Journal of International Education and Development
2025, VOL. 9, NO. 8, 93-96
DOI: 10.47297/wspiedWSP2516-250021.20250908

Research on the Application of AI Image Recognition in Online Education

Xiangyu Zhao

School of Communication Technology, Communication University of China Nanjing, Nanjing, Jiangsu, 211172, China

ABSTRACT

The rapid expansion of online education is facing core challenges in terms of teaching quality and learning outcomes. Traditional classroom teachers struggle to monitor students' learning status, and the teaching process has blind spots, with limited dimensions for assessing learning outcomes. Breakthroughs in artificial intelligence technology have opened up new possibilities for educational digitization, with image-intelligent analysis demonstrating unique value. Therefore, this paper explores the practical application of computer vision technology in online education scenarios, focusing on three core directions: dynamic tracking of learning behavior, intelligent management of teaching environments, and multi-dimensional assessment of skill development. The aim is to provide a feasible path for constructing the next-generation innovative education platform that is "perceptible, warm, and capable of evolution."

KEYWORDS

AI image recognition; Online education; Learning behavior analysis; Personalized learning

1 Introduction

The evolution of digital technology has profoundly reshaped the educational ecosystem. Online education, with its ability to transcend the limitations of time and space, is increasingly becoming an essential medium for knowledge transfer. Artificial intelligence image recognition technology has made breakthrough progress, with its powerful visual analysis capabilities enabling the precise identification of multi-dimensional information such as facial expressions, posture, and movements, and learning environments without interfering with the learning process. This capability aligns closely with the requirements of online education scenarios, providing a technical foundation for learning immersion monitoring, classroom participation assessment, question-answering process assistance, and automated resource adaptation. While there have been occasional attempts at localized technical applications, there remains a lack of systematic research to outline integration pathways and future development directions. This paper aims to integrate cutting-edge AI image recognition methods with key business scenarios in online education to provide insights for constructing an innovative educational environment characterized by strong interactivity, adaptability, and practical guidance.

2 Overview of AI image recognition technology

2.1 Basic concepts of image recognition technology

Image recognition technology refers to a multidisciplinary field that simulates the human visual system using computers to automatically recognize and understand objects, scenes, actions, and features in images or videos. Its essence lies in endowing machines with the ability to "see-recognize-judge," i. e., the process of extracting semantic information from raw pixel data. Traditional pattern recognition methods rely on manually designed feature extractors combined with classification algorithms for object identification [1]. Modern image recognition has evolved from passive perception to active understanding, capable of identifying object categories and spatial positions in static images, capturing dynamic behavioral features, and even inferring implicit intentions. Notably, its capabilities are expanding from structured scenes to complex unstructured environments, such as simultaneously recognizing diverse heterogeneous information like board writing text, student postures, and device statuses in a classroom setting. This technological evolution not only transforms human-machine interaction forms but also continuously expands the analytical dimensions of the visual world, laying the cognitive foundation for cross-domain applications.

2.2 Core technologies and methods of AI image recognition

The breakthrough progress in AI image recognition stems from the revolutionary advancements in deep learning technology, with its core being a multi-layer feature abstraction architecture dominated by convolutional neural networks. This architecture simulates the hierarchical processing mechanism of biological visual pathways, progressively constructing high-level semantic representations from low-level pixels: the lower-level convolutional layers extract edge and basic texture features, the middle-level network fuses simple features to form local patterns, and the higher-level abstract network generates interpretable semantic labels (e. g., identifying the complete scene of "a student solving trigonometric functions on a blackboard"). Key algorithms such as RNN and Transformer further enhance spatio-temporal association modeling capabilities, enabling the detection of continuous writing actions by students [2]. The latest trend is the maturation of multimodal fusion algorithms, such as combining camera input with microphone audio streams to achieve dual-channel cognition of teacher-student interaction scenarios. When the system detects a student raising their hand but fails to recognize their spoken words, it can automatically trigger the teaching assistant prompt module. Additionally, technological innovations such as efficient feature pre-training through self-supervised learning and model lightweighting for edge devices are continuously improving the real-time performance and accuracy of online scenario deployment.

2.3 Application areas and trends of AI image recognition

Image recognition technology has deeply penetrated highly specialized fields such as industrial manufacturing, autonomous driving, medical diagnosis, and security surveillance. It is accelerating its expansion into more everyday and social applications. In educational settings, its evolution is undergoing three major transformations: first, shifting from specialized devices to ubiquitous sensing, using ordinary cameras for contactless monitoring; second, transitioning from single-action recognition to behavioral chain analysis, inferring teaching effectiveness through temporal analysis of students' note-taking speed and teachers' movement trajectories; and finally, upgrading from an auxiliary tool to a decision-making hub, such as generating a heatmap of academic anxiety indices in smart classrooms by combining human pose estimation and micro-expression recognition. Technological trends are focusing on four key dimensions: first, the integration of real-time recognition and edge computing, which reduces the response delay of online education systems through algorithm lightweighting; second, cross-device federated learning mechanisms, which enable multi-terminal collaborative training while protecting student privacy; third, the deep embedding of generative image models, such as using image synthesis technology to generate visual materials adapted to different teaching stages automatically; fourth, enhanced explainability driven by social and ethical considerations, with a focus on mitigating algorithmic bias risks in educational settings. These evolutionary directions collectively point toward a paradigm shift in the core of technology, from a mere sensing tool to infrastructure that participates in the reconstruction of educational elements.

3 Advantages and challenges of online education

3.1 Advantages of online education

3.1.1 Flexibility and convenience

The core breakthrough of online education lies in fundamentally eliminating the constraints of traditional learning time and space. Its flexibility and convenience have completely reshaped the paradigm of educational participation. Fixed schedules no longer constrain learners. They can choose their learning times based on their schedules, particularly benefiting groups with significant work-study conflicts, those living in remote areas, or those with mobility limitations, by providing them with equal educational opportunities. Efficient knowledge acquisition in fragmented scenarios has become possible, such as completing micro-courses on mobile devices during commutes or children continuing their learning plans at home during sudden health issues. This flexible mechanism also extends to the interactive level: real-time live streaming classrooms meet the need for immediate answers to questions, while the post-class playback function supports repeated digestion of difficult points at one's own pace [3]. Teachers also benefit from this transformation, as the shift from fixed podiums to digital lecturers has significantly simplified teaching organization processes—the efficiency revolution in homework grading, resource sharing, and feedback dissemination has freed up time previously spent on repetitive tasks, enabling educators to focus on core teaching design and personalized care.

3.1.2 Widespread access to resources

Online platforms have reshaped the resource allocation ecosystem by breaking down physical barriers, transforming the scarcity of high-quality educational resources into scalability. Digital media enables individuals to access knowledge repositories that were previously inaccessible due to the obstacles of elite universities or geographical constraints: a rural student can enroll in open courses to study quantum mechanics lectures from top universities. At the same time, self-learners can use international platforms to access the latest programming framework tutorials from Silicon Valley. The breadth and depth of resource aggregation have significantly expanded, manifested both in cross-institutional shared course repositories (integrating teaching modules from hundreds of universities across different disciplines) and in the cross-supply of diverse learning materials. This open ecosystem has catalysed a multi-layered resource derivative network:

teachers can integrate multi-source videos to create specialized course packages, while students can compare different instructors' varied interpretations of the same knowledge points to deepen their understanding.

3.2 Challenges of online education

3.2.1 Student motivation and self-discipline issues

As the educational environment transitions from physical campuses to virtual spaces, the lack of mechanisms to sustain student motivation has become a critical pain point. The constraints imposed by teachers' authority, peer pressure, and the ritualistic nature of learning in physical classrooms are significantly weakened in online settings. When interacting with a screen in a single-point manner, students are easily distracted by multitasking environments, leading to fragmented attention spans and difficulty in establishing a state of deep concentration. This impact is particularly pronounced among groups with underdeveloped self-directed learning abilities—adolescents may selectively ignore complex assignments when lacking direct supervision. At the same time, adult learners may postpone their study plans due to occupational pressures, eventually abandoning courses altogether [4]. Particularly concerning is the dilution of social presence: the enthusiasm for learning sparked by teacher-student physical interaction and collective psychological fields in traditional classrooms cannot be effectively transmitted in fragmented, isolated learning environments, leading some users to develop feelings of isolation and cognitive loneliness. This psychological mechanism poses a deep threat to course completion rates, necessitating the reconstruction of the motivational system through external intervention and internal incentives.

3.2.2 Challenges in learning assessment and tracking

The distortion and lag in teaching evaluation during the virtualization process have become key bottlenecks hindering precise teaching. In traditional classrooms, teachers' ability to identify immediate learning blind spots through continuous observation of body language is simplified to lost video signals online. The anti-cheating mechanisms of paper-based exams are difficult to transplant into virtual environments, and online tests are prone to academic misconduct, such as proxy testing or using multiple devices to look up information. A deeper issue lies in the blurring of formative assessment: offline teachers review students' thought processes through annotation traces on assignments, but this has been replaced by video platforms' single completion rate statistics and mechanical multiple-choice question scoring, leading to the fundamental causes of conceptual misunderstandings being obscured by data appearances. Although platforms have assignment submission functions, the cognitive complexity of text and images poses a significant challenge to automatic grading technology. The combination of these assessment mechanism flaws and the absence of learning trajectories (such as the inability to record the duration of video pauses or the paths of literature page jumps) ultimately creates an information black hole in teaching decisions—teachers cannot dynamically optimize course structures based on precise data. Students fail to receive customized feedback integrated into cognitive contexts, turning educational upgrades into exploratory practices based on vague assumptions.

4 AI image recognition technology in online education

4.1 Application of AI image recognition in learning behavior analysis

In virtual classrooms, image recognition technology acts as an observer, helping teachers understand students' states across the screen. It captures natural learning scenes through ordinary cameras. When students' eyes drift away from the screen for an extended period or they repeatedly rub their eyes, the system adjusts the pace of teaching content, for example, breaking down complex concepts into engaging interactions or inserting immediate questions during key segments. During group discussions, this technology analyzes members' facial orientations and speaking patterns to visually display participation activity maps, enabling teachers to encourage quiet students to express their views promptly [5]. This technology is particularly effective in hands-on courses, where students operating experimental equipment receive subtle prompts such as "moving the tool two centimeters to the left yields better results" based on standard imagery. The entire process requires no additional devices, maintaining the natural flow of the course. All processing of personal video feeds is completed within the device, with only brief prompts generated regarding the overall status.

4.2 Application of AI image recognition in classroom management

AI image recognition technology provides an invisible safeguard for order in online teaching spaces. The teacher's management interface can detect abnormal classroom activities. When students repeatedly switch to unrelated programs or multiple students leave the camera's field of view simultaneously, a gentle vibration reminder is triggered to maintain focus. During essential exams, the system continuously monitors students' movement patterns. Sudden frequent head turns or prolonged downward glances are recorded as data points, which are used in conjunction with screen recordings to maintain a fair environment. This is particularly important in safety-sensitive courses. If chemical equipment operation exceeds safety boundaries or art tools are swung with excessive force, immediate alerts are sent to both teachers and students. In daily teaching, most monitoring functions are displayed as semi-transparent cartoon pop-ups, with analysis modules only briefly activated when necessary to avoid the tension caused by continuous monitoring, ensuring learning safety only at critical moments.

4.3 Application of AI image recognition in assessment and feedback

Image recognition is driving the transformation of online evaluation systems toward a process-oriented approach. When students display their math problem-solving drafts in front of the camera, the system can assess the validity of key derivation steps without relying on manual input of formula documents. Language assignment analysis focuses more on constructive guidance, such as observing the connection between English cursive letters and pointing out specific letters that could be written more smoothly and neatly, rather than simply judging the aesthetic appeal of the handwriting. Skill development courses benefit significantly: during drawing exercises, the system provides color mixing suggestions based on the naturalness of adjacent color areas; in mechanical training, it compares the learner's hand gesture paths with standard images to generate training records with directional guidance. Feedback responses have become more sensitive—when a confused expression is detected, related case analyses are pushed; when a confident expression is captured, in-depth extension questions are provided. This allows teaching evaluations to go beyond simply marking the correctness of final answers, better reflecting the value of the thinking development process.

5 Conclusion

AI image recognition technology has opened up new possibilities for online education. By dynamically analyzing learning behavior, optimizing classroom management mechanisms, and innovating assessment and feedback models, it has significantly improved the accuracy and interactive experience of teaching. This technological transformation is driving the evolution of education from standardization to personalization, providing new perspectives for addressing long-standing challenges such as insufficient learning motivation and weak process supervision.

About the Author

Xiangyu Zhao, master's degree, lecturer, research direction: computer science and technology.

References

[1] Sun Yi, Zhou Junhui, Ren Jie. A Study on the Practical Application of Intelligent Selection of Experimental Instruments and Chemicals in Junior High School Chemistry Based on Image Recognition [J]. China Modern Educational Equipment, 2024, (24): 66-69.

[2] Zheng Jiping, Wang Meijing, Leng Duanjie. Research and Application of Image Recognition Technology in Online Teaching [J]. Industrial and Information Technology Education, 2023, (12): 85-88.

[3] Lan Xuwen. Integration of Labor Education in Primary School Artificial Intelligence Experiential Practice: Taking "Ability to See and Distinguish—Image Technology Experience" as an Example [J]. China Educational Technology Equipment, 2022, (23): 106-108+115.

[4] Jiang Liping, Sun Chunyan. Application of Image Recognition Technology in Biology Experiments [J]. Middle School Biology Teaching, 2023, (04): 49-51.

[5] Huang Yutong. Scenario Applications of Artificial Intelligence in Higher Education [J]. Journal of Harbin Engineering University, 2022, 43(07): 1066-1072.