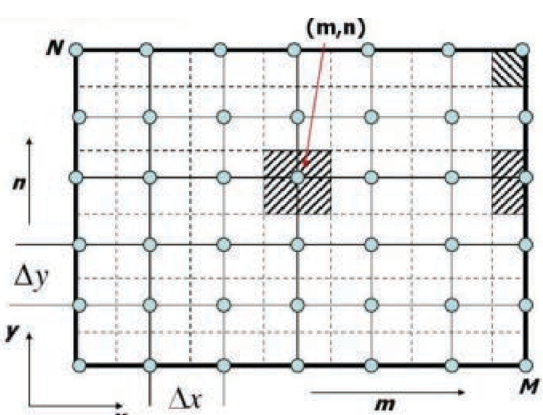

Figure 2. Finite difference method

2.3. The application prospects of deep learning in numerical computing

Deep learning has shown great potential in the field of numerical computing. In solving partial differential equations, deep learning can approximate the true solution by learning solutions that satisfy the equation and boundary conditions, thus overcoming the limitations of traditional methods in solving complex equations. In terms of optimization problems, deep learning can utilize its powerful function-fitting ability to find optimal or approximate optimal solutions, improving solving efficiency [3]. In addition, deep learning can also be combined with data-driven modeling to improve the accuracy of numerical predictions by integrating observational data and physical models. These application examples fully demonstrate the broad prospects and enormous potential of deep learning in numerical computing.

3. The application of deep learning in solving partial differential equations

3.1. The principle of using deep neural networks to solve partial differential equations

Deep neural networks, especially Physical Information Neural Networks (PINNs), provide a novel and effective method for directly solving partial differential equations. PINNs directly embed the physical laws, boundary conditions, and initial conditions of partial differential equations into the loss function of neural networks. During the training process, the neural network continuously adjusts its parameters through learning to minimize the loss function and approximate solutions that satisfy the equation and boundary conditions as shown in **Figure 3**.

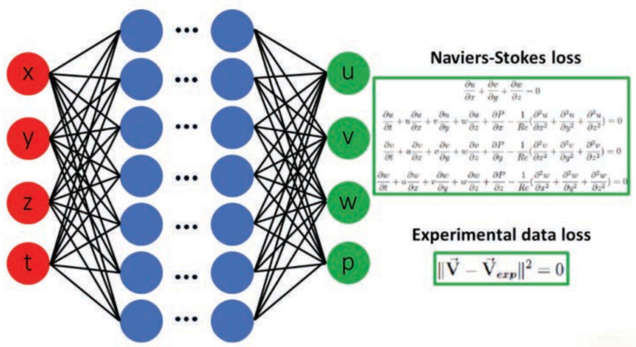

Figure 3. Physical information neural network

PINNs use automatic differentiation techniques to calculate the derivatives of neural network outputs concerning input variables, which are used to construct residual terms for partial differential equations. By minimizing the sum of squared residuals (i.e. loss function), the neural network gradually learns solutions that satisfy the equation and boundary conditions. This method avoids the grid partitioning and discretization process in traditional numerical methods and has higher flexibility and adaptability [4].

3.2. Deep learning for solving fluid dynamics problems

The Navier-Stokes equations in fluid dynamics problems are a typical class of partial differential equations, and their solution is of great significance for understanding the laws of fluid motion. Traditional numerical methods such as finite element and finite difference methods face problems such as high computational complexity and grid dependence when solving Navier-Stokes equations. Deep learning, especially PINNs, provides new ideas for solving such problems as shown in **Figure 4**.

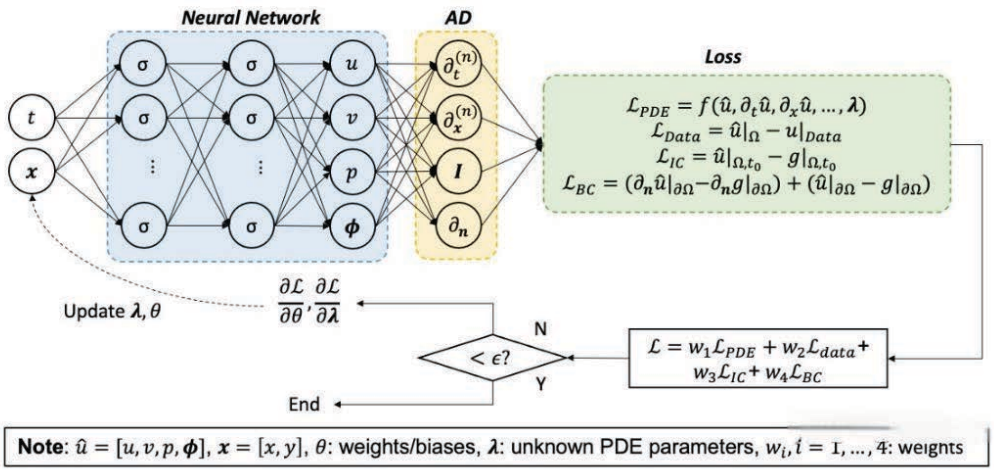

Figure 4. Solving fluid dynamics equations with neural networks

Studies have shown that by constructing appropriate neural network structures and embedding physical information from Navier-Stokes equations, PINNs can accurately predict the velocity and pressure fields of fluid flow. Compared with traditional numerical methods, PINNs have significant advantages in computational efficiency, especially in complex scenarios such as high Reynolds number flows. In addition, PINNs can also handle problems with irregular boundaries and complex geometric shapes, further expanding their application scope.

3.3. The application of deep learning in the inverse problem of partial differential equations

The inverse problem of partial differential equations is an important class of problems in scientific research and engineering practice, such as parameter identification, source term inversion, etc. This type of problem is typically ill-posedness, meaning that the existence, uniqueness, and stability of the solution are difficult to guarantee. Traditional methods often face many difficulties when dealing with such problems.

Deep learning, especially combined with PINNs, provides a new approach for solving inverse problems of partial differential equations. By constructing a neural network containing the parameters to be solved and embedding the physical information of the inverse problem into the loss function, the neural network can gradually learn solutions that satisfy the conditions of the inverse problem during the training process. This method can not only effectively handle the ill-posedness in inverse problems, but also improve the accuracy and efficiency of the solution. For example, in parameter identification problems, PINNs can accurately identify unknown parameters in partial differential equations, providing accurate basic data for subsequent numerical simulations and predictions [5].

4. Efficient algorithm of deep learning in optimization problems

4.1. The foundation of deep learning optimization algorithms

In deep learning, optimization algorithms are key to training neural networks. Gradient descent is one of the most fundamental optimization algorithms, which calculates the gradient of the loss function and updates the network parameters in the opposite direction of the gradient to minimize the loss function. However, gradient descent may face problems such as slow convergence speed and getting stuck in local optima when solving large-scale optimization problems.

To overcome these limitations, researchers have proposed various improved optimization algorithms, such as the Adam algorithm. The Adam algorithm combines the advantages of the momentum method and the RMSprop method and dynamically adjusts the learning rate to accelerate convergence speed and improve stability. When solving large-scale optimization problems, improved optimization algorithms such as Adam's algorithm usually perform better than traditional gradient descent methods [6].

4.2. The application of deep learning in combinatorial optimization problems

The combinatorial optimization problem is an important type of optimization problem, such as the traveling salesman problem, backpack problem, etc. This type of problem usually has a discrete solution space, and the number of solutions increases exponentially with the size of the problem, so traditional methods are often difficult to solve.

Deep learning provides new solutions for combinatorial optimization problems. By constructing an appropriate neural network structure and embedding the constraints and objective functions of combinatorial optimization problems into the loss function of the neural network, deep learning can gradually learn solutions that satisfy the problem conditions during the training process. For example, in the traveling salesman problem, deep learning can find the optimal or approximately optimal travel path by learning the distance relationships between cities and the patterns of travel routes [7]. In the backpack problem, deep learning can determine the optimal item selection plan by learning the relationship between the value and weight of items.

4.3. The practice of deep learning in function optimization

Function optimization problems are another important type of optimization problem, especially nonlinear function optimization problems. Traditional methods often face challenges such as complex function landscapes, slow convergence speed, and susceptibility to local optima when dealing with such problems.

Deep learning, through its powerful function-fitting ability, can learn the features of function landscapes, thereby accelerating convergence to the global optimal solution. For example, when solving complex nonlinear function optimization problems, a deep neural network can be constructed to approximate the objective function, and the gradient information of the neural network can be used to guide the optimization process. By continuously learning the changing patterns of the function landscape, deep learning can gradually adjust the optimization direction, avoid falling into local optima, and accelerate convergence to the global optimum. This practice has achieved significant results in multiple fields, demonstrating the enormous potential of deep learning in function optimization problems.

5. Data-driven deep learning numerical computation methods

5.1. Basic principles of data-driven modeling

Data-driven modeling is a method of constructing and training numerical models based on large amounts of data. Unlike traditional methods that rely on physical laws and mathematical equations for modeling, data-driven modeling mainly relies on statistical patterns and pattern recognition of data. In deep learning, this is typically achieved by building deep neural networks and training them using large amounts of data. These networks are capable of automatically extracting features from data and learning complex mapping relationships behind the data [8].

Compared to traditional physical models, data-driven models are more flexible and able to adapt to more complex and changing environments. However, they also depend on the quality, and quantity of data, and may lack physical interpretability. However, data-driven models and physical models are not completely opposed, and the two can be integrated to leverage their respective strengths.

5.2. The application of deep learning in data assimilation

Data assimilation is a technique that combines observational data with physical models to improve the accuracy of numerical predictions. Deep learning plays an important role in data assimilation. By constructing deep neural networks, observation data can be integrated into the prediction process of physical models, thereby correcting the initial conditions or parameters of the model and improving the accuracy of predictions as shown in **Figure 5**.

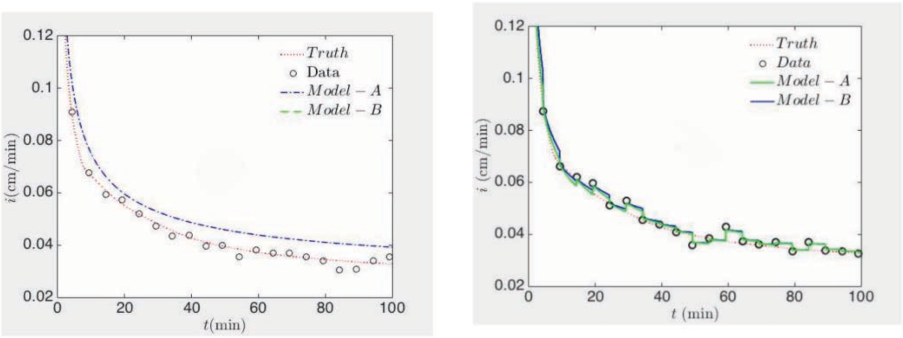

Figure 5. Data assimilation

Data assimilation techniques have been widely applied in fields such as meteorological forecasting and ocean simulation. For example, in meteorological forecasting, deep learning can integrate satellite observation data, ground observation station data, and meteorological models to provide more accurate weather forecasts. In ocean simulation, deep learning can integrate buoy data, satellite remote sensing data, and ocean circulation models to improve the accuracy of ocean state prediction.

5.3. Practice of data-driven deep learning in uncertainty quantification

Uncertainty, such as randomness and fuzziness, often accompanies numerical calculations. Deep learning has shown great potential in dealing with these uncertain problems. By constructing appropriate neural network structures, deep learning can estimate the uncertainty of prediction results and provide confidence intervals or probability distributions [3].

For example, in financial risk assessment, deep learning can estimate the uncertainty of investment returns, providing decision-makers with risk assessment and decision support. In environmental science, deep learning can estimate the uncertainty of pollutant concentration prediction and provide a scientific basis for environmental policy formulation. These practices indicate that data-driven deep learning has broad application prospects in uncertainty quantification.

6. Challenges and future prospects of deep learning numerical computing methods

6.1. Current challenges in deep learning numerical computation methods

Although deep learning has shown great potential in the field of numerical computing, it still faces many challenges. Firstly, computational efficiency is a prominent issue. Deep learning models, especially deep neural networks, typically require a significant amount of computational resources and time to train, which is particularly evident when dealing with large-scale numerical computing problems. Secondly, the lack of interpretability is also a major drawback of deep learning numerical computation methods. Deep learning models are often seen as 'black boxes', and their decision-making processes lack clear physical or mathematical explanations, which limits their application in certain fields that require interpretability. Finally, generalization ability is also a concern for deep learning numerical computation methods. Although deep learning performs well on training data, its performance may significantly decrease when faced with unseen data [9].

These challenges have had a profound impact on the development of deep learning numerical computation methods. The issue of computational efficiency limits the application of deep learning in real-time or large-scale numerical computing; Insufficient interpretability hinders the promotion of deep learning in certain fields; The issue of generalization ability requires researchers to pay more attention to the generalization performance and robustness of the model.

6.2. Fusion strategy of deep learning and other numerical methods

To overcome the challenges of deep learning numerical computation methods, researchers have begun to explore ways to integrate deep learning with traditional numerical methods. For example, deep learning models can be combined with traditional numerical methods such as finite element and finite difference, utilizing the powerful learning ability of deep learning to improve the computational efficiency and accuracy of traditional methods.

In specific cases, this fusion method has demonstrated significant advantages. For example, in fluid dynamics simulations, introducing deep learning models to accelerate the solving process can significantly improve computational efficiency; In solid mechanics analysis, using deep learning models to optimize the generation of finite element meshes can improve computational accuracy [10]. These practices indicate that the integration of deep learning and traditional numerical methods is an effective way to improve numerical computing performance.

6.3. Future prospects of deep learning numerical computation methods

Looking ahead to the future, the development of deep learning in the field of numerical computing will show the following trends: firstly, computational efficiency will continue to improve. Through algorithm optimization and hardware acceleration, the computational efficiency of deep learning numerical computing methods will be significantly improved; Secondly, interpretability will gradually increase, and researchers will strive to develop more interpretable deep learning models to meet the demand for interpretability in certain fields; Thirdly, the generalization ability will receive more attention, and researchers will place greater emphasis on the model's generalization performance and robustness to improve the practicality of deep learning numerical computation methods [11].

To promote the development of deep learning numerical computing methods, it is recommended to strengthen interdisciplinary cooperation and promote the deep integration of deep learning and traditional numerical methods; At the same time, research and development investment should be increased to support the application and promotion of deep learning numerical computing methods in more fields.

7. Conclusion

Deep learning has shown great potential in the field of numerical computing, bringing innovation to traditional numerical methods. Despite facing challenges such as computational efficiency, interpretability, and generalization ability, the application prospects of deep learning numerical computing methods are broad through the integration with traditional numerical methods and optimization of algorithms and hardware. In the future, with the continuous advancement of technology, deep learning will play an important role in more scientific computing and engineering applications, providing efficient and accurate solutions for solving complex problems.

Disclosure statement

The author declares no conflict of interest.

References

[1] Liu D, Chen Q, Wang X, 2024, Deep Learning Method for Solving Linear Integral Equations with Primitive Function Transformation. Computational Physics, 41(5): 651–662.

[2] Li Q, Zhu R, Li S, et al., 2024, Overview of Numerical Calculation Methods for Electrothermal Coupling of Converter Transformers. High Voltage Technology, 50(10): 4734–4748.

[3] Niu Q, Dong S, 2023, Research on Numerical Calculation Method of Infrared Radiation from Rocket Engine Spray Flames (Invited). Optoelectronic Technology Application, 38(1): 1–10.[4] Hu J, Zhu H, Huang X, et al, 2024, Forest Mobile Robot Tree Trunk Detection based on Deep Learning. Forest Engineering, 40(4): 109–114.

[5] Yin H, Nan Z, 2024, Investor Sentiment Mining based on Deep Learning and its Impact on Stock Price Collapse Risk. Journal of Central University of Finance and Economics, (3): 36–56.[6] Zhang H, Zhang H, 2022, Research on the Prediction Method of Effective Proliferation Factor of Core Based on Convolutional Neural Network Model. Modern Applied Physics, 13(2): 6.

[7] Yang C, 2021, Review on Theoretical Model and Numerical Methods of Hydraulic Fracturing. Yunnan Chemical Industry, 48(1): 17–20.[8] Li X, Zhang P, Chen G, et al., 2019, The Rising Behavior of Bubbles and Interfacial Mass Transfer in Liquid: Experimental Study and Numerical Simulation. Advances in Chemical Engineering, 38(2): 740–751.

[9] Dinsmore D, Alexander P, 2012, A Critical Discussion of Deep and Surface Processing: What It Means, How It Is Measured, the Role of Context, and Model Specification. Educational Psychology Review, (24): 499–567.

[10] Chang Z, Martin V, Tammy S, 2009, A Cross-cultural Study of Online Collaborative Learning. Multicultural Education & Technology Journal, 3(1): 33–46.

[11] Gerry G, Desiree J, Paul A. Kirschner, 2016, Changing Learning Behavior: Self-efficacy and Goal Orientation in PBL Groups in Higher Education. International Journal of Educational Research, 75: 146–158.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research Progress on Particle Behavior and Dynamics in Optical Tweezers

Jiali Pan*

China Jiliang University, Hangzhou 310018, Zhejiang, China

***Corresponding author:** Jiali Pan, panjiali0206@163.com

Abstract: Optical tweezers technology utilizes the optical potential well generated by a focused laser beam to achieve precise manipulation of micro and nanoparticles. Based on the optical tweezers platform, the motion behavior and dynamic laws of particles are deeply studied, which can reveal the transport mechanism of complex systems. Based on summarizing the principles and experimental methods of optical tweezers technology, this article systematically summarizes the typical force characteristics of particles in optical tweezers, focusing on the dynamic research progress of single particle non-equilibrium state, double particle coupling, and multi-particle cluster system, laying a theoretical foundation for expanding the application of optical tweezers technology in physics, chemistry, biology, and other fields.

Keywords: Optical tweezers; Particle manipulation; Brownian motion; Non-equilibrium state; Coupling dynamics

Online publication: April 2, 2025

1. Introduction

Since its inception in 1986, optical tweezers technology has been widely used in micro nanoparticle manipulation and property characterization due to its unique non-contact manipulation capability and nanoscale spatial resolution. Especially in recent years, the combination of optical tweezers with pump-probe spectroscopy, microfluidic chips, and other technologies has provided new research ideas and experimental methods for studying the transport behavior of particles in complex environments and revealing their underlying dynamic mechanisms. Based on reviewing the development of optical tweezers technology, this article will focus on exploring the typical behavioral characteristics and dynamic models of particle systems in optical tweezers, to provide a reference for a deeper understanding of the physical mechanisms of particle manipulation and to expand the application of optical tweezers technology in cutting-edge interdisciplinary fields.

2. Principles and experimental methods of optical tweezers technology

2.1. Basic composition of optical tweezers system

The core of the optical tweezers system is a highly focused laser beam. When the laser beam is focused on a small particle, the refractive index difference between the particle and the surrounding medium causes the incident photon to interact with the particle, resulting in a change in momentum. According to the law of conservation of momentum, the particle will experience a force in the opposite direction of light propagation. When the optical tweezers force reaches equilibrium with the gravity of the particles, the particles are confined near the optical focal point, forming a stable three-dimensional optical potential well [1].

A typical optical tweezers system mainly consists of a laser, focusing objective lens, sample cell, position-sensitive detector, etc. The laser provides a continuous Gaussian laser beam, which is focused by the objective lens a few microns above the particles to be captured, forming an optical potential well. The particles are captured and suspended in the sample pool under the action of optical tweezers. By controlling the polarization state, power, wavelength, and other parameters of the incident beam, precise control of the particle capture position and force can be achieved [2].

2.2. Theoretical calculation model of optical tweezers force

In optical tweezers systems, accurately estimating the magnitude of the optical tweezer force acting on particles is crucial for achieving particle manipulation. According to the Mie scattering theory, when the particle size is much smaller than the wavelength of the incident light, the Rayleigh approximation can be used to describe the interaction between particles and the optical field. Assuming the polarizability of a single isotropic spherical particle is α, the environmental dielectric constant is ε, and the incident light intensity is *I*, the optical tweezer force *F* acting on the particle can be expressed as a gradient of light intensity: $\vec{F} = \frac{1}{2}\alpha\nabla\left(\frac{I}{c\varepsilon}\right)$ Where c is the speed of light.

The optical tweezer force exerted on particles in the Rayleigh region is directly proportional to the intensity of the incident light. When the particle size is equal to or larger than the wavelength, strict Mie scattering theory needs to be used for calculations, simplifying the incident light into an infinitely extended plane wave. When the light is incident on the surface of the particle, some of it is reflected, while the rest is refracted into the interior of the particle. Considering the absorption and multiple scattering effects of particles on light, the resultant force exerted on particles can be further decomposed into two components: scattering force directed towards the direction of light propagation and gradient force directed towards the direction of light intensity gradient. Scattering force drives particles to move along the direction of light propagation. Gradient force provides lateral constraint, trapping particles at the point of maximum light intensity, and the combined force of the two forces determines the equilibrium position of the particles in the optical tweezers system [3].

2.3. Experimental methods for particle capture and manipulation

Based on the optical tweezers system, precise manipulation of single and multiple particles can be achieved. Its core is to modulate the spatiotemporal distribution of the incident light field, apply precise and controllable optical tweezer forces, and guide the migration and assembly of particles.

For single particle systems, directional transport of particles can be achieved by controlling the polarization state of the incident beam and utilizing the anisotropy of the optical tweezers potential well. If radially polarized light is used to form an asymmetric optical tweezers array in the horizontal plane of the sample cell, it can induce particles to migrate in a specific direction [4].

For multi-particle systems, spatial light modulators can be used to dynamically control the incident light wavefront. Real-time modification of the shape, size, and arrangement of the optical tweezers array enables precise assembly and manipulation of particles. For example, computer-generated holograms can be used to modulate the incident Gaussian light wave, projecting light intensity distributions of any shape such as circular or spiral on the surface of the sample pool, thereby controlling the spatial arrangement and combination of particles. Constructing microstructures with specific patterns, introducing time-dependent terms to incident light using acousto-optic modulators, and dynamically controlling the assembly process of particles in real-time [5]. Machine vision feedback is also an important means to improve the spatiotemporal resolution of particle manipulation. Real-time particle motion images are collected by high-speed CCD cameras and combined with digital image processing algorithms to track particle position and velocity in real-time, obtain information such as particle spatiotemporal trajectories, and integrate visual feedback signals into closed-loop control systems. By adjusting the magnitude and direction of the optical tweezers force in real-time, precise servo control of individual and multiple particles can be achieved, enabling them to move along predetermined trajectories and significantly improving the spatiotemporal resolution of particle manipulation [6].

3. Typical behavioral characteristics of particles in optical tweezers

3.1. Brownian motion and fluctuations of particles

Small particles in the suspended liquid will be randomly impacted by the thermal motion of surrounding solvent molecules, exhibiting irregular Brownian motion. Even if a single particle is confined in an optical tweezers potential well, it will still undergo restricted Brownian motion near the center of the potential well. Its position and velocity fluctuate over time, and the movement of particles is influenced by the combined effects of solvent viscosity resistance, optical tweezers binding force, and random collisions of solvent molecules. By analyzing these factors, we can better understand the restricted Brownian motion characteristics of particles in potential wells and provide a theoretical basis for related research.

Solving the stochastic differential equation can obtain the root mean square value of the particle displacement χ, which reflects the amplitude of the fluctuation of the particle around the binding center: $\sqrt{\langle x^2 \rangle} = \sqrt{\frac{k_B T}{\kappa}}$. It can be seen that the fluctuation amplitude of particles is directly proportional to the square root of the system temperature and inversely proportional to the square root of the stiffness of the optical tweezers. By analyzing the time series of particle displacement, it can characterize the characteristics such as the binding strength of the optical tweezers potential well, and can also be used to detect the spatiotemporal fluctuations of local environmental temperature [7].

The restricted diffusion motion of particles in the potential well also satisfies the Einstein-Smoluchowski relationship: $D = \frac{k_B T}{\gamma} = \frac{k_B T}{6\pi\eta a}$ Among them, D is the diffusion coefficient of the particle, η is the fluid viscosity, and α is the particle radius. By measuring the change of the particle's mean azimuthal shift with time, the diffusion coefficient of the particle can be directly calculated, and the microenvironment parameters of the local fluid can be quantitatively characterized, providing a new idea for quantitative analysis of particle fluid interactions.

3.2. Interaction forces between particles

When capturing multiple particles, additional interaction forces are generated between adjacent particles due to the interaction of the scattering field. For two spherical dielectric particles with a radius of a and a distance of *r*, the classical dipole approximation theory provides that the interaction potential energy *U(r)* between particles can be expressed as: $U(r) =-\frac{A}{r^6}+\frac{B}{r^{12}}$ The first term is the attractive potential caused by dispersion force, the second term is the repulsive potential, and A and B are constants related to particle properties.

This potential energy form is similar to the Lanner-Jones potential, with an equilibrium distance where two particles are balanced and form a stable bound state. By adjusting the particle spacing, the particles can be induced to self-assemble and form composite structures such as dimers [8].

When the particle spacing is small to a certain extent, the high-order multipole moment effect of the scattering field becomes non-negligible. The latest research shows that high dielectric constant particles will excite significant beam self-focusing effect inside the particles under the action of a strong focusing laser field, resulting in uneven distribution of the scattering field, which induces the particles to produce additional dipole moments and exhibits excellent scattering force characteristics along the optical axis. Theoretical analysis shows that this force is proportional to the fourth power of the particle radius and plays a key role in the stable confinement of submicron particles. The multiple polar moment interactions induced by scattering fields have become a new factor affecting the stability of complex particle systems [9].

3.3. Particle migration induced by optical tweezers potential well

Based on the asymmetry of optical tweezers force, optical tweezers technology can be used to manipulate the directional migration of particles. According to the Maxwell stress tensor method, the optical force acting on particles can be expressed as the integral of the electromagnetic field on the particle surface: where T is the Maxwell stress tensor and S is any closed surface surrounding the particle.

For an ideal symmetric optical field, the particle experiences zero resultant force and is in a state of force equilibrium, but in practical optical tweezers systems. When the focused light field is transmitted in a medium, distortion occurs, and the axial light intensity distribution becomes asymmetric, resulting in additional scattering forces along the direction of light propagation on the particles. When the pushing effect of the scattering force exceeds the binding effect of the gradient force, the particles will detach from the potential well and migrate along the direction of light propagation.

To achieve flexible and controllable direction and rate of particle migration. Dynamic control of the spatial distribution and polarization state of the incident light field is required, such as using a liquid crystal spatial light modulator to spiral phase modulate the wavefront of a Gaussian beam, forming a vortex optical trap array with rotational symmetry of light intensity distribution in the sample plane. The captured particles will migrate along a circular orbit under the action of rotational light torque, and the migration speed can be dynamically controlled by adjusting the light intensity and polarization state.

4. Research progress on particle dynamics of optical tweezers

4.1. Single particle non-equilibrium dynamics

By binding a single particle in an optical tweezers potential well and modulating the binding potential over time, the particle can move away from the thermodynamic equilibrium state and exhibit unique transport behavior. Theoretical studies have shown that in a potential well where the stiffness of the optical tweezers is periodically modulated over time. The motion of particles satisfies the generalized Langevin equation:$m \cdot \frac{d^2x(t)}{dt^2} =- \gamma \cdot \frac{dx(t)}{dt} - \kappa \cdot x(t) + \sqrt{2k_BT\gamma} \cdot \zeta(t)$ The potential well stiffness *κ(t)* exhibits a periodic function.

Numerical simulations have found that there is a synchronization phenomenon between the migration direction of particles and the modulation frequency of the driving force. At a specific frequency, particles mainly undergo directional migration along the stiffness gradient direction of the optical tweezers, and there is an optimal modulation frequency at which the particle transport efficiency is highest. Unbalanced modulation optical tweezers provide a new driving mechanism for particle transport [10].

Further research has shown that periodic modulation optical tweezers can also induce random resonance phenomena in particles by periodically modulating the electric field component of incident polarized light. A bistable potential well that evolves can be constructed, and under appropriate noise intensity, the transition rate of particles between bistable states is synchronized with the driving modulation, thereby significantly enhancing transport efficiency. The synergistic effect of non-equilibrium driving and noise can greatly amplify the weak transport response inherent in the system, which provides a new idea for using optical tweezers to construct artificial brown noise environments and study the stochastic resonance effect of particle transport.

4.2. Double particle coupled motion behavior

When capturing two particles in adjacent optical tweezers potential wells, the dual particle system exhibits distinct linkage behavior compared to a single particle due to the interaction of scattering fields and the coupling effect of surrounding fluids. Experimental studies have found that when two identical micron-sized polystyrene spheres are captured in dual-beam optical tweezers at a certain angle. The two spheres exhibit synchronous periodic oscillations along the optical axis, and their oscillation amplitude and frequency are closely related to the angle between the light beams.

Theoretical analysis shows that when two beams of light overlap in space, the scattered light field produces a significant self-focusing effect between the two microspheres. Causing additional light pressure, thereby inducing oscillatory motion between the two balls.

The fluid coupling effect in the dual particle system is also an important factor affecting the dynamic behavior of particles. The latest research used optical tweezers technology to study the Brownian motion of two colloidal microspheres in viscoelastic fluids and found that when the two microspheres approach a certain distance [11]. Their Brownian motions are no longer independent of each other, but show a clear correlation, exhibiting a synergistic diffusion phenomenon. The motion of each microsphere will excite viscoelastic stress waves in the surrounding fluid. When two microspheres are very close, the stress waves will couple with each other, transferring momentum between the two microspheres, thus causing the correlation between the motion of the two microspheres. By measuring the cross-correlation function of particle displacement, microrheological parameters such as the viscoelastic relaxation time of fluid media can be quantitatively extracted. The dual particle system has opened up new research ideas for the micro-viscoelastic characterization of complex fluid materials.

4.3. Dynamics model of multi-particle cluster

Under the action of a strong focusing optical tweezers field, a large number of particles spontaneously form complex cluster structures under the interaction of multiple scattering forces and gradient forces [12]. In the experiment, a large number of polystyrene microspheres were captured by an optical tweezers array, and it was found that the microspheres moved randomly in the potential well, gradually aggregated, and finally formed island-shaped cluster structures.

Theoretical analysis reveals that the fluctuations in particle density satisfy the Cahn-Hilliard equation:

$\frac{\partial \phi}{\partial t} = \nabla \cdot [M \cdot \nabla \mu]$ Among them, φ is the particle density, *M* is the mobility, and μ is the chemical potential. Further consideration is given to the interaction force exerted by optical tweezers on particles and the density fluctuations caused by Brownian motion. The dynamic process of particle cluster formation can be quantitatively described. Studies have shown that by optimizing the optical field distribution parameters of the optical tweezers array, the morphology of the cluster structure formed by particle self-assembly can be controlled, providing a new approach for the preparation of functionalized colloidal materials.

Introducing active particles into the optical tweezers system provides a new approach for constructing thermodynamic open systems. By modifying the surface of the particles with light-driven molecular motors, the particles can continuously absorb energy from the environment under illumination, overcoming viscous dissipation. To maintain the system in a steady state far from equilibrium, the dynamic behavior of this active particle system can be quantitatively described by a continuum model as follows:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0$$

$$\rho\left(\frac{dv}{dt} + v \cdot \nabla v\right) = -\nabla p + \eta \nabla^2 v + f$$

Among them, ρ is the particle density field, v is the velocity field, ρ is the pressure, η is the viscosity, and f is the random force including self-driving force and noise, in a steady state. The density and velocity fluctuation spectra of particle clusters are directly related to the statistical characteristics of particle self-driving velocity and noise. Active particle systems exhibit novel dynamic behaviors that differ from classical Brownian particle systems, such as sustained directional transport, large density fluctuations, and self-organizing pattern formation. Optical tweezers technology is an ideal model for studying statistical physical laws far from equilibrium systems. It provides a powerful tool for constructing and manipulating thermodynamic open systems at the microscale and is expected to deepen people's understanding of the physical mechanisms of non-equilibrium phenomena commonly present in living systems as shown in **Figure 1**.

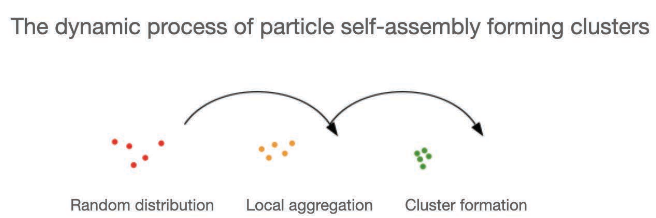

Figure 1. Schematic diagram of the formation process of multi-particle clusters under the action of optical tweezers

5. Conclusion

Optical tweezers technology, with its unique single particle manipulation and ultra-high spatial resolution imaging capabilities, has become a powerful tool for studying particle dynamics behavior. Based on a review of the development of optical tweezers technology, this article focuses on exploring the typical characteristics of particle forces in optical tweezers. Summarizing the research progress in single particle non-equilibrium transport, dual particle coupled motion, and dynamic models of multi-particle systems, these research works greatly expand people's understanding of the transport mechanism of complex particle systems, provide new ideas for developing new colloidal materials, constructing intelligent drug delivery systems, and revealing the working mechanism of intracellular molecular motors.

Looking ahead to the future, the integration of optical tweezers technology with micro-nano processing, surface plasmon resonance, microfluidic chips, and other technological means is expected to achieve multi-field coupling control of particle dynamics behavior, achieve integrated and intelligent particle manipulation, and promote innovative applications of optical tweezers in interdisciplinary fields such as physics, chemistry, and biology. Introducing optical tweezers into novel physical environments such as mixed dimensional systems and active substance systems provides a new research paradigm for studying non-equilibrium statistical physics developing biomimetic intelligent materials, and the development of optical tweezers technology. It will undoubtedly encourage people to reveal the laws of motion of the material world, manipulate the structure and function of matter at the micro and nano scales, and continuously expand the cognitive boundaries of humanity towards the microscopic world.

Disclosure statement

The author declares no conflict of interest.

References

[1] Zhang Y, Zhang S, Min C, et al., 2021, Research and Application Progress of Femtosecond Optical Tweezers Technology. Chinese Journal of Lasers, 48(19): 205–218.

[2] Cui L, Yang M, Lu C, et al., 2020, Research on Optical Sorting Theory, Technology, and Application Based on Optical Tweezers. China Laser, 1–36.

[3] Zhong H, Chen J, Chen J, et al., 2024, Progress in Optical Trapping and Spectroscopic Measurements of Airborne Particles. Chinese Laser, 51(03): 215–232.

[4] Tao Y, Zhong W, Wu X, et al., 2023, The Basic Principles and Applications of Optical Torque. Acta Optica Sinica, 43(16): 175–192.

[5] Chen D, 2023, Research on Functional Composite Capillary Fiber Optic Tweezers, dissertation, Guilin University of Electronic Science and Technology.

[6] Yang J, Deng R, Wang X, et al., 2024, Research Progress on Photothermal Tweezers Technology and Applications. China Laser, 51(18): 1801005.

[7] Huo C, Qiu S, Liang Q, et al., 2023, Progress in the Trapping and Manipulation Volume of Optical Tweezers. Chinese Journal of Physical Chemistry, 40(1): 11–27.

[8] Zhong Y, Peng Y, Chen J, et al., 2023, Optical Temperature Field-driven Tweezers: Principles and Biomedical Applications. Acta Optica Sinica, 43(14): 9–29.

[9] Cheng H, 2023, Research on Single Particle Suspension and Scattering Measurement Technology Based on Fiber Optic Tweezers, dissertation, Beijing Jiaotong University.

[10] Sun J, Zhang Y, 2024, Measurement Technology of Single Droplet Optical Tweezers with Micrometer Sub Micrometer Size. Journal of Light Scattering, 36(4): 363–374.

[11] Reisenbauer M, Rudolph H, Egyed L, et al., 2024, Non-Hermitian Dynamics and Non-reciprocity of Optically

Coupled Nanoparticles. Nature Physics, 20(10): 1629–1635.

[12] Deng H, Song N, Luo Z, et al., 2023, Single Fiber Optical Tweezers Based on Optical Sharp Edge Diffraction, CN202111136433.6A, CN113764116B.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Exploring the Application of Blockchain and IoT Technology in Commodity Management

Zhangjun Jiang[1]*, Xiaojian Zhu[2]

[1]Seenew Medical technology (Zhejiang) Co., Ltd., Hangzhou 311111, China
[2]Hangzhou Kuowei Information Technology Co., Ltd., Hangzhou 311111, China

Corresponding author: Zhangjun Jiang, jiangzhangjun@come-future.com

Abstract: This study provides a detailed analysis of the application of blockchain and Internet of Things (IoT) technologies in various aspects of commodity management, addressing issues such as information asymmetry, data security and privacy challenges, insufficient supply chain transparency, and difficulties in regulation. The study also explores the challenges and strategies associated with the implementation of these technologies. Through this analysis, the article aims to provide theoretical support and practical reference for improving the efficiency and quality of commodity management, thereby promoting the digital transformation of commodity management.

Keywords: Blockchain; Internet of Things (IoT); Commodity management; Pricing; Supply chain

Online publication: March 27, 2025

1. Introduction

In the global economic system, commodities serve as basic raw materials and the efficiency and stability of their management play a crucial role in the development of various industries. From energy to metals and from agricultural products to chemical products, the smooth circulation of commodities affects the cost and efficiency of the entire industrial chain. However, with the increasing complexity of the market and the deepening of globalization, traditional commodity management models are facing many severe challenges. In recent years, the rapid development of blockchain and Internet of Things (IoT) technologies has provided new ideas and methods for solving these problems. Blockchain technology, with its decentralized, tamper-proof, and traceable characteristics, is expected to reshape the trust foundation of commodity management, while IoT technology, with its powerful data collection and transmission capabilities, can realize real-time monitoring and management of the entire lifecycle of commodities. Conducting in-depth research on the application of these two technologies in commodity management has important practical significance for enhancing industry competitiveness, optimizing resource allocation, and ensuring market stability.

2. Application of blockchain in commodity pricing

2.1. Achieving transparent pricing

The distributed ledger characteristic of blockchain technology provides strong support for transparent pricing of bulk commodities. The distributed ledger is maintained by multiple nodes, and all transaction data is publicly transparent and tamper-proof. In commodity trading, information such as the price, quantity, and transaction time of each transaction will be recorded on the distributed ledger, which can be viewed and verified by market participants. This allows the price formation process to be based on real and accurate transaction data, avoiding price distortions caused by information asymmetry or market manipulation. Participants can make price analysis and decisions based on these public data, improving the transparency and fairness of pricing, and making prices more reflective of the true value of commodities [1].

2.2. Automation of pricing mechanisms

The smart contract function of blockchain can automate the pricing mechanism of bulk commodities. A smart contract is an automatically executed contract that exists in the form of code on the blockchain, with its terms and conditions preset. In the commodity trading scenario, smart contracts can automatically perform pricing operations based on preset quality standards, quantity requirements, delivery time, and other conditions. When the transaction conditions are met, the smart contract automatically performs price calculations and transaction settlements according to the agreed price formula, without manual intervention. This not only improves pricing efficiency, reduces human errors and price disputes, but also enhances the accuracy and reliability of the pricing process [2].

2.3. Decentralized data sharing and supply chain traceability

The decentralized nature of blockchain breaks the traditional information islands and enables data sharing across various links in the commodity supply chain. In the blockchain network, there is no centralized control node, and all participants can equally record and access data. Data generated in each link, such as production data, transportation records, and warehousing information, will be encrypted and stored on the blockchain, and the consistency and accuracy of the data will be ensured through a consensus mechanism. At the same time, the traceability of the blockchain allows the entire process information of bulk commodities from the source of production to the sales terminal to be clearly recorded and queried. Through blockchain browsers or related applications, regulatory agencies, enterprises, and consumers can obtain detailed information about commodities, achieving transparent management of the supply chain [3].

2.4. Advantages in derivative pricing and cross-border payments

In terms of commodity derivative pricing, blockchain technology can provide more accurate and real-time data support. Derivative prices are closely related to underlying commodity prices. The underlying commodity transaction data, inventory information, and market supply and demand data recorded on the blockchain can reflect market dynamics in real-time, providing precise input parameters for the derivative pricing model. This helps to build a more reasonable derivative pricing model, improve the accuracy and rationality of pricing, and reduce market risks. In the field of cross-border payments, blockchain technology can solve the problems of tedious procedures, high costs, and long settlement cycles in traditional cross-border payments. The distributed ledger and encryption technology of the blockchain enable cross-border payments to be made directly between the two parties to the transaction, without relying on multiple intermediate financial institutions. The transaction parties transmit value through the blockchain platform, and smart contracts automatically execute payment terms, ensuring payment security and accuracy. This significantly reduces the payment process and settlement time, lowers cross-border payment costs, and improves capital efficiency [4].

3. Application of Internet of Things technology in commodity management

3.1. Device identity authentication and data security

The application of Internet of Things (IoT) technology in the field of commodity management is particularly prominent in its device identity authentication mechanism and data security guarantees. This mechanism ensures the integrity and credibility of data in a complex supply chain environment. Under the extensive connectivity of the IoT, each device connected to the network is assigned a unique digital identity, which serves as their "network ID card". This identity not only ensures the traceability of the device but also lays a solid foundation for subsequent data interaction [5].

Before data interaction, the IoT platform utilizes advanced encryption technology to authenticate devices. This process ensures that only authorized devices can access the network and upload data. This strict identity authentication mechanism acts as a solid line of defense, effectively blocking access from unauthorized devices, thus preventing data tampering and forgery. The implementation of this mechanism greatly enhances the reliability of data sources, providing a strong guarantee for the accuracy of data in commodity management [6].

IoT technology also employs various data encryption techniques to ensure secure data transmission and storage. During data transmission, whether through wired or wireless means, the IoT uses symmetric or asymmetric encryption algorithms to encrypt the data. This encryption method ensures that even if the data is intercepted during transmission, it cannot be easily decrypted, thus maintaining data confidentiality. In terms of data storage, the IoT platform encrypts data stored in the cloud or locally. Only authorized entities with the corresponding decryption key can access and read this data, further enhancing data security.

3.2. Data sharing and exchange mechanisms

IoT technology has established efficient data sharing and exchange mechanisms, enabling real-time data circulation across various stages of commodity management. Serving as a data aggregation and management center, the IoT platform connects IoT devices throughout the supply chain. These devices collect various types of data such as temperature, humidity, location, and status in real-time, uploading it to the IoT platform. Through standardized data interfaces and protocols, the platform integrates and processes data from different formats and sources, subsequently sharing it with authorized participants. For instance, manufacturing enterprises can share production data with transportation and warehousing companies, while transportation enterprises can share cargo location and status data with sales companies and regulatory bodies. This data-sharing mechanism breaks down information barriers, enhances collaboration efficiency across supply chain stages, and enables participants to make decisions based on real-time, accurate data.

3.3. Application of smart contracts in automated operations

The integration of IoT and blockchain allows smart contracts to play a crucial role in automating commodity management operations. Smart contracts can automatically trigger corresponding actions based on data collected by IoT devices. In warehouse management, IoT devices continuously monitor environmental parameters such as temperature, humidity, and air quality within the warehouse. When monitored data exceeds preset ranges, smart contracts automatically activate ventilation systems, air conditioning units, or dehumidifiers to adjust the warehouse environment, ensuring safe storage of goods. In the supply chain logistics stage, when IoT devices detect that goods have reached their designated location and passed quality inspection, smart contracts automatically execute payment operations, completing the delivery process. This automation reduces manual intervention, improves management efficiency, lowers operational risks, and enhances the intelligence and automation level of the supply chain [7].

4. Technical application challenges and coping strategies

4.1. Challenges of Technology Integration

The significant differences in the technical architectures of blockchain and the Internet of Things (IoT) pose numerous challenges for their integration. Blockchain operates based on a distributed ledger and consensus mechanism, emphasizing data consistency and immutability. Conversely, the IoT focuses on device perception and data transmission, requiring real-time capabilities and the ability to handle large amounts of data. Regarding data format, blockchain data structures are typically organized in blocks with specific encryption and storage methods. However, the IoT collects data in various formats, lacking a unified standard, making direct data interaction between the two technologies difficult.

In terms of communication protocols, there are compatibility issues between the peer-to-peer (P2P) communication protocol commonly used in blockchain and the various IoT communication protocols such as Wi-Fi and Bluetooth. This creates obstacles for communication between devices and blockchain nodes. Additionally, blockchain data processing speeds are relatively slow. For instance, the Bitcoin blockchain has a limited number of transactions it can process per second, whereas the IoT needs to handle large amounts of real-time data generated by numerous devices. This makes it challenging to synchronize data transmission and processing efficiency between the two technologies. To address these challenges, the industry needs to establish unified technical standards, regulating data formats and communication protocols for blockchain and the IoT to facilitate interconnection and interoperability. Simultaneously, the development of compatible middleware can serve as a bridge between the two, enabling effective data conversion and transmission, enhancing the feasibility of technology integration.

4.2. Security and privacy risks

With the application of blockchain and IoT technologies in commodity management, security and privacy risks are increasingly prominent. While blockchain smart contracts offer the advantage of automated execution, they also carry the risk of vulnerabilities. Once the code of a smart contract is written, it is difficult to modify. If there are vulnerabilities and they are exploited by hackers, it could lead to severe consequences such as tampered transaction data or stolen funds. IoT devices also face security risks due to their limited computing and storage capabilities, making it difficult to implement complex security protection mechanisms. These devices can be easily hijacked, resulting in data theft.

Regarding data privacy protection, as the scope of data sharing expands, protecting the privacy of businesses and users while ensuring data circulation has become a significant challenge. To address these issues, it is essential to strengthen security audits, regularly detect and fix vulnerabilities in blockchain smart contracts, and enhance the security of smart contracts. For IoT devices, multiple encryption techniques and access control mechanisms, such as device identity authentication and encrypted data transmission, should be employed to prevent device attacks and data leaks. In terms of data privacy protection, technologies like homomorphic encryption and zero-knowledge proofs can enable data computation and sharing in an encrypted state, ensuring that data privacy is not compromised. Simultaneously, establishing and improving relevant laws and regulations to clarify the boundaries of data usage and privacy protection is crucial to provide legal safeguards for technology application.

5. Technical application challenges and coping strategies

5.1. Challenges of technology integration

The significant differences in the technical architectures of blockchain and the Internet of Things (IoT) pose numerous challenges for their integration. Blockchain operates based on a distributed ledger and consensus mechanism, emphasizing data consistency and immutability. Conversely, the IoT focuses on device perception and data transmission, requiring real-time capabilities and the ability to handle large amounts of data. Regarding data format, blockchain data structures are typically organized in blocks with specific encryption and storage methods. However, the IoT collects data in various formats, lacking a unified standard, making direct data interaction between the two technologies difficult.

In terms of communication protocols, there are compatibility issues between the peer-to-peer (P2P) communication protocol commonly used in blockchain and the various IoT communication protocols such as Wi-Fi and Bluetooth. This creates obstacles for communication between devices and blockchain nodes. Additionally, blockchain data processing speeds are relatively slow. For instance, the Bitcoin blockchain has a limited number of transactions it can process per second, whereas the IoT needs to handle large amounts of real-time data generated by numerous devices. This makes it challenging to synchronize data transmission and processing efficiency between the two technologies. To address these challenges, the industry needs to establish unified technical standards, regulating data formats and communication protocols for blockchain and the IoT to facilitate interconnection and interoperability. Simultaneously, the development of compatible middleware can serve as a bridge between the two, enabling effective data conversion and transmission, and enhancing the feasibility of technology integration.

5.2. Security and privacy risks

With the application of blockchain and IoT technologies in commodity management, security and privacy risks are increasingly prominent. While blockchain smart contracts enable automated execution, they also pose risks due to potential vulnerabilities [8]. Once a smart contract's code is written, it is difficult to modify. If there are vulnerabilities and they are exploited by hackers, it could lead to severe consequences such as tampered transaction data or stolen funds. IoT devices also face security risks due to their limited computing and storage capabilities, making it difficult to implement complex security protection mechanisms. These devices can be easily hijacked, resulting in data theft. Regarding data privacy protection, as the scope of data sharing expands, protecting the privacy of businesses and users while ensuring data circulation has become a significant challenge.

To address these issues, it is essential to strengthen security audits, regularly detect and fix vulnerabilities in blockchain smart contracts, and enhance the security of smart contracts. For IoT devices, multiple encryption techniques and access control mechanisms, such as device identity authentication and encrypted data transmission, should be employed to prevent device attacks and data leaks. In terms of data privacy protection, technologies like homomorphic encryption and zero-knowledge proofs can enable data computation and sharing in an encrypted state, ensuring that data privacy is not compromised. At the same time, establishing and refining relevant laws and regulations to define the boundaries of data usage and privacy protection is essential for ensuring legal safeguards in technology applications.

5.3. Cost-benefit balance

The initial application of blockchain and IoT technologies requires significant capital investment. In terms of equipment procurement, IoT devices have high acquisition costs and need to be continuously updated to meet technological advancements and business needs. Meanwhile, setting up and maintaining blockchain nodes also demands certain hardware and software investments. For system development, considerable manpower and time are required for the development and integration of blockchain and IoT systems to achieve their collaborative work. Regarding technical training, to enable employees to skillfully use the new technologies, enterprises need to organize professional training, which also increases costs. However, in the short term, the benefits brought by technology application may not be apparent and enterprises may face an imbalance between costs and benefits. For instance, the running-in period of the new system may temporarily reduce business efficiency and it takes time for the cost reduction and efficiency improvement brought by technology application to become evident. To achieve a cost-benefit balance, enterprises should conduct a comprehensive cost-benefit analysis, reasonably select technology application solutions based on their actual needs and business scale, optimize technology selection, and reduce unnecessary cost inputs. Simultaneously, they should develop a long-term technology application plan, evaluate the benefits of technology application from a long-term perspective, such as cost reduction, efficiency improvement, and enhanced market competitiveness, to achieve a balance between costs and benefits. Additionally, enterprises can reduce costs and increase efficiency by sharing technical resources and jointly developing applications with partners.

5.4. Talent shortage and skill improvement

Currently, there is a shortage of professionals in blockchain and IoT, and the skills of existing staff within enterprises are difficult to meet the needs of technology application. Blockchain technology involves knowledge in multiple fields, such as cryptography, distributed systems, and consensus algorithms, while IoT technology covers various aspects, such as sensor technology, communication technology, and data analysis. Compound talents who understand both blockchain and IoT technologies and are familiar with the commodity management business are even more scarce. This talent shortage restricts the promotion and application of technology. Therefore, it is necessary to strengthen school-enterprise cooperation, and universities and vocational colleges should offer relevant majors and courses to cultivate blockchain and IoT technical talents. Enterprises should carry out on-the-job training to improve the technical level and business capabilities of existing employees, encourage employees to learn new technologies and master new skills to adapt to the business changes brought by technology application [9]. Meanwhile, enterprises can introduce external professional talents to enrich the technical team and improve the enterprise's technological innovation ability and application level. In addition, industry associations can organize professional training and technical exchange activities to promote talent cultivation and technology dissemination [10].

6. Conclusion

Blockchain and IoT technologies have brought innovative solutions to commodity management, effectively addressing many challenges faced in current commodity management. In the future, with the continuous development and improvement of technology, blockchain and IoT technologies are expected to play a greater role in the field of commodity management, pushing the commodity market towards a more efficient, transparent, and safe direction. Enterprises and regulatory agencies should actively keep pace with technological advancements, increasing investment in research, development, and the application of new technologies to strengthen their competitiveness and regulatory capabilities in the global commodity market. Simultaneously, the academic community also needs to further study the application effects and optimization strategies of these two technologies in commodity management to provide stronger theoretical support for practice.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Tan Z, Zhao W, 2017, Application and Development of Blockchain Technology in China's Commodity Trading Market. Business Observation, 2017(5): 48–51.

[2] Liu Z, Zhang G, 2021, "Blockchain + Internet of Things" Boosts Food Supply Chain Traceability. Agricultural Engineering, 11(02): 51–55.

[3] Zou D, Jia Q, 2020, Review of the Application Progress of Blockchain in the Field of Commodity Trading. Business Economics Research, 2020(10): 154–157.

[4] Du H, Wu J, Li C, 2021, Optimization of Commodity Listing Transaction Mode Based on Consortium Blockchain. Innovation and Technology, 21(3): 69–77.

[5] Zhang L, Wu X, Li S, et al., 2023, Application of Blockchain + Internet of Things Technology in the Traceability of Agricultural Products in Western Hunan. Economic Research Guide, 2023(10): 42–44.

[6] Fan Y, Wu S, Gao X, 2023, Integration of Information Technology and Commodity Lifecycle Inspection. Quality and Certification, 2023(11): 57–59.

[7] Jiang Y, Yao X, Li B, 2020, Application of Internet of Things and Blockchain Technology in Improving Supply Chain Financial Risk Management. Business Exhibition Economy, 2020(6): 51–54.

[8] Fu Y, Kuang Z, 2023, Application of Blockchain Technology in China's Customs Supervision: Practical Problems and Implementation Paths. Customs and Economic and Trade Research, 44(2): 30–42.

[9] Zhu H, Chen L, 2019, Research on Commodity Spot Warehouse Receipt Financing and Trading Based on Blockchain Technology. Network Space Security, 10(11): 26–31 + 47.

[10] Zhu H, 2020, Research on the Digital Transformation of Commodity Trading Based on the Free Trade Zone Strategy. Future and Development, 44(05): 92–98 + 91.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Electronic Research and Application, 2025, Volume 9, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

LACC-RCE: A Local Adaptive Color Correction and Rayleigh-Based Contrast Enhancement Method for Underwater Image Enhancement

Tiancheng Liu*

School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, Beijing 100044, China

***Corresponding author:** Tiancheng Liu, liutiancheng1998@163.com

Abstract: Underwater images are inherently degraded by color distortion, contrast reduction, and uneven brightness, primarily due to light absorption and scattering in water. To mitigate these challenges, a novel enhancement approach is proposed, integrating Local Adaptive Color Correction (LACC) with contrast enhancement based on adaptive Rayleigh distribution stretching and CLAHE (LACC-RCE). Conventional color correction methods predominantly employ global adjustment strategies, which are often inadequate for handling spatially varying color distortions. In contrast, the proposed LACC method incorporates local color analysis, tone-weighted control, and spatially adaptive adjustments, allowing for region-specific color correction. This approach effectively enhances color fidelity and perceptual naturalness, addressing the limitations of global correction techniques. For contrast enhancement, the proposed method leverages the global mapping characteristics of the Rayleigh distribution to improve overall contrast, while CLAHE is employed to adaptively enhance local regions. A weighted fusion strategy is then applied to synthesize high-quality underwater images. Experimental results indicate that LACC-RCE surpasses conventional methods in color restoration, contrast optimization, and detail preservation, thereby enhancing the visual quality of underwater images. This improvement facilitates more reliable inputs for underwater object detection and recognition tasks.

Keywords: Underwater; Image enhancement; Local adaptive color correction; Rayleigh distribution stretching; Contrast enhancement

Online publication: March 28, 2025

1. Introduction

Underwater imaging is frequently degraded by blurring, color distortion, and reduced contrast, primarily due to the distinct optical properties of the underwater environment. The selective absorption of light in water leads to severe attenuation of red wavelengths, whereas blue and green wavelengths penetrate more effectively, resulting in a characteristic bluish-green appearance in underwater images. Furthermore, suspended particles and water molecules contribute to multiple scattering of light, which significantly diminishes image contrast. Additionally, underwater currents and environmental disturbances often induce camera instability, further degrading image quality and leading to blurring of object edges and contours. These degradation effects severely compromise feature representation in underwater images, creating significant challenges for downstream computer vision tasks, such as object detection, image segmentation, and recognition. Consequently, underwater image restoration and enhancement remains a topic of substantial practical significance. Existing approaches in this field can be generally categorized into three main types: non-physical model-based enhancement methods, physical imaging model-based restoration methods, and deep learning-based enhancement techniques.

Physical model-based underwater image restoration methods reconstruct degraded images by formulating underwater optical imaging models and estimating key parameters based on predefined priors. The restoration process is then conducted through inverse computation techniques. These methods primarily follow two technical pathways: one approach exploits the statistical characteristics of the darkest pixels to estimate and invert the dark channel, thereby facilitating image restoration. Notable algorithms include the Dark Channel Prior (DCP) [1] and its enhanced variant, the Underwater Dark Channel Prior (UDCP) [2]. Another approach to underwater image restoration relies on parameter estimation within underwater optical imaging models. Yu *et al.* [3] proposed a dehazing algorithm incorporating dual transmission maps, designed to adapt to varying underwater environments. Liu and Liang [4] employed grayscale morphological closing operations to estimate background light, effectively mitigating interference from white objects. Furthermore, they introduced a new underwater light attenuation prior (NULAP) and an adjusted reverse saturation map (ARSM) to enhance the accuracy and refinement of transmission map (TM) estimations. While these methods have demonstrated effectiveness in recovering color fidelity and fine details, their stability and consistency across diverse underwater conditions require further refinement.

Non-physical model-based underwater image enhancement methods focus on direct pixel-level enhancement without relying on physical imaging models. For instance, Song and Wang [5] employed white balance-based color correction to compensate for color distortions induced by medium attenuation. Additionally, they incorporated contrast and spatial cues through a saliency-weighted coefficient update strategy, aiming to achieve high-quality image fusion and enhancement. Zhang *et al.* [6] proposed an approach that enhances both global and local contrast using a dual-histogram-based iterative thresholding method and a limited histogram approach. To further refine the enhanced images, they employed a multi-scale fusion (MSF) strategy and a multi-scale unsharp masking (MSUM) technique. However, these methods may exhibit limited adaptability to diverse and complex underwater environments, often resulting in over-enhancement artifacts.

In recent years, deep learning-based image enhancement approaches have proliferated. For instance, Yan *et al.* [7] introduced a model-driven CycleGAN that integrates a physical model, enhancing both the effectiveness and generalization capability of traditional GAN-based methods in complex real-world underwater environments. Additionally, Wang *et al.* [8] developed UPGformer, a physics-guided transformer architecture designed to improve depth estimation accuracy. Ren *et al.* [9] introduced an enhanced Swin-Convs Transformer Block (RSCTB) designed to reinforce local attention mechanisms across both channel and spatial domains. This approach enhances the model's ability to perceive and restore images degraded by non-uniform medium distributions. However, deep learning-based methods demand extensive training data and impose high computational costs, while acquiring high-quality underwater datasets remains a critical challenge.

Underwater images frequently suffer from color distortion, contrast degradation, and uneven brightness distribution due to the selective absorption and scattering of light in water. To address these challenges, this study introduces a novel Local Adaptive Color Correction (LACC) method, integrated with a contrast enhancement framework leveraging adaptive Rayleigh distribution stretching and Contrast Limited Adaptive Histogram Equalization (CLAHE). This approach establishes a comprehensive underwater image enhancement pipeline. Unlike traditional global color correction techniques, which often fail to effectively adapt to spatially varying color distortions, the proposed method provides localized corrections, mitigating issues of under-correction or over-correction. Experimental results indicate that the proposed approach surpasses conventional methods in color fidelity and detail preservation.

2. Methods

Underwater images are inherently degraded by color distortion, contrast attenuation, and uneven brightness distribution, primarily caused by light absorption and scattering in water. To mitigate these effects, this study introduces a LACC method, integrated with a contrast enhancement technique that leverages adaptive Rayleigh distribution stretching and CLAHE.

Traditional color correction techniques primarily rely on global adjustments, which are often insufficient for addressing spatially varying color distortions. This limitation may result in some regions retaining noticeable color shifts while others experience excessive correction. The proposed LACC method incorporates local color analysis, tone-weighted control, and spatially adaptive adjustments, enabling region-specific correction and producing more perceptually natural and realistic color restoration. For contrast enhancement, the proposed method integrates adaptive Rayleigh distribution stretching and CLAHE. Initially, the global mapping properties of the Rayleigh distribution are leveraged to enhance overall contrast, followed by an adaptive truncation strategy to regulate the pixel stretching range. Subsequently, CLAHE is applied to refine local contrast adaptively, and a weighted fusion strategy is employed to synthesize the final enhanced image. The methodological flowchart is illustrated in **Figure 1**.

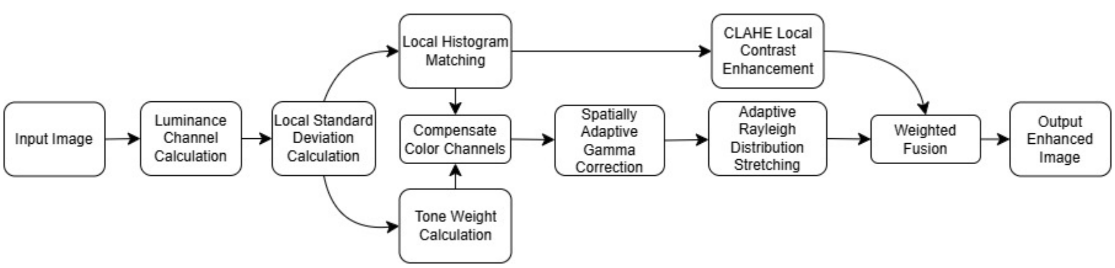

Figure 1. Methodological flowchart

2.1. Local Adaptive Color Correction (LACC) method

To enhance the accuracy of color distortion correction, this study employs a local standard deviation ratio approach to compensate for attenuation across color channels. The entire image is segmented into multiple small regions, where local statistical characteristics of both the luminance and color channels are computed independently. The luminance channel is defined as follows:

$$I_l(i,j) = max\{I_r(i,j), I_g(i,j), I_b(i,j)\} \quad (1)$$

where I_r, I_g, and I_b, correspond to the pixel values of the red, green, and blue channels, respectively, while

$I_l(i,j)$ represents the luminance of the given pixel.

Subsequently, the standard deviation of each color channel within the local window is calculated as follows:

$$\sigma_c^{local} = \sqrt{\frac{1}{K}\sum(I_c - \bar{I}_c)^2} \tag{2}$$

where K is the total number of pixels within the window, and $\bar{I}_c$ represents the mean intensity of the corresponding color channel.

Using the standard deviation ratio, the color channel values are modified to align more closely with the luminance channel:

$$I_c^{new} = I_c + (I_l - I_c) \times \frac{\sigma_c^{local}}{\sigma_l^{local}} \tag{3}$$

A smaller standard deviation in a color channel indicates lower variation, allowing for stronger alignment with the luminance channel. Conversely, a larger standard deviation suggests greater variability, warranting less adjustment to maintain the channel's original characteristics.

Despite compensating for color channel attenuation, regional variations in color distribution may still persist. To mitigate this issue, this study adopts a Local Histogram Matching (LHM) approach to ensure a more uniform color distribution across different regions of the image.

A reference set of high-quality underwater images is selected, from which the Cumulative Distribution Function (CDF) is computed as a standard. The histogram of each local window in the target image is then calculated and adjusted through CDF-based matching:

$$I_c^{match} = H_{ref}^{-1}\left(H_{input}(I_c)\right) \tag{4}$$

where H_{input} denotes the histogram of the input image, and H_{ref}^{-1} represents the inverse cumulative mapping of the reference image histogram.

During color correction, some regions may undergo overcompensation, resulting in color distortion. For example, blue tones may become overly dominant in certain areas, while red tones may be excessively enhanced. To mitigate this effect, this study introduces a Tone Weighting mechanism, ensuring that the correction magnitude is proportional to the deviation of the original tone. The tone weighting is formulated as follows:

$$W_c(i,j) = 1 - \left|\frac{I_c(i,j) - \bar{I}_c^{local}}{\bar{I}_c^{global}}\right| \tag{5}$$

where $\bar{I}_c^{local}$ denotes the mean intensity of the local window, represents the mean intensity of the entire image, and $W_c(i,j)$ regulates the extent of color compensation.

The final color adjustment is expressed as:

$$I_c^{final} = W_c \cdot I_c^{match} + (1 - W_c) \cdot I_c \tag{6}$$

In areas with significant color distortion, a lower W_c value leads to stronger color adjustments. Due to the uneven illumination in underwater environments, this study applies Spatially Adaptive Gamma Correction (SAGC) to enhance darker regions while mitigating overexposure in brighter areas.

2.2. Underwater image contrast enhancement based on adaptive Rayleigh distribution stretching and CLAHE

This study proposes a contrast enhancement approach that integrates adaptive Rayleigh distribution stretching with CLAHE. The method first leverages the global mapping properties of the Rayleigh distribution to improve overall contrast, while an adaptive truncation strategy is employed to constrain the pixel stretching range. Subsequently, CLAHE is applied for localized contrast enhancement, and a weighted fusion strategy is utilized to synthesize the final enhanced image.

To improve global contrast and ensure a more uniform brightness distribution, the Rayleigh distribution is employed for pixel value stretching. The probability density function (PDF) of the Rayleigh distribution is given by:

$$PDF(I) = \frac{I}{\sigma^2}\exp\left(-\frac{I^2}{2\sigma^2}\right) \quad (7)$$

where I denotes the pixel value of the input image, and σ serves as the distribution control parameter.

To accommodate diverse underwater lighting conditions, an adaptive computation method is utilized for determining σ:

$$\sigma = k \times \frac{I_{max} - I_{min}}{I_{avg}} \quad (8)$$

Where I_{max}, I_{min} denote the maximum and minimum pixel values of the image, respectively, I_{avg}represents the average pixel value, and k serves as the adjustment factor.

The Rayleigh distribution mapping function is formulated as follows:

$$I_{SR} = (I_{max} - I_{min}) \times \left(\frac{I - I_{min}}{I_{max} - I_{min}}\right)^{\sigma} + I_{min} \quad (9)$$

This transformation adjusts the pixel values to follow the Rayleigh distribution after enhancement, effectively enhancing overall contrast.

In the pixel stretching process, direct mapping may cause excessive enhancement of certain pixel values, leading to brightness distortion or loss of details. To address this, the Otsu thresholding method is utilized to determine the optimal pixel adjustment range, ensuring that the enhanced image preserves a well-balanced brightness distribution. The adjustment range is formulated as follows:

$$I_{out,min} = \max\left(I_{c,min}, T_{low}\right) \quad (10)$$

$$I_{out,max} = \min\left(I_{c,max}, T_{high}\right) \quad (11)$$

where $I_{c,min}$ and $I_{c,max}$ denote the minimum and maximum pixel values following the Rayleigh distribution transformation. T_{low} and T_{high} are derived from the Otsu thresholding method, representing the adaptive brightness range tailored to the current image.

This strategy effectively mitigates local overexposure and detail loss that may arise from global stretching, thereby improving the overall visual quality of the image.

Given that global contrast enhancement focuses on adjusting the overall brightness distribution, certain localized details may remain under-enhanced. To compensate for this limitation, CLAHE is incorporated to enhance local contrast adaptively.

$$I_{CLAHE} = \text{CLAHE}(I_{SR}) \tag{12}$$

where I_{CLAHE} denotes the CLAHE-enhanced image, and CLAHE(I_{SR}) represents the CLAHE transformation applied to the input image.

CLAHE improves local contrast by applying independent histogram equalization within small regions while restricting pixel values to mitigate over-enhancement artifacts, which are often observed in traditional histogram equalization. Given the complementary strengths and limitations of global and local contrast enhancement, a weighted fusion strategy is implemented to synthesize the final enhanced image:

$$I_{final} = \alpha \cdot I_{SR} + (1 - \alpha) \cdot I_{CLAHE} \tag{13}$$

Where α is the fusion weight, and I_{SR} corresponds to the image enhanced through Rayleigh distribution stretching.

This strategy ensures that the image retains global contrast improvements while simultaneously enhancing local details, producing a sharper and more naturally enhanced visual representation.

3. Experimental results and analysis

To assess the effectiveness of the proposed algorithm, comparative evaluations were performed from both subjective and objective perspectives in comparison with existing methods. The experiments were conducted on datasets collected from two representative underwater environments. The first dataset originates from the Jingyuan Shipwreck site in the Yellow Sea, where the seafloor is primarily composed of silty sand, a condition that promotes suspended particle formation, resulting in high turbidity and low visibility. The second dataset was collected from the Xisha underwater trench, where images predominantly suffer from severe color distortion. A total of 246 images were processed, all standardized to a 1024 × 1024 pixel resolution. Performance evaluation was conducted using both qualitative and quantitative approaches. Qualitative assessment involved visual comparisons to subjectively evaluate image quality, whereas quantitative analysis utilized a comprehensive set of evaluation metrics to objectively assess the algorithm's effectiveness.

Four existing underwater image enhancement methods were selected for comparison with the proposed approach: Retinex, UW-CycleGAN, UDCP, and UWCNN. These methods provide a comprehensive evaluation of the impact of different underwater image processing techniques on image quality.

To ensure objective assessment, three quantitative evaluation metrics were employed: Patch-Based Contrast Quality Index (PCQI), Underwater Color Image Quality Evaluation Metric (UCIQE), and Underwater Image Quality Measure (UIQM).

Five representative images were selected for visualization. The first dataset, as illustrated in **Figure 2**, was collected from the Jingyuan Shipwreck site.

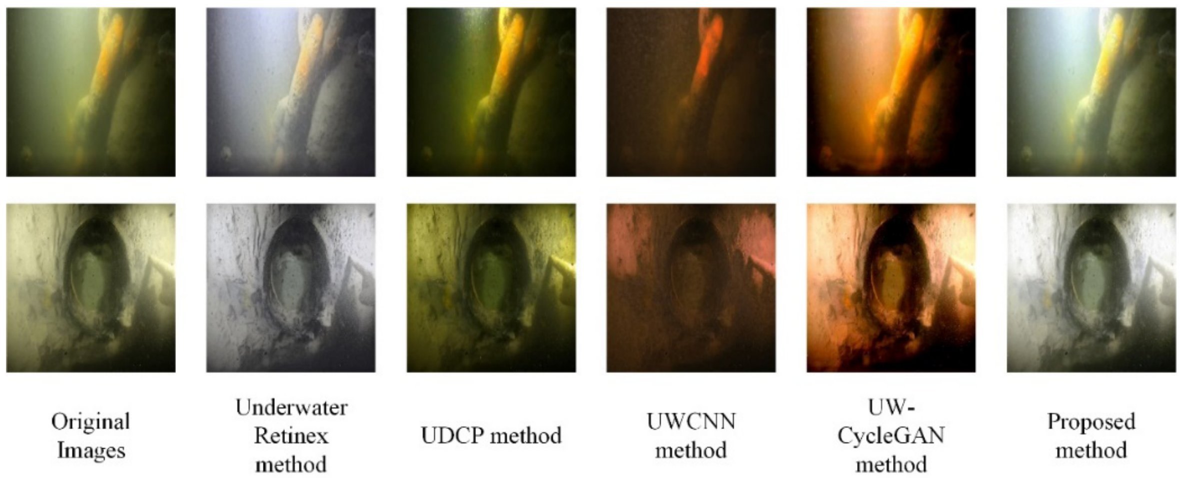

Figure 2. Images enhancement results of the proposed method compared to other methods

The turbidity of the underwater environment, combined with the presence of artificial light sources, causes significant light scattering, leading to a notable reduction in image details and contrast while also affecting color fidelity. Furthermore, overexposure artifacts are observed near the artificial light sources on both sides of the image. The UWCNN method leads to a complete loss of fine details. The Retinex method demonstrates limited effectiveness in underwater environments, with some images exhibiting noticeable sharpness degradation. The UDCP method results in the over-enhancement of colors in specific regions, while the UW-CycleGAN method suffers from excessive color amplification, causing color distortion. In contrast, the proposed method, specifically optimized for underwater conditions, effectively enhances text clarity, preserves image details and contrast, and improves color fidelity, while minimizing over-enhancement artifacts (**Figure 3**).

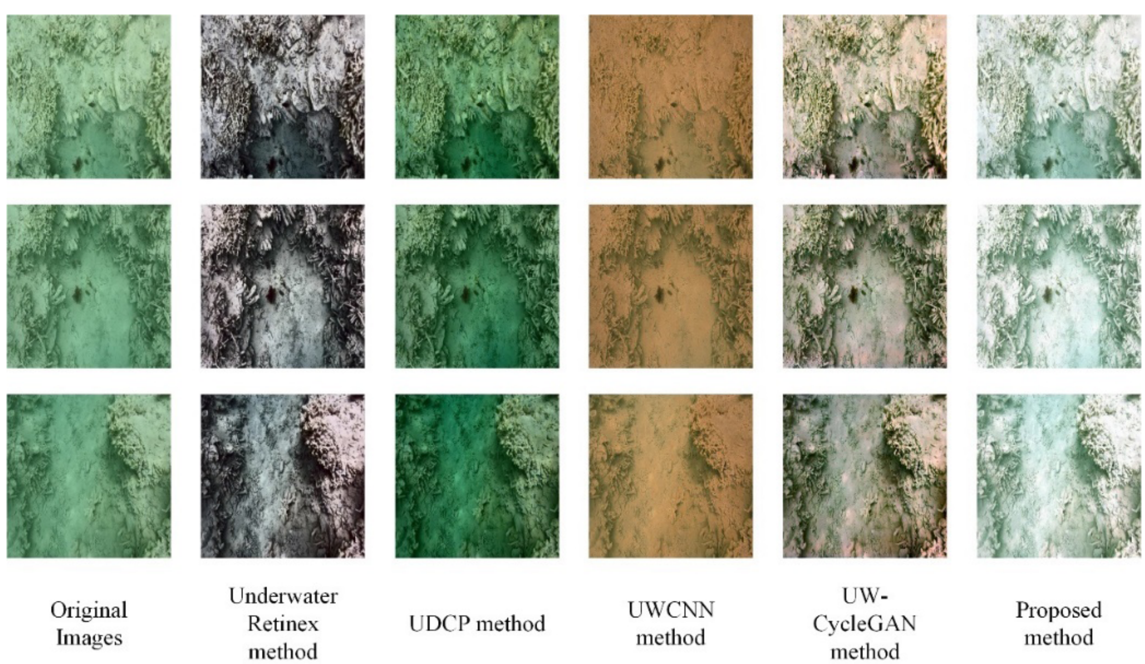

Figure 3. Images enhancement results of the proposed method compared to other methods

In the Xisha underwater trench, the optical properties of light absorption and scattering result in significant color distortions, predominantly characterized by a strong cyan-green tint, which substantially degrades color fidelity and visual clarity. Various enhancement methods exhibit distinct advantages and limitations in addressing this issue. The UWCNN method applies excessive compensation to the red channel, leading to oversaturated red regions. The UDCP method overcorrects the green channel, causing the image to appear unnaturally greenish. The Retinex method over-enhances dark regions, resulting in underexposure and loss of fine details. Similarly, the UW-CycleGAN method suffers from overcompensated red hues, distorting the overall color balance. By comparison, the proposed method demonstrates superior performance in color correction and image enhancement, effectively alleviating underwater color distortions, improving overall visual quality, and achieving a more natural and realistic restoration.

To quantitativelyw assess image quality, UIQM, UCIQE, and PCQI values were computed for the five selected images. The results are summarized in **Table 1**.

Table 1. Performance of evaluation metrics for different image processing methods

Images	Methods	UIQM	UCIQE	PCQI
1	Retinex	7.5709	24.2249	0.4744
	UDCP	6.3653	33.6903	0.5135
	UWCNN	1.6955	18.9805	0.3197
	UW-CycleGAN	4.9672	32.6752	0.4904
	Ours	7.4270	33.9983	0.6329
2	Retinex	7.3989	23.7096	0.5133
	UDCP	4.3207	21.5981	0.4306
	UWCNN	1.7468	17.1665	0.3547
	UW-CycleGAN	4.8150	21.5271	0.5017
	Ours	7.4704	24.4932	0.6250
3	Retinex	7.6713	26.0896	0.4855
	UDCP	3.2267	32.9838	0.5318
	UWCNN	6.4588	15.8973	0.4200
	UW-CycleGAN	7.0313	31.6088	0.7346
	Ours	7.7864	33.0733	0.7470
4	Retinex	7.5230	25.0243	0.5369
	UDCP	4.0981	24.9454	0.617
	UWCNN	5.2463	15.2347	0.4844
	UW-CycleGAN	7.7680	25.7651	0.7790
	Ours	7.8279	25.0277	0.7857
5	Retinex	4.8094	26.9719	0.5062
	UDCP	3.9349	25.3797	0.5339
	UWCNN	5.3996	16.5229	0.4460
	UW-CycleGAN	8.3393	30.7937	0.7238
	Ours	8.7182	30.4943	0.7119

The experimental results reveal significant differences in evaluation metrics among the tested image processing methods. Notably, the proposed method consistently outperforms others across all metrics. In particular, for the first test image, the Retinex method achieves a relatively high UIQM score, indicating its effectiveness in enhancing image details and overall visual quality. However, its lower performance in UCIQE and PCQI suggests limitations in color restoration and contrast optimization, highlighting areas for further improvement. Meanwhile, the UW-CycleGAN method exhibits distinctive performance trends in the fourth and fifth test images. It attains higher scores in UCIQE and PCQI, demonstrating strong capabilities in color enhancement and perceptual contrast improvement. Nevertheless, its relatively low UIQM scores expose deficiencies in overall image quality enhancement and detail preservation, indicating challenges in achieving a well-balanced and comprehensive enhancement.

4. Conclusion and future directions

This study presents LACC-RCE, a novel underwater image enhancement framework that combines LACC with adaptive Rayleigh distribution stretching and CLAHE-based contrast enhancement. Experimental results indicate that the method exhibits strong adaptability in color correction, contrast enhancement, and detail preservation, leading to a notable improvement in the visual quality of underwater images. Additionally, in terms of quantitative evaluation metrics, the proposed approach demonstrates consistent and superior performance across diverse underwater environments. Nevertheless, under extreme conditions, such as high turbidity or low-light environments, the proposed method may still exhibit insufficient enhancement in certain local regions or color shifts, suggesting that its performance in highly complex underwater scenarios requires further improvement. Future research can focus on the following directions, Refining color correction and contrast enhancement strategies to improve the method's adaptability and robustness in varying underwater conditions. Incorporating deep learning approaches to explore data-driven enhancement techniques, thereby improving the model's generalization capability across different underwater environments.

Funding

Graduate Student Innovation Projects of Beijing University of Civil Engineering and Architecture (No. PG2024121)

Disclosure statement

The author declares no conflict of interest.

References

[1] He K, Sun J, Tang X, 2011, Single Image Haze Removal Using Dark Channel Prior. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(12): 2341–2353. https://doi.org/10.1109/TPAMI.2010.168

[2] Drews PLJ, Nascimento ER, Botelho SSC, et al., 2016, Underwater Depth Estimation and Image Restoration Based on Single Images. IEEE Computer Graphics and Applications, 36(2): 24–35. https://doi.org/10.1109/MCG.2016.26

[3] Yu H, Li X, Lou Q, et al., 2020, Underwater Image Enhancement Based on DCP and Depth Transmission Map. Multimedia Tools and Applications, 79(27–28): 20373–20390. https://doi.org/10.1007/s11042-020-08701-3