AboulHassane, C. et al. (Eds.) (2024). Proceedings of the 32nd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Real-time Adaptive Learning Environments Using Gaze and Emotion Recognition Engagement and Learning Outcomes

AboulHassane CISSE[a*]
[a]*Graduate School of Informatics, Osaka Metropolitan University, Japan*
*sn22869p@st.omu.ac.jp

Abstract: This study explores the integration of real-time adaptive learning environments with gaze tracking and emotion recognition technologies to enhance student engagement and learning outcomes. By leveraging artificial intelligence (AI) and machine learning, the research aims to develop a framework that dynamically adjusts instructional strategies based on students' cognitive and emotional states. The focus is on three objectives: integrating these technologies, evaluating their impact, and assessing the effectiveness of automated communication strategies—Affective Backchannels (AB), Conversational Strategies (CS), and their combination (AB+CS). Using a mixed-methods approach, data from approximately 30 university students in digital learning environments will be analyzed. The findings are expected to provide valuable insights into the application of these technologies in education, potentially informing future educational policies and practices.

Keywords: Adaptive Learning, Gaze Tracking, Emotion Recognition, Student Engagement, Artificial Intelligence in Education

1. Introduction

In the evolving landscape of educational technology, integrating advanced analytical tools to enhance learning environments has become a focal point of research. This study explores the integration of gaze tracking and facial emotion recognition technologies to improve student engagement and learning outcomes. Recent advancements in AI and machine learning facilitate the development of systems capable of interpreting human behaviors, such as eye movements and facial expressions, which are critical in understanding student engagement and comprehension (D'Mello, Duckworth, & DiCerbo, 2017). This research aims to fill the gap by developing an adaptive learning framework that responds to students' cognitive and emotional states, fostering a more engaging learning experience.

Additionally, previous research by Ayedoun, Hayashi, and Seta (2019) has demonstrated the effectiveness of incorporating automated communication strategies, specifically Affective Backchannels (AB), Conversational Strategies (CS), and their combination (AB+CS), in enhancing learners' willingness to communicate in second language acquisition. Affective Backchannels (AB) refer to non-verbal cues that provide emotional feedback, while Conversational Strategies (CS) involve verbal techniques aimed at maintaining and encouraging dialogue. The combination of these strategies (AB+CS) has been shown to significantly improve interaction and engagement in educational settings.

Building on previous work by AboulHassane, Seta, and Hayashi (2024), which explored the use of multimodal data like gaze tracking and emotion recognition in human-agent communication, this research aims to develop an adaptive learning framework that responds to students' cognitive and emotional states for a more personalized and engaging learning experience (AboulHassane, Seta, & Hayashi, 2024).

The objectives of this research are as follows:

- **Integrate Real-time Eye-tracking and Emotion Recognition Technologies**: Develop and implement a system that dynamically adjusts instructional strategies in classroom settings (Shaun, De Baker & Inventado, 2014).
- **Evaluate the Impact on Student Engagement and Learning Outcomes:** Assess how these environments influence student engagement, concentration, and overall learning outcomes (D'Mello et al., 2017).
- **Assess the Impact of Automated Communication Strategies**: Investigate how automated communication strategies—AB, CS, and AB+CS—enhance instructional effectiveness.

2. Research Questions and Hypotheses:

How does integrating real-time eye-tracking and emotion recognition technologies, along with automated communication strategies (Affective Backchannels and Conversational Strategies), into adaptive learning environments impact student engagement, concentration, and overall learning outcomes compared to traditional instructional methods?
It is hypothesized that the integration of real-time eye-tracking and emotion recognition technologies, coupled with automated communication strategies (Affective Backchannels and Conversational Strategies), into adaptive learning environments will significantly enhance student engagement, concentration, and overall learning outcomes compared to traditional instructional methods (D'Mello, Duckworth, & DiCerbo, 2017; Shaun, De Baker & Inventado, 2014; Darling-Hammond, Flook, Cook-Harvey, Barron, & Osher, 2020).

Table 1. *Table of RQs and Hs*

	Research Question (RQ)	Hypotheses (H)
1	How effectively can real-time eye-tracking and emotion recognition technologies be integrated into classroom environments to dynamically adjust instructional strategies?	Integration of real-time eye-tracking and emotion recognition technologies will result in more personalized and responsive teaching, leading to higher student engagement and better comprehension (D'Mello, Duckworth, & DiCerbo, 2017).
2	What are the impacts of adaptive learning environments on student engagement, concentration, and overall learning outcomes?	Adaptive learning environments that utilize real-time gaze and emotion data will enhance student engagement, concentration, and result in superior learning outcomes compared to traditional learning environments (Shaun, De Baker & Inventado, 2014).
3	How do automated communication strategies—Affective Backchannels (AB), Conversational Strategies (CS), and their combination (AB+CS)—integrated with real-time eye-tracking and emotion recognition technologies influence student engagement and learning outcomes in adaptive learning environments?	The use of automated communication strategies—AB, CS, and AB+CS—in conjunction with real-time eye-tracking and emotion recognition technologies significantly enhances student engagement and learning outcomes, offering an optimized instructional approach tailored to the individual's communication dynamics (Darling-Hammond et al., 2020).

3. Methods

3.1 Design

A mixed-methods research design combining quantitative and qualitative approaches will be employed to comprehensively analyze the impact of real-time eye-tracking, emotion recognition technologies, and automated communication strategies on student engagement and learning outcomes.

3.2 Participants

Approximately 30 university students enrolled in digital learning courses across multiple classrooms will participate, ensuring a diverse and generalizable sample.

3.3 Procedures

The procedures for this study involve several key steps. First, eye-tracking devices (e.g., Tobii) and emotion recognition software (e.g., OpenFace) will be integrated into the Peter Conversational Agent to capture real-time gaze and emotional data. These data will be saved into a database for storage, and a dashboard will be used to analyze and display the information in real time. For data collection, both quantitative and qualitative methods will be employed. Quantitative data will include eye-tracking metrics, emotion recognition outputs, and pre-test/post-test academic performance measures. Qualitative data will be gathered through interviews and observational records to provide contextual insights.

The collected data will address the three research questions as follows:

i. **Integration of Real-time Technologies**: Eye-tracking metrics and emotion recognition outputs will provide real-time feedback on students' attention and emotional states, enabling dynamic adjustments to instructional strategies based on how accurately the Peter Conversational Agent interprets gaze patterns and emotional cues.

ii. **Impact on Engagement and Learning Outcomes:** Pre-test and post-test performance, along with eye-tracking and emotion recognition data, will evaluate how the Peter Conversational Agent influences student engagement and learning outcomes by assessing attention, emotional responses, and improvements in concentration, understanding, and performance.

iii. **Role of Automated Communication Strategies (AB, CS, AB+CS):** The effectiveness of Affective Backchannels (AB), Conversational Strategies (CS), and their combination (AB+CS) will be assessed by how the Peter Conversational Agent adapts to students' cognitive and emotional states, with qualitative insights from interviews and observational data and quantitative analysis of eye-tracking and emotion recognition outputs to evaluate instructional effectiveness and student interaction.

3.4 System Design

The system architecture integrates biometric and behavioral data collection, including heart rate monitoring, eye tracking, and facial emotion recognition. The Peter Conversational Agent adapts its strategies based on this data to enhance language learning engagement.

Figure 1: Screenshot of Peter Conversational Agent

3.5 Algorithm for Real-Time Analysis

The algorithm for real-time analysis involves several critical steps. Initially, biometric data will be cleaned and normalized to remove noise and ensure accuracy. Key features indicating engagement, emotional state, and interaction patterns will then be identified through feature extraction. Supervised learning algorithms will be employed for classification and regression, enabling the classification of emotional states and the prediction of engagement, with time-series analysis used for dynamic response adaptation. Based on the data analysis, the agent will adjust dialogue strategies by modifying language complexity, feedback, and emotional tone. Finally, evaluation metrics, including model accuracy, precision, recall, F1 scores, and mean squared error (MSE), will be used to assess the system's performance, alongside evaluations of user experience and language learning outcomes.

3.6 Data Storage and Database

Biometric and conversational data are stored locally in CSV and TXT formats to facilitate offline analysis and easy parsing. The system employs PostgreSQL, hosted on Google Cloud, to manage and analyze large volumes of multimodal interaction data, ensuring secure and efficient storage, processing, and retrieval, which is crucial for enabling real-time adaptive learning interventions. To maintain compliance with privacy regulations, data anonymization is implemented, and secure data transfer protocols, such as Bluetooth and cables, are used. Users are also informed about data usage to ensure transparency and adherence to privacy standards.

4. Potential Impact

This ongoing research holds significant potential to influence various aspects of education. The findings are expected to inform educators about the benefits of adaptive learning technologies, thereby enabling more personalized teaching approaches that cater to individual student needs (Darling-Hammond et al., 2020). The study aims to demonstrate how adaptive learning environments can enhance student engagement and concentration, ultimately leading to improved academic performance (Pane et al., 2017). Moreover, as the development of the system to collect and process real-time data continues, this research will provide valuable data-driven insights that can guide future educational policies regarding the integration of technology-enhanced learning environments. By contributing to the understanding and application of these advanced technologies, the research has the potential to shape the future of educational practices and policies significantly.

Acknowledgments

This work was supported by JST SPRING, Grant Number JPMJSP2138.

References

AboulHassane, C., Seta, K., & Hayashi, Y. (2024). Integrative analysis of multimodal interaction data: Predicting communication dynamics and willingness to communicate (WtC) in human-agent interaction. *Proceedings of LASI-Europe 2024*, CEUR-WS. https://ceur-ws.org/Vol-3738/paper4.pdf

Ayedoun, E., Hayashi, Y., & Seta, K. (2019). Adding communicative and affective strategies to an embodied conversational agent to enhance second language learners' willingness to communicate. International Journal of Artificial Intelligence in Education, 29, 29-57.

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. Applied developmental science, 24(2), 97-140.

D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, analytic, automated (AAA) measurement of engagement during learning. Educational psychologist, 52(2), 104-123.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. Learning and Instruction, 22(2), 145-157.

Pane, J. F., Steiner, E. D., Baird, M. D., Hamilton, L. S., & Pane, J. D. (2017). Informing Progress: Insights on Personalized Learning Implementation and Effects. Research Report. RR-2042-BMGF. Rand Corporation.

Plass, J. L., & Kaplan, U. (2016). Emotional design in digital media for learning. In Emotions, technology, design, and learning (pp. 131-161). Academic Press.

Shaun, R., De Baker, J., & Inventado, P. S. (2014). Chapter 4: Educational Data Mining and Learning Analytics.