Learning by Participation with AI Enhancement

Takunda Makanda[1, a]

[1]Xi'an Technological University, Xi'an, Shaanxi, China

[a]milesrevenx@gmail.com

Abstract. With major strides in innovations in industries across the board it was only a matter of time where AI's impact on education will be fully recognized. The integration of Artificial Intelligence (AI) in education has revolutionized traditional learning paradigms, from the one size fits all design of classroom setups AI is offering personalized and efficient solutions to enhance student engagement and academic outcomes. This paper explores the application of YOLO11, a state-of-the-art real-time object detection model together with Classroom Discourse Analyzer and Generative Artificial Intelligence, in educational settings. By leveraging YOLO11's advanced capabilities in object detection and analytics, pose estimation, and instance segmentation, we propose innovative methods for interactive learning, classroom monitoring, and adaptive teaching strategies. The study highlights YOLO11's potential to identify and analyses student behavior, optimize resource allocation, and foster inclusive learning environments. Preliminary results demonstrate YOLO11's efficiency in processing educational data with unparalleled accuracy and speed, paving the way for its broader adoption in smart classrooms. This research underscores the transformative role of AI and the already paved groundwork in data collection with CDA, in shaping the future of education and emphasizes the need for further exploration of YOLO11's applications in diverse educational contexts.

Keywords: You Only Look Once 11 (YOLO11); Classroom Discourse Analyser (CDA); Generative Artificial Intelligence (GAI)

1. Introduction

Throughout the centuries the practice of teaching has spearheaded the sharing of ideas among likeminded individuals. Methods in how to convey or deliver the teachings has always been under development, technologies like the CDA addresses the difficulties regarding data set, coding, visualization, and tracking and comparison that teachers face in analyzing their classroom discourse data [1] [2]. The advancement of Artificial Intelligence (AI) technologies has significantly reshaped various sectors, including education. Along history of innovations and increments towards perfecting the method of delivering materials to the students is outlined in the paper 'A History of Instructional Media, Instructional Design, and Theories' [4], from visual instruction movement to audio instruction movement to the present efforts. As each new medium was introduced to the world, there was a great deal of initial interest and enthusiasm about its anticipated impact on instructional practices [5].

With the emergence of innovative tools like YOLO11, a cutting-edge real-time object detection model, educators and researchers are exploring novel methods to enhance teaching and learning experiences. Since YOLOv5 [6], the algorithm has significantly improved through the introduction of the CSPNet framework other approaches to applying artificial intelligence in learning includes the use of Generative Artificial Intelligence, a quintessential example is the Intelligent Tutoring System (ITS), meticulously designed to cater to students' unique learning levels and progress [5] [7]. YOLO11's capabilities extend beyond traditional applications, offering potential solutions for behavior analysis, classroom monitoring, and interactive learning environments.

Given the transformative power of these tools this paper aims to explore the best way to fit the enhancements of these tools and how they can integrate with the other available technologies. This paper investigates the integration of YOLO11 and GAI's in educational contexts, highlighting its ability to analyses visual data efficiently and provide actionable insights. Generative AI on the other hand presents a better reasoning capacity than human, this is a vital resource when analyzing the classroom engagements and providing feedback on how best the improvements can assist the students and classroom interaction [7].

The Intent is to collaborate Generative AI's analyzing skills, CDA's classroom environment data and analysis approach and YOLO11's, precision and speed, educators can address challenges such as returning student attention, personalized learning, classroom engagement and resource management [9]. This research aims to demonstrate the transformative potential of AI models and other technologies in education, paving the way for AI-driven strategies that foster inclusivity, adaptability, and innovation in modern learning environments.

2. Backgrounds

Understanding these technologies how they incremented the classroom experience can better establish our basis on how their collaboration can be a greater impact on the material delivery. In this section we take a deeper dive on these milestones and the features this paper would highlight as the key focus.

2.1 Generative Artificial Intelligence. Generative Artificial Intelligence makes use of large language models in areas that were particularly regarded as human purview, these include analytical abilities, fraud detection and creative work [10]. With much evolution since its introduction, there has been much changes that has seen generative Artificial Intelligence improving personalized learning.

2.2 GAI Overview. GAI's can transform personalized learning in educational settings, especially classrooms. It provides a comprehensive analysis of how GAI can enhance teaching strategies, materials, environments, and student outcomes, all the while addressing the practical and ethical considerations of its implementation [11].

2.3 Highlighted Features. GAI's can transform personalized learning in educational settings, especially classrooms. This section provides a comprehensive analysis of how GAI can facilitate enhanced teaching strategies, material delivery, environment analysis, and student outcomes, all while addressing the practical and ethical considerations of its implementation.

Personalized Learning Strategies – by providing Socratic and tailored questioning, students are encouraged to engage in deep thinking and creative thinking.

Customized Learning Strategies – By implementing GAI's student learning materials and learning pace can be customized to a student's needs.

Automated and Adaptive Teaching Materials – GAI's can generate specific subject content/resources in a variety of subjects to suit the needs of a student, also turning textbook content into more interactive forms.

Intelligent Learning Environment – Students can receive real-time support by creating smart adaptive classrooms

These are among the many features that are valuable in a classroom setting that can be supported by GAI's with the development of hybrid intelligence systems, where human and machine roles complement each other to improve teaching and learning outcomes [12] [10].

2.4 You Only Look Once (YOLO). YOLOv11 is the latest version in a series of real-time object detection models. It brings significant architectural improvements, including the C3k2 block, C2PSA attention module, and SPPF enhancements [13], allowing for: Higher accuracy, Faster processing and Broader task support (e.g., object detection, pose estimation, segmentation, classification).

2.5 YOLOv11 Overview. With improved real-time detection, there are significant applications in education as outlined above, not just limited to that but also expands to situations where attention is of the utmost importance like special needs classes [14]. With its expansive span in applicability and scalability it is feasible to deploy to schools that might not have high end equipment using Nano and small models of YOLOv11 are ideal for low-cost edge devices like Raspberry Pi or NVIDIA Jetson.

2.6 Highlighted Features. Automated Attendance and Participation Tracking - YOLOv11's oriented bounding box (OBB) capabilities enable recognition of students from different camera angle, Face and posture detection can infer presence and alert for absenteeism

Smart Classroom Monitoring – by monitoring the students using object detection [15] [16], face orientation and pose estimation the model can provided insights in the student interaction and engagement with three class teachings and materials.

Real-Time Feedback – This feature is most useful in areas like practical lessons (e.g. Lab experiments, physical education), to help enforce rules for safety and in special education classrooms it comes in handy in pose detection when a student might need assistance.

Enhanced Accessibility and Inclusion – as mentioned in the above statement students with motor functional challenges my benefit from detection of the different postures when help is needed.

2.7 Classroom Discourse Analyzer (CDA). The practicality of CDA in assessing student engagement over classroom discussions this tool helps with assessing the patterns and quality of verbal interactions between teachers and students. It transforms raw classroom dialogue into coded, visualized, and trackable data for analysis [1].

2.8 CDA Overview. The attention provided by the teacher makes CDA more adaptible to the teachers needs as they code the data themselves and turns the classroom discussions more manageable and easier to track.

2.9 Highlighted Features. Structured Data Input & Transformation – Teachers enter data in classified attributes to track the Session, turn, speaker.

Multi-Dimensional Coding – CDA codes each conversational turn using three dimensions, evaluation, knowledge content and participation invitation.

Visualization of Discourse – utilizing graphical representation of the discourse coded times the tool helps the teachers see who dominates the discussions, which students are silent.

Tracking & Comparison Tools – Over time the accumulated data can be used to track the improvements and student contribution levels. With the accumulated data it can be used to track participation, compare sessions and contrast different classrooms.

3. Modelling and Integration

Each of these technologies focus on a specific area that they excel exceptionally well in a classroom; how best can they work together in enhancing the overall performance of the classroom. For example, with YOLO it gives a comprehensive analysis on the posture and attentiveness of the student, While CDA gives the depth of contribution conversationally. In this example the two technologies can depend on each other in the form presented by the Table 1.

3.1 Model. The core idea is to harness the core strengths from these technologies to create an environment that pioneers and fosters inclusivity. Active participation and active listening are key contributors towards this model and the key features extracted from these models. As highlighted in the second section above the features to be aimed at from these contributors are massive improvements and the proposed outlay of management between the models is displayed in the Figure 1.

Table 1 The performance of CDA and YOLOv11 in class

Component	CDA	YOLOv11
Input Type	Textual transcripts of classroom discourse	Live/stored video feeds
Data Focus	Verbal participation patterns, discourse quality	Visual behavior, gesture, posture, object/person tracking
Output	Turn-taking, talk moves, content quality	Student engagement levels, hand-raising, movement, attention
Role in Classroom	Understand what and how students speak	Observe who, when, and how students behave non-verbally

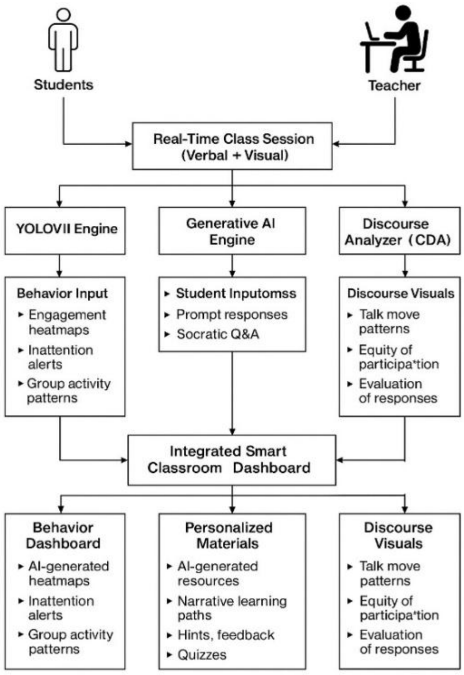

Figure 1. The proposed outlay of management between Student and Teacher

3.2 Training. It is without doubt that the AI models in both in LLMs and pattern recognition have a bias proportionality and only by training it with the correct data can help mitigate this issue, the human component in this also needs to be properly trained so as to limit the human errors.

3.3 Model Training. While many factors affect biases in machine learning, one of the things that we can directly effect change in is by training our model with the correct information, this is done in labelling of the training data. When training our AI models for classrooms it is also important to remember the minorities these can greatly affect our bias as the minorities are underrepresented [17]. In our case these can be the communities like the disabled and the marginalized groups of the society, by the virtue of identifying them it can be easier to find data that directly correlates to them specifically and making sure it is correct.

3.4 Teaching Staff Training. While Model training has its fair share of human error and human bias the most interactive group of individuals with the system are the teachers and their continual involvement needs to be accounted for. The CDA system requires and relies on the teachers to code the conversations correctly, with this information the dependence of the system on this information cannot afford these errors. Appropriate training and simplification of the system should be an aim when designing the system.

3.5 Proposed Reliance levels. Given the biases and errors discussed above here is a proposed reliance probability chart that aims to deliver successful results, as shown in Figure 2.

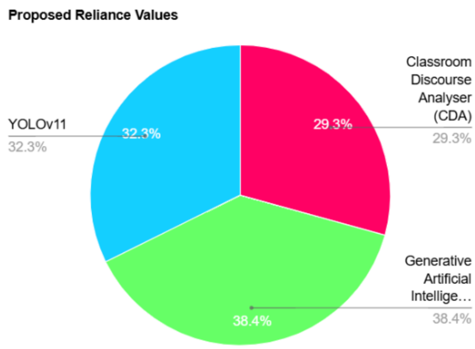

Figure 2. The dependency probability diagram of the successful delivery result

4. Analysis

While each feature in this in this integration has its limitation in this section the aim is to figure out how best can we mitigate these shortcomings. By weighing the advantages and disadvantages of each we can have a better understanding on which feature we can lean on more to improve our efficiency, as shown in Table 2.

Table 2 Comparative analysis of three technologies

Technology	Best Use Cases	Strengths	Limitations / Bias Risks
Generative AI (GAI)	- Personalized learning - Dynamic content generation - Conversational agents	- Scalable - Adaptive - Subject diverse - Easy to deploy	- Hallucination risk - Cultural/language bias - Over-reliance
YOLOv11	- Visual engagement tracking - Safety in labs and P.E - Attention/posture analytics	- Real-time - High accuracy on visuals - Pose aware	- Camera/setup sensitive - Demographic bias (e.g. skin tone)
CDA	- Teacher PD - Dialogue quality tracking - Equity analysis	- Nuanced discourse insight - Supports reflection	- Manual coding overhead -Speech-to-text dependency

4.1 Feature Reliance Analysis. Section 3.2.3 provided a suggestion of the proposed reliance on these features and here we dive a little deeper how we come to this suggestion. CDA has the lowest reliance though its importance is realized, its reliability is reduced by the continuous need for human input that can have some errors.

According to research from MIT there are better methods of handling the biases in AI models and due to these new and basic safeguards AI models do commend more reliability. With that being said we can model our system as follows.

$$P(overall) = P(GAI) \bullet w(GAI) + P(YOLO) \bullet w(YOLO) + P(CDA) \bullet w(CDA) \quad (1)$$

Using the above approach, we can assess the probability of a successful launch with our dependence on each feature. Where weights sum to 1 (i.e., 0.38 + 0.33 + 0.29 = 1):

$$P(overall) \approx (0.85 \times 0.38) + (0.78 \times 0.33) + (0.70 \times 0.29) \approx 78.3\% \quad (2)$$

Thus, taking into consideration the errors that can arise from these features under ideal but realistic conditions, the system has a 78% probability of functional success.

4.2 Impact Assessment. With these technologies working together there is an expected vast improvement in the overall classroom performance as a multimodal system they complement each other in areas where one feature cannot excel.

Comparatively deployment of assisted AI systems in education has been overtly reliant on one model more than taking into consideration the contribution the impact others can have when deployed all together. Around 44% of students interact with GAI's in classrooms [18], with 60% of teachers have integrated AI into their daily teaching.

With the current trajectory AI usage in all aspects of our live has shown to be a powerful tool and rightfully so the integration of these systems can provide a broader coverage as shown in the Table 3.

Table 3 The deployment of artificial intelligence systems in education

Impact Area	Current Baseline	Projected Improvement	Justification
Student Engagement	50–60% active participation per class	+20–35%	YOLOv11 detects disengagement in real-time; GAI personalizes instruction
Learning Outcomes	Mixed: ~30–50% proficiency in key subjects	+10–25%	GAI boosts formative feedback; CDA helps improve reasoning-based discussion

Equity of Participation	20–30% students dominate 70% of talk	+30–50% improvement	CDA provides visibility into participation gaps, enabling teacher rebalancing
Teacher Workload Efficiency	High (manual prep, slow feedback cycles)	-20–40% time burden	GAI automates materials; CDA automates feedback loops
Student Retention / Dropout	~15% dropout in some systems (avg)	-5–10% relative reduction	More personalization + engagement = lower risk of disengagement/dropout
Teacher Professional Growth	Slow, intuition-based	+30–60% insight gain	CDA visualizes practice; GAI simulates feedback loops for teacher PD

5. Conclusion

While Deployment of these models separately has seen great improvements in the overall performance in classrooms, this paper introduces a different view on how these models can be integrated into features. The system relies on the strengths of each feature to produce a wholesome outlook of a classroom-oriented education, not to mention the exciting developments that can come from this integration such as online learning with assisted virtual reality with eye-tracking assistance to detect attention levels.

This paper has contributed to the broader view of the classroom dynamics encompassing classroom behavior detection, personalized education and interaction tracking these as per this papers concern are quintessential. It also highlights the pitfalls to be expected when working with each feature and how to mitigate these pitfalls.

It is in this paper's belief that future works can improve on the efficiency and deployment strategies to create a more interactive and inclusive learning experience.

References

[1] Chen, Gaowei & Clarke, Sherice & Resnick, Lauren. (2015). Classroom Discourse Analyzer (CDA): A Discourse Analytic Tool for Teachers.

[2] Clarke, S., Chen, G., Stainton, K., Katz, S., Greeno, J., Resnick, L., ... & Rosé, C. P. (2013). The impact of CSCL beyond the online environment. In Proceedings of Computer Supported Collaborative Learning.

[3] Yuang, W. Yuan-Hao, J. Jiayi, L. Changyong, Q. Rui, J. (2024) The Advancement of Personalized Learning Potentially Accelerated by Generative AI.

[4] An, Y. (2021). A history of instructional media, instructional design, and theories. International Journal of Technology in Education (IJTE), 4(1), 1-21.

[5] Yuang, W. Yizhou, Z. Yuan-Hao, J. Bo, J. (2024). Enhancing Explainability of Knowledge Learning Paths: Causal Knowledge Networks.

[6] Zijian, H. Kang, W. Tian, F. Lei, S. Rui, C. Xihong, F. (2024). Comprehensive Performance Evaluation of YOLOv11, YOLOv10, YOLOv9, YOLOv8 and YOLOv5 on Object Detection of Power Equipment.

[7] B. Jiang, Y. Wei, M. Gu, and C. Yin. Understanding students' backtracking behaviours in digital textbooks: a data-driven perspective. Interactive Learning Environments, pages 1–18, 2023.

[8] Kenneth R Koedinger, Emma Brunskill, Ryan SJd Baker, Elizabeth A McLaughlin, and John Stamper. New potentials for data-driven intelligent tutoring system development and optimization. AI Magazine, 34(3):27–41, 2013.

[9] Yang F, Wang T, Wang X. Student Classroom Behaviour Detection based on YOLOv7-BRA and Multi-Model Fusion[J]. arXiv preprint arXiv:2305.07825, 2023.

[10] Jeffrey L Elman. Finding structure in time. Cognitive science, 14(2):179–211, 1990.

[11] John R. Anderson and Edward Skwarecki. The automated tutoring of introductory computer programming. Communications of the ACM, 29(9):842–849, 1986.

[12] Gwo-Jen Hwang. Definition, framework and research issues of smart learning environments-a context-aware ubiquitous learning perspective. Smart Learning Environments, 1(1):1–14, 2014.

[13] Glenn Jocher and Jing Qiu. Ultralytics yolo11, 2024.

[14] Rahima, K. Muhammad, H. (2024) YOLOV11: AN OVERVIEW OF THE KEY ARCHITECTURAL ENHANCEMENTS.

[15] FAN, Y. Xiaofei, W. (2024) Student Classroom Behaviour Detection and Analysis Based on Spatio-Temporal Network and Multi-Model Fusion.

[16] FAN, Y. (2024) SCB-Dataset: A Dataset for Detecting Student Classroom Behavior

[17] TAHSIN A. KHEYA, MOHAMED R. BOUADJENEK, SUNIL, A. (2024) The Pursuit of Fairness in Artificial Intelligence Models: A Survey.

[18] AIRPM, AI in Education Statistics, AI in Education Statistics · AIPRM, 2025, Accessed: 15 May 2025.