

AIGC-Supported Intelligent Collection and Application Research of Classroom Teaching Behaviors

Zhenyue Sun¹, Maoyue Zhang², Yun Cheng^{3,*}, Ming Yu⁴

¹Huanggang Normal University, Huanggang, China

²Dezhou No.2 Middle School, Dezhou, China

³Huanggang Normal University, Huanggang, China

⁴Huanggang Normal University, Huanggang, China

*Corresponding author: yunccnu@126.com126.com

Abstract: The analysis of teaching behaviors is of great significance for teaching diagnosis and quality improvement. It can also serve as a basis for assisting teachers in reflection and evidence-based teaching. However, most of the previous automated research methods have used deep learning models to train on specific samples. The coding categories that can be automatically analyzed are single, which cannot be applied to different coding systems required by complex teaching scenarios. Nor can the coding be modified according to the needs of the scenarios or multi-coding analysis be carried out. Based on the review of the methods for analyzing classroom teaching behaviors based on artificial intelligence, this study proposes an automated collection process of teaching behaviors, which is "determining the analysis objectives—selecting the coding system—splicing prompts", and trains a large model in the vertical domain to judge the feasibility of AIGC (Artificial Intelligence Generated Content) in recognizing teaching behaviors.

Keywords: Classroom Analysis; AIGC; Collection of Teaching Behaviors; Classroom Teaching

1. Introduction

Promoting high-quality development in the classroom under the new era background is both the due meaning of promoting the construction of a high-quality system and the practical requirement for solving teaching problems and improving educational quality ^[1]. The core of high-quality educational development lies in the classroom. Analyzing and mining the typical characteristics of high-quality classrooms based on classroom behavior, and optimizing existing classroom teaching based on big data evidence-based methods are effective ways to achieve high-quality classrooms. Previous studies on the realization of classroom behavior analysis are mostly data-intensive and labor-intensive projects, with high costs and inability to flexibly adjust the coding architecture. AIGC can complete data denoising tasks through autoencoders, achieve data optimization, further enhance the accuracy of data analysis, improve the effectiveness of data processing, and provide technical support for large-scale automated collection of classroom behaviors. On this basis, by means of text analysis supported by AIGC, we can explore the learning mechanisms of individual and group learners, and deeply integrate the advantages of quantitative research in natural sciences into educational research activities ^[2]. Based on the possibility of AIGC bringing new paradigms to educational research, this study attempts to explore the feasibility of AIGC in automatically collecting classroom teaching behaviors, and addresses the shortcoming of output instability of large models by using regularization for filtering to enable output according to usage requirements.

2. Overview of Classroom Teaching Behavior Analysis

2.1 Overview of Teaching Behavior Coding

The theory and practice of analyzing classroom teaching based on coding scales are relatively mature domestically and internationally. Typical classroom teaching analysis methods include the Flanders Interaction Analysis System (FIAS), the Verbal Interaction Category System (VICS), the Information Technology-based Interaction Analysis System (ITIAS), the Student-Teacher (S-T)

analysis method, etc. [3]. These codings can comprehensively reflect the main behaviors of teachers and students in the classroom, but the granularity is relatively coarse. Subsequent research has gradually begun to focus on students' cognition and higher-order thinking levels. For example, Song Yu et al. proposed a classroom dialogue coding system oriented to knowledge construction based on the Scheme for Educational Dialogue Analysis created by the Cambridge University team and the Knowledge Building System proposed by the University of Hong Kong team, endowing discourse categories with the connotation of knowledge construction [4]. Ma Ruxia et al. adopted the cognitive process dimension in Bloom's Taxonomy of Educational Objectives (2001 edition) as the coding framework for classroom dialogue, dividing the cognitive goals of learning into six major categories: memory, understanding, application, analysis, evaluation, and creation, with the first three categories regarded as low cognitive levels and the last three as high cognitive levels, each containing multiple subclasses [5]. Hu Ju et al. combined Zhang Delu's comprehensive framework of multimodal discourse analysis and the framework for dividing teaching stages in the teaching procedure, refining the types of modalities according to the characteristics of music classes, and forming an analysis framework for music classroom teaching behavior that includes 18 modalities [6]. For example, Qian Wenjing reconstructed a dual-coding analysis system for classroom interaction suitable for Chinese subjects because the interaction analysis system of FIAS is insufficient to explain the transition relationships of classroom teaching behaviors [7]. The coding system for classroom behavior analysis also needs to be able to adapt to changes in teaching scenarios and teachers' needs. Therefore, there is an urgent need for an automatic collection method for teaching behaviors that can be applied to different coding systems.

2.2 Overview of Automatic Collection of Teaching Behaviors

Although classroom behavior analysis is gradually becoming automated, multiple studies have contributed to this development: Liu Qingtang et al. [8]. Applied intelligent technology to automatically collect and code teaching process data, followed by analytical and visual representation. The YOWO model, adapted for classroom student behavior recognition, was verified to achieve high accuracy in identifying students' learning behaviors from video recordings [9]. Software-based automatic speech recognition of teachers' near-field voice enabled precise extraction of speaking periods, supporting automated S-T analysis [10]. A deep learning method incorporating human body skeleton features was developed to recognize students' classroom behaviors by extracting key skeletal information from images [11]. Due to being trained on datasets with manually labeled specific categories, previous automated analysis methods can generally only analyze single-category codings, cannot perform multi-category coding analysis, and cannot meet teachers' needs to modify the coding system.

3. Construction of Intelligent Collection Model for Teachers' Classroom Teaching Behaviors

3.1 Fine-Tuning the Base Large Model

This paper uses a classroom teaching behavior coding dataset to fine-tune the Large Language Models (LLM), training a vertical domain large model suitable for classroom teaching behavior coding tasks to improve its performance in this specific domain. The model demonstrates more reliable coding knowledge responses. Generally, large models have already acquired a certain level of behavior understanding ability during pre-training. Through instruction-supervised fine-tuning, these abilities can be optimized for specific classroom behavior coding tasks in the target domain. Instruction-supervised fine-tuning is a simple, direct, and effective solution. During the fine-tuning process, this method introduces explicit coding instructions, guiding the large model to perform specific coding tasks in the target domain. The large model generates corresponding coding outputs based on the instructions and the input content to be coded, and optimizes training by calculating the cross-entropy loss between the generated coding outputs and the actual coding.

The experimental process of this paper is shown in Figure 1, using the Qwen1.5-7B-Chat model as the base model and fine-tuning the vertical domain large model, significantly improving automatic coding performance. To save training costs, this paper adopts LoRA for fine-tuning, which freezes the pre-trained model parameters and uses low-rank matrix representation for parameter updates, greatly reducing the computational volume of fine-tuning and not increasing additional inference costs, making it a popular method for fine-tuning large models.

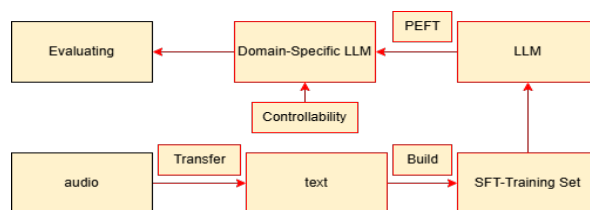


Figure 1: Fine-Tuning the Base Large Model for Teaching Behavior Collection

This paper constructs a vertical large model suitable for the field of classroom behavior analysis, and realizes the processes of automatic transcription, segmentation, and splicing of teaching audio through automated Python scripts, sending prompts to the large model, obtaining and visualizing the return results, and being able to automatically collect teaching behaviors in any batch is the prerequisite for mining classroom rules.

3.2 Controllable Application Workflow for Large-Scale Models

In this paper, a vertical large model trained on manually labeled categories is used for automatic collection of classroom behaviors. In terms of controllability, regular expression filtering and extraction are adopted to further extract valid information from the results returned by the large model, for example, only needing to obtain the final coding number by filtering out text using regular expressions and obtaining the return result. In terms of application, simply replace the required coding table according to the prompt words, as shown in Figure 2.

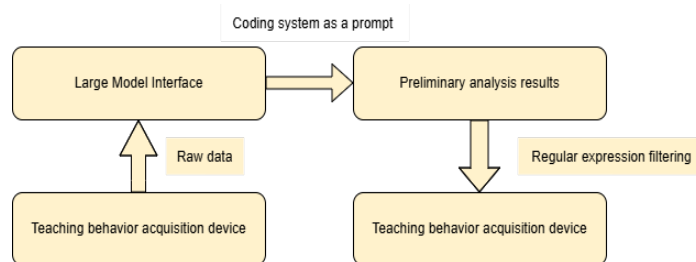


Figure 2: Application Process of the Large Model

4. Collection and Analysis of Teaching Behavior Data

4.1 Selection of Coding System

The basic ideas of related research on the classification of teaching behaviors mainly originate from four types of coding systems: the Flanders Interaction Analysis System^[12], the Verbal Interaction Category System^[13], the Information Technology-based Interaction Analysis System (ITIAS)^[3], and S-T analysis^[14], which provide important guidance and reference for research on teaching behavior analysis. Previous classification systems for teaching behaviors have the following issues: incompleteness of behavior classification, a trend of "emphasizing teaching over learning," coarseness of behavior division, fuzziness of behavior division. Based on a comparative analysis of four types of classroom teaching behavior coding systems, Cheng Yun et al. constructed a new classroom teaching behavior analysis coding system from the two dimensions of behavior mode and behavior subject, dividing classroom teaching behaviors into four major categories and 16 subcategories^[15]. This coding system basically covers all categories of teaching behaviors, so this study selects this coding framework for research, which can be shown in Table 1.

Table 1: Classroom Teaching Behavior Coding Table

Behavior Type	Behavior Description	Behavior Subject	Coding	Behavior Description
Verbal Behavior	Lecturing	Teacher	1	Offering factual information or insights on the content or steps, expressing one's own opinions as a teacher, presenting personal explanations, or quoting the viewpoints of an authority (not students).
Verbal Behavior	Questioning	Teacher	2	Asking students questions based on one's own opinions or thoughts as a teacher, and anticipating

				their responses; often requiring students to retain certain facts or guiding them to imagine and analyze.
Verbal Behavior	Instruction	Teacher	3	Directing or commanding students to perform certain actions, expecting them to comply, accept the teacher's will, or attempt to modify their behavior (such as through criticism); guiding students in engaging in learning activities.
Verbal Behavior	Feedback and Evaluation	Student	4	Concluding the conversation with a definitive answer as a teacher; for students' responses, encouraging them to further ponder or discuss, develop or supplement their opinions or thoughts; evaluating the learning content, students' viewpoints, or the effectiveness of their learning.
Verbal Behavior	Initiating Questions	Student	5	Students inquiring with the teacher about what they should do and how to do it; presenting their own questions to the teacher, anticipating solutions.
Verbal Behavior	Responding	Student	6	Students providing responses to the teacher's questions. For closed-ended questions with a single correct answer, the teacher typically requires students to remember certain facts; for open-ended questions with multiple possible answers, the teacher often prompts students to imagine or analyze.
Verbal Behavior	Dialogue	Student	7	Students actively initiating conversations, expressing their own thoughts; introducing new topics; freely elaborating on their opinions and ideas; transcending existing knowledge frameworks.
Verbal Behavior	Discussion	Student	8	Students engaging in discussions with each other, freely exchanging viewpoints.
Activity Behavior	Observation	Student	9	Students actively observing the activities of the teacher or classmates, or independently carrying out learning activities by observing multimedia and other information.
Activity Behavior	Note-taking or Practice	Student	10	Students taking notes in notebooks or textbooks, or engaging in in-class exercises.
Activity Behavior	Practice or Experimentation	Student	11	Teachers organizing practical activities and teaching experiments such as games, singing, and dancing, with student participation; students independently operating computers and other tools to engage in learning activities.
Activity Behavior	Reflection	Student	12	Students contemplating questions; without verbal or overt activity.
Activity Behavior	Whiteboard Writing	Teacher	13	Displaying teaching content through methods such as copying and calculating on blackboards, whiteboards, and other tools.
Activity Behavior	Demonstration or Presentation	Teacher	14	Teachers using physical objects, models for demonstrations, or multimedia for presentations.
Activity Behavior	Observation / Patrolling	Teacher	15	Teachers monitoring or observing students' learning progress.
Activity Behavior	Individual Guidance or Participation in Activities	Teacher	16	Teachers and students interacting through questions and answers, requests and responses, evaluations and feedback; during students' practice and engagement in experimental activities, teachers providing real-time guidance for any issues or difficulties encountered by students.

4.2 Preparation of Dataset

This study analyzes six "One Teacher, One Excellent Lesson" samples from the National Smart Education Platform, which feature comprehensive observational data on classroom teaching behaviors. Based on a coding framework for classroom teaching behaviors, this research employs manual coding methods combined with instructional videos to analyze the transcribed structured text. To ensure coding reliability, the two researchers involved underwent stages of coding training, pre-coding, and coding consistency checks. The internal consistency of the coding was high ($Kappa = 0.95$, $p < 0.01$), and inconsistent content was reviewed and revised to a unified code. The results of this coding were used as standard data for training the automatic collection model and as benchmark data for comparing

the accuracy of AI-based automatic collection.

4.3 Intelligent Collection and Application Pathway

The development of intelligent and accompanying collection tools has made it possible to collect classroom teaching behaviors in a multidimensional, comprehensive manner. This paper proposes a pathway as illustrated in Figure 1 for intelligent collection and application of teaching behaviors based on AI-Generated Content (AIGC). After obtaining the most basic data at the collection layer, the data is processed using AIGC, such as transforming raw data into meaningful "information" through feature extraction. At the collection layer, it is necessary to integrate theories related to classroom behavior analysis with AIGC for collaborative collection. For example, AI can be trained to learn different coding systems through "cue words" and training, enabling it to distinguish the coding categories of different information. At the analysis layer, feature analysis is conducted on the data subjects based on the data sources, thereby quantifying the characteristics of the subjects. At the application layer, these representations are used to meet the specific needs of teachers, students, administrative departments, etc., and the classroom behavior analysis theory is continuously optimized and refined through feedback during use. Machine intelligence and human intelligence form a clear division of labor. Machine intelligence excels in data processing, so AI's advantages are leveraged at the collection layer. Humans excel in abstraction and holistic perception, so AI-based data is used to refine theories at the application and theoretical levels. The role of teaching behavior analysis is leveraged through the pathway of "AI intelligent collection and subject-based evidence-based application." (As shown in Figure 3)

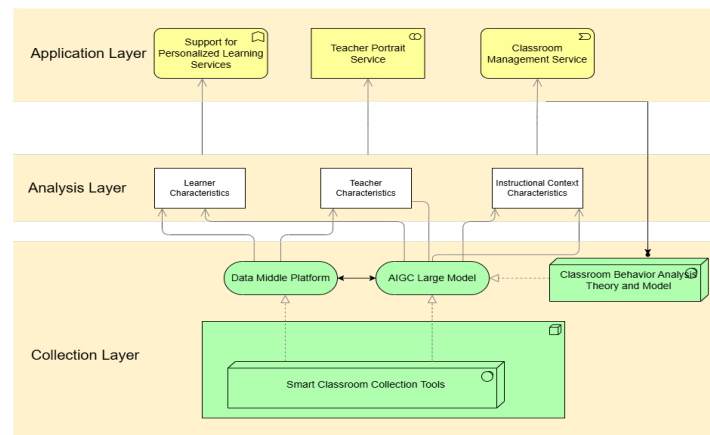


Figure 3: Intelligent Collection and Application Pathway of Teaching Behaviors

5. Application and Analysis

5.1 SFT Dataset

In this paper, the aforementioned coded dataset of classroom teaching behaviors is randomly divided into a training set and a test set in a 9:1 ratio. For the training set, it is constructed as a Supervised Finetuning (SFT) dataset. This dataset comprises a total of 426 meticulously crafted finetuning instructions, containing samples of various codes. It consists of (instruction, output) pairs, where the instruction represents the human requirement for the model's task, and the output represents the expected output following the instruction, used to optimize the large model's output to align with human-expected instruction compliance results.

5.2 Experimental Settings

In the finetuning experiments, the settings are as follows: the learning rate is set to $5e-5$, the training batch size per device is 1, the warm-up ratio is 0.1, and the number of training epochs is 5. To efficiently train the model, LoRA is used for finetuning, with LoRA's rank set to 8. A series of detailed experiments are conducted to validate the effectiveness of the vertical domain large model.

5.3 Result Analysis

To verify the validity of the vertical domain large model, this paper compares the accuracy of classroom behavior coding between the base large model and manual coding. The metrics include accuracy, recall, precision, and F1 score, as shown in Table 2.

Table 2 The automatic classification results of different methods.

Model Method	Accuracy %	Recall %	Precision %	F1 Score %
Base Large Model	0.6113	0.889	0.604	0.724
Manual Coding	0.901	0.963	0.902	0.932
Vertical Domain Large Model	0.689	0.765	0.725	0.745

After practical testing, the accuracy of the large model after fine-tuning training has slightly improved, but the improvement is not significant. This may be due to the insufficient amount of manually coded data, as well as the limitation of the parameter count for all large models discussed in this paper. In terms of accuracy, the model still cannot compete with human coding.

5.4 Differential Analysis

For the inconsistent parts of the coding, we analyzed the differences between AI coding and manual coding in each category, examining possible reasons for the discrepancies. The categories are ranked from highest to lowest based on the percentage of differences in data collection for each coding category, as shown in Table 3. Categories with a percentage below 2.38% are not listed.

Table 3: Comparison of Differences in Teaching Behavior Collection between Human and AI

Coding Category and Meaning	Manual Coding	AIGC Coding	Proportion
14 Demonstration or Presentation	14	1	23.81%
14 Demonstration or Presentation	14	3	19.05%
1 Lecturing	1	3	11.90%
2 Questioning	2	1	9.52%
2 Questioning	2	3	9.52%
4 Feedback and Evaluation	4	1	9.52%
3 Instruction	3	2	4.76%
14 Demonstration or Presentation	14	16	2.38%

Based on the analysis of the table, it can be observed that the primary category of misidentification by AI is category 14, accounting for nearly half of the errors. The main reason for this is speculated to be that category 14 involves teachers' demonstrations, which require video modal support for accurate identification. However, this study did not utilize a multimodal large model or incorporate video modalities for data collection. Secondly, AI frequently confuses category 2, which involves questioning, and misclassifies it as lecturing or instruction by teachers. This type of coding requires human cognition and contextual analysis, which large models are not adept at. The differences in other coding categories are relatively low. If multimodal methods are integrated into the collection process, the accuracy of AIGC in teaching behavior collection should improve by at least half or more. Alternatively, if the coding system selected for practical application relies solely on verbal or textual modalities, which can be readily achieved, the accuracy of AIGC should also increase.

5.5 Visualization Analysis of Coding Results

A 15-minute micro-lecture sample from the experimental dataset was selected to compare the time-series of manual coding results and AI coding results. As shown in Figure 4, the overlapping sections remain largely consistent. There are slight differences in category 14 coding within the first four minutes, and in the final few minutes, AI tends to code more as category 1 while manual coding is category 4. The middle section remains largely consistent.

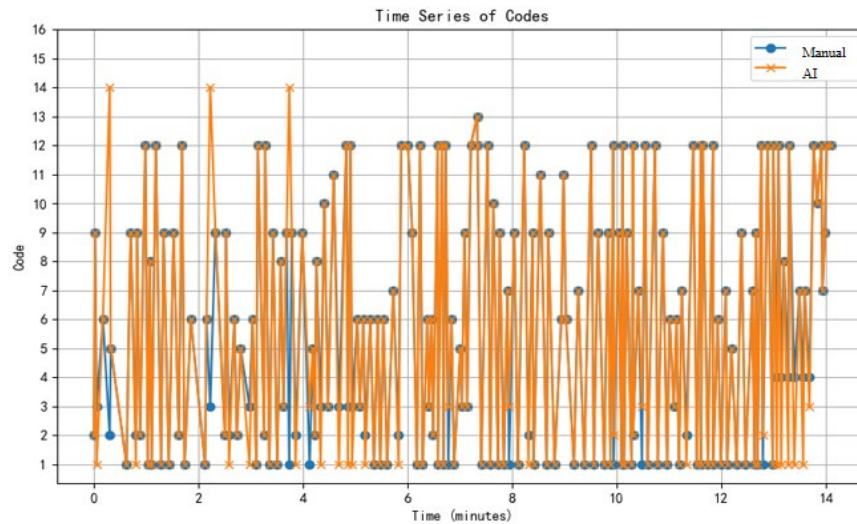


Figure 4: Comparison Chart of Coding Time Series

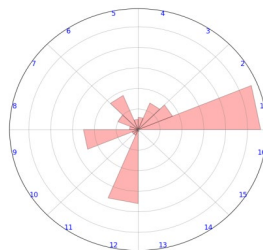


Figure 5: Cloud Model Diagram of Manual Coding

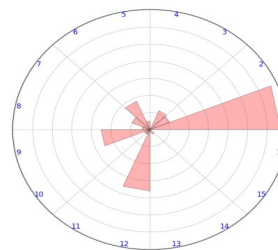


Figure 6: Cloud Model Diagram of AI Coding

The cloud models for the coding results are depicted in Figures 5 and 6. The AI coding exhibits a very small cloudlet in coding category 14, followed by slightly smaller cloudlets in coding categories 2 and 3 compared to manual coding. The overall cloud model parameters are shown in Table 4.

Table 4: Differences in Cloud Model Parameters between AI and Manual Coding

	Manual Coding	AI Coding
Perimeter	365.2	391.2
Total Area	1159.9	1371.5
Centroid	(16.1, -0.4)	(24.4,1.5)

From Table 4, it can be observed that there are insignificant differences between AI coding and manual coding in terms of the cloud model parameters obtained based on coding categories. The centroids of the cloud models reflect the emphases of the teaching modes throughout the entire class, and the centroids of both are relatively close in coordinates. Therefore, a preliminary conclusion can be drawn: AI coding and manual coding exhibit slight differences in visualization results and parameters of the visualization model, but these do not affect the overall outcomes.

6. Summary and Outlook

This study explored the application of AIGC in automatic collection of classroom teaching behaviors, constructed an intelligent collection model for teaching behaviors based on AIGC, and verified its feasibility under a specific coding system through experiments. The research results indicate that the fine-tuned large vertical model demonstrates certain accuracy in the task of classroom teaching behavior coding, but further optimization is still required. Future research directions can include: 1) Multimodal Data Fusion: The current study primarily analyzes teaching behaviors based on text data. In the future, it can explore the integration of multimodal data such as video and audio into AIGC models to improve recognition accuracy and comprehensiveness. 2) Highly Personalized and Automated Coding Systems: Through the AIGC automatic collection model for multimodal data,

highly personalized and automated coding systems can be realized to meet the needs of different disciplines and teaching scenarios. In terms of model interpretability, the current AIGC models lack transparency in their decision-making processes. In the future, research can explore ways to improve model interpretability so that users can better understand the working principles and results of the models.

This study preliminarily explored the application of AIGC in the analysis of teaching behaviors and can further expand its application scenarios in the future, such as teaching diagnosis, teaching evaluation, and teaching improvement. This study provides new ideas and methods for the analysis of classroom teaching behaviors and technical support for improving classroom teaching quality. The application of AIGC technology will help promote the informatization and intelligentization of education, facilitating educational equity and quality improvement. AIGC will provide more intelligent, efficient, and personalized solutions for the analysis of classroom teaching behaviors, contributing to the construction of a smart education ecosystem.

Acknowledgement

This project is supported by This research was supported by the National Natural Science Foundation of China (No. 71974073), The Education Science Planning Major Projects of Huanggang (No. 2023ZD02, 2023ZD04), Collaborative Education Project of the Ministry of Education (No. 202102085011), the key Project of Philosophy and Social Science Research of Hubei Provincial Education Department (Project No.24D067), the High-level Cultivation Program Project of Huanggang Normal University in 2024 (No. 202420804), the "Four New" Research and Reform Practice Project of Huanggang Normal University (No. 2022CN01), and The Graduate Research Station of Huanggang Normal University(NO.5032024001, 5032024002) .

References

- [1] Xie Youru, Xia Wan, Qiu Yi, et al. *Empowering High-quality Classrooms through Digital Transformation: Logical Essentials, Practical Dimensions and Typical Patterns*[J]. *China Educational Technology*, 2023(9): 50-58.
- [2] Liu Sanya, Hao Xiaohan, Li Qing. *A New Paradigm of Educational Research: Educational Science Research Driven by Artificial Intelligence*[J]. *Educational Research*, 2024, 45(3): 147-159.
- [3] Gu Xiaoqing, Wang Wei. *New Exploration of Classroom Analysis Technology to Support Teachers' Professional Development*[J]. *China Educational Technology*, 2004(7): 18-21.
- [4] Song Yu, Wu Baoxian, Hao Tianyong. *An Analysis of the Laws of Classroom Dialogue Oriented to Knowledge Construction*[J/OL]. *E-Education Research*, 2021, 42(3): 111-119. DOI:10.13811/j.cnki.eer.2021.03.016.
- [5] Ma Ruxia, Wang Lu, Peng Le. *Knowledge Discovery from Big Data: Strategies for Promoting Deep Learning in the Classroom*[J/OL]. *E-Education Research*, 2022, 43(5): 84-91, 108. DOI:10.13811/j.cnki.eer.2022.05.012.
- [6] Hu Ju, Lu Tian, Zhao Jian. *Analysis of Multimodal Teaching Behaviors in the Cloud Music Classroom of Student Teachers' Internship and Teaching Support*[J/OL]. *E-Education Research*, 2022, 43(09): 114-119+128. DOI:10.13811/j.cnki.eer.2022.09.014.
- [7] Qian Wenjing, Li Hongmei. *Development of a Dual Coding Analysis System for Chinese Classroom Interaction in a Smart Classroom Environment and Mining of Teaching Behavior Patterns*[J/OL]. *China Modern Educational Equipment*, 2023(16): 63-66. DOI:10.13492/j.cnki.cmee.2023.16.002.
- [8] Liu Qingtang, He Haoyi, Wu Linjing, et al. *Methods and Applications of Classroom Teaching Behavior Analysis Based on Artificial Intelligence*[J]. *China Educational Technology*, 2019(9): 13-21.
- [9] Xie Wei, Tao Yaping, Gao Jie, et al. *Real-time Recognition of Classroom Learning Behaviors Based on YOWO*[J]. *Modern Educational Technology*, 2022, 32(6): 107-114.
- [10] Luo Zuying, Zhao Qiqi, Duan Fuqing. *Automatic Analysis of Classroom Teaching Process Based on Teachers' Near-field Speech*[J/OL]. *Modern Educational Technology*, 2021, 31(08): 76-84. DOI:CNKI:SUN:XJJS.0.2021-08-010.
- [11] He Xiuling, Yang Fan, Chen Zengzhao, et al. *Recognition of Students' Classroom Behaviors Based on Human Skeleton and Deep Learning*[J]. *Modern Educational Technology*, 2020, 30(11): 105-112.
- [12] Ding Ying, Yang Kaicheng. *A Comparative Study of Teaching Analysis Methods*[J]. *Modern Educational Technology*, 2012, 22(9): 12-17.

- [13] Yang Pingzhan, Liu Juanwen, Luo Ping. *Discussion on the Classroom Interaction Coding System (VICS)[J]*. *Teacher*, 2009(6): 3-4.
- [14] Fu Derong, Zhang Huimin. *Educational Information Processing[M]*. *Educational Information Processing*, 2021.
- [15] Wang Yanli, Cheng Yun, Wang Feng, Tong Sanhong, Huang Kebin. *Exploration of Observation Methods of Classroom Teaching Behaviors Supported by Technology[J/OL]*. *Modern Educational Technology*, 2016, 26(09): 39-45. DOI: CNKI:SUN:XJJS.0.2016-09-007.