Research on the Construction of an AI Image Recognition-Based Evaluation System for Middle School Physics Experiment Operational Compliance

Ruiqi Zhang

The First Middle School of Chongqing University Town Shapingba District，Chongqing，401331;

Abstract： With the rapid development of artificial intelligence (AI) technology, its application in the educational field is becoming increasingly widespread. This study focuses on research concerning the construction of an AI image recognition-based evaluation system for middle school physics experiment operational compliance. Through an analysis of the current state of evaluation in middle school physics experimental teaching, the necessity of introducing AI technology is elaborated. Key technologies involved in system construction are introduced in detail, including image recognition algorithms, data acquisition, and processing. Practical case studies demonstrate the system's significant effectiveness in improving evaluation accuracy, objectivity, and teaching efficiency, providing valuable insights for the intelligent development of middle school physics experimental teaching.

Keywords: AI Image Recognition; Middle School Physics Experiments; Operational Compliance; Evaluation System

DOI:10.69979/3041-0843.25.03.033

1 Introduction

Experimental teaching in middle school physics is crucial for cultivating student competencies. Accurately and objectively evaluating the compliance of experimental operations is key to enhancing teaching quality. However, traditional evaluation relies on teachers' manual observation, presenting problems such as difficulty in comprehensively monitoring student operations, susceptibility to subjective influences, and inconsistent evaluation standards. AI image recognition technology presents an opportunity to address these issues. An evaluation system built upon this technology can monitor experiment operations in real-time, analyze them precisely, and automatically generate results according to predefined standards. This improves efficiency and objectivity in evaluation, holding significant importance for the optimization of middle school physics experimental teaching.

2 Overview of AI Image Recognition Technology

2.1 Fundamental Principles of AI Image Recognition Technology

AI image recognition technology is a vital branch of artificial intelligence, aiming to enable computers to understand image content and automatically identify and classify objects, scenes, behaviors, and other information. It is based on theories such as deep learning and computer vision. By constructing

neural network models and training them on large datasets of images, the models learn to recognize patterns and features withi images. During recognition, images are first preprocessed (e.g., grayscale conversion, noise reduction) before being input into models like Convolutional Neural Networks (CNN). CNNs consist of convolutional layers, pooling layers, and fully connected layers, responsible for feature extraction, dimensionality reduction, and classification decision-making, respectively. Model parameters are adjusted during training to minimize error. Once trained, the model can recognize unknown images and output relevant information.

2.2 Current Application Status of AI Image Recognition Technology in Education

In recent years, the application o AI image recognition technology in education has emerged, bringing numerous innovations. In intelligent teaching assistance, it can identify student classroom behaviors to help teachers adjust strategies, and automatically grade homework and tests, reducing teacher workload and providing error analysis. Its potential in experimental teaching is immense. For example, in chemistry, it can identify apparatus assembly and reactions; in biology, it can recognize cells; and in physics, attempts at monitoring and analysis exist. However, applications are mostly exploratory. Evaluation systems specifically for middle school physics experiment operational compliance are yet to be perfected, and widespread adoption faces challenges.

3 Analysis of Current Evaluation Methods for Middle School Physics Experiment Operations

3.1 Traditional Evaluation Methods and Their Limitations

Traditional evaluation of middle school physics experiment operations primarily relies on teachers' on-site observation and subjective judgment. Due to limited energy, teachers find it difficult to comprehensively monitor the operational details of every student, leading to some non-compliant operations going uncorrected promptly, which hinders the development of students' experimental skills. Simultaneously, teacher subjectivity has a significant impact. Differing understandings of operational standards and varying levels of strictness in evaluation reduce the objectivity and fairness of results. Furthermore, manual evaluation is inefficient, becoming time-consuming and labor-intensive in large-scale scenarios, affecting teaching progress and the timeliness of feedback.[2]

3.2 Necessity of Introducing AI Image Recognition Technology

Introducing AI image recognition technology to construct an evaluation system for middle school physics experiment operational compliance is highly necessary. Such a system can monitor experiment operations comprehensively and in real-time using multiple cameras, offering broader coverage than teacher observation and enabling the timely detection of non-compliant operations. This technology also ensures objective and accurate evaluation. By analyzing operations according to unified standards,it eliminates subjective influences and provides students with precise feedback. Moreover, it significantly improves evaluation efficiency, rapidly processing data to generate results, facilitating teaching improvement and enriching content. It represents a crucial approach to solving evaluation problems and enhancing teaching quality.[3]

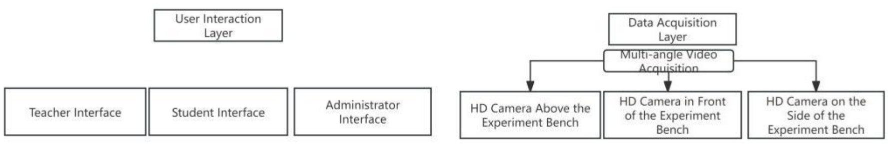
4 Construction of the AI Image Recognition-Based Evaluation System

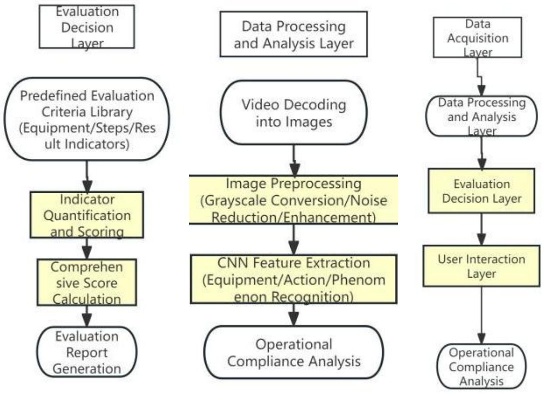
4.1 Overall System Architecture Design

The system comprises four layers: Data Acquisition Layer, Data Processing & Analysis Layer, Evaluation Decision Layer, and User Interaction Layer.

Data Acquisition Layer: Utilizes high-definition cameras positioned at critical points on the experiment bench to capture operations from multiple angles. Video streams are transmitted to the next layer.Data Processing & Analysis Layer: Decodes videos into image sequences. After preprocessing (e.g., grayscale conversion), it uses CNN models to extract features, identifying apparatus, actions, etc., and analyzes operational compliance.

Evaluation Decision Layer: Quantifies evaluation indicators (e.g., apparatus usage) based on predefined standards, calculates a comprehensive score, and generates reports.User Interaction Layer: Provides interfaces for teachers and students to view relevant information and for administrators to maintain the system.





4.2 Key Technology Implementation

4.2.1 Image Acquisition and Preprocessing

Image Acquisition: Employs high-definition cameras with good low-light performance to adapt to varying lighting conditions. Cameras are positioned at multiple angles based on bench layout and operational characteristics (e.g., overhead for apparatus/hands, front/side for overall posture), enabling 360-degree monitoring through camera collaboration.Image Preprocessing:Grayscale Conversion: RGB images are converted to grayscale using the weighted average method to reduce data volume.Noise Reduction: Gaussian filtering is applied, using a Gaussian kernel to perform weighted averaging on pixels, suppressing noise.Enhancement: Histogram equalization adjusts the grayscale distribution to enhance contrast, improving the discernibility of key information and laying the foundation for subsequent processing.

4.2.2 Deep Learning-Based Image Recognition Algorithm

Deep learning algorithms exhibit outstanding capabilities in the field of image recognition. This system employs the Convolutional Neural Network (CNN) as its core recognition algorithm. CNN is specifically designed for processing grid-structured data like images. Through the combination of convolutional layers, pooling layers, and fully connected layers, it can automatically learn image features.

The specific network architecture of the CNN model needs to be designed based on the characteristics of middle school physics experiment operation images and recognition requirements. It primarily includes the following key components:

1.Convolutional Layers:The number of layers is determined according to actual needs.Multiple convolution kernels of different sizes are used alternately to extract multi-scale features from the experiment operation images (e.g., detailed features of experimental apparatus, overall characteristics of student operational actions).

2.Pooling Layers:Max pooling layers are employed.They perform downsampling on the feature maps output by the convolutional layers.This reduces the data volume while preserving important features.It enhances model processing efficiency.

3.Fully Connected Layers:These layers are used to integrate the extracted features and perform classification decisions.They map the learned features onto recognition results relevant to the experiment operations (e.g., determining the type of experimental apparatus, or whether an operational action is compliant).

Overall, the combination of convolutional layers, pooling layers, and fully connected layers enables the model to automatically learn the features of middle school physics experiment operation images, achieving accurate recognition of experiment-related information.

When constructing the CNN model:Based on the characteristics of middle school physics experiment operation images and recognition requirements, a suitable network architecture must be designed.Key parameters to determine include the number of convolutional layers, convolution kernel parameters, and pooling layer types (e.g., using convolution kernels of different sizes to extract multi-scale features, and utilizing max pooling layers to reduce data volume while preserving important features).

During model training:A large number of experiment image samples containing both correct and incorrect operations are collected and annotated.The annotations clearly specify information such as the experimental apparatus, operational actions, and their compliance within the samples.The data is divided into a training set, a validation set, and a test set.The backpropagation algorithm is used during training to adjust model parameters in order to minimize prediction error.After extensive iterative training, the model masters the relevant image features and acquires the capability to accurately recognize experimental apparatus, operational actions, and judge their compliance.

4.2.3 Establishment and Quantification of the Evaluation Index System

The evaluation index system is crucial for the accuracy and scientific rigor of the system. It is constructed based on the middle school physics curriculum standards and experimenta teaching syllabi, covering three aspects: Apparatus Preparation, Operation Process, and Result Presentation.Apparatus Preparation Indicators: Include correctness of selection, completeness, cleanliness, and placement rationality (e.g., scoring for selecting correct apparatus and arranging it properly in an electricity experiment).Operation Process Indicators: Include step correctness, operational proficiency/compliance, and safety awareness (e.g., scoring for following steps correctly an prioritizing safety in a mechanics experiment).Result Presentation Indicators: Include measurement accuracy, completeness/standardization of data recording, and conclusion correctness (e.g., scoring for accurate measurements and correct conclusions in an optics experiment).Quantification: Indicators are refined, and detailed scoring rules are established. For example, experimental steps are decomposed into key operational points, each assigned a score, ensuring more objective and accurate evaluation results.

4.3 Introduction to System Functional Modules

4.3.1 Experiment Operation Monitoring and Real-time Feedback

This module collects image data of student experiment operations in real-time and utilizes the image recognition algorithm for real-time analysis. Upon detecting non-compliant operations (e.g., incorrect apparatus use, steps performed out of order), the system immediately sends feedback to the student via the User Interaction Layer, prompting timely correction. Simultaneously, the system flags images or video clips of non-compliant operations for subsequent teacher review and guidance.Example: During a circuit connection experiment, if the system detects a student connecting a wire to the wrong terminal, it immediately pops up an alert indicating the error location and the correct connection method, helping the student correct the mistake promptly and improve operational compliance.

4.3.2 Evaluation Result Generation and Report Output

Upon completion of an experiment, the system automatically generates evaluation results based on the predefined index system and image recognition analysis. Results include a comprehensive score, scores for

individual indicators, and a detailed evaluation report. The report not only highlights strengths and weaknesses in the student's operation but also provides specific improvement suggestions and reference materials to aid understanding and mastery of operational standards.Example: A report might indicate that a student lacked proficiency in balancing a beam balance during a mass measurement experiment. It would recommend watching a relevant demonstration video and provide the link. Teachers and students can easily view results and reports via the User Interaction Layer. Students can engage in targeted learning based on the report, while teachers can adjust teaching strategies based on class-wide evaluation data, enhancing teaching relevance and effectiveness.

4.3.3 Data Management and Analysis

This module is responsible for storing, managing, and analyzing student experiment operation data, archiving videos, images, and evaluation results categorically. The system enables multi-dimensional data analysis (e.g., performance across different experiments, class/grade differences). Teachers can leverage these analyses to understand teaching effectiveness and inform decision-making.Example: Identifying a class with a high error rate in electricity experiments could prompt targeted reinforcement.

5 System Application Case Study andEfectiveness Analysis

5.1 Application Case Introduction

Two parallel Grade 8 (equivalent to US Grade 8/Year 9) classes in a middle school were selected. One class (Experimental Class) used the AI image recognition-based evaluation system during physics experiment teaching. The other class (Control Class) used the traditional teacher-led manual evaluation method. Over one semester, both classes completed the same physics experiment curriculum, covering mechanics, electricity, optics, etc.

Experimental Class: During experiments, the system automatically captured operation images, performed real-time analysis and evaluation. Students received immediate feedback to correct non-compliance. Post-experiment, students accessed detailed reports. Teachers used system-generated results and data analysis reports for summaries, feedback, and targeted instruction on common issues.

Control Class:Teachers observe and evaluated operations during experiments and provided scores and summaries afterward, following traditional methods and standards.

5.2 Efectiveness Analysis

5.2.1 Improvement in Evaluation Accuracy and Objectivity

Comparative analysis revealed that the Experimental Class using the system achieved significantly higher evaluation accuracy and objectivity. Under traditional manual evaluation, teacher subjectivity led to inconsistencies; for instance, in evaluating the "Measurement of Sliding Friction" experiment, scores given by two different teachers to the same student differed by more than 5 points (out of 100) in 23% of cases. In contrast, the AI system's

evaluations were highly consistent, with score variations for the same student/same experiment below 2 points across different evaluations, effectively eliminating evaluator bias. Furthermore, the system precisely identified subtle non-compliance. For example, in the "Ohm's Law Experiment (Measuring Resistance with Volt-Ampere Method)", it accurately detected students failing to open the switch before connecting the circuit - an error missed by teachers in about 35% of cases under traditional evaluation - demonstrating the system's accuracy advantage.

5.2.2 Efect on Teaching Quality Enhancement

Teaching practice over the semester showed that the Experimental Class exhibited significantly better operational compliance than the Control Class. In the end-of-semester practical skills assessment, the Experimental Class's average score was 12.5 points higher. The gap was particularly notable in the operational compliance score item. Specifically, circuit connection errors in electricity experiments decreased by 40% in the Experimental Class, and instrument usage compliance in mechanics experiments improved by 35%. This improvement is attributed to the system's real-time feedback, allowing students to correct errors immediately and form correct habits. Additionally, the evaluation reports provided teachers with data for targeted instruction. For example, after identifying inconsistent data recording in the "Convex Lens Imaging Law" experiment within the Experimental Class, teachers conducted focused training, reducing the occurrence rate of this issue by 50%.

5.2.3 Improvement in Evaluation Eficiency

The AI system demonstrated substantial efficiency gains. Manually evaluating a class of 45 students typically required 3 class periods (~ 135 minutes). The system evaluated a class of the same size in just 20 minutes, representing an efficiency improvement of over 85%. The advantage is even more pronounced in large-scale assessments: evaluating 8 full-grade classes took 3 teachers 2 days manually, whereas the system completed all evaluations and generated detailed reports within 1 hour. This drastically reduces teacher workload, freeing up time for lessonplanning and personalized guidance.

6 System Limitations and Improvement Directions

6.1 Analysis of Existing Problems

Despite significant results, the system faces challenges:Technical Limitations: Adaptability to complex scenarios needs improvement. Sudden lighting changes or occlusions can reduce image recognition accuracy by 10%-15%.

Data Limitations: Training data primarily covers common operations, lacking sufficient samples of rare errors, leading to lower recognition accuracy for such cases.Evaluation Scope: While the index system covers main operational aspects, it currently struggles to evaluate innovativeactions performed within compliance standards, making it difficult to distinguish and reward students' unique, compliant approaches.[1]

6.2 Discussion on Improvement Directions

Future improvements could focus on:Technical Optimization: Introduce multi-sensor fusion (e.g., combining infrared sensors, depth cameras) to compensate for limitations of pure image recognition in complex environments, potentially boosting accuracy above 95%.Data Expansion: Collaborate with multiple schools to collect more diverse operation samples, especially rare error cases, building a more comprehensive dataset. Utilize generative AI to synthesize virtual samples, enhancing data

diversity.Evaluation System Enhancement: Add an "Innovation" dimension. Analyze optimized steps or unique approaches within compliant operations, awarding bonus points to encourage exploration and innovation within established standards.

7 Conclusion and Outlook

This research successfully constructed an AI image recognition-based evaluation system for middle school physics experiment operational compliance. It effectively addresses the accuracy, objectivity, and efficiency issues inherent in traditional evaluation methods, enhancing evaluation precision and student operational compliance. The proposed approach and system design are feasible and practical. Future work will focus on developing a more intelligent and personalized system, potentially integrating technologies like Natural Language Processing (NLP) and Virtual Reality (VR), to further advance the overall intelligence of experimental teaching evaluation.

References

[1] Li Fangzhou, Xu Jinze, Fang Lei, et al. Perception Quality Evaluation System Based on AI Image Recognition Technology [C]// Proceedings of the 31st Annual Conference of the China Society of Automotive Engineers (Volume 1). 2024.
[2] Wang Jie. Research on Garbage Classification System Design and Software Quality Assurance Based on AI Image Recognition Technology [J]. Mobile Information, 2024, 46(8): 367-369.
[3] Long Yan, Ji Yaojiu, Wang Yang. An AI Classroom System for Physics, Chemistry,and Biology Experiments Based on Image Recognition: CN202010239809.5 [P]. CN111314390A. [Accessed: 2025-07-03].