A Model for Analyzing the Behavior of Classroom Teacher-Student Interaction Based on Deep Learning

Mingyue Zhou[a], Wei Pan*, Zelin Zhang[b]

Capital Normal University, Beijing,100048, China

[a]2211002049@cnu.edu.cn
*Corresponding author:5263@cnu.edu.cn
[b]5263@cnu.edu.cn

Abstract. In the context of combining education and artificial intelligence, it is important to identify and analyze interactive behaviors in intelligent classroom behaviors. The main contribution of the article is the design and implementation of an intelligent analysis model of classroom teacher-student interaction behavior based on deep learning. In this study, we provide a method for encoding common teacher-student interactional behaviors, and we employ the YOLOv8 deep learning network to recognize these behaviors and to perform temporal and statistical analyses of the data: classroom videos are classified and analyzed for teacher-individual interaction, teacher-group interaction, teacher-class interaction, cross interaction and student-student interactions. The experimental results show that the detection results of this model can basically cover the detection results of the manual observation method with an average difference rate of less than 5%, which is of practical value for classroom teaching evaluation and teacher-student interaction behavior assessment.

Keywords: educational analytics; video analytics; deep learning; intelligent analytics

1 Introduction

Analysis of teacher-student interaction behavior is a crucial metric for assessing the effectiveness of instruction in the classroom since it can give teachers insight into how well their charges engage in their lessons, foster better classroom inter-action, and improve instruction overall. It is difficult to gather objective evaluation data, particularly information about teacher-student interaction, and peer or supervisor auditing is the primary method used to evaluate the quality of traditional classroom instruction.

With the quick advancement of artificial intelligence technology, it is now possible to evaluate classroom instruction quality intelligently [1]. The applicability and gen-

G. Guan et al. (eds.), *Proceedings of the 2023 3rd International Conference on Education, Information Management and Service Science (EIMSS 2023)*, Atlantis Highlights in Computer Sciences 16,
https://doi.org/10.2991/978-94-6463-264-4_27

eralizability of study findings are, however, constrained because many present studies rely on smart classrooms and the issues of research and discussion are very narrow, focused either on teacher conduct or student behavior. In recent years, intelligent class-room teacher-student interaction behavior analysis has become a popular rsearch direction, and there are some research materials and cases in this regard: A study of student classroom behavior based on deep learning[2], extracting key infmation of human skeleton to improve the accuracy of identifying student behavior; based on face detection and gaze estimation techniques[3], recording students' concentration on classroom; through multimodal techniques combining sound and image to compute the motivation of evaluating teacher interactions[4]; based on speech recognition techniques[5], S-T (student-teacher) behavior coding is proposed to classify different types of teaching and interaction behaviors and to compute the time of interaction behaviors. Similarly, more and more studies have started to use multiple techniques for classroom behavior analysis, but most of them tend to consider only student or teacher subjects, and lack more in-depth intelligent analysis in the field of teacher-student interaction behavior analysis oriented.

The Ministry of Education's 2022 Compulsory Education Syllabus and Curriculum Standards highlighted the significance of observing, documenting, and analyzing the classroom learning process [6] and the necessity of actively enhancing dialogue and communication by concentrating on teachers' and students' typical behavioral performance. In order to better achieve this goal, a code for various teacher-student interactions in classrooms is proposed [7]. The code can be di-vided into five categories: teacher-student interaction, teacher-individual inter-action, teacher-class interaction, cross-interaction, and student-student interaction. This coding system, however, is only appropriate for conventional manual statistical analysis techniques; intra-intelligent analysis and identification, which demand more precise targets and data, do not. The ability of the model to recognize and categorize student-teacher interactions may be constrained by the inadequacy of simple coding techniques to capture complex human interaction data. In order to get over these restrictions and enhance the models' capacity to recognize and categorize student-teacher interaction patterns, intelligent analysis and recognition must combine more advanced coding techniques with cutting-edge deep learning models and training on enormous datasets.

In conclusion, the article explores how to utilize behavior analysis to analyze teacher-student interactions in elementary and secondary school classroom teaching videos while taking into account real-world requirements. The teacher-student interaction behavior classification theory [8], which is currently widely accepted in the educational community, is used as the foundation for coding teacher and student behaviors separately. Next, deep learning networks are used to detect and identify teacher-student behaviors, and the data related to each teacher-student interaction are statistically calculated based on the pertinent theory. Finally, the temporal and statistical perspectives on the behavior of teacher-student interaction are shown for the entire class. The results of the experiments demonstrate that the teacher-student interaction behavior analysis can adequately cover the outcomes of the manual observation detection method with an average difference rate of less than 5%, demonstrating its applicability in the assessment of teacher-student interaction behavior and intelligent classroom teaching.

2 The construction of intelligent analysis model of teacher-student interaction behavior

Teacher-student interaction behaviors have a significant impact on the effectiveness of instruction and student learning outcomes. The article's primary research focuses on using classroom video recordings to automatically identify teacher-student interaction behaviors, visualize and analyze these interactions once they have been identified, and then give teachers reference data so they can enhance their teaching and develop better teaching strategies. First, the processing of instructional videos is set up for post-processing, which includes annotating movies into frame pictures and datasets for image annotation. Second, the teacher and student activities in the movies were identified using the YOLOv8 model, which is capable of automatically distinguishing interactive behaviors. The teacher-student interaction behaviors were then categorized in accordance with the observed interaction behaviors, and matching visual reference data, including the incidence and duration of the interaction, the identities of the participants, and the type of interaction, were provided.

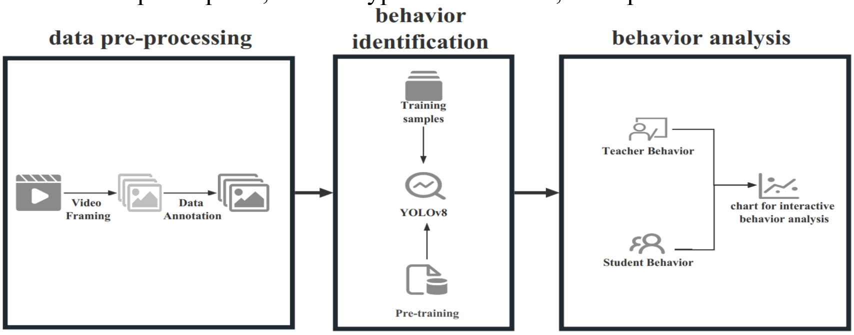


Fig. 1. Diagram of an intelligent analysis model for classroom interaction behavior

Figure 1 illustrates the division of the model analysis framework into three sections: data pre-processing, behavior identification, and behavior analysis.

2.1 Data pre-processing

The video clips are first subjected to data pre-processing before the deep learning model is trained. After the video has been frame-separated, which involves taking one frame per second, it is converted into an image file. Next, data annotation is carried out, which includes details on the interaction style, location, and number of all teachers and students present in the classroom.

2.2 Behavioral identification

To assist teachers and students in evaluating and improving the teaching process in order to determine the effectiveness of the interaction between teachers and students, the types of interaction behaviors are determined based on the teacher-student behaviors detected through YOLOv8. The types of interaction behaviors are coded based on the types of interaction behaviors suitable for deep learning, and the classification results are quantified and analyzed.

2.3 Behavioral analysis

The teacher-student behavior detected by YOLOv8 is used to determine the type of interaction behavior based on the coding of the type of interaction behavior suitable for deep learning. The classification results are quantitatively analyzed to obtain visualization data, which is convenient for teachers and students to evaluate and improve the teaching process, as well as to assess the effectiveness of the interaction between teachers and students.

3 The design of intelligent analysis model of classroom interaction behavior

3.1 Interactive behavior classification

The definition of classroom interaction subject type codes is proposed and summarized into the following main typical interaction behaviors, which lead to the classification of interaction types:

Teacher-individual interactions (TI interactions): teacher-student interactions in which the teacher directs their behavior toward particular students, typically take the form of one-on-one tutoring, question-and-answer sessions, one-on-one guidance, and feedback. For example, a typical teacher-individual interaction might involve the teacher helping students with their homework.

Teacher-group interaction (TG interactions): teacher-student interaction in which the teacher's behavior is directed at groups of students. This behavior typically takes the form of the teacher explaining to, tutoring, and evaluating student groups; for example, a typical teacher-group interaction would be the teacher's discussion with a student group.

Teacher-class interaction (TC interactions): teacher's behavior is intended to promote interaction among all students in the class. This is typically demonstrated by the teacher planning the lesson, delivering the lecture, and assessing the entire class. For example, a typical teacher-class interaction might involve the teacher lecturing and the students paying attention to the lecture. Formatting: Insert one hard return immediately after the last character of the last affiliation line. Then paste down the copy of affiliation 1. Repeat as necessary for each additional affiliation.

Cross interaction (TR interactions): a teacher-student contact in which the teacher exhibits conduct that is intended for both the student body as a whole and for particular pupils. An example of a typical cross-interaction is when a teacher stands in front of one student while explaining to another student or inviting all students to send questions.

Student-student interaction (SS interactions): Interactions with other students, such as group discussions and independent writing tasks, are interactions with individuals or groups of students.

The article proposes a coding list of interaction types based on the actual classroom situation and the adaptation of the intelligent analysis model for classroom interaction behaviors, as shown in table 1. The teacher-student interaction behavior types can be divided and quantitatively analyzed, the classroom interaction types can be analyzed in time sequence, and the percentage of each type of interaction in the total interaction can be counted to show the interaction process of a real classroom. The behavior type code applicable to deep learning can be designed according to the two subject types, teacher and students.

Table 1. Classroom interaction type coding table

Type Of Interaction	**Code Interpretation**	**Interaction behavior**	
		Teacher	***Student***
TI interaction	The teacher's actions point to individual students.	guide	write
TG interaction	The teacher's behavior is directed at student-teacher interaction in small groups.	guide	discuss
TC interaction	the teacher's behavior is directed at the whole class.	teach	
CR interaction	the teacher's behavior is directed at individual students and groups of students.	ask, invite, teach	hand up, stand
SS interaction	Student behavior is directed at individual student or group student interaction.		discuss

3.2 Interactive behavior coding

The normative criteria for behavior coding are shown in table 2, including four interactive behaviors of students hand up, standing, discussing, and writing, and four interactive behaviors of the teacher inviting, asking questions, guiding, and teaching, for a cumulative total of eight interactive behaviors.

Table 2. Interactive behavior code description table

Interactive Roles	**Interaction behavior**	**Action state**
Teacher	Ask	Facing all students + raising hands to demonstrate the action

Interactive Roles	**Interaction behavior**	**Action state**
	Invite	Facing all students + pointing fingers at students
	Teach	Facing all students + finger towards the board/screen
	Guide	Facing individual students + hand pointing to the table
Student	Stand	Students stand up to answer questions
	Hand up	Students raise their hands to answer questions
	Write	Students look down + pen in hand
	Discuss	Students gather in small groups

3.3 YOLOv8 model.

The YOLOv8 network structure diagram is illustrated in figure 2 and is primarily separated into two parts: feature extraction network and detection network. The behavior detection portion of the model chooses YOLOv8 as the deep learning behavior recognition network.

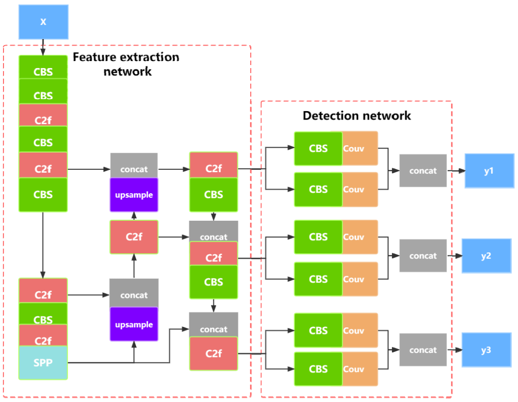


Fig. 2. Diagram of the YOLOv8 structure

Feature extraction network: Convolutional neural networks are used to extract features from images. Specifically, through multiple convolution, pooling and normalization operations, integrated into CBS module and C2f module, feature images with different number of channels are concatenated as well as upsampled to transform the input image x into a feature map for multiple scales, capable of extracting features at different scales of student and teacher behavior targets under real classrooms.

Detection network: the feature maps are used as inputs and the predicted target frames, categories and confidence levels are output. Specifically, the location detection and classification detection are separated and then connected, and the feature maps are subjected to convolution and pooling operations, i.e., the CBS module and the Conv module, respectively, to output information on the location and size of the target box at each location, as well as the probability of the category to which the target box at each location belongs. Splitting the detection task into multiple subtasks can alleviate the box position and category errors due to the serious front-to-back occlusion problem among students, and improve the efficiency and accuracy of detection.

Overall, YOLO v8's network structure allows it to perform well in terms of detection effect and running speed, particularly for teacher-student interaction behavior recognition and detection, and real classrooms with many targets, different sizes, and serious obscuration problems can be detected quickly.

4 Experiments and analysis

4.1 Data set

The four interaction behaviors of teacher inviting, asking, teaching, and guiding as well as the four interactive behaviors of students standing, discussing, and writing in class are all included in the created dataset. To increase the model's precision, the dataset needs to be sufficiently large and inclusive of normal instructional interaction behaviors. The videos in this collection were created from 14 authentic classroom videos, each of which has a 40-minute video of a lesson being taught. The dataset has a total of 10047 photos after being divided and cleaned by taking one frame every three seconds, and the training set has a total of 8037 images after dividing the validation set and test set by 8:1:1.

4.2 Experiments in training

The hardware environment of this experiment is CPU Intel(R) Xeon(R) Platinum 8255C, GPU RTX 2080 Ti (11GB), 64GB memory, and PyTorch framework.

Results of all student actions detected by the YOLOv8 algorithm are shown in Table 3. In the article model, this methodology was applied. All forms of interaction behaviors are more accurately identified when compared to the somewhat less accurate detection of the teacher's invitation to answer, which occurs mostly because the variations between the teacher's invitation to answer actions in the real classroom are smaller. Furthermore, the instructor's lectures and in-class instruction are more accurate. In order to be more precise, the article combines two sub-subjects and a variety of teacher-student interaction behaviors to jointly assess the type of interaction.

Table 3. Interactive behavior recognition Precision table

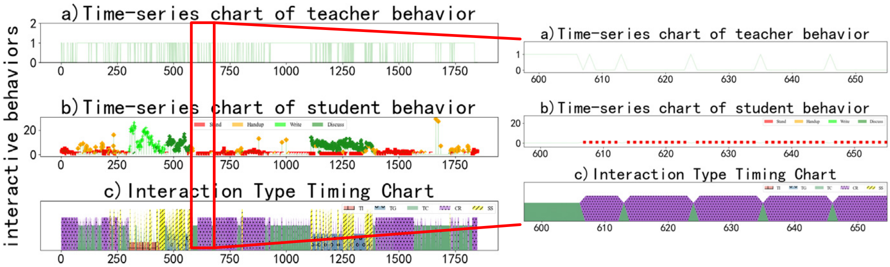
Interaction behavior	**Student**				**Teacher**			
	Stand	***Hand up***	***Write***	***Dis-cuss***	***Invite***	***Ask***	***Guide***	***Teach***
Precision(%)	94.30	85.00	90.40	87.70	57.60	88.70	92.50	97.20

The results of YOLOv8 experiments on student behavior and teacher behavior under real classroom images are shown in figure 3, which covers the behaviors that can be recognized by the manual observation method.

Fig. 3. Diagram of Experimental results

4.3 Experimental Analysis

Course A is from a high-quality example course of mathematics class in the elementary section. In figure 4, which has the full-time series diagram of one class on the left and the partially enlarged version on the right, from top to bottom, are the teacher behavior time series diagram, student behavior time series diagram, and interaction behavior time series results for the 40-minute one-class A.



a) Timing Chart b) Partial enlargement

Fig. 4. Interaction type timing Chart & Partial enlargement

The student behavior timing diagram displays four different student actions: standing, hand up, writing, and discussing. The teacher behavior timing diagram displays all observed teacher interaction behaviors. It is suggested that the behavior of the majority of students in the class at any given time prevails, and the important degree of student behavior is formulated as follows: students standing > students discussing > students writing > students raising hands. Based on the teacher timing diagram and student timing diagram, the type of teacher-student behavioral interaction at the current moment can be determined by combining the time and duration of student behavior occurrence, and the time and duration of teacher behavior occurrence, and then based on the coding of teacher-student interaction types shown in Table 1. As shown in Figure b, the type of classroom interaction is known to be alternating between student standing and teacher lecture between 600 and 650 seconds, and alternating between teacher-class interaction (TC) and cross interaction (CR).

The manual observation approach employed the same coding methods to record the interaction duration through observation. The experiment used one sample per second, and the interaction duration was counted according to the results every second.

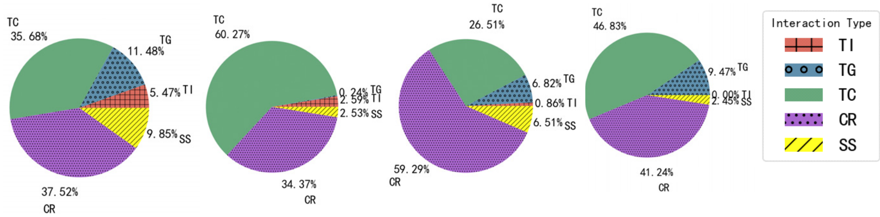
Table 4 displays the outcomes of the comparison. According to the proportion of total interactions to the course (the ratio of the sum of the five categories of interaction duration to the overall course duration), there is an error of roughly 1% between the types of interaction behaviors suggested by the model and the manual observation approach.

Table 4. Interactive behavior recognition Precision table

Course	Method	Interactive time(seconds)				
		TI	*TG*	*TC*	*CR*	*SS*
Course A	*model Method*	102	213	657	688	174
	Manual Statistics Method	102	212	659	693	179
Course B	*model Method*	44	4	1024	584	43
	Manual Statistics Method	43	4	1046	594	43
Course C	*Model Method*	11	87	338	756	83
	Manual Statistics Method	11	87	337	754	82
Course D	*model Method*	0	166	821	723	43
	Manual Statistics Method	0	168	819	720	42
Average variance rate(%)		2.27	0.47	2.75	2.70	4.08

It can be said that the maximum difference between the interaction behavior coding type suggested by the model and the results of the manual observation method was less than 3% and that the determination of the teacher-student interaction was made using the average difference rate (the ratio of the difference between the manual observation method and the model algorithm to the manual observation method), which was calculated based on the duration of each interaction type compared.

Excellent courses should have both types of interactions and appropriate proportions, such as Course A course interaction behavior in the cross-interaction accounted for more, while in the course arranged teacher-group interaction and student-student interaction, mobilized the subjectivity of students, and deepen the knowledge learned. Interaction types accounted for the total interaction time effect as shown in figure 5, Course A and Course C course balanced course compared to Course B and Course D course, the proportion of various types of interaction is moderate, so it is an excellent course while proving that the analysis results provided by the model approach can indeed provide a reference role.



a) Course A b) Course B c) Course C d) Course D

Fig. 5. Interaction Type Percentage Chart

5 Conclusions

Under the background of education intelligence, the article designs an intelligent analysis model of teacher-student interaction behavior in primary and secondary school classrooms based on teacher-student interaction classification theory, which is widely recognized in the education field, and combined with existing artificial intelligence technology. Teachers only need to provide teaching videos, and the model can judge and analyze teacher-student interaction behaviors and give visual statistical results.

References

1. Liu, Z. Chen, N. (2021) Research on the Interactive Behaviors in Classroom Teaching under Smart Classroom Environment-Taking 10 High-quality Junior High School Mathematics Teaching Courses as Observation Objects. Modern Educational Technology (09),28-36.
2. R. Fu, T. Wu, Z. Luo, F. Duan, X. Qiao and P. Guo, (2019) "Learning Behavior Analysis in Classroom Based on Deep Learning," Tenth International Conference on Intelligent Control and Information Processing (ICICIP), Marrakesh, Morocco, 2019, pp. 206-212, doi: 10.1109/ICICIP47338.2019.9012177
3. Ngoc Anh, B., Tung Son, N., Truong Lam, P., Phuong Chi, L., Huu Tuan, N., Cong Dat, N., Van Dinh, T. (2019). A computer-vision based application for student behavior monitoring in classroom. Retrieved from https://doi.org/10.3390/app9224729
4. Li, H., Kang, Y., Hao, Y., Ding, W., Wu, Z., & Liu, Z. (2021, July 15). A Multimodal Machine Learning Framework for teacher vocal delivery evaluation. arXiv.org. https://arxiv.org/abs/2107.07956v1

5. Luo, Z. Zhao, Q. Duan, F. (2021) Automatic analysis of classroom teaching process based on teacher's near-field speech. Modern Educational Technology (08), 76-84.
6. The Ministry of Education, "Notice on the issuance of 'Compulsory Education Curriculum Program and Curriculum Standards'," http://www.moe.gov.cn/srcsite/A26/s8001/202204/t20220420_619921.html.
7. Wang, X. Wu, L. (2012) A study on the changes of elementary school mathematics classroom teaching in ten years of curriculum reform. China Electro-Chemical Education (08),111-114+128.
8. Reis, D. Kupec, J. Hong, J. Daoudi, A. " Real-Time Flying Object Detection with YOLOv8". https://doi.org/10.48550/arXiv.2305.09972.