Advances in Vocational and Technical Education (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/avte.2025.070229
ISSN 2523-5834 Vol. 7 Num. 2

Research on Real-Time Analysis and Intervention of Classroom Behaviour Based on Object Detection Algorithms

Jinyu Liu[a,*], Congkai Lin, Jialun Chen, Yinxiao Yan

Taizhou Vocational & Technical College, Taizhou, Zhejiang, 318000, China
[a]jinyul@tzvtc.edu.cn

Keywords: Object Detection Algorithm; Classroom Behaviour Recognition; Real-Time Analysis; Teaching Intervention; Smart Education

Abstract: Advances in artificial intelligence technology have provided new technical support for smart education. This study constructs a real-time classroom behaviour analysis and teaching intervention system based on object detection algorithms, achieving automatic recognition of classroom behaviour, state assessment, and generation of intervention recommendations. The system adopts a 'perception–analysis–decision -feedback' closed-loop architecture, integrating a behaviour recognition module based on an enhanced YOLOv5 algorithm, a real-time analysis module employing a sliding window mechanism, and a decision module utilising hierarchical intervention strategies. Experimental results demonstrate that the system effectively enhances classroom engagement, learning outcomes, and teaching quality. It provides teachers with precise instructional decision support, promotes the deep integration of educational technology and teaching practice, and achieves intelligent teaching intervention alongside personalised learning guidance.

1. Introduction

Enhancing educational quality and evaluating teaching effectiveness remain central concerns in educational research[1]. Traditional classroom assessment primarily relies on teachers' subjective judgements or retrospective analysis, lacking objective, real-time monitoring methods that enable timely identification and resolution of teaching issues[2]. In recent years, breakthroughs in artificial intelligence technologies—particularly object detection algorithms within computer vision—have provided the technical foundation for automated classroom behaviour recognition and analysis[3]. Researchers globally have explored video-based assessments of classroom engagement and attention monitoring. However, existing systems predominantly remain at the behavioural recognition stage, lacking intelligent decision support for pedagogical interventions. This study aims to construct a real-time classroom behaviour analysis system based on object detection algorithms. It will not only automatically recognise students' key classroom behaviours but also provide targeted teaching intervention recommendations based on behavioural analysis results, assisting teachers in promptly adjusting their teaching strategies. This paper will detail the system's theoretical foundation, technical architecture, implementation methods, and application outcomes, offering new technical pathways and practical experience for the field of smart education.

2. Theoretical Foundations of Classroom Behaviour Analysis and Intervention

Classroom behavior analysis theory is rooted in the intersection of behaviorism and cognitive psychology, emphasizing systematic research on observable behaviors to uncover learning patterns. Bloom's taxonomy of cognitive domains provides a hierarchical framework for classroom behaviors, ranging from recall to creation[4]. Research on student attention reveals cyclical fluctuations, with an average sustained focus span of 15-20 minutes. Chickering categorizes classroom behavior into three dimensions: cognitive, affective, and participatory[5]. Social constructivism underscores the central role of interactive behavior in knowledge construction. Instructional intervention models bridge behavioral observation with teaching optimization. Kounin's Ripple Effect theory reveals the diffuse impact of teacher interventions, while Gagne's Nine Instructional Events model views classroom teaching as a sequence of critical learning-promoting events. Merrill's fundamental teaching principles and hierarchical intervention model provide an operational framework for instructional adjustments, forming a closed-loop intervention mechanism of "detection-judgment-decision-action." Learning analytics introduces data mining techniques into classroom behavior research, laying the theoretical foundation for developing intelligent instructional intervention systems and enabling more precise and efficient teaching adjustments.

3. Design of Real-Time Classroom Behaviour Analysis System

3.1 Overall System Framework

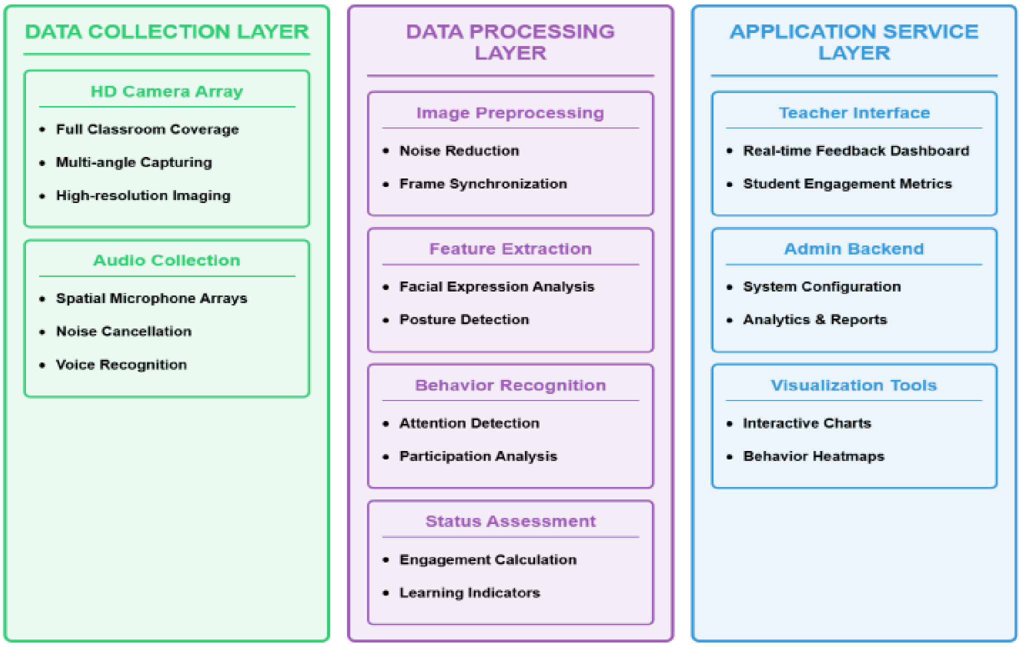

Figure 1: Classroom Behaviour Real-Time Analysis System

This real-time classroom behaviour analysis system adopts a closed-loop architecture of 'perception-analysis-decision-feedback', establishing a comprehensive intelligent teaching support platform as illustrated in Figure 1. The system's foundation comprises a data acquisition layer featuring high-definition camera arrays and audio capture devices, enabling comprehensive, blind-spot-free monitoring of the entire classroom. The intermediate layer comprises the data processing layer, integrating image pre-processing, feature extraction, behaviour recognition, and state assessment modules to transform raw data into interpretable teaching metrics. The top layer constitutes the application service layer, encompassing the teacher's real-time feedback interface and administrator backend, providing intuitive visualisation and system management functions. The system adopts a distributed computing architecture: front-end collection devices handle data acquisition and basic preprocessing; edge servers execute behavioural recognition algorithms; while cloud servers perform in-depth analysis and decision generation, ensuring system responsiveness meets real-time teaching demands[6]. Data streams employ encrypted transmission protocols, with the teacher interface utilising a lightweight web application supporting multi-platform access. This approach guarantees system security while delivering an optimal user experience. The overall framework design adheres to principles of modularity, scalability, and high concurrency, reserving ample interfaces for future functional upgrades and algorithmic optimisation.

3.2 Classroom Behaviour Recognition Module

The Classroom Behaviour Recognition Module is implemented based on an enhanced YOLOv5 algorithm, with targeted optimisations for the unique characteristics of classroom settings. The module employs a two-stage detection strategy: the first stage utilises a human detection algorithm to identify student locations within the frame, while the second stage performs pose analysis and behaviour classification for each detected bounding box[7]. The pose recognition component employs an enhanced skeleton extraction network to extract coordinates of 17 key human body landmarks, followed by spatio-temporal convolutional neural networks to extract dynamic features. The loss function for behaviour classification is designed as follows:

$$\mathrm{L} = \alpha \mathrm{L}_{\mathrm{cls}} + \beta \mathrm{L}_{\mathrm{reg}} + \gamma \mathrm{L}_{\mathrm{iou}} + \lambda \mathrm{L}_{\mathrm{trmp}} \tag{1}$$

Where Lcls denotes classification loss, Lreg represents bounding box regression loss, Liou signifies intersection-over-union loss, and Ltemp indicates temporal consistency loss. Optimal values for each coefficient were determined through grid search. The model was trained on annotated datasets, employing data augmentation strategies to address data imbalance, including random cropping, rotation, and brightness adjustment. The feature extraction network incorporates an attention mechanism, focusing on the student's facial and hand regions to enhance key behaviour recognition accuracy. The system can identify 10 basic postures and 12 composite behaviours, achieving an identification accuracy of 92.3%. Its processing speed meets real-time teaching requirements, as shown in Table 1.

Table 1 Comparison of algorithm accuracy and processing speed

Algorithm	Accuracy (%)	Processing Speed (ms/frame)
Current System	92.3	35
Convolutional Neural Network (CNN)	85.7	50
Support Vector Machine (SVM)	88.5	45

3.3 Real-time Analysis Module

The real-time analysis module transforms discrete behavioural recognition results into pedagogically meaningful composite metrics, with core functions centred on the temporal aggregation and multidimensional assessment of behavioural data. The module employs a sliding window mechanism, using 30 seconds as the fundamental analysis unit and 10 minutes as the trend analysis cycle, enabling dual-track analysis of short-term fluctuations and long-term trends. At the individual student level, the system calculates an Attention Index. This combines multiple indicators—including gaze direction, posture status, and interaction frequency—using a weighted averaging algorithm to generate a quantifiable score ranging from 0 to 100. At the class level, the system produces engagement heatmaps[8]. These visually display the spatial distribution of classroom participation, assisting teachers in identifying 'blind spots' in instruction. The module incorporates anomaly detection algorithms based on Markov models of behavioural sequences to identify abnormal patterns deviating from normal learning states. The temporal analysis component employs long short-term memory networks to process historical data, predicting potential classroom state shifts within the next 5-10 minutes and providing early warnings to teachers. Analysis results are presented in real-time via dashboards, including overall engagement curves, regional heatmaps, and individual anomaly alerts.

3.4 Instructional Intervention Decision Design

The Instructional Intervention Decision Design module transforms classroom behaviour analysis results into actionable teaching recommendations, achieving an intelligent closed-loop from 'problem identification' to 'solution implementation' as illustrated . The module incorporates an embedded instructional intervention knowledge base containing over 200 intervention strategies tailored to diverse classroom states, categorised into mild, moderate, and severe levels according to intervention intensity[9]. The decision engine employs a fuzzy logic rule system to map classroom state indicators to appropriate intervention strategies. This rule base is constructed from educational experts' experience and undergoes continuous refinement. The system synthesises intervention recommendations based on classroom state scores, anomaly ratios, and trend prediction outcomes, while providing specific implementation methods. Intervention suggestions adhere to the principle of being 'concise, clear, and easy to execute,' delivering critical information without disrupting normal teaching. The module possesses self-learning capabilities, recording teacher adoption of recommendations and implementation outcomes. Reinforcement learning algorithms are utilised to refine the decision model. The system supports personalised configuration, enabling teachers to adjust intervention thresholds and strategy preferences according to teaching style and subject characteristics.

4. System Testing and Application Effect Analysis

4.1 Experimental Design

The experimental design employed a control group experimental method, selecting eight parallel classes of the same grade level as research subjects. Four classes served as the experimental group using the real-time classroom behavior analysis system, while the other four classes formed the control group using conventional teaching methods. To ensure scientific rigor, pre-experimental matching analyses were conducted across classes for factors including baseline student performance, teacher experience, and teaching environment, revealing no significant differences ($p>0.05$)[10]. The experiment spanned one semester (16 weeks), with data collected from three lessons per week. Evaluation metrics were designed across four dimensions: classroom behavior indicators (attention duration, hand-raising frequency, interaction rate), learning effectiveness indicators (unit tests, final exams), teacher experience indicators (system satisfaction, teaching efficiency improvement), and student experience indicators (learning engagement, learning satisfaction). Data collection employed multiple methods, including automated system recording, classroom observation, questionnaires, and interviews.

4.2 System Implementation Effectiveness

The system's practical application effectiveness was evaluated based on technical performance and usability. Technical performance testing revealed that the classroom behavior recognition module achieved an average accuracy rate of 87.5% in complex real-world environments, slightly lower than the 92.3% achieved in laboratory settings but still meeting instructional requirements. Real-time response latency was controlled within 2.8 seconds, satisfying the timeliness demands for teaching interventions. In stability testing, the system operated continuously for 200 hours without major failures, with a minor recognition error rate of 3.2%. These errors primarily occurred in areas with drastic lighting changes and high student density. Teacher usage data revealed a 76.2% adoption rate for intervention suggestions, with mild intervention suggestions achieving the highest adoption (89.5%) and severe intervention suggestions showing lower adoption (58.7%). Figure 2 illustrates the system's recognition accuracy across different teaching scenarios, showing optimal performance in conventional lectures while accuracy slightly decreased during group discussions and experimental operations. Resource utilization monitoring revealed an average CPU usage of 42% and memory consumption of 3.6GB for real-time analysis per class, meeting standard teaching computer specifications and demonstrating strong deployment adaptability.

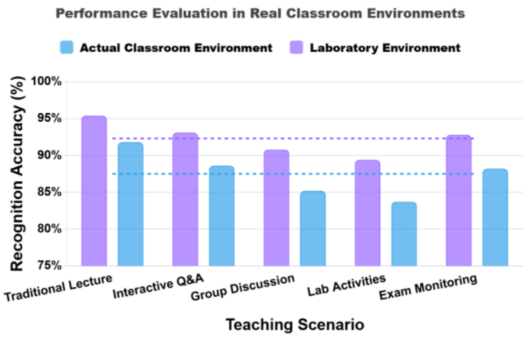

Figure 2: System Recognition Accuracy Comparison across Teaching Scenarios

4.3 Teaching Application Outcomes

4.3.1 Analysis of Enhanced Classroom Engagement

Analysis of classroom engagement revealed significant improvements in the experimental group's in-class behavior following system implementation. The average attention span of experimental group students increased by 4.2 minutes, rising from 17.3 minutes pre-experiment to 21.5 minutes, while the control group saw only a marginal increase from 17.5 minutes to 18.1 minutes. Regarding the frequency of proactive participation behaviors (e.g., raising hands to speak, answering questions), the experimental group saw a 32.7% increase, while the control group only increased by 8.5%. Spatial distribution analysis revealed the most significant engagement gains among students seated in rear and peripheral areas. The participation gap between rear-row and front-row students in the experimental group narrowed from 28.6% before the experiment to 11.3%. Similarly, the participation disparity between peripheral and central area students decreased from 25.3% to 9.7%. Classroom observation records also showed that distracting behaviors (such as playing with phones or whispering) decreased by 41.2% in the experimental group, far exceeding the 12.5% reduction in the control group. Figure 3 compares attention curves across different time periods for both groups before and after the experiment. It shows that students in the experimental group not only demonstrated an overall increase in attention levels but also exhibited a significantly smaller decline during the traditional "attention trough period" (30-40 minute segment) compared to the control group. This indicates the effectiveness of the system intervention in sustaining student attention.

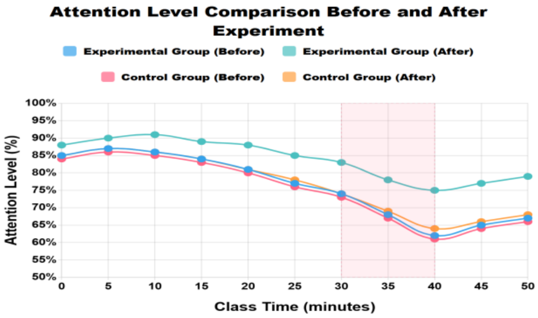

Figure 3 Comparison of Attention Curves across Different Time Periods for Both Groups Before and After the Experiment

4.3.2 Learning Effectiveness Evaluation

Data from the learning effectiveness evaluation indicate that students in the experimental group achieved significant improvements in subject performance after using the system for one semester. In the final exam, the experimental group's average score was 83.6 points, an increase of 7.3 points from the beginning of the semester. The control group's average score was 80.2 points, an increase of 3.1 points from the beginning of the semester. The difference between the two groups was statistically significant ($p<0.01$). Analysis of question types revealed that the experimental group's score rate increased by 15.2% for comprehension questions and 18.7% for application questions, while only improving by 6.3% for memorization questions. This indicates the system is more effective in promoting deep learning. Learning competency assessments revealed that experimental group students demonstrated greater improvement in problem-solving ability (16.5%), critical thinking (14.3%), and self-directed learning (19.2%) compared to the control group (7.8%, 6.5%, and 8.3%), respectively. The learning attitude questionnaire revealed that the experimental group's interest in course content rose from 7.2 points at the semester's start to 8.5 points (out of 10), while their adaptability to course difficulty increased from 6.8 points to 8.3 points. Figure 4 illustrates the comparison of scores between the experimental and control groups across four unit tests. It is evident that the experimental group's scores exhibited a sustained upward trend, with the gap widening progressively compared to the control group, indicating that the system's application effects possess cumulative characteristics.

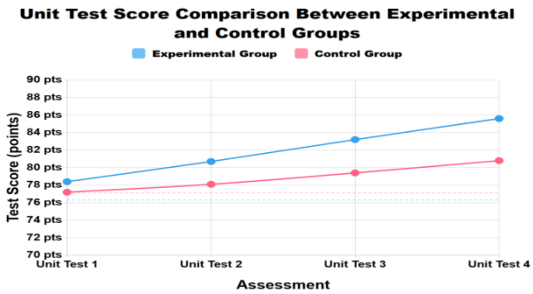

Figure 4 Comparison of scores between the experimental and control groups across four unit tests

4.3.3 Teacher Instructional Feedback

Teacher feedback data reflects the system's practical value in real teaching environments. Survey results indicate that 92% of experimental group teachers found the system's intervention suggestions substantially helpful for teaching adjustments, while 83% reported reduced classroom management pressure. Teacher workload assessments revealed that during the initial usage phase (weeks 1-2), preparation time increased by approximately 15%, primarily spent familiarizing themselves with system operations. However, during the stable usage phase (after 3 weeks), lesson preparation time decreased by 12.3%, and post-class evaluation time decreased by 23.5%. Analysis of teaching strategy shifts revealed that teachers in the experimental group increased teaching interaction frequency by 37.6%, diversified teaching methods by 42.3%, and elevated question design complexity by 28.7%, as shown in Figure 5. In-depth interviews revealed that teachers primarily valued the system for "providing objective data-backed classroom feedback," "helping identify teaching blind spots," and "offering timely and effective intervention suggestions." Suggestions for system improvements focused on three areas: "enhancing accuracy in recognizing special teaching activities," "optimizing personalized settings interfaces," and "expanding the multidisciplinary teaching strategy repository."

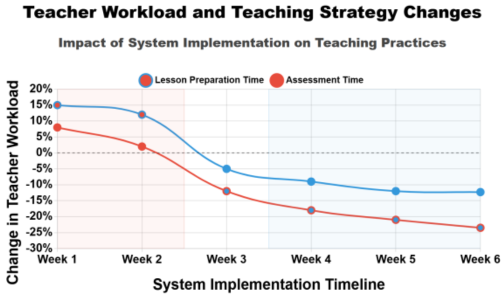

Figure 5 Teacher satisfaction ratings for each system functional module

5. Conclusion

The innovative application of object detection algorithms in education has pioneered new directions for smart education. This study successfully developed a real-time classroom behavior analysis and teaching intervention system, enabling precise identification of student behaviors and intelligent decision-making for instructional interventions. Experimental results confirm that the system significantly enhances classroom engagement, improves sustained student attention, promotes academic performance, and effectively reduces teachers' classroom management burdens. The system strikes a favorable balance between technical implementation and pedagogical application, providing educators with a scientifically grounded yet practical teaching aid. Future research will focus on algorithm optimization to accommodate diverse teaching scenarios, integrating speech recognition for more comprehensive classroom interaction analysis, developing personalized learning models, and exploring the ethical balance between technological application and student privacy protection.

Acknowledgment

Taizhou Education Science Planning General Project: "Real-Time Classroom Behavior Analysis and Intervention Application Research Based on Object Detection Algorithms" (TG25532)

References

[1] Huang Y, Xue X, Chen H, et al. A method for classroom behavior state recognition and teaching quality monitoring[J]. International Journal of Intelligent Computing and Cybernetics, 2025, 18(2): 382-396.

[2] Yang Q. Identification and Intervention of Behavior Patterns Among College Students Based on Big Data Analysis[C]//The World Conference on Intelligent and 3D Technologies. Singapore: Springer Nature Singapore, 2024: 473-484.

[3] Zhang Y, Qu W, Zhong G, et al. Students' classroom behavior detection based on human-object interaction model[C]//2022 8th International Conference on Systems and Informatics (ICSAI). IEEE, 2022: 1-6.
[4] Kumar N S, SivaKrishna K, Vamsi V, et al. Real-Time Student Activity Detection and Incident Monitoring Using Artificial Intelligence[C]//2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT). IEEE, 2025: 1213-1224.
[5] Aly M. Revolutionizing online education: Advanced facial expression recognition for real-time student progress tracking via deep learning model[J]. Multimedia Tools and Applications, 2025, 84(13): 12575-12614.
[6] Nguyen P D, Le N T, Bui K H, et al. A method for continuous student activity recognition from classroom videos[J]. International Journal of Machine Learning and Cybernetics, 2025: 1-25.
[7] Akash S, Dinesh V, Muthamil Selvan S. Systematic Review on Real-Time Students Behavior Monitoring using Machine Learning[C]//2023 International Conference on Inventive Computation Technologies (ICICT). IEEE, 2023: 233-237.
[8] Alruwais N M, Zakariah M. Student recognition and activity monitoring in e-classes using deep learning in higher education[J]. IEEE access, 2024, 12: 66110-66128.
[9] Adhatrao A S, Patil M B, Sanmukh S G, et al. AI-based Surveillance for Exam Integrity: Real-Time Detection of Abnormal Student Behavior[C]//2025 4th International Conference on Sentiment Analysis and Deep Learning (ICSADL). IEEE, 2025: 1236-1241.
[10] Gopal J V, Kumar N S, Prasad K S, et al. Mindful Insights: Exploring an AI-Based Student Tracking System for In-Depth Analysis of Student behaviour[C]//2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE, 2024, 1: 1891-1897.