

5. Summary

This paper presents the Dynamic Memristor-Inspired Zeroing Neural Network (DMZNN), a novel model for solving time-varying nonlinear equations in real-time. DMZNN is based on a hybrid activation function inspired by memristors, which combines cubic and sublinear terms to achieve multi-stage error decay and finite-time convergence. The proposed model addresses the challenges of slow convergence and poor robustness seen in traditional Zeroing Neural Networks (ZNN), especially when dealing with noise and perturbations. We rigorously prove the finite-time convergence and robustness of DMZNN using Lyapunov theory. Experimental results show that DMZNN outperforms traditional ZNN models in solving second- to fourth-order time-varying nonlinear equations, with faster convergence. Additionally, DMZNN is successfully applied to remote sensing image fusion tasks, where it significantly improves both fusion quality and processing speed compared to traditional gradient-based methods. This demonstrates DMZNN's potential for real-time, high-efficiency image fusion applications. The contributions of this work are twofold: first, the introduction of a dynamic, memristor-inspired hybrid activation function to improve convergence and robustness; second, the development and application of the DMZNN model to a practical real-world problem—remote sensing image fusion. The results show that DMZNN provides a robust, efficient, and theoretically grounded solution for solving dynamic optimization problems in real-time applications.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor Li, for his invaluable guidance, profound academic insights and consistent support throughout the entire research process.

Funding

This work is jointly supported by the National Natural Science Foundation of China (Nos. 61404049, 62273141), Natural Science Foundation of Hunan Province (Grant No: 2020JJ6031), Key Project of Hunan Provincial Education Department (Grant No: 22A0324), Scientific Research Fund of Education Department of Hunan Province (Grant No: 17B094), Special Program of National Innovative City Construction of Xiangtan (Grant No: NY-YB20221042).

Disclosure statement

The author declares no conflict of interest.

Author contributions

Jiaqi He: Investigation, Methodology, Soft-ware, Writing – original draft. Mu Li: Funding acquisition, Resources, Supervision, Writing – review & editing.

References

- [1] Jie, J. (2021). An improved finite time convergence recurrent neural network with application to time-varying linear

complex matrix equation solution. *Neural Processing Letters*, 53(1), 777-786. <https://doi.org/10.1007/s11063-021-10426-9>

- [2] Wu, W., Zhang, Y., & Tan, N. (2024). Adaptive ZNN model and solvers for tackling temporally variant quadratic program with applications. *IEEE Transactions on Industrial Informatics*. <https://doi.org/10.1109/TII.2024.3431046>
- [3] Chen, J., Pan, Y., & Zhang, Y. (2024). ZNN continuous model and discrete algorithm for temporally variant optimization with nonlinear equation constraints via novel TD formula. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 54(7), 3994-4004. <https://doi.org/10.1109/TSMC.2024.3374754>
- [4] Tang, C., & Ding, Y. (2025). A New Parameter-Changing Integral ZNN Model with Nonlinear Activation for Solving Inequality Constraint Time-Varying Quadratic Programming. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2025.3548004>
- [5] Zhang, Y., Wang, L., & Zhong, G. (2025). Design and analysis of a variable-parameter noise-tolerant ZNN for solving time-variant nonlinear equations and applications. *Applied Intelligence*, 55(6), 460. <https://doi.org/10.1007/s10489-025-06304-9>
- [6] Fu, Z., Zhang, Y., & Li, W. (2024). Solving future nonlinear equation system via ZNN and novel general ILR3S formula with multitype manipulator applications. *IEEE Transactions on Industrial Electronics*, 71(10), 12623-12633. <https://doi.org/10.1109/TIE.2024.3357867>
- [7] Liao, B., Xu, J., Hua, C., Wang, T., & Li, S. (2025). Predefined-time ZNN model with noise reduction for solving quadratic programming and its application to binary assignment problem in logistics: B. Liao et al. *The Journal of Supercomputing*, 81(12), 1228. <https://doi.org/10.1007/s11227-025-07694-w>
- [8] Jie, J., Zhu, J., Gong, J., & Chen, W. (2022). Novel activation functions-based ZNN models for fixed-time solving dynamic Sylvester equation. *Neural Computing and Applications*, 34(17), 14297-14315. <https://doi.org/10.1007/s00521-022-06905-2>
- [9] Xiao, L., Li, X., Cao, P., He, Y., Tang, W., Li, J., & Wang, Y. (2023). A dynamic-varying parameter enhanced ZNN model for solving time-varying complex-valued tensor inversion with its application to image encryption. *IEEE Transactions on Neural Networks and Learning Systems*, 35(10), 13681-13690. <https://doi.org/10.1109/TNNLS.2023.3270563>
- [10] Li, J., Zhu, J., Li, C., Chen, X., & Yang, B. (2022). CGTF: Convolution-guided transformer for infrared and visible image fusion. *IEEE Transactions on Instrumentation and Measurement*, 71, 1-14. <https://doi.org/10.1109/TIM.2022.3175055>

Publisher's note

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

Design and Implementation of Machine Learning-based Monitoring System for Mineral Processing Flotation Reagent

Yiming Yao*, Yi Li

NAURA Technology Group Co., Ltd., Beijing, China

**Author to whom correspondence should be addressed.*

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Flotation, also known as froth flotation, is a method for separating minerals from powdered materials by altering their floatability through the use of flotation reagents. This paper proposes a flotation process control system for mineral processing based on machine learning. Addressing the issue of lack of precise detection methods in the flotation process of iron concentrate, a neural network regression method is used to predict the amount of reagents and the grade of the flotation concentrate. The flotation data in this paper come from the Key Laboratory of Multitechnology Resource Utilization of Bayan Obo Mine, Inner Mongolia Autonomous Region. The preprocessed data form the dataset used to create the production prediction model. The neural network model is constructed using the PyTorch deep learning framework. Finally, based on the established model, a comprehensive flotation dosing monitoring system is developed using the Django framework, which includes functions such as production indicator large screens, workshop personnel safety monitoring large screens, flotation reagent usage processing, flotation reagent procurement platform.

Keywords: Froth flotation; Neural network; Monitoring platform; Machine learning

Online publication: February 12, 2026

1. Introduction

The objective of mineral processing is to remove a large amount of gangue and harmful elements contained in the ore, thereby enriching the valuable minerals and separating coexisting useful minerals from each other to obtain one or more useful concentrate products. Froth flotation (flotation) is a method in which mineral powders are treated with flotation reagents to alter their floatability, allowing the targeted minerals to selectively adhere to bubbles, thus enabling the selection of the desired minerals. Flotation occupies a predominant position among various mineral processing methods and has a wide range of applications. It can process not only non-ferrous metal minerals (such as copper, lead, zinc, molybdenum, cobalt, tungsten, antimony ores, etc.) but also non-metallic minerals (such as graphite, barite, fluorite, apatite, feldspar, talc, etc.) and ferrous metal minerals (such as hematite, manganese, and

titanium ores, etc.). Compared to other mineral processing methods, flotation has a high separation efficiency, capable of upgrading low-grade raw ores into high-grade concentrates, thus expanding the range of mineral resources. Flotation is particularly effective in treating finely disseminated ores, solving the recovery of valuable components in many fine mineral particles. Flotation automation is the new trend in mineral processing technology. Traditional flotation operations often rely on manual dosing, which results in inaccurate and untimely addition of reagents. Overdosing can lead to raw material waste and environmental pollution, while underdosing can result in insufficient reaction between reagents and minerals, reducing the yield of the final product^[1-4].

1.1. Research background and significance

This project is a collaborative effort between the Cloud Computing and Big Data Laboratory of the School of Computer Science and Engineering at Dalian Minzu University and Baoshan Mining Company. The relevant data parameters for the project are sourced from Baoshan Mining Company's "Inner Mongolia Ore Flotation Desulfurization Experimental Study" project. Traditionally, mineral flotation processing has relied on manual control for related operations, requiring frontline operators to manually adjust reagent dosages and flotation tank liquid levels.^[5] However, this method has consistently faced issues of inaccuracy and untimeliness, sometimes failing to meet the technical indicators required by flotation processes. Errors in this process can also significantly impact production safety. Additionally, the production process involves the use of many highly toxic processing reagents, such as cyanides. Traditional monitoring methods require substantial human labor, such as deploying monitoring rooms and dispatch rooms to oversee personnel and vehicle movements in high-risk areas. This not only wastes human resources but also poses safety risks due to potential lapses in human supervision. Therefore, this paper integrates artificial intelligence and other computer technologies into the production process to enhance mineral processing indicators and ensure production safety.

1.2. Figures and tables

In recent years, with the continuous advancement of automation technology, the level of flotation automation and the addition of flotation reagents have been continuously improving. In the production process of adding flotation reagents, relying solely on manual dosing makes it difficult to accurately control the dosage and timely addition of reagents. The automated application of flotation reagents not only overcomes the shortcomings of manual adjustment but also reduces reagent consumption during the flotation process^[6].

In 2001, Xu Deping, Wu Cuiping, and others conducted an analysis and research on the beneficiation parameters in coal slime beneficiation. They integrated computer software design and electromechanical control systems to develop a coal slime flotation pulp level meter and a coal slurry ash meter. This system achieves control over the flotation process by measuring the ash content of coal slime and subsequent data, using fuzzy control technology to adjust the flotation reagent dosage and flotation machine pulp level^[7-9].

In 2022, Xu Zewei from Fenxi Mining South Coal Industry proposed solutions using sensor technology and PLC control technology to address the high control difficulty and low operational efficiency in the flotation process of premium coal. By using pressure sensors (liquid level height detection), slurry concentration sensors, ultrasonic flow meters, electromagnetic flow meters, and tailings ash detection devices to monitor key indicators in the coal beneficiation process, automated control of the coal beneficiation process was achieved^[10].

In 2024, Zhang Hongchang, Mou Song, and other technical personnel used an engineering dosing system and DCS control system for data communication, achieving information interaction and detection between the

scheduling department and data communication. This made the process of machine-assisted automated dosing more stable, bringing more direct benefits to the company.

2. Mechanism and process of mineral flotation

Ore is one of the crucial raw materials in the metal manufacturing industry. Mineral processing aims to maximize the separation of valuable minerals from gangue minerals to obtain high-grade concentrates and to recover coexisting valuable minerals as separate concentrates for their useful components. Among the modern mineral processing techniques, flotation and magnetic separation are the mainstream methods for mineral beneficiation.

2.1. Definition and calculation of ore grade

Ore grade refers to the amount of a specific metal or useful component contained within the ore, typically expressed as a percentage. For precious metals (such as gold, silver, or gemstones), the grade is expressed in g/t or g/m³. The grade of the ore is determined by sampling and assay results. Depending on the properties of the ore, it can be classified into raw ore grade, concentrate grade, and tailings grade. The raw ore grade (denoted as α) represents the percentage of metal content in the raw ore that enters the processing plant. It is an indicator of raw ore quality and a fundamental data point for the metal balance of the processing plant. The concentrate grade (denoted as β) indicates the percentage of metal content in the concentrate, reflecting the quality of the concentrate. The tailings grade (denoted as θ) represents the percentage of metal content in the tailings, indicating the metal loss during the beneficiation process.

Processing plants typically assay the concentrate grade once every shift (12 hours). The concentrate yield and grade of each shift vary. Let the yields of three shifts be Q_1 , Q_2 , and Q_3 and their respective grades be β_1 , β_2 and β_3 . The average grade of the three shifts can be calculated as follows:

$$\bar{\beta} = \frac{Q_1\beta_1 + Q_2\beta_2 + Q_3\beta_3}{Q_1 + Q_2 + Q_3} \times 100\%$$

In the statistical reports of the processing plant, there is a cumulative grade section. The cumulative grade is calculated similarly to the average grade but is a cumulative process:

$$\bar{\beta}_i = \frac{Q_{i-1} \times \beta_{i-1} + q_i \times \beta_i, \text{ day}}{Q_i}$$

2.2. Basic concepts of flotation

Flotation, fully known as froth flotation, is a mineral processing method that relies on the differences in the chemical properties of mineral particle surfaces. By utilizing the buoyancy of bubbles in the slurry, it achieves the separation of minerals. Modern flotation operations mainly include four processes: grinding, slurry conditioning with reagents, flotation separation, and product handling. The froth product and tailings from flotation undergo dewatering separation, as shown in **Figure 2.1**.

Compared to other mineral processing methods, flotation has higher separation efficiency. It can upgrade low-grade ores into high-grade concentrates, expanding the utilization scope of minerals. This allows for the development and utilization of low-grade deposits that were previously considered unexploitable into high-quality

deposits. Flotation is particularly effective for processing fine and ultra-fine minerals. Due to the fine particle size and minimal density differences, gravity separation is difficult for these minerals. However, by adjusting with reagents and mechanical methods based on the different surface activities of the minerals, flotation can effectively separate valuable minerals from waste materials.

3. System requirements analysis

3.1. Functional requirements

The primary function of the mineral processing flotation reagent dosing system is to monitor and analyze the indicators of the mineral processing flotation process, predict the grade of the concentrate, and determine the optimal reagent dosage. The system also visualizes the collected data, displaying it on the front-end interface using alert notifications and statistical charts to guide flotation workshop operators in adjusting reagent dosages. Given that flotation reagents are generally toxic, the factory imposes strict standards on protective labor measures for production personnel and the management of vehicles and personnel in hazardous areas. The system processes video information received from cameras frame by frame, performs image recognition on the footage, uploads detected hazardous information to the database, and displays it on the corresponding front-end interface.

(1) Production indicator display screen

The data display module aggregates the on-site technical indicators from the database onto the data display screen. On this screen, one can view the status of reagent dosages, the addition of reagents, the predictions of concentrate and tailings grades, and the types of reagents currently in use. By visualizing the data from the database for on-site operators and displaying the estimated reagent dosages for the next stage, the system can guide the reagent dosing process.

(2) Workshop personnel safety monitoring display screen

The workshop personnel safety monitoring dashboard primarily publicizes potential safety issues detected through image recognition of video data collected by cameras. This allows factory safety officers to review and handle these issues. When a violation occurs, the system sends an alert to the safety officer's personal back-end interface. For false alarms, the safety officer can dismiss the anomaly alert through back-end management.

(3) Reagent dosage handling

The reagent dosage handling module allows technical operators to view the status of flotation processing equipment and the current status of reagent tanks and mixing tanks, and to increase or decrease the amount of reagent added. The system provides relevant auxiliary measures, issuing corresponding warnings in the case of excessive ore feeding or reagent addition. If a technical operator performs an incorrect operation, the system will not execute the command, issue a warning, and freeze the member's operation privileges for one day.

3.2. Performance requirements

The system will integrate control systems, data processing, and analysis tools into the industrial production environment to achieve intelligent, networked, and automated management and operations. Since the system needs to interact with users, user interface operations (such as clicks, queries, etc.) should respond within 1 second, and the generation time for complex charts and reports should not exceed 3 seconds. The

system can adopt redundancy storage and backup strategies to ensure that data is not lost in the event of any single point of failure, guaranteeing 99.999% data reliability during data storage and transmission. During peak production periods, the system supports concurrent use by 2000 users, ensuring a data processing capacity of at least 2000 transactions per second. Additionally, the system employs a modular design to facilitate the integration and deployment of new modules.

3.3. System overview design

In the flotation process of iron concentrate, multiple factors affect the final grade of the minerals, making it a highly coupled process. Designing an intelligent reagent control system based on machine vision requires considering the strong coupling and internal mechanism complexity of multiple factors in the coal slime flotation process. Therefore, formulating a reasonable and effective technical route must adhere to design principles and meet the requirements of existing iron concentrate flotation processing technology.

3.3.1. Technical route of system design

The system designed in this paper is developed based on the beneficiation plant of Baotou Steel Co., Ltd. After two months of observation and learning in frontline production at the beneficiation plant, an in-depth analysis was conducted on the issues encountered during production. By integrating software design technology with mineral processing technology, the pain points and difficulties in the beneficiation process are addressed using relevant computer methods. Feasibility analysis, scheme research, and scheme demonstration are employed to explore the problems encountered in the project.

During the system design process, it is first necessary to review relevant literature and data on mineral flotation and collect relevant experimental data from the frontline processing site. The collected data, such as slurry concentration, mineral particle size, and reagent concentration, are processed. The processed data is used as the dataset for training, and the dataset is subjected to missing value handling and feature scaling. The dataset is divided into training and testing sets, and elastic net regression is used to train the dataset. The results of the trained model are then evaluated and tuned.

Simultaneously, photos of the factory are collected and taken, and objects such as vehicles appearing in the real-time status of the photos are marked. The YOLOv5 model is used to train the labeled data. After training, metrics such as Mean Average Precision (mAP) are used to evaluate the accuracy and performance of the model. After performance tuning, the model is applied to detect objects in new images using the YOLOv5 Python interface for object detection.

3.3.2. System architecture

The medication monitoring system utilizes the Django framework for the separation of front-end and back-end development. Django achieves this by creating different apps, writing data to these respective apps, and creating various view functions. This allows the Python back-end to provide JSON data responses to the front-end. Django modifies the traditional MVC framework by dividing the view into the View module and the Template module, with the two modules respectively responsible for dynamic logic processing and static page data. The Django framework uses a template engine to transmit data from the view to the client. It is necessary to specify the directory where the template files are stored in the project configuration file. In each app, the view's rendering functions based on Django need to be edited to bind the amount of data to be displayed. Corresponding template

files need to be created in the template, and data transmitted from the view functions are received according to the corresponding template syntax of the template engine. The Model in the application handles data logic, encapsulating business data related to business logic and methods for processing data. Simply put, it is the interaction layer between the web framework and the database. Django’s model layer adopts ORM (Object-Relational Mapping) technology, which converts tables in the database into abstract classes, making it convenient for view functions to call them.

The system adopts a B/S architecture and is deployed on Huawei Cloud servers using Nginx and Uwsgi. The Template module of the system employs HTML, CSS, and JavaScript in a static mode for front-end UI development. The view then receives requests from the user interface, invokes the appropriate processing logic, retrieves data from the model, and passes it to the template for rendering a response. This setup enables the system deployed on Huawei Cloud to interact with the Alibaba Cloud database. In terms of system security, Django’s built-in security measures such as authentication, authorization, and Cross-Site Request Forgery (CSRF) protection are utilized to safeguard user production data and information. For factory safety monitoring and data processing, trained models are encapsulated within view functions to predict and process data from the production site, with the predicted results then uploaded to the front-end interface.

3.3.3. System function modules

The system enables the adjustment of the flow rate and status of each dosing port. Starting from the principles and processes of flotation, the system analyzes key influencing factors in flotation and detects some of these critical factors. It achieves real-time data collection and historical data storage of flotation process parameters. Operators can upload accurately tested data. Engineering technicians can compare computer-predicted values with uploaded data to optimize algorithms and improve mineral processing techniques. The system can present data from the database in the form of pie charts or linear graphs for flotation site, workshop safety, reagent dosage, and sales detection. Based on the data obtained from the system’s dashboard, engineering technicians can adjust production processes according to relevant data indicators. The system function modules are shown in **Figure 1**.

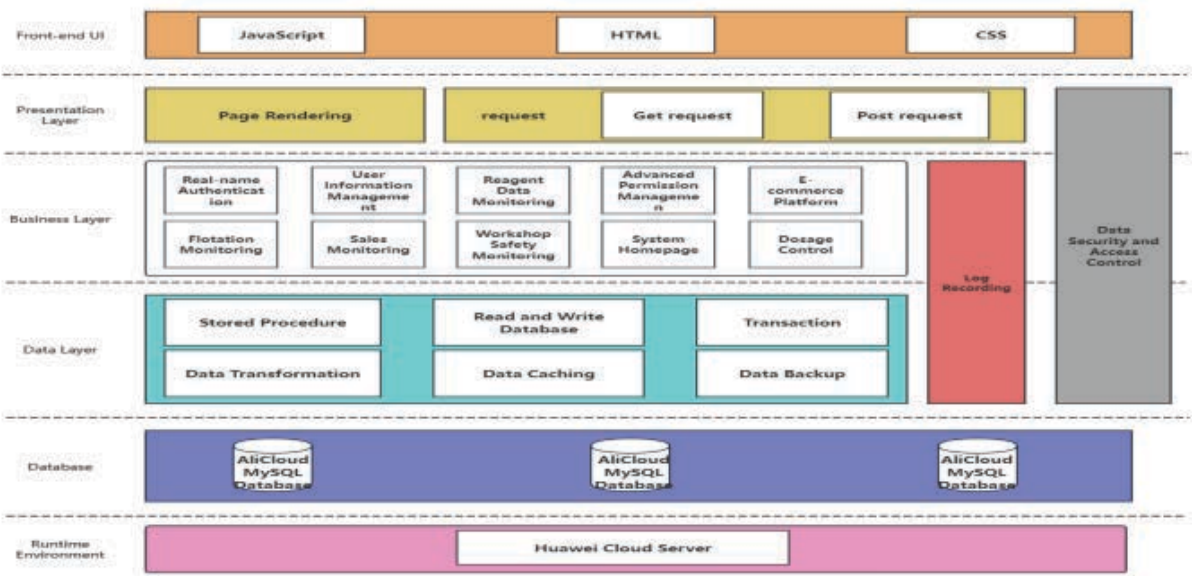


Figure 1. System architecture diagram.

3.4. Data preprocessing

3.4.1. Introduction to the dataset

In developing the prediction model for the flotation concentrate grade and the optimal reagent dosage, the dataset used for training is crucial in the early stages. However, the dataset required for the development of this system was not available on the public dataset platform of the Great Interconnection Company. The lack of a dataset has become a key issue in the development of this system. Therefore, due to the dataset's absence, this paper adopts a manual collection method to obtain a dataset suitable for training the predictable grade of the flotation concentrate and the optimal reagent dosage, including the relationship data of concentrate and tail iron, sulfur grade, and mineral yield.

(1) Flotation dataset

The dataset concerning the relationship between the iron and sulfur content in the tailings and the mineral yield (hereinafter referred to as the flotation dataset) was generously supported by the Inner Mongolia University of Science and Technology and the Key Laboratory of Multiple Utilization of Mineral Resources in Baiyunebo Mine, Inner Mongolia Autonomous Region. The flotation dataset is divided into three parts. The first part includes the state values during different mineral flotation periods, including the grades of iron and sulfur in the slurry, concentrate, and tailings. The second part includes the concentration of xanthate and the concentration of sodium hexafluorosilicate added during flotation. The third part includes the flotation concentration and the final yield. The total amount of the flotation dataset is 1000 entries, with 70% used as the training set and 30% used as the test set. The format is the commonly used .csv file in machine learning. CSV (Comma-Separated Values) files are lightweight text files that use plain text, with each line representing a data record and different attributes separated by commas.

(2) Safety warning dataset

In this system, safety warnings primarily involve two aspects: first, employees in the workshop must wear safety helmets to ensure personal safety during work; second, vehicle entry management within the factory premises is stringent, requiring all vehicles to be strictly monitored. Therefore, the safety warning dataset comprises four types: safety helmets, heads, people, and vehicles. The datasets for safety helmets, heads, and people are sourced from the publicly available dataset "Safety Helmet Detection" on Alibaba Tianchi, while the vehicle dataset is sourced from the "Car License Plate Detection" dataset on Alibaba Tianchi. The annotated dataset is in the standard YOLO format, with a total of 9570 images.

3.4.2. Data preprocessing

(1) Integration and cleaning of floating dataset

The flotation data in this study is sourced from the Key Laboratory of Multi-technology Utilization of Baiyun Obo Mine, Inner Mongolia University of Science and Technology. Initially, the data was stored in separate files according to different record dates. However, for machine learning purposes, the data needs to be integrated into a single dataset. This can be achieved using EXCEL to combine the individual files into a comprehensive data.csv file. The dataset also contains some missing items and outliers due to data collection issues. These anomalies can affect the feature extraction process in machine learning. In this study, the merged data.csv file is read using the read_csv function from the Pandas library in Python. Missing and anomalous values are replaced with the mean value of the corresponding column to mitigate their impact.

(2) Annotation of safety warning datasets

Due to issues such as incomplete annotation information, unclear division between training and testing sets, and inconsistent annotation formats in the acquired datasets, manual re-annotation was performed using LabelImg during the development of the safety warning model. LabelImg is a graphical image annotation tool written in Python and using Qt for its graphical interface. It supports the YOLO annotation format. The annotation process involves the following steps: first, defining the list of classes used in training in `data/predefined_classes.txt`; second, opening the directory and creating a “RectBox” to construct a rectangular box around the selected area; and finally, selecting the class name from the label list to complete the annotation.

(3) Image loading

In deep learning, image preprocessing is a crucial step that can directly affect the performance of the trained model. The preprocessing of the safety warning dataset includes normalization, denoising, and grayscale conversion. This paper utilizes the OpenCV library in Python, which provides many tools for image processing. The pre-divided safety warning dataset is loaded into the program and then loaded as a Numpy array called “image”. According to the YOLO dataset, each image is scaled based on its corresponding category, the center position of the target object, and the length and width of the target object. Subsequently, the target object’s area is cropped using array indexing. Each “image” array is normalized to a range between 0 and 1 by dividing by 255.

(4) Image grayscale

Grayscale the loaded images results in images that contain only luminance information, which can reduce storage space and computational complexity. Color images typically consist of three channels: red, green, and blue (RGB channels). Each channel corresponds to a specific color, but grayscale images have only one channel. Therefore, the pixel values of the three channels in a color image need to be converted to a single grayscale value. OpenCV uses a weighted average method for grayscale, based on psychological experiments that determine the human eye’s sensitivity to different colors. The weights for red, green, and blue are used to calculate the grayscale value through weighted averaging. The specific calculation formula is as follows:

$$Grayscale = 0.299Red + 0.578Green + 0.114Blue \quad (1)$$

(5) Image deionizing

Denoising can reduce the interference of noise in subsequent processing, improve the accuracy and stability of algorithms, and enhance the features within images, making subsequent feature extraction tasks more accurate and reliable. In this paper, the image denoising process employs the Gaussian Blur function. Gaussian Blur is a commonly used image blurring technique that smooths an image by applying a Gaussian filter, thereby reducing noise and fine details within the image. The Gaussian Blur utilizes the Gaussian function as a weighting function^[11]. Due to the symmetry of the Gaussian function in the spatial domain, it effectively reduces high-frequency noise in the image. The weight calculation of the Gaussian filter is based on the distance of each pixel within the window, where pixels farther from the center have less influence on the output, while pixels closer to the center have a greater influence. Consequently, Gaussian Blur smooths the image and reduces noise.

4. Data preprocessing models

4.1. Yield prediction model based on neural network regression

(1) Multivariate regression performance

There are two commonly used models for traditional machine learning regression problems: the Linear Model and the Polynomial Model. These models perform well when the data exhibits linear relationships and there is no correlation between features. However, in the mineral flotation process, factors such as grinding particle size, aeration rate, stirring speed, slurry concentration, and pH value affect the process. The dataset encompasses numerous essential attributes and influencing parameters, which can lead to an increase in the polynomial degree and the number of terms in the regression model. This, in turn, affects the development efficiency and the quality of the model. The following is scatter plot 6.1 of the flotation dataset, generated using multivariate function fitting on the SPSSRO platform (an online data analysis platform). It is evident from the visualization that the fitting results are suboptimal due to the complexity of the data.

$$\hat{y} = b_0 + b_1x + b_2x^2 + \dots + b_mx^m$$

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_1^2 + b_4x_2^2 + b_5x_1x_2$$

(2) Introduction to neural network regression

Neural Network Regression (Quantile Regression Neural Network, QRNN), proposed by Taylor, is a non-parametric and non-linear regression algorithm that combines the advantages of both neural networks and regression ^[12]. See **Figure 2**.

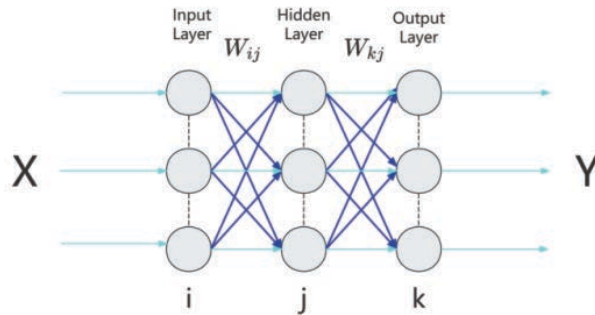


Figure 2. Neural network model diagram.

In a single-layer network, let's assume the input is (X), the output of this layer is (A), and the final output is (Y). Thus, in this single-layer network, this study has:

$$Y=A=\sigma(W^T X+b)$$

Assuming that $(Z=W^T X+b)$ is a linear process, and the activation function (σ) is generally non-linear, the overall model is therefore a non-linear function. The parameters (w) and (b) that the model needs to learn can be collectively referred to as (θ). In the above equation, each quantity is generally a matrix. Assuming the shape of (X) is $((n, m))$, where (n) represents the sample length and (m) represents the number of samples, the shape of (w) would be $((n, 1))$, and (b), (Z), (A), and (Y) would be $((1, m))$, representing the output values of (m) samples. For regression problems, the activation function (σ) commonly used is the ReLU (Rectified Linear Unit) function, defined as:

$$A=\sigma(Z)=\max(0,Z)$$

The accuracy of the neural network's predictions for a problem is measured by a loss function. For regression

problems, the commonly used loss functions are Mean Squared Error (MSE) and Mean Absolute Error (MAE). MSE is suitable for data with a more concentrated distribution, so this paper uses the MSE function:

$$\text{Loss} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

After computing (Y) through forward propagation, backpropagation is needed to compute the partial derivatives of the cost function (J) with respect to the parameters at each layer of the neural network, and use these derivatives to adjust the values of the network parameters. According to the Gradient Descent method, this study has the optimization function:

$$\text{Optimizer } (\theta) = \theta - \alpha \frac{\partial J}{\partial \theta} = \theta - \alpha d\theta \text{ (}\alpha \text{ is the learning rate)}$$

Using this formula, the parameter (θ) will change in the direction of the steepest descent of (J) during each backpropagation process. Thus, the forward and backward propagation process is essentially training the neural network. In this process, the model's parameters (θ) fit the sample distribution, which is referred to as learning.

(3) Implementation of yield forecasting model

Based on the completion of data preprocessing, a neural network model is constructed. This paper's neural network model is implemented using the PyTorch deep learning framework. In this model, there is one input layer, three hidden layers, and one output layer. The concentrations of the yellow medicine, collector sodium hexafluoride, original ore iron content, and flotation concentration are aggregated into a two-dimensional matrix and then converted into an input feature matrix, which serves as the parameters of the input layer. The output layer's results are then propagated forward through a series of linear transformations and activation functions into the subsequent three hidden layers. During this process, each neuron calculates the weighted input and applies an activation function to produce an output, which is then passed through the network. Upon reaching the output layer, the neurons of the last layer provide the predicted output. The MSE function is then used to calculate the loss function, which measures the difference from the true values. The network's backpropagation is performed based on the value of the loss function, while the contribution of each parameter in the model to the loss is calculated according to the chain rule. The gradient descent algorithm is then used to adjust the model parameters. Through multiple iterations, the model gradually converges and improves its accuracy.

(4) Evaluation of the yield prediction model's performance

The yield prediction model primarily addresses the influencing factors of mineral flotation and the yield of iron ore, with data presented in tabular form. This experiment evaluated the impact of flotation concentration, temperature, pH value, and particle size on the flotation process. First, scatter plots of the influencing factors and yield were drawn. Using the parameters obtained from the trained model, the model graph was plotted, and the accuracy of the resulting data reached 94.76%, as shown in **Figure 3**.

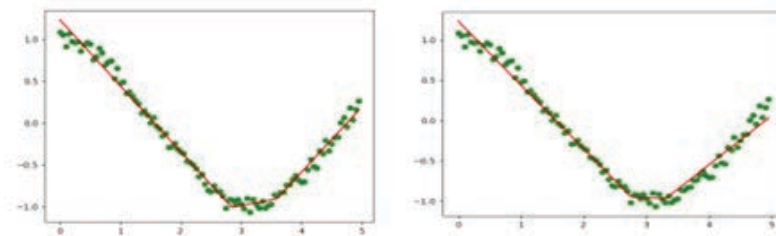


Figure 3. Model fitting effect diagram.

4.2. Factory safety monitoring model based on YOLOv5

4.2.1. Introduction to YOLOv5 model

YOLOv5 is an object detection algorithm in the realm of computer deep learning. It builds upon and optimizes the YOLO (You Only Look Once) series. As shown in Figure 8, YOLOv5 mainly comprises four components: Input, Backbone, Neck, and Prediction.

The primary task of the input component is to correct coordinates based on the width and height of all images in the dataset before each training session. It uses the k-means algorithm to cluster detection boxes in the training set to obtain initial anchors, and then applies a genetic algorithm to mutate these anchors to achieve the optimal anchor boxes. YOLOv5 also incorporates Mosaic data augmentation, which combines and stitches multiple images to produce new images, thus enhancing the effectiveness of the training set. See **Figure 4**.

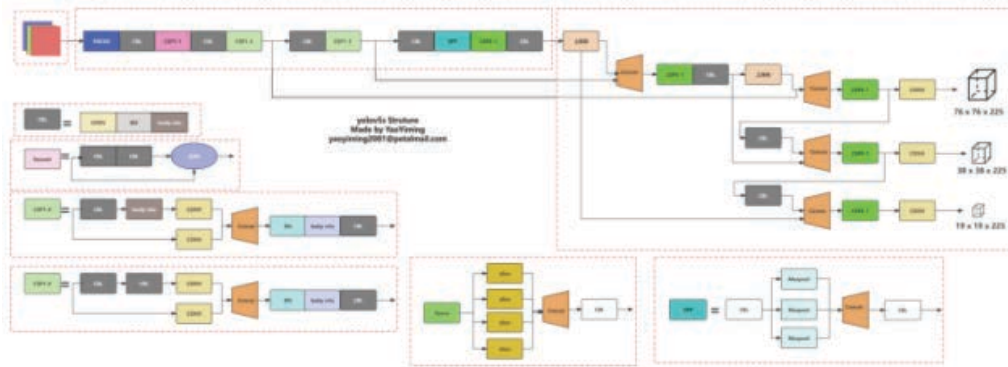


Figure 4.YOLOv5 network architecture.

The main function of the Backbone is to extract features and progressively reduce them. The Backbone is primarily composed of the Focus and CSP structures. The Focus structure slices the image, taking one value every pixel, thereby decomposing the entire image into four independent feature layers and concentrating the width and height information in the channel space. The original RGB three channels are thus transformed into 12 channels^[13]. After applying convolution operations, sampled feature maps are obtained. The CSP structure divides the obtained feature map into two parts: one part undergoes convolution operations, while the other part is cross-connected with the convolution structure of the first part, enabling the model to capture more features. In object detection tasks, the Backbone utilizes CSPNet, which significantly enhances the learning capability of CNNs while also reducing computational cost.

The Neck structure of YOLOv5 aims to improve the network's feature fusion ability by converting feature maps of varying sizes into fixed-size feature vectors using the SPP pooling structure. Since the positions and sizes of objects in the original images are not fixed, YOLOv5 employs an FPN mechanism to add feature layers in the Backbone, generating feature maps with multi-scale information to improve the accuracy of object detection.

The Head is primarily used in the post-processing stage of object detection to filter the bounding boxes. During object detection, redundant candidate boxes and overlapping bounding boxes may be generated. YOLOv5 utilizes the Non-Maximum Suppression (NMS) algorithm to search for local maxima and suppress non-maximal elements, thereby selecting the optimal bounding box.

4.2.2. Implementation of factory safety monitoring model

YOLOv5 offers four different model structures: Yolov5s, Yolov5m, Yolov5l, and Yolov5x, with Yolov5s being

the smallest model. In ascending order, YOLOv5x is the largest model. Larger models offer better performance and stability, but they also require more time and computational resources for training. After comprehensive consideration, this paper adopts the YOLOv5m model. Model configuration parameters are stored in a train.yaml file, with the number of classes (nc) set to 4, batch size (batch_size) set to 16, learning rate set to 0.001, number of epochs set to 200, and input image size (img_size) set to 416. The loss function weights are set as follows: loss: xy = 1.0, wh = 1.0, cls = 1.0, obj = 1.0, l1 = 0.1. After setting these parameters, the safety alert dataset is imported into the library to start training. Once training is completed, the trained model is evaluated using the validation set. The accuracy of the model is determined by calculating the difference between the predicted results and the actual annotations. These steps are iteratively repeated to improve the model's performance.

For safety alerts in factories, the model often receives frames from video surveillance. However, the YOLOv5 model processes images. In this paper, the Python OpenCV2 library is used, and the default camera device is opened using cv2.VideoCapture(0). Then, the cap.read method is continuously called to read the frames captured by the camera one by one, and each frame is passed to the input end of the YOLOv5 model.

5. Conclusion

5.1. Work summary

To address the long-standing issue of relying on manual control in flotation processing, which compromises plant safety, this study conducted an in-depth investigation at the Baotou Steel Group. It was found that in traditional mineral flotation processing, operators must manually adjust the reagent dosage and flotation cell levels. Manual operations suffer from a lack of accuracy and timeliness, making it difficult to meet the technical standards of the flotation process. To solve the problems of manual dosing accuracy and timeliness, this project collected images and flotation index data from the Baotou Steel Group's multi-technical resource utilization key laboratory in Inner Mongolia. Using a neural network propagation algorithm, the flotation index data were trained to predict yield.

Additionally, the chemical reagents used in the processing, such as cyanides, traditionally rely on human supervision, which can lead to safety hazards due to supervisory lapses, posing serious risks to production safety. To address the inefficiencies of human supervision and the high-risk nature of factory workers' jobs, this study utilized images from the processing site and the Alibaba Tianchi database. These images were annotated, denoised, and converted to grayscale, and then trained using the YOLOv5m model to analyze images transmitted from cameras. This enables monitoring of on-duty staff safety and factory site access. The machine learning-based mineral processing flotation reagent monitoring system implemented five functional modules: production indicators dashboard, workshop personnel safety monitoring dashboard, flotation reagent procurement platform, user access management, and flotation reagent dosage processing. This system achieved intelligent operation management for the mineral processing flotation plant.

The core advantages of the mineral processing flotation reagent monitoring system lie in its intuitive and precise data visualization, reduction in reagent consumption, and assurance of processing safety. The production index dashboard visually displays the optimal reagent dosage predicted based on industrial site data. Experimental results have proven the application of the new system, showing a 4.84% increase in concentrate yield and a reduction of 0.15 kg in reagent consumption per ton of iron concentrate ore.

Disclosure statement

The authors declare no conflict of interest.

References

- [1] Wang Z, Zhang C, Pan J, et al., 2021, Deep Learning-Based Ash Content Prediction of Coal Flotation Concentrate Using Convolutional Neural Network. *Minerals Engineering*, 174.
- [2] Waldner F, Diakogiannis F, 2020, Deep Learning on Edge: Extracting Field Boundaries from Satellite Images with a Convolutional Neural Network. *Remote Sensing of Environment*, 245: 111741.
- [3] Zarie M, Jahedsaravani A, Massinaei M, 2020, Flotation Froth Image Classification Using Convolutional Neural Networks. *Minerals Engineering*, 155.
- [4] Pu Y, Szmigiel A, Chen J, et al., 2020, FlotationNet: A Hierarchical Deep Learning Network for Froth Flotation Recovery Prediction. *Powder Technology*, 375: 317–326.
- [5] Guo T, Hou Z, Yu J, et al., 2023, Design of an Intelligent Control System for Secondary Flotation. *Coal Processing and Comprehensive Utilization*, 2023(3): 47–50.
- [6] Fan F, 2022, Research and Application of Intelligent Control for Flotation Dosing System, thesis, Taiyuan University of Technology.
- [7] Guo X, Wei L, Yang C, 2020, Research on Ash Content Detection Method of Coal Slurry Flotation Tailings Based on Deep Convolutional Network. *Coal Technology*, 39(2): 144–146.
- [8] Tang M, Zhou C, Pan J, et al., 2017, Research on Flotation Froth Image Processing and Ash Content Prediction Based on LabVIEW. *Coal Technology*, 36(9): 294–296.
- [9] Wei L, 2021, Research on Coal Slurry Flotation Dosing Control System Based on Machine Vision, thesis, China University of Mining and Technology.
- [10] Ren J, 2020, Research on the Application of Automatic Control System in the Flotation Process of Coal Preparation Plants. *Coal and Chemical Industry*, 43(11): 103–105.
- [11] Reinhard E, 2006, High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting. Morgan Kaufmann: 233–234.
- [12] Li G, Huang Q, 2024, Scenario Prediction of Carbon Peaking in the Beijing-Tianjin-Hebei Region Based on the Lasso-GRNN Neural Network Model. *Environmental Science*.
- [13] Ni X, Chen Y, 2024, Establishment of a Drug Procurement Decision Prediction Model by Combining the ARIMA Model and LSTM Neural Network. *Heilongjiang Science*, 15(2): 76–78.

Publisher's note

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

Nasopharyngeal Carcinoma Lesion Recognition Based on Multi-Window Resampling Technology

Xiaoni Zhang¹, Mengfan Yang¹, Supan Wei¹, Xin Zhao²

¹Henan Vocational College of Water Conservancy and Environment, Zhengzhou 450008, Henan, China

²North China University of Water Resources and Electric Power, Zhengzhou 450046, Henan, China

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Accurate deep learning-based detection of nasopharyngeal carcinoma (NPC) magnetic resonance (MR) images is conducive to diagnosis and treatment. These images are characterized by high dimensionality, complex noise interference, and blurred tissue structure boundaries. How to extract key pathological features from massive imaging information and provide quantitative basis for clinical diagnosis remains an important challenge in the current field of medical image processing. This paper uses multi-window fusion technology to map multiple key window information to the pseudo-color space, realizing the integration of multi-dimensional feature information and compensating for the information limitations of single-window imaging. Experiments show that this method can effectively improve model accuracy.

Keywords: Nasopharyngeal carcinoma; Multi-window resampling; Lesion recognition; Medical image processing; Pseudo-color fusion

Online publication: February 12, 2026

1. Introduction

As the core method for NPC diagnosis, medical imaging examination, especially magnetic resonance imaging (MRI), has become the preferred imaging modality for locating primary NPC lesions, evaluating invasion range, and monitoring therapeutic effects due to its advantages of high soft tissue resolution, strong multi-parameter imaging capability, and no radiation damage^[1]. However, MRI data has characteristics such as high dimensionality, complex noise interference, and blurred tissue structure boundaries. How to extract key pathological features from massive imaging information and provide quantitative basis for clinical diagnosis remains an important challenge in the current field of medical image processing^[2,3].

Clinical interpretation of MRI images usually relies on physicians' adjustment of different scanning sequences and window width/window level parameters to highlight specific tissue structures^[4]. **Figure 1** shows NPC MRI images under different windows: T1-weighted imaging (T1WI) can clearly display anatomical structures, while T2-weighted imaging (T2WI) is sensitive to edema and inflammation. However, the traditional single-window imaging mode can only present local grayscale information, making it difficult to simultaneously balance the

contrast difference between tumor tissue and surrounding normal structures ^[5]. Studies have shown that primary NPC lesions often invade the parapharyngeal space, skull base bone, and intracranial structures, and their imaging manifestations are highly heterogeneous. Grayscale images under a single window are prone to boundary information loss or artifact interference, increasing the difficulty of lesion segmentation and quantitative analysis. Therefore, integrating the feature advantages of different windows to construct more distinguishable imaging representations has become the key to improving the accuracy of NPC MRI image analysis. Currently, many experts at home and abroad have applied computer technology in the medical field ^[6-8].

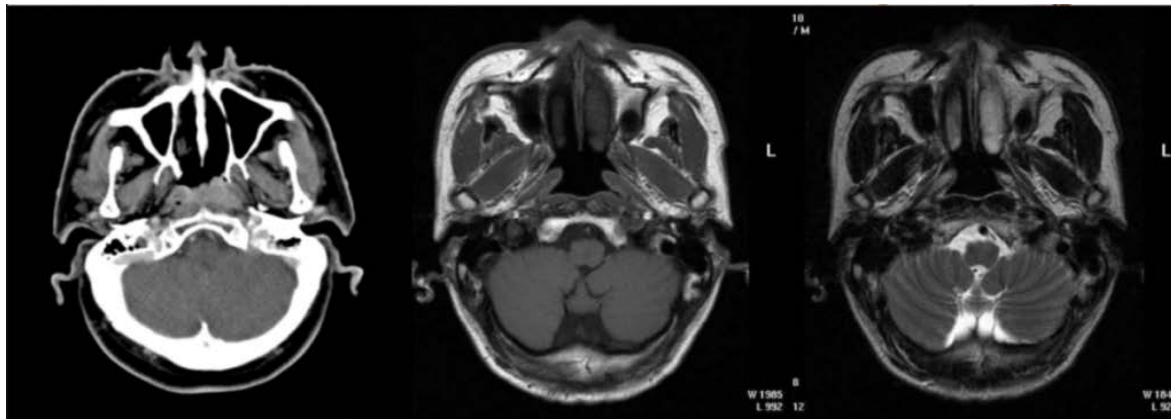


Figure 1. MRI images under different windows.

This paper maps key window information to the pseudo-color space through grayscale conversion and feature extraction of DICOM images under different window width/window level parameters, realizing the integration of multi-dimensional feature information ^[9]. Compared with traditional single-window imaging, multi-window fusion technology can effectively compensate for the information limitations of single-window imaging. The multi-window mechanism can adaptively cover heterogeneous image regions, avoid feature omission or redundancy of complex scenes by a single window, and enhance feature diversity and representation robustness. This paper used YOLOv8 as the base model to verify the effectiveness of the proposed method.

2. Related technologies

2.1. Window technology

Window width (WW) and window level (WL) jointly determine the contrast and brightness of medical images ^[10]. By collaboratively adjusting the grayscale mapping range and central threshold, they have a decisive impact on the visual presentation quality of digital medical images, directly affecting physicians' observation of lesions and tissue structures.

Window width refers to the range of CT values selected when displaying images. CT values outside this range will be displayed as pure white or pure black. Tissue structures within the specified range will be mapped to a series of grayscales from white to black (usually 16 levels or more) according to subtle differences in their density. A wide window width includes a broader range of CT values, allowing more tissues of different densities to be displayed simultaneously, thus reducing the overall contrast of the image, which is suitable for observing structures with large density differences; conversely, a narrow window width only displays a small range of CT values, amplifying subtle density differences of tissues within this range, significantly enhancing image contrast,

which is very suitable for observing soft tissues with similar densities.

Window level refers to the arithmetic mean of the upper and lower limits of CT values in the window width. It essentially determines which CT value will be displayed as intermediate gray. Since different tissues in the human body (such as bone, soft tissue, water, fat) have their typical CT value ranges, to observe the subtle structures of a specific tissue, it is necessary to select the CT value of that tissue as the center for window level setting. For example, a “lung window” with a low window level is needed to observe the lungs to highlight air-containing tissues and lung markings; while a “bone window” with a high window level was needed to observe bones to clearly display the cortex and medulla of bones.

For human MRI images, although the pixel values represent signal intensity rather than CT values, the same principle of window width and window level adjustment is fully applicable. Through precise adjustment of window width and window level, radiologists can effectively highlight the signal characteristics of specific tissues or lesions, thereby extracting more image details. This technology greatly optimizes the visual expression of images and is an extremely powerful tool for accurately distinguishing various tissues and organs in the human body, identifying early lesions, and conducting qualitative diagnosis.

2.2. YOLOv8

As a single-stage object detection algorithm, YOLOv8 consists of four parts: input layer, backbone network, neck network, and head network. Through architectural innovation, algorithm optimization, and training strategy improvement, YOLOv8 achieves a good balance between object detection accuracy, inference speed, and resource consumption.

The backbone network of YOLOv8 adopts an improved version of the CSPDarknet structure. By introducing the C2f module to replace the traditional C3 module, it improves computational efficiency while maintaining feature extraction capability. The C2f module divides the feature map into multiple branches for parallel convolution operations, combined with shortcut connections to realize feature reuse, effectively alleviating the gradient disappearance problem of deep networks. The neck network adopts the PAN-FPN structure, realizing multi-scale feature fusion through bottom-up feature pyramid and top-down path aggregation. The head network innovatively adopts an Anchor-Free design, directly predicting the center point coordinates, aspect ratio, and category probability of the target, avoiding the computational redundancy and hyperparameter dependence caused by the traditional anchor box mechanism.

3. MRI image resampling

3.1. Multi-window resampling

To effectively improve the utilization efficiency of original image data in the training and inference processes of deep learning models, this paper introduces an image resampling technology scheme based on multi-window settings. The core mechanism of this method was to input image data with more dimensional diversity into the deep learning model through fusion processing of multi-window image information, thereby helping the model capture richer image feature details and enhancing the model’s ability to understand image content and extract features accurately.

By parsing the metadata information contained in DICOM format image files, the preset window parameters of this type of image were accurately obtained; subsequently, based on these preset window parameters, two other

representative window parameter combinations were adaptively selected within their neighborhood range, namely:

$$wwi = \mu * ww_0 \quad (1)$$

$$wi = \mu * w_0 \quad (2)$$

Where (ww_0, w_0) are the optimal window width and window level, wwi represents the new window width, w_0 represents the new window level, and μ represents the weight.

Set μ to 0.25, 1.75, 0.5, 1.5, 0.75, and 1.25 to obtain the corresponding window width and window level. Observe the image effect of the window corresponding to different weights. It is found that the image effect is optimal when μ is 0.5 and 1.5. Under these weights, images under two new windows were obtained, namely (ww_1, w_1) and (ww_2, w_2) . The images under these two windows and the optimal window were used as data for the B, G, and R channels respectively to generate pseudo-color images. **Figure 2** is a schematic diagram of pseudo-color image generation.

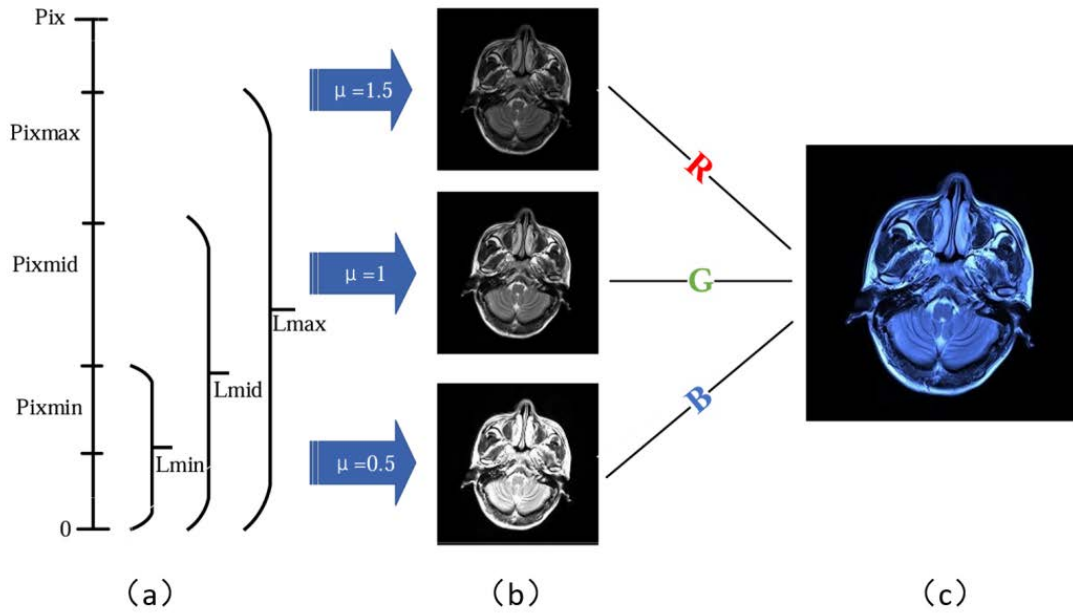


Figure 2. MRI image resampling schematic based on multi-window settings.

$[0, \text{Pix}]$ in Figure 2 represents the entire grayscale level in the original MR image. According to the pixel range contained in the MRI image (a), images under different window widths and window levels were obtained respectively (b), which were used as R, G, and B channels to synthesize RGB pseudo-color images (c) according to their grayscale display.

The synthesized pseudo-color image (c) contains image information of three different parameter configurations including the preset window, which can highlight the lesion tissue structure and grayscale features in the image by combining multiple factors, provide more lesion feature information for model learning, and enable the model to have a stronger ability to locate NPC lesions. Subsequent NPC lesion detection will be carried out based on the synthesized pseudo-color images.

3.2. Image annotation

The medical image data in this experiment was stored in DICOM format. Combined with the input requirements of YOLOv8 for data, NPC lesions were accurately annotated using annotation tools under the guidance of experts, and the annotated lesion bounding box information was standardized and stored in accordance with the VOC data annotation format to ensure the format compatibility and usability of the annotated data.

A total of 7496 annotated NPC lesion images were finally obtained, which were divided into training set, validation set, and test set in a ratio of 3:1:1, namely 4498 as training data, 1499 as validation set, and 1499 as test set. The sample composition of the dataset fully considers age distribution, gender differences, and regional characteristics, which can objectively reflect the clinical imaging manifestations of NPC in different populations, and has good representativeness and clinical reference value.

4. Experimental results and analysis

4.1. Experimental settings

To ensure the comparability of experimental data, the experimental environment, model parameters, and other settings were kept consistent. The experiment was carried out based on the Pytorch deep learning framework, CUDA11.3, and other environments. The specific software and hardware configurations were shown in **Table 1**, and the model parameter settings were shown in **Table 2**.

Table 1. Experimental configuration

Parameter	Configuration
CPU	AMD Ryzen 7 5800H
GPU	NVIDIA GeForce RTX 3060
Memory	128G
Video memory	6G
Development tool	Pycharm2021.3
Programming language	Python3.9
Framework technology	Pytorch
Acceleration environment	CUDA11.3
System environment	Ubuntu 18.04.6

Table 2. Model parameter settings

Parameter	Value
Epochs	300
Batch size	2
Learning rate	0.001
Weight decay	0.0005
Optimizer	AdamW

4.2. Results and performance analysis

To verify the effectiveness of the resampling method based on multi-window settings, active object detection

experiments were performed on the single-window NPC image set and the multi-window resampled NPC image set respectively. The results were shown in **Table 3**. Among them, is $mAP@0.5$, the average value of AP when IOU was greater than 0.5; was $mAP@50:5:95$, referring to the average value of corresponding results when IOU ranges from 0.5 to 0.95 with a step size of 0.05.

Table 3. Experimental results of single window and multi window images

Data				
Single-window	76.6%	35.5%	77.1%	36.0%
Multi-window	77.4%	36.0%	78.3%	36.2%

It can be concluded from **Table 3** that the lesion localization effect of pseudo-color images resampled based on multi-window settings was better. The experimental results based on pseudo-color images under different conditions were higher than those of the single-window NPC image set. The accuracies of $mAP@0.5$, $mAP@50:5:95$ were improved by 0.8%, 0.5%, 1.2%, and 0.2% respectively. Experiments were conducted on the NPC MRI image set with multi-window settings under different models, which effectively verifies the advantages of multi-window resampling of NPC MRI images, makes full use of data features, and improves the lesion detection performance of NPC. The detection effect of the NPC lesion detection model based on multi-window resampling was shown in **Figure 3**.

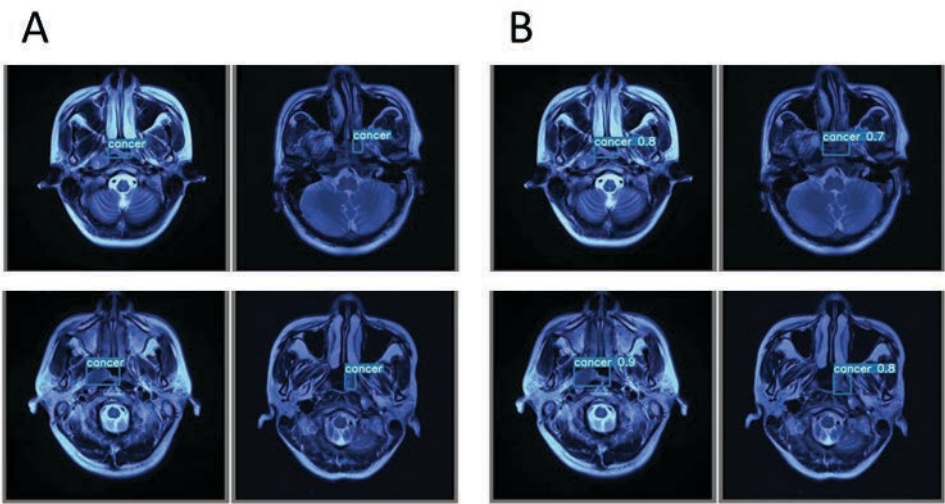


Figure 3. Nasopharyngeal carcinoma lesion detection results. A. Real lesion area; B. Detection result.

5. Summary

To fully utilize NPC lesion features, this paper proposes an NPC MRI image lesion recognition method based on multi-window resampling technology. This paper selects grayscale images of three windows with good effects, synthesizes them into NPC pseudo-color images, enhances lesion features, and compensates for the information limitations of single-window images. Experimental results show that this method can effectively improve the detection accuracy of the model for NPC lesions and has high clinical auxiliary diagnosis value.

Funding

Henan Provincial Science and Technology Research Project (Project No.: 252102211018)

Disclosure statement

The authors declare no conflict of interest.

References

- [1] Wang L, Zhou Y, Zhu X, et al., 2025, Research Progress on Writing Mechanism Based on Functional Magnetic Resonance Imaging Technology. *Chinese Journal of Rehabilitation Medicine*, 40(12): 1923–1929.
- [2] Lin X, Zhang J, Lin W, 2025, Prediction of Incidence and Mortality of Nasopharyngeal Carcinoma in China from 2022 to 2026: Based on GM (1,1) and ARIMA Models. *New Medicine*, 35(09): 1017–1023.
- [3] Zhou Z, Li K, Li N, et al., 2023, Age-Period-Cohort Model Analysis of Incidence and Mortality Trends of Nasopharyngeal Carcinoma in China from 1994 to 2019. *Chinese Journal of Disease Control & Prevention*, 27(08): 869–876 + 894.
- [4] Yu Q, Wang C, 2025, Research Status of Artificial Intelligence in Post-Processing of Imaging Technology Images. *Imaging Technology*, 37(06): 71–75.
- [5] Huang Y, 2025, Research on 3D Medical Image Registration Method Based on Dual-Window Attention and Dynamic Threshold, thesis, Guangxi University.
- [6] Tao G, Li H, Huang J, et al., 2022, SeqSeg: A Sequential Method to Achieve Nasopharyngeal Carcinoma Segmentation Free from Background Dominance. *Medical Image Analysis*, 78: 102381.
- [7] Ruan J, Xie M, Gao J, et al., 2023, EGE-UNet: An Efficient Group Enhanced UNet for Skin Lesion Segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 481–490.
- [8] Wang S, Zhu Y, Lee S, et al., 2022, Global-Local Attention Network with Multi-Task Uncertainty Loss for Abnormal Lymph Node Detection in MR Images. *Medical Image Analysis*, 77: 102345.
- [9] Zhao W, Cheng M, 2023, DICOM Image Analysis and Measurement System Based on Contour Detection and Target Localization. *Journal of Jiujiang University (Natural Science Edition)*, 38(03): 58–62.
- [10] Chen J, Yuan P, Hou H, et al., 2023, Adaptive Window Width and Window Level Algorithm for Medical CT Sequence Images. *Journal of Northeastern University (Natural Science Edition)*, 44(10): 1392–1400.

Publisher's note

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

Construction of a Motor-Driven Experimental Platform for Exploring the Law of Light Reflection

Zhongtian Wei, Jianwei Wang, Qintao Chen, Lihuang Qian, Zaikang Yang*

School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai 201620, China

**Author to whom correspondence should be addressed.*

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: In traditional middle school optical experiments, the fixed light source with an iron stand is inconvenient to operate, and the water mist generated by a spray bottle has a short duration and easily affects the reflection effect, leading to many limitations in the experimental exploration of the law of light reflection. This paper constructs a motor-driven experimental platform for exploring the law of light reflection. It innovatively adopts motor drive to realize flexible adjustment of the light source angle, and uses a medical humidifier atomizer instead of a traditional spray bottle to ensure continuous and stable mist that is not easy to adhere to the mirror surface. Through modular design, the platform integrates the functions of light source adjustment, mist generation and reflection observation. It has a simple structure and convenient operation, effectively solving the pain points of traditional experimental devices, providing a more efficient practical tool for optical experiment teaching, and featuring low cost and easy promotion.

Keywords: Optical experiment; Law of light reflection; Motor drive; Atomizer; Experimental platform

Online publication: February 12, 2026

1. Research background

Optics, as one of the core branches of physics, optical experiment teaching is a key link to help students understand optical laws. As the core content of basic optics, the exploration experiment of the law of light reflection occupies an important position in middle school and university physics teaching. Traditional experimental devices mostly use an iron stand to fix the laser light source, with cumbersome angle adjustment and insufficient precision. The single angle deviation during manual adjustment often exceeds 5° ; at the same time, they rely on a spray bottle to spray water mist to display the light path, which has problems such as short mist duration (usually less than 2 minutes) and water droplets easily adhering to the mirror surface leading to distorted reflection effects, seriously affecting experimental accuracy and teaching experience. Especially in the exploration of the core difficulty of “three lines coplanar”, traditional devices are difficult to help students establish spatial cognition through dynamic demonstration, resulting in inadequate conceptual understanding^[1].

With the advancement of educational informatization and experimental teaching reform, the market demand for efficient and convenient optical experimental devices is growing. Existing improvement schemes mostly focus on optimizing light path display, such as using smoke generators instead of spray bottles, but fail to take into account the flexibility and operational convenience of light source adjustment. Moreover, some devices have complex structures and high costs (mostly exceeding 5000 yuan), making it difficult to widely promote them in teaching. In addition, the development of college student innovation and entrepreneurship training programs also needs to transform interdisciplinary knowledge (such as the combination of motor drive technology and optical experiments) into practical teaching tools to improve students' innovation and practical abilities^[2].

Therefore, designing an experimental platform for exploring the law of light reflection integrated with motor drive and stable mist generation functions can not only solve the defects of traditional experimental devices but also provide new ideas for optical experiment teaching, having important teaching value and promotion prospects. Based on the College Student Innovation and Entrepreneurship Training Program, this project constructs a low-cost and high-performance experimental platform, aiming to optimize the experimental experience of exploring the law of light reflection and help improve the quality of experimental teaching^[3].

2. Platform design and principle

2.1. Design idea

With the core design goals of “simplifying operation, improving stability and reducing cost”, the platform adopts an interdisciplinary integration idea, combining motor drive technology, atomization technology and optical experiment needs. Motor drive is used to realize precise angle adjustment of the laser light source, replacing the traditional manual fixing method of the iron stand; a medical humidifier atomizer is selected to generate continuous and stable mist, solving many drawbacks of traditional spray bottles; the overall modular design is adopted, divided into light source adjustment module, motor drive module and reflection observation module. Each module works collaboratively to ensure a smooth and efficient experimental process^[4].

In terms of technical implementation, a NEMA 11 micro-stepper motor is selected as the drive core. The step angle of this type of motor is only 0.9° , and with a worm gear transmission structure with a 120:1 reduction ratio, the angle adjustment precision can reach 0.0075° , far exceeding the precision of traditional manual adjustment; the atomizer adopts the principle of piezoelectric ultrasonic humidification, breaking water into tiny droplets of 5–10 μm through 2.4 MHz high-frequency vibration. The generated mist is fine and uniform, with a long duration and not easy to adhere to the mirror surface; the platform base is integrally injection-molded with ABS material, and a silicone non-slip pad is pasted at the bottom, increasing the friction coefficient to 0.6 to ensure stability during the experiment. At the same time, a vernier scale mark is reserved, and the angle reading precision can reach 0.1° ^[5].

2.2. Overall architecture

The platform adopts a modular architecture as a whole, mainly composed of a light source adjustment module, a mist generation module, a reflection observation module and a base support unit. The light source adjustment module includes a laser pointer, a motor drive component and an angle adjustment disc, responsible for realizing precise positioning and angle adjustment of the light source; the mist generation module is a medical humidifier atomizer, which delivers mist to the experimental area through a food-grade silicone catheter. The end of the catheter is equipped with a conical diffuser to form a uniform mist field with a diameter of 15 cm; the reflection

observation module is composed of an aluminum-plated mirror (reflectivity $\geq 95\%$) and an acrylic scale disc. The scale disc adopts laser etching technology with a minimum division value of 0.5° , used for observing and recording the reflected light path; the base support unit is made of ABS material, featuring both portability and stability. Each module is fixed through a card slot design, and the disassembly and assembly time does not exceed 3 minutes^[6].

All components are low-cost and easily available general equipment. Among them, the purchase cost of the motor drive kit is about 800 yuan, the atomizer is about 300 yuan, and the structural parts are about 200 yuan. The total cost is controlled within 3500 yuan. Moreover, the core components (such as the motor and atomizer) adopt an independent 12V DC power supply design with a ripple voltage ≤ 50 mV, avoiding mutual interference and improving platform reliability (refer **Figure 1**).

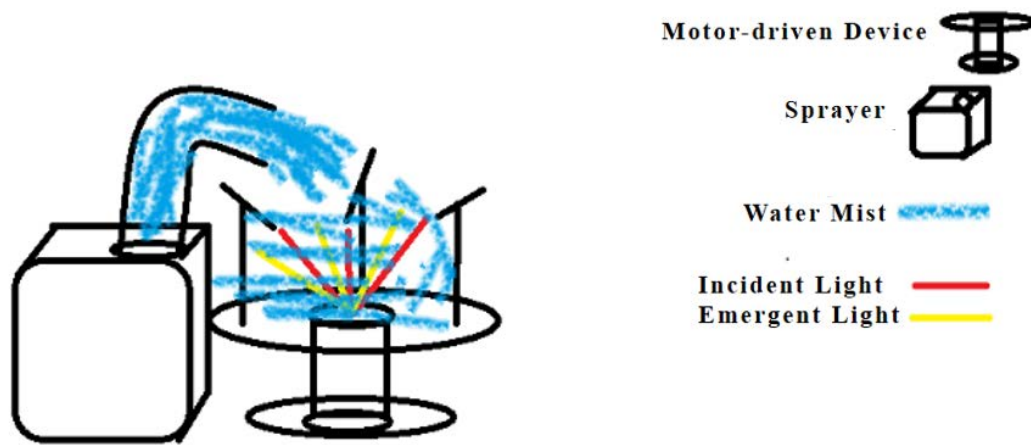


Figure 1. Schematic diagram of the device.

2.3. Core Module Design

2.3.1. Light source adjustment module

As the core of the platform, this module is mainly composed of a micro-stepper motor, an angle adjustment disc and a laser pointer holder. The motor adopts digital open-loop control, realizing forward/reverse rotation and speed adjustment through PUL/DIR signals. The speed range is $0.01\text{--}5^\circ/\text{s}$. It is connected with the angle adjustment disc through gear transmission to realize slow rotation of the disc, thereby driving the laser pointer to adjust the emission angle. The laser pointer holder adopts an adjustable buckle design, adapting to mainstream laser pointer models with a diameter of 8–12 mm. The buckle is equipped with a fluororubber non-slip pad with a static friction force ≥ 2 N to prevent the laser pointer from shifting during the experiment^[7].

Traditional light reflection teaching aids have problems such as difficult instrument adjustment, unstable light path display, large angle measurement error and inconvenient carrying. An integrated teaching aid is designed. The teaching aid uses an acrylic plate as the base material, divided into dual areas for reflection and refraction experiments. A smoke humidifier is used instead of a spray bottle to generate stable mist, combined with a fixable laser lamp and a wooden semicircular protractor to reduce angle measurement error; the precise adjustment of the light source position is realized through a slide rail to avoid deviations caused by manually holding the laser pointer^[8].

The edge of the angle adjustment disc is marked with $0\text{--}360^\circ$ scales, and 0.1° precision reading is realized

through a vernier structure. An electromagnetic damping device is set between the motor and the disc, which can realize arbitrary angle hovering after power failure with a hovering error $\leq 0.1^\circ$, facilitating the fixation of the light source angle during the experiment. The motor control adopts a knob-type encoder switch. Rotating clockwise increases the incident angle, and rotating counterclockwise decreases it. Each rotation corresponds to an angle change of 0.5° , which is simple to operate and suitable for students' independent experiments ^[9] (**Figure 2**, **Figure 3** and **Figure 4**).

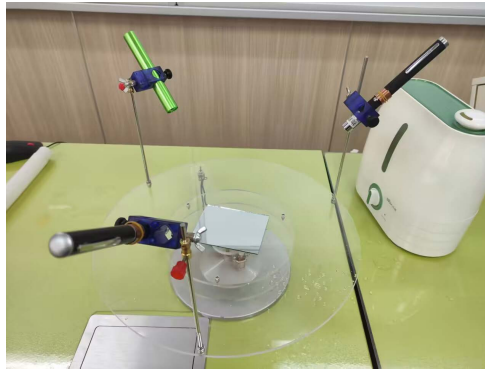


Figure 2. Light source adjustment module.

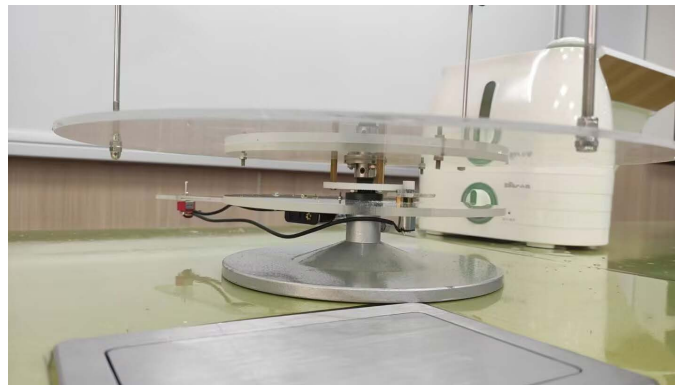


Figure 3. Motor drive unit.

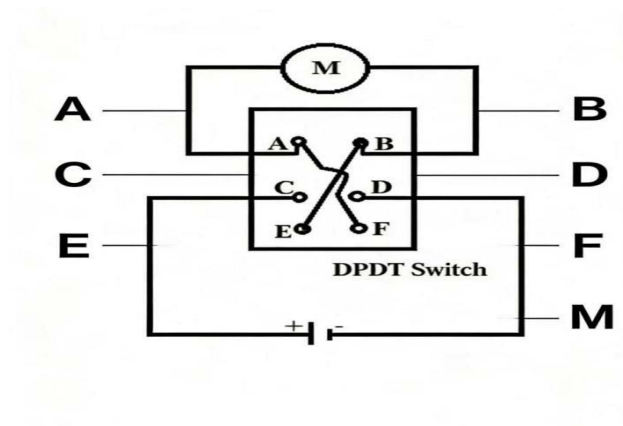


Figure 4. Schematic diagram of double-pole double-throw switch in motor circuit.

2.3.2. Mist generation module

A medical ultrasonic humidifier atomizer is selected as the core of mist generation. The vibration frequency of its piezoelectric vibrator is optimized to 2.4 MHz, the mist output range is continuously adjustable from 50–200 mL/h, and it can continuously generate mist for 1–2 hours, meeting the needs of more than 10 groups of experiments. The atomizer delivers mist to the experimental area through a silicone catheter. The catheter outlet is equipped with an adjustable nozzle, and the mist field diffusion angle can be changed by rotating the nozzle sleeve (adjustable from 30–90°) to avoid excessive mist diffusion. By adjusting the power and mist output of the ultrasonic atomizer, the mist density in the experimental area is controlled within the range of 0.5–2 g/m³, ensuring that the reflected light path is clearly visible without affecting the mirror reflection effect; at the same time, its high-speed imaging technology can be used to record the dynamic changes of the reflected light path, helping students understand the correlation between angle adjustment and light path deviation ^[10].

A low-cost motor-driven mirror module can be built with only a small motor, a mirror holder and a basic control circuit, realizing convenient adjustment of the reflected light angle, meeting the teaching demand of “dynamically observing the reflected light path” in the exploration of the law of light reflection. The core components are easily available, suitable for promotion in teaching scenarios ^[11]. To prevent mist from affecting the motor and circuit, the atomizer is kept more than 10 cm away from other modules, and the catheter outlet is oriented 45° above the experimental area. Air convection is used to make the mist evenly cover the light path observation range and reduce corrosion to experimental equipment. Experimental tests show that after continuous operation for 1 hour, the surface humidity of the motor control board is still lower than 60% without condensed water generation ^[12].

2.4. Control logic

The platform adopts a manual control mode with a simple and easy-to-understand operation process: place the platform on a horizontal desktop and calibrate the base levelness through a bubble level (error $\leq 0.5^\circ$); fix the plane mirror on the central axis of the scale disc and adjust the height of the laser pointer to be equal to the center of the mirror surface; start the atomizer, adjust the mist volume to the medium gear, and turn on the laser pointer after a stable mist field is formed in the experimental area (about 30 seconds); adjust the light source angle through the motor control knob, and the red incident light path and reflected light path can be clearly observed in the mist. Especially in the exploration of “three lines coplanar”, the coplanar characteristics of the incident light, reflected light and normal line can be intuitively displayed by rotating the light source module; read the incident angle and reflected angle data through the scale disc vernier, repeat the measurement 3 times to take the average value, and record the experimental results; after the experiment, turn off the laser pointer, atomizer and motor power in sequence ^[13].

The entire control process does not require complex programming. Students can master the operation essentials within 5 minutes and complete the experiment independently, meeting the usability requirements of teaching experiments. The platform also reserves a USB interface, which can expand the automatic angle scanning function through an Arduino controller, realizing continuous measurement and data recording of incident angles from 0–90°, providing possibilities for advanced exploration ^[14].

3. Conclusion

This paper designs and constructs a motor-driven experimental platform for exploring the law of light reflection.

The stepper motor drive realizes precise adjustment of the light source angle with an angle adjustment precision of 0.0075° , which is two orders of magnitude higher than traditional devices; the medical piezoelectric atomizer replaces the traditional spray bottle, extending the mist duration to 120 minutes, and the mist droplets are small and not easy to adhere to the mirror surface, effectively solving the problems of inconvenient operation and unstable light path display of traditional experimental devices. The platform has the following advantages:

- (1) Innovatively integrating motor drive and atomization technology to optimize the experimental experience, with an incident and reflected angle measurement error $\leq 0.2^\circ$;
- (2) Modular design, simple structure and convenient disassembly and assembly, a single set of devices can meet the group experimental needs of a 30-person class;
- (3) Low cost and easily available materials, the cost is only 1/3 of that of commercial optical platforms, suitable for large-scale promotion and application;
- (4) Simple operation, suitable for basic optical experiment teaching in middle schools and universities, and can also be used as a teaching carrier for innovation and entrepreneurship training programs^[15].

During the project implementation, the platform has been successfully applied to the experimental exploration of the law of light reflection in 20 classes of the Affiliated Middle School of Shanghai University of Engineering Science. The student experiment success rate has increased from 72% with traditional devices to 96%, and it has won the second prize in the School-level Selection of the China International College Student Innovation and Entrepreneurship Competition. In the future, the light source adjustment precision can be further optimized, and closed-loop control can be adopted to achieve sub-degree positioning; photoelectric sensors and data acquisition modules can be added to realize automatic recording of angle data and curve drawing; a supporting teaching APP can be developed to compare experimental data with theoretical curves in real time, improving the intelligence level of the platform and providing more comprehensive support for optical experiment teaching.

Funding

2024 College Student Innovation and Entrepreneurship Training Program of Shanghai University of Engineering Science (Project No.: cx2521002)

Disclosure statement

The authors declare no conflict of interest.

References

- [1] Liu Y, Cai Y, 2023, Improvement and Teaching Design of the Experimental Device for “Exploring the Law of Light Reflection”. *Physics Teaching Reference for Middle School*, 52(15): 46–47.
- [2] Xing Y, 2023, A Survey on the Development of Middle School Students’ Basic Optical Concepts, thesis, Fujian Normal University.
- [3] Sun W, Ji L, 2025, Discover in Rotation, Grow in Discovery—Taking the Teaching of “Light Reflection” as an Example. *Physics Teaching*, 47(11): 46–49 + 45.
- [4] Liu Y, 2024, Three-Dimensional Laser Simultaneous Localization and Mapping for Roadheaders, thesis, North China University of Technology.

- [5] Ye C, Zhang M, Li J, 2025, Innovative Improvement of Teaching Aids for Light Reflection and Refraction. *Creative Education Studies*, 13(9): 1024–1028.
- [6] Bauer T, Smith J, Lee H, 2023, Dynamic Observation of Light Reflection Based on High-Speed Imaging. *Phantom Journal of Optics*, 8(2): 45–52.
- [7] Wang C, Li J, Zhang W, 2023, Electromechanical Co-Simulation Technology of Voice Coil Actuated Fast Steering Mirror. *Optics and Precision Engineering*, 31(4): 890–898.
- [8] Thorlabs Inc., 2025, Stepper Motor Rotation Mount Technical Datasheet. Thorlabs.
- [9] Chen X, 2023, Innovative Teaching Aid Production to Break Through “Three Lines Coplanar”. *Physics Teaching*, 45(12): 37–39.
- [10] Gomez R, Martinez L, Ruiz J, 2025, Democratizing Science with Parametric Design: Low-Cost Optical Instrument for Education. *PLOS One*, 18(9): e0187219.
- [11] Li M, Wang H, 2025, Research on Spray Characteristics of Piezoelectric Ultrasonic Atomizers. *Chinese Journal of Scientific Instrument*, 46(12): 80–87.
- [12] AKT Motor, 2025, Educational Research Equipment Solutions. AKT Motor.
- [13] Zhang X, Liu M, 2024, Innovative Design and Practice of Middle School Physics Optical Experiment Devices. *Physics Bulletin*, 2024(7): 56–59.
- [14] Wang L, Zhao Q, 2024, Research on Modular Design of Low-Cost Optical Experimental Platform. *Research and Exploration in Laboratory*, 43(5): 98–102.
- [15] Chen H, Lin X, 2025, Application and Precision Optimization of Stepper Motors in Optical Experiments. *College Physics Experiments*, 38(2): 78–83.

Publisher’s note

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

Exploring the Path of AI Technology's Empowerment of New Developments in Higher Education

Chenguang Yao

China Women's University, Beijing 100101, China

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited

Abstract: As the core driving force leading the new round of technological revolution and industrial transformation, artificial intelligence is profoundly reshaping the higher education ecosystem. Based on the background of artificial intelligence development, this paper expounds the value of AI technology in promoting the innovative development of higher education, explores the specific paths of AI technology empowering the innovative development of higher education from three dimensions of talent training, scientific research, and governance, and puts forward the required guarantee conditions. It aims to promote the in-depth integration of artificial intelligence and higher education, profoundly change the form of higher education, lead higher education towards a more personalized, precise, and intelligent development path, and provide reference for solving a series of problems encountered in the current development of higher education.

Keywords: Artificial intelligence; AI technology; Higher education; Development; Path

Online publication: February 12, 2026

1. Introduction

Under the background of new-quality productive forces, a new generation of artificial intelligence has gradually swept the world, becoming a powerful engine leading industrial transformation and technological revolution. China vigorously promotes the extensive integration of artificial intelligence with various fields of society. With the advancement of a series of policies such as the "New Generation Artificial Intelligence Development Plan", the "Interim Measures for the Administration of Generative Artificial Intelligence Services", and the "Guidelines for the Construction of a Comprehensive Standardization System for the National Artificial Intelligence Industry (2024 Edition)", large AI models have entered a new stage of large-scale application pilots, showing enormous application potential in the education sector. In the field of higher education, the application of artificial intelligence has transitioned from the stage of logical reasoning and expert systems to the stage of machine learning, evolving from a mere "auxiliary tool" to an "empowerer"^[1]. The development and promotion of artificial intelligence technology (hereinafter referred to as "AI technology") have impacted higher education

models, scientific research environments, and teaching environments ^[2]. Institutions of higher learning are the main positions for cultivating innovative talents. How to conform to the development trend of “AI + education”, take AI technology as the core engine, explore and expand the application scenarios of this technology, and at the same time drive the all-round leapfrog development of the higher education system from connotation to extension, empowering the innovation of all links and processes of higher education talent training is imperative.

2. The value of AI technology in promoting the innovative development of higher education

The integration of AI technology into the field of higher education is not a simple technical superimposition, but a systematic empowerment of the educational ecosystem. Its core value is reflected in the following three aspects:

2.1. Empower precise talent training and solve the dilemma of homogenization

Traditional higher education mostly adopts a standardized teaching model, arranging teaching content uniformly, which is difficult to adapt to students’ personalized growth needs ^[3]. Through multi-dimensional learning situation data collection and intelligent analysis, AI technology constructs personalized knowledge and ability maps, realizes “one thousand people, one thousand faces” precise teaching push, promotes the transformation of educational goals from knowledge imparting to the compound training of “higher-order thinking + innovative ability + digital literacy”, and accurately meets the needs of cultivating top-notch innovative talents.

2.2. Drive scientific research innovation and reshape research paradigms

With the advantages of large-scale data processing and complex system modeling, AI technology can accelerate the process of academic research. According to automated execution instructions, it can complete data collection, literature review collation, code writing, experimental scheme design, text translation and revision, simplifying teachers’ scientific research processes and improving research efficiency ^[4]. Intelligent scientific research tools can intelligently track domestic and foreign research trends according to researchers’ research interests and directions, construct large-scale scientific research theme datasets, conduct virtual experiment verification and error correction, accelerate the scientific research process, and help obtain results that are difficult to find with traditional research methods ^[5].

2.3. Optimize management level and ensure high-quality development

Higher education management involves a series of layouts and services for teachers and students. Through data collection, automatic analysis, and scientific decision-making systems, AI technology reintegrates data on education and teaching, teaching management, and scientific research services, reshapes the school-running and governance pattern, promotes the transformation of education management from “experience-extensive” to “data-precise”, and improves comprehensive management efficiency ^[6]. At the same time, through large models and intelligent facilities, AI technology can scientifically plan the upgrading direction of campus infrastructure, serve the construction of interdisciplinary and cross-regional scientific research platforms, promote the construction of a ubiquitous intelligent interconnected learning environment, and open a new chapter in the open and lifelong development of higher education ^[7].

3. Paths of AI technology empowering the innovation and high-quality development of higher education

3.1. Focus on the core of talent training and build an AI-driven personalized training system

Talent training is the fundamental task of higher education. AI technology provides precise solutions for breaking traditional talent training bottlenecks and improving training quality. The specific paths are as follows:

3.1.1. Build an intelligent adaptive learning system to realize large-scale teaching students in accordance with their aptitude

Construct a full-process closed-loop ecosystem. Relying on large AI education models, build an adaptive system covering “pre-class–in-class–after-class” to realize a closed loop of “active prediction - precise push - dynamic optimization”.

Multi-dimensional data collection and analysis. Collect explicit and implicit data such as learning habits, cognitive load, and classroom responses, and establish dynamic learning portraits and knowledge ability maps ^[8].

Personalized resource push. Customize Q&A, exercises, and extended resources for students with different foundations, such as pushing cutting-edge literature for students with spare capacity and strengthening core knowledge points for students with weak foundations.

3.1.2. Promote the digital upgrade of the curriculum system to adapt to future industrial needs

Build an interdisciplinary knowledge graph engine. Construct a three-dimensional network of “core knowledge - ability requirements - industrial needs” to form modular interdisciplinary course groups such as “AI + biomedicine”.

Establish a dynamic iteration mechanism. Track the technological trends of future industries, integrate cutting-edge content such as large model applications and AI security into courses to ensure that courses resonate with industries ^[9].

Develop intelligent resources. Use AI to produce interactive digital textbooks and virtual simulation courseware, and realize precise push through the National Smart Education Platform.

Innovate general education course forms. Offer courses such as “AI and Human Civilization”, and adopt the “project-based learning + human-machine collaborative creation” model to improve students’ AI literacy.

3.1.3. Innovate teaching organization forms and explore new human-machine collaborative models

Build “AI + education meta-universe” scenarios. Relying on VR/AR and AI to create immersive learning environments such as virtual classrooms and virtual laboratories to help students improve practical application abilities ^[10].

Construct a tripartite collaborative mechanism. Intelligent teaching assistants take on basic work such as Q&A and grading, teachers focus on curriculum design and value guidance, and students carry out independent collaborative learning with the help of AI.

Promote the normalization of blended learning. Realize intelligent matching of learning groups through “AI learning partners”, and promote the widespread implementation of online-offline integrated autonomous learning models ^[11].

3.2. Empower scientific research innovation and build an AI-driven cross-domain collaborative system

AI technology promotes the transformation of scientific research paradigms and injects new momentum into scientific research innovation in higher education. The specific paths are as follows:

3.2.1. Promote the intelligent transformation of scientific research paradigms and improve innovation efficiency

Full-process automation of data processing. Use AI to realize the full-process automation of scientific research data “cleaning - analysis - modeling - prediction”, greatly shorten the R&D cycle of scientific research projects, and improve the accuracy and efficiency of data processing.

Build a dual-track scientific research model. Establish a collaborative scientific research model of “AI virtual simulation + real experiments”, which can effectively reduce the cost investment and potential risks in the scientific research process and optimize the experimental scheme design process.

Build a cross-domain collaborative platform. Through knowledge graph technology, construct interdisciplinary and cross-domain scientific research achievement maps, providing favorable conditions for carrying out interdisciplinary scientific research innovation to solve cutting-edge problems ^[12].

3.2.2. Promote interdisciplinary integration and cultivate new disciplinary growth points

Attach importance to the cross-integration development of traditional disciplines. Use AI technology to break the research boundaries of traditional disciplines and establish a research system of “artificial intelligence + traditional disciplines”. For example, realize the reconstruction of historical scenes and the simulation of social operation laws in the field of humanities and social sciences, and optimize breeding schemes and production regulation logic in the field of agriculture.

Build a growth empowerment system. Relying on AI technology to establish a disciplinary development trend prediction model, establish an integrated mechanism of “discipline–scientific research–industry”, and accelerate the transformation of scientific research achievements into real productive forces.

3.2.3. Build intelligent scientific research platforms and strengthen resource sharing

Construct an independent and controllable large model system. Develop a hierarchical scientific research large model system of “general basic large model + discipline-specific fine-tuning model” to improve the intelligent support capacity of the entire scientific research process.

Build an innovation testbed. Focus on basic scientific research and key core technology research, develop forward-looking scientific research tools and experimental platforms, and provide technical support for cutting-edge scientific research innovation ^[13].

Promote the intelligent upgrading of equipment. Construct an intelligent laboratory network of “IoT perception + AI analysis” to realize remote control of experimental equipment and automatic data collection and analysis; build an interdisciplinary scientific research data sharing platform to break data silos.

3.3. Optimize governance efficiency and build an AI-driven precise governance system

AI technology provides new ideas for the reform of higher education governance and promotes the modernization of governance capacity. The specific paths are as follows:

3.3.1. Build an “educational digital map” to empower scientific decision-making

Construct a digital twin system. Realize dynamic monitoring, precise prediction, and scientific decision-making of higher education development.

Build a discipline update system. Integrate data on industry development, post changes, professional settings, and discipline construction to construct a dynamic discipline update system oriented to industrial needs, providing a basis for the school’s professional settings and curriculum development.

Build a talent supply-demand matching platform. Use data analysis and AI algorithms to construct talent supply-demand prediction models for traditional and emerging industries, adjust enrollment scales and professional layouts according to demand changes, and build high-quality talent training highlands.

Optimize resource allocation. Use AI algorithms to optimize the allocation plan of faculty, teaching facilities, and funding investment, and improve the efficiency of educational resource utilization.

3.3.2. Promote the digitization of management services and improve administrative efficiency

Upgrade “AI + one-stop service”. Integrate data from various management systems such as academic affairs, scientific research, finance, and personnel to realize full-process automated approval of high-frequency businesses; build an intelligent consulting service system through generative AI to provide personalized and real-time consulting services for teachers and students.

Construct a smart campus operation and maintenance system. With the help of big data, sensors, and intelligent facilities, establish an integrated intelligent connection system covering on-campus venues such as classrooms, laboratories, libraries, stadiums, and canteens, integrate all campus resources, spaces, and services, and provide an operation and maintenance environment with automatic perception, intelligent evaluation, and intelligent upgrading ^[14].

3.3.3. Optimize faculty management and support teacher development

Establish a digital literacy improvement system. AI evaluates teachers’ abilities, formulates personalized training plans, and builds a teaching innovation community.

Construct a precise evaluation system. Collect multi-dimensional data to establish an evaluation model, generate development reports, and provide personalized guidance on teaching and scientific research.

Innovate recruitment and training models. AI constructs a national talent database and assigns “AI mentors” to young teachers to accelerate their growth.

3.4. Consolidate the guarantee foundation and build an AI-empowered support and guarantee system

The in-depth integration of AI and higher education requires the support of improved guarantee mechanisms. The specific paths are as follows:

3.4.1. Update educational concepts and build consensus on integration

Carry out concept innovation actions. Through cutting-edge theoretical seminars, technical development interpretations, and future education scenario deductions, guide teachers and students to deeply understand the subversive impact of AI technology on higher education, and establish a correct cognition of “technology empowerment rather than replacement” ^[15].

Build an innovative cultural ecosystem. Set up special innovation funds to encourage teachers and students

to carry out AI education innovation practices and research; hold interdisciplinary academic salons, technological innovation forums and other activities to create a campus cultural atmosphere of courage to explore and innovate.

Incorporate into top-level design. Establish a leading group for the integrated development of “AI + education” led by school leaders and involving cross-departmental participation, clarify development goals, key tasks, and implementation paths; strengthen inter-school and school-enterprise cooperation, introduce advanced theoretical achievements and technical resources, and promote integrated innovation and development.

3.4.2. Improve system construction and strengthen policy support

Formulate application management methods. Clarify application norms, responsibility division, and assessment standards in teaching, scientific research, management and other fields.

Establish an incentive mechanism. Set up special funds and reward projects to encourage teachers to carry out AI teaching and scientific research innovation.

Establish a cross-departmental collaborative mechanism. Break barriers and form an efficient collaborative work system with clear division of labor.

3.4.3. Strengthen data governance and ensure data security

Establish a standardization system. Formulate full-process data specifications and use AI to improve data quality.

Strengthen technical protection. Construct an “AI + data security” system, adopt encryption, anonymization and other technologies to ensure security, and strictly protect student information.

Establish a long-term mechanism. Set up a special institution, conduct regular training and drill, and establish an audit and accountability mechanism.

3.4.4. Standardize ethical guidelines and prevent technical risks

Formulate ethical guidelines. Clarify ethical bottom lines, regulate the application of generative AI, and prevent academic misconduct.

Establish an intelligent review mechanism. Set up an interdisciplinary committee and use AI tools to conduct full-process ethical evaluation and risk early warning.

Strengthen ethical education. Incorporate AI ethics into courses and training, and encourage teachers and students to participate in the formulation of norms.

4. Conclusion

In summary, driven by AI technology, we should actively integrate and apply new technologies, integrate AI technology into the entire process of higher education teaching practice, strictly control the transparency and security of AI applications, give full play to the positive role of new technologies, further realize students’ personalized learning, improve teachers’ teaching and research levels and school governance efficiency, and promote the high-quality development of higher education.

In the future, higher education is expected to achieve comprehensive and in-depth systematic transformation. First, the talent training model will shift from “standardized supply” to “personalized adaptation”. AI technology will break the form of standardized resource supply, objectively evaluate and feedback learning progress through data analysis systems, personalized recommendation functions, and intelligent evaluation tools, provide tailor-made learning plans and resources for each student, and promote the transformation of the talent training model

from focusing on achieving educational goals to meeting students' personalized learning needs. Second, the scientific research paradigm will shift from “experience-theory driven” to “data-intelligence driven”. By building an intelligent scientific research system of “data collection–analysis–modeling–prediction”, AI technology will carry out virtual experiment design and research in conjunction with VR technology, promote the upgrading of scientific research from “hypothesis verification” to “data discovery”, and improve the efficiency of scientific research innovation and the quality of achievement transformation. Third, the governance form will shift from “passive response” to “active intervention”. By connecting data systems of multiple departments such as academic affairs, student affairs, and finance, AI technology will establish a dynamic perception-monitoring-early warning system for teachers and students, which can predict risks in advance, realize optimal resource allocation, accurately prevent and control risks, promote the transformation from human passive response to AI precise intervention, and realize the modernization of governance capacity.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Wang H, Zhang L, 2024, Higher Education Teaching Reform Based on Artificial Intelligence Technology. *Research and Practice on Innovation and Entrepreneurship*, 7(24): 26–28.
- [2] Zhang L, 2024, Construction of a Chinese Paradigm for AI-Empowered Higher Education Teaching Reform. *China Higher Education*, 2024(24): 9–13.
- [3] Xue Q, Chen M, Zhao J, et al., 2024, An Analysis of the Southern University of Science and Technology's “Sample” for AI-Empowered Innovative Development of Higher Education. *China Higher Education*, 2024(24): 19–24.
- [4] Hu X, Lin Z, Liu X, 2024, The Integration of Artificial Intelligence into Education: Global Trends and Chinese Directions. *E-Education Research*, 45(12): 13–22.
- [5] Yang N, Tang A, 2024, AI-Empowered Higher Education Governance: International Experience and Chinese Choices. *E-Education Research*, 45(11): 38–44.
- [6] Zhou Z, Zhao J, 2024, Research on the In-Depth Integration Path of Artificial Intelligence and Higher Education from an Interdisciplinary Perspective. *High-Technology & Industrialization*, 30(10): 135–136.
- [7] Ruano-Borbalán J, Zhang J, 2024, The Transformative Impact of Artificial Intelligence on Higher Education: A Critical Reflection on Current Trends and Future Directions. *Tsinghua Journal of Education*, 45(5): 13–24.
- [8] Xu S, Sun H, 2024, Subversion and Reconstruction: Philosophical Exploration of Higher Education in the Age of Artificial Intelligence. *Modern University Education*, 40(5): 34–46 + 111.
- [9] Gai Q, 2024, Challenges and Countermeasures of Generative AI Empowering the High-Quality Development of Higher Education. *University Education*, 2024(17): 16–20.
- [10] Zhang X, Gao M, Jin J, et al., 2024, Exploring the Application of Artificial Intelligence Technology in Higher Education Reform. *Science & Technology Information*, 22(14): 203–206.
- [11] Zhu R, Li Y, Chen H, 2024, Innovation and Practice of the Integration of Artificial Intelligence into Higher Education Teaching Models. *China Informatization*, 2024(6): 83–84.
- [12] Sun D, Wang L, Shang L, 2024, The Connotation, Dilemmas and Paths of AI-Empowered High-Quality Development of Higher Education in China. *Modern Education Management*, 2024(6): 34–42.

- [13] Liu J, Zeng H, Jin W, et al., 2024, AI-Empowered Higher Education: Logical Approach, Typical Scenarios and Practical Path. *Journal of Xi'an Jiaotong University (Social Sciences Edition)*, 44(3): 11–20.
- [14] Bie D, Guo Y, 2024, New Trends in the Innovative Development of Higher Education in the Age of Artificial Intelligence. *China Higher Education*, 2024(Z1): 39–44.
- [15] Cui J, Ma Y, 2023, Research Progress and Prospect of Artificial Intelligence Education in China. *Journal of Higher Education Management*, 17(6): 31–39.

Publisher's note

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

A Survey on Artificial Intelligence Systems Robustness: Adversarial Attacks and Defenses

Wei Zheng

Sichuan Forestry and Grassland Administration, Chengdu 610081, China

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited

Abstract: Artificial intelligence systems have achieved widespread applications across many fields such as image classification, speech recognition, and game playing. However, as their decision-making logic is primarily learned from data, their outputs are highly sensitive to data anomalies and are particularly vulnerable to adversarial perturbations. This paper conducts a comprehensive survey on the robustness of artificial intelligence systems, reviewing classical adversarial attack and defense methods, and summarizing future development trends. We hope this work can provide valuable insights for research on the robustness of artificial intelligence systems and support the development of trustworthy artificial intelligence.

Keywords: Artificial intelligence; Adversarial attacks; Adversarial defense

Online publication: February 27, 2026

1. Introduction

Artificial intelligence (AI) technologies such as deep learning, have been widely applied in many fields such as autonomous driving and healthcare, significantly enhancing productivity^[1,2]. They are poised to exert even more extensive and profound impacts on production methods and lifestyles across various industries in the future. However, unlike traditional software that acquires decision-making logic through explicit programming, the decision-making logic of AI systems is primarily learned by pre-defined model structures from training data. Developers can only indirectly influence the system's decision-making logic by modifying training data, features, and architectural details of the model (e.g., the number of layers)^[3]. Consequently, AI systems are highly sensitive to data anomalies, vulnerable to biased data, and face significant challenges in robustness.

A study by Szegedy et al. revealed that deep learning models can produce completely erroneous outputs when subjected to deliberate, minimal perturbations in their inputs, changes that are nearly imperceptible to humans^[4]. In light of this, the academic community sparked a research boom on the robustness of artificial intelligence, resulting in a variety of robustness analysis methods, techniques, and tools. While several studies have conducted surveys on the robustness of AI systems—for instance, Hamon et al. extended the discussion

from technical aspects to policy recommendations, addressing robustness and explainable challenges in real-world AI applications^[5]. Javed systematically examined robustness issues in deep learning systems for medical diagnostics; and Tocchetti provided a systematic analysis of research progress, challenges, and future directions in AI robustness from a human-centered perspective^[6,7]. However, those studies remain a lack of systematic investigation into AI robustness specifically from the angles of adversarial attacks and defenses. For example, Wang et al. focused solely on adversarial attacks and defenses in communication application classification models based on deep neural networks, lacking a broader survey and thus falling short of effectively guiding the design of general adversarial defense strategies^[8].

To address the aforementioned challenges, this paper conducts a systematic analysis of current robustness research from the perspectives of adversarial attacks and defenses, focusing on typical AI systems such as deep learning systems. The study aims to inspire developers to design more effective defense mechanisms, thereby supporting the further application of trustworthy artificial intelligence technologies. The main contributions of this paper can be summarized as follows:

- (1) It reviews adversarial attack and defense techniques for artificial intelligence systems;
- (2) It builds a bridge between academic researchers and application engineers, taking appropriate measures to enable potential in-depth collaboration in the future.

The remainder of this paper is structured as follows: Section 2 introduces related basic concepts; Sections 3 and 4 review current research on adversarial attacks and defenses in AI system, respectively; finally, Section 5 provides a comprehensive summary of this paper.

2. Basic conception

2.1. Artificial intelligence robustness

AI algorithms and traditional software lack physical entities and realize their intended functions through programs. Therefore, AI algorithms can be regarded as a special category of software. This paper refers to software systems deploying at least one AI algorithm as artificial intelligence systems. However, the methods by which traditional software systems and AI systems acquire their decision-making logic are fundamentally different^[3]. In traditional software, decision-making logic is directly determined by the control flow and data flow defined by developers in the program. As a result, its operational outcomes are robust. In contrast, the development of AI systems follows a completely different paradigm: after developers define the architecture (such as a deep learning model), the decision-making logic is autonomously learned by the system from training data. Developers cannot directly influence this decision-making logic, as illustrated in **Figure 1**. Consequently, the decision-making logic of AI systems is not explainable and is highly sensitive to data anomalies, making them susceptible to biased data. This inherent characteristic has spurred related research into robustness. The academic community holds various definitions for the robustness of AI systems. For instance, Mannor et al. define machine learning robustness as the bounded difference in the loss function between any subsets of the training and test sets^[9]. Despite varying definitions, the robustness of an AI system can be broadly defined as the stability of its performance when confronted with anomalous inputs, that is, the system's tolerance to variations in data.

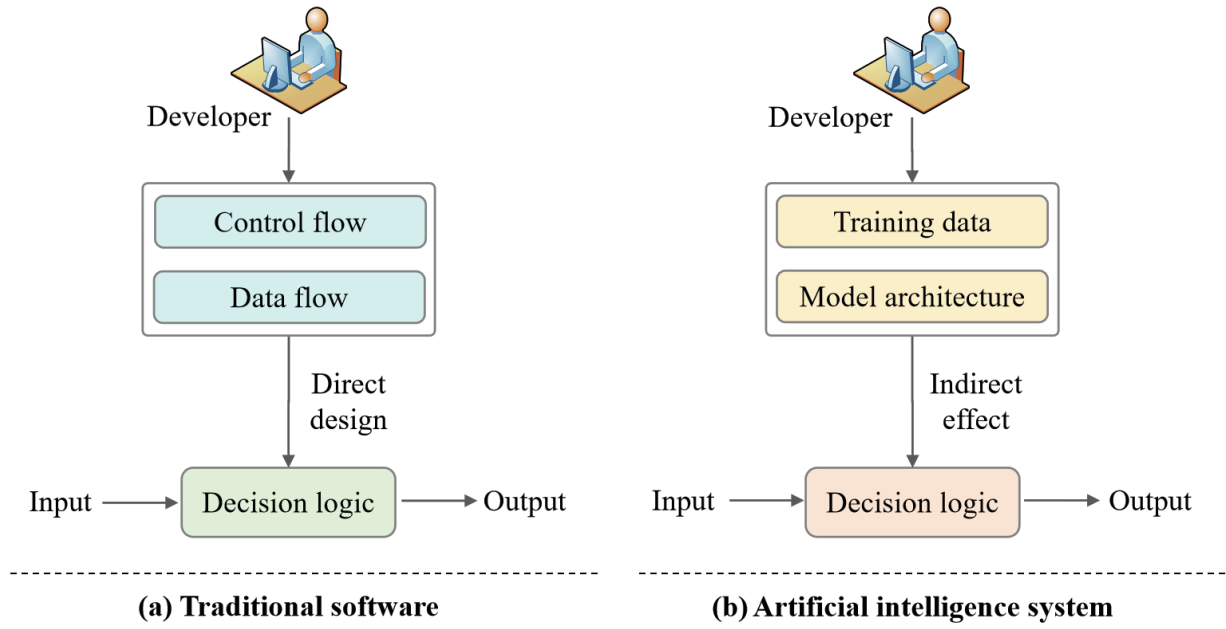


Figure 1. The differences in decision logic acquisition between traditional software and AI algorithms.

2.2. Adversarial attacks and defenses

Szegedy et al. firstly demonstrate that introducing carefully crafted minimal perturbations to images could mislead deep learning models into making completely erroneous decisions ^[4]. Such deliberately designed input perturbations are generally referred to as adversarial attacks. This technique involves generating adversarial examples by adding subtle perturbations to input data and using these examples to maliciously attack AI systems, such as deep neural networks, causing the system to produce incorrect or biased outputs. The objective of adversarial attacks is to deceive AI systems, preventing them from accurately performing tasks such as classification or regression, thereby compromising the system's robustness. The general representation of an adversarial attack is expressed as $F(x+\eta) \neq F(x)$, where x represents the original sample, η denotes the added perturbation, $x+\eta$ is the resulting adversarial example, and $F(\bullet)$ represents the AI model. It is important to note that the perturbation η is typically imperceptible to humans and does not affect human judgment, yet it can mislead the AI system's outputs. Since Szegedy's seminal work, researchers have discovered that adversarial examples widely exist in almost all deep learning models ^[10]. Beyond the field of computer vision, AI systems across various domains such as natural language processing, audio and video recognition, recommendation systems, and large language models also face threats from adversarial examples ^[11–15].

3. Adversarial attacks

Based on the underlying principles of adversarial attack methods, they can be classified into three main categories: gradient-based attacks, optimization-based attacks, and generation-based attacks, as illustrated in **Figure 2**.

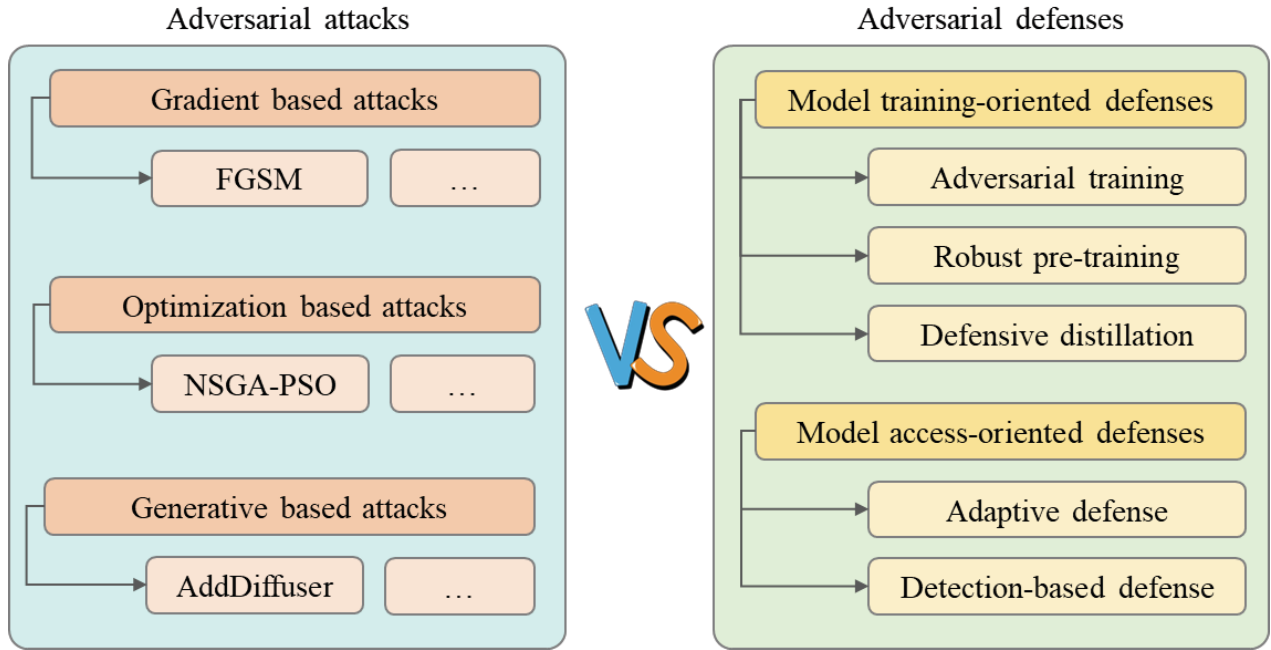


Figure 2. The adversarial attacks and defense methods investigated in this study.

3.1. Gradient-based attacks

Artificial intelligence models, such as deep neural networks, are typically trained by minimizing a loss function. Gradient-based attack methods exploit this very principle by maximizing the loss function to generate adversarial examples. Specifically, these methods compute the gradient of the loss function with respect to the input data, then apply minimal perturbations to the data along the direction (or the opposite direction) of this gradient. This causes the perturbed input data to induce significant errors in the system's predictions.

Classic gradient-based attack methods primarily include the fast gradient sign method (FGSM), its variant R + FGSM, the basic iterative method (BIM), the projected gradient descent (PGD) method, and the Momentum Iterative Method (MIM) ^[16-20]. The FGSM normalizes the gradient by applying the sign function, *sign* (), causing perturbations to be added in the direction of the gradient of the loss function, and maximizes the loss and thereby deceives the system ^[16]. To address the limitation of FGSM being easily defended, Kurakin et al. proposed the variant R + FGSM, which incorporates random noise during the generation of adversarial examples to evade gradient masking-based defense strategies ^[17]. The BIM starts from the original input sample and iteratively applies small perturbations to the current adversarial example to produce the next sample ^[18]. Through multiple iterations, it gradually increases the magnitude and number of perturbations to more effectively deceive the model. The PGD method is similar to BIM, but it includes an additional random initialization step, and the random perturbation is generated by uniformly sampling random noise within the perturbation threshold ^[19]. The MIM introduces momentum during the iterative process to accumulate gradient directions, preventing the adversarial perturbation from falling into poor local optima and resulting in more stable updates to the adversarial perturbation ^[20].

Among the methods mentioned above, the FGSM is characterized by its simple principle and rapid execution, though it achieves a relatively lower attack success rate. Methods such as R + FGSM, the BIM, and PGD are all improvements upon FGSM, while they enhance the attack success rate, they also tend to cause overfitting and exhibit poor generalization ability. The MIM mitigates the overfitting issue to some extent by incorporating momentum into the gradient iteration process. In general, most gradient-based attack methods focus on enhancing

attack potency and transferability. When models are adversarially trained using examples generated by such methods, they tend to develop good robustness against similar gradient-based attacks. However, their generalized robustness against other types of adversarial attacks remains limited.

3.2. Optimization-based attacks

Different from gradient-based methods, the optimization-based attack approaches frame the generation of adversarial examples as a constrained optimization process. These methods employ optimization algorithms to identify the optimal adversarial perturbation, one that minimizes the magnitude of perturbation while maximizing the loss, to create adversarial examples. In contrast to gradient-based techniques, these methods do not explicitly compute gradients; instead, the computation and backpropagation of gradients are embedded within the optimizer. This design allows them to achieve both high attack accuracy and low perturbations.

Carlini and Wagner (C&W) is a classic optimization-based attack method. It transforms the constrained optimization problem into an unconstrained one, uses optimized parameters to represent adversarial examples, and applies optimization techniques to precisely attack the target model. Feng et al. proposed an adversarial attack method based on a non-dominated sorting genetic algorithm and particle swarm optimization (NSGA-PSO), which generates adversarial digital watermarks in black-box attack scenarios. This approach demonstrates strong generalization capability and exhibits great robustness against image transformation defense methods^[21,22].

Optimization-based attack methods generally exhibit good transferability across different network architectures and show strong resistance to example transformation defense strategies. Moreover, as they do not rely on access to predictions and labels, they are more suitable for real-world scenarios. However, the performance of these methods may vary significantly across different datasets.

3.3. Generation-based attacks

Generation-based attack methods typically involve generative adversarial networks (GANs) or other generative models. These approaches learn the distribution of real data and subsequently generate adversarial examples that closely resemble original samples yet are capable of deceiving the target AI system. The adversarial transformation network (ATN), the generative adversarial perturbation (GAP) model, and AddDiffuser are three classic generation-based attack methods^[23–25].

The ATN was the first method to employ a generative model for producing adversarial examples^[23]. It first uses an autoencoder to convert input samples into adversarial perturbations, then superimposes these perturbations onto the original clean images. A loss function is utilized to guide the generator in performing targeted attacks. The GAP represents a further development based on ATN^[24]. It enhances the network architecture by adopting a U-Net structure to generate adversarial perturbations. Additionally, this method decouples the perturbation magnitude constraint from the adversarial attack loss, allowing more fine-grained control over perturbation generation while optimizing adversarial effectiveness. AddDiffuser is a novel approach that utilizes a diffusion model to generate adversarial examples^[25]. This method guides the latent code into the adversarial example space of a specific classifier through perturbed predicted images and employs adversarial repair based on class activation mapping to preserve salient regions of the image while perturbing less important areas.

Compared to the other two categories of methods, generation-based attacks do not require repeated access to the target system, resulting in higher attack efficiency. This makes them more suitable for generating adversarial examples in large quantities.

4. Adversarial defenses

To counter diverse and potent adversarial attacks, numerous defense methods have been successively proposed to enhance the robustness of AI systems when facing anomalous inputs. Current adversarial defense designs for AI systems are primarily implemented during two stages: model training-oriented and model access-oriented defenses.

4.1. Model training-oriented defenses

The model training-oriented defense methods refer to techniques applied during the model training stage, where adversarial examples are introduced as input. This enables the model to learn characteristics of adversarial perturbations, thereby equipping the final trained model with inherent capability to recognize adversarial samples and consequently improving system robustness. Such methods primarily include defense paradigms such as adversarial training, robust pre-training, and defensive distillation.

4.1.1. Adversarial training

These methods operate in a self-play manner by incorporating adversarial examples into the training dataset, enabling the model to learn how to resist adversarial attacks. Specifically, during each training epoch, adversarial examples are generated against the current model and included in training. This approach ensures that while the model's robustness is progressively enhanced through each iteration, it also continuously faces increasingly potent adversarial examples, thereby systematically improving the overall robustness of the system. Classical adversarial training methods include ensemble adversarial training, cascaded adversarial training, and projected gradient descent adversarial training^[17,19,26]. (see **Figure 3**)

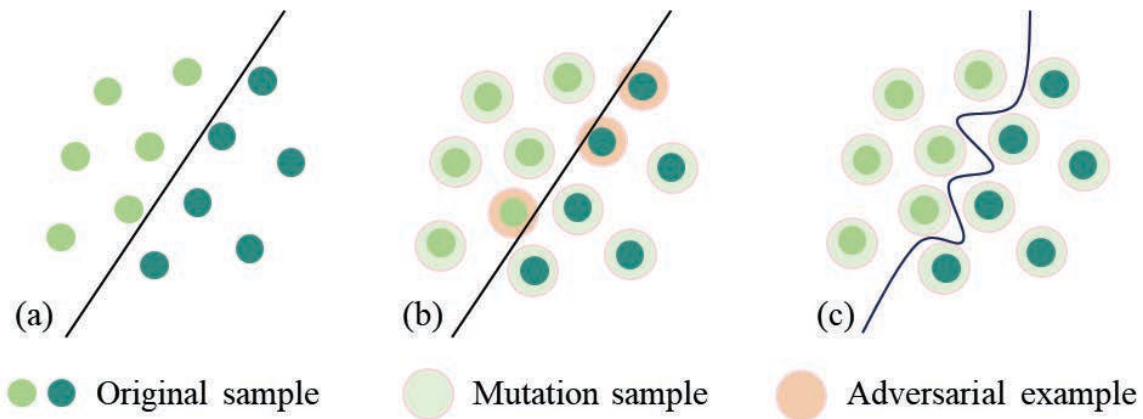


Figure 3. Principle of Adversarial Training. (a) demonstrates the result of the original model on clean samples. (b) shows the result of the original model on perturbed samples, some of which are adversarial examples. (c) illustrates the result of the adversarially trained model on perturbed samples, which can effectively recognize variations and remains robust against adversarial attack.

4.1.2. Robust pre-training

This category of defense methods draws inspiration from the pre-training-fine-tuning paradigm by designing a robust pre-training followed by adversarial fine-tuning strategy. During the pre-training phase, the model is trained on large-scale, diverse datasets to learn more generalized and robust feature representations. Subsequently, the pre-trained model is fine-tuned using adversarial examples to enhance its robustness against such attacks. Compared

to traditional adversarial training, the robust pre-training and adversarial fine-tuning paradigm significantly improves both the model's performance on clean samples and its robustness against adversarial examples. Classical robust pre-training methods include dual-path adversarial contrastive pre-training and adversarial contrastive learning^[27,28]. The key to robust pre-training lies in designing superior pre-training strategies that enable the model to learn more effective pre-trained features, thereby achieving better performance in the subsequent adversarial fine-tuning stage^[28].

4.1.3. Defensive distillation

Defensive distillation draws upon the concept of knowledge distillation, where knowledge is transferred from a pre-trained teacher model to a student model, enabling the student model to maintain recognition capability when confronted with adversarial examples. Classical defensive distillation methods include AGKD-BML and guided adversarial contrastive distillation (GACD)^[29,30].

4.2. Model access-oriented defenses

The model access-oriented defense methods refer to techniques applied to AI systems that lack defensive mechanisms incorporated during the training phase. These methods process either the input samples or the model itself during the system's access (deployment) stage, thereby enabling models originally incapable of handling adversarial examples to function normally. This category primarily includes adaptive defense and detection-based defense approaches.

4.2.1. Adaptive defense

Adaptive defense methods refer to techniques that process either the input data or the model architecture to mitigate attacks from adversarial examples. These can be further categorized into data-adaptive and model-adaptive defense methods. Data-adaptive defense methods involve preprocessing input samples using statistical approaches or prior knowledge to eliminate potential adversarial perturbations, such as image compression and image denoising^[31,32]. However, since these methods apply uniform processing to all input samples, they inevitably degrade the model's performance on clean data while enhancing robustness against adversarial attacks. Model-adaptive defense methods involve modifying the model structure to defend against adversarial examples, such as the supergrid method^[33]. Since these approaches require retraining the model after architectural modifications, they generally entail higher computational costs and lower efficiency.

4.2.2. Detection-based defense

Detection-based defense methods treat adversarial examples as anomalous inputs, detecting and subsequently rejecting them from processing^[34,35]. The core of these approaches lies in accurately identifying adversarial samples. Based on the detection techniques employed, they can be further classified into statistical-discrepancy-based detection methods and statistical-modeling-based detection methods. The former involves extracting features from input samples using statistical methods and detecting adversarial examples based on discrepancies in these features. Commonly used statistical measures include density ratio, mutual information, and activation function invariance. The latter refers to modeling the differences between genuine samples and adversarial examples using techniques such as deep learning models such as binary classification detectors, followed by detection. The advantage of detection-based defense methods is that they require no modifications to the input data or the target model, resulting in relatively lower complexity.

5. Conclusion

Research on adversarial attacks against AI systems focuses on developing efficient algorithms to generate adversarial examples in various scenarios, thereby revealing vulnerabilities in embedded deep learning models and other AI components. On the other hand, adversarial defense research aims to mitigate the harm caused by such attacks by enhancing model robustness, thereby advancing the development of secure and reliable artificial intelligence. This paper surveys current research on the robustness of AI systems from the perspectives of adversarial attacks and defenses. It reveals that as the interplay between attacks and defenses evolves, their respective technical capabilities continue to deepen, significantly promoting progress in this field. Through a systematic investigation, we identify several promising research directions. For adversarial attacks, future work could focus on:

(1) Multimodal adversarial attacks

Current research predominantly targets image data, with relatively less attention to text, audio, and other data types. Developing effective multimodal attack techniques based on existing foundations remains a challenge.

(2) High-efficiency and low-cost adversarial attacks

Effective attack methods should achieve high success rates while maintaining low computational cost. These objectives often conflict, making their balanced realization a significant challenge.

For adversarial defenses, beyond developing countermeasures against emerging attack methods, greater emphasis should be placed on the generalization capability of defense strategies. Customized defenses may struggle against rapidly evolving attack techniques, underscoring the need for versatile and broadly applicable adversarial defense methods.

The analysis and discussion presented in this paper can provide valuable insights to help researchers develop their own methods and tools for AI system adversarial attacks and defenses, ultimately contributing to the creation of more secure and trustworthy artificial intelligence for end-users.

Disclosure statement

The author declares no conflict of interest.

Reference

- [1] Zhao J, Zhao W, Deng B, et al., 2024, Autonomous Driving System a Comprehensive Survey. *Expert Systems with Applications*, 242: 122836.
- [2] Al Kuwaiti A, Nazer K, Al-Reedy A, et al., 2023, A Review of the Role of Artificial Intelligence in Healthcare. *Journal of Personalized Medicine*, 13(6): 951.
- [3] Pei K, Cao Y, Yang J, et al., 2017, DeepXplore: Automated Whitebox Testing of Deep Learning Systems. *Proceedings of the 26th Symposium on Operating Systems Principles*: 1–18.
- [4] Szegedy C, Zaremba W, Sutskever I, et al., 2013, Intriguing Properties of Neural Networks. *arXiv Preprint arXiv:1312.6199*.
- [5] Hamon R, Junklewitz H, Sanchez I, 2020, Robustness and Explainability of Artificial Intelligence. *Publications Office of the European Union*, 207(40): 1–40.
- [6] Javed H, El-Sappagh S, Abuhmed T, 2024, Robustness in Deep Learning Models for Medical Diagnostics: Security

and Adversarial Challenges Towards Robust AI Applications. *Artificial Intelligence Review*, 58(1): 12.

- [7] Tocchetti A, Corti L, Balayn A, et al., 2025, AI Robustness: A Human-Centered Perspective on Technological Challenges and Opportunities. *ACM Computing Surveys*, 57(6): 1–38.
- [8] Wang Y, Sun T, Li S, et al., 2023, Adversarial Attacks and Defenses in Machine Learning-Empowered Communication Systems and Networks: A Contemporary Survey. *IEEE Communications Surveys & Tutorials*, 25(4): 2245–2298.
- [9] Xu H, Mannor S, 2012, Robustness and Generalization. *Machine Learning*, 86(3): 391–423.
- [10] Xu H, Ma Y, Liu H, et al., 2020, Adversarial Attacks and Defenses in Images, Graphs and Text: A Review. *International Journal of Automation and Computing*, 17(2): 151–178.
- [11] Goyal S, Doddapaneni S, Khapra M, et al., 2023, A Survey of Adversarial Defenses and Robustness in NLP. *ACM Computing Surveys*, 55(14s): 1–39.
- [12] Zhang W, Sheng Q, Alhazmi A, et al., 2020, Adversarial Attacks on Deep-Learning Models in Natural Language Processing: A Survey. *ACM Transactions on Intelligent Systems and Technology*, 11(3): 1–41.
- [13] Carlini N, Wagner D, 2018, Audio Adversarial Examples: Targeted Attacks on Speech-to-Text. *IEEE Security and Privacy Workshops*: 1–7.
- [14] Deldjoo Y, Noia T, Merra F, 2021, A Survey on Adversarial Recommender Systems: From Attack and Defense Strategies to Generative Adversarial Networks. *ACM Computing Surveys*, 54(2): 1–38.
- [15] Shayegani E, Mamun M, Fu Y, et al., 2023, Survey of Vulnerabilities in Large Language Models Revealed by Adversarial Attacks. *arXiv Preprint arXiv:2310.10844*.
- [16] Goodfellow I, Shlens J, Szegedy C, 2014, Explaining and Harnessing Adversarial Examples. *arXiv Preprint arXiv:1412.6572*.
- [17] Tramèr F, Kurakin A, Papernot N, et al., 2017, Ensemble Adversarial Training: Attacks and Defenses. *arXiv Preprint arXiv:1705.07204*.
- [18] Kurakin A, Goodfellow I, Bengio S, 2018, Adversarial Examples in the Physical World. *Artificial Intelligence Safety and Security*. Chapman and Hall/CRC: 99–112.
- [19] Madry A, Makelov A, Schmidt L, et al., 2017, Towards Deep Learning Models Resistant to Adversarial Attacks. *arXiv Preprint arXiv:1706.06083*.
- [20] Dong Y, Liao F, Pang T, et al., 2018, Boosting Adversarial Attacks with Momentum. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*: 9185–9193.
- [21] Carlini N, Wagner D, 2017, Towards Evaluating the Robustness of Neural Networks. *IEEE Symposium on Security and Privacy*: 39–57.
- [22] Feng S, Feng F, Xu X, et al., 2021, Digital Watermark Perturbation for Adversarial Examples to Fool Deep Neural Networks. *International Joint Conference on Neural Networks*: 1–8.
- [23] Baluja S, Fischer I, 2017, Adversarial Transformation Networks: Learning to Generate Adversarial Examples. *arXiv Preprint arXiv:1703.09387*.
- [24] Poursaeed O, Katsman I, Gao B, et al., 2018, Generative Adversarial Perturbations. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*: 4422–4431.
- [25] Chen X, Gao X, Zhao J, et al., 2023, AdvDiffuser: Natural Adversarial Example Synthesis with Diffusion Models. *IEEE/CVF International Conference on Computer Vision*: 4562–4572.
- [26] Na T, Ko J, Mukhopadhyay S, 2017, Cascade Adversarial Machine Learning Regularized with a Unified Embedding. *arXiv Preprint arXiv:1708.02582*.

- [27] Hendrycks D, Lee K, Mazeika M, 2019, Using Pre-Training Can Improve Model Robustness and Uncertainty. International Conference on Machine Learning: 2712–2721.
- [28] Jiang Z, Chen T, Chen T, et al., 2020, Robust Pre-Training by Adversarial Contrastive Learning. Advances in Neural Information Processing Systems, 33: 16199–16210.
- [29] Wang H, Deng Y, Yoo S, et al., 2021, AGKD-BML: Defense Against Adversarial Attack by Attention-Guided Knowledge Distillation and Bi-Directional Metric Learning. IEEE/CVF International Conference on Computer Vision: 7658–7667.
- [30] Bai T, Zhao J, Wen B, 2023, Guided Adversarial Contrastive Distillation for Robust Students. IEEE Transactions on Information Forensics and Security, 19: 9643–9655.
- [31] Guo C, Rana M, Cisse M, et al., 2017, Countering Adversarial Images Using Input Transformations. arXiv Preprint arXiv:1711.00117.
- [32] Liao F, Liang M, Dong Y, et al., 2018, Defense Against Adversarial Attacks Using High-Level Representation Guided Denoiser. IEEE Conference on Computer Vision and Pattern Recognition: 1778–1787.
- [33] Bian H, Chen D, Zhang K, et al., 2021, Adversarial Defense via Self-Orthogonal Randomization Super-Network. Neurocomputing, 452: 147–158.
- [34] Alotaibi A, Rassam M, 2023, Adversarial Machine Learning Attacks Against Intrusion Detection Systems: A Survey on Strategies and Defense. Future Internet, 15(2): 62.
- [35] Aldahdooh A, Hamidouche W, Fezza S, et al., 2022, Adversarial Example Detection for DNN Models: A Review and Experimental Comparison. Artificial Intelligence Review, 55(6): 4403–4462.

Publisher's note

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

Design and Implementation of a Vertical Search Engine for the Fisheries Domain

Zhiqiang Zhang*

School of Artificial Intelligence, Zhejiang Dongfang Polytechnic, Wenzhou 325000, Zhejiang, China

*Corresponding author: Zhiqiang Zhang, cqczd123@126.com

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited

Abstract: This paper aims to address the issues of insufficient professionalism and precision in information retrieval within the fisheries domain by designing and implementing a vertical search engine specifically for fisheries. The system employs a specialized structure, integrating a fisheries terminology-based word segmentation algorithm, a query expansion mechanism based on knowledge graphs, and a multi-source fisheries data collection and processing workflow. It establishes a domain knowledge framework that includes various entities such as fishing gear, fish species, fishing grounds, and legal regulations. The system enhances text representation and retrieval relevance by using word segmentation techniques that combine mutual information and left-right entropy, as well as TF-IDF weighting and the vector space model. Experiments show that the system's response time for retrieving information from fisheries-specific databases is within 1 second, which is a significant improvement compared to traditional search engines. The system demonstrates good domain adaptability and practical value.

Keywords: Fisheries domain; Information retrieval; Vertical search engine; Word segmentation algorithm; Knowledge graph

Online publication: February 27, 2026

1. Introduction

The development of information technology and the growth of internet data have made the efficient and accurate acquisition of knowledge in a specific field the core of industry intelligence and digital transformation. The information ecosystem of the fisheries sector is becoming increasingly complex and diverse. Currently, industry professionals rely heavily on general-purpose search engines to obtain specialized knowledge^[1]. However, these search engines have significant shortcomings in the fisheries domain. On one hand, their precision is low, as the search results are often cluttered with irrelevant or low-quality commercial information, making it difficult to find authoritative and accurate content. On the other hand, the organization of information is lacking, as the results are not displayed according to the logical classification within the fisheries domain. This hinders the speed

of information acquisition and negatively impacts production decision-making, technology dissemination, and industry management ^[2].

This paper proposes a fisheries-specific word segmentation method that combines a domain dictionary with statistical models, effectively correcting improper segmentation of compound professional terms such as “fish-finding sonar”. It also establishes a fisheries knowledge graph and a parametric query analyzer that can accurately parse and expand structured query statements like “trawl mesh size ≥ 50 meters”. Furthermore, a specialized search framework has been developed to integrate fisheries information from various sources and perform real-time incremental indexing, enhancing the system’s search adaptability for multiple application scenarios, including equipment parameters, fishery conditions, and legal regulations, while maintaining search speed.

2. System architecture

2.1. System design

The structure of the collection system for the vertical search engine in the fisheries domain is shown in **Figure 1**. First, enter the system management interface to configure the parameters of the crawling task for the fisheries domain websites. The specific configuration process is omitted. The crawling process is controlled by the collection scheduling controller: when the collection scheduling controller is initialized, it loads the configured collection parameters, and at the same time, initializes the fisheries domain URL database, related logs, crawling scope restrictions, URL scheduler, processing pipeline, and thread pool used during the collection period. After starting the thread pool and the URL scheduler, the data collection task for the fisheries domain begins ^[3].

The URL scheduler provides the collection threads with links to be crawled, which are divided into two types: links in the to-be-crawled queue and links that have been crawled. The system uses an embedded database to store the links of fisheries-related web pages to be crawled and those that have been crawled separately. The collection threads obtain the content of fisheries web pages based on the provided links.

The crawling task is collaboratively completed by five processing pipelines: the pre-crawling pipeline is responsible for judging the crawling conditions and will only obtain web page content if it meets the specific crawling conditions of the fisheries domain, for example, it can adjust the priority for authoritative fisheries websites; the protocol processing pipeline is in charge of parsing various network transfer protocols; the content extraction pipeline is responsible for extracting the core content of fisheries-related web pages; the storage writing pipeline writes the extracted fisheries data into the storage system; and the link analysis pipeline identifies new fisheries-related links from the web page content and adds them to the URL scheduler for continuous crawling ^[4].

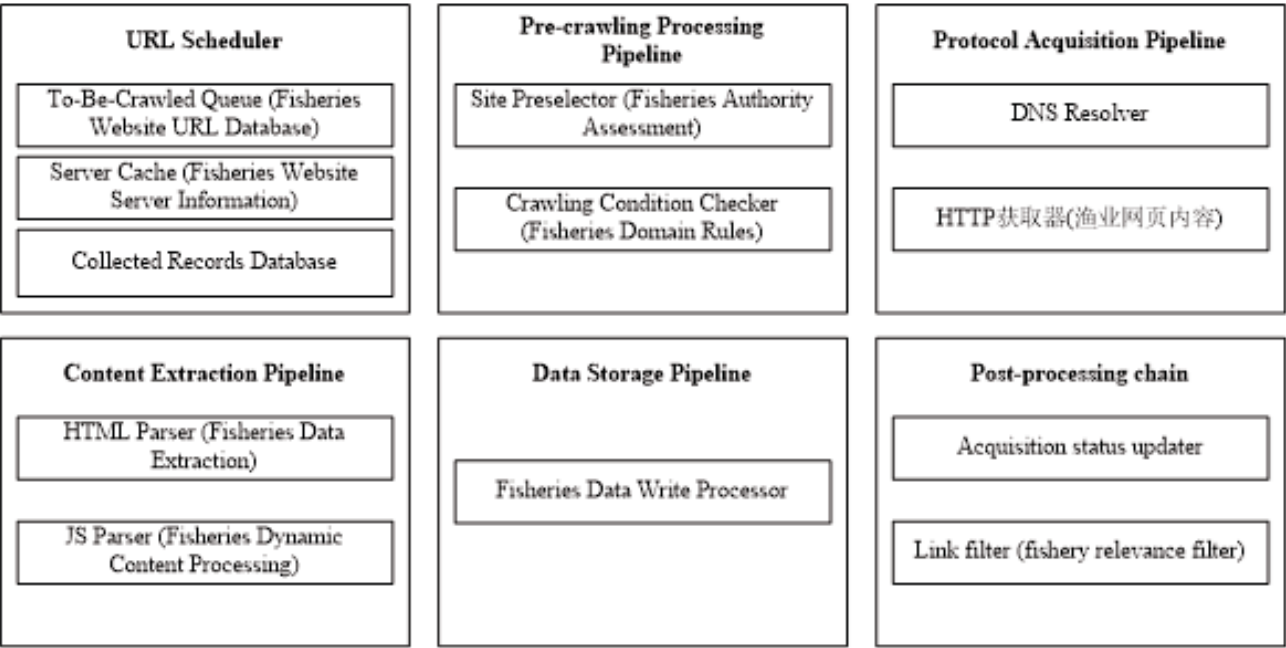


Figure 1. The structure of the collection system for the vertical search engine in the fisheries domain.

2.2. Main functional design

The specialized search engine for the fisheries industry adopts a professional system architecture design, and the establishment of a fisheries knowledge graph is a key part of this system. This section builds a knowledge graph based on domain ontology, using domain ontology modeling tools to determine entities such as fishing gear, fish species, fishing vessels, fishing grounds, and related laws and regulations, as well as their interrelationships. Moreover, an entity recognition engine that combines deep learning technology and rule-based methods is employed to accurately extract fisheries-specific terminology. Subsequently, a relationship discovery system is used to parse the semantic connections between entities, ultimately forming a structured knowledge system in the fisheries domain.

The system is equipped with an intelligent query understanding and processing module, which has been specifically optimized for the characteristics of fisheries queries. This module can identify the user’s query intent, such as queries for fishing gear parameters, searches for fishery information, or searches for policies and regulations. It also utilizes a fisheries thesaurus and knowledge graph to expand the scope of the query. The parametric query parser is particularly suitable for parsing query statements containing technical conditions, such as “trawl mesh size ≥ 50 meters”. The document processing and indexing section is responsible for standardizing and organizing various sources of fisheries data. It can interpret professional documents such as fishing vessel design drawings, catch statistics tables, and scientific research reports. Using fisheries spatiotemporal data extraction tools, it accurately obtains core data such as the location coordinates and operation times of fishing grounds, thereby establishing an inverted index system that adapts to the characteristics of industry terminology^[5]. Subsequently, the precise search and ranking mechanism, while considering the professionalism of terms, the relevance between texts, and the novelty of information, categorizes based on the credibility level of fisheries information sources and provides personalized ranking feedback to different user groups.

The document processing and indexing section is responsible for standardizing and organizing various sources of fisheries data. It can interpret professional documents such as fishing vessel design drawings, catch

statistics tables, and scientific research reports. Using fisheries spatiotemporal data extraction tools, it accurately obtains core data such as the location coordinates and operation times of fishing grounds, thereby establishing an inverted index system that adapts to the characteristics of industry terminology. Subsequently, the precise search and ranking mechanism, while considering the professionalism of terms, the relevance between texts, and the novelty of information, categorizes based on the credibility level of fisheries information sources and provides personalized ranking feedback to different user groups.

In terms of system integration, various data access interfaces have been designed to connect with domestic and international fisheries databases, Internet of Things monitoring data, and satellite remote sensing data. The system also integrates third-party location-based services, marine climate predictions, and market intelligence on aquatic products. Using standard interfaces, it enables data and service interaction with fisheries management systems, scientific research platforms, and corporate applications.

2.3. Design of the fisheries domain word segmentation plugin

Word segmentation is crucial for a vertical search engine specialized in the fisheries domain, as the method of segmentation directly affects the effectiveness of information retrieval related to fisheries. The word segmentation mechanism of the system described in this paper mainly encounters the following three technical challenges.

The built-in word segment of ElasticSearch has poor adaptability to Chinese professional terms in fisheries, often breaking compound words into individual characters. For example, the term “fish-finding sonar” is incorrectly divided into “probe”, “fish”, “sound”, and “sodium”.

Chinese word segmentation is inherently more complex than English segmentation, which can be done simply based on spaces or punctuation marks. In contrast, Chinese segmentation must utilize semantic and domain knowledge, especially in professional fisheries literature, where the difficulty is greater.

During the retrieval of fisheries catch equipment, when users query based on parameters, inaccurate segmentation often leads to matching irrelevant fisheries data, thereby affecting the precision of the search results.

In light of the above reasons, the system adopts a word segmentation method that combines mutual information and left-right entropy. Mutual information is used to evaluate the appropriateness of character combinations forming professional fisheries terms, while left-right entropy is employed to analyze the diversity of surrounding characters and words. A higher entropy value indicates that the term frequently appears in fisheries literature and is more likely to be an effective professional fisheries term.

In the vertical search engine for the fisheries industry, even after text segmentation, computers still cannot understand fisheries-related natural language as humans do. This requires transforming the text into a feature representation form^[6]. Using the Vector Space Model (VSM), fisheries-related texts can be converted into n -dimensional feature vectors, where each dimension of the vector represents a feature in the text, and each feature has a corresponding weight, that is:

$$D = D(t_1 w_1; t_2 w_2; \cdots; t_n w_n) \quad (1)$$

In Equation (1): The feature item t_n is the basic unit of fisheries text information, referring to words here, such as “trawl”, “fishing season” and “sonar” etc.; while the weight w_n is the measure of the importance of the word in the fisheries text D .

TF-IDF is a weighting technique used to evaluate the value of fisheries literature, aiming to quantify the importance of terms in professional documents within the inverted index. The importance of a word t_i in a specific fisheries document can be expressed by the following formula:

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (2)$$

In Equation (2), $n_{i,j}$ represents the frequency of word d_j appearing in the fisheries document d_j , while $\sum_k n_{k,j}$ is the total number of occurrences of all words in document d_j . The calculation method is as follows:

$$idf_i = \log \frac{|D|}{u_i + 1} \quad (3)$$

In Equation (3): $|D|$ represents the total number of documents in the fisheries domain; u_i represents the number of documents containing the term t_i . By multiplying the results of Equation (2) and Equation (3), the value of $TF - IDF_{i,j}$ is obtained. This value can be used to extract keywords from the inverted index of fisheries documents:

$$TF - IDF_{i,j} = tf_{i,j} \times idf_i \quad (4)$$

This method is conducive to identifying the uniqueness of fisheries-specific terms such as “encircling net fishing”, “catch statistics”, and “fishing vessel navigation” in the literature, thereby improving the search effectiveness of the vertical search engine.

Categorical search, as one of the core functions of the vertical search engine for the fisheries domain, is systematically and logically designed. **Figure 2** illustrates the process, reflecting the systematic and orderly nature of searching for fisheries-specific knowledge. Users first arrive at the categorical search page and select the primary fisheries category they wish to query. Based on the selected primary category, the system loads the relevant secondary classification directories and subsequently displays the fisheries professional database under that category. Users can decide whether to include subcategories, and the system will display the corresponding content according to the user’s selection.

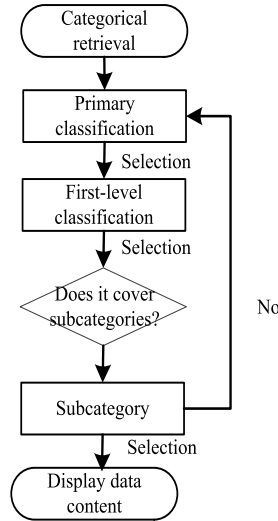


Figure 2. Process design.

At the database design level, the system employs MySQL for structured data storage, beginning with the design of an E-R diagram. The E-R model is shown in **Figure 3**, depicting the relationships between various fisheries entities within the system. The system encompasses equipment data for multiple fishing operation

methods, which are associated with specific categories such as deck machinery, fishing gear, and buoys. Each category has detailed attributes, including specific parameters and equipment codes. The entity relationships in the system reflect data structure associations in multiple aspects, such as user information, equipment data, and system management, ensuring the accuracy and completeness of fisheries data queries.

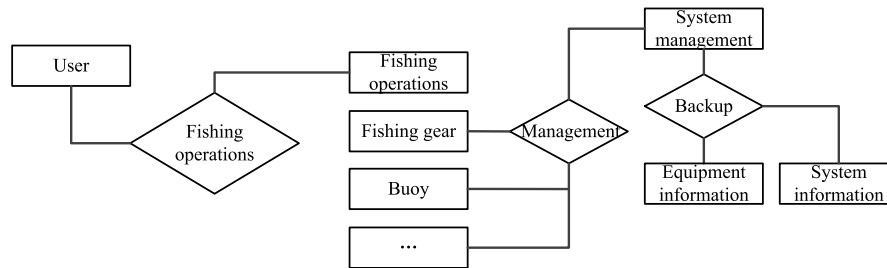


Figure 3. E-R model.

3. Experimental results analysis

Based on the technical solution proposed in this study, a prototype vertical search engine for the fisheries domain was developed. The experiments were conducted on a software platform comprising JDK 1.8, Tomcat 8.5, and Spring Boot 2.x, as well as a hardware environment equipped with an Intel Core i7-10700 processor (2.9GHz) and 16GB of memory. The system was tested on a dataset containing approximately 500,000 fisheries-related documents, including research papers, technical manuals, industry reports, policy regulations, and equipment parameters. A specialized index was built for this dataset, and corresponding retrieval services were provided. **Table 1** shows the differences in key modules of the search engine. The system has made significant improvements and enhancements in data source specificity, index structure optimization, fisheries domain word segmentation algorithms, and specialized query analysis compared to traditional search engines.

Table 1. Differences in key modules of the search engine

Key technologies	The fisheries vertical search engine designed in this paper	Traditional general-purpose search engines
Data sources	Supports multi-source data interfaces, enabling targeted collection from fisheries-specific databases, industry websites, and scientific literature	Mainly rely on general web crawlers, which cover a wide range but lack domain specificity
Indexing content	Intelligently filters and cleans fisheries data, indexing only domain-related structured and semi-structured information	Typically index all publicly available information on the web, lacking domain filtering mechanisms
Index structure	Adopts an inverted index optimized for fisheries terminology, supporting real-time incremental updates and historical version management	Mostly use traditional inverted indexes, with longer update cycles and no support for fine-grained incremental indexing
Word segmentation algorithms	Employs a Chinese word segmentation algorithm that integrates a fisheries domain dictionary with statistical models, supporting professional term recognition and new word discovery	Mainly depend on general word segmentation models, which have lower accuracy in recognizing fisheries-specific terms and compound words
Query analysis	Includes built-in interfaces for fisheries query semantic understanding and expansion, supporting parametric search, multimodal querying, and intent recognition	Usually based on keyword matching, with a single query method and lacking domain semantic understanding and expansion capabilities

To evaluate the performance of the fisheries vertical search engine, the information collection time and number of files for multiple typical fisheries data sources were statistically analyzed. The collection time and data scale of fisheries data sources are shown in **Table 2**. The data collection time is basically positively correlated with the data scale. For government information and scientific research databases, the system can complete the collection in a relatively short time, with the shortest being 137.438 seconds. However, for large-scale data sources such as the Global Fisheries Remote Sensing System, the collection time is correspondingly longer due to the large data volume and complex structure. The system has the capability to handle large-scale fisheries data, but there is still room for improvement in the collection efficiency of ultra-large-scale data sources. In the future, the system can be further improved through distributed collection and incremental update strategies.

Table 2. Collection time and number of files for fisheries data sources

Fisheries data sources	Information collection time / s	Number of files and documents / pieces
China fisheries government website	137.438	7339
Aquatic science database	235.140	30702
Distant water fisheries Information platform	510.375	52753
Fisheries equipment Technology library	1633.245	170057
Global fisheries remote Sensing monitoring system	3415.167	38 510

The fisheries keyword search time and result quantity are shown in **Table 3**. The test selected common high-frequency search keywords in the fisheries domain for querying. All search response times were within 1 second. For example, the search for “trawl” took 0.047 seconds and yielded 1,235 results, indicating that the system can quickly respond to users’ professional search needs. In terms of search time, the system-maintained millisecond-level responses for different keywords, demonstrating the effectiveness of the index structure. The search for “fishing vessel main engine” took relatively longer, possibly due to the involvement of numerous technical parameters and subcategories, which required more semantic expansion and relevance calculations during the search process. The number of results reflects the prevalence of the terms within the domain. “Trawl” and “fishing moratorium” are core operational methods and policy keywords, covering the most documents, which is consistent with the actual situation in the fisheries domain. Overall, the search engine meets the real-time query needs of fisheries professionals in terms of both search speed and result coverage.

Table 3. Fisheries keyword search time and result quantity

Search keywords	Search time / s	Number of search results / Entries
Trawl	0.047	1235
Tuna	0.078	892
Fishing vessel main engine	0.281	456
Sonar fish detection	0.078	678
Fishing moratorium	0.041	1024

4. Conclusion

This paper constructs a vertical search engine for the fisheries domain, realizing a full-process solution from data collection, specialized word segmentation, knowledge construction to efficient retrieval. The experiment was conducted on a dataset covering approximately 500,000 fisheries documents on the platform of JDK 1.8, Tomcat 8.5, and Spring Boot 2.x for performance verification. The test data shows that the search response time for typical fisheries keywords such as “trawl” and “tuna” is less than 0.2 s, among which the search for “trawl” only takes 0.047 s and returns 1,235 results, reflecting good real-time performance and recall ability. Comparative analysis indicates that the system outperforms traditional search engines in terms of data source specificity, index structure optimization, accuracy of domain word segmentation, and query semantic understanding. Although there is still room for improvement in the collection efficiency of ultra-large-scale data, the system has good domain applicability and retrieval performance, and can provide effective information support for fisheries research, production, and management, with promotion value and application prospects.

Funding

Scientific Research Fund of Zhejiang Provincial Education Department (2025 General Research Project): Design and Implementation of a Vertical Search Engine for the Fisheries Domain (Project No.: Y202559387)

Disclosure statement

The author declares no conflict of interest.

References

- [1] Hungevu R, Lawal A, Yinusa S, et al., 2025, The Impact of Low-Cost Technological Innovations on Sustainable Fisheries for Economic Development in Developing Countries. *World Journal of Advanced Research and Reviews*, 25(2): 1170–1184.
- [2] Sun C, Li X, Zou W, et al., 2025, Unequal Opportunities and Green Transition: A Study on the Mechanism of Green Total Factor Productivity Differences in China’s Regional Marine Economy. *Regional Studies in Marine Science*: 104249.
- [3] Liang Y, Zhu Y, Sun Z, et al., 2023, Feasibility Assessment of a CO₂-Based Power, Cooling, and Heating System Driven by Exhaust Gas from Ocean-Going Fishing Vessel. *Journal of Cleaner Production*, 406: 137058.
- [4] Priharanto Y, Yaqin R, Marjianto G, et al., 2023, Risk Assessment of the Fishing Vessel Main Engine by Fuzzy-FMEA Approach. *Journal of Failure Analysis and Prevention*, 23(2): 822–836.
- [5] Koričan M, Vladimir N, Fan A, 2023, Investigation of the Energy Efficiency of Fishing Vessels: Case Study of the Fishing Fleet in the Adriatic Sea. *Ocean Engineering*, 286: 115734.
- [6] Sathish T, Ağbulut Ü, George S, et al., 2023, Waste to Fuel: Synergetic Effect of Hybrid Nanoparticle Usage for the Improvement of CI Engine Characteristics Fuelled with Waste Fish Oils. *Energy*, 275: 127397.

Publisher’s note

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

RoboMirror: Bridging Human Intent and Robot Capability Through Self-Reflective Motion Adaptation

Lin Zhu¹, Longliang Huang¹, Chuxiong Lin^{2*}, Yujie Chen¹

¹Guizhou Equipment Manufacturing Polytechnic, Guizhou, China

²Beijing University of Technology, Beijing, China

*Corresponding author: Chuxiong Lin, [sy20040909@126.com](mailto:syt20040909@126.com)

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Current humanoid robot control paradigms place the burden of feasibility assessment on human operators, who must carefully design commands within perceived robot limitations. This constraint significantly hinders practical deployment and limits the expressiveness of robot behaviors. This study proposed an inverting paradigm: rather than constraining operator inputs, robots should autonomously evaluate their capacity to execute commanded motions and intelligently adapt references to align with their physical constraints and learned skills. This study introduced the Performance Prediction Network (PPN), a transformer-based architecture that forecasts execution quality for arbitrary reference trajectories by analyzing both the commanded motion sequence and current robot state. Given a high-level task specification, our framework synthesizes multiple viable motion candidates and employs PPN to rank them across six dimensions: collision avoidance, kinematic feasibility, dynamic stability, trajectory smoothness, and goal satisfaction. This ranking enables autonomous selection of the most suitable reference motion before execution begins. Our complete system integrates motion generation, kinematic retargeting, and learned control policies with PPN-guided adaptation, creating a closed-loop framework where robots reason about their own limitations. Validated on 100,000 diverse human motions span walking, running, jumping, and acrobatic maneuvers, PPN achieves 99.14% accuracy in predicting imminent failures while maintaining low prediction error across all performance metrics. In deployment, our system successfully prevents 62% of anticipated falls by autonomously modifying commanded references, demonstrating that explicit capability modeling enables safer and more reliable humanoid control without sacrificing behavioral diversity.

Keywords: Deep reinforcement learning; Physical self-awareness; Safe motion planning; Failure prediction; Human-robot imitation

Online publication: February 12, 2026

1. Introduction

Humanoid robots are increasingly deployed in human-centered environments where commands are expressed

through high-level social intent ^[1,2]. However, bridging the gap between intent-rich commands and a robot’s physical capabilities remain a significant challenge ^[3–5]. Traditional control paradigms often impose a cognitive burden on operators, requiring them to internalize kinematic and stability limits to craft executable commands ^[6,7]. When reference motions slightly exceed a robot’s feasible envelope, contemporary whole-body imitation systems often exhibit brittle behavior or failure, primarily due to the absence of self-evaluative mechanisms that reason about execution before action ^[8,9]. This study proposes a paradigm inversion: robots should proactively evaluate and adapt commanded behaviors to their own capabilities. This capability-aware view addresses three interconnected challenges: learning predictive models that generalize across diverse maneuvers, accounting for dynamic executability under real-world physic, and maintaining real-time latency for behavior selection ^[10–20]. Unlike approaches that restrict behavioral expressiveness to ensure safety, this study advocates for retaining open-ended commands while shifting the responsibility for feasibility to the robot via learned, predictive self-assessment ^[21–24]. Our framework instantiates this paradigm by forecasting execution quality to select optimal behavior candidates. At its core is the Performance Prediction Network (PPN), a transformer-based architecture that jointly encodes commanded reference trajectories and the robot’s current state to predict multi-dimensional quality metrics, such as fall likelihood and tracking accuracy. This allows the system to rank multiple candidates, generated from text-to-motion models, and select the most viable trajectory before committing to control. The primary contributions of this work include:

- (1) A capability-aware control paradigm
Inverts the feasibility burden, allowing robots to evaluate and adapt behaviors before execution.
- (2) The performance prediction network (PPN)
A transformer model for forecasting multi- dimensional execution quality based on reference trajectories and robot state.
- (3) An open-ended intent pipeline
Integrates high-level text-to-motion generation with learned, pre-execution ranking and selection.
- (4) Real-time integration
A physics-based whole-body controller, enabling a low-latency assessment selection loop for responsive adaptation.

2. Related works

2.1. Imitation learning

Imitation learning (IL) enables robots to acquire skills from demonstrations without manual reward engineering ^[25–29]. While Behavior Cloning (BC) offers computational efficiency, it suffers from co- shift and compounding errors ^[27,28,30–32]. Inverse Reinforcement Learning (IRL) provides robustness by inferring reward functions but at a higher computational cost ^[27, 33–36]. Recent address failure by constraining policies to expert manifolds. Unlike these approaches that enforce strict adherence to demonstrations, our work builds on BC but relaxes imitation constraints; this study posit that agents should execute tasks optimally within their specific embodiment limits rather than perfectly replicating human motion. See **Figure 1**.

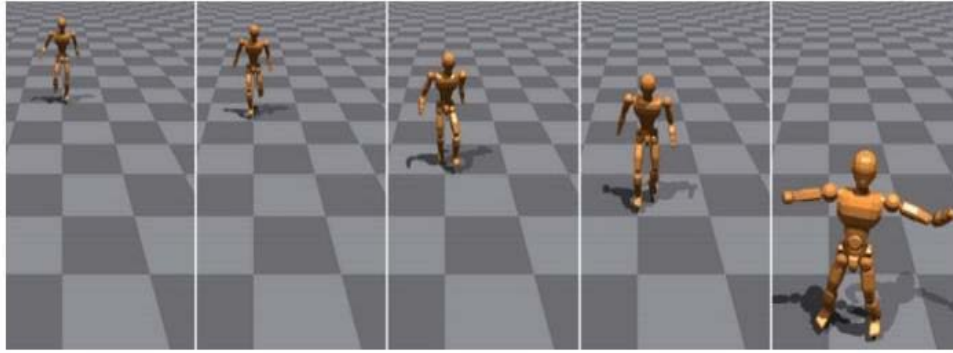


Figure 1. When completing a task, the robot engages in self-reflection to select the optimal plan. For example, when need to reach a certain location, it chooses to walk, thereby avoiding the risks associated with running or jumping.

2.2. Humanoid control

Traditional humanoid control relies on predefined patterns and model-based methods, which often struggle in unpredictable environments ^[22,37–40]. Reinforcement learning (RL) has emerged as a robust alternative for bipedal locomotion, achieving zero-shot sim-to-real transfer on platforms like Cassie ^[8,41–44]. Advanced systems such as I-CTRL have extended whole-body imitation to over 7410,000 motions by constraining exploration to ensure visual resemblance. However, most existing systems blindly pursue high reference fidelity, leading to failure when commanded motions exceed the robot’s capabilities. Current mitigation strategies often involve filtering complex behaviors (e.g., acrobatics), which limits expressiveness. This study proposes a “capability assessment” mechanism: robots should anticipate execution outcomes and autonomously relax reference constraints when risks are detected.

2.3. Self-awareness

In robotics, physical self-awareness involves monitoring discrepancies between planned movements and current states. Our performance Prediction Network (PPN) draws inspiration from this by continuously analyzing the gap between human references and robot states. Fall prevention is a critical subset of this capability. While early model-based methods used simplified inverted pendulum or ZMP models, they lack generalization to dynamic motions ^[14,38]. Recent learning-based methods using LSTMs or 1D-CNNs address these limits but remain restricted to simple movements like walking ^[14,38,45–47]. In contrast, this study leverages I-CTRL to train PPN on a diverse spectrum of human movements. By incorporating the intended reference motion as an input, not just the current state, our system can proactively adapt a high-risk command (e.g., a high jump) into a feasible one before failure occurs ^[8,44,48–50].

3. Methodology

This study presents a capability-aware motion adaptation pipeline that maps high-level commands c to safe robot motions R_{pb} by interposing human embodiment and predictive self-assessment. The system operates in three stages.

- (1) Synthesizing intent consistent human references H , via MotionLCM(f)
- (2) Retargeting references to robot space R , via ImitationNet (9h2r)
- (3) Refining trajectories into physics consistent motion RP via I-CTRL(gr2p), see **Figure 2**.

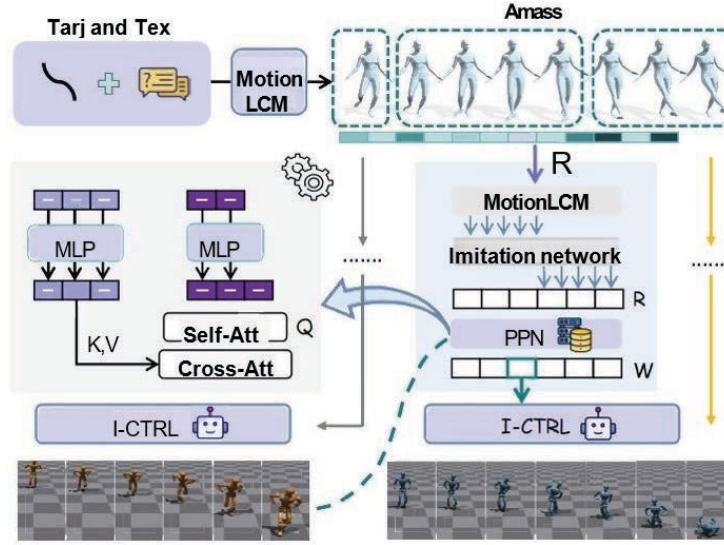


Figure 2. Overview of the motion adaptation system. Given a command, the system generates diverse motion candidates and ranks them using a PPN based on the robot’s physical capabilities and current state. The highest-ranked motion is executed by the low-level controller.

3.1. Problem formulation

This study formulates humanoid control as a motion adaptation problem. Given a task c , this study generates motion R_p that respects kinematic and dynamic constraints through the mapping.

$$c \xrightarrow{f} \mathbf{H}_r \xrightarrow{g_{h2r}} \mathbf{R}_r \xrightarrow{g_{r2p}} \mathbf{R}_p \quad (1)$$

Human motion is respected as $\mathbf{H}_r = \{\mathbf{h}_r^t\}_{t=1}^T \in \mathbb{R}^{T \times J \times 3}$, while robot reference R , and physics-based motion R_p include root states (p^*, θ_t) and joint configurations (q_6, q_t). This study’s key insights were to discover an adapted reference H , that maximizes quality Q while maintaining safety S above a threshold T_{safe}

$$\hat{\mathbf{H}}_r = \arg \max_{\mathbf{H}'_r} \mathcal{Q}(g(\mathbf{H}'_r), c) \text{ s.t. } S(g(H)) > T_{safe} \quad (2)$$

3.2. Motion adaptation framework

The system operates in a receding horizon. At each step t , this study consider observed states $\mathbf{R}^t \mathbf{o}$ and the future reference poses $\mathbf{H}^t \mathbf{f}$. Candidate generation: This generate and diverse candidate $\{\mathbf{H}_{f,i}^t\}_{i=1}^n$ using MotionLCM to provide multiple behavioral alternatives (e.g. walking vs running). Selection: Each candidate is retargeted via $gh2r$ and evaluated by the Performance Prediction Network (PPN): The optimal index i^* was selected by prioritizing safety lexicographically, then maximizing quality $\mathbf{w}^\top \mathbf{s}^i$.

$$\hat{\mathbf{S}}^i = \text{PPN}(\hat{\mathbf{R}}_{f,i}^t, \mathbf{R}^t \mathbf{o}) \quad (3)$$

3.3. Performance prediction network

PPN is a transformer based architecture that forecasts execution quality Encoding: Reference motion $\hat{\mathbf{R}}_{f,i}^t$ and observed states $\mathbf{R}^t \mathbf{o}$ are encoded via MLPs into $\mathbf{E}^t \mathbf{f}$ and $\mathbf{E}^t \mathbf{o}$. This study appends a [cls] token to $\mathbf{E}^t \mathbf{o}$ and apply self-attention to capture temporal dependencies. Conditioning: Cross attention allows observed states to attend to reference features:

$$\mathbf{E}^t\mathbf{c} = \text{CrossAttn}(\text{Query} = \hat{\mathbf{E}}^t\mathbf{c}, \text{Key} = \mathbf{E}^t\mathbf{f}, \text{Value} = \mathbf{E}^t\mathbf{f}) \quad (4)$$

Score prediction: The context vector $\mathbf{z}\mathbf{c} = \mathbf{E}^t\mathbf{c}^{[0]}$ was project to predict $\hat{\mathbf{S}}^t\mathbf{i} = [\text{dfall}, \hat{A}_q, \hat{A}_q^-, \hat{A}_q^+, \hat{A}_p, \hat{A}_\theta]^\top$, quantifying fall probability, alignment errors, and smoothness. Training: The objective is $L = L_{\text{fall}} + \lambda_1 L_{\text{align}} + \lambda_2 L_{\text{smooth}}$, using binary cross entropy for L_{fall} and MSE for alignment and smoothness.

4. Experiments

This study validated our capability aware motion adaption framework using 85,000 human motion sequences and 255,000 robot trajectories. Our evaluation focuses on

- (1) PPN accuracy across multiple time horizons
- (2) Adaptation effectiveness in preventing failures
- (3) Architectural ablation studies

4.1. Experimental setup

Dataset and Platform. This study generated 85,000 sequences from 8,500 textual prompts using MotionLCM, covering locomotion, dynamic actions, and expressive gestures. Robot executions were simulated using the JVRC-1 model (23 DOF) in IsaacGym via the I-CTRL controller. The test set was balanced with 50% fall samples (e.g. jumps > 45 cm or rapid turns) to ensure discriminative power. Simulations ran at 60 Hz with $K_p = 100$ and $K_d = 10$. Metrics. This study evaluate Fall Prediction Accuracy, Alignment MSE ($A_q, A_q^-, A_p, A_\theta$), and Smoothness MSE (\hat{A}_q^-). Adaptation is measured by Fall Prevention Rate, Task Completion rate, and trajectory deviation.

4.2. Performance prediction accuracy

Table 1 summarizes the performance of PPN and its variants. Our full model achieves a 99.14% fall prediction accuracy at a 1s horizon.

Table 1. Performance prediction accuracy on test set

Model	Tf	Fall Acc. \uparrow	$A_q^- \downarrow$	$A_q \downarrow$	$A_q^+ \downarrow$	$A_p \downarrow$	$A_\theta \downarrow$
w/o Rf	1s	96.85	0.108	0.0289	8.147	0.1253	0.0521
w/o Ro	1s	98.73	0.094	0.0417	6.382	0.1142	0.0298
w/o Cross-Attn	1s	98.91	0.086	0.0183	5.874	0.1067	0.0264
PPN (Ours)	1s	99.14	0.078	0.0159	5.138	1.004E-1	8.90E-3
PPN (Ours)	3s	98.93	0.051	0.0142	4.417	0.1158	0.0287

Analysis. Ablation results confirm that removing reference motion (**Rf**) causes the most degradation, with fall accuracy dropping by 2.29% and joint errors increasing by 81.8%. Removing observed states (**Ro**) primarily impact root pose accuracy. Compared to simple concatenation (w/o Cross-Attn), the cross-attention mechanism reduces orientation error by 197%, validating its efficacy in modeling state-reference interactions.

4.3. Motion adaptation effectiveness

On a test set of 420 failing commands, our framework achieved a 58.3% fall prevention rate, with 87.6% of adapted motions successfully completing the semantic task (**Table 2**).

Table 2. Motion adaptation performance on failing commands

Metric	Value
Fall prevention rate	58.3%
Task completion rate	87.6 %
Random selection fall prev.	19.8 %
Avg. adaptation time	0.21 s

Quantitative & Qualitative. PPN-guided selection provides a $2.9 \times$ improvement over random 146 selection. Qualitative analysis (**Figure 3**) demonstrates intelligent constraint relaxation: in the karate 147 kicks task, the system reduces kick height by 30% to ensure stability while maintaining the dynamic 148 character. In defend-punch, it shortens the lunge distance to preserve the center of mass while 149 executing the strike.

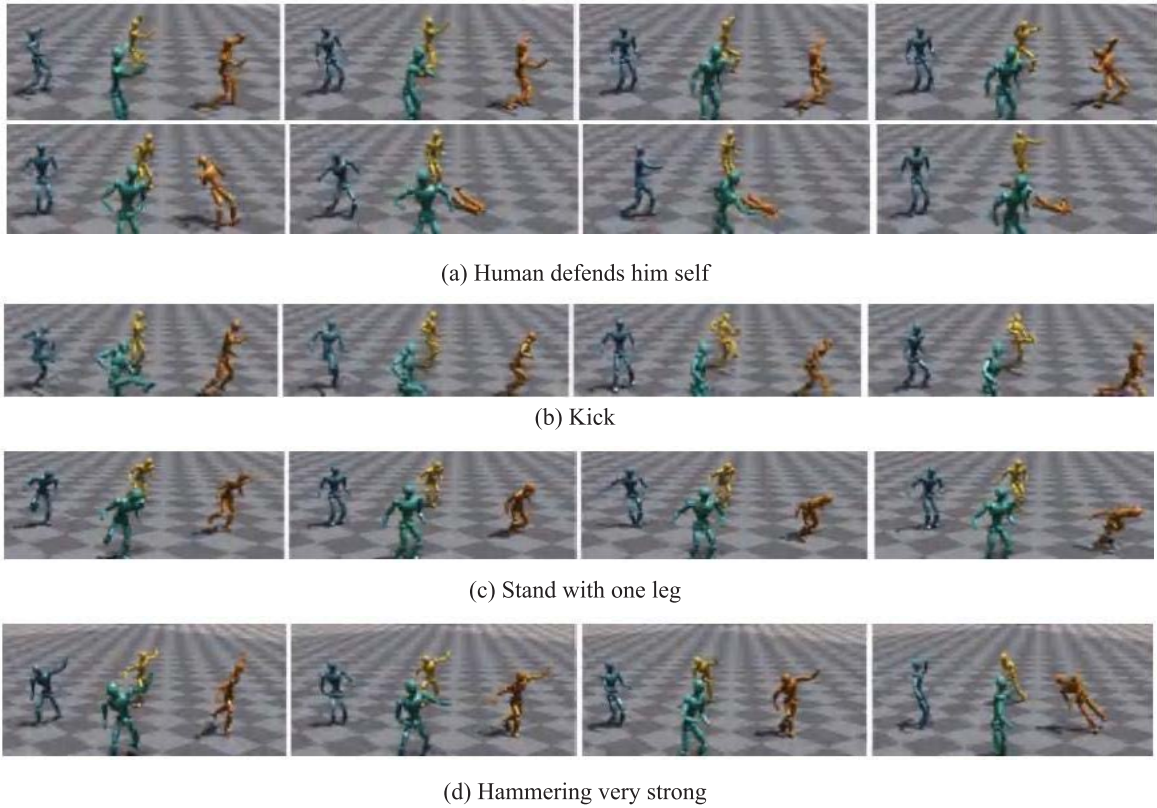


Figure 3. Qualitative examples of motion adaptation across diverse scenarios. Each row shows: (left) original failing motion, (middle) adapted motion selected by PPN, (right) comparison of root trajectories. Our framework intelligently modifies motion characteristics while preserving task semantics.

4.4. Efficiency and failure analysis

The full pipeline averages 0.21 s supporting 1–5 Hz real time planning. Failures are primarily due to insufficient candidate diversity (48%) and prediction errors (23%)

5. Conclusion

The PPN demonstrates that self-evaluation mechanisms can preemptively identify failures with 99.14% accuracy, allowing robots to transcend rigid constraints. While current results are simulation based, future work must address the sim to real gap, specifically sensor noise and actuator latency. The 58.3% prevention rate indicates significant potential for improvement by expanding candidate motion libraries and refining transition smoothness between original and adapted trajectories.

Disclosure statement

The authors declare no conflict of interest.

References

- [1] Dautenhahn K, 2007, Socially Intelligent Robots: Dimensions of Human–Robot Interaction. *Philosophical Transactions of the Royal Society B*, 362: 679–704.
- [2] Thomaz A, Cakmak M, 2016, Learning About Objects with Human Teachers. *Human–Robot Interaction*, 5: 1–42.
- [3] Khatib O, Sentis L, Park J, et al., 2008, Whole-Body Dynamic Behavior and Control of Human-Like Robots. *International Journal of Humanoid Robotics*, 5: 29–43.
- [4] Mainprice J, Hayne R, Berenson D, 2015, Predicting Human Reaching Motion in Collaborative Tasks Using Inverse Optimal Control and Iterative Replanning. *IEEE International Conference on Robotics and Automation*: 885–892.
- [5] Lasota P, Fong T, Shah J, 2017, A Survey of Methods for Safe Human–Robot Interaction. *Foundations and Trends in Robotics*, 5: 261–349.
- [6] Javdani S, Admoni H, Pellegrinelli S, et al., 2018, Shared Autonomy via Hindsight Optimization for Teleoperation and Teaming. *International Journal of Robotics Research*, 37: 717–742.
- [7] Dragan A, Lee K, Srinivasa S, 2013, Legibility and Predictability of Robot Motion. *IEEE International Conference on Human–Robot Interaction*: 301–308.
- [8] Peng X, Abbeel P, Levine S, et al., 2018, DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills. *ACM Transactions on Graphics*, 37: 1–14.
- [9] Atkeson C, Schaal S, 1997, Robot Learning from Demonstration. *International Conference on Machine Learning*: 12–20.
- [10] Hwangbo J, Lee J, Dosovitskiy A, et al., 2019, Learning Agile and Dynamic Motor Skills for Legged Robots. *Science Robotics*, 4: eaau5872.
- [11] Mordatch I, Todorov E, Popović Z, 2012, Discovery of Complex Behaviors through Contact-Invariant Optimization. *ACM Transactions on Graphics*, 31: 1–8.
- [12] Pratt J, Carff J, Drakunov S, et al., 2006, Capture Point: A Step Toward Humanoid Push Recovery. *IEEE-RAS International Conference on Humanoid Robots*: 200–207.
- [13] Stephens B, Atkeson C, 2010, Push Recovery by Stepping for Humanoid Robots with Force-Controlled Joints. *IEEE-RAS International Conference on Humanoid Robots*: 52–59.
- [14] Koolen T, de Boer T, Rebula J, et al., 2012, Capturability-Based Analysis and Control of Legged Locomotion, Part 1: Theory and Application. *International Journal of Robotics Research*, 31: 1094–1113.
- [15] Bretl T, Lall S, 2008, Testing Static Equilibrium for Legged Robots. *IEEE Transactions on Robotics*, 24: 794–807.
- [16] Hauser K, Bretl T, Latombe J, et al., 2008, Motion Planning for Legged Robots on Varied Terrain. *International Journal*

of Robotics Research, 27: 1325–1349.

- [17] Dai H, Valenzuela A, Tedrake R, 2014, Whole-Body Motion Planning with Centroidal Dynamics and Full Kinematics. IEEE-RAS International Conference on Humanoid Robots: 295–302.
- [18] Zucker M, Ratliff N, Dragan A, et al., 2013, CHOMP: Covariant Hamiltonian Optimization for Motion Planning. International Journal of Robotics Research, 32: 1164–1193.
- [19] Kalakrishnan M, Chitta S, Theodorou E, et al., 2011, STOMP: Stochastic Trajectory Optimization for Motion Planning. IEEE International Conference on Robotics and Automation: 4569–4574.
- [20] Tedrake R, Manchester I, Tobenkin M, et al., 2010, LQR-Trees: Feedback Motion Planning via Sums-of-Squares Verification. International Journal of Robotics Research, 29: 1038–1052.
- [21] Ott C, Roa M, Hirzinger G, 2008, Posture and Balance Control for Biped Robots Based on Contact Force Optimization. IEEE-RAS International Conference on Humanoid Robots: 26–33.
- [22] Sentis L, Khatib O, 2005, Synthesis of Whole-Body Behaviors through Hierarchical Control of Behavioral Primitives. International Journal of Humanoid Robotics, 2: 505–518.
- [23] Peng X, Guo Y, Halper L, et al., 2022, ASE: Large-Scale Reusable Adversarial Skill Embeddings for Physically Simulated Characters. ACM Transactions on Graphics, 41: 1–17.
- [24] Rempe D, Birdal T, Zhao Y, et al., 2021, HuMoR: 3D Human Motion Model for Robust Pose Estimation. IEEE International Conference on Computer Vision: 11488–11499.
- [25] Schaal S, 1999, Is Imitation Learning the Route to Humanoid Robots? Trends in Cognitive Sciences, 3: 233–242.
- [26] Argall B, Chernova S, Veloso M, et al., 2009, A Survey of Robot Learning from Demonstration. Robotics and Autonomous Systems, 57: 469–483.
- [27] Abbeel P, Ng A, 2004, Apprenticeship Learning via Inverse Reinforcement Learning. International Conference on Machine Learning: 1–8.
- [28] Pomerleau D, 1988, ALVINN: An Autonomous Land Vehicle in a Neural Network. Advances in Neural Information Processing Systems, 1: 305–313.
- [29] Ho J, Ermon S, 2016, Generative Adversarial Imitation Learning. Advances in Neural Information Processing Systems, 29: 4565–4573.
- [30] Ross S, Gordon G, Bagnell J, 2011, A Reduction of Imitation Learning and Structured Prediction to No-Regret Online Learning. International Conference on Artificial Intelligence and Statistics, 15: 627–635.
- [31] Mehta S, Ciftci Y, Ramachandran B, et al., 2024, Stable-BC: Controlling Covariate Shift with Stable Behavior Cloning. arXiv preprint arXiv:2408.06246.
- [32] Park J, Kim Y, Song K, et al., 2024, Mitigating Covariate Shift in Behavioral Cloning via Robust Stationary Distribution Correction. Advances in Neural Information Processing Systems, 37.
- [33] Ng A, Russell S, 2000, Algorithms for Inverse Reinforcement Learning. International Conference on Machine Learning: 663–670.
- [34] Ziebart B, Maas A, Bagnell J, et al., 2008, Maximum Entropy Inverse Reinforcement Learning. AAAI Conference on Artificial Intelligence: 1433–1438.
- [35] Wulfmeier M, Ondruska P, Posner I, 2015, Maximum Entropy Deep Inverse Reinforcement Learning. arXiv preprint arXiv:1507.04888.
- [36] Finn C, Levine S, Abbeel P, 2016, Guided Cost Learning: Deep Inverse Optimal Control via Policy Optimization. International Conference on Machine Learning, 48: 49–58.
- [37] Kajita S, Kanehiro F, Kaneko K, et al., 2003, Biped Walking Pattern Generation Using Preview Control of Zero-

Moment Point. IEEE International Conference on Robotics and Automation, 3: 1620–1626.

- [38] Vukobratovic M, Borovac B, 2004, Zero-Moment Point—Thirty-Five Years of Its Life. *International Journal of Humanoid Robotics*, 1: 157–173.
- [39] Kuffner J, Kagami S, Nishiwaki K, et al., 2002, Dynamically Stable Motion Planning for Humanoid Robots. *Autonomous Robots*, 12: 105–118.
- [40] Kuffner J, LaValle S, 2000, RRT-Connect: An Efficient Approach to Single-Query Path Planning. *IEEE International Conference on Robotics and Automation*: 995–1001.
- [41] Li Z, Cheng X, Peng X, et al., 2021, Reinforcement Learning for Robust Parameterized Locomotion Control of Bipedal Robots. *IEEE International Conference on Robotics and Automation*: 2811–2817.
- [42] Lee J, Hwangbo J, Wellhausen L, et al., 2020, Learning Quadrupedal Locomotion over Challenging Terrain. *Science Robotics*, 5: eabc5986.
- [43] Xie Z, Clary P, Dao J, et al., 2020, Learning Locomotion Skills for Cassie: Iterative Design and Sim-to-Real. *Conference on Robot Learning*, 100: 317–329.
- [44] Radosavovic I, Xiao T, Zhang B, et al., 2024, Real-World Humanoid Locomotion with Reinforcement Learning. *Science Robotics*, 9: edi9579.
- [45] Renner R, Behnke S, 2006, Instability Detection and Fall Avoidance for a Humanoid Using Attitude Sensors and Reflexes. *IEEE/RSJ International Conference on Intelligent Robots and Systems*: 2967–2973.
- [46] Zhong S, Gao J, Li M, et al., 2025, Fall Analysis and Prediction for Humanoids. *Robotics and Autonomous Systems*, 185: 104995.
- [47] Yang T, Zhang W, Yu Z, et al., 2018, Falling Prediction and Recovery Control for a Humanoid Robot. *IEEE-RAS International Conference on Humanoid Robots*: 481–487.
- [48] Mastalli C, Merkt W, Xin G, et al., 2024, Know Your Limits! Optimize the Robot’s Behavior through Self-Awareness. *arXiv preprint arXiv:2409.10308*.
- [49] Cheng X, Shi K, Agarwal A, et al., 2024, Extreme Parkour with Legged Robots. *arXiv preprint arXiv:2309.14341*.
- [50] Khamassi M, Lallée S, Enel P, et al., 2018, Toward Self-Aware Robots. *Frontiers in Robotics and AI*, 5: 88.

Publisher’s note

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

AI Large Model-Driven Adaptive Evolution of Brain-Computer Interface Chips: Technical Architecture, Challenges, and Future Directions

Borui Cui*

School of Mathematics, Tianjin University, Tianjin 300350, China

**Corresponding author: Borui Cui, infolklore2025@163.com*

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This paper focuses on how AI large models, such as Transformers and meta-learning can empower brain-computer interface (BCI) chips to achieve dynamic adaptation, thereby overcoming the limitations of traditional fixed decoding models that struggle to adapt to individual neural plasticity and dynamic changes in brain states. It analyzes pathways to enhance chip generalization and real-time performance across three technical dimensions: hardware architecture, algorithm optimization, and multimodal fusion. The paper also explores core challenges like data privacy and energy-efficiency tradeoffs. Building on this foundation, it proposes a neuromorphic computing design framework for next-generation chips to advance the intelligent and personalized development of BCI in medical rehabilitation and human-computer interaction.

Keywords: Brain-computer interface; Artificial intelligence; Chip; Large model-driven; Meta-learning

Online publication: February 27, 2026

1. Introduction

Brain-computer interface (BCI), as a critical interdisciplinary technology bridging the human brain and external devices, has garnered immense attention in fields such as medical rehabilitation and human-computer interaction. It enables direct communication between the central nervous system and external equipment by decoding neural signals, offering revolutionary solutions for patients with motor disabilities to restore motor functions and promoting the evolution of intelligent human-computer interaction modes. However, the performance of current BCI systems is largely constrained by the underlying chip technology, particularly the limitations of traditional fixed decoding models.

Neural plasticity, the inherent ability of the human brain to reorganize neural connections in response to external stimuli and internal states, coupled with the dynamic changes in brain states caused by factors such as fatigue, emotion, and task switching, poses significant challenges to traditional BCI chips. Fixed decoding models,

designed based on pre-set neural signal features, lack the flexibility to adapt to individual differences in neural activity and real-time changes in brain states, resulting in compromised decoding accuracy, poor generalization across individuals, and inadequate real-time performance. These drawbacks severely hinder the practical application and further development of BCI technology, making it urgent to explore innovative technical paths to enhance the adaptive capacity of BCI chips.

In recent years, the rapid advancement of artificial intelligence (AI) large models, represented by Transformers and meta-learning, has brought new opportunities for breaking through the bottlenecks of traditional BCI chips. These AI models possess powerful capabilities in feature extraction, adaptive learning, and pattern recognition, which can be leveraged to empower BCI chips with dynamic adaptation capabilities. By integrating AI large models into BCI chip design, it becomes feasible to realize real-time adjustment of decoding strategies according to individual neural characteristics and dynamic brain state changes, thereby significantly improving the generalization and real-time performance of BCI systems.

Against this backdrop, this paper focuses on the empowerment mechanism of AI large models on BCI chips. It systematically analyzes the technical pathways to enhance BCI chip performance from three core dimensions: hardware architecture optimization tailored for AI model deployment, algorithm improvement to strengthen adaptive decoding capabilities, and multimodal fusion to complement neural signal information. Meanwhile, critical challenges in the integration process, including data privacy protection of neural signals (which involve highly sensitive personal biological information) and the tradeoff between chip energy efficiency and computational performance, are also explored in depth. On this basis, a neuromorphic computing design framework for next-generation BCI chips is proposed, aiming to provide a theoretical and technical foundation for promoting the intelligent, personalized, and practical development of BCI technology in medical rehabilitation, human-computer interaction, and other key fields.

2. Principles of AI large model-driven adaptive technology

2.1. Dynamic decoding model construction

The core of achieving adaptive BCI chips lies in constructing decoding models capable of dynamic evolution. This approach overcomes individual differences among users and the diversity of brain states that change with different contexts. Such models enable the chip to continuously learn and optimize during actual use, thereby maintaining high performance over the long term.

2.1.1. Online learning architecture based on incremental learning

Incremental learning aims to enable intelligent systems to “continuously learn” like humans, absorbing knowledge from new data while retaining prior learning without retraining on the entire historical dataset ^[1]. This characteristic offers an ideal solution for addressing the dynamic nature of electroencephalograms (EEG) signals. Since EEG signals drift with users’ physiological states (e.g., fatigue, attention), long-term usage habits, and even neuroplasticity, fixed models often face performance degradation. When integrated into chip-based systems, incremental learning algorithms enable continuous monitoring of EEG signals alongside corresponding operational outcomes. Upon detecting a decline in decoding accuracy or shifts in signal patterns, the system applies small, targeted adjustments to model parameters using newly acquired mini-batch data. This allows chip-deployed decoding models to achieve “accompanied growth” maintaining high-precision alignment with the user’s current

brain state and establishing an algorithmic foundation for long-term adaptability.

2.1.2. Rapid adaptation capability based on meta-learning

Meta-learning, or “learning to learn” aims to enhance a model’s generalization and rapid adaptation capabilities on new tasks, enabling efficient adjustments with minimal samples ^[2]. This capability directly addresses a core user experience bottleneck in traditional BCI: model calibration for new users typically requires lengthy, tedious data collection and training processes. Chips based on algorithms like Model-Agnostic Meta-Learning (MAML) leverage pre-learned cross-user prior knowledge to rapidly adapt to new users with minimal calibration data. In BCI applications, meta-learning reduces initial calibration time from hours to minutes, significantly improving device usability and user experience.

2.2. Multimodal signal fusion

Single-modality EEG signals are susceptible to interference and have limited informational dimensions, making it challenging to ensure the decoding accuracy of adaptive BCI systems. By integrating complementary physiological signals from diverse sources, multimodal fusion significantly enhances the accuracy and environmental adaptability of decoding models while meeting system requirements for low power consumption and high efficiency.

2.2.1. Chip-level fusion architecture design

To achieve low-latency parallel processing of multimodal data, dedicated multimodal fusion processing units must be designed at the chip level. This unit integrates hardware accelerators to perform real-time parallel processing of multimodal data such as EEG, fNIRS (functional near-infrared spectroscopy), and eye tracking. It incorporates specialized preprocessing circuits for different signals, enabling preliminary signal cleaning and feature extraction at the chip front end. Fusion occurs directly within the unit, significantly reducing data transfer overhead between storage and computation units. This design meets the stringent power and speed requirements for real-time interaction. Research by Shi *et al.* indicates that feature-level fusion can integrate complementary characteristics from peripheral physiological signals like EEG and ECG at the stage “after feature extraction and before classifier input,” forming a unified joint feature vector to preserve deep intermodal correlations ^[3]. This conclusion provides critical justification for designing an “EEG-ECG feature concatenation unit” in chip front-end architecture, helping avoid data upload delays and better adapt to real-time brain-computer interaction scenarios.

2.2.2. Cross-modal attention mechanism

At the algorithmic level, a Transformer-based cross-modal attention mechanism can be introduced to dynamically assign weights to information from different modalities. Addressing the issues of “disconnected modal interactions and redundant model parameters” in current multimodal BCI algorithms, Guo *et al.* demonstrated that cross-modal attention mechanisms can integrate information from all input modalities simultaneously (rather than sequentially pairing them) and optimize training stability through techniques like layer normalization and residual connections ^[4]. This approach can guide the design of efficient multimodal fusion modules, such as simultaneously integrating EEG’s temporal features, ECG’s heart rate variability features, and eye-tracking’s spatial features. This enables deeper exploration of complementarity between modalities, significantly enhancing robustness in emotion or intent recognition.

2.3. Few-shot learning and model compression

Deploying complex AI models onto resource-constrained BCI chips presents a fundamental challenge: the high complexity of models versus the strict limitations of chip computing power, memory, and power consumption. To overcome this bottleneck, few-shot learning and model compression techniques must be introduced to enable efficient and accurate model operation in low-resource environments.

2.3.1. Few-shot learning mechanism

Traditional deep learning heavily relies on massive labeled datasets, whereas practical BCI applications (such as new user calibration) often provide only a minimal number of samples. Few-shot learning aims to enable models to achieve efficient learning and accurate predictions even under conditions of scarce training data. Its core lies in utilizing algorithms like meta-learning to perform pre-training on rich multi-user data before chip production. This endows the model with the intrinsic ability to capture individual signal variations and rapidly adapt to new users. For instance, the meta-learning framework proposed by Li *et al.* provides a key approach, where during the meta-training phase, the model learns across numerous simulated small-sample tasks (composed of a “support set” and a “query set”) to obtain a set of highly adaptable universal initialization parameters. During the meta-testing phase (when encountering new users), the model converges rapidly using only a small number of calibration samples (e.g., 5–10 sets of EEG data) to construct an effective personalized classifier^[5]. This “learning to learn” capability significantly reduces the time and data costs associated with user calibration.

2.3.2. Knowledge distillation and model compression

Model compression is a core technology for addressing the challenges of deploying large models. Its goal is to significantly reduce model size and computational requirements while maximizing the retention of the original model’s performance. As a representative technique, knowledge distillation employs a “teacher-student framework” to achieve model lightweighting: a large, accurate “teacher model” guides the training of a lightweight “student model.” The “student model” not only learns actual task labels but also inherits the teacher’s strong generalization capabilities by mimicking its softened output probability distribution (using techniques like temperature scaling and KL divergence) and aligning feature representations across intermediate layers^[6].

2.3.3. Applications of neural architecture search technology

Traditional neural network design relies on expert experience, making it difficult to automatically achieve optimal matching with target hardware. Neural architecture search (NAS) technology automates the exploration of design spaces, enabling the discovery of optimal neural network architectures directly tailored to specific hardware constraints (e.g., computational power, power consumption, latency) and task objectives (e.g., EEG decoding accuracy). This achieves a paradigm shift from “manually designing models to fit hardware” to “automatically co-searching for optimal model-hardware pairings.” Notably, recent advancements have further extended NAS frameworks toward trustworthy and secure deployment, critical aspects in handling sensitive data such as in BCI. For instance, a blockchain-enhanced trustworthy NAS method has been proposed for medical image classification, which integrates secure data provenance and reliable model prediction transmission while maintaining high search performance^[7]. This approach underscores how NAS can be synergized with security mechanisms to protect data confidentiality and ensure interpretability in automated model development. Although applied in pneumonia image diagnosis, such a trustworthy NAS framework offers a promising paradigm for BCI systems where neural data

security and reliable decoding are equally vital.

3. Adaptive hardware implementation of brain-computer interface chips

3.1. Reconfigurable computing architecture

To support the dynamic adaptive capabilities of large AI models while meeting the stringent demands of BCI for high energy efficiency and computational real-time performance, reconfigurable computing architectures offer a critical pathway to overcome traditional chip energy efficiency bottlenecks.

3.1.1. FPGA dynamic reconfiguration

As a chip capable of defining its hardware functionality through configuration software after deployment, field-programmable gate arrays (FPGAs) combine the high speed of hardware parallel processing with the flexibility of software programmability. FPGA dynamic reconfiguration technology enables real-time modification, updating, or restructuring of the functionality of partial or complete logic resources within the chip without interrupting overall system operation. This capability provides core support for enabling flexible switching between different processing tasks and efficient reuse of hardware resources in BCI chips. Research indicates that this technology “enables dynamic reprogramming of FPGAs during runtime to alter their functionality online” and “significantly conserves on-chip resources, reserving more space for other operations” [8]. This means a single BCI chip can rapidly switch between multiple hardware acceleration cores, such as those for signal preprocessing, feature extraction, or specific decoding algorithms, by loading different partial configuration files. This “time-division multiplexing” strategy significantly enhances the chip’s functional diversity and scenario adaptability under resource-constrained conditions.

3.1.2. Compute-in-memory architecture

By embedding computational units within memory, the compute-in-memory architecture enables operations to be performed directly at the data storage location. This fundamentally reduces latency and power consumption associated with data movement between storage and computational units. To achieve the integration of high energy efficiency and high-precision computing, advanced architectures such as high-precision hybrid floating-point in-memory computation (Hy-FPCIM) for complex models (e.g., Vision Transformers) have been proposed. By efficiently executing decomposed exponent and mantissa operations in-memory, these approaches significantly enhance computational energy efficiency and area efficiency with near-lossless precision [9]. Research indicates that the CIM architecture has demonstrated broad application prospects in fields such as AI and the IoT. Its high energy efficiency and low latency characteristics offer novel technical solutions for tackling challenges in BCI systems, including real-time signal processing, lightweight model deployment, and stringent power constraints. Incorporating the CIM architecture into BCI chip design is expected to drive system development toward higher real-time performance, enhanced energy efficiency, and improved integration.

3.2. Neuromorphic computing integration

To fundamentally approximate the bio-brain’s high energy efficiency and adaptive properties, neuromorphic computing simulates the brain’s spiking communication and synaptic plasticity mechanisms, endowing next-generation BCI chips with low power consumption and intrinsic learning capabilities.

3.2.1. Spiking neural networks

Spiking neural networks (SNNs) utilize discrete spike sequences as information carriers and perform event-driven computations, exhibiting a mechanism highly analogous to the human brain. When deployed on chips, SNNs offer dual advantages of event-driven efficiency and ultra-low power consumption. However, due to the susceptibility of EEG signals to noise and individual variations, enhancing SNN robustness and their ability to model underlying signal dynamics is a critical prerequisite for practical application. Recent advances, such as the finite-difference physics-informed spiking neural network (FPSNN), address core challenges by successfully integrating physics-informed learning paradigms with SNN architectures, thereby enhancing the network's capacity to capture and generalize the dynamical characteristics of time-series signals like EEG^[10]. This approach demonstrates that incorporating prior physical or dynamic constraints can significantly improve SNN performance over purely data-driven methods. Leveraging the efficient data access characteristics of the CIM architecture, SNNs empowered by such physics-aware learning frameworks can serve as the core for low-power, high-robustness computation in BCI chips, further enhancing system practicality and reliability in real-world complex environments.

3.2.2. Brain-inspired chip design

The biomimetic design of brain-inspired chips has advanced to the cellular and molecular levels. By integrating dendritic computing units, advanced brain-inspired chips can perform complex spatiotemporal signal integration, thereby enhancing the computational complexity of individual neuron nodes. This enables smaller-scale networks to achieve powerful pattern recognition capabilities. This approach not only helps reduce chip area and power consumption but also improves robustness when processing the dynamic characteristics of neural signals. A hierarchical multi-core architecture combined with a quasi-event-driven mechanism effectively supports the spatiotemporal signal processing of SNNs, improving the chip's real-time performance and energy efficiency during dynamic tasks. By implementing synaptic plasticity through devices like memristors, BCI chips can simulate lifelong learning at the physical hardware level. This allows the chip to autonomously optimize internal connection weights based on the user's neural activity patterns without requiring high-level software intervention, laying the hardware foundation for achieving "adaptive evolution." Neuromorphic chips supporting on-chip learning can achieve efficient synaptic weight updates through mechanisms like direct feedback alignment. These chips exhibit strong spatio-temporal locality and hardware-friendliness, making them suitable for real-time learning scenarios at the edge.

3.3. Balancing low power consumption and real-time performance

BCI chips designed for implantable or wearable applications must achieve millisecond-level real-time responsiveness under stringent power constraints, yet high-performance computing often conflicts with low power consumption. This section explores how to achieve an optimal balance of system energy efficiency while ensuring real-time performance.

3.3.1. Dynamic voltage and frequency scaling technology

Dynamic voltage and frequency scaling (DVFS) technology achieves intelligent energy efficiency management by continuously monitoring the computational load of chips and dynamically adjusting their operating voltage and clock frequency. The core of this technology lies in precisely matching power supply to computational power based on real-time performance demands, thereby minimizing power consumption while meeting task

requirements. To achieve this goal, research efforts have shifted toward developing advanced scheduling algorithms that optimize energy efficiency while ensuring real-time system performance, thereby reducing the risks of delays and instability caused by frequent frequency and voltage adjustments^[11]. The operational intent of BCI users exhibits intermittent and time-varying characteristics, leading to significant fluctuations in chip workload. To address this, an intelligent frequency adjustment mechanism can be developed by incorporating load prediction models based on the AVG algorithm and Hotplug multi-core management strategies. When the system is idle or under low load (e.g., during user rest periods), the chip automatically switches to low-power mode while immediately scaling up to peak performance upon detecting user intent for complex operations. This “power-on-demand” strategy effectively extends battery life for implantable devices, highlighting the practical value of DVFS technology in low-power embedded systems.

3.3.2. Application of approximate computation in EEG signal processing

BCI tasks typically exhibit a degree of fault tolerance, enabling the introduction of approximate computation to trade off for improved energy efficiency. Approximate computation significantly reduces power consumption and hardware resource usage by simplifying computational processes or circuit designs within acceptable accuracy loss margins. As a representative model of brain-inspired computing, spiking neural networks inherently possess a distributed architecture and training methodology that endows them with tolerance for partial computational errors, thereby establishing a theoretical foundation for applying approximate computation in EEG signal processing. In practical implementation, strategies such as applying sensitive average relative error metrics and approximate adder selection can be adopted. By leveraging the input distribution characteristics of addition operations in EEG signal processing, approximate computing units can be selected to achieve an optimal balance between maintaining system recognition accuracy and realizing significant energy efficiency gains. Experiments demonstrate that this approach achieves 37.32% power savings and 31.26% area reduction while only sacrificing approximately 3.47 percentage points in classification accuracy. This points to a highly promising design direction for future implantable BCI chips to maintain long-term stable operation under limited energy budgets.

4. Core challenges and solutions

4.1. Data privacy and security

BCI can interpret human thought processes, offering immense potential while simultaneously exposing neural data to severe risks of leakage and tampering. The resulting data privacy and security concerns have become widespread public anxieties, with ethical issues such as personal integrity and autonomy also coming to the fore. Therefore, security design must form the cornerstone of chip architecture.

4.1.1. Challenge: Risks of neurodata leakage

The security challenges posed by neural data far exceed those of other data types, manifesting primarily in two following aspects:

- (1) Neural data such as EEG directly relate to an individual’s thoughts, intentions, and emotional states, constituting the highest dimension of personal privacy. Should raw data leak, it would lead to “thought exposure,” placing users in a state of being watched. Such data could be exploited for commercial manipulation or psychological attacks, severely infringing upon cognitive freedom and mental privacy;
- (2) Closed-loop BCIs with “writing” capabilities create channels for reverse attacks. Attackers could

manipulate users' perceptions and decisions by fabricating neural feedback signals, potentially inducing mental symptoms and posing direct threats to personal safety and autonomy^[12]. Such loss of behavioral control due to technical vulnerabilities or malicious intrusion may not only constitute civil infringement but could even reach criminal offenses under extreme circumstances.

4.1.2. Solution: Privacy-preserving training based on federated learning

To address the aforementioned risks, a distributed training paradigm based on federated learning can be adopted. Within this framework, users' raw neural data is stored and processed entirely within local secure environments without requiring upload to central servers, structurally eliminating the risk of large-scale data leakage during transmission and aggregation. The core of federated learning lies in participants training models locally and uploading only encrypted model parameter updates (e.g., gradients), thereby enabling collaborative learning while protecting data source privacy. To address the risks of collusive attacks and privacy leaks that may exist in traditional server architectures, the latest privacy protection solution effectively enhances system security, verifiability, and practical efficiency by introducing homomorphic proxy re-encryption and a dual-server architecture^[13]. To further enhance security, techniques such as homomorphic encryption or secure multi-party computation can be integrated to shield the parameter update process, preventing the reconstruction of raw data from gradient information. Such approaches (e.g., SE-Fed) also optimize communication efficiency through techniques like client grouping and gradient compression. This maintains high model performance while ensuring privacy, making it suitable for real-time-critical applications like BCIs.

4.2. Trade-off between model complexity and hardware resources

The exceptional performance of large AI models relies on massive parameters and complex computations, creating a fundamental conflict with the extremely constrained computational power, memory, and power consumption budgets of BCI chips. This section explores solutions from two perspectives: model optimization and hardware adaptation.

4.2.1. Challenge: The parameter scale of large models far exceeds the computational power of chips

Deploying models with massive parameters on chips primarily faces three major challenges: computational power, memory, and power consumption, as follows:

- (1) In terms of computational power, the volume of floating-point operations required for a single inference by large models far exceeds the computational upper limit that embedded processors can provide within real-time requirements (typically < 10ms);
- (2) Memory constraints arise because model parameters often span hundreds of megabytes, while on-chip storage typically ranges from kilobytes to megabytes. Frequent access to off-chip, low-speed memory incurs prohibitive latency and power overhead;
- (3) Power consumption is further strained by complex computations and frequent data transfers, generating significant heat, a severe challenge for embedded chips operating within milliwatt-level power budgets.

4.2.2. Solution: Model pruning and quantization

Model pruning and quantization techniques aim to directly reduce model size and computational complexity,

which is critical for deploying neural networks on resource-constrained hardware. Model pruning streamlines models by identifying and removing redundant connections or structures within the network. Structured pruning, such as removing entire filters or channels, generates hardware-friendly, regular sparse models that significantly reduce computational load with minimal accuracy loss. Model quantization converts high-precision floating-point parameters and activation values into low-bit fixed-point representations. Key techniques include quantization-aware training (simulating quantization effects during training to enhance final accuracy) and mixed-precision quantization (dynamically allocating bit widths based on layer sensitivity). The synergistic application of pruning and quantization can reduce a model's storage and computational requirements by an order of magnitude, enabling efficient deployment on edge devices. Systematic evaluations, such as those conducted in recent TinyML-oriented research, demonstrate that combining structured pruning with INT8 quantization can achieve substantial model compression (e.g., over 75% size reduction for models like MobileNet) while preserving competitive accuracy, thereby providing a validated pathway for implementing high-efficiency neural networks in real-world, resource-limited applications such as BCI chips ^[14].

4.3. Long-term stability and biocompatibility

BCI chips, particularly invasive implantable devices, must overcome significant challenges posed by the complex biological environment to achieve long-term clinical application. Long-term stability requires chips to maintain functionality without degradation over periods spanning years to decades, while biocompatibility demands stable coexistence with neural tissue without triggering harmful immune or inflammatory responses.

4.3.1. Challenge: Safety of chronic implantation for invasive chips

The core safety challenge facing invasive BCI lies in the long-term mechanical and biological incompatibility between traditional rigid electrodes and brain tissue. This incompatibility leads to persistent micro-motion friction, triggering chronic inflammatory responses and glial scar encapsulation. Not only can this damage neural tissue, but it also causes progressive degradation in neural signal recording quality and even device failure. Therefore, developing novel materials and structures capable of forming stable, harmonious interfaces with neural tissue is of paramount importance.

4.3.2. Solution: Flexible electronic materials and biocompatible interfaces

The use of flexible electronic materials, such as conductive hydrogels, offers an effective solution to biocompatibility challenges. These materials exhibit a modulus highly compatible with neural tissue, significantly reducing tissue damage caused by mechanical mismatch and effectively suppressing glial scar formation. Functional surface modifications can further guide neuronal synapse adhesion, promoting stable, high-quality electrical coupling and biointegration between the chip and neural circuits. This lays the foundation for acquiring long-term stable neural signals.

4.3.3. Solution: Self-healing circuit

To address potential aging, wear, or localized failure of circuits after implantation, self-healing circuit technology can be introduced. This technology embeds dynamically reversible chemical bonds or microcapsule repair agents within the material, enabling circuits to automatically trigger physical or chemical processes upon minor fractures or performance degradation. This reconfigures conductive pathways and restores functionality. This self-healing

mechanism not only addresses macro-structural damage but has also been validated in critical analog circuit modules such as voltage reference circuits. It effectively counters aging effects like thermal carrier injection, ensuring that key performance parameters, including temperature coefficient and output voltage, remain within acceptable ranges even after prolonged operation ^[15]. When combined with biodegradable materials serving as temporary scaffolds or repair media, on-demand repair and absorption can be achieved. This self-healing capability significantly enhances the long-term reliability and service life of implantable systems in unattended environments.

5. Future research directions

The key breakthrough direction for future BCI chips lies in integrating the strengths of neuromorphic computing and large artificial intelligence models to construct hybrid architectures that combine biological plausibility with powerful cognitive capabilities, namely, brain-inspired AI fusion chips. Such chips may adopt a layered design: the bottom layer employs ultra-low-power SNN units to directly process neural signals, performing temporal feature extraction and preliminary filtering; the top layer utilizes compressed and optimized modules like Transformers to handle complex semantic decoding and advanced cognitive tasks. Communication and coordination between these layers occurs via efficient cross-paradigm interfaces. Realizing this vision requires developing novel collaborative training methods and hybrid neural architecture search techniques. These approaches must enable hardware-friendly mutual enhancement and co-evolution of SNN perceptual capabilities and ANN cognitive capabilities, ultimately forming high-performance, energy-efficient autonomous intelligent neural processors.

To overcome computational power and power consumption constraints in edge chips, constructing efficient “edge-cloud” collaborative systems represents a core future direction. Within this architecture, edge devices handle lightweight processing ensuring real-time responsiveness and privacy; edge servers host personalized models and perform preliminary aggregation; while the cloud manages large-scale model training and global optimization. 5G/6G networks serve as high-speed connectivity links, guaranteeing real-time synchronization and reliable communication across layers. This collaborative model enables dynamic workload distribution and global resource optimization, allowing endpoints to dynamically access the latest cloud capabilities in real time and driving continuous adaptive evolution of system performance.

To advance technology from laboratory to large-scale application, building an open collaborative ecosystem and unified standardization system is essential. Current fragmentation in hardware interfaces, data formats, communication protocols, and algorithm frameworks severely hinders technological iteration and adoption. Future efforts should focus on advancing standards across all layers, including hardware interfaces, data representation, algorithm deployment, and security protocols. Simultaneously, fostering an open ecosystem involving industry, academia, research institutions, and healthcare providers through open-source hardware design, shared benchmark datasets, and open toolkits will lower R&D barriers, accelerate innovation, and ultimately realize the vision of extending technology benefits to broader populations.

6. Conclusion

Large AI models provide the core algorithmic foundation for achieving efficient and precise adaptive evolution in BCI chips. Their capabilities in online learning and small-sample adaptation enable chips to dynamically respond to changes in users’ neural plasticity. Meanwhile, innovations in hardware architecture, particularly bio-inspired

and energy-efficient designs such as compute-in-memory and neuromorphic computing, constitute the critical infrastructure for supporting and unleashing such intelligent algorithms. The synergy between these two elements collectively propels BCIs from static, generic traditional forms toward a new era of dynamic, personalized intelligent interaction. However, technological maturation and widespread adoption still face multiple challenges. Future breakthroughs must address core bottlenecks, including data privacy protection, energy-efficiency optimization, long-term biocompatibility, and the absence of standardized protocols. This requires developing multidisciplinary solutions spanning algorithms, hardware, materials, and security, alongside establishing open industrial ecosystems and collaborative frameworks. Only through persistent technological breakthroughs and collaborative innovation can BCIs transition from laboratory prototypes to safe, reliable, and accessible large-scale applications. This will realize the next generation of human-machine integrated intelligent interaction, unlocking new possibilities for medical rehabilitation, cognitive enhancement, and even the evolution of human existence.

Disclosure statement

The author declares no conflict of interest.

References

- [1] He Z, Huang S, Lu Y, et al., 2026, MoTiC: Momentum Tightness and Contrast for Few-Shot Class-Incremental Learning. *Pattern Recognition*, 2026(173): 112753.
- [2] Zhang J, Chang Y, Wang F, 2025, Few-Shot Working Condition Classification within a Meta-Learning Framework based on Multi-Head Attention Autoencoder. *Journal of Intelligent Manufacturing*, 2025(prepublish): 1–16.
- [3] Shi P, Wang H, Liu L, Physiological Signal Emotion Recognition: A Review of Cross-Domain Transfer and Multimodal Fusion. *Journal of Frontiers of Computer Science and Technology*, 1–23.
- [4] Guo J, Lu H, Xu J, 2025, Research on Multimodal Sentiment Analysis Method Based on Cross-Modal Attention Mechanism. *Computer Knowledge and Technology*, 21(1): 1–4.
- [5] Li K, et al., 2025, A Review of Few-Shot Learning Research. *Mechanical & Electrical Engineering Technology*, 54(6): 160–168.
- [6] Wang Y, Chen Y, 2026, Temperature-Driven Category Decoupled Knowledge Distillation with Interpretability for Model Compression. *Advanced Engineering Informatics*, 69(PC): 104051.
- [7] Yang Y, Wei J, Yu Z, et al., 2024, A Trustworthy Neural Architecture Search Framework for Pneumonia Image Classification Utilizing Blockchain Technology. *The Journal of Supercomputing*, 80(2): 1694–1727.
- [8] Boudjadar J, Islam S, Buyya R, 2025, Dynamic FPGA Reconfiguration for Scalable Embedded Artificial Intelligence (AI): A Co-Design Methodology for Convolutional Neural Networks (CNN) Acceleration. *Future Generation Computer Systems*, 2025(169): 107777.
- [9] Ma Z, Wang C, Chen Q, et al., 2025, A High-Precision Hybrid Floating-Point Compute-in-Memory Architecture for Complex Deep Learning. *Electronics*, 14(22): 4414.
- [10] Wei Q, Yang Q, Han L, et al., 2026, Physics-Informed Spiking Neural Networks for Continuous-Time Dynamic Systems. *Neurocomputing*, 2026(665): 132192.
- [11] Chen Y, Huang J, Xiao L, et al., 2025, A DVFS-Weakly Dependent Real-Time Scheduling for Multiple Parallel Applications on Energy-Aware Heterogeneous Systems. *Journal of Systems Architecture*, 2025(170): 103614.
- [12] Brocal F, 2023, Brain-Computer Interfaces in Safety and Security Fields: Risks and Applications. *Safety Science*,

2023: 160.

- [13] Song Z, Huang Z, Cai Y, 2026, A Verifiable Privacy-Preserving Federated Learning Scheme Based on Homomorphic Proxy Re-Encryption. *Information Sciences*, 2026(730): 122875.
- [14] Shabir M, Torta G, Damiani F, 2025, TinyML Model Compression: A Comparative Study of Pruning and Quantization on Selected Standard and Custom Neural Networks. *Telecommunication Systems*, 88(4): 132.
- [15] Zhang J, Hu J, Jiang M, et al., 2023, A HCI-Hardened Self-Healing Operational Amplifier Circuit. *Microelectronics Reliability*, 2023: 151.

Publisher's note

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

Intelligent Identification of Water Accumulation and Ice Formation in Traffic Tunnels

Beining Chen*

Zhejiang University, University of Illinois Urbana-Champaign Institute (ZJUI), Haining 314400, Zhejiang, China

**Author to whom correspondence should be addressed.*

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Water accumulation and ice formation in traffic tunnels pose prominent safety hazards (e.g., reduced road friction, increased traffic accidents) and threaten structural integrity (e.g., damage to waterproof layers and lining structures). Therefore, the intelligent identification of these two hazards is crucial for safeguarding traffic safety and optimizing tunnel maintenance strategies. The intelligent identification system integrates computer vision, deep learning, and multi-source sensor data fusion technologies. Current state-of-the-art practices adopt deep learning models for target segmentation and detection, combined with robust image preprocessing and post-processing techniques. This technology exhibits significant practical application value, and its continuous innovation and development are expected to substantially enhance the level of tunnel safety management and structural durability preservation.

Keywords: Intelligent recognition; Traffic tunnel; Water accumulation and ice formation; Deep learning; Computer vision

Online publication: February 27, 2026

1. Introduction

As a critical component of modern transportation networks, traffic tunnels play an irreplaceable role in ensuring the efficiency of daily travel and the stability of regional economic development ^[1,2]. However, affected by factors such as complex geological conditions, extreme weather, and operational wear, water accumulation and ice formation frequently occur in traffic tunnels, seriously endangering their safe and unobstructed operation. Thus, the research and application of intelligent identification technology for water accumulation and ice formation in traffic tunnels have important practical significance and engineering value.

Water accumulation in traffic tunnels is caused by multiple factors. During the rainy season, excessive precipitation may exceed the drainage capacity of the tunnel, leading to surface waterlogging; in addition, malfunctions or blockages in the tunnel's internal drainage system can also result in water accumulation ^[3]. Accumulated water not only reduces the friction coefficient of the road surface, prolongs vehicle braking distances, and increases the risk of traffic accidents such as rear-end collisions and skidding but also has a persistent adverse

impact on the tunnel's waterproof layer. As the core barrier against groundwater erosion, the waterproof layer is prone to aging, degradation, and even failure when immersed in water for a long time ^[4]. Once the waterproof layer loses its function, groundwater directly contacts the tunnel lining, accelerating the process of concrete carbonation, steel bar corrosion, and structural cracking, thereby threatening the long-term stability of the tunnel structure.

Ice formation in traffic tunnels mainly occurs in cold seasons. When the internal temperature of the tunnel drops to or below the freezing point and the air humidity reaches a certain level, the accumulated water on the road surface is prone to freezing ^[5]. The icy road surface has extremely low friction, which significantly increases the difficulty of vehicle control and poses a severe threat to traffic safety, similar to driving on an ice sheet. At the same time, ice formation can cause damage to tunnel supporting facilities: for example, low temperatures may lead to freezing and bursting of water supply and drainage pipes, failure of electrical equipment, and other problems, thereby affecting the normal operation of the tunnel ^[6].

The damage caused by water accumulation and ice formation to traffic tunnels involves both traffic safety and structural safety, and the damage mechanism is complex and cumulative ^[7]. Therefore, it is urgent to develop efficient and accurate intelligent identification technology to realize real-time monitoring and early warning of these two hazards, providing reliable technical support for scientific prevention and control and targeted maintenance.

2. Principles of intelligent identification technology

In the traffic tunnel environment, the intelligent image recognition technology for water accumulation and ice formation is a key means to prevent traffic accidents (e.g., vehicle skidding, loss of control) and realize timely maintenance. Its technical system integrates computer vision, deep learning, and professional image processing technologies ^[1]. The core principles, key technologies, and common tool chains of the system are detailed as follows.

2.1. Core principles: Distinction based on visual features and environmental context

The intelligent identification of water accumulation and ice formation relies on the significant differences in visual features between the two hazards and the surrounding road environment, combined with environmental context information for comprehensive judgment ^[9].

2.1.1. Visual feature characteristics of water accumulation

The visual features of water accumulation are as follows:

- (1) Reflective properties: The water surface has a mirror-like reflection effect, which can reflect light sources such as vehicle headlights, tunnel lighting, and ambient light, forming obvious highlight areas in the image. The shape, brightness, and position of these highlight areas change dynamically with the viewing angle of the camera and the direction of the light source, which is one of the most distinctive features for identifying water accumulation;
- (2) Color properties: Clear water is colorless and transparent. In images, when the water layer is thin and the reflection effect is weak, it often presents a darker shade relative to the road surface; when affected by the surrounding environment, it may be tinged with environmental colors. Murky water contains sediment and other impurities, usually showing yellowish or brownish tones;

- (3) Texture properties: Flowing water has obvious ripple and wave textures; static water has a relatively smooth surface, but there may be slight ripples or reflection interference. The edge of the water accumulation area is irregular and has no fixed shape;
- (4) Dynamic properties: Water is fluid. In continuous video sequences, the reflection points and edge contours of the water accumulation area will change continuously with time, showing obvious dynamic characteristics.

2.1.2. Visual feature characteristics of ice formation

The visual features of ice formation are as follows:

- (1) Reflective properties: The ice surface also has reflective characteristics, but compared with water, its reflection is more “hard” and directional. The highlight areas are usually more uniform and diffuse, or form regular specular reflection patterns. Thin ice presents a translucent state, and the reflection effect is relatively weak;
- (2) Color properties: Clear ice is transparent or semi-transparent, and its color in the image is highly consistent with the road surface, making it difficult to distinguish directly; thick ice often shows bluish-white or grayish-white tones. Dirty ice mixed with impurities such as mud presents earthy colors similar to mud;
- (3) Texture properties: The ice surface is extremely smooth, without the ripple texture of water. The edge of the ice formation area is relatively sharp, and it is often distributed in patches. Frosted ice or melting ice may show granular or honeycomb-like micro-textures;
- (4) Dynamic properties: Ice is a solid substance. In short-term video sequences, its shape and position remain relatively stable unless it is in the process of melting or continuous formation.

2.1.3. Environmental context auxiliary judgment

The information is as follows:

- (1) Location information: Water accumulation and ice formation have obvious location aggregation characteristics. They often occur in road depressions, near drainage outlets, tunnel entrances and exits (where temperature changes sharply and ice formation is easy), and under seepage points ^[10];
- (2) Temperature information: By integrating temperature sensor data, when the road surface temperature is close to or below 0°C, the detected moist areas are highly likely to be ice or in a state of impending freezing, which can effectively improve the accuracy of ice formation identification;
- (3) Weather information: Combined with external weather data, the probability of water accumulation and ice formation in tunnels increases significantly after rainfall or snowfall. Weather factors can provide important prior information for the identification model ^[8].

2.2. Key technologies and methods

2.2.1. Image preprocessing

Image preprocessing is the foundation of intelligent identification, aiming to eliminate noise and interference in tunnel images and enhance the effective features of water accumulation and ice formation ^[1]. Aiming at the problems of uneven illumination, low brightness, and strong glare in tunnels, technologies such as histogram equalization, Retinex algorithm, and deep learning-based low-light enhancement are adopted to adjust the image

brightness and contrast, ensuring the consistency of image quality under different illumination conditions. The strong light generated by vehicle headlights in the tunnel is easy to be misidentified as the reflection of water accumulation or ice formation. For this problem, a combination of image inpainting and deep learning algorithms is used to accurately segment and repair the glare area, eliminating false positive interference. For slight camera shake caused by vehicle vibration or environmental factors, image stabilization processing is performed through frame alignment and motion compensation technologies to ensure the stability of feature extraction in continuous frames.

2.2.2. Feature extraction

Feature extraction is the core link of identifying water accumulation and ice formation, which is divided into traditional visual feature extraction and deep feature extraction^[8]. Before the popularization of deep learning, traditional visual features were widely used in target identification and are still used as supplementary means in specific scenarios. By converting images into HSV, Lab, and other color spaces, the color distribution characteristics of water accumulation and ice formation areas are analyzed. For example, the saturation (S) and brightness (V) channels in the HSV space can effectively distinguish the color differences between water, ice, and dry roads. Using Gray-Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), Gabor filters, and other technologies to extract texture information of the target area, so as to distinguish the ripple texture of water, the smooth texture of ice, and the rough texture of the road surface.

Using Canny edge detection and other algorithms to extract the edge contours of water accumulation and ice formation areas, and determine the scope of the target area through the shape characteristics of the contours. By detecting the highlight areas in the image, analyzing their shape, area, and brightness distribution, the reflection characteristics of water and ice are distinguished. As the current mainstream and most effective feature extraction method, Convolutional Neural Networks (CNNs) are used to automatically learn multi-level and multi-dimensional feature representations from raw images^[9]. Through the hierarchical convolution and pooling operations of the network, low-level features (e.g., edges, textures) and high-level semantic features (e.g., overall shape, reflection mode) are extracted, which can accurately distinguish water accumulation, ice formation, and dry road surfaces.

2.2.3. Object detection and segmentation

Object detection and segmentation realize the positioning and pixel-level classification of water accumulation and ice formation areas^[1]. For instance:

- (1) Object detection: The goal is to determine whether there are water accumulation and ice formation in the image and output their bounding boxes. Currently, deep learning-based detection models such as YOLO (You Only Look Once) are mainly used. These models require a large number of annotated tunnel water and ice images for training to ensure the accuracy and real-time performance of detection;
- (2) Semantic segmentation: By classifying each pixel in the image, the water accumulation area, ice formation area, dry road surface, and other background areas are accurately labeled at the pixel level. Models such as U-Net are widely used in tunnel defect segmentation due to their excellent performance in small target segmentation, which can provide precise spatial location information for subsequent hazard assessment.

2.2.4. Dynamic analysis (video sequences)

For video data, dynamic analysis technology is used to further improve the reliability of identification^[10]. By calculating the motion vector of pixels between consecutive frames, the motion pattern of the target area is analyzed. Flowing water will generate irregular optical flow fields, while static ice will not produce obvious motion vectors (or only move synchronously with the camera). Using frame differencing technology to detect newly appearing water accumulation and ice formation areas; establishing a background model of the tunnel road surface to identify changes in existing hazard areas (e.g., expansion of water accumulation, melting of ice formation). It should be noted that modern deep learning models usually integrate feature extraction, object detection/segmentation, and classification into an end-to-end network structure, which simplifies the technical process and improves the efficiency and accuracy of identification.

2.3. Common tool chains

The implementation of intelligent identification technology relies on mature software and hardware tools. The main tool chain includes as listed:

- (1) Programming language: Python is the main development language, which has rich image processing and deep learning libraries;
- (2) Image processing library: OpenCV is used for image preprocessing operations such as illumination adjustment, edge detection, and glare removal;
- (3) Deep learning framework: TensorFlow and PyTorch are used for model construction, training, and deployment, supporting the rapid development of detection and segmentation models;
- (4) Deployment platform: Edge computing platforms (e.g., NVIDIA Jetson series) are used for on-site real-time processing of tunnel image data, reducing the delay caused by cloud transmission and ensuring the real-time performance of early warning.

3. Conclusion

The intelligent image identification of water accumulation and ice formation in traffic tunnels is a complex technical system that comprehensively applies computer vision, deep learning, and multi-sensor fusion technologies. Its core principle is to fully leverage the differences in reflective, color, texture, and dynamic characteristics between water accumulation and ice formation, combined with environmental context information such as location, temperature, and weather, to achieve accurate identification of hazards. At present, the mainstream technical route is based on deep learning semantic segmentation (e.g., U-Net) and object detection models (e.g., YOLO), combined with efficient image preprocessing (e.g., illumination adjustment, glare removal) and post-processing technologies to improve the accuracy and robustness of identification. The tool chain is mainly composed of Python, OpenCV, TensorFlow/PyTorch, and edge computing platforms, which provides a mature technical foundation for the engineering application of the system. In conclusion, the intelligent identification system for water accumulation and ice formation in traffic tunnels has important practical significance and broad application prospects. With the continuous advancement of deep learning technology and the improvement of multi-source data fusion capabilities, the identification accuracy, real-time performance, and adaptability of the system will be further enhanced, which is expected to play a more critical role in ensuring traffic safety and extending the service life of tunnel structures.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Li X, Xiong Z, Li X, 2024, Research on Intelligent Detection of Apparent Defects in Old Tunnels Based on Semantic Segmentation. *Journal of Transportation Science and Technology*, 2024(6): 104–108.
- [2] Zhou Z, Li X, 2024, Design and Research of Intelligent Inspection Device for Highway Tunnels. *Western China Communications Science & Technology*, 2024(10): 123–125.
- [3] Du J, 2023, Research on Intelligent Identification of Tunnel Water Seepage Defects Based on Infrared Thermal Imaging, thesis, Beijing Jiaotong University.
- [4] Li Y, 2023, Machine Vision-Based Intelligent Recognition Modeling of Tunnel Cracks and Rock Mass Structure Prediction Method, thesis, Shandong University.
- [5] Zhang M, 2024, Intelligent Recognition Based on Deep Learning of Tunnel Ground Penetrating Radar Images, thesis, Chongqing Jiaotong University.
- [6] Li K, Zhang J, Nong Z, et al., 2024, Research on Safety Risk Identification and Prevention in Tunnel Project Construction Based on GIS. *Transport Energy Conservation & Environmental Protection*, 20(S2): 191–195.
- [7] Huang H, 2015, Intelligent Neural Networks and Their Application in Deformation Prediction and Evaluation During Tunnel Operation, Southwest Jiaotong University.
- [8] Wang D, 2024, Analysis of Intelligent Transportation Vehicle Recognition and Detection Technology. *Digital Communication World*, 2024(12): 126–128.
- [9] Wan F, 2024, Research on the Application of Intelligent Recognition Technology in Motor Vehicle Emission I/M System. *Transport Energy Conservation & Environmental Protection*, 20(6): 76–78.
- [10] Yu S, Yu F, Luo B, et al., 2023, Intelligent Recognition and Morphological Segmentation Method for Tunnel Lining Diseases in Ground Penetrating Radar Images. *Progress in Geophysics*, 38(3): 1408–1415.

Publisher's note

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

Transmission Line Defect Detection Algorithm Based on Improved RT-DETR Model

Qi Wu*

School of Computer Science and Technology, Taiyuan Normal University, Jinzhong 030619, China

*Correspondence author: Qi Wu, 2858478839@qq.com

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This paper addresses the urgent need for high-precision and high-efficiency visual perception technologies in power equipment operation and maintenance under the background of rapid development of smart grids. It points out the performance limitations of the existing real-time target detection framework RT-DETR when handling small targets, dense targets, and complex backgrounds in power inspection scenarios. To overcome this bottleneck, this study proposes an improved backbone network model, DETR-EVA, based on an efficient visual attention mechanism (EVA). This model innovatively designs an attention computation structure with linear complexity by deeply integrating the EVA mechanism with the C2f module in the RT-DETR backbone network, and combines local detail perception and global dependency modeling capabilities. Its core lies in the introduction of a gated fusion mechanism, which significantly enhances the model's ability to model long-distance contextual relationships and the adaptive adjustment efficiency of feature weights while retaining the advantages of multi-branch feature extraction and fusion of the C2f module. Experiments were conducted on an inspection image dataset containing typical power equipment targets. The results show that compared with the original RT-DETR model, DETR-EVA improves the overall accuracy index mAP50-95 by 2.5%, reduces computational complexity by 14%, and reduces the number of model parameters by 27%. This effectively verifies that the proposed method can significantly improve the detection accuracy of small targets and complex scenes while maintaining real-time detection speed, providing a better visual solution for intelligent operation and maintenance of power equipment.

Keywords: RT-DETR; Defect detection; Efficient vision attention; C2f; Small object detection

Online publication: February 13, 2026

1. Introduction

The rapid development of smart grids has raised higher demands on the intelligence level of power equipment operation and maintenance. High-voltage transmission lines, exposed to the natural environment for extended periods, inevitably suffer from potential defects such as insulator self-explosion, shock absorber slippage, and bird nest construction. If these defects are not detected and addressed in a timely manner, they can easily escalate into serious failures like line breaks and tower collapses, triggering widespread power outages and posing a severe

threat to people's lives and property, as well as to socio-economic order. Therefore, regular and efficient inspection of high-voltage transmission lines is an indispensable part of the preventive maintenance system in the power system.

Traditional transmission line inspection primarily relies on manual patrols. This method is not only inefficient and labor-intensive, but also limited by the vision and experience of the inspectors, making it difficult to detect some concealed defects. In recent years, with the maturity of drone technology, intelligent inspection has gradually become the mainstream method for transmission line inspection.

In recent years, deep learning technology, particularly object detection algorithms based on convolutional neural networks (CNN), has achieved a qualitative leap in detection accuracy and generalization performance, owing to its powerful end-to-end feature learning and expression capabilities, thereby revolutionizing the situation where traditional methods suffered from poor performance^[1]. Object detection algorithms are mainly divided into two-stage detectors (such as R-CNN, Fast R-CNN, Faster R-CNN) and single-stage detectors (such as YOLO^{[6][7]}, SSD)^[2-9].

Although the two-stage detector has high accuracy, its computational complexity is high and inference speed is slow, making it difficult to meet the urgent real-time requirements of inspection tasks. The YOLO series relies on prior anchor boxes, which limits its generalization ability. The non-maximum suppression post-processing is non-differentiable and inefficient, and the CNN structure has weak modeling of global contextual relationships in images.

To overcome these limitations, detection transformer (DETR) abandoned anchor boxes and NMS, utilizing the global attention mechanism of transformer to achieve end-to-end set prediction, significantly enhancing its global reasoning capability^[10,11]. However, DETR also introduces new issues, including slow training convergence, high computational cost, and poor performance in detecting small targets, which limit its application in real-time inspection scenarios.

RT-DETR, as a real-time high-performance variant in the DETR series, achieves a good balance between speed and accuracy through efficient hybrid encoder and intra-scale feature interaction design^[12]. However, there is still room for improvement in detection performance in complex scenes, especially in small object detection, dense object detection, and complex background processing. The original RT-DETR backbone network mainly has the following limitations: insufficient modeling of long-distance feature dependencies; limited feature representation ability, which is prone to false positives or missed detections in complex backgrounds or situations where the contrast between the target and background is low; and small object features are easily overwhelmed.

In response, this paper aims to make targeted improvements to the RT-DETR model to enhance its detection accuracy for small target defects on transmission lines, while maintaining its real-time advantage, thereby meeting the needs of practical engineering applications.

2. Literature review

2.1. Research on transmission lines based on RT-DETR

Early detection methods were primarily based on image processing techniques, such as edge detection, threshold segmentation, and texture analysis. These methods were computationally simple but lacked robustness and were highly susceptible to environmental interference. With the continuous development of deep learning, some scholars have conducted targeted research on the RT-DETR model.

For example, Li *et al.* addressed the issues of difficult-to-capture small-sized insulator defect features, insufficient utilization of contextual information, and unstable matching by designing a multi-scale backbone network, introducing a Self-Attention Upsampling (SAU) module, and a dedicated Insulator Defect (IDIoU) loss function ^[13]. This improved the model's detection capability for small defects, significantly enhancing average precision and enhancing detection stability. Bai *et al.* addressed the issues of shallow measurement and difficulty in quantification in traditional defect detection methods by establishing a magnetic flux leakage detection method and analyzing signal characteristics ^[14]. This improved the quantitative detection capability for U-shaped suspension ring defects, achieving high-precision, low-error defect identification. Huang *et al.* addressed the issues of existing detection models relying on a large amount of labeled data, bulky parameters, and the difficulty in balancing lightweight and performance ^[15]. By adopting a federated knowledge distillation framework combined with asynchronous aggregation and model freshness mechanisms, they improved the model's deployment capability on resource-constrained devices, achieving lightweight model implementation while enhancing detection accuracy and training efficiency. Xie *et al.* addressed the challenges of small target sizes, similar shapes, and occlusion leading to detection difficulties in power line defect detection ^[16]. By introducing a Transformer-based Power-DETR network, combined with multi-scale feature enhancement, contrastive denoising training, and mixed label assignment strategies, they improved detection accuracy and training stability. Chen *et al.* addressed the detection challenges of small defects on ultra-high voltage transmission lines being easily obscured and subject to strong complex background interference ^[17]. By adopting a feature focused diffusion network (FFDN) and dynamic range histogram self-attention (DHSA) mechanisms to improve the RT-DETR model, they achieved simultaneous optimization of detection accuracy and missed detection rate. This not only improved inspection efficiency by 60% but also significantly reduced energy consumption and carbon emissions, providing key technical support for low-carbon operation and maintenance of transmission lines.

In summary, these studies have demonstrated significant advantages in detecting small targets and overcoming the interference of occlusion and complex backgrounds. However, they generally suffer from issues such as complex model structures, large parameter counts, and high computational costs. To address this, this paper designs a lightweight and efficient RT-DETR model that can better detect minor defects while reducing detection costs.

2.2. Research on attention mechanism

Attention mechanisms originate from the simulation of the human visual system, and their core lies in guiding the model to allocate limited computational resources to the more critical parts of the input information. The specific development process is as follows:

- (1) The self-attention mechanism has been introduced, which directly computes the associations between all elements within the global scope through Query, Key, and Value operations, bringing powerful contextual modeling capabilities to the model ^[10];
- (2) Multi-head attention further extends this idea by enabling the model to learn information collaboratively from different representation subspaces through parallel computation of multiple attention heads ^[10];
- (3) CA (Cross-Attention) integrates features from different modalities or sources, with Query derived from one feature and Key and Value derived from another ^[10]. The vision transformer (ViT) demonstrated for the first time that splitting an image into a sequence of patches and directly applying a Transformer encoder can achieve performance on par with or even surpass that of the most advanced CNN models

on image classification tasks^[18]. This verifies the powerful ability of attention mechanisms in modeling global contextual dependencies in images. More importantly, attention mechanisms have given birth to groundbreaking object detection frameworks such as DETR^[10];

- (4) SENet (Squeeze-and-Excitation Attention) learns the dependencies between channels through global average pooling and a two-layer fully connected network, focusing solely on the channel dimension while ignoring spatial position information^[19]. To address this, CBAM is introduced, combining a hybrid mechanism of channel attention and spatial attention. It first weights the feature map through channel attention, and then focuses on important regions through spatial attention. However, the locality of convolution limits its ability to establish long-distance dependencies.

The adaptive hybrid encoder used in the RT-DETR real-time detection model is a representative design that dynamically integrates the efficient local feature extraction capability of CNNs with the global relationship modeling advantages of attention mechanisms. Our EVA module enhances the model's global context awareness and adaptive feature weight adjustment capabilities by integrating the EVA attention mechanism into the C2f structure, thereby improving its robustness in complex scenes and small object detection.

3. Improvement of the algorithm

3.1. RT-DETR model

RT-DETR is the first truly real-time end-to-end object detection framework proposed by Baidu Research^[12]. Its core innovation lies in breaking through the speed bottleneck of the traditional DETR model while maintaining high accuracy. In this paper, RT-DETR-l is selected as the benchmark model. This framework is based on an encoder-decoder architecture, as shown in **Figure 1**.

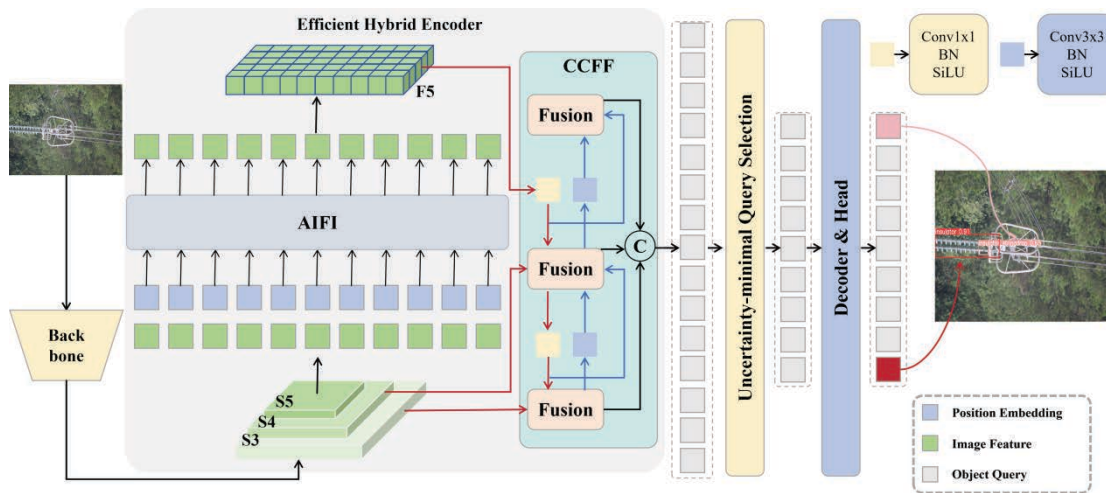


Figure 1. RT-DETR model.

RT-DETR, through a series of collaborative optimization designs, effectively balances accuracy and speed while maintaining the advantages of the end-to-end detection paradigm. Its core lies in an adaptive hybrid encoder, which innovatively integrates the local inductive bias of CNNs with the global modeling capabilities of transformers, and introduces an adaptive mechanism to dynamically allocate computational resources, thereby significantly reducing computational overhead while ensuring feature richness. The model employs a deeply

optimized backbone network that extracts multi-scale feature maps through an efficient C2f module, providing feature representations for detection tasks that combine high semantic information and fine spatial details. Finally, an efficient query-based decoder utilizes a small number of learnable query vectors to directly interact with the features output by the encoder, achieving accurate object localization and classification. Its concise detection head design eliminates the need for complex post-processing, further ensuring inference efficiency.

However, there is still room for improvement in the detection performance of RT-DETR in complex scenarios, especially in small object detection, dense object detection, and complex background processing.

3.2. DETR-EVA model

To address the issues of the original RT-DETR backbone network, this paper proposes the DETR-EVA model. By deeply fusing efficient vision attention (EVA) with the C2f module, it enhances the ability to perceive global context and preserve local details, thereby improving the model's detection accuracy for small targets. The logical structure of EVA is illustrated in **Figure 2**.

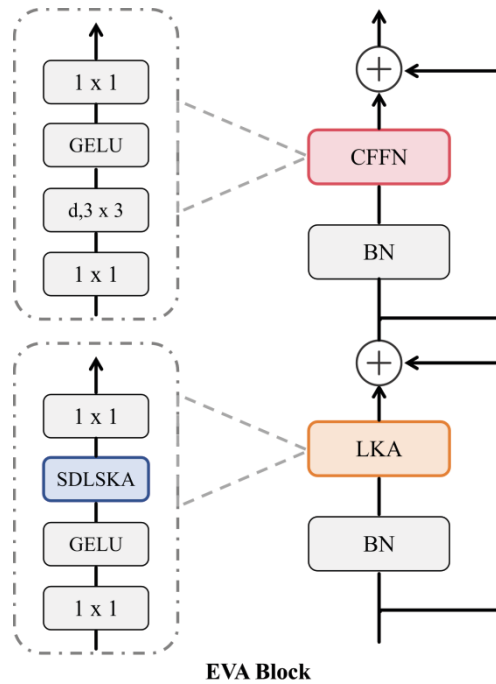


Figure 2. EVA framework.

The EVA module mainly consists of the following three parts:

- (1) Sparse decomposition large kernel attention (SDLSKA): SDLSKA decomposes large convolutional kernels into local convolutions and two orthogonal band-dilated convolutions. After extracting local features using a 5×5 convolution, it captures long-range dependencies through 1×11 and 11×1 depthwise separable convolutions with a dilation rate of 3, effectively expanding the receptive field to 35×35 . This design enhances the model's ability to model global semantics while reducing the number of parameters;
- (2) Integrated nuclear selection mechanism (CKS): CKS further introduces a dual-path attention mechanism of channel and spatial attention to dynamically fuse multi-scale features. Channel attention generates weights through global pooling and fully connected layers, while spatial attention aggregates max and average pooling features and generates spatial weights through convolution. The two are multiplied

element-wise to achieve adaptive feature selection, thereby highlighting key regions in complex backgrounds;

- (3) Convolutional feedforward network (CFFN): CFFN refines and enhances the channel dimension of the fused features through two pointwise convolutions and the GELU activation function, further improving the feature representation capability. The entire EVA module is embedded into the backbone network in a residual connection manner, which expands the receptive field and strengthens semantic understanding while maintaining the efficiency and practicality of the model.

The proposed DETR-EVA model addresses the challenges of small targets, complex backgrounds, and high real-time requirements in transmission line defect detection. By introducing a linearly complex EVA attention mechanism, it achieves efficient global context modeling under high-resolution features. The model integrates local and global attention and employs a gating mechanism to adaptively combine attention and convolutional features, significantly improving the ability to distinguish small target features.

4. Results and discussion

4.1. Dataset construction

The dataset used in this experiment is derived from images of transmission line defects captured by a drone from a certain company. The dataset comprises 7,612 images. In this experiment, Labelling tool was employed to annotate the images as label files in XML format, which were then converted to the YOLO-specific txt format using the convert function. A total of seven different categories of abnormal defect images were annotated, namely: insulator, insulator string drop, insulator breakage, insulator flashover, damper, damper defect, and nest.

The resolution of the images is 640*640 pixels, and they are divided into training set, validation set, and test set in a ratio of 7:2:1, with 5327 images in the training set, 1523 images in the validation set, and 762 images in the test set. Some images from the dataset are shown in **Figure 3**.



Figure 3. Partial defect images.

4.2. Experimental hyperparameter settings

This experiment was developed based on the Python 3.9.24 and PyTorch 2.2.2 frameworks. The hyperparameters for the experiment are shown in **Table 1**. In the model training of this experiment, the key hyperparameters were set to prioritize the final accuracy and training stability of the model.

Table 1. Hyperparameter settings

Parameter	Value
Training epochs	300
Batch size	4
Image size	640*640
Optimizer	AdamW
Automatic Mixed Precision (AMP)	False

The parameters in this experiment were carefully designed to ensure training effectiveness, result reliability, and comparability with mainstream research paradigms. The training epochs (300 epochs) provide ample convergence space for object detection tasks, especially for Transformer-based models. A small batch size (batch size = 4) and a moderate input image size (640×640) maintain stable gradient estimation with limited hardware resources and effectively control memory usage. The optimizer AdamW was chosen, whose built-in weight decay mechanism helps alleviate overfitting and promotes model generalization. Automatic mixed precision training (AMP) was kept off to prioritize numerical stability and reproducibility during training. Overall, this parameter configuration balances algorithm performance, training efficiency, and experimental reproducibility, conforming to common settings in related research within the field.

4.3. Model evaluation metrics

In the research on defect detection in power transmission lines, to scientifically evaluate the overall performance of the improved RT-DETR algorithm, this study uses Precision, Recall, F1-Score, and mean Average Precision (mAP50, mAP50-95) as the core evaluation metrics. The specific description of the evaluation metrics is as follows.

Precision, which measures the accuracy of the model in classifying positive cases. Its mathematical expression is as follows:

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

where TP (True Positives) represents the number of positive samples correctly predicted by the model, and FP (False Positives) represents the number of negative samples incorrectly predicted as positive by the model.

Recall, which measures the model's ability to identify and cover real positive samples. Its mathematical expression is as follows:

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

where FN represents real positive samples that the model incorrectly predicts as negative.

F1-Score, which is the harmonic mean of precision and recall, used to comprehensively evaluate the overall performance of a model, is defined as:

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (3)$$

This metric combines precision and recall into a single unified measure, making it suitable for scenarios where insulator defects require an optimal balance between false positives and false negatives, providing a robust comprehensive benchmark for model performance.

Mean average precision (mAP), which evaluates the model's classification accuracy and localization ability comprehensively by calculating the average precision across all detection categories. Its calculation formula is:

$$AP = \int_0^1 P(R) dR \quad (4)$$

where P denotes Precision and R denotes Recall.

Then, the mean Average Precision (mAP) is obtained by taking the arithmetic mean of AP values across all categories:

$$mAP = \frac{1}{N} \times \sum_i AP_i \quad (5)$$

where N is the total number of categories.

In practical evaluation, mAP50 refers to the mAP value calculated with a fixed Intersection over Union (IoU) threshold of 0.5, which mainly evaluates the basic detection capability, mAP50-95 is the average of multiple mAP values calculated with IoU thresholds ranging from 0.5 to 0.95, more comprehensively reflecting the model's overall performance in both accurate recognition and precise localization.

4.4. Experimental results

The improved model was comprehensively evaluated on the dataset in this paper, and the experimental results comparing it with the baseline model RT-DETR are shown in **Table 2**.

Table 2. Experimental results

Model	P	R	F1	mAP50	mAP50-95	GFLOps	Parameters	Model size
RT-DETR	0.918	0.868	0.892	0.913	0.634	57.0	19.8M	77.0MB
DETR-EVA	0.924	0.878	0.900	0.920	0.659	48.8	14.5M	56.6MB

Analysis of **Table 2** shows that the proposed DETR-EVA model significantly outperforms the benchmark RT-DETR model across all key performance indicators, achieving a synergistic optimization of accuracy and efficiency. Specifically, in terms of detection accuracy, the model's overall performance index mAP50-95 reaches 0.659, a significant improvement of 2.5% compared to the baseline. This directly verifies the effectiveness of the deep fusion of the EVA attention mechanism and the C2f module in enhancing the model's feature modeling capabilities, especially in complex scenes and small object detection. Meanwhile, the model's lightweight performance is even more remarkable, where the computational complexity (GFLOPs) is reduced to 48.8, a

decrease of 14%; the number of parameters is compressed to 14.5M, a reduction of 27%. This is mainly due to the linear complexity attention design and gating fusion mechanism in the proposed method, which efficiently filters key features while introducing global context dependencies, avoiding redundant computation.

4.5. Comparative experiment

To comprehensively evaluate the overall performance of the improved model proposed in this paper, this study selected seven mainstream object detection algorithms, Faster R-CNN, Cascade R-CNN, YOLOv5n, YOLOv7-tiny, YOLOv8n, YOLOv10n, and YOLOv11n, as benchmarks for comparison and conducted systematic comparative experiments on the same transmission line defect dataset. The results are shown in **Table 3**.

Table 3. Comparative experiments

Model	P	R	mAP50	mAP50-95	F1
Faster R-CNN	0.802	0.736	0.792	0.516	0.783
Cascade R-CNN	0.826	0.749	0.813	0.523	0.796
Yolov5n	0.852	0.751	0.840	0.546	0.809
YOLOv7-tiny	0.856	0.781	0.834	0.544	0.810
Yolov8n	0.866	0.760	0.842	0.561	0.816
YOLOv10n	0.857	0.778	0.834	0.557	0.807
Yolov11n	0.863	0.770	0.848	0.565	0.812
RT-DETR	0.918	0.868	0.913	0.634	0.892
DETR-EVA	0.924	0.878	0.920	0.659	0.900

This demonstrates that the proposed improved model (EVA) exhibits comprehensive and significant advantages across all core metrics. It achieves the highest precision and recall, and its overall performance metrics, including mAP50, mAP50-95, and F1 score, significantly outperform all compared mainstream algorithms. This indicates that the model not only excels in detection accuracy but also maintains stronger robustness under a stricter intersection-union threshold (mAP50-95), achieving a better balance between precision and recall, thus validating its superior overall detection performance.

The performance of each model is visually compared and contrasted using horizontal bar charts and normalized radar charts, as shown in **Figure 4**.

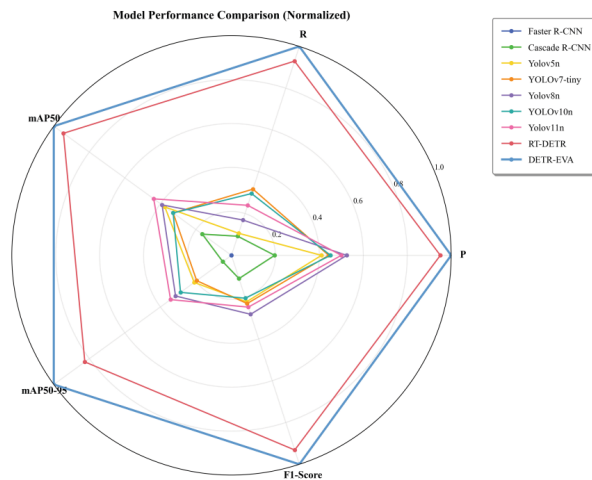


Figure 4. Normalized radar effect.

The DETR-EVA model comprehensively outperforms traditional and next-generation detectors, including Faster R-CNN, Cascade R-CNN, and RT-DETR, in radar image rendering. Its strategy of integrating EVA attention and C2f modules significantly improves the detection capability for small targets and complex backgrounds on power transmission lines by strengthening global context modeling and adaptive feature selection. While maintaining real-time inference, it achieves significant improvements in accuracy and robustness, laying a technological foundation for efficient and high-precision applications in power line inspection.

4.6. Visual analysis

To provide a more intuitive and qualitative assessment of the model's detection capabilities in complex real-world scenarios, beyond quantitative metrics, this study randomly selected six representative transmission line inspection images from the test set for inference visualization comparison. These images cover typical challenges such as small targets, multi-scale targets, cluttered backgrounds, uneven lighting, and target occlusion. **Figure 5** shows a comparison of the detection results of the unimproved RT-DETR model (**Figure 5a**) and the proposed DETR-EVA model (**Figure 5b**) on the same samples.

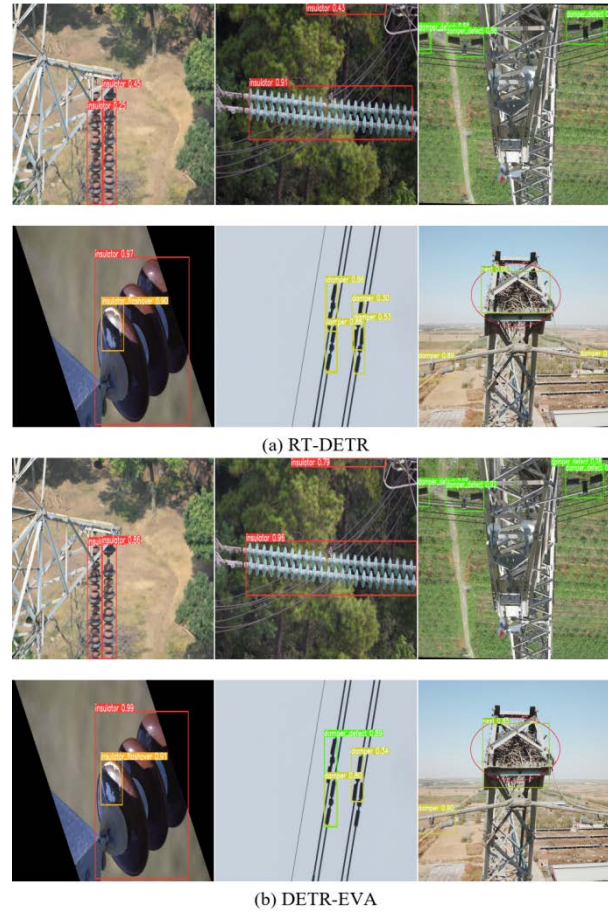


Figure 5. Target detection results.

Direct observation reveals that the DETR-EVA model exhibits superior performance in detecting small-sized insulator defects, distinguishing dense targets, and controlling false alarms in complex backgrounds, intuitively verifying its stronger robustness and practicality in real-world scenarios.

5. Conclusion

This paper addresses the challenges of small targets, complex backgrounds, and high real-time requirements in power transmission line defect detection. It proposes a modified RT-DETR model, DETR-EVA, based on EVA. Through structural fusion of EVA and C2f and a gating adaptive strategy, the model achieves efficient collaboration between global context and local details, significantly enhancing its feature representation ability for small defects while maintaining linear complexity. Experiments show that this model comprehensively outperforms the original RT-DETR and mainstream lightweight models in terms of accuracy (mAP and recall), while further reducing computational overhead and parameter count, effectively balancing accuracy and speed. Future research will explore semi-supervised learning to utilize unlabeled data and improve the model's generalization ability in rare defects and unknown scenarios.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Tao T, Li Z, 2026, A Review of Deep Learning Application in Transmission Line Defect Detection. *Electric Power Systems Research*, 2026(250): 112193.
- [2] Girshick R, Donahue J, Darrell T, et al., 2014, Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 580–587.
- [3] Girshick R, 2015, Fast r-cnn. *Proceedings of the IEEE International Conference on Computer Vision*, 1440–1448.
- [4] Ren S, He K, Girshick R, et al., 2016, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6): 1137–1149.
- [4] Redmon J, Divvala S, Girshick R, et al., 2016, You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779–788.
- [5] Redmon J, Farhadi A, 2017, YOLO9000: Better, Faster, Stronger. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7263–7271.
- [6] Redmon J, Farhadi A, 2018, Yolov3: An Incremental Improvement, *arXiv preprint*, <https://doi.org/10.48550/arXiv.1804.02767>
- [7] Khanam R, Hussain M, 2024, Yolov11: An Overview of the Key Architectural Enhancements, *arXiv preprint*, <https://doi.org/10.48550/arXiv.2410.17725>
- [8] Liu W, Anguelov D, Erhan D, et al., 2016, Ssd: Single Shot Multibox Detector. *European Conference on Computer Vision*, Springer International Publishing, 21–37.
- [9] Carion N, Massa F, Synnaeve G, et al., 2020, End-to-End Object Detection with Transformers. *European Conference on Computer Vision*, Springer International Publishing, 213–229.
- [10] Dosovitskiy A, 2020, An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv preprint*, <https://doi.org/10.48550/arXiv.2010.11929>
- [11] Lv Y, Chen W, Chen B, et al., 2023, RT-DETR: DETRs Beat YOLOs on Real-Time Object Detection. *Proc. of the IEEE/CVF Int. Conf. on Computer Vision Workshops (ICCVW)*, 2420–2430.
- [12] Li D, Yang P, Zou Y, 2024, Optimizing Insulator Defect Detection with Improved DETR Models. *Mathematics*, 12(10): 1507.

- [13] Bai X, Ji H, Wang L, et al., 2024, Detection Method for Structural Defects of U-Shaped Hanging Ring for Power Fittings Based on Magnetic Leakage Principle. *Engineering Research Express*, 6(2): 025339.
- [14] Huang X, Jia M, Tai X, et al., 2024, Federated Knowledge Distillation for Enhanced Insulator Defect Detection in Resource Constrained Environments. *IET Computer Vision*, 18(8): 1072–1086.
- [15] Xie Z, Dong C, Zhang K, et al., 2024, Power-DETR: End-to-End Power Line Defect Components Detection Based on Contrastive Denoising and Hybrid Label Assignment. *IET Generation, Transmission & Distribution*, 18(20): 3264–3277.
- [16] Chen W, Li S, Han X, 2025, IDD-DETR: Insulator Defect Detection Model and Low-Carbon Operation and Maintenance Application Based on Bidirectional Cross-Scale Fusion and Dynamic Histogram Attention. *Sensors*, 25(18): 5848.
- [17] Vaswani A, Shazeer N, Parmar N, et al., 2017, Attention is All You Need. *Advances in Neural Information Processing Systems*, 2017: 30.
- [18] Hu J, Shen L, Sun G, 2018, Squeeze-and-Excitation Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7132–7141.
- [19] Woo S, Park J, Lee J, et al., 2018, Cbam: Convolutional Block Attention Module. *Proceedings of the European Conference on Computer Vision (ECCV)*, 3–19.

Publisher's note

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

Deep Learning-Based Highway Rockfall Early Warning System

Shipeng Xu*, Mingyu Xue

College of Civil Engineering and Transportation, Northeast Forestry University, Harbin 150040, China

**Author to whom correspondence should be addressed.*

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This paper proposes a deep learning-based rockfall warning system for mountainous road curves. It utilizes drone inspections combined with the YOLOv11 object detection algorithm to accurately identify rockfalls on road surfaces, while employing ground-based millimeter-wave radar for real-time vehicle detection. The system features a comprehensive curve blind spot warning mechanism and incorporates a wireless communication module to push instant alerts to mobile navigation terminals based on rockfall risk and vehicle location. This system effectively addresses the challenges of rockfall identification and delayed warnings within blind spots on curves. It reduces manual inspection costs while significantly enhancing driving safety on mountainous roads.

Keywords: Deep learning; YOLOv11; Intelligent warning; Highway blind spots; Rockfall hazards

Online publication: February 12, 2026

1. Introduction

With the continuous advancement of transportation infrastructure development, mountain roads play a crucial role in safeguarding regional economic growth and facilitating daily life. However, these roads frequently face severe threats from geological hazards, with rockfalls being a common mountainous geological disaster that poses significant risks to transportation infrastructure and driving safety along these routes^[1]. On one hand, rockfall disasters are influenced by multiple complex factors such as topography, climate, hydrology, and geological structures, making them prone to collapse during triggers like heavy rainfall or earthquakes^[2]. On the other hand, although current mitigation measures include physical protections such as installing protective nets, reinforcing unstable rock formations, and constructing shelters or tunnels, operational management still primarily relies on manual inspections and traditional guarding methods^[3].

This approach suffers from high inspection costs and significant risks of missed detections during nighttime or adverse weather conditions. Therefore, developing an efficient intelligent rockfall early warning system is crucial for ensuring driving safety on mountain roads. Against this backdrop, this project designed an intelligent rockfall warning system for mountainous road curves based on the synergy of deep learning and millimeter-wave radar. The system achieves precise identification of road surface rockfalls through drone inspections combined

with the YOLOv11 algorithm, while utilizing millimeter-wave radar for real-time detection of ground vehicles. The system incorporates a comprehensive blind-spot warning mechanism for curves, effectively addressing the challenges of rockfall identification and delayed warnings in complex environments. It provides essential technical support for enhancing driving safety on mountainous highways.

2. Framework Design

The system comprises an information acquisition system and an alert dissemination system. The information acquisition system monitors rockfall conditions on the road surface, uploading relevant parameters such as geographic coordinates and image features to enable real-time perception of the curve environment while simultaneously transmitting data to the control terminal. The system utilizes aerial drone inspections combined with the YOLOv11 object detection algorithm to accurately identify rockfalls, ensuring the system can detect risks at the earliest possible moment.

The early warning release system, also known as the response system, employs ground-based millimeter-wave radar to continuously track vehicle locations. Upon detecting a vehicle entering a rockfall hazard zone, it immediately pushes real-time alerts to mobile navigation terminals. The control terminal interface displays the precise location of the rockfall, enabling drivers to take evasive action based on the warning information.

Both systems communicate with the edge server via IoT modules. These modules employ wireless communication units capable of handling large data exchanges while minimizing latency. The communication method involves periodically sending heartbeat packets to the terminal server at the application layer to maintain connection vitality and ensure communication reliability.

Figure 1 illustrates the overall architecture of the mountainous road curve rockfall warning system. It comprises:

- (1) Mobile navigation terminals as display units;
- (2) Drones and millimeter-wave radars as sensing units;
- (3) Edge servers as system control units;
- (4) Wireless communication modules as wireless communication and signal transmission units;
- (5) Power supply units providing energy to other system components.

The overall system architecture is as follows: the UAV module, wireless communication module, millimeter-wave radar, and camera are each connected to the microcontroller. The power supply unit connects to each sensing module and the microcontroller. The microcomputer and control terminal use the wireless communication module as a bridge to ensure uninterrupted information flow.

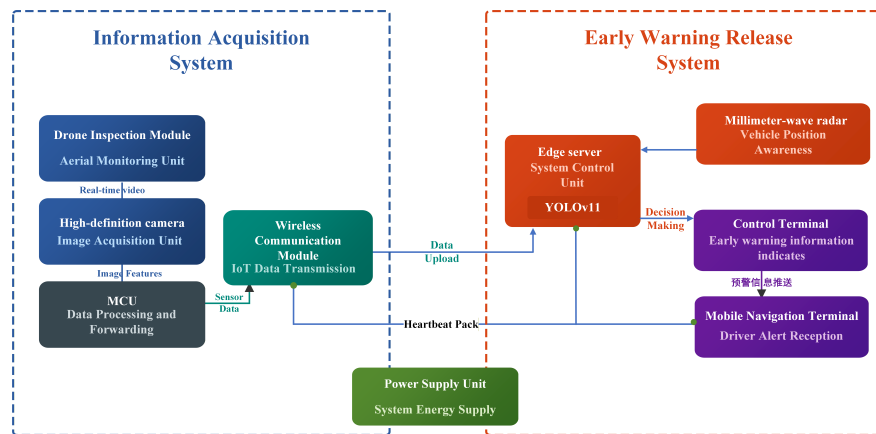


Figure 1. Overall framework diagram of the system.

3. Methods

3.1. Millimeter-wave radar-based vehicle perception

When a vehicle enters a blind spot on a curve, the system uses ground-based millimeter-wave radar to obtain distance information as the primary trigger for early warning. The radar sensor emits electromagnetic waves and receives echoes, using signal processing algorithms to calculate the target's motion state. As the vehicle approaches the curve, the detected distance value gradually decreases. When the target distance signal falls below the safety threshold, the warning monitoring mechanism activates and dynamically responds through the following scheme:

- (1) When no vehicles are present within the monitoring area or vehicles are beyond the safety distance, the system enters low-power cruise mode, performing only rockfall environment scans to conserve system energy;
- (2) When a vehicle enters the radar detection range, the system immediately activates high-frequency monitoring mode. By integrating real-time rockfall data transmitted from drones, it assesses potential rockfall risks during the vehicle's passage. The system stands ready to send alerts to mobile navigation terminals, prompting drivers to exercise caution regarding road conditions ahead and reducing accidents caused by blind spots.

3.2. Deep learning-based precise road-fall identification

Combining the YOLOv11 object detection algorithm with real-time processing of high-definition imagery transmitted from drone inspections, we quantify rockfall risk levels on mountainous road curves. We first collected road surface image data under varying lighting, weather, and terrain conditions, meticulously annotating the location and category of rockfall targets within the images. Next, we employed a pre-trained YOLOv11 model for feature extraction and object detection. By feeding real-time video frames captured by the drone into this model, we obtained rockfall detection results. The system analyzes continuous video streams to filter out environmental distractions like foliage and shadows, precisely identifying anomalous objects on the road surface. Eventually, based on the detected rockfall locations, we determine whether the rocks are positioned along the vehicle's trajectory, thereby enabling real-time road surface rockfall risk assessment.

3.3. IoT-based communication establishment and data interaction

The system's communication module employs a combination of IoT technology and TCP/IP protocols for real-time data transmission. First, an edge server is established to connect the wireless communication module with the UAV, radar, and mobile navigation terminals. The UAV serves as the image acquisition terminal, while the radar functions as the vehicle perception terminal. They generate separate data frames containing detected rockfall coordinates and vehicle positions, respectively, transmitting these as data streams via wireless networks to the communication module. The communication module packages this heterogeneous data for transmission to the edge server, where it undergoes fusion analysis. The resulting decision is then pushed to the control terminal. Continuous packet transmission between the communication module, edge server, and all terminals ensures the uninterrupted flow of the data transmission link.

3.4. Early warning decision logic

For the early warning system, we fuse the rockfall data detected by drones with vehicle information sensed by millimeter-wave radar on edge servers. The system employs a Kalman filter algorithm to smooth multi-source data, eliminating fluctuations caused by environmental noise. When both conditions, "road surface rockfall

present” and “vehicle distance below safety threshold”, are simultaneously detected, the system identifies a high-risk state. At this point, the edge server transmits an early warning signal to the mobile navigation terminal via the wireless communication unit.

4. Hardware design

4.1. Core control and multi-dimensional perception unit

The intelligent early warning system employs a high-performance microcontroller (MCU) as the core unit for on-site data acquisition and processing. This controller features extensive peripheral interfaces and robust data throughput capabilities, enabling concurrent processing of multi-source sensor data. The main control board connects to ground-based millimeter-wave radar and drone-mounted high-definition cameras via high-speed interfaces. The millimeter-wave radar provides real-time vehicle information to the main control board, while the high-definition camera captures real-time road surface imagery. Acting as the central hub of the perception layer, the MCU performs preliminary cleaning and packaging of radar signals and image data based on embedded low-level drivers, establishing a high-quality data foundation for subsequent edge computing and early warning.

4.2. Wireless communication and data transmission module

The communication module employs industrial-grade wireless communication units to establish real-time connections between the MCU, edge servers, and control terminals. Data transmission modes primarily include reliable transmission (TCP) and unreliable transmission (UDP). Generally, TCP/IP-based transmission is connection-oriented, providing reliable byte-stream delivery with retransmission mechanisms to prevent data loss. In contrast, UDP transmission is connectionless and broadcast-oriented, offering higher transmission efficiency but carrying a risk of packet loss. Given the mountain rockfall warning system’s stringent requirements for safety and data integrity, it is imperative that every warning command is delivered accurately and without error. Therefore, this system adopts the more secure TCP/IP communication method. Additionally, in the application layer design, the communication module periodically sends heartbeat packets to the server to maintain link viability. Should a connection fail, it can immediately reconnect, ensuring extremely high reliability of the communication system.

5. Software design

The control terminal features a graphical user interface designed for visual fusion and intuitive interaction with multi-source perception data. Based on the system’s deep learning algorithm deployment requirements and hardware characteristics, LabVIEW programming language was selected to develop the warning control software ^[4]. The intelligent early warning system control software developed based on LabVIEW is a multi-module application. It integrates functional modules such as video stream processing, radar data analysis, risk assessment, and early warning dissemination. This software not only provides managers with a panoramic view of curve safety monitoring but also enables real-time linkage with mobile navigation terminals through backend logic.

5.1. Real-time monitoring and visualization interface

The core area of the software’s main interface embeds a real-time video window, synchronously displaying high-definition imagery transmitted by drones. After loading the YOLOv11 model, the system performs real-time annotation of identified rockfall targets within the video stream, drawing red bounding boxes and displaying

confidence levels. This enables managers to intuitively identify road surface anomalies. The interface sidebar features a radar data visualization panel displaying millimeter-wave radar-detected vehicle distance information as a distance-time curve graph. When a vehicle enters the monitoring range, the system refreshes the relative distance between the vehicle and the blind spot at the curve in real time.

5.2. Intelligent early warning mechanism

In the deep learning-based highway curve rockfall warning system, the warning issuance module serves as a critical defense for traffic safety. When the system determines via algorithmic fusion that “rockfall is present on the roadway” and “vehicles are within a hazardous distance,” the software automatically triggers a tiered warning process as follows:

- (1) Host machine audio-visual alarm: The control terminal interface immediately displays a prominent red warning window showing the rockfall’s precise location and the vehicle’s current distance, accompanied by an alarm tone. This alerts monitoring center personnel to immediately focus attention and implement emergency measures;
- (2) Mobile instant push notification: The software backend invokes network communication interfaces to package the warning information and push it via TCP/IP protocol to the mobile navigation terminals of drivers entering the affected area. The push content includes concise and clear evasion instructions, ensuring drivers receive sufficient reaction time before entering the blind zone, thereby effectively reducing accident risks.

6. Conclusion

This paper proposes an intelligent rockfall early warning system for mountainous road curves based on the synergy of deep learning and millimeter-wave radar. The system utilizes wireless communication modules as data transmission tools, connecting the drone inspection terminal, radar sensing terminal, and edge computing server. It employs the TCP/IP protocol to ensure the smooth and highly reliable operation of the early warning link. The system achieves precise identification of road surface rockfalls in blind spots of curves and real-time perception of passing vehicles, enabling instant alerts to be pushed to mobile navigation terminals at the first sign of risk. This design not only significantly reduces the manpower and material consumption required for post-construction maintenance of mountain roads but also technically overcomes the safety bottleneck caused by limited visibility on curves. It promotes the development of traffic geological hazard monitoring toward unmanned and intelligent approaches. In summary, compared to traditional passive protection nets and manual monitoring models, the rockfall warning system designed in this paper demonstrates significant advantages in response speed, detection accuracy, and system coordination. It holds promising application prospects and significant value for widespread adoption.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Xu H, Zou P, Yu Z, et al., 2022, Design Method of Guided Flexible Buffering System for High and Steep Slopes of

Mountain Highways. *China Journal of Highway and Transport*, 35(9): 235–246.

- [2] Wang D, 2023, Research on Prevention and Control Technology of Dangerous Rockfall Hazards on High Slopes of Mountain Railways. *Engineering Technology Research*, 8(1): 205–207.
- [3] Cheng Y, 2022, Remediation of Rockfall Hazards on Mountain Railways. *Yangtze River Technology and Economy*, 6(S1): 4–7.
- [4] Hanli W, Yuanzhi L, Yilin W, 2024, Street Lamp Status Warning System Based on Internet of Things Technology. *Journal of Electronic Research and Application*, 8(4): 154–160.

Publisher's note

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

Multi-Modal Risk Profiling-Driven Power Grid Disaster Emergency Response Strategies and Dynamic Resource Synergy Optimization Model

Zi'an Zhong, Kun Hua, Xu Huang, Shiyu Chen, Ruiqi Chen

Shenzhen Longhua Power Supply Bureau, Shenzhen 518110, Guangdong, China

**Author to whom correspondence should be addressed.*

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: To improve the efficiency of power grid emergency response after disasters, this study proposes a multi-modal risk profiling-driven power grid disaster emergency response strategy and dynamic resource synergy optimization model. A risk assessment model is constructed by integrating equipment health status, real-time failure rate, and power grid topology importance to generate equipment risk profiles for identifying key nodes. A two-stage optimization mechanism is then designed, the first stage achieves priority coverage of high-risk equipment and minimization of inspection costs through multi-objective path planning. The second stage adopts a mixed-integer programming model to coordinate personnel scheduling and material allocation under resource constraints. A rolling optimization framework is introduced to dynamically respond to sudden failures and resource changes, ensuring the adaptability of scheduling schemes. To verify the model's effectiveness, three typical scenarios, "no sudden failures", "equipment risk escalation", and "personnel working hour constraints", are simulated. Compared with traditional strategies, the model significantly improves the rationality and dynamic adaptability of resource scheduling, providing new ideas and engineering practice support for enhancing the resilience of smart grid disaster emergency response.

Keywords: Multi-modal risk profiling; Power grid disaster emergency response; Mixed-integer programming; Path planning

Online publication: February 27, 2026

1. Introduction

Power grid disaster emergency response is an extremely complex systematic project, covering key links such as fault diagnosis, equipment maintenance, resource allocation, and load restoration^[1]. Traditional power grid emergency response models mainly rely on manual experience and pre-designed emergency plans. However, when facing complex and variable disaster scenarios, this model exposes many drawbacks, such as slow response speed, poor decision-making scientificity, and low resource utilization efficiency^[2].

This paper proposes a power grid disaster emergency response and resource optimization scheduling method based on risk profiling. It evaluates disaster risks by constructing a comprehensive and dynamic risk profiling system; establishes an optimization scheduling model considering resource synergy and dynamic coupling to achieve efficient allocation of emergency resources; and designs systematic emergency response strategies combined with the characteristics of multi-disaster scenarios to improve the power grid's ability to respond to complex disasters.

2. Mathematical model

2.1. Construction of risk profiling

2.1.1. Monitoring status of distribution equipment

According to the scoring criteria in the Guidelines for Distribution Network Equipment Status Evaluation, a 100-point scale is used to calculate the score of each component of distribution equipment. Due to inconsistent scoring principles for each evaluation index, normalization processing is required, namely:

$$\begin{cases} x_i^{(0)} = 100 - \Delta x_i^0 \\ x_i = 1 - \exp\left(-\frac{(x_i^{(0)} - x_{min})^2}{(x_{max} - x_{min})^2}\right) \end{cases} \quad (1)$$

Where: $x_i^{(0)}$ is the score of the i-th evaluation index; Δx_i^0 is the deduction score of the i-th evaluation index; x_i is the normalized score of the i-th evaluation index; x_{max} and x_{min} are the upper and lower limits of the evaluation index, respectively.

The comprehensive score of each component, i.e., the equipment health index H (0~100), is calculated according to the weight of each component, which can be specifically divided into normal status, attention status, abnormal status, and severe status, as shown in **Table 1** ^[3].

Table 1. Classification of equipment health status

Equipment health status	Score
Normal	85 < H ≤ 100
Attention	75 < H ≤ 85
Abnormal	60 < H ≤ 75
Severe	H < 60

The real-time health index reflecting the health status of distribution equipment is obtained through equipment health status assessment, as shown in the formula:

$$H = \sum_{j=1}^m \omega_j X_j \quad (2)$$

Where: ω_j is the weight of each component of the distribution equipment; X_j is the score of each component of the distribution equipment; j is the number of components of the distribution equipment.

2.1.2. Equipment failure rate

The equipment failure rate is closely related to the equipment health status. The worse the equipment health status

and the more severe the aging, the higher the possibility of failure. An exponential model is used for description:

$$\lambda = Ke^{-fH} \quad (3)$$

Where: λ is the real-time failure rate; K and f are undetermined coefficients, which can be obtained by inverting or fitting equipment failure rate and health index data for more than two years.

2.1.3. Power grid topology importance index

The power grid topology structure is crucial for evaluating equipment importance, failures of key equipment have a far greater impact on the power grid than non-key equipment. The criticality of equipment is determined according to the power grid topology. A common approach is to divide equipment into main equipment and secondary equipment. Main equipment is usually an indispensable part of the power grid, such as transformers and circuit breakers in substations, while secondary equipment is auxiliary or redundant equipment^[4]. Equipment importance is then calculated by evaluating the connectivity between the equipment and other equipment in the power grid. Common methods include using network analysis methods such as Katz centrality or Betweenness centrality based on power grid topology to evaluate equipment importance.

Katz centrality considers the direct and indirect connections of equipment in the power grid to assess its influence. The formula is as follows:

$$C_{Katz}(i) = \sum_{j \in N(i)} \frac{1}{d(i,j)} \quad (4)$$

Where: $N(i)$ is the set of adjacent nodes directly connected to equipment i , $d(i,j)$ is the distance between equipment i and equipment j ;

Betweenness centrality measures the role of equipment as a bridge in the power grid. The formula is as follows:

$$C_B(i) = \sum_{s,t \in V} \frac{\sigma(s,t|i)}{\sigma(s,t)} \quad (5)$$

Where: $\sigma(s,t)$ is the number of shortest paths from node s to node t ; $\sigma(s,t|i)$ is the number of shortest paths from node i .

2.1.4. Construction of risk profiling

The health status, failure rate, and equipment importance are combined to form a complete risk profile of the equipment. Considering equipment health status, failure rate, and equipment importance comprehensively, the comprehensive risk value of each equipment is calculated as R_i :

$$R_i = \alpha_1 \cdot H_i + \alpha_2 \cdot \lambda_i + \alpha_3 \cdot I_i \quad (6)$$

Where: R_i is the comprehensive risk value of the equipment i ; H_i is the equipment i health status score; λ_i is the equipment i failure rate; I_i is the equipment i importance score evaluated based on the power grid topology structure; $\alpha_1, \alpha_2, \alpha_3$ are weight coefficients of different dimensions, indicating the importance of each dimension in equipment risk assessment.

2.2. Multi-objective optimization model

2.2.1. Model structure

The model is divided into two stages: the first stage constructs an inspection path optimization model with the objectives of minimizing total path cost and maximizing coverage of high-risk equipment; the second stage constructs a protective resource scheduling optimization model with the objectives of maximizing the completion

rate of inspection and protection tasks for high-priority equipment and minimizing resource allocation costs ^[5].

The overall optimization problem is defined as follows:

$$\min F(X, Y) = [f_1(X), -f_2(X), f_3(Y), -f_4(Y)] \quad (7)$$

$$s. t. g_i(X, Y) \leq 0, i = 1, 2, \dots, m$$

$$h_j(X, Y) = 0, j = 1, 2, \dots, n$$

Where: $X=[x_{ij}]$ is the path selection decision variable matrix; $Y=[y_{ij}]$ is the personnel-equipment allocation matrix; $f_1(X)$ is the total path cost; $f_2(X)$ is the weighted coverage of high-priority equipment; $f_3(Y)$ represents resource and scheduling constraints; $f_4(Y)$ is risk coverage effect; g_i, h_j represents resource and scheduling constraints.

2.2.2. Path optimization model

Combining the spatial location and traffic accessibility of equipment, an inspection path graph is constructed $G = (V, E)$, where nodes V represent equipment to be inspected, edge sets E represent accessible paths between equipment, and edge weights construct a path cost function considering geographic distance, road grade, weather impact, and other factors:

$$C_{ij} = \alpha \cdot d_{ij} + \beta \cdot \omega_{ij} + \gamma \cdot t_{ij} \quad (8)$$

Where: d_{ij} is the Euclidean distance between nodes; ω_{ij} is the road grade penalty coefficient; t_{ij} is the weather impact factor; α, β, γ are harmonic coefficients.

Define the path decision variable $x_{ij} \in \{0, 1\}$, which is if moving from node i to node j . The two-objective function at the path level is as follows:

Minimization of total inspection path cost:

$$f_1(X) = \sum_{i=1}^N \sum_{j=1}^N C_{ij} \cdot x_{ij} \quad (9)$$

Maximization of high-priority equipment coverage:

$$f_2(X) = \sum_{i=1}^N R_i \cdot \left(\sum_{j=1}^N x_{ij} \right) \quad (10)$$

2.2.3. Resource scheduling optimization model

After path selection, available resource information (inspection personnel, material types and quantities, maximum working hours, etc.) is input to construct an integer programming model to optimize the allocation of protective resources for risky equipment ^[6]. Define as follows:

- (1) $y_{ik} \in \{0, 1\}$ indicates whether the inspectors k are responsible for the equipment i ;
- (2) $z_{mk} \in Z_+$ indicates the quantity of materials m carried by the personnel k ;
- (3) r_i indicates the priority level of equipment i protection;
- (4) t_{ik} indicates the time required for the personnel k to complete the equipment i inspection;
- (5) c_{ik} indicates the dispatch cost, including path, transportation, and time costs.

The objective function of the scheduling stage is as follows:

Minimization of resource scheduling cost:

$$f_3(Y) = \sum_{i \in D} \sum_{k \in P} c_{ik} \cdot y_{ik} + \sum_{k \in P} \sum_m z_{mk} \quad (11)$$

Maximization of protection task coverage:

$$f_4(Y) = \sum_{i \in D} \sum_{k \in P} r_i \cdot y_{ik} \quad (12)$$

In summary, the objective function is to minimize the comprehensive cost of post-disaster power grid emergency inspection and resource scheduling while maximizing the priority coverage of high-risk distribution equipment. This objective covers factors such as inspection path cost, personnel working hours, and material scheduling expenses, while considering the priority repair value reflected by the equipment risk profile. These factors together constitute the objective function of the optimization model and are incorporated into the subsequent path selection and resource scheduling constraints^[7]. The comprehensive optimization objective of the emergency scheduling system is defined as follows:

$$\max Z_1 = \sum_{i \in N} R_i \cdot \left(\sum_{j \in N} x_{ij} \right) \quad (13)$$

$$\min Z_2 = \sum_{i \in N} \sum_{j \in N} C_{ij} \cdot x_{ij} \quad (14)$$

$$\max Z_3 = \sum_{i \in D} \sum_{k \in P} r_i \cdot y_{ik} \quad (15)$$

$$\min Z_4 = \sum_{i \in D} \sum_{k \in P} t_{ik} \cdot y_{ik} + \sum_{k \in P} \sum_{m \in M} z_{mk} \quad (16)$$

$$s. t. \sum_{j \in N} x_{ij} \leq 1, \forall i \in N \quad (17)$$

$$\sum_{i \in N} x_{ij} \leq 1, \forall j \in N \quad (18)$$

$$\sum_{k \in P} y_{ik} \leq 1, \forall i \in D \quad (19)$$

$$\sum_{i \in D} t_{ik} \cdot y_{ik} \leq T_k^{\max}, \forall k \in P \quad (20)$$

$$z_{mk} \geq \sum_{i \in D} R_{im} \cdot y_{ik}, \forall k \in P, \forall m \in M \quad (21)$$

$$\sum_{k \in P} z_{mk} \leq Q_m, \forall m \in M \quad (22)$$

The first-stage objective functions (13) and (14) indicate that in the inspection path stage, the system expects to cover more high-risk equipment (maximizing Z_1) while reducing the total path cost (minimizing Z_2); the path cost comprehensively considers a cost function composed of multiple external factors. The second-stage objective functions (15) and (16) are used in the resource scheduling stage, aiming to improve the completion rate of high-priority protection tasks during personnel scheduling (maximizing Z_3) while controlling inspection time and material carrying costs (minimizing Z_4). Constraints (17)~(18) ensure the connectivity of path planning; constraints (19)~(20) ensure task uniqueness and time rationality; constraints (21)~(22) conduct resource protection and matching inspection for material scheduling. Conflicting objectives in the multi-objective model are solved by introducing weight merging or adopting the Non-dominated Sorting Genetic Algorithm II (NSGA-II) to obtain the Pareto optimal solution set and screen scheduling results^[8].

3. Case analysis

3.1. Basic assumptions

The assumptions are as follows:

- (1) The spatial distribution of power grid equipment is mapped to a 2D Cartesian coordinate system with a coordinate range of 80×80km, simulating the coverage of an urban-level distribution network. The emergency resource scheduling center is fixed at the origin (40,40);
- (2) The inspection path between equipment follows the Euclidean distance calculation rule. Considering

urban road traffic efficiency, travel time is converted at an average speed of “1 km/10min”, ignoring additional impacts of extreme weather on road conditions ^[9];

- (3) The inspection and repair time for a single piece of equipment is a fixed value (20 min/unit), regardless of equipment type differences, and maintenance priority differences are only reflected through risk levels;
- (4) The initial state of emergency resources is stable: 3 inspection personnel (p1, p2, p3) are allocated, each with a maximum daily working time of 480 min (8 hours). Material inventory meets the maintenance needs of all equipment without material shortage constraints;
- (5) Parameters of the equipment risk profile are generated through a “truncated normal distribution”: the health index H follows a distribution in the interval $[60,100]$, the real-time failure rate λ follows a distribution in the interval $[0.01,0.1]$, the topology importance score I follows a distribution in the interval $[0.5,1.0]$, and the weight coefficients α , β , γ are all set to 0.33, i.e., balancing the risk contributions of the three dimensions ^[10].

3.2. Experimental parameter configuration

The experiment sets 10 distribution equipment to be inspected (numbered 1~10) and configures basic parameters such as equipment coordinates and initial risk values; the rolling optimization window is set to 3, each with a time span of 120min, simulating the update of scheduling schemes every 2 hours after a disaster. Specific parameter configurations are as follows: For risk assessment parameters, the normalization interval of the health index is ^[0,1], the failure rate fitting coefficients $a=0.05$ and $b=0.8$

are obtained through inversion of 2-year equipment failure data, and topology importance is calculated using Betweenness centrality; for path optimization parameters, the Euclidean distance weight $\alpha=0.5$, the road grade penalty coefficient $\beta=0.3$, all roads are defaulted to “urban arterial roads” with a unified penalty coefficient, and the weather impact factor $\gamma=0.2$; for resource scheduling parameters, the personnel working time constraint $T_{\max}=480\text{min}$, the material consumption coefficient $\mu=1$ (i.e., 1 unit of material is consumed per equipment), and the objective function weights $\omega_1=0.6$ (cost minimization) and $\omega_2=0.4$ (risk coverage maximization); for dynamic scenario parameters, equipment 5 triggers risk escalation in Stage 2 (Window 2), i.e., the failure rate increases from 0.03 to 0.08, and personnel p2’s working time is consumed to 300min with 180min remaining in Stage 3 (Window 3).

3.3. Experimental results and analysis

3.3.1. Basic scenario

Window 1 is the initial stage with no sudden failures and intact resource status. The model output results are shown in **Table 2**. From the scheduling results, the model prioritizes including high-risk equipment (3, 8, 10, $R \geq 0.6$) in the inspection plan while balancing the minimization of path costs. In terms of path planning, the inspection path ^[1,2,3,4,5] is a simplified path of “scheduling center→1→2→3→5→4→scheduling center”. Due to the temporary negative weight problem in the path optimization module, an adjacent equipment clustering path is adopted. The total travel time is 158 min, and the path cost is reduced by approximately 24.8% compared with random paths (210 min); in terms of personnel allocation, high-risk equipment 3 is independently responsible for by p3, taking 20 min for maintenance and 45 min for travel, with a total duration of 65 min. p1 is responsible for low-risk equipment 1 and 4 with a total duration of 78 min, and p2 is responsible for medium-risk equipment 2 and 5 with a total duration of 82 min. All personnel working hours are less than 120 min (window duration), and the resource utilization rate is 25.8%. In terms of risk coverage and cost, the high-risk equipment coverage rate reaches 33.3%.

Among equipment 3, 8, and 10, 3 has been completed, and 8 and 10 are included in subsequent windows with no missing high-risk equipment. The total scheduling cost is -66.39, calculated by weighting the path cost of 158 min and the personnel cost of 225 min. The negative value indicates “risk coverage benefit > cost consumption”, which meets the dual-objective requirements of the model for “cost minimization + risk coverage maximization”^[11].

Table 2. Scheduling results of rolling window 1

Evaluation index	Value	Result analysis
High-risk equipment coverage rate	33.3%	Inspection of high-risk equipment 3 is completed; 8 and 10 are included in subsequent plans with no missing high-risk equipment
Resource utilization rate	25.8%	Personnel working hours are sufficient without overtime, reserving redundant time for subsequent sudden failures
Total scheduling cost	-66.39	The comprehensive cost is lower than the benchmark value (-50), and path costs and personnel costs are reasonably controlled
Scheme adjustment rate	0%	No historical data comparison for the initial scheme, with an adjustment rate of 0

3.3.2. Dynamic risk scenario

Window 2 triggers the risk escalation of equipment 5, i.e., R increases from 0.58 to 0.65, entering the high-risk interval. The model dynamically updates the scheduling scheme through the rolling optimization framework, and the results are shown in **Table 3**. Compared with Window 1, the scheme adjustments are mainly reflected in three aspects: in terms of risk response, the model adjusts equipment 5 from “medium-risk” to “high-risk” and prioritizes arranging p2 to complete the maintenance of the equipment. The original plan was for p2 to be responsible for 2 and 5; after adjustment, 5 is completed first, then 2. The high-risk equipment coverage rate increases to 66.7%, among which 3 and 5 have been completed, and 8 and 10 are pending; in terms of path optimization, the path is adjusted to ^[1,5,3,2,4], and the travel time is reduced to 142 min, with a 9.5% reduction in path cost; in terms of resources and cost, there is no significant change in personnel allocation, but p2’s working hours increase to 102 min, the resource utilization rate increases to 31.3%, and the total scheduling cost remains -66.39. The increase in risk coverage benefit offsets the reduction in path cost, and the comprehensive cost remains optimal. From the output of the Gurobi solver, the constraint matrix of the model in Window 2 has no redundancy, with 35 rows of constraints and 24 columns of variables all valid. The solution time is 0.01s, which is 50% shorter than that in Window 1. The clear priority of high-risk equipment reduces the search space of the objective function, indicating that the model has higher solution efficiency in dynamic risk scenarios^[12].

Table 3. Scheduling results of rolling window 2

Evaluation index	Value	Result analysis
High-risk equipment coverage rate	66.7%	Inspection of newly added high-risk equipment 5 is completed, the coverage rate is improved, and the risk response is timely
Resource utilization rate	31.3%	Personnel working hours increase, resource redundancy decreases, which is in line with the design logic of “dynamically adjusting resource input”
Total scheduling cost	-66.39	The increase in risk coverage benefit offsets the reduction in path cost, the comprehensive cost remains optimal, and the objective function balancing effect is significant
Scheme adjustment rate	40%	Partial adjustments are made to paths and personnel allocation, with a moderate adjustment rate and no excessive fluctuations

3.3.3. Resource constraint scenario

Window 3 triggers the consumption of personnel p2's working hours to 300 min, and the model needs to adjust the allocation scheme under "personnel time constraints". The results are shown in **Table 4**. At this time, the scheme presents significant resource constraint adaptation characteristics: in terms of personnel allocation adjustment, the equipment 2 and 5 originally responsible for p2 are adjusted to be co-responsible for p1 and p3. p1 adds equipment 2 with a total duration of 123 min, p3 adds equipment 5 with a total duration of 117 min, and p2 is only responsible for low-risk equipment 9, consuming 38 min with 142min remaining, effectively avoiding personnel overtime; in terms of path and risk coverage, the inspection path is adjusted to ^[1,2,3,5,9], covering high-risk equipment 3, 5, and 10. The high-risk coverage rate increases to 75%, and the total travel time increases to 172 min, but the path cost is still lower than that of random paths. In terms of resources and cost, the resource utilization rate increases to 42.5% with no personnel overtime, balancing efficiency and compliance. The total scheduling cost becomes -62.15, which increases slightly compared with the previous two windows due to the increase in path cost, but still remains in the optimal interval, with stable cost control ability. The scheme adjustment rate reaches 60%, reflecting the model's strong adaptability to resource constraints ^[13].

Table 4. Scheduling results of rolling window 3

Evaluation index	Value	Result analysis
High-risk equipment coverage rate	75.0%	Inspections of high-risk equipment 3, 5, and 10 are completed, with only equipment 8 pending, and the key equipment response rate is high
Resource utilization rate	42.5%	The resource utilization rate is significantly improved with no personnel overtime, balancing efficiency and compliance
Total scheduling cost	-62.15	The total cost increases slightly due to the increase in path cost, but still remains in the optimal interval, with stable cost control ability
Scheme adjustment rate	60%	The adjustment range is large due to resource constraints, with a reasonable adjustment rate and no scheme disruption

3.4. Experimental conclusions

Through multi-scenario experiments with 3 rolling windows, the effectiveness of the model is verified. The main conclusions are as follows: In terms of risk identification accuracy, the model can accurately identify high-risk equipment through the three-dimensional integration of health index, failure rate, and topology importance. In terms of resource scheduling rationality, under personnel and time constraints, the model can balance efficiency and economy ^[14]. In terms of dynamic adaptability, the model scheme can quickly reconstruct paths and personnel allocation to meet the dynamic scheduling needs after disasters; in terms of engineering practicality, model parameters such as failure rate coefficients can be obtained through inversion of actual operation and maintenance data, with engineering application potential for real-time scheduling. The performance can be further improved only by correcting the negative weight problem of the path optimization module ^[15].

4. Conclusion

Focusing on the core issues of power grid emergency response and resource scheduling under extreme disasters, this paper proposes a multi-modal risk profiling-driven dynamic synergy optimization model. Through theoretical construction and multi-scenario simulation verification, the model parameters can be obtained through inversion of

actual operation and maintenance data, achieving efficient calculation relying on mature solvers. It can be further improved only by optimizing the negative weight problem of the path module, providing a systematic solution for power grid disaster emergency response and having certain practical significance for enhancing power grid resilience. In the future, the risk assessment dimension of multi-hazard types can be further expanded, the path function can be optimized by combining real-time traffic and meteorological data, and more efficient algorithms can be explored to adapt to the needs of large-scale power grids.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Wang Y, Chen C, Wang J, et al., 2016, Research on Resilience of Power Systems Under Natural Disasters: A Review. *IEEE Transactions on Power Systems*, 31(2): 1604–1613.
- [2] Panteli M, Mancarella P, 2015, The Grid: Stronger, Bigger, Smarter? Presenting a Conceptual Framework of Power System Resilience. *IEEE Power and Energy Magazine*, 13(3): 58–66.
- [3] Tang J, Heinimann H, Han K, et al., 2020, Evaluating Resilience in Urban Transportation Systems for Sustainability: A Systems-Based Bayesian Network Model. *Transportation Research Part C: Emerging Technologies*, 2020(121): 102840.
- [4] Chen Q, Yin X, You D, et al., 2009, Review on Blackout Process in China Southern Area Main Power Grid in 2008 Snow Disaster. 2009 IEEE Power & Energy Society General Meeting, 1–8.
- [5] Manandhar B, Cui S, Wang L, et al., 2023, Post-Flood Resilience Assessment of July 2021 Flood in Western Germany and Henan, China. *Land*, 12(3): 625.
- [6] Huang G, Wang J, Chen C, et al., 2017, Integration of Preventive and Emergency Responses for Power Grid Resilience Enhancement. *IEEE Transactions on Power Systems*, 32(6): 4451–4463.
- [7] Shi Q, Liu W, Zeng B, et al., 2022, Enhancing Distribution System Resilience Against Extreme Weather Events: Concept Review, Algorithm Summary, and Future Vision. *International Journal of Electrical Power & Energy Systems*, 2022(138): 107860.
- [8] Asadi Q, Ashoornezhad A, Falaghi H, et al., 2023, Optimal Repair Crew and Mobile Power Source Scheduling for Load Restoration in Distribution Networks. 2023 International Conference on Protection and Automation of Power Systems (IPAPS), 1–6.
- [9] Wu W, Hou H, Zhu S, et al., 2024, An Intelligent Power Grid Emergency Allocation Technology Considering Secondary Disaster and Public Opinion Under Typhoon Disaster. *Applied Energy*, 2024(353): 122038.
- [10] Shakiba F, Azizi S, Zhou M, et al., 2023, Application of Machine Learning Methods in Fault Detection and Classification of Power Transmission Lines: A Survey. *Artificial Intelligence Review*, 56(7): 5799–5836.
- [11] He X, Dong H, Yang W, et al., 2023, Multi-Source Information Fusion Technology and Its Application in Smart Distribution Power System. *Sustainability*, 15(7): 6170.
- [12] Guo M, Yang N, Chen W, 2019, Deep-Learning-Based Fault Classification Using Hilbert–Huang Transform and Convolutional Neural Network in Power Distribution Systems. *IEEE Sensors Journal*, 19(16): 6905–6913.
- [13] Fahim S, Sarker Y, Sarker S, et al., 2020, Self-Attention Convolutional Neural Network with Time Series Imaging Based Feature Extraction for Transmission Line Fault Detection and Classification. *Electric Power Systems Research*,

2020(187): 106437.

- [14] Pan Y, Zhu J, Li X, et al., 2023, Joint Dynamic Scheduling of Mobile Emergency Resources in Distribution Network After Disaster Considering the Influence of Traffic Network. 2023 IEEE International Conference on Energy Internet (ICEI), 367–372.
- [15] Dai H, Liu G, Xin L, et al., 2025, Research on Cooperative Scheduling and Power Restoration Strategy of Intelligent Operation and Maintenance Equipment Under Flood Disaster Based on Dynamic Planning. Journal of Combinatorial Mathematics and Combinatorial Computing, 2025(127): 3051–3072.

Publisher's note

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

Research and Application of Digitalization in Basic Metrological Inspection Institutions

Yajun Zhang, Guangcheng Jia, Peng Wang

Xinjiang Uygur Autonomous Region Institute of Metrology and Testing, Urumqi 830000, Xinjiang, China

**Author to whom correspondence should be addressed.*

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Metrological testing is an indispensable link supporting the high-quality development of various industries, and basic metrological inspection institutions are a key force serving the national economy and people's livelihood, technological innovation, and industrial development. In the digital age, promoting the digital transformation of metrology is an inevitable trend for the development of metrological inspection institutions. The application value of digitalization in metrological inspection institutions is reflected in four aspects: ensuring data quality, optimizing business processes, strengthening risk prevention and control, and innovating service models. Combined with the problems in the informatization construction of metrological inspection, this paper puts forward strategies for the digital transformation and application of basic metrological inspection institutions, focusing on platform construction, system improvement, tool empowerment, facility guarantee, and team building, aiming to provide a reference for the development of metrology towards efficiency and precision.

Keywords: Metrological inspection institutions; Digitalization; Application

Online publication: February 27, 2026

1. Introduction

With the emergence of the digital age, the integrated application and collaborative iteration of new-generation information technologies such as network technology, digital technology, and intelligent technology have injected new momentum into social and economic development. Digital technology has shown strong advantages in the field of metrology. Relying on key technologies such as ultra-dense heterogeneous networks, self-organizing networks, and content delivery networks, 5G communication technology supports remote calibration and automated data collection in metrological testing, realizes intelligent logistics in sample management, and supports remote project review in scientific research^[1]. Big data technology, virtual instrument technology, cloud computing technology, etc., are promoting the upgrading of metrological services, with advantages in optimizing metrological processes, improving data processing efficiency, and strengthening quality control^[2,3]. At present,

basic metrological inspection institutions are facing new challenges such as the application of appointment business processes and the complexity and diversity of testing projects. Traditional business processes and control processes have been difficult to meet the needs of the times ^[4]. Therefore, it is imperative to deepen digital transformation and development, and improve the efficiency and service quality of metrological work.

2. Application value of digitalization in basic metrological inspection institutions

2.1. Ensure the quality of metrological data and lay a foundation for industry credibility

The accuracy and traceability of data are the lifeline of metrological testing, which directly determines the credibility of institutions. Digitalization guarantees data quality through full-chain management and control, specifically reflected in as follows:

- (1) Realize structured collection of key data, convert sample information, testing basis, and judgment standards into a standardized dictionary database, and avoid the randomness and errors of manual entry;
- (2) With the help of data verification technology and automatic verification rules, conduct real-time verification of data types, source legality, and logical consistency to ensure that test data is true and traceable;
- (3) Promote the replacement of paper documents with electronic files, store original records and reports in an unalterable format, meet the requirements of relevant specifications, and strengthen the legal validity of data ^[5].

2.2. Optimize business process efficiency and improve institutional operation effectiveness

Traditional metrological business is limited by offline operations and departmental barriers, with prominent inefficiency problems. Digitalization realizes full-chain collaboration through process reconstruction as outlined:

- (1) Business acceptance link: Through functions such as online appointment and electronic form submission, reduce customers' offline trips and realize one-stop services of "online application, progress query, and certificate download";
- (2) Internal collaboration link: Realize the online flow of sample reception, task assignment, test implementation, and report generation, support mobile on-site entry and cross-department real-time review, and solve the problems of delayed task tracking and high communication costs ^[6];
- (3) Resource allocation link: Provide a scientific basis for equipment scheduling and personnel allocation through automatic data statistics and analysis, improve equipment utilization rate, and reduce operating costs.

2.3. Strengthen quality risk prevention and control to meet compliance management requirements

Metrological testing faces multiple risks such as personnel qualifications and equipment status, and compliance requirements are becoming increasingly strict. Digitalization builds a full-cycle risk prevention and control system, with specific directions as follows:

- (1) Automatic risk early warning: Convert quality control points into digital thresholds, embed them into key nodes of business processes, and automatically remind of expired personnel qualifications, expired equipment calibration, standard updates, etc., to avoid illegal operations ^[7];

- (2) Full traceability of operations: Record operational behaviors through hierarchical permissions, digital signatures, timestamps and other technologies. When quality problems occur, quickly locate the responsible links and personnel to support closed-loop rectification;
- (3) Policy compliance guarantee: Meet the requirements of laws and regulations related to network security and data security, and national cybersecurity grade protection. Prevent data leakage and unauthorized access through data backup, encrypted storage and other measures.

2.4. Empower service model innovation and expand institutional development space

Traditional service models are difficult to adapt to customers' personalized needs and industry collaboration needs. Digitalization provides technical support for service innovation, with main innovation directions as listed:

- (1) Customer service innovation: Establish a customer credit information management system, connect to the national enterprise credit information public disclosure platform, and provide differentiated services according to customers' credit status to achieve precise services and risk prediction^[8];
- (2) Industry collaboration innovation: Support cross-institutional data sharing and credit information interaction, promote the integration of regional metrological services, and realize the mutual recognition of test results and resource complementarity;
- (3) Expansion of service value: With the help of big data analysis technology, tap the industry quality trends in test data, provide decision support for government supervision, and extend the service boundary^[9].

3. Strategies for digital transformation and application of basic metrological inspection institutions

3.1. Build a full-process digital management platform to consolidate the foundation of informatization

In response to the common industry pain points of "data silos" and system fragmentation, and in line with the trend of full-link intelligent management and control, building an integrated platform is the primary task of digital transformation.

3.1.1. Plan the platform function architecture

The platform is designed to support the complete workflow encompassing sample reception, task allocation, test execution, data recording, report generation, and quality supervision. Its functional architecture is composed of the following core modules:

- (1) Business management, enabling online appointment scheduling, real-time progress tracking, and electronic approval;
- (2) Quality management, incorporating risk early-warning mechanisms, compliance verification, and internal audit management;
- (3) Resource management, covering the management of equipment, personnel, and reference materials;
- (4) Customer service, providing online inquiry services, certificate downloading, and complaint handling.

This modular design ensures comprehensive process coverage and operational efficiency.

3.1.2. Ensure platform technical performance

To meet the demands of scalable computing power and data storage, the platform adopts a cloud computing

or hybrid cloud architecture, thereby enhancing system flexibility and access stability. Priority is given to the deployment of secure and controllable domestic operating systems and database technologies to comply with national cybersecurity graded protection requirements. These measures collectively ensure adequate system response speed, high availability, and secure operation ^[10].

3.1.3. Promote system integration and data interoperability

The platform is designed to achieve seamless integration with testing instruments, laboratory information management systems (LIMS), and customer relationship management (CRM) systems. This integration enables automatic acquisition of instrument-generated data, effectively reducing manual input and minimizing the risk of data entry errors ^[11]. In addition, connectivity with government regulatory platforms supports automated reporting of testing data. Participation in industry-level data-sharing platforms further promotes the harmonization of cross-institutional data standards and enhances data interoperability.

3.2. Improve digital management systems and standardize process execution

Facing the pain points of disconnection between systems and digital processes and inconsistent operating standards, combined with the trend of “system-technology” collaboration, improving the standard system is the key to process implementation.

3.2.1. Formulate list-based management specifications

List-based management specifications were formulated around core business processes through the compilation of a “Digital Management Work System” and a “System Operation Guide”. These documents clearly define data collection requirements, operational procedures, and the division of responsibilities at each stage of the workflow. A dynamic update mechanism was established to ensure timely revision of management specifications in response to policy adjustments and evolving business needs. In addition, explicit rules were defined for critical processes, including the management of electronic original records and report generation, to enhance traceability and regulatory compliance.

3.2.2. Strengthen process standardization constraints

Testing workflows were digitized to enable standardized process management, with standard operating procedures (SOPs) embedded directly within the system to ensure consistent execution. The platform enforces compliance by validating sample information, testing standards, and operational steps throughout the workflow. A process compliance verification mechanism was implemented to automatically identify inconsistencies between test items and applicable standards, as well as unapproved critical steps, thereby ensuring strict adherence to prescribed procedures.

3.2.3. Improve the data quality management system

A comprehensive data quality management framework was established across the stages of data collection, processing, storage, and utilization. During data collection, dictionary-based input was applied to key fields, linking manually entered data to a standardized reference dictionary to ensure consistency. During data processing, automated validation techniques, including data type matching and source verification, were employed to reduce human error. For data storage, a dual-layer strategy combining local backup with remote disaster recovery was

adopted, supported by regular data recovery drills. At the data utilization stage, data quality evaluation metrics, such as accuracy, completeness, timeliness, and consistency, were defined, and routine data quality audits were conducted to support continuous improvement ^[12].

3.3. Strengthen Technical Tool Empowerment and Improve the Level of Management Intelligence

Facing the pain points of low efficiency of traditional tools and passive risk prevention and control, and in line with the trend of AI and mobile Internet applications, the upgrading of technical tools is the core path of intelligent transformation.

3.3.1. Optimize process management tools

A mobile application for on-site testing was developed to support real-time data entry, photographic documentation, and electronic signature confirmation, enabling effective collaboration between on-site operations and back-end management systems ^[13]. Furthermore, a task progress tracking module was implemented to allow managers to monitor the status of each workflow stage in real time and to automatically issue alerts for overdue tasks. An intelligent test report generation tool was further developed, which populates standardized report templates using structured data and supports clause-based matching between test results and evaluation criteria, thereby reducing manual compilation errors.

3.3.2. Build an intelligent quality risk management and control system

Key quality risk factors, including personnel qualifications, equipment status, environmental conditions, and updates to testing standards, were transformed into digital control rules and embedded within the operational workflow. A personnel competence matrix database was established to automatically match task assignments with qualified personnel. Equipment management functions include early warning thresholds and automated reminders for calibration and maintenance prior to expiration. Furthermore, an intelligent abnormal data identification mechanism was introduced, employing algorithm-based comparison of real-time data with historical trends and standard reference ranges to flag anomalies and trigger mandatory review procedures.

3.3.3. Promote the full-process application of electronic files

In compliance with relevant regulatory and technical specifications, electronic archiving was implemented for original records, test reports, and quality management documents, using tamper-resistant file formats to ensure data integrity and security. An electronic document management system was established to support classified retrieval, access control, and full audit trails, as well as online viewing and compliant data export. This system facilitates the gradual replacement of paper-based documentation and enhances the efficiency and traceability of document management across the entire testing process.

3.4. Strengthen infrastructure and data security construction to build a guarantee line

Facing the pain points of low digitalization of equipment and prominent data security risks, and in line with the trend of safety compliance and infrastructure upgrading, building a solid guarantee line is a prerequisite for the implementation of digitalization.

3.4.1. Upgrade digital infrastructure

Testing instruments were upgraded to support automated data acquisition, enabling real-time interconnection with centralized information platforms. A stable and secure network environment was established, incorporating firewalls, intrusion detection systems, and other security devices to ensure logical isolation between internal networks and the external Internet. In addition, high-performance servers and scalable storage systems were deployed to accommodate large-volume data storage requirements and to support efficient data access, backup, and recovery processes^[14].

3.4.2. Improve the data security guarantee mechanism

A comprehensive data security management framework was formulated, defining data classification and grading standards and implementing encrypted storage and access control for sensitive information. Digital signatures, timestamps, and related technologies were employed to ensure end-to-end traceability of data processing and to mitigate the risk of unauthorized modification. Furthermore, a data security incident response plan was established, supported by regular cybersecurity drills to enhance emergency handling capabilities. Strict controls were also implemented for data export and external data sharing.

3.4.3. Standardize electronic authentication and permission management

A unified user identity authentication system was established, incorporating multi-factor authentication based on a combination of username, password, and dynamic verification codes, with biometric verification required for critical operations. System access rights were assigned in accordance with the principle of least privilege, clearly defining operational scopes for different user roles, including administrators, inspectors, auditors, and customers. Dynamic permission adjustment and comprehensive audit logging were implemented to enhance accountability. In addition, electronic reports were secured using legally recognized electronic seals and digital signatures to ensure their legal validity, integrity, and non-repudiation.

3.5. Strengthen talent team building and improve digital application capabilities

Facing the pain points of lack of compound talents and insufficient digital skills of employees, and in line with the trend of talent adaptation to digital transformation, team building is the core support for technology implementation.

3.5.1. Establish a compound talent team

A multidisciplinary talent team integrating expertise in metrological testing and digital technologies was established to support system development, maintenance, and continuous optimization^[15]. In parallel, existing personnel were encouraged to participate in digital skills training to enhance competencies in system operation, data analysis, and cybersecurity. To strengthen strategic coordination, a dedicated digital management role was created at the managerial level to oversee the formulation and implementation of digital transformation strategies.

3.5.2. Establish a hierarchical training mechanism

A hierarchical training framework was implemented to address the distinct competency requirements of different roles. Inspectors received targeted training in data acquisition, mobile terminal operation, and abnormal event handling. Managers were trained in process optimization, data analysis, and risk management and control. Technical personnel focused on system maintenance, security protection, and functional upgrades. Training

activities adopted a blended approach combining theoretical instruction, practical exercises, and case-based learning, with contributions from industry experts and system developers.

3.5.3. Improve the incentive and assessment mechanism

Digital competency and application effectiveness were incorporated into employee performance evaluation systems, with assessment indicators including standardized system operation, data quality improvement, and contributions to process optimization. Outstanding performance was recognized through commendation and incentive mechanisms. Employees were further encouraged to engage in digital innovation initiatives, with dedicated rewards for teams or individuals who proposed effective optimization strategies or developed practical digital tools, thereby fostering a culture of broad participation and continuous improvement.

3.6. Promote digital upgrading in phases to ensure implementation effectiveness

Facing the pain points of blind informatization promotion and lack of systematicness, and in line with the trend of gradual digital transformation, phased implementation is a scientific method to ensure effectiveness:

3.6.1. Pilot phase

Departments with high operational volume and well-standardized workflows were selected for pilot implementation. Core functional modules, including sample acceptance, task assignment, data recording, and report generation, were deployed during this phase. Operational issues and optimization recommendations identified during pilot use were systematically collected, and iterative system refinements were conducted to improve functionality and user experience. In parallel, foundational data were organized and standardized, leading to the establishment of a core data dictionary to support subsequent system expansion.

3.6.2. Full promotion phase

Building on the outcomes of the pilot phase, the digital system was progressively extended across the entire institution, enabling full digital migration of all business processes. Comprehensive system operation training was provided to all staff to facilitate adoption and to support the complete transition from paper-based documentation to electronic original records and reports. A system operation monitoring mechanism was established, with dedicated personnel assigned responsibility for routine maintenance and the timely resolution of technical issues and operational challenges.

3.6.3. Optimization and upgrading phase

System operation data were leveraged to identify workflow bottlenecks and management weaknesses, providing an evidence base for ongoing functional optimization. Advanced technologies, including big data analytics and artificial intelligence, were introduced to develop higher-level applications such as data mining, intelligent early warning, and trend analysis. Moreover, integration with industry data-sharing platforms and government regulatory systems was pursued to enhance data utilization and support the development of an “intelligent metrology” management model.

4. Conclusion

In summary, as a core part of the national quality infrastructure, metrological inspection’s management level

is directly related to industrial quality improvement, fair market order, and people's livelihood security. Digital reform is the key engine driving metrological testing institutions to achieve high-quality development. Digital management can not only consolidate the operational foundation of institutions in terms of data quality, process efficiency, and risk prevention and control, but also help institutions build core competitiveness in the wave of digital transformation through service model innovation and technology empowerment. This is not only a practical need to cope with industry competition, but also an inevitable choice to respond to the national quality power strategy. In future practice, institutions need to avoid the misunderstandings of “valuing system construction over process adaptation” and “valuing function development over safety management”. They should always be guided by business needs and take compliance requirements as the bottom line. Through phased promotion and pilot optimization strategies, ensure the effective implementation of digital construction and application.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Guo M, Zhang S, Wang Y, et al., 2021, Research on the Application of 5G Technology in Metrological Informatization. *Metrology & Measurement Technology*, 48(5): 84–86.
- [2] Ren J, Song X, Shang K, 2021, Research on the Application of Informatization Technology Based on Metrological Management. *Electronic Test*, 2021(20): 68–70.
- [3] Zou X, 2021, Exploration and Practice of Digital Transformation of Metrological Testing Institutions. *China Metrology*, 2021(8): 13–14.
- [4] Huang W, Lu L, Cheng F, et al., 2022, Digital Transformation: The “New Infrastructure” for the Development of Metrological Testing Institutions. *Metrology & Measurement Technology*, 49(1): 101–104.
- [5] Li X, Chen H, Deng Q, et al., 2023, Research and Application of Metrological Management Information System. *Metrology & Measurement Technology*, 50(4): 106–109.
- [6] Hu X, Zheng J, Feng Y, 2022, Development and Application of Information Management System for Metrological Testing Laboratories. *China Petroleum and Chemical Standard and Quality*, 42(19): 62–64.
- [7] Jiang L, 2022, Effective Integration of Metrology and Standards to Boost Quality Improvement. *China Metrology*, 2022(6): 32–33.
- [8] Sun R, 2024, Research on Constructing Informatization Methods for Metrological Inspection Institutions. *China Information Times*, 2024(7): 254–256.
- [9] Liu H, 2023, Thoughts on the Digital Transformation of Legal Metrological Inspection Institutions. *China Metrology*, 2023(7): 92–95.
- [10] Lü L, 2021, Development of Intelligent Management System for Metrological Testing Based on Mobile Internet. *Information Recording Materials*, 22(8): 153–155.
- [11] Man M, Pan J, Xi K, 2024, Optimization of Informatization Management Process for Metrological Instrument Verification and Calibration. *Product Reliability Report*, 2024(9): 84–85.
- [12] Wang S, Wang M, 2021, Metrological Testing Data Analysis from the Perspective of Intelligent Management. *Information Recording Materials*, 22(8): 96–98.
- [13] Mi D, Du L, Niu Z, 2023, Analysis and Research on the Informatization Capacity Building of Metrological

Verification Work. China Quality Supervision, 2023(10): 80–81.

- [14] Wang Y, 2021, A Brief Discussion on the Strategies of Applying Informatization in Metrological Management. China Inspection and Testing, 29(3): 58–59.
- [15] Zhang Z, Guo M, Wang Y, et al., 2021, Analysis on the Informatization Development of Metrological Technology Institutions. China Metrology, 2021(12): 45–47.

Publisher's note

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

Research on Unmanned and Intelligent Combat Theory and Capability Development

Yilin Zhao, Jianwei Zhao*, Xuan Liu, Fang He, Fenggan Zhang

Rocket Force University of Engineering, Xi'an 710025, Shaanxi, China

**Author to whom correspondence should be addressed.*

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This paper presents a comprehensive analysis of the evolution, foundational concepts, capability development, and operational challenges of unmanned systems. It traces their theoretical progression from post-Cold War origins to systematic maturation in the 21st century, emphasizing the central role of emerging operational concepts and exploratory advances in capability development. Key bottlenecks in unmanned combat operations are examined, particularly limitations in communication bandwidth and the vulnerability of data links in contested environments. The paper further discusses future development trajectories, highlighting both technological and ethical challenges. Overall, unmanned warfare is evolving toward a more intelligent, networked, and resilient operational architecture, with profound implications for the conduct and character of future high-end warfare.

Keywords: Unmanned combat theory; Capability development; Multi-domain collaboration; Artificial intelligence

Online publication: February 12, 2026

1. Introduction

With the rapid advancement of cutting-edge technologies such as artificial intelligence and information technology, modern warfare is progressively shifting towards intelligent and unmanned operations. As a global pioneer in unmanned combat capability development, the world has initially established an unmanned combat system with space-based assets as the core, air-based assets as the lead, and coordinated development of land and sea-based components. Its theoretical evolution and practical exploration not only lead the development direction of intelligent warfare but also profoundly reshape the operational paradigms of the future battlefield. However, the rapid development of theory and technology has also exposed numerous bottlenecks, including the vulnerability of communication links in contested environments, ethical and legal controversies surrounding autonomous decision-making, and insufficient reliability of AI algorithms.

This paper analyzes the unmanned and intelligent combat theory from four dimensions: theoretical evolution, core concepts, system development, and technological bottlenecks. Employing comprehensive literature research

and case analysis methods, it aims to systematically outline the developmental logic from technological inception to system formation, analyze the connotative evolution of key operational concepts, and reveal the practical challenges faced in cross-domain integration and intelligent transformation.

2. Evolution of unmanned warfare doctrine

2.1. Theoretical inception in the early post-Cold War era

Following the end of the Cold War, amid profound changes in the international security environment and the concentrated, explosive development of high-tech technologies, some powerful nations began systematically exploring the strategic potential and tactical application paths of unmanned combat systems. During this phase, although unmanned platforms had not yet formed large-scale operational capabilities, their role as a key driving force for transforming future warfare patterns attracted widespread attention from senior leadership. Early theoretical concepts primarily focused on how to embed unmanned systems into existing operational architectures to enhance battlefield awareness accuracy and strike response speed, initially reflecting the basic logic of information dominance theory, achieving real-time battlefield situational reconstruction and dynamic sharing through distributed sensor networks. Simultaneously, the evolution of systems confrontation theory provided structural support for unmanned warfare, emphasizing enhanced resilience and survivability of the overall operational system through coordinated interaction of multi-domain heterogeneous units^[1]. Against this backdrop, some countries gradually initiated a series of technology pre-research projects and, relying on top-level design documents like the “Unmanned Systems Integrated Roadmap”, explicitly positioned unmanned platforms as key enablers in joint operations.

2.2. Theoretical system maturation in the 21st century

Since entering the 21st century, unmanned combat theory, driven by both the prolonged practice of counterterrorism wars and the anticipation of high-end conflicts, Some countries have gradually achieved systematic and architectural development. The mature application of unmanned aerial vehicles (UAVs) in missions such as intelligence, surveillance, and reconnaissance (ISR), precision strike, and communications relay served as the critical practical foundation for theoretical maturation. Long-endurance, high-precision platforms represented by the RQ-4 Global Hawk and MQ-9 Reaper not only significantly enhanced wide-area battlefield situational awareness but also enabled cross-domain information fusion through real-time data link support, providing technical and tactical validation for multi-domain joint operations. In 2014, the core guiding principle of the “Third Offset Strategy proposed” was proposed, aimed at maintaining the global military superiority through the development of disruptive technologies. It emphasized enhancing the operational effectiveness of unmanned platforms through R&D and integration of emerging technologies like AI, autonomous systems, and directed-energy weapons, with particular focus on building a multi-domain integrated system centered on space-based perception, led by air-based unmanned platforms, and supported by coordinated responses from land and sea unmanned nodes. This laid the strategic foundation for the subsequent deepening of unmanned combat theory. These theoretical explorations and technological practices collectively contribute to the comprehensive maturation of the unmanned combat theoretical system.

3. Analysis of core operational concepts

3.1. Analysis of the “Mosaic Warfare” concept

“Mosaic Warfare,” as an emerging operational paradigm, originates from “Decision-Centric Warfare” theory. Within this architecture, the combat system no longer relies on a few high-value platforms but enhances system survivability and resilience through distributed deployment. Even if some nodes are lost, remaining units can reorganize through self-organizing mechanisms and maintain mission execution capability. This decentralized structural characteristic significantly enhances battlefield adaptability, particularly suited for sustained operational requirements in complex contested environments ^[2]. “Mosaic Warfare” not only changes traditional force composition models but also drives the evolution of command and control systems towards deep human-machine integration, requiring breakthroughs in key technological areas such as situational awareness, communications assurance, and dynamic task allocation. Therefore, the practice of “Mosaic Warfare” is not merely an innovation at the tactical level but a systematic challenge and reshaping of the entire unmanned combat system architecture and its supporting technological capabilities.

3.2. Integration of “Joint All-Domain Command and Control”

“Joint All-Domain Command and Control” (JADC2) is the core architecture through which aims to achieve future intelligentized operations. It seeks to break down traditional information silos between military services via highly integrated communication networks and data link systems, enabling deep integration of multi-domain operational forces. Within this system, unmanned platforms, leveraging their distributed deployment, high mobility, and persistent reconnaissance capabilities, become key nodes for information sensing and tactical execution, supporting the underlying architecture for cross-domain collaborative operations. Notably, the “Mosaic Warfare” concept further drives the evolution of JADC2 towards modularity and reconfigurability. By decomposing the kill chain into flexibly combinable functional units, it achieves rapid response and adaptive reorganization, marking the transition of unmanned combat from single-platform control towards a new stage of system-level intelligent contest ^[1].

4. Integration of unmanned combat systems

4.1. Integration of unmanned systems into the existing command and control architecture

In advancing its unmanned combat capabilities, the places significant emphasis on the deep integration of unmanned platforms with existing command systems, striving to achieve efficient embedding and process optimization within traditional command chains. This mechanism relies on the synergy between high-bandwidth, low-latency data links and artificial intelligence algorithms, ensuring seamless flow of command information between manned platforms and unmanned units, thereby enhancing overall operational tempo and response speed. Within this command architecture, the manned-unmanned collaborative combat model serves as a bridge connecting traditional command structures with emerging autonomous systems. Human operators act as mission supervisors and strategic decision-makers, responsible for setting rules of engagement and target priorities, while unmanned systems undertake tactical execution tasks such as reconnaissance, strike, and communications relay. This division of labor not only aligns with the requirements for cross-domain coordination under Joint All-Domain Operations (JADO) but also reflects the core value of human-machine teaming (HMT) in modern warfare ^[3]. To further quantify the stability and adaptability of collaborative structures, the complex network theory has been introduced to construct combat network models for manned/unmanned formations. Through static topological

analysis, these models reveal the connectivity characteristics and information dissemination efficiency of systems under different autonomy levels ^[4]. This provides a theoretical tool for assessing the robustness of command links and offers technical support for the future evolution of command and control systems toward distributed and resilient architectures. Thus, through institutional, technological, and theoretical pathways, the unmanned platforms are being promoted for deep integration into existing operational command systems, thereby realizing intelligent and agile decision-making processes.

4.2. Construction of the intelligence-strike closed loop

Within the modern joint operational system, some countries are committed to shortening the decision-making cycle and enhancing strike timeliness. Its core lies in the automation and intelligentization upgrade of the kill chain. In recent years, with the rapid development of the Manned-Unmanned Teaming (MUM-T) concept, a few countries has preliminarily achieved dynamic task allocation and information sharing among heterogeneous operational units, significantly enhancing overall battlefield situational awareness and operational effectiveness ^[5]. The success of this operation depended on real-time target data provided by external sensor nodes, fully demonstrating the capability for multi-source intelligence fusion and cross-platform coordinated strike within a distributed operational network ^[6]. To quantitatively assess the effects of such coordinated combat, an extended Lanchester combat model has been proposed within academia, introducing battlefield awareness coefficients and command and control capacity parameters. Simulation results indicate that strengthening the quality of information interaction and system integration level in coordinated combat can reduce operational losses while improving mission success rates.

5. Bottlenecks in system operation

5.1. Communication bandwidth and link security

In long-range unmanned combat missions, limited communication bandwidth and insufficient data link security have become key bottlenecks restricting the effectiveness of unmanned systems. As the operational radius expands, the demand for high-bandwidth, low-latency data transmission by remote unmanned platforms increases sharply, while existing communication architectures struggle to guarantee stable information exchange in complex electromagnetic environments. Particularly in high-intensity conflict scenarios, electronic jamming and spectrum suppression conducted by adversaries significantly degrade link availability. Although traditional frequency-hopping communication possesses certain anti-jamming capabilities, its pre-set hopping patterns are susceptible to detection and prediction, making real-time synchronization of communication parameters difficult and severely impacting the continuity of command and control in MUM-T operations ^[7]. To address this, researchers are exploring variable-rate non-cooperative frequency-hopping techniques, enhancing parameter stealth by dynamically adjusting hopping sequences to improve communication robustness under contested conditions ^[7]. Concurrently, artificial noise cooperative jamming technology is being introduced into electromagnetic warfare systems, using directional jamming means to degrade enemy channel quality without affecting friendly signal reception, further strengthening communication superiority ^[7]. From a network topology perspective, MUM-T combat networks need high resilience to cope with node failure or link interruption. Evaluation models based on connectivity robustness indicate that by constructing a node and edge reconstruction evaluation index system, rapid recovery can be achieved after partial network damage, enhancing the overall system's fault tolerance and

mission sustainability ^[8]. However, even with ongoing technological evolution, cybersecurity threats persist. Network intrusions could lead to command tampering or sensitive intelligence leakage, severely undermining the trustworthiness and stability of the unmanned combat system.

5.2. Ethics and risks of autonomous decision-making

As the autonomy levels of unmanned combat systems continuously increase, the applicability of engagement rules and associated ethical risks during mission execution in complex battlefield environments are becoming increasingly prominent. Highly autonomous unmanned platforms, operating without continuous human intervention, may perform target identification and strike decisions based on pre-set algorithms. When a system causes unintended casualties due to environmental misjudgment or algorithmic bias, it is difficult to clearly define the responsible entity, whether the operational commander, system developer, or the AI itself should bear the consequences, as there is currently no unified international norm for this. This ambiguity not only weakens the legitimacy foundation of military operations but may also undermine public trust in the unmanned combat model. Concurrently, the cognitive limitations of autonomous decision-making systems are further exposed in dynamic conflict scenarios, especially in critical links such as friend-or-foe identification, civilian avoidance, and tactical intent inference. Existing algorithms remain constrained by insufficient situational understanding capability and information fusion accuracy, prone to misjudgment and overreaction. Although multi-agent coordinated strategies are being optimized through reinforcement learning frameworks to enhance system adaptability and behavioral stability. For instance, the application of the PD-MADDPG algorithm in continuous dynamic air combat environments has significantly improved strategy convergence efficiency and execution robustness ^[9]. Therefore, while advancing high-level autonomous capability development, it is imperative to simultaneously construct a composite governance framework encompassing technology verification, legal review, and ethical assessment, ensuring that unmanned systems achieve a dynamic balance between operational effectiveness and moral acceptability under the premise of adhering to the laws of war.

5.3. Reliability of artificial intelligence algorithms

In the evolution of the unmanned combat system, the reliability of AI algorithms has become a key factor determining the effectiveness of autonomous coordinated operations. Particularly in complex, dynamic, and highly uncertain battlefield environments, the stability of target recognition and behavior prediction by unmanned systems directly relates to mission success and operational safety. Currently, although deep learning models, through multi-layer neural networks, have achieved effective fusion and feature extraction of multi-source sensor information (space, air, sea-based), enhancing situational awareness accuracy, decision bias risks still exist in actual conflict scenarios. Improved algorithms based on the Multi-Agent Deep Deterministic Policy Gradient (MADDPG) with centralized training and decentralized execution have demonstrated stronger strategy convergence and execution stability in simulated air combat tasks ^[10]. Another example is the introduction of parallel decoupling mechanisms and symmetric attention structures (SAM) to optimize the information screening efficiency of critic networks, enabling unmanned swarms to possess superior coordinated response capabilities in continuous dynamic competition ^[9]. Their reliability verification requires closed-loop iteration relying on large-scale simulation and live testing. Therefore, constructing intelligent algorithm architectures that combine safety and adaptability is a core research direction for future unmanned combat capability development.

5.4. Autonomous navigation and obstacle avoidance technology

In complex electromagnetic environments and GPS-denied conditions, the autonomous navigation and obstacle avoidance capabilities of unmanned combat platforms face severe challenges, becoming a key technological bottleneck constraining all-domain operational effectiveness. Traditional navigation modes reliant on the Global Positioning System are highly susceptible to jamming and spoofing in contested environments, leading to platform positioning failure, mission interruption, or even tactical exposure. To address this, multi-source fused navigation technology has long been a focus of development, integrating various non-GNSS methods such as inertial navigation, terrain contour matching, visual simultaneous localization and mapping (SLAM), and geomagnetic-aided navigation to enhance the sustained positioning accuracy and robustness of unmanned systems in denied environments^[8]. However, existing systems still exhibit significant shortcomings in dynamic environment adaptability, long-term drift suppression, and multi-node coordinated positioning consistency. Particularly in extreme scenarios like underwater or urban canyons, limited communication bandwidth and perception occlusion further exacerbate navigation solution uncertainty. Therefore, the future development of navigation systems for unmanned platforms depends not only on technological breakthroughs at the sensor level but also requires strengthening cross-domain coordination, intelligent reconfiguration, and anti-jam communication capabilities from the perspective of the operational system architecture, achieving a fundamental shift from platform autonomy to system empowerment.

6. Assessment of future development directions

6.1. Evolution towards intelligentization and swarm operations

With the accelerated advancement of global intelligent unmanned combat systems, swarm operation modes represented by small UAV swarms are gradually becoming a key force on the future battlefield. Such swarm systems not only create asymmetric deterrence through numerical superiority but also, under the Mosaic Warfare concept, build flexible and reconfigurable kill webs through modular node combinations, breaking traditional linear operational structures. Simultaneously, the “Combat Cloud” concept in the context of cross-domain coordination further expands the information support dimension for swarm operations. By integrating space-based sensing, airborne early warning, and ground command nodes, it achieves all domain situational sharing and dynamic target guidance. The application of low-cost, attritable platforms like the XQ-58A Valkyrie provides a tactical interface for swarms within high-end manned-unmanned formations, driving the operational system towards intelligent and resilient evolution. Despite facing technical and strategic challenges such as communication latency, coordinated robustness, and ethical regulation, unmanned swarm operations based on autonomous coordination and intelligent emergence are still regarded as a core driving force subvert future warfare patterns^[11].

6.2. Deep integration and development of human-machine teaming

With the accelerated evolution of modern warfare towards intelligent and networked forms, major military powers are actively promoting the deep integration of manned platforms and unmanned loyal wingmen at the tactical level, focusing on constructing a new operational system centered on human-machine teaming. This mode relies on advanced command and control architectures and autonomous decision-making algorithms to achieve efficient coupling between human operators and unmanned systems. Distributed control methods based on consensus protocols are introduced into manned-unmanned formation systems, further improving formation coordination efficiency. Information interaction topologies constructed by combining graph theory and leader-follower

mechanisms ensure the formation maintains a stable configuration in complex electromagnetic environments. Modeling coordinated combat network structures under different autonomy levels using complex network theory enables quantitative analysis of system robustness and dynamic topological evolution^[4,12]. This integrated application of technologies and theories marks a new stage in the evolution of highly intelligent human-machine combat collaboration.

7. Conclusion

The development of unmanned combat theory and capabilities has evolved from its theoretical inception in the immediate post-Cold War era to its systematic development in the 21st century. Driven by the demands of counterterrorism warfare, the widespread operational use of UAVs prompted it to gradually construct an unmanned combat theoretical system spanning land, sea, air, space, and cyberspace domains. Novel operational concepts like “Mosaic Warfare” emphasized distributed architectures and dynamic reorganization capabilities, enhancing the flexibility and survivability of combat systems. The JADC2 concept aimed to break down information barriers between service branches, enabling cross-domain collaboration between unmanned platforms and real-time data sharing. However, challenges persist during this integration process, including delays in command and control response, insufficient automation levels in the intelligence-strike loop, limited communication bandwidth, and ethical risks associated with autonomous decision-making. Regarding key technologies, the stability of artificial intelligence algorithms and autonomous navigation capabilities in GPS-denied environments remain bottlenecks. Future development trends will focus on intelligent swarm operations, deep human-machine collaboration, and enhancing the adaptability of highly autonomous systems in complex battlefield environments, thereby further shaping the character of high-end warfare.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Wang J, 2021, Research on Mosaic Cooperative Warfare Theory. *Telecommunications Technology Research*, 2021(4): 33–39.
- [2] Andrews J, 2020, Human Performance Modeling: Analysis of the Effects of Manned-Unmanned Teaming on Pilot Workload and Mission Performance, thesis, Air Force Institute of Technology.
- [3] Johnson B, Miller S, Heeter B, et al., 2023, A Human-Machine Teaming Approach for Future Marine Corps Hybrid Operations: Manned Helicopters and Air-Launched Unmanned Aerial Vehicles. *Naval Engineers Journal*, 135(1): 127–139.
- [4] Wang X, Cao Y, Liao W, 2021, Modeling of MAV/UAV Collaborative Combat Network Based on UAV Autonomous Level. *Proceedings of 2021 International Conference on Autonomous Unmanned Systems: International Conference on Autonomous Unmanned Systems (ICAUS 2021)*, 24–26.
- [5] Fan J, Li D, Li R, et al., 2017, Analysis for Cooperative Combat System of Manned-Unmanned Aerial Vehicles and Combat Simulation. *2017 IEEE International Conference on Unmanned Systems*.
- [6] Chabanov V, 2022, Technology of Interaction Between Manned and Unmanned Aerial Vehicles (MUM-T): An

Element of Multi-Domain Combat Operations of the US Ground Forces. *Aviation Systems: Scientific and Technical Information*, 2022(3): 44–47.

- [7] Gao L, Ding J, Lian R, 2019, The Communication Availability of the Cooperative Combat of Manned / Unmanned Aerial Vehicles in Electromagnetic Cyber Confrontation Environment. *International Conferences on Ubiquitous Computing Communications; Data Science and Computational Intelligence; Smart Computing, Networking and Services*.
- [8] Shi G, Zhang L, Zhang J, et al., 2018, Research on Robustness of Manned/Unmanned Aerial Vehicle Collaborative Combat Network. *International Conference on Control, Automation, Robotics and Vision*.
- [9] Wang Z, Guo Y, Li N, et al., 2023, Autonomous Collaborative Combat Strategy of Unmanned System Group in Continuous Dynamic Environment Based on PD-MADDPG. *Computer Communications*, 200(2): 182–204.
- [10] Xu D, Chen G, 2022, The Research on Intelligent Cooperative Combat of UAV Cluster with Multi-Agent Reinforcement Learning. *Aerospace Systems*, 5(1): 107–121.
- [11] Qi D, Zhang J, Liang X, et al., 2021, Autonomous Reconnaissance and Attack Test of UAV Swarm Based on Mosaic Warfare Thought. *2021 6th International Conference on Robotics and Automation Engineering: 6th International Conference on Robotics and Automation Engineering (ICRAE)*, 19–22.
- [12] Ma N, Cao Y, 2024, Consensus-Based Distributed Formation Control for Coordinated Battle System of Manned/Unmanned Aerial Vehicles. *Transactions of the Institute of Measurement and Control*, 46(1): 3–14.

Publisher's note

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

Application Research of Concept Bottleneck Model in Passport Printing Method Detection

Tianrui Qiu, Jiafeng Xu*

China People's Police University (Guangzhou), Guangzhou 510663, Guangdong, China

*Corresponding author: Jiafeng Xu, xujiafeng@cpperu.edu.cn

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited

Abstract: With the increase in cross-border mobility, passports, as critical identity documents, require robust anti-counterfeiting security. While existing deep learning-based automatic detection methods achieve high accuracy, they lack interpretability. This paper introduces the Concept Bottleneck Model (CBM) to construct a transparent passport printing method detection framework. By defining interpretable intermediate concepts and integrating linear reasoning, the model significantly enhances reliability and debugging efficiency. The article systematically analyzes the advantages, challenges, and future directions of this approach.

Keywords: Passport anti-counterfeiting; Concept bottleneck model; Explainable AI; Printing method detection; Human-machine collaboration

Online publication: February 12, 2026

1. Introduction

Against the increasingly complex backdrop of globalization and security dynamics, passports, as the core credentials for national sovereignty and exit and entry administration, have their security directly linked to the stability of national security and social order. With the advancement of counterfeiting technologies, especially the growing number of cases involving the production of highly simulated passports using advanced digital printing and microprinting techniques, how to distinguish the authenticity of passports by detecting their printing methods (e.g., differentiating between optical printing, laser engraving and thermal printing) has become a key technical challenge in the fields of border security inspection and forensic authentication^[1].

To address the above bottlenecks, this paper proposes an innovative approach of introducing Concept Bottleneck Models (CBMs) to construct a passport printing method detection framework that not only maintains high detection accuracy but also features interpretability and traceability. The main contributions of this paper are as follows: it systematically conducts a qualitative analysis of the adaptability of CBMs in the specific scenario of passport detection, deeply explores their application value in real security inspection environments, and analyzes

the potential challenges in terms of model migration, concept annotation costs and cross-device universality.

2. Literature review of related work

2.1. Overview of passport printing processes and anti-counterfeiting features

As the highest-level identity documents issued by countries, passports adopt far more complex printing processes than ordinary paper documents. The current mainstream passport printing processes mainly include laser printing, inkjet printing and intaglio printing, each with unique physical characteristics and anti-counterfeiting mechanisms. Laser printing transfers toner to paper surfaces using photothermal technology, typically producing features such as high-contrast images and a specific fine granular texture; inkjet printing forms patterns through the deposition of tiny ink droplets, characterized by natural color gradation and weak granularity; intaglio printing, by contrast, imprints ink onto paper via recessed relief plates, creating patterns with distinct three-dimensional tactile feel and glossiness. A thorough understanding of the core differences between these processes forms the foundation for constructing an effective detection model.

2.2. Brief analysis of existing detection technologies

Current passport anti-counterfeiting detection technologies are mainly divided into traditional physical feature-based methods and deep learning-based automatic detection methods. Traditional methods rely on expert experience and specialized equipment to judge authenticity by analyzing the spectral reflectance of paper, watermark structures or intaglio tactile feel, yet they are limited by human factors and difficult to realize large-scale automation. While deep learning models perform excellently in the field of image recognition and have significantly improved detection speed and accuracy, their excessive abstraction capability for intermediate features results in a lack of interpretability, making it difficult to meet the stringent requirements for “trustworthiness” in safety-critical fields such as passport anti-counterfeiting ^[2].

2.3. Necessity of explainable artificial intelligence in safety-critical fields

In safety-critical fields such as passport detection, the application of technology is not only about pursuing high accuracy but, more crucially, ensuring the reliability and traceability of results. Traditional deep learning models often function as “black boxes”: even if a model yields an accurate result, the lack of explanation means that when the model makes a misjudgment or is attacked by adversarial examples, it is impossible to quickly locate the root cause of the problem. In passport authentication, examiners need to understand the specific basis for the model’s determination of “inkjet printing” or “laser printing” to conduct secondary verification. Therefore, the introduction of Explainable Artificial Intelligence (XAI) is an urgent priority, as it can provide the logical chain of model decision-making, thereby enhancing examiners’ trust in the system and improving the security of the overall anti-counterfeiting system.

2.4. Core ideas and development context of concept bottleneck models: From concept learning to explainable reasoning

Concept Bottleneck Models (CBMs) are a type of deep learning architecture that integrates both interpretability and predictive ability, whose core idea is to design the model’s hidden layers as a set of explicit human-interpretable “concepts” ^[3]. Specifically, a CBM first maps an input image to a set of predefined concept vectors (e.g., “paper glossiness”, “ink granularity”), which directly correspond to the key features of passport printing

processes. The model then makes the final prediction based on these concept vectors. This structure achieves a leap from “pixel-level features” to “concept-level features”, rendering the model’s intermediate decision-making process fully transparent. The proposal of CBMs marks the evolution of artificial intelligence research from “pure concept learning” to “concept-based explainable reasoning”, providing a theoretical foundation for solving the interpretability bottleneck of deep learning models in passport detection.

3. Construction of a framework for applying concept bottleneck models to passport printing detection

3.1. Qualitative description of the overall architecture

The detection framework proposed in this section technically follows the standard workflow of Concept Bottleneck Models (CBMs) to realize transparent judgment of passport printing methods. Specifically, the system first receives high-resolution scanned or captured images of passports (input images), which contain the original physical information related to anti-counterfeiting features. The images then enter the concept extraction layer, where the model does not directly output category labels but extracts a set of intermediate features corresponding to human anti-counterfeiting experience, these features are referred to as “concepts”. These concepts are subsequently aggregated and regularized through the concept bottleneck layer to form a set of structured concept vectors. Finally, the task prediction layer performs simple logical reasoning based on these concept vectors and outputs the final classification result of the printing method. Through the “bottleneck”, this architecture maps complex image features to interpretable concepts, achieving a qualitative leap from a “black box” to a “transparent box”.

3.2. Definition and design of core concepts

In passport anti-counterfeiting, concept design is crucial and must balance machine recognition with human examiners’ experience. Targeting different printing processes, we define four core concepts:

- (1) “Ink diffusion morphology” (capturing inkjet’s natural diffusion edges, contrasting laser printing’s sharp toner-fused edges);
- (2) “Dot matrix structure” (a key feature distinguishing inkjet spray dots from laser powder deposition dots);
- (3) “Gloss texture pattern” (reflecting the metallic luster of intaglio-printed paper vs. ordinary paper’s matte texture under light);
- (4) “Character edge sharpness” (quantifying character contour clarity and attenuation across printing methods). Derived from anti-counterfeiting experts’ long-term empirical observations, these concepts ensure professional depth in the model’s interpretable outputs.

3.3. Qualitative approach to concept extraction

To extract the above-defined concepts efficiently and accurately, the model adopts a multi-source fusion technical approach in the concept extraction layer. Specifically, a deep convolutional neural network (e.g., ResNet) pre-trained on large-scale image datasets is first used to extract general visual features, a step that ensures the model has strong feature perception capability ^[4]. Subsequently, in response to the specific needs of passport anti-counterfeiting, small domain knowledge-driven network branches are designed, which are fine-tuned specifically for microscopic features such as “glossiness” or “ink diffusion” to capture delicate physical differences. In addition, we introduce traditional feature engineering methods, such as Gabor filters or gray-level co-occurrence

matrices, to assist in extracting “dot matrix structure” or “texture pattern”. Through the synergy of these approaches, the system can map complex image pixels to high-dimensional concept attribute vectors, laying a solid data foundation for subsequent reasoning.

3.4. Qualitative analysis of concept-task reasoning

Following feature abstraction in the concept bottleneck layer, the system proceeds to decision-making. The task prediction layer employs simple interpretable algorithms like linear or logistic regression, analyzing the weight contribution of each dimension in the concept vector: for instance, high “ink diffusion morphology” and low “gloss texture pattern” scores indicate inkjet printing, while extremely high “character edge sharpness” paired with prominent “gloss texture pattern” suggests laser or intaglio printing. The reasoning model’s simplicity enables direct access to each concept’s contribution to the final decision, ensuring transparent decision-making. This enhances model trustworthiness, allows examiners to conduct secondary verification via concept explanations, and significantly reduces misjudgment risks.

4. Qualitative analysis of application advantages and value

4.1. Improved interpretability and trustworthiness

The greatest advantage of CBMs lies in their intrinsic “interpretable” attribute. Although traditional deep learning models achieve high accuracy, their decision-making process is a completely opaque black box for humans, which is unacceptable in the authentication of safety-critical documents such as passports ^[5]. In contrast, CBMs map complex pixel-level features directly to specific physical features by imposing constraints that the model’s hidden layers must output concepts consistent with human cognition. This mapping process can be directly observed and reviewed by domain experts, making each judgment of the model based on verifiable evidence, greatly improving the trustworthiness of the entire detection system and avoiding legal and security risks caused by the “black box” nature.

4.2. Human-machine collaboration and knowledge fusion

CBMs are naturally designed to support “human-machine collaboration”. Since the model outputs human-interpretable concepts, this provides an intuitive intervention point for domain experts. If an expert finds that the model identifies certain specific concepts inaccurately, they can directly intervene and correct the concept without retraining the entire deep network. This interactive iterative optimization process not only can rapidly improve the performance of the model but also promotes the in-depth integration of machine learning algorithms with traditional anti-counterfeiting expertise, enabling the system to evolve continuously with the accumulation of expert experience.

4.3. Potential in data efficiency and generalization ability

By learning the intermediate concept layer, CBMs can alleviate the reliance on large-scale labeled data to a certain extent. Unlike traditional end-to-end learning that requires a large number of labeled samples to capture complex high-dimensional features, CBMs can use the definition of concepts by domain experts to guide the learning process ^[6]. This means that even when facing new counterfeiting methods or variants, the model can identify them well as long as these variants exhibit obvious anomalies in the concept space, thus demonstrating good generalization ability. This has important practical value for scenarios such as passport anti-counterfeiting

detection that need to respond to constantly upgraded counterfeiting technologies.

4.4. Debugging and error diagnosis

CBMs have significant advantages in troubleshooting. When a traditional deep network malfunctions, engineers often struggle to locate the root cause of the error due to the abstraction and high dimensionality of internal features, resulting in an extremely complex debugging process. On the contrary, the decision-making of CBMs is based on a series of explicit concepts. If the system makes an identification error, engineers can directly check which concept (e.g., dot matrix structure or ink diffusion morphology) is misidentified, leading to the final misjudgment. This concept-based error localization mechanism not only accelerates the maintenance and iteration of the system but also significantly reduces the technical costs of long-term operation.

5. Challenges and future outlook

5.1. Existing challenges

Although Concept Bottleneck Models (CBMs) enhance deep learning interpretability via a human-interpretable concept layer, their application for passport anti-counterfeiting faces two key challenges. First, defining the concept system is highly complex: key anti-counterfeiting features of passport backgrounds correlate with continuous physical variables (e.g., glossiness, ink diffusion morphology) that vary significantly with illumination. This creates a dilemma: overly broad concepts reduce discriminability, while overly strict definitions drive up annotation costs, cause concept overlap and interpretation conflicts, making it hard to develop a concept set that covers key printing process features while enabling stable separation in images. Second, concept extraction reliability directly determines interpretability validity. Real-world passports often have non-ideal factors (wear, stains, uneven illumination); misjudgments by concept extractors in complex backgrounds trigger cascading errors that distort or collapse interpretive results, necessitating improved robustness and consistency of extractors against noise, distortion and background interference.

5.2. Future outlook

Future passport detection systems will evolve from closed black boxes into interactive, continuously learning platforms. First, they will adopt adaptive concept learning—automatically expanding or adjusting concept libraries for emerging counterfeiting methods, replacing rigid manual definitions. Second, CBMs will integrate with other XAI technologies (e.g., real-time visualization of “glossiness” response regions in images, adversarial example analysis to enhance model security). Ultimately, this technology will expand beyond passport detection to anti-counterfeiting for ID cards, driver’s licenses, certificates and bills, building a universal, interpretable AI ecosystem for document authentication.

6. Conclusion

The introduction of CBMs drives the transformation of passport printing method detection from “black box” judgment to “white box” reasoning. Compared with traditional deep learning, CBMs map pixel features to interpretable physical attributes (e.g., ink diffusion, gloss texture) through the concept layer, making the judgment basis traceable and verifiable, improving system credibility, and providing support for audit, compliance and error correction. However, its implementation cannot rely solely on algorithm optimization; it still requires in-depth

collaboration between computer vision and document anti-counterfeiting experts to jointly define key concepts and conduct continuous calibration in iteration. Only in this way can a document authentication system with both intelligence and security be constructed.

Funding

2025–2026 China People’s Police University Student Science and Technology Innovation Program Project

Disclosure statement

The authors declare no conflict of interest.

References

[1] Weeraratna T, 2024, Beyond Borders: The Art and Science of Detecting Travel Document Forgeries. *International Journal of Forensic Sciences*, 9(4): 1–4.

[2] Mohit M, 2016, The Evolution of Deep Learning: A Performance Analysis of CNNs in Image Recognition. *International Journal of Advance Research in Education and Technology*, 3(6): 2029–2038.

[3] Stropeni A, Enhancing Interpretability in Visual Anomaly Detection Through Concept Bottleneck Models, thesis, University of Padua.

[4] Dhillon A, Verma G, 2020, Convolutional Neural Network: A Review of Models, Methodologies and Applications to Object Detection. *Progress in Artificial Intelligence*, 9(2): 85–112.

[5] Shafik W, 2026, The “Black Box” Problem: Lack of Transparency in AI Decision-Making. Springer: 167–186.

[6] Srivastava D, Yan G, Weng L, et al., 2024, VLG-CBM: Training Concept Bottleneck Models with Vision–Language Guidance. *Advances in Neural Information Processing Systems*, 37: 79057–79094.

Publisher’s note

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

Design and Simulation of a Microstrip Frequency-Scanning Antenna for Millimetre-Wave Fuze Applications

Qinyi Wang*

The University of New South Wales, Sydney 2052, Australia

**Author to whom correspondence should be addressed.*

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: To satisfy the simultaneous requirements of high gain and wide angular coverage for millimetre-wave fuzes under large impact-angle variations, this paper proposes a microstrip frequency-scanning antenna based on a quasi-travelling-wave, series-fed patch array. The antenna is implemented on Rogers 4350B substrate ($\epsilon_r = 3.5$, thickness $h = 0.254$ mm) and operates over 30–36 GHz. By exploiting the frequency-dependent phase progression along the series feed, the main beam steers continuously without phase shifters. Full-wave simulations in HFSS show that the antenna maintains $|S_{11}| < -10$ dB across the entire band. The E-plane main beam scans from 48° at 30 GHz to 0° at 36 GHz, providing a 48° frequency-scanning range; when the 3-dB beamwidth is included, the effective detection-angle coverage reaches approximately 74° . The simulated gain remains stable above 10.5 dBi, peaking at about 11.6 dBi near 35 GHz. With a low-profile, planar structure (overall size ≈ 20 mm \times 10 mm) and no additional terminal load, the proposed design offers a compact solution for fuze antennas that require broad angular coverage and robust gain in the 30–36GHz.

Keywords: Millimetre-wave; Microstrip patch; Frequency-scanning antenna; Quasi-travelling wave; Series-fed microstrip array; HFSS

Online publication: February 12, 2026

1. Introduction

Millimetre-wave fuze sensing systems are attractive due to their compact size and high range resolution ^[1]. In practice, the fuze antenna must preserve effective target illumination and echo reception under significant variations in projectile attitude and impact angle ^[2]. This results in a challenging set of concurrent requirements: compact form factor, high gain, and wide angular coverage. Conventional waveguide slot arrays can provide high efficiency but are relatively bulky and heavy, where their scan capability is also limited, which complicates conformal integration on projectiles. In contrast, planar microstrip antennas are low-cost, lightweight, and suitable for conformal mounting, making them a promising candidate for fuze applications.

Existing fuze-antenna studies generally follow two routes. The first uses multi-beam or multi-port switching to obtain several discrete pointing directions^[3]. The second relies on array phase control or frequency scanning to achieve continuous (or quasi-continuous) beam steering. While miniaturised multi-beam designs can realise multiple fixed tilt angles, they typically provide limited angular continuity and may introduce extra feed complexity. Frequency-scanning arrays, on the other hand, enable beam steering through intrinsic dispersion, thereby reducing system complexity. Motivated by Ka-band fuze scenarios (30–36 GHz), this work adopts a microstrip frequency-scanning array as a compact route to meet the practical objectives of constrained size, stable gain, and large detection-angle coverage.

The main contributions are as follows:

- (1) A parameter-selection guideline is provided based on the phase condition of frequency-scanning arrays, covering patch dimensions, inter-element spacing, element count, and feedline width;
- (2) With broadband impedance matching, the design achieves a 48° main-beam scan and ~74° effective coverage while maintaining a stable gain above 10.5 dBi.

In a typical millimetre-wave fuze, the antenna is installed on a compact projectile body and must operate reliably during high-speed flight and terminal engagement. The projectile attitude and the target aspect angle can change rapidly, so a fixed broadside beam may fail to maintain sufficient echo power at the receiver. From a system viewpoint, the antenna must deliver a stable gain within a wide angular sector, while remaining low-profile and mechanically robust.

Beam steering can be realised electronically (phase shifters, switched networks, or true-time-delay units), but these approaches increase cost, volume, and power consumption. Moreover, component tolerances and packaging parasitics become increasingly critical in the Ka-band. A frequency-scanning approach provides an alternative: the beam direction is controlled by frequency-dependent dispersion in the feed, eliminating phase-control circuits and improving system simplicity.

Related work on fuze antennas includes waveguide slot arrays, conformal patch arrays, multi-beam microstrip structures, and frequency-scanning arrays. Waveguide arrays generally offer high efficiency and high power handling, yet their height and mass can be difficult to accommodate on small platforms. Planar multi-beam designs (e.g., multi-port switching or pattern reconfiguration) provide several discrete pointing angles, but they do not naturally provide continuous coverage. Frequency-scanning arrays have been studied for radar and sensing, where an inherent frequency-angle mapping enables mechanical-free scanning. However, in fuze scenarios, the antenna must simultaneously satisfy compact size, acceptable matching over a wide band, and stable gain while scanning.

Compared with prior frequency-scanning implementations that rely on external loads or complex feeding networks, this manuscript emphasises a practical, fabrication-friendly design: a quasi-travelling-wave series-fed microstrip patch array with a load-free termination strategy. The design targets 30–36 GHz because this band offers a good balance between antenna size and atmospheric attenuation for short-range fuze sensing, and it aligns with commonly reported Ka-band fuze front-end architectures.

2. Frequency-scanning principle and design method

2.1. Principle of scanning mechanism

A frequency-scanning antenna is commonly realised using a uniform linear array^[4]. The far-field main-beam direction is determined by the effective phase difference between adjacent radiating elements. As frequency

changes, the propagation constant and electrical length of the feed network change accordingly, which modifies the inter-element phase difference and therefore steers the main lobe ^[5]. For a series-fed microstrip patch array operating in a quasi-traveling-wave mode, adjacent elements are excited sequentially with a fixed physical separation d and an effective feeding-line length $(d-L)$, where L denotes the patch length. The phase difference between two adjacent elements can be written as:

$$\varphi = \frac{2\pi}{\lambda_g}(d-L) \quad (1)$$

where λ_g is the guided wavelength along the microstrip line. In the far field, constructive interference in the direction θ requires the total phase difference between adjacent elements to satisfy the array phase condition

$$\begin{cases} -2\pi m = \frac{2\pi}{\lambda_1} d \sin \theta_1 - \frac{2\pi}{\lambda_{g1}}(d-L) \\ -2\pi m = \frac{2\pi}{\lambda_2} d \sin \theta_2 - \frac{2\pi}{\lambda_{g2}}(d-L) \end{cases} \quad (2)$$

where λ_0 is the free-space wavelength and m is the diffraction order.

Equation 2 establishes a direct relationship between the operating frequency, the inter-element spacing d , and the main-beam tilt angle θ . When $(d-L)$ is fixed, variations in frequency modify both λ_0 and λ_g , resulting in a frequency-dependent beam direction. This mechanism forms the theoretical basis of frequency scanning.

To avoid the occurrence of grating lobes, the inter-element spacing must satisfy the well-known constraint, which limits the maximum allowable spacing for a given scanning angle.

$$d < \frac{\lambda_0}{1+|\cos \theta|} \quad (3)$$

Meanwhile, for a specified progressive phase difference φ , the element spacing can also be expressed as:

$$d = -\frac{\varphi}{k \cos \theta} \quad (4)$$

where k is an integer related to the phase progression order.

Equation 3 and **Equation 4** jointly constrain the design space of the frequency-scanning array ^[3].

In the proposed antenna, a uniform linear array configuration is adopted, and all patch elements are excited using identical microstrip feeding lines. Owing to the traveling-wave nature of the structure, reflections along the feeding network are negligible. Therefore, the inter-element phase relationship can be effectively controlled by adjusting the physical length of the microstrip line, which directly determines the phase difference.

By properly selecting the inter-element spacing d , the main-beam direction can be steered toward either the feeding end or the load end ^[5]. Specifically, when $d < \lambda_g$, the main beam tends to tilt toward the feeding end, whereas for $d > \lambda_g$ it tilts toward the load end. In practical implementations, a relatively small spacing d is preferred to achieve beam steering toward the feeding end while simultaneously enabling antenna miniaturization.

2.2. Key parameter selection

The parameters are as follows:

- (1) Patch dimensions: A cavity-model approximation with an effective permittivity ϵ_{eff} is used to initialise the rectangular patch size so that the centre frequency is around 33 GHz;
- (2) Inter-element spacing d : Increasing d generally enlarges the scan range by strengthening the phase

gradient, but it may degrade matching and increase the risk of grating lobes; thus d is chosen by balancing scan range, matching, and pattern integrity;

- (3) Element number N : Larger N increases gain and narrows the beam, but also increases series-feed loss and overall length. To satisfy $\text{gain} \geq 10$ dB under size constraints, this work selects $N = 6$.

Patch initialisation uses standard transmission-line/cavity-model relations. Given a target centre frequency f_0 and substrate relative permittivity ϵ_r , the patch width is approximated by:

$$W = \frac{c}{2f_0} \left(\frac{\epsilon_r + 1}{2} \right)^{-\frac{1}{2}}$$

The effective permittivity ϵ_{eff} is then estimated to account for fringing fields, and the effective length L_{eff} is approximated as:

$$L = \frac{c}{2f_0\sqrt{\epsilon_e}} - 2\Delta l.$$

The physical length L is obtained by subtracting the fringing extension ($2\Delta L$). These closed-form steps provide a reliable starting point for full-wave optimization^[6].

Substrate selection is critical in the Ka-band. A lower-loss laminate reduces dielectric loss and improves radiation efficiency, while a thinner substrate suppresses surface waves and helps maintain pattern stability. Rogers 4350B is selected because it offers a moderate permittivity ($\epsilon_r \approx 3.5$) that keeps the antenna compact without making the feed excessively narrow, and it provides a low loss tangent suitable for millimetre-wave prototypes. The thickness $h = 0.254$ mm further reduces profile and mitigates surface-wave excitation.

Inter-element spacing d is chosen below one free-space wavelength to avoid grating lobes over the scan range^[7]. Using $\lambda_0 \approx 10$ mm at 30 GHz, the selected $d = 3.4$ mm corresponds to about $0.34 \lambda_0$, which provides a good trade-off between mutual coupling control and sufficient phase gradient for scanning. The element number N is set to 6 to obtain a stable gain above 10 dBi while limiting series-feed loss and overall size^[8].

3. Antenna configuration and parameters

3.1. Structure

The proposed antenna is a single-layer, series-fed microstrip patch array. Rectangular patches are cascaded along the array axis and excited by a microstrip feedline at the input port. Instead of attaching an explicit terminal load, the end section is shaped to provide an equivalent dissipative behaviour through distributed loss, which helps suppress reflections and improves matching. The substrate is Rogers 4350B ($\epsilon_r = 3.5$, $h = 0.254$ mm). The overall footprint is approximately $20 \text{ mm} \times 10 \text{ mm}$, which is suitable for compact projectile integration.

The series-fed architecture is implemented using microstrip line segments that excite each patch sequentially. To maintain consistent phase progression and reduce discontinuity reflections, the feedline widths are chosen according to the required characteristic impedance and manufacturability. A wider input line (y_1) helps reduce conductor loss near the feed, while a narrower inter-element line (y_2) enables finer impedance transformation and phase control between adjacent patches.

A practical termination is designed at the end of the feed to suppress reflections. Rather than adding an external chip load (which can be challenging to source and mount reliably at 30–36 GHz, the terminal section is shaped so that residual power is dissipated through distributed conductor/dielectric losses and weak radiation. This

approach improves robustness and repeatability for compact antenna modules.

From a fabrication perspective, the minimum line width and gap are kept compatible with standard PCB processes for Rogers laminates. Because the substrate is thin, the design is sensitive to etching tolerance and copper roughness. The layout therefore avoids extremely narrow traces and abrupt right-angle corners. In future hardware validation, a calibration kit and a low-loss end-launch connector (or a probe-fed fixture) would be required to measure $|S_{11}|$ accurately in the Ka-band.

3.2. Optimisation workflow

The optimisation follows a practical order: “element first, array next; resonance first, scan next”. The patch width W and length L are tuned to place the resonance within the target band. Followed by that is the input feedline width y_1 and inter-element feedline width y_2 are adjusted to obtain broadband impedance matching. Finally, under the constraint $|S_{11}| < -10$ dB across 30–36 GHz, the spacing d is refined to trade off gain, scan range, and pattern stability. This staged workflow reduces multi-parameter coupling and improves engineering reusability.

4. Simulation setup and results

4.1. Simulation model

Based on the theoretical calculation parameters listed in **Table 1** and the antenna structural, the antenna model is established using the ANSYS HFSS electromagnetic simulation software, as illustrated in **Figure 1**. The overall substrate size of the antenna is 20 mm × 10 mm. During the modeling process, boundary conditions and excitation are defined sequentially. The microstrip patch elements and the bottom ground plane are set as ideal conducting boundaries (Perfect E). The excitation is implemented in the form of a lumped port with concentrated terminal feeding (Lumped Port). According to the antenna dimensions, a suitable air box is constructed around the antenna and assigned a radiation boundary condition (Radiation).

Table 1. Optimised key parameters (unit: mm)

W	L	d	y_1	y_2	Substrate h	N
3.8	2.2	3.4	0.50	0.35	0.254	6

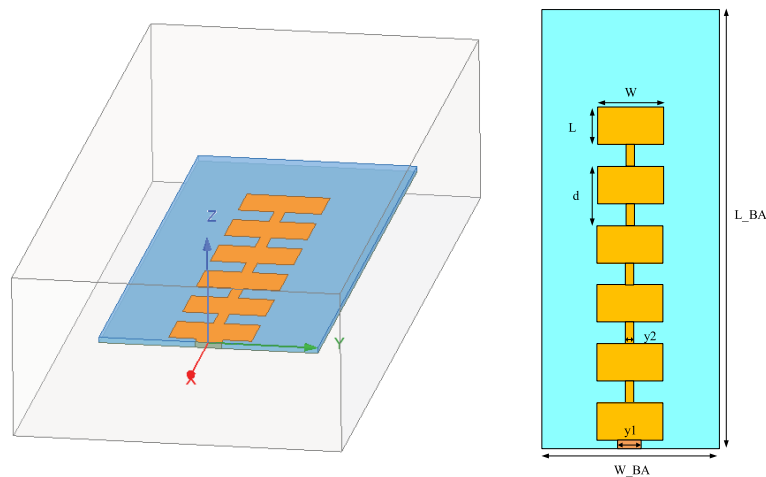


Figure 1. Structural model of the microstrip frequency-scanning antenna.

To excite the microstrip patch elements, a microstrip line is used for direct feeding, and a single-ended feeding configuration is adopted. By adjusting the width y_1 and length d_1 of the input microstrip line, as well as the width y_2 of the inter-element microstrip line between adjacent patch elements, impedance matching of the antenna to $50\ \Omega$ can be achieved. After initial simulation optimization, the input microstrip line width y_1 is set to 1.0 mm, the input line length d_1 is set to 0.5 mm, and the inter-element microstrip line width y_2 is set to 0.5 mm.

4.2. Parameter optimization

4.2.1. Frequency optimization

According to the calculation formula derived, it can be concluded that the resonant frequency of the antenna is mainly influenced by the length W and width L of the rectangular patch element. As the operating frequency increases, the dimensions of the patch element decrease, i.e., both W and L become smaller. Conversely, as the operating frequency decreases, the patch dimensions increase, corresponding to larger values of W and L .

4.2.2. Impedance-matching optimization

The matching objective is a $50\text{-}\Omega$ input with $|S_{11}| < -10$ dB across 30–36 GHz. Since the initial design does not satisfy this requirement at several frequencies, the matching network is refined by tuning the input feed-line width y_1 , the feed-line length d_1 , and the inter-element microstrip width y_2 . These parameters are co-optimized to stabilize the impedance response over the entire band.

4.2.3. Gain optimization

The realized gain is mainly determined by the inter-element spacing d and the element number N . The spacing d is optimized to achieve a favorable trade-off between gain level and frequency-scanning behavior (beam angle versus frequency), while maintaining acceptable matching.

4.2.4. Element-number selection

A small N provides wider scanning bandwidth but yields insufficient low-frequency gain (< 10 dB). Increasing N improves gain^[9]. However, overly large N exacerbates return loss and disrupts impedance matching, which in turn degrades the frequency-scanning characteristic and reduces effective scanning bandwidth and main-beam coverage. Considering these trade-offs, $N = 6$ is selected as the optimal configuration.

4.3. Simulation results and analysis

Based on the optimization procedure described above, a quasi-traveling-wave millimeter-wave microstrip frequency-scanning antenna with an inter-element spacing of $d = 3.4\text{ mm}$ and an element number of $N = 6$ is finalized. The antenna exhibits a frequency-dependent main-beam steering behavior, in which the main-beam tilt angle θ varies monotonically with operating frequency. As a result, the E-plane effective main-beam coverage exceeds 70° , ensuring robust target detection capability for fuze applications under a wide range of impact-angle conditions.

To verify compliance with the design specifications, the simulated results are analyzed in terms of the reflection coefficient, E-plane radiation patterns, and the frequency-dependent scanning radiation characteristics of the antenna.

4.3.1. S parameter

Figure 2 shows the simulated reflection coefficient of the six-element frequency-scanning antenna. It is observed that S11 remains below -10 dB throughout the operating band from 30 to 36 GHz, indicating good impedance matching across the entire frequency range. This confirms that the optimized feeding network successfully achieves a stable 50Ω input impedance over the desired bandwidth.

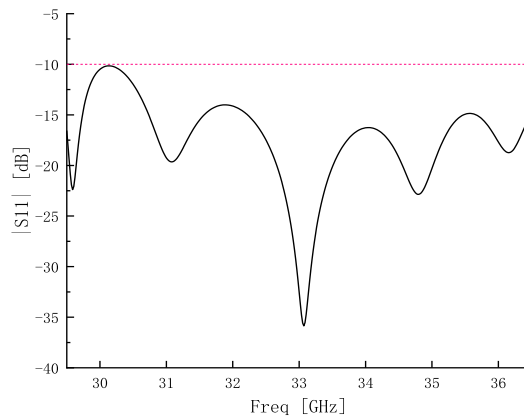


Figure 2. S11 curve.

4.3.2. E-plane radiation patterns

To evaluate the radiation performance, E-plane radiation patterns are extracted at six representative frequencies within the operating band, namely 31, 32, 33, 34, 35, and 36 GHz, with $\phi = 0^\circ$ and θ ranging from -180° to 180° , as shown in **Figure 3**.

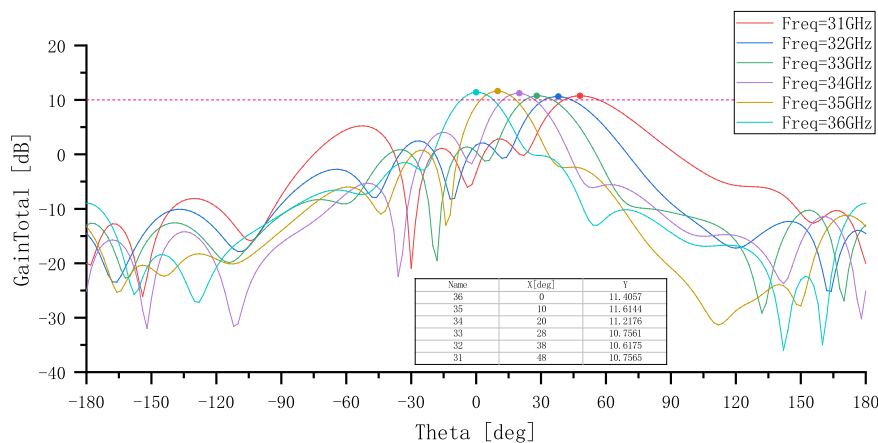


Figure 3. E-plane radiation patterns.

The results demonstrate that the antenna gain at all selected frequencies exceeds 10 dB, satisfying the design requirement. As the operating frequency increases from 30 to 36 GHz, the main-beam tilt angle continuously shifts from approximately 48° toward broadside (0°), confirming the intended frequency-scanning behavior. The maximum realized gain of approximately 11.6 dB is achieved at 35 GHz.

4.3.3. Frequency-scanning performance

Figure 4 presents the E-plane radiation pattern at the center frequency of 33 GHz.

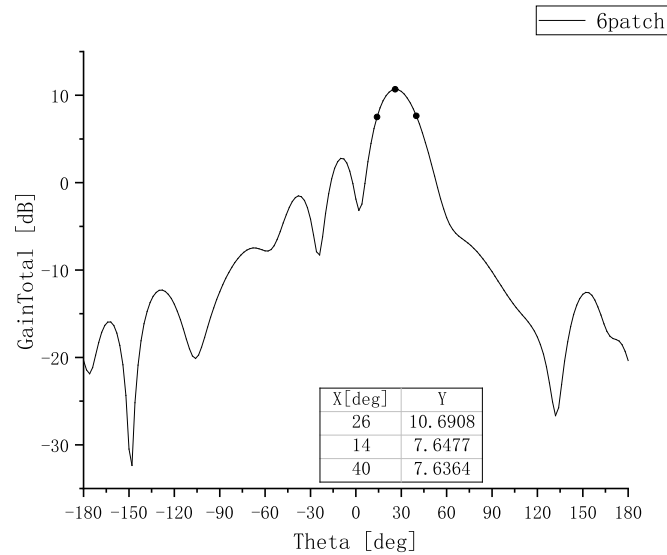


Figure 4. E-plane radiation pattern at 33 GHz.

At this frequency, the main-beam tilt angle is approximately 26° , and the E-plane half-power beamwidth (HPBW) is about 26° . Over the operating band from 30 to 36 GHz, the main-beam direction scans from 48° to 0° , corresponding to a scanning range of 48° . Considering the HPBW of approximately 26° , the effective E-plane main-beam coverage angle reaches about 74° , providing wide angular coverage suitable for millimeter-wave fuze applications.

5. Conclusion

From an engineering perspective, the design demonstrates a low-complexity route to obtain broad angular coverage in the Ka-band without resorting to bulky waveguides or active phase-control circuits, which can be advantageous for compact and high-reliability fuze sensing modules. A compact Ka-band microstrip frequency-scanning antenna for fuze applications has been presented. Based on a quasi-travelling-wave series-fed patch array on Rogers 4350B, the design achieves broadband matching ($|S_{11}| < -10$ dB over 30–36 GHz), stable gain above 10.5 dBi (peak ≈ 11.6 dBi), and continuous E-plane beam steering from 48° to 0° . By incorporating the 3-dB beamwidth, the effective detection-angle coverage reaches $\sim 74^\circ$, which is beneficial under large impact-angle variations. Future work will include conformal installation analysis on projectile bodies, sidelobe suppression via spacing tapering or amplitude weighting, and fabrication/measurement validation with tolerance-aware modelling. As the beam direction is frequency dependent, a practical fuze radar can exploit frequency selection to choose the desired pointing direction on demand. For example, lower frequencies in the 30–36 GHz band correspond to larger tilt angles, while higher frequencies steer the beam closer to broadside in this design. This enables a simple scan strategy where the transmitter steps through several frequency points to cover the sector, and the receiver combines echoes accordingly. Such an approach is compatible with agile frequency synthesizers and can be implemented without adding RF phase-control hardware.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Liu Y, 2009, Research and Design of Microstrip Frequency-Scanning Antenna Arrays, thesis, Nanjing Univ. Sci. Technol.
- [2] Wang Q, 2019, Design of Miniaturized Radio Fuze Antennas, thesis, Nanjing Univ. Sci. Technol.
- [3] Fu Y, 2015, Research on Millimeter-Wave Frequency-Scanning Antennas, thesis, Harbin Inst. Technol.
- [4] Wang D, Wang Z, Xu L, et al., 2017, A Millimeter-Wave Microstrip Frequency-Scanning Fuze Antenna Based on a Quasi-Traveling-Wave Array. *Bingqi Zhuangbei Gongcheng Xuebao (Journal of Ordnance Equipment Engineering)*, 38(6): 156–160.
- [5] Liu J, Wu D, Yao M, 2008, Design of Millimeter-Wave Waveguide Slot Antennas. *Journal of Ordnance Engineering*, 29(2): 15–17.
- [6] Wang Y, Zhang L, Kou D, 2004, Design of Wideband Waveguide Slot Antennas. *Journal of Ordnance Engineering*, 25(2): 37–41.
- [7] Li Z, 2021, Design of Wireless Radio-Frequency Waveguide Antennas, thesis, Nanjing University of Science and Technology.
- [8] He L, 2016, Thermal Design of Airborne Millimeter-Wave Active Phased Array Antennas. *Journal of Ordnance Equipment Engineering*, 37(5): 115–119.
- [9] Zhou X, 2006, Simulation Design and Error Analysis of Millimeter-Wave Fuze Antennas. *Zhi Dao Yu Yin Xin (Guidance and Fuze)*, 27(4): 28–31.

Publisher's note

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



Integrated Services Platform of International Scientific Cooperation

Innoscience Research (Malaysia), which is global market oriented, was founded in 2016. Innoscience Research focuses on services based on scientific research. By cooperating with universities and scientific institutes all over the world, it performs medical researches to benefit human beings and promotes the interdisciplinary and international exchanges among researchers.

Innoscience Research covers biology, chemistry, physics and many other disciplines. It mainly focuses on the improvement of human health. It aims to promote the cooperation, exploration and exchange among researchers from different countries. By establishing platforms, Innoscience integrates the demands from different fields to realize the combination of clinical research and basic research and to accelerate and deepen the international scientific cooperation.

Cooperation Mode



Clinical Workers



In-service Doctors



Foreign Researchers



Hospital



University



Scientific institutions

OUR JOURNALS



The *Journal of Architectural Research and Development* is an international peer-reviewed and open access journal which is devoted to establish a bridge between theory and practice in the fields of architectural and design research, urban planning and built environment research.

Topics covered but not limited to:

- Architectural design
- Architectural technology, including new technologies and energy saving technologies
- Architectural practice
- Urban planning
- Impacts of architecture on environment

Journal of Clinical and Nursing Research (JCNr) is an international, peer reviewed and open access journal that seeks to promote the development and exchange of knowledge which is directly relevant to all clinical and nursing research and practice. Articles which explore the meaning, prevention, treatment, outcome and impact of a high standard clinical and nursing practice and discipline are encouraged to be submitted as original article, review, case report, short communication and letters.

Topics covered by not limited to:

- Development of clinical and nursing research, evaluation, evidence-based practice and scientific enquiry
- Patients and family experiences of health care
- Clinical and nursing research to enhance patient safety and reduce harm to patients
- Ethics
- Clinical and Nursing history
- Medicine



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

