Multimodal Physiological Sensing for Adaptive Learning Environments

Alessio Ferrato[1,*], Lorenzo Battisti[1], Giorgio Biancini[1], Marco Napoleone[1], Sabrina Fagioli[2], Carla Limongelli[1], Mauro Mezzini[2], Davide Nardo[2] and Giuseppe Sansonetti[1]

[1]Department of Engineering, Roma Tre University, Rome, Italy
[2]Department of Education, Roma Tre University, Rome, Italy

Abstract

The COVID-19 pandemic has highlighted critical challenges in maintaining student engagement across educational settings, with decreased attendance and participation affecting learning outcomes. The multidimensional nature of student engagement, encompassing emotional, behavioral and cognitive components, presents significant methodological challenges in measurement despite its established role as a crucial predictor of academic resilience. Although computer science research has integrated sensors and cameras for monitoring student states, existing solutions face privacy concerns, environmental dependencies, and scalability issues, with no freely available frameworks for researchers. This research proposes an open-source intelligent interface system that integrates multimodal physiological sensing with real-time logging and analytics to study and enhance teaching effectiveness. By monitoring physiological indicators through a distributed architecture, the system enables classroom-wide monitoring while preserving privacy. The framework utilizes wearable biosensors and ambient audio analysis to create a comprehensive engagement monitoring platform that can record, detect, and respond to the complex interplay between emotional states and learning processes. Beyond its technological contributions, this research aims to enhance educational accessibility and inclusion, particularly for disadvantaged and at-risk populations. The system ability to provide objective and real-time insights supports data-driven teaching adaptations, representing a step toward more effective and adaptive learning environments.

Keywords

Adaptive Learning, Embedded Systems, Inclusive Education Interfaces, Internet of Things

1. Introduction

The COVID-19 pandemic has profoundly affected traditional education, decreasing attendance in conventional and online university courses. Students frequently cite a lack of engagement and motivation as significant barriers to active participation, which impacts learning outcomes. Student engagement, defined as the level of interest, curiosity, and participation in learning activities [1, 2], is widely acknowledged as a key predictor of academic resilience and persistence in university courses. Consequently, it is crucial in mitigating student dropout rates [3]. Numerous studies have endeavored to identify the profiles of students at risk of dropout by examining the determinants of engagement [1, 4]. However, the multifaceted nature of engagement presents challenges for researchers in clearly defining the construct, particularly in disentangling the specific dimensions or components—such as emotional, behavioral, and cognitive—that constitute engagement.

Developing a comprehensive profile of engaged students can provide valuable insights into the key factors that promote successful learning outcomes. This information can also help adopt educational strategies tailored to individual needs to enhance academic resilience and reduce student attrition, especially for disadvantaged and at-risk populations. These challenges in understanding and measuring student engagement highlight the critical need for a comprehensive framework to collect and analyze

Joint Proceedings of the ACM IUI Workshops 2025, March 24-27, 2025, Cagliari, Italy
*Corresponding author.
alessio.ferrato@uniroma3.it (A. Ferrato); lor.battisti5@stud.uniroma3.it (L. Battisti); gio.biancini@stud.uniroma3.it (G. Biancini); mar.napoleone3@stud.uniroma3.it (M. Napoleone); sabrina.fagioli@uniroma3.it (S. Fagioli); carla.limongelli@uniroma3.it (C. Limongelli); mauro.mezzini@uniroma3.it (M. Mezzini); davide.nardo@uniroma3.it (D. Nardo); gsansone@dia.uniroma3.it (G. Sansonetti)
0000-0001-5405-3018 (A. Ferrato); 0000-0003-4953-1390 (G. Sansonetti)

real-time physiological and behavioral data from students in order to merge quantitative metrics with qualitative methods [5]. Even if multiple research studies have demonstrated the effectiveness of sensor-based and camera-based systems in monitoring students [6, 7, 8] and other categories of users [9], existing solutions face significant challenges, including privacy concerns, environmental dependencies, scalability issues and none of them is freely available for researchers. These challenges are also perceived in other domains, such as recommender systems [10, 11], where personalization plays a key role.

In this paper, we present our idea of an open-source framework for real-time collection and analysis of multimodal physiological data in educational settings. Our system will integrate non-invasive sensors and a distributed architecture for classroom-wide monitoring, enabling scalable and privacy-preserving data collection and engagement assessment. This framework addresses the critical need for accessible tools to help researchers and educators understand and respond to student engagement patterns, ultimately supporting the development of more effective and adaptive learning environments for all students, including those from vulnerable and disadvantaged backgrounds.

2. Research Background

In recent years, classroom environments have undergone significant transformation by integrating sensing technologies to understand and quantify student behavior. These smart classroom systems represent a convergence of sensor technologies, computer vision, and machine learning [12]. While traditional classroom observation relies heavily on teacher perception, modern approaches leverage multimodal sensing architectures to capture fine-grained behavioral data. This review focuses on technological approaches for recognizing and analyzing student actions and behaviors in classroom settings. We examine systems utilizing different sensing modalities to capture and classify student activities, particularly implementations demonstrating scalability and real-world applicability.

Woolf et al.[6] pioneered emotion-aware tutoring systems by implementing a scalable multi-sensor architecture for classroom environments. Their system architecture integrated four sensor streams (galvanic skin response, pressure-sensitive mouse input, posture detection chair, and facial expression recognition) with an intelligent tutoring system, enabling parallel affect detection for up to 25 concurrent users. Their follow-up work [13] evaluated the system's predictive performance, demonstrating that their machine learning model achieved an accuracy of 60% in emotion classification when combining multimodal sensor data with computer vision, significantly outperforming baseline models that relied solely on tutoring interaction metrics.

Building on this foundation, Zaletelj et al. [14] advanced the field with a Kinect-based attention monitoring system. Their system tracked facial and body behaviors through single-camera setups focused on groups of 3-4 students. Testing on 18 subjects achieved 75.3% accuracy for three-level attention classification. The researchers defined engagement levels based on observable behaviors: high engagement (writing notes most of the time, leaning forward), medium engagement (observing slides, head movements), and low engagement (leaning back, yawning, looking away). Key features tracked were body posture (head position, lean angle), gaze direction (toward slides/board/away), and facial signals (eyes, yawning). Human observers provided ground truth data by rating attention levels 1-5, then normalized to three levels. While effective for small groups, the system faced challenges with feature detection reliability and capturing comprehensive behavioral differences.

Ahuja et al. [7] addressed previous limitations with EduSense, a comprehensive classroom sensing system utilizing computer vision and audio processing for automated classroom analytics. Their system employed a distributed architecture with dual-camera setups (student-facing and instructor-facing) in each classroom, processing multiple data streams simultaneously. The researchers achieved notable accuracy rates in their controlled study (n=30), with 95.5% accuracy for body detection and 94.6% for hand raise detection. Their subsequent real-world deployment across 22 courses (n=687) maintained robust performance, achieving 92.2% accuracy for student detection and 99.6% for instructor detection. While the system demonstrated promising capabilities in affect detection and classroom dynamics analysis, it faced challenges with processing speed (0.3-2.0 FPS) and facial analysis at greater distances, highlighting the ongoing technical challenges in scaling such systems.

More recently, Wang et al. [8] introduced an AI-powered motion sensor system for monitoring student classroom behaviors. The researchers developed a system using motion sensors (placed on students' backs and shoulders) combined with a novel algorithm called Voting-Based Dynamic Time Warping (VB-DTW) to identify different classroom behaviors. The authors tested it with 13 participants who simulated 14 common classroom behaviors (e.g., sitting still, writing, shaking legs). While the system performed excellently for most movements, it struggled with detecting subtle, prolonged actions like writing and recognition of near-static data. Additionally, the sensor position requires a longer setup than less intrusive solutions (e.g., smartbands).

While the above research demonstrates promising capabilities in monitoring student engagement, they predominantly rely on camera-based solutions or intrusive sensor placements that raise privacy concerns and implementation challenges. Furthermore, none of these solutions are freely available to researchers, limiting reproducibility and broader adoption. We propose a new framework that addresses these limitations by utilizing non-invasive wearable biosensors and ambient audio analysis, offering a privacy-preserving and scalable solution for engagement monitoring. By providing an open-source platform, we aim to enable researchers to study the complex relationships between physiological indicators and learning processes, ultimately supporting the development of more adaptive and inclusive educational environments.

3. System Design

Figure 1 is a high-level illustration of how we plan to design the system. It comprises data collection, processing, and interface layers that work together to create a comprehensive engagement monitoring platform.

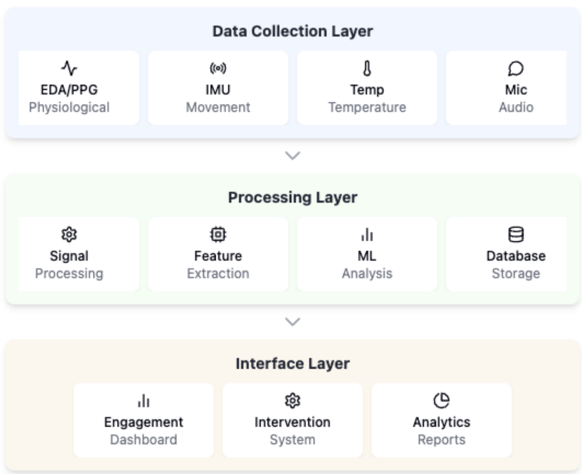

Figure 1: System diagram of the proposed classroom monitoring platform. The Data Collection Layer includes multimodal sensors (EDA/PPG for physiological data, IMU for movement tracking, temperature monitoring, and microphone for audio input). The Processing Layer shows the main components: signal processing, feature extraction, machine learning analysis, and database storage. The Interface Layer presents the user-facing components: an engagement dashboard, an intervention system, and analytics reports.

3.1. Hardware Implementation

At the core of our data collection infrastructure lies the EmotiBit board [1], developed by OpenBCI[2] Corporation. This compact biosensing board integrates multiple measurement modalities into a unified wearable device, offering a robust physiological and kinematic data acquisition solution. The system's primary sensing component consists of an Electrodermal Activity (EDA) measurement system that captures variations in skin conductance through integrated electrodes. This works with a Photoplethysmography (PPG) module, which employs optical sensors to measure dynamic blood volume changes for cardiovascular monitoring. Additionally, the board incorporates a precision temperature sensor for thermal measurements and a 9-axis Inertial Measurement Unit (IMU) that combines triaxial accelerometer, gyroscope, and magnetometer capabilities for comprehensive motion tracking.

We will integrate a microphone into the board for ambient audio processing. This microphone captures individual audio features while maintaining privacy through processing techniques that extract only participation-related metrics without storing raw audio data. The audio processing focuses on detecting general classroom dynamics, such as discussion intensity and turn-taking patterns, rather than recording specific conversations or identifying individual voices.

A key aspect of our implementation is its operation through Bluetooth connectivity. Each device is a Bluetooth peripheral, connecting to a central classroom computer that handles data collection and processing. This wireless approach eliminates the need for complex wiring while enabling flexible deployment in various classroom settings. The system's architecture is designed to handle multiple concurrent Bluetooth connections, allowing for seamless classroom-wide monitoring.

Students randomly take a smart band upon entering the classroom in our vision to ensure further privacy preservation. This randomization means no specific device is permanently associated with any individual student, effectively anonymizing the collected data. At the end of each session, the devices are returned and sanitized, ready for random redistribution in the next class.

This system effectively addresses several limitations inherent in traditional camera-based monitoring systems. While cameras typically require controlled lighting conditions and fixed mounting positions, our wearable approach enables data collection in naturalistic environments. Furthermore, the EmotiBit provides direct physiological measurements through its contact-based sensors, eliminating the need to infer physiological states through visual processing algorithms. This direct measurement approach significantly reduces common sources of error found in camera-based systems (e.g., motion artifacts).

3.2. System Architecture

The system architecture follows a layered data processing and analysis approach, beginning with the Data Collection Layer. This foundation manages the acquisition of multimodal sensor data from the EmotiBit devices.

The Processing Layer builds upon this data foundation through a sophisticated pipeline that transforms raw sensor inputs into meaningful insights. This transformation begins with signal processing modules that apply filtering and noise reduction techniques. The cleaned signals then undergo feature extraction to derive relevant characteristics from the multimodal data. These features feed into machine learning models that analyze patterns and classify engagement levels. Throughout this process, a database stores the processed and raw data. Raw data will serve for further offline analysis while processed for the real-time dashboard.

The Interface Layer is the primary means for researchers and educators to interact with and interpret the collected data in real time. The engagement dashboard, mocked in Figure 2, presents a comprehensive view of classroom dynamics through four essential metrics cards. The Engagement card displays an aggregate score derived from the physiological and behavioral indicators, providing a quick assessment of overall classroom engagement. The Activity Level card shows the intensity of student involvement over the last 5 minutes, categorized as High, Medium, or Low, based on movement and

[1] https://emotibit.com/
[2] https://openbci.com/

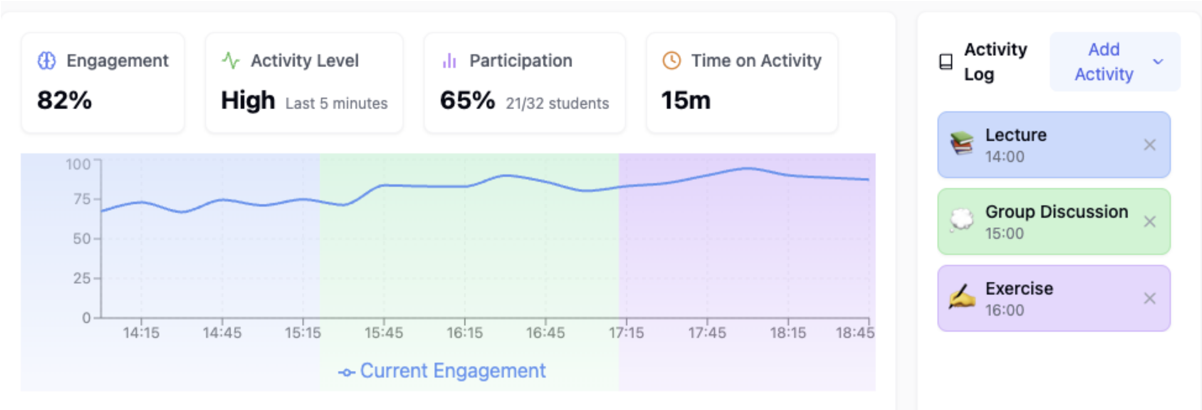

Figure 2: Mockup of a real-time classroom engagement dashboard interface. The visualization presents key engagement metrics and a timeline of student engagement patterns across different learning activities (Lecture, Group Discussion, and Exercise). Color-coded sections in the graph and activity section provide a quick visual correlation between activities and engagement levels.

physiological patterns. The Participation card tracks active involvement through movement patterns and physiological arousal levels. The Time on Activity card monitors the duration of the current learning segment, helping educators pace their sessions effectively. Below these metrics, an interactive timeline visualization demonstrates engagement patterns across different learning activities. The Activity Log component, positioned on the right side of the dashboard, allows educators to mark different learning segments and their timestamps. This feature is crucial for post-session analysis, enabling researchers to study how different teaching modalities affect student engagement. The log can be exported alongside the sensor data for detailed offline analysis, supporting immediate teaching adjustments and longer-term research goals.

In our opinion, the interface layer supports real-time monitoring and facilitates deeper research into the relationship between teaching methods and student engagement. The system's ability to correlate physiological data with specific learning activities provides researchers with valuable insights into the effectiveness of different pedagogical approaches while maintaining an intuitive and accessible interface for classroom use.

4. Discussion and Conclusion

Our system is designed as an accessible research tool for investigators with different academic backgrounds, including psychology, education, computer science, and human-computer interaction. The framework provides a straightforward protocol for data collection and analysis, eliminating the need for extensive technical expertise in sensor technology or signal processing.

Researchers can easily deploy the system in their specific educational contexts for data collection through a simple setup process. The Bluetooth-based connectivity and random device distribution protocol can be implemented without specialized technical knowledge, making it accessible for various research scenarios. The system automatically handles device pairing, data synchronization, and preliminary processing, allowing researchers to focus on their experimental design and observation protocols.

The framework supports both online and offline analysis approaches. For real-time analysis, the system provides immediate feedback through the dashboard interface, enabling researchers to monitor engagement patterns as they emerge and make timely interventions or annotations. The offline analysis capabilities include comprehensive data export features in standard formats, supporting integration with common research tools and statistical software packages. Researchers can access raw sensor data and processed metrics, allowing for custom analyses ranging from basic statistical tests to advanced machine learning applications.

The system's modular architecture allows researchers to focus on specific areas of interest while leveraging the platform's robust data collection and processing capabilities. This approach democratizes access to sophisticated engagement monitoring tools, enabling a wider range of researchers to contribute to our understanding of educational dynamics and student engagement.

This approach is promising for promoting social and cultural inclusion in educational settings, especially for vulnerable populations [15]. The system's ability to detect subtle physiological indicators of engagement and emotional states could be particularly valuable for students with communication difficulties, including those with autism spectrum disorders, non-native speakers, and students with learning disabilities. By providing objective, real-time insights into student engagement patterns, the system can help educators better understand and respond to the needs of students who may struggle to express their learning experiences through traditional channels. Furthermore, the aggregated nature of the data collection ensures privacy while enabling evidence-based adaptations that can benefit diverse learning groups, including elderly learners, children with special needs, and students from disadvantaged backgrounds. Moreover, this framework could be effectively adapted for a variety of contexts, including galleries and museums, where the emotional dimension of the experience plays a crucial role [16].

Understanding engagement's neural and physiological correlates can offer valuable insights into creating learning environments that rekindle students' interest and enhance their learning experiences. By identifying at-risk students, universities can adopt proactive policies to prevent dropouts and significantly increase the number of graduates. Additionally, teachers can leverage Large Language Models to craft novel class materials [17]. The accessibility of our framework to researchers from various disciplines can accelerate this understanding, fostering interdisciplinary collaboration and innovative approaches to educational challenges.

This research thus represents a technological advancement and a step toward more equitable and inclusive educational environments that can better serve all members of our increasingly diverse society. By providing researchers with accessible tools for studying engagement, we enable evidence-based improvements in educational practices that benefit learners across the spectrum of abilities and backgrounds.

References

[1] O. Alrashidi, H. P. Phan, B. H. Ngu, Academic engagement: an overview of its definitions, dimensions, and major conceptualisations., International Education Studies 9 (2016) 41–52.

[2] M. Perkmann, R. Salandra, V. Tartari, M. McKelvey, A. Hughes, Academic engagement: A review of the literature 2011-2019, Research policy 50 (2021) 104114.

[3] S. Abreu Alves, J. Sinval, L. Lucas Neto, J. Marôco, A. Gonçalves Ferreira, P. Oliveira, Burnout and dropout intention in medical students: the protective role of academic engagement, BMC Medical Education 22 (2022) 83.

[4] S. Reyes-de Cózar, A. Merino-Cajaraville, M. R. Salguero-Pazos, Avoiding academic burnout: academic factors that enhance university student engagement, Behavioral Sciences 13 (2023) 989.

[5] C. Gena, F. Cena, F. Vernero, P. Grillo, The evaluation of a social adaptive website for cultural events, User Modeling and User-Adapted Interaction 23 (2013) 89 – 137. doi:10.1007/s11257-012-9129-9, cited by: 16.

[6] B. Woolf, T. Dragon, I. Arroyo, D. Cooper, W. Burleson, K. Muldner, Recognizing and responding to student affect, in: Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction: 13th International Conference, HCI International 2009, San Diego, CA, USA, July 19-24, 2009, Proceedings, Part III 13, Springer, 2009, pp. 713–722.

[7] K. Ahuja, D. Kim, F. Xhakaj, V. Varga, A. Xie, S. Zhang, J. E. Townsend, C. Harrison, A. Ogan, Y. Agarwal, Edusense: Practical classroom sensing at scale, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 3 (2019) 1–26.

[8] H. Wang, C. Gao, H. Fu, C. Z.-H. Ma, Q. Wang, Z. He, M. Li, Automated student classroom behaviors' perception and identification using motion sensors, Bioengineering 10 (2023) 127.

[9] M. Mezzini, C. Limongelli, G. Sansonetti, C. De Medio, Tracking museum visitors through convolutional object detectors, in: Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization, UMAP '20 Adjunct, ACM, New York, NY, USA, 2020, pp. 352–355. doi:10.1145/3386392.3399282.

[10] N. Sardella, C. Biancalana, A. Micarelli, G. Sansonetti, An approach to conversational recommendation of restaurants, in: C. Stephanidis (Ed.), HCI International 2019 - Posters, volume 1034, Springer International Publishing, Cham, 2019, pp. 123–130. doi:10.1007/978-3-030-23525-3_16.

[11] A. Ferrato, C. Limongelli, M. Mezzini, G. Sansonetti, The meta4rs proposal: Museum emotion and tracking analysis for recommender systems, in: Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization, UMAP '22 Adjunct, Association for Computing Machinery, New York, NY, USA, 2022, pp. 406–409. doi:10.1145/3511047.3537664.

[12] M. K. Saini, N. Goel, How smart are smart classrooms? a review of smart classroom technologies, ACM Computing Surveys (CSUR) 52 (2019) 1–28.

[13] I. Arroyo, D. G. Cooper, W. Burleson, B. P. Woolf, K. Muldner, R. Christopherson, Emotion sensors go to school, in: Artificial intelligence in education, Ios Press, 2009, pp. 17–24.

[14] J. Zaletelj, A. Košir, Predicting students' attention in the classroom from kinect facial and body features, EURASIP journal on image and video processing 2017 (2017) 1–12.

[15] D. Macis, S. Perilli, C. Gena, Employing socially assistive robots in elderly care, in: UMAP '22: 30th ACM Conference on User Modeling, Adaptation and Personalization, Barcelona, Spain, July 4 - 7, 2022, Adjunct Proceedings, ACM, 2022, pp. 130–138. URL: https://doi.org/10.1145/3511047.3537687. doi:10.1145/3511047.3537687.

[16] A. Lieto, M. Striani, C. Gena, E. Dolza, A. M. Marras, G. L. Pozzato, R. Damiano, A sensemaking system for grouping and suggesting stories from multiple affective viewpoints in museums, Hum. Comput. Interact. 39 (2024) 109–143. URL: https://doi.org/10.1080/07370024.2023.2242355. doi:10.1080/07370024.2023.2242355.

[17] G. Biancini, A. Ferrato, C. Limongelli, Multiple-choice question generation using large language models: Methodology and educator insights, in: Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization, UMAP Adjunct '24, Association for Computing Machinery, New York, NY, USA, 2024, p. 584–590. URL: https://doi.org/10.1145/3631700.3665233. doi:10.1145/3631700.3665233.