Proceedings of the 4th International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2026.36220

Research on Intelligent Analysis and Decision Support Method for Kindergarten Play Behavior Oriented to Teacher Observation

Weili Yan

School of Education, Jilin International Studies University, Changchun, China
wellington589125@gmail.com

Abstract. Kindergarten free play is an important context in which children's social interaction, language expression, emotional regulation, and problem solving abilities naturally emerge. Teacher observation provides a key basis for identifying children's developmental status and offering educational support. Traditional observation mainly depends on teacher experience, field records, and post activity reflection. It can preserve the richness of educational contexts, but it may lead to incomplete records, inconsistent judgment, and delayed suggestions when peer interaction is frequent and play behavior changes rapidly. This study develops an intelligent analysis and decision support method for kindergarten play behavior oriented to teacher observation. A teacher support system is constructed to integrate video observation data, teacher records, and manual behavior coding into a computable, explainable, and feedback based analytical process. Based on free play data collected from middle and senior classes in two kindergartens from 2024 to 2025, the system identifies solitary play, parallel play, cooperative play, role play, conflict regulation, and help seeking behavior. It further generates teacher oriented support suggestions. The results show that the multimodal model performs better than the single vision model in recognizing complex interactive behaviors, especially cooperative play and conflict regulation. The recommendation priority matrix also presents the matching relationship between play behavior types and teacher support strategies. The proposed method helps transform teacher observation from experience based judgment to data assisted guidance. It provides an operational technical path for play observation, educational intervention, and professional reflection.

Keywords: Kindergarten play behavior, Teacher observation, Intelligent analysis, Decision support, Explainable recommendation

1. Introduction

Kindergarten play is a natural context in which children's cognitive development, social interaction, emotional expression, and problem-solving ability become visible. The observation of teachers serves as the foundation for the identification of the development stage of children and for providing educational assistance accordingly. With the gradual transition of quality evaluation in early

DOI: 10.54254/2755-2721/2026.36220

childhood education from outcomes assessment to process assessment, the participation of children, their relations with peers, utilization of materials, and emotional engagement in playing activities have been acquiring increasing pedagogical importance [1]. Teacher observation in contemporary practices relies largely on field notes, professional experience, and after-the-fact reflection. While the former two ways keep intact contextual complexity, it is challenging to continuously record the fast-changing behaviors in frequent interactions between children. There is hope to find an intelligent solution through the structure of observational data using computer vision, multimodal recognition, and recommendable interpretation of video footage, textual notes, and behavioral codes [2]. Therefore, this study aims to build a teacher assistance system which consists of play behavior recognition, organization of observational data, and generation of pedagogical suggestions.

2. Literature review

2.1. Observation of kindergarten play behavior

Research on kindergarten play behavior observation usually focuses on children's participation patterns, social relationships, and developmental clues in free play. Playing is regarded as more than just a method of doing things since playing is viewed as a process through which children gain experience by means of bodily movements, speech acts, interactions with peers and objects. Teachers make conclusions concerning children's developmental needs based on their interest, role negotiation, conflict regulation and cooperative construction while observing them [3]. Existing literature reveals that observations have to stay contextualized since the context in which children's behavior appears plays a very important role. The method of manual observation demonstrates teachers' professionalism but, at the same time, in the case when there are many children engaged in various activities and their actions develop rapidly, records may become incomplete and evaluations may become inconsistent. Behavior coding system offers a classification framework that makes it possible to convert children's solitary playing, parallel playing, cooperative playing and conflict behaviors into quantifiable data [4]. This conversion does not decrease the contextuality of analysis; conversely, it creates a reliable base for teacher's reflection.

2.2. Intelligent behavior analysis in educational contexts

Intelligent analysis of behavior in education is based on methods of computer vision, speech recognition, sensors, and machine learning to detect movements, locations, interaction rate, and emotions during learning. Researches on learning engagement, teacher-student interaction, and cooperation among students in classroom settings give us the methodological framework for detection of kindergarten play behavior [5]. Kindergarten setting is more free and dynamic because children move much, bodily movements are less noticeable, and occlusions happen very often. It means that one visual feature is not enough to interpret the pedagogical sense of play behavior. Multimodal integration will allow us to include video trajectories, distance between peers, activity of voices, and teacher's written observations into common analysis which will enhance recognition of interaction states [6]. The main problem of intelligent analysis is not only the accuracy of the classification but also if it is possible to interpret the recognition results from the teacher's point of view. In case if there is no clear connection between labels and the observation facts, technical outputs cannot enter the teacher's decision-making process.

DOI: 10.54254/2755-2721/2026.36220

2.3. Teacher decision support and explainable recommendation

Research on decision support for teachers focuses on transforming educational data into usable judgments. The key importance of such a research field is that it helps teachers to comprehend the children, change the environment, and find the best ways to help. Explainable artificial intelligence highlights the presentation of reasoning of models that allows teachers to see sources of the evidence, feature weights, and context used for recognizing behavior [7]. Generated suggestions in the kindergarten play observation have to be structured in terms of educational actions, i.e., reminders for teachers to keep watching underparticipating children, increase collaborative tools, rearrange spatial layout, or foster verbal negotiation [8]. Moreover, the teacher's feedback influences the quality of system suggestions because teacher evaluation of relevance, feasibility, and context of recommendation could change recommendation rules. Data-driven guidance could create practical value if it is used together with the professional experience of teachers and without violating the children's privacy [9].

3. Experimental methods

3.1. Observation data collection and behavior coding

Data were collected in kindergarten free-play settings from middle and senior classes in two kindergartens between 2024 and 2025. Building, role-play, and construction areas were selected as observation fields. Data were collected twice weekly for 30 minutes per session, yielding 60 valid five-minute video clips, 300 minutes of observation materials, and 360 teacher observation records. The dataset included children's spatial positions, peer interactions, material use, verbal negotiation, emotional responses, and teacher field notes. Parental informed consent was obtained, and identifying information was de-identified. Behavior coding used 10-second windows with six labels: solitary play, parallel play, cooperative play, role play, conflict regulation, and help-seeking. Two coders annotated the data independently, and inconsistent clips were reviewed before the final annotations were used for model training and evaluation.

3.2. Construction of a multimodal behavior recognition model

Multimodal behavior recognition used video clips, teacher records, and manual coding as inputs. Video data extracted movement trajectories, peer distance, body orientation, and material contact frequency, while teacher records provided semantic features such as behavioral intention, interaction partners, and emotional descriptions [10]. Each 10-second window generated a fused feature vector combining visual, temporal, and textual information, as shown in Formula (1). The fused vector was then input into the classifier to predict behavior categories, as shown in Formula (2). The dataset was divided into 70% training and 30% testing data, and performance was evaluated using accuracy, recall, and F1 score. This design ensured that recognition results remained consistent with teacher observation evidence.

$$z_i = \alpha x_{v,i} + \beta x_{t,i} + \gamma x_{s,i} \tag{1}$$

In formula 1, z_i represents the fused feature of the i_{th} time window. $x_{v,i}$ represents visual features. $x_{t,i}$ represents temporal features. $x_{s,i}$ represents semantic features from teacher texts. α, β, and γ represent the fusion weights of different modalities.

DOI: 10.54254/2755-2721/2026.36220

$$\hat{y}_i = \arg\max_k \mathrm{softmax}(W_k z_i + b_k) \tag{2}$$

In formula 2, $\widehat{y}_i$ represents the predicted behavior category of the i_{th} time window. k represents the candidate behavior category. W_k represents the classification parameter of the k_{th} behavior category. b_k represents the bias term.

3.3. Teacher-oriented decision support mechanism

The teacher support decision-making mechanism was built on behavior recognition results and converted behavior category, duration, peer distance, verbal interaction intensity, and teacher feedback into recommendation priorities. The system generated executable actions, including continued observation of low-participation children, adding collaborative materials, adjusting play space, guiding verbal negotiation, and recording emotional changes [11]. Recommendations were ranked by combining behavioral concern, developmental support needs, and teacher feedback fit rather than relying on a single behavior label, as shown in Formula (3). To improve interpretability, feature contributions were further calculated and high-contribution features were converted into evidence statements in observation reports, as shown in Formula (4). This process linked behavior recognition, explanatory evidence, and teacher action suggestions.

$$S_j = \lambda_1 R_j + \lambda_2 D_j + \lambda_3 F_j \tag{3}$$

In formula 3, S_j represents the priority score of the j_{th} suggestion. R_j represents the concern level of the corresponding behavior. D_j represents the developmental support need. F_j represents the fit of teacher feedback. λ_1, λ_2, and λ_3 represent weight parameters.

$$E_{j,m} = \frac{\phi_{j,m}}{\sum_{m=1}^{M} |\phi_{j,m}|} \tag{4}$$

In formula 4, $E_{j,m}$ represents the explanatory contribution of the mt_h feature to the j_{th} suggestion. $\phi_{j,m}$ represents the local explanation value of this feature. M represents the total number of features entering the recommendation generation module.

4. Results

4.1. Behavior recognition performance

The prototype validation used 60 de identified free play videos, with 70% for training and 30% for testing. The consistent annotations of two coders served as the reference. As shown in Table 1, the multimodal model achieved an overall accuracy of 0.83 and a macro average F1 value of 0.81. Compared with the single vision model, it performed better in cooperative play, role play, and conflict regulation. The F1 value of cooperative play increased from 0.74 to 0.84, while conflict regulation increased from 0.68 to 0.78. These improvements indicate that teacher records helped supplement visual features, especially for behaviors involving verbal negotiation, material sharing, and interaction continuity. Help seeking behavior reached an F1 value of 0.82, though partial occlusion still caused some errors. These results support the feasibility of the multimodal approach in a small scale prototype test.

DOI: 10.54254/2755-2721/2026.36220

Table 1. Comparison of behavior recognition model performance

Behavior category	F1 value of single vision model	F1 value of multimodal model
Solitary play	0.86	0.89
Parallel play	0.79	0.83
Cooperative play	0.74	0.84
Role play	0.72	0.80
Conflict regulation	0.68	0.78
Help seeking behavior	0.76	0.82

4.2. Application effect of decision support

The application effect of the teacher support system was presented through a recommendation priority matrix. The system matched six types of play behavior with five teacher support strategies and generated priority scores ranging from 0 to 1, with higher scores indicating stronger support needs. As shown in Figure 1, cooperative play achieved the highest score of 0.86 for adding collaborative materials, indicating the importance of material sharing and peer collaboration in extending cooperative play. Conflict regulation scored 0.88 for guiding verbal negotiation and 0.81 for recording emotional changes, showing that the system could transform conflict recognition into process-oriented support. Solitary play scored 0.82 for continued observation, helping avoid treating short-term solitary behavior directly as a developmental risk. Overall, the results were consistent with the recommendation priority calculation and demonstrated the transition from behavior recognition to teacher action suggestions.

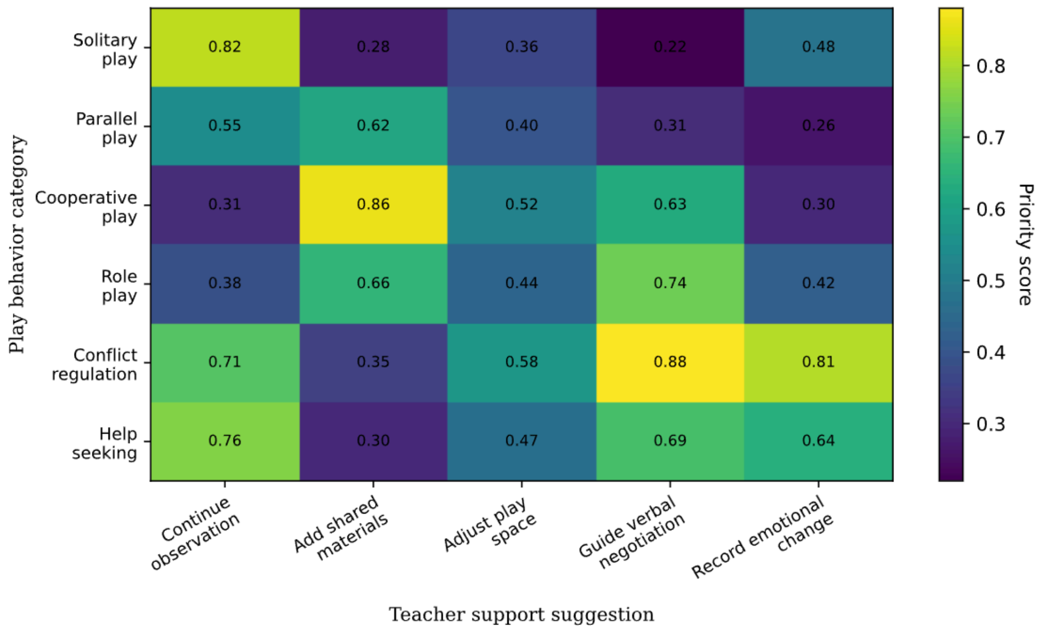

Figure 1. Recommendation priority heat matrix for play behavior support

DOI: 10.54254/2755-2721/2026.36220

5. Discussion

The value of intelligent analysis of kindergarten play behavior lies in transforming scattered observation information into traceable evidence, which supports more stable teacher judgment. The results show that complex behaviors, such as cooperative play, role play, and conflict regulation, need to be identified through spatial relations, behavioral continuity, and teacher records. A single vision model cannot fully explain children's interaction intentions. Multimodal fusion improves recognition performance while preserving the explanatory role of teacher field records. The decision support module converts recognition results into specific suggestions, so observation can directly support activity adjustment and child guidance. System outputs still need to be interpreted according to children's differences, classroom contexts, and teacher experience. Privacy protection, data authorization, and algorithmic transparency should also be emphasized.

6. Conclusion

The intelligent analysis and decision-support method for kindergarten play behavior forms a complete process from observation data collection and behavior coding to multimodal recognition and explainable recommendation generation. The results show that integrating video, temporal, and teacher textual features improves the recognition of complex play behaviors, particularly cooperative play, conflict regulation, and help-seeking. Based on recommendation priority and feature contribution, the system converts model predictions into practical teacher actions, including continued observation, adding collaborative materials, adjusting play areas, and guiding verbal negotiation. This improves the completeness of observation and the specificity of educational support. Future research may expand sample size and age coverage and improve robustness under occlusion, limited verbal interaction, and overlapping interactions

References

[1] Chu, C., et al. (2024). Digital play in the early years: A systematic review. *International Journal of Child-Computer Interaction, 40*, 100652.
[2] Alotaibi, M. S. (2024). Game-based learning in early childhood education: A systematic review and meta-analysis. *Frontiers in Psychology, 15*, 1307881.
[3] Kervin, L., et al. (2024). Establishing a digital technology play space for children. *Journal of Museum Education, 49*(4), 525-537.
[4] Vartiainen, J., Sormunen, K., & Kangas, J. (2024). Relationality of play and playfulness in early childhood sustainability education. *Learning and Instruction, 93*, 101963.
[5] Tang, K.-S., et al. (2024). Exploring the multimodal affordances of digital coding devices in fostering creative thinking in early childhood education. *Thinking Skills and Creativity, 53*, 101602.
[6] Leung, S. K. Y., & Choi, K. W. Y. (2024). Teachers' perceptions of technical affordances in early visual arts education. *European Early Childhood Education Research Journal, 32*(1), 147-166.
[7] Ocak, C., Kopcha, T. J., & Dey, R. (2023). An AI-enhanced pattern recognition approach to temporal and spatial analysis of children's embodied interactions. *Computers and Education: Artificial Intelligence, 5*, 100146.
[8] Ouhaichi, H., Spikol, D., & Vogel, B. (2023). Research trends in multimodal learning analytics: A systematic mapping study. *Computers and Education: Artificial Intelligence, 4*, 100136.
[9] Caskurlu, S., Ocak, C., & Dai, C.-P. (2025). The scope of multimodal learning analytics in K-8: A systematic review. *Journal of Learning Analytics, 12*(2), 224-236.
[10] Mohammadi, M., et al. (2025). Artificial intelligence in multimodal learning analytics: A systematic literature review. *Computers and Education: Artificial Intelligence, 8*, 100426.
[11] Alfredo, R., et al. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence, 6*, 100215.