



PAPER

Systematic Insights and Trends in AI-Based Student Engagement Detection: A Systematic Review and Bibliometric Analysis

Shatha Radeef¹  ,
Ayham Zaitouny²,
Negmeldin Alsheikh³,
Shayma Alkobaisi⁴,
Nazar Zaki¹

¹Department of Computer Science and Software Engineering, College of Information Technology, UAE University, Al Ain, United Arab Emirates

²Department of Mathematical Sciences, College of Science, UAE University, Al Ain, United Arab Emirates

³Department of Curriculum and Instruction, College of Education, UAE University, Al Ain, United Arab Emirates

⁴Information Systems & Security, College of Information Technology, UAE University, Al Ain, United Arab Emirates

202090002@uaeu.ac.ae

ABSTRACT

This systematic review critically examines the growing field of artificial intelligence (AI) applications in tracking student engagement and disengagement in educational settings. We synthesize current literature, employing bibliometric analysis to understand the complexities of technology-integrated teaching methods and their effectiveness in creating engaging learning environments. This study employs a rigorous methodological framework, incorporating the preferred reporting items for systematic reviews and meta-analyses (PRISMA) model and the population, intervention, comparison, outcomes, and study design (PICOS) criteria to ensure a structured and comprehensive review. A systematic search strategy was implemented to identify relevant studies from authoritative academic databases. The research findings indicate a significant use of new datasets and virtual learning environments, particularly emphasizing higher education. Despite the promising advancements in AI-driven engagement detection, our analysis reveals critical research gaps, such as the lack of detailed demographic information, especially the age factor that greatly influences engagement behaviors. This absence highlights the need for more specific engagement detection tools suitable for different educational levels. Another key observation is the limited research on early education, a critical area where engagement is crucial yet subtly indicated. Considering these points, we offer recommendations for future research, calling for a comprehensive approach that includes detailed demographics, integration of various learning settings, ensuring broad technology access, improving multimodal techniques, and maintaining privacy and ethical standards. The study's practical implications underscore the need for more adaptable, inclusive, and ethically responsible technological contributions to education, benefiting educators, policymakers, and AI developers.

KEYWORDS

artificial intelligence (AI) in education, classroom engagement analysis, disengagement detection techniques, educational data mining, machine learning in student monitoring, student engagement

Radeef, S., Zaitouny, A., Alsheikh, N., Alkobaisi, S., Zaki, N. (2025). Systematic Insights and Trends in AI-Based Student Engagement Detection: A Systematic Review and Bibliometric Analysis. *International Journal of Emerging Technologies in Learning (iJET)*, 20(3), pp. 19–40. <https://doi.org/10.3991/ijet.v20i03.55133>

Article submitted 2025-02-25. Revision uploaded 2025-06-02. Final acceptance 2025-06-02.

© 2025 by the authors of this article. Published under CC-BY.

1 INTRODUCTION

Student engagement in the classroom is a multifaceted concept that includes both behavioral and emotional components. It represents students' positive, active, and enthusiastic involvement in learning activities, contrasted with disaffection, which involves passive, withdrawn, or negative responses to the learning environment [1]. Student engagement is crucial in education due to its direct correlation with academic achievement and positive developmental outcomes. Teachers often face challenges in maintaining student engagement, evident through various indicators such as attendance issues, incomplete homework, boredom, disaffection, and school drop-out rates [2]. Fortunately, engagement is malleable and can be influenced by situational changes and opportunities. Therefore, understanding and supporting student engagement is a critical area of focus for educators, offering hope for improvement and adaptation in teaching strategies. This systematic review aims to explore existing research on identifying student engagement and disengagement in classroom settings. It delves into the various techniques and methods employed, examining their effectiveness in recognizing any form of student disengagement. A key emphasis of this review is on integrating technology, with a particular focus on artificial intelligence (AI), to facilitate the detection of student disengagement. The primary concentration is on student demographics and the real-time detection of disengagement. Through this comprehensive review, we aim to address critical questions:

- Firstly, how does technology, especially AI, contribute to monitoring student engagement and disengagement in educational environments?
- Secondly, what are the emerging avenues for future research to enhance student engagement and disengagement detection?

The structure of this document is outlined as follows: Section 2 provides a detailed explanation of the methodology employed in this study. In Section 3, we delve into the bibliometric analysis conducted. Section 4 is dedicated to presenting the results and findings of our research. This is followed by Section 5, where we engage in a thorough discussion of the results obtained. Section 6 concludes the document, offering our conclusions.

2 METHODOLOGY

The methodology comprises several steps, which will be explained in detail below.

2.1 Search strings and database

For the formulation of the keywords, the search terms involved two parts: disengagement detection and student participants. As a result, the following search strings were employed: (“disengagement detection” OR “engagement detection” OR “Mind wandering detection” OR “inattentiveness detection” OR “attentiveness detection”) AND (student). Scopus scholarly database was used to retrieve the literature. The Scopus is a comprehensive and expertly curated abstract and citation database that provides access to linked scholarly literature across various disciplines. The database quickly finds relevant and authoritative research, identifies experts, and provides access to reliable data, metrics, and analytical tools. Scopus is a major commercial database that covers scholarly literature from various disciplines and

offers features such as academic journal rankings, author profiles, and an H-index calculator. It contains approximately 71 million items and 1.4 billion references. Several research scholars consider Scopus a high-quality source for contemporary data analyses. It indexes content from more than 25,000 active titles. The database is multidisciplinary, covering a broad range of subjects. Scopus also offers an author lookup feature, allowing users to find and view author profiles [3].

2.2 Inclusion and exclusion criteria

The population, intervention, comparison, outcomes, and study characteristics (PICOS) framework was adopted to create more structured eligibility criteria. The PICOS framework is a structured approach widely used in systematic reviews, especially in educational and healthcare settings. It stands for population, intervention, comparison, outcomes, and study characteristics. The PICOS components are [4]:

Population (P): Population refers to the specific group of subjects of interest. In our study, it represents the students' population.

Intervention (I): Intervention represents the methods for detecting student engagement. These might include technological tools (such as eye-tracking devices and engagement detection software), teacher observation methods, student self-report measures, or other innovative assessment techniques. Although our primary focus is on AI-based methods, we have broadened our search to include various methods, aiming to compare their effectiveness with that of AI-based approaches.

Comparison (C): In some studies, different methods of engagement detection might be compared, such as the effectiveness of technological tools versus traditional observation methods. In other cases, the comparator might be the absence of any specific engagement detection method.

Outcomes (O): The outcomes would be the results of the engagement detection methods. This could include these methods' accuracy, reliability, and effectiveness in detecting student engagement and their impact on educational outcomes, student participation, and learning experiences.

Study design (S): Our focus is mainly on empirical research that relies on real-world, observable, and measurable evidence to investigate a hypothesis or answer a specific question related to detecting student engagement in class.

In educational settings, the PICOS framework helps define the scope of a systematic review clearly. It aids in developing a focused research question and establishing inclusion and exclusion criteria for studies, ensuring that the review is comprehensive and relevant to the targeted educational context [5]. Additionally, it guides the literature search strategy and the analysis of the gathered data, ensuring a methodical approach to synthesizing evidence in the field of education [6]. Accordingly, the following inclusion and exclusion criteria were developed as shown in Table 1.

Table 1. Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
Students' participants	Non-students' participants
The focus is on the majority of students	The focus is on a minority of students, such as students with special educational needs
The intervention is mainly focused on engagement detection	The intervention is focused on any other educational purposes
The study is empirical	The study is not empirical
The document type is a journal or a conference paper	The document type is any other type. Such as book chapters or reviews

2.3 PRISMA model

A systematic literature review was conducted to retrieve information about the methods and approaches used to detect students’ classroom engagement. The preferred reporting items for systematic reviews and meta-analysis (PRISMA) are used to identify relevant literature and synthesize findings to answer the proposed research questions. PRISMA is an evidence-based minimum set of items aimed at helping authors report systematic reviews and meta-analyses. The PRISMA statement is designed to improve the quality of systematic reviews and meta-analyses by providing a checklist of items to be included in the report, such as the search strategy, inclusion criteria, and risk of bias assessment [7]. The identification phase is the initial step in our systematic review process. During this phase, we collected all studies potentially relevant to our research. We searched the Scopus database for articles that met our predefined criteria. As a result, 79 studies were relevant, with one duplicated study identified. In the screening phase, we applied initial screening criteria to the identified studies to determine their relevance to our review topic. This process involved reviewing titles and abstracts to exclude studies that did not meet our inclusion criteria. Consequently, 14 studies were excluded after reviewing their titles and abstracts. The remaining 64 papers underwent full-text review, excluding nine studies due to irrelevant populations, such as adults or non-student groups. Additionally, two studies were excluded due to wrong intervention; the first focused on automated assessment rather than engagement detection, while the other focused on the impact of learning instructions on student engagement. We also considered the type of document in assessing the eligibility of including a document, leading to the exclusion of 11 studies: three were lecture notes, two were preprints, and six were book chapters. Furthermore, three studies were excluded for not being empirical. This phase’s primary goal was to narrow the list of studies to potentially relevant ones for a more detailed examination. Finally, 39 studies were deemed eligible for inclusion according to the inclusion and exclusion criteria discussed in Table 1. The PRISMA diagram illustrating this process is presented in Figure. 1.

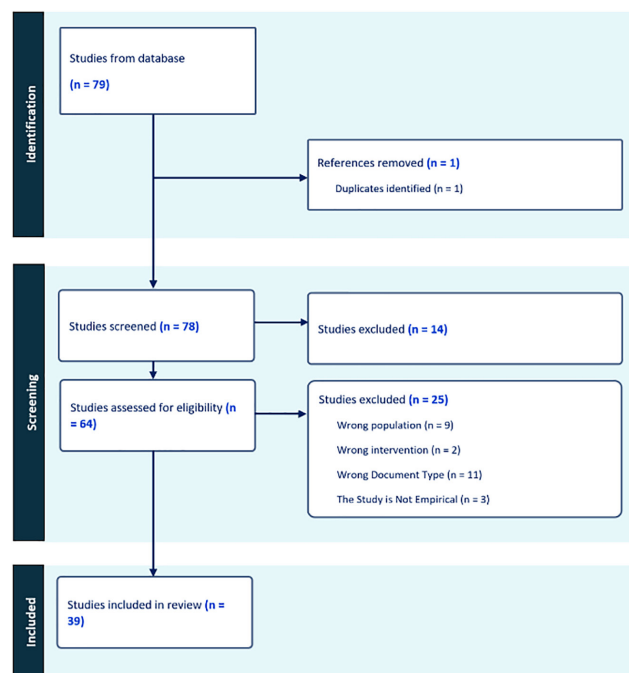


Fig. 1. PRISMA illustrating the data retrieval process

2.4 Data extraction

An extraction form is developed to extract data from selected papers. To ensure consistency, each selected paper is examined using the same criteria, including the objective, educational level of participants, engagement detection method, outcomes, type of learning environment, sample size, data collection method, variables, results, and limitations. Two reviewers independently extracted data from each study to reduce the risk of bias and increase the accuracy of the extracted data. Any discrepancies between the reviewers are resolved by consensus.

3 BIBLIOMETRIC ANALYSIS

Bibliometric analysis is a quantitative approach to assessing and analyzing scientific publications based on statistical methods. It involves evaluating various aspects of academic literature, such as publication and citation counts, authorship patterns, and the evolution of research topics over time. This method helps identify influential researchers, key research trends, and emerging areas within a specific field [8]. The analysis of Google Trends, combined with a synthesis of the literature and a bibliometric examination of the selected articles, is presented in the following sections.

3.1 Google trends analysis

A Google Trends analysis was conducted on selected keywords to assess their global search popularity, emphasizing the terms relevant to this study. The analysis compared three key phrases: “Artificial Intelligence in Education,” “Student Engagement,” and “Educational Data Mining” to gauge interest and trends in these areas. Figure 2 shows a significant increase in web search interest for “Artificial Intelligence in Education” from 2019 to 2023, indicating growing curiosity or relevance in the field. “Student Engagement” has consistently high interest, with a notable jump in 2022 and 2023, reflecting its importance in educational discussions. “Educational Data Mining” shows fluctuating interest with a slight decrease in 2020, then stabilizing, suggesting a steady but niche interest area. Overall, the trends indicate a rising interest in integrating AI into education, a consistent focus on student engagement, and data-driven approaches in educational research and practice.

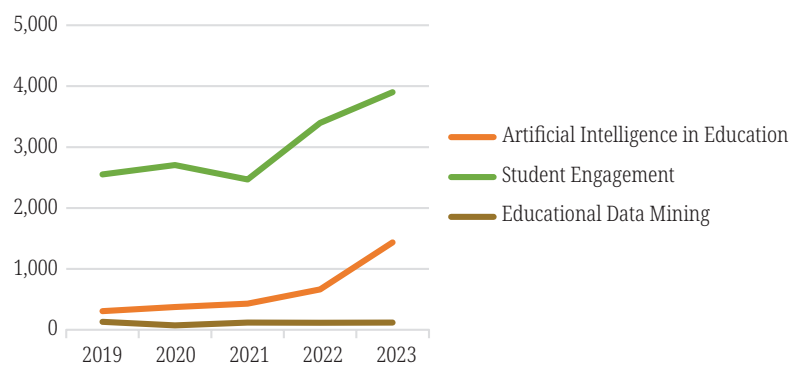


Fig. 2. Web search trends

Figure 3 shows interest over time in “artificial intelligence in education,” “Student Engagement,” and “Educational Data Mining” based on Google Trends data for

news search. Interest in “artificial intelligence in education” and “Educational Data Mining” initially decreased from 2019 to 2021 but significantly rose in 2022 and 2023, with “Educational Data Mining” peaking in 2023. “Student Engagement” saw a drop in 2020, an increase in 2021, and a moderate fluctuation after that. The overall trend indicates growing interest in AI and data mining in education, while interest in student engagement shows variability.

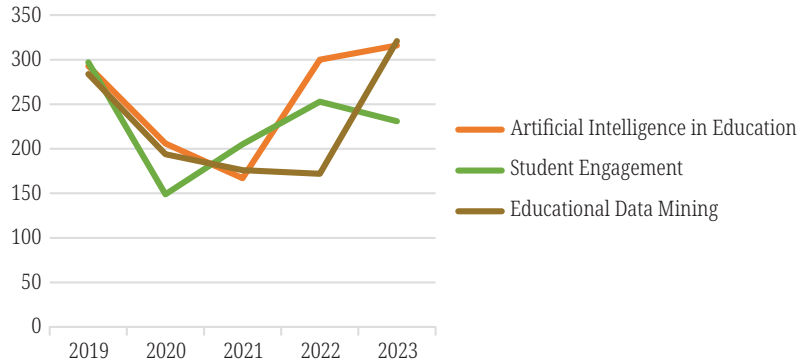


Fig. 3. News search trends

According to Figure 4, image search, interest in “Student Engagement” consistently surpasses the other terms, peaking significantly in 2021. “Artificial Intelligence in Education” shows fluctuating interest, with a notable increase in 2023. “Educational Data Mining” exhibits a steady increase until 2021, followed by a decline. In conclusion, the analysis of Google Trends data reveals a rising interest in the use of artificial intelligence in education, a consistent emphasis on student engagement, and a steady but specialized interest in educational data mining. These trends reflect the evolving focus in education, showcasing the growing importance of technology and data-driven strategies in improving educational practices.

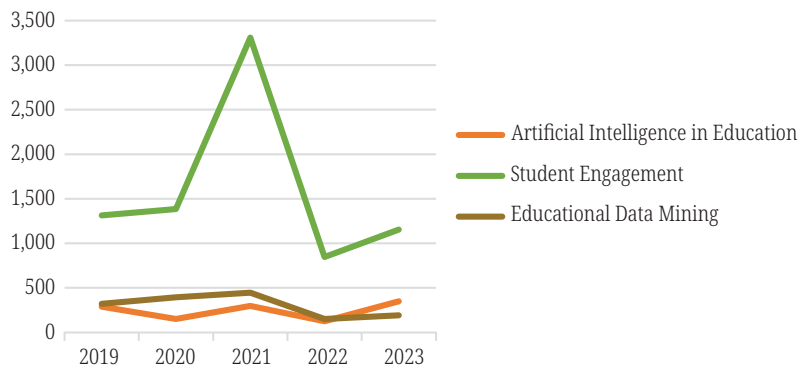


Fig. 4. Image search trends

3.2 Literature synthesis

In this study, the collected articles were analyzed bibliometrically. Below, we detail the specific procedures undertaken during the literature synthesis phase.

Articles classification. Figure 5 shows the classification of articles by subject area, revealing that computer science is the predominant field with 37 articles, suggesting a solid technological focus in AI-based student engagement detection. Engineering

follows with 18 articles, indicating significant interdisciplinary research involving technical design aspects. Social Sciences contributes 10 articles highlighting the role of human and societal factors in the research area. This distribution emphasizes the intersection of technology and education, with a notable presence of social considerations. Notably, there is some overlap in the articles (totaling more than 39), which clearly highlights the interdisciplinary nature of this study topic. There is some overlap in the articles (totaling more than 39), which clearly highlights the interdisciplinary nature of this study topic.

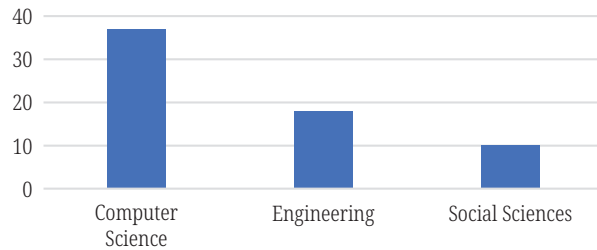


Fig. 5. The collected articles classification based on the subject area

Publication trend (Yearly). Figure 6 outlines the yearly distribution of articles from 2014 to 2023. The number of articles has increased noticeably, with a significant jump to 16 in 2023. This trend suggests growing interest and research output in the field, particularly in recent years, which may reflect AI’s increasing importance and application in educational settings.

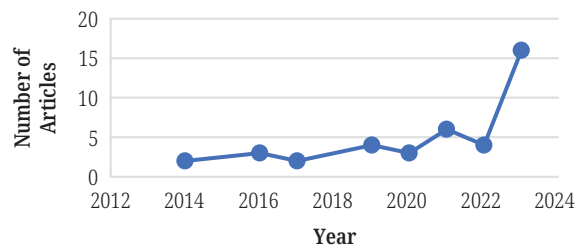


Fig. 6. Publication trends of the reviewed articles over the past 10 years

Article Types. Figure 7 indicates a close distribution between conference papers and journal articles, with 19 conference papers and 20 journal articles analyzed. This suggests a balanced mix of venues where research on AI-based student engagement detection is being presented and published. The substantial number of conference papers might also point to the rapidly evolving nature of this field, where researchers are eager to share and discuss their latest findings.

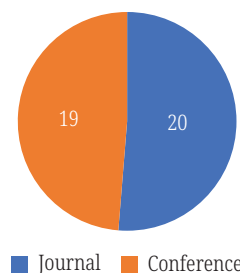


Fig. 7. Distribution of the article types

Top sources. Table 2 highlights that “The ACM International Conference Proceeding Series” is a notable source with three documents contributing to this field and 11 citations overall, suggesting a foundational or pivotal role in AI-based student engagement research. “Education And Information Technologies” has fewer documents (2) but a high citation count (62), indicating the influential nature of the work published there. “Electronics Switzerland” and “IEEE Global Engineering Education Conference Educon” each have two documents with eight and two citations, respectively, pointing to their emerging contribution to the domain. The citation counts reflect the academic impact and recognition within the research community.

Table 2. Top publishing venues in student engagement detection research

Source Title	Documents
ACM International Conference Proceeding Series	3
Education and Information Technologies	2
Electronics Switzerland	2
IEEE Global Engineering Education Conference Educon	2

Keyword analysis and scientometric mapping. This study uses VOSviewer for keyword analysis and scientometric mapping of the collected studies. This software is renowned for delineating connections among academic elements, such as keywords, authors, and institutions. It excels in performing quantitative analysis and presents data in an organized, comprehensible format. The co-occurrence analysis of keywords is conducted through VOSviewer, and the findings are consolidated in Table 3. The analysis begins with 2467 keywords from Scopus, refined to 56 by enforcing a threshold where only keywords appearing in 10 or more articles are included.

Table 3. Keywords co-occurrence analysis

Analysis and Counting Method	Criteria	Outcome
Type of Analysis	Co-occurrence	2467
Counting Method	Full Counting	
Unit of Analysis	Keywords	
Minimum Number of Occurrences of a Keyword	10	56

In addition to the co-occurrence analysis of keywords, Table 4 displays the most frequently occurring terms. The total link strength reflects how strongly the keyword is connected to all other keywords in the network. Based on the total link strength metric from VOSviewer, the most prominent keywords have a total link strength of 100 or more. ‘Students,’ a term used in the search strings of our study, was mentioned 114 times, accounting for approximately 14% of the top keywords. ‘E-learning’ ranks as the second most prevalent keyword, with 104 mentions and a 13% share, while ‘Learning Systems’ is mentioned 86 times, representing 11%. Conversely, the keyword ‘engagement detection,’ also used in our search strings, appears only 23 times, highlighting a scarcity of literature on this topic. This also indicates a primary focus on virtual learning environments as opposed to traditional ones with physically present students.

Co-authorship analysis. The co-authorship analysis, displayed in Table 5, was conducted using VOSviewer. The threshold for inclusion was set at a minimum of five documents authored by an individual. Only 16 authors met this criterion. Given that the articles were primarily sourced from Scopus, these authors are recognized as significant contributors to the field within this repository.

Table 5. Co-authorship analysis

Analysis and Counting Method	Criteria	Outcome
Type of Analysis	Co-authorship	1741
Counting Method	Full Counting	
Unit of Analysis	Authors	
Minimum Number of Documents of an author	5	16

Organizational affiliation analysis. The analysis of the organizational affiliation of the authors sets the inclusion criterion at a minimum of one article per organization and at least five citations per organization, as detailed in Table 6. Among the retrieved articles, 79 organizations are affiliated, with 40 meeting the inclusion criteria.

Table 6. Organizational affiliation analysis

Analysis and Counting Method	Criteria	Outcome
Type of Analysis	Co-authorship	79
Counting Method	Full Counting	
Unit of Analysis	Organization	
Minimum Number of Documents of an Organization	1	
Minimum Number of Citations of an Organization	5	40

Country of origins analysis. The text outlines an analysis of the contributions of various countries to a set of retrieved articles, focusing on the number of documents provided. A minimum criterion of three documents per country was set for inclusion. Out of 20 countries contributing to Scopus, only five met this criterion. The document and citation count of the top contributing countries are summarized in Table 7, and their global distribution is illustrated in Figures 9 and 10 respectively.

Table 7. Number of documents and citations count of top contributing countries

Country	Documents	Citations
India	14	166
United States	7	362
China	6	89
Turkey	4	75
Saudi Arabia	3	18

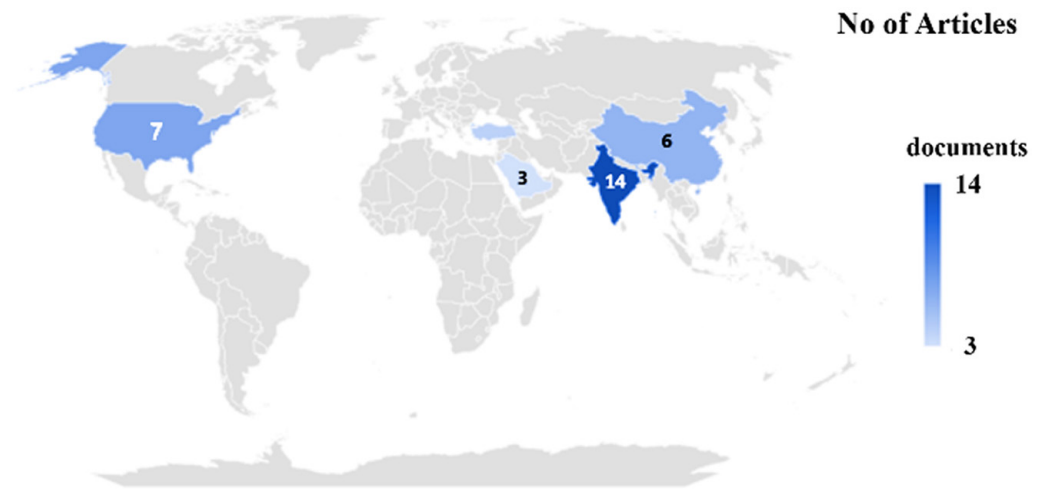


Fig. 9. Country of origins analysis based on number of articles

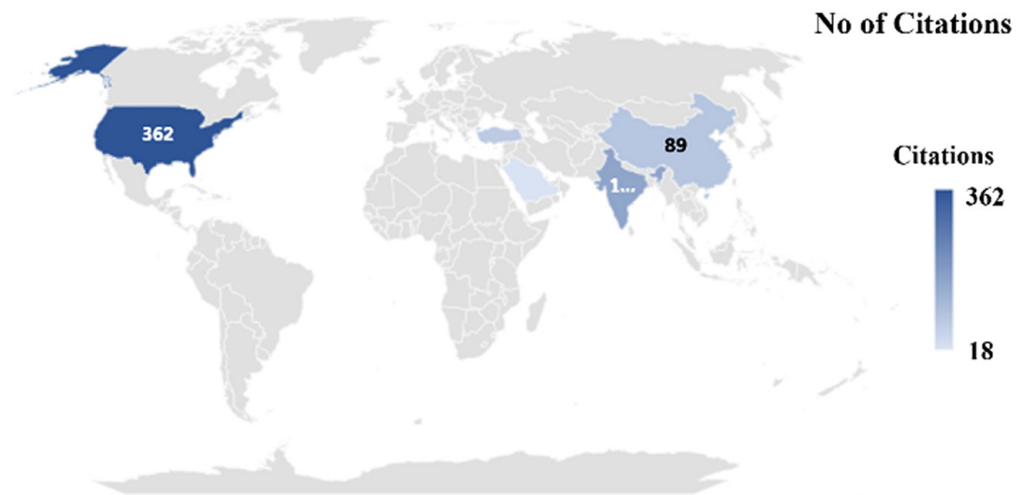


Fig. 10. Country of origins analysis based on number of citations

4 RESULTS AND DISCUSSION

In this section, we aim to present and analyze the findings synthesized from the literature review and discuss these findings. We categorize the engagement detection methods described in the included articles into categories based on the technologies and approaches used. Table 8 offers a consolidated summary of the included literature, emphasizing pivotal shared characteristics across studies. It highlights the learning environment type, predominantly virtual settings, focusing on higher education students. The datasets employed range from novel data collected through direct interaction to existing online video material. A recurring theme in these studies is their challenges, such as small sample sizes, limited demographic information, and the intricate task of accurately interpreting engagement from facial expressions or other observable behaviors.

Table 8. Summary of included literature

Category	References	Learning Environment	Age Group	Dataset	Limitations
Augmented and Assisted Technologies	[9], [10]	Physical	Primary	Novel	Tech dependency, data representatively
		Virtual	College	Common	
Behavioral and Engagement Pattern Analysis	[11], [12], [13], [14]	Virtual	College	Novel	Accuracy, generalizability, data dependency on the quality of used technology
		Virtual	College	Novel	
		Virtual	Unspecified	Common	
		Virtual	College	Novel	
Computer Vision and Image Processing	[15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25]	Virtual	Unspecified	Novel + Common	Generalizability, data diversity, scalability
		Virtual	Unspecified	Common	
		Physical	Unspecified	Novel	
		Physical	Unspecified	Novel	
		Virtual	Unspecified	Common	
		Virtual	Unspecified	Common	
		Virtual	Unspecified	Novel	
		Virtual	Unspecified	Novel	
		Physical	Primary	Novel	
		Virtual	Unspecified	Common	
		Physical	College	Novel	
Machine Learning and Data Analytics	[26], [27], [28], [29], [30], [31], [32], [33]	Virtual	Unspecified	Novel	Data, generalizability, demographics, bias
		Virtual	Unspecified	Novel	
		Virtual	Unspecified	Novel	
		Mixed	Unspecified	Common	
		Virtual	Unspecified	Common	
		Virtual	Unspecified	Novel	
		Virtual	College	Common	
		Virtual	Unspecified	Common	
Multimodal Systems and Fusion Techniques	[34], [35], [36], [37], [38], [39]	Virtual	High School	Novel	Sample size, data complexity, generalizability
		Virtual	College	Novel	
		Physical	College	Novel	
		Physical	Middle School	Novel	
		Mixed	Unspecified	Common	
		Virtual	High School	Novel	

(Continued)

Table 8. Summary of included literature (Continued)

Category	References	Learning Environment	Age Group	Dataset	Limitations
Sensor Technologies and Biometrics	[40], [41], [42], [43], [44], [45]	Physical	High School	Novel	Sample size, privacy, generalizability, resources
		Virtual	College	Novel	
		Virtual	College	Novel	
		Virtual	Unspecified	Novel	
		Mixed	College	Novel	
		Physical	High School	Novel	
Miscellaneous/Unique Approaches	[46], [47]	Virtual	College	Novel	Data Reliability, Emotional Depth
		Virtual	Unspecified	Novel	

Based on Figure 11, it appears that most of the literature on student engagement detection in class utilizes novel datasets, with 27 out of 39 studies employing datasets that are presumably unique or newly created for each study. Meanwhile, 11 studies have used common datasets, which are likely well-established and publicly available resources researchers often use for engagement detection analysis. Only one study has reported using a mixed dataset, indicating a combination of both novel and common data sources. This distribution suggests a significant leaning toward the innovation of new data collection methods and possibly tailored datasets to suit specific research needs in student engagement detection.

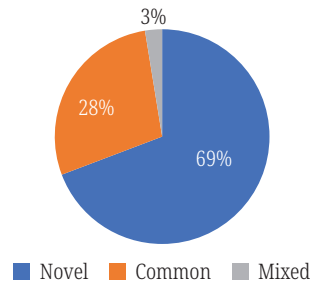


Fig. 11. Articles according to dataset

Moreover, the data from the literature on student engagement detection reveals possible limitations, most notably the significant number of studies that do not specify the age group of participants. This lack of detail suggests a potential oversight in adapting engagement detection methods to various educational levels, as depicted in Figure 12.

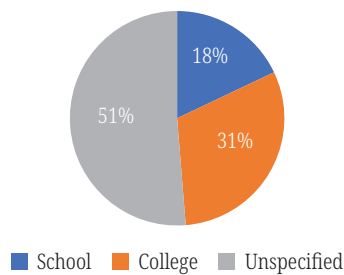


Fig. 12. Articles according to students' age

Furthermore, the skew towards research on college students, with comparatively fewer studies focusing on school-aged children, points to a gap in understanding younger learners' engagement behaviors. Younger students, still developing their ability to articulate their needs, might not benefit as much from methods optimized for older students, who typically express themselves more clearly. This imbalance could lead to a shortfall in engagement detection approaches that are sensitive to the subtler cues of engagement in early education, where they are most crucial for fostering educational development (see Figure 12). The distribution of the studies according to the learning environment suggests a strong preference for research conducted in virtual settings, as seen in Figure 13, with 27 studies collecting data from such environments. This could indicate the growing interest in online learning and the technological advancements that facilitate virtual education research. In comparison, only nine studies focused on physical classrooms, which could reflect the challenges of instrumenting such environments for data collection or perhaps a shift in educational trends towards digital platforms. Studies utilizing a mixed-method approach, combining physical and virtual settings, are the least represented, with only three studies. This underrepresentation may indicate the complexity of designing research that effectively integrates and compares these distinct environments. The dominance of virtual environments in the literature could also suggest potential biases towards the types of engagement and interaction that are more readily observable or measurable online, possibly at the expense of insights that can only be gained in physical settings (see Figure 13).

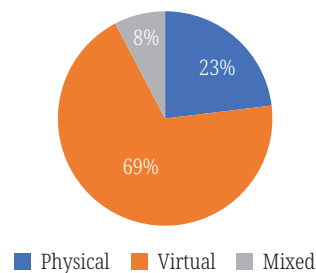


Fig. 13. Articles according to the type of the learning environment

Further details and discussion on each category are provided in the following subsections to answer our research questions. Specifically, to address RQ1—‘How does technology, particularly AI, contribute to monitoring student engagement and disengagement in educational environments?’

4.1 Augmented and assisted technologies

Recent advancements in augmented and assisted Technologies for education have demonstrated significant potential in enhancing student engagement. One such example is a study that developed an innovative system using computer vision, augmented reality, and haptic feedback to bolster student involvement in STEM education at the primary school level. The system successfully integrated these technologies, resulting in improved engagement metrics in a physical classroom setting with a sample of 30 students. It highlighted the effectiveness of combining augmented reality and haptic feedback over traditional learning methods by monitoring and analyzing engagement through students' facial expressions, body postures, and head movements. However, challenges such as the availability of

technology, complexity of integration, and scalability remain potential limitations [9]. Furthermore, using an AI-based system known as iSEEDS in higher education allows for real-time detection of emotions and engagement in virtual learning environments. This system employs convolutional neural networks (CNNs) and accessible datasets such as FER 2013 and Eye Gaze for training. It enables educators to adapt teaching strategies based on immediate feedback on students' emotional states and focus levels. Nonetheless, its limitations include the representativeness of the datasets for diverse student populations and the system's performance in varying real-world classroom conditions [10].

4.2 Behavioral and engagement pattern analysis

The innovative realm of behavioral and engagement pattern analysis within higher education has seen the development of various systems aimed at enhancing online learning experiences. For instance, the E-Learn Detector, a smart monitoring system, has successfully recognized signs of disengagement and cheating by analyzing behaviors such as eye states and unauthorized object presence, providing educators with real-time alerts [11]. On the other hand, a study exploring disengaged reading patterns found that using scrolling data can predict disengagement with a reasonable degree of accuracy, suggesting that simple engagement indicators can be effectively utilized for timely intervention [12]. Additionally, real-time image processing technologies have been applied to recognize student engagement during online classes, achieving substantial accuracy in classifying 'Engaged' and 'Not Engaged' statuses [13]. Furthermore, integrating facial expression and mouse behavior recognition has been shown to enhance the accuracy of engagement detection, underscoring the value of multimodal data in understanding student interactions with digital learning platforms [14]. These systems, however, are not without limitations; they include challenges in behavior detection reliability, potential biases in data interpretation, dependency on consistent webcam quality, and the subjective nature of engagement labeling, which could affect generalizability across different educational settings and populations.

4.3 Computer vision and image processing

In a series of explorations within computer vision and image processing, researchers have sought to refine engagement detection methods in e-learning. These studies have shown exceptional results, such as a method leveraging random forest algorithms achieving 99.2% accuracy in attentiveness detection [15] and a critical frame-based approach surpassing traditional frame selection methods in reflecting online learners' engagement [16]. Deep CNN-based learning approaches have also been validated in classroom settings with an accuracy of 93%, providing reliable student engagement data [17]. In addition, developing a hybrid CNN model has led to 86% accuracy in predicting students' affective states in classrooms by analyzing visual cues [18]. In comparison, implementing a shallow residual CNN has yielded classification accuracies of over 91%, marking significant progress in the field [19]. Furthermore, applying the DAiSEE dataset in a CNN model has enhanced the accuracy of engagement detection models for online learning environments, achieving 77.97% accuracy [20]. Studies have also demonstrated the efficacy of multi-dimensional feature fusion for real-time engagement detection, reaching 62.03% accuracy and validating its effectiveness against NSSE-China survey

scores [21]. Deep learning algorithms developed to analyze students' emotions and engagement levels in real time have provided valuable insights for improving online education quality [22]. In physical classrooms, the ICAPD framework and simAM-YOLOv8n model have shown improved performance in detecting cognitive engagement [23]. Further advancements include the YOLO-v5-based single-step student affect state detection system, which exhibited high precision and recall metrics [24], and a video-based method for estimating facial expressions and heart rate to measure engagement with moderate accuracy levels [25]. While these studies present promising approaches to engagement detection, they consistently highlight limitations such as data generalizability, demographic detail absence, and the variability of environmental factors affecting detection accuracy.

4.4 Machine Learning and Data Analytics

Various machine learning and data analytics methods have been explored to detect and predict student engagement in virtual learning environments. One approach utilized a machine learning algorithm, particularly the CATBoost model, which accurately predicted student engagement based on interaction data within a virtual learning environment (VLE). However, it faced limitations due to the lack of detailed demographic data on participants [26]. Another study incorporated students' personality traits into sensor-free affect detection systems, finding only modest improvements and questioning the cost-effectiveness of including such data [27]. A hybrid algorithm combining particle swarm optimization (PSO) with Naive Bayes was shown to be effective in identifying critical attributes for predicting learner disengagement [28]. Domain adaptation in topic modeling was also employed to enhance engagement detection across datasets, using behavioral cues from video data [29]. A multi-rate Attention-Based GRU model significantly improved engagement prediction accuracy despite potential biases due to uncontrolled recording conditions [30]. Further efforts involved developing a system to predict and visualize student engagement in real-time, aiding educators in dynamic teaching adaptations [31]. The SEPN model highlighted the effectiveness of sequential engagement-based predictions in academic performance, utilizing a comprehensive dataset from the Open University [32]. Lastly, a two-stage algorithm incorporating behavioral and emotional dimensions via CNNs demonstrated high accuracy in engagement detection [33]. These advancements, while promising, often contend with challenges of data completeness, generalizability across different educational settings, and the interpretational complexity of engagement indicators.

4.5 Multimodal systems and fusion techniques

Multimodal systems and fusion techniques are at the forefront of educational technology research, offering novel ways to gauge student engagement and emotional states. A semi-supervised model personalization approach has been shown to significantly enhance the performance of emotional engagement detection systems in ITS, achieving notable F1 measures for different instructional sections [34]. Similarly, a system using multimodal facial cues has achieved impressive engagement detection accuracy, effectively integrating facial expressions, eye movements, and head positions in real-time e-learning environments [35]. Researchers have also explored the efficacy of a Fast-Slow Neural Network (FSNN), combining video and EEG data

to better recognize student engagement by maintaining sequential multimodal data integrity, resulting in superior performance compared to traditional methods [36]. In the context of collaborative game-based learning, multimodal models utilizing facial and textual data have outperformed unimodal models, providing a more comprehensive view of student engagement [37]. A study using a single LSTM model to fuse facial and body features concluded with high accuracy, suggesting the model's potential for capturing the complexities of engagement [38]. Finally, a study assessing a multimodal approach combining appearance, context-performance, and mouse modalities yielded better engagement detection, particularly during digital learning content assessment sections [39]. However, these studies recognize limitations such as small participant samples, generalizability issues, and the computational complexity of real-time multimodal data processing. Integrating multiple data sources and developing innovative architectures suggest a promising future for educational technology, although careful consideration of their practical application remains essential.

4.6 Sensor technologies and biometrics

The innovative intersection of sensor technologies and biometrics has spawned a collection of studies aiming to quantify and improve student engagement and attentiveness in educational settings. Research has yielded a system capable of measuring engagement with 85–95% accuracy using machine learning models, eye-tracking, 2D/3D camera sensors, and eye trackers in high school classrooms [40]. Another study in higher education utilized the XGBoost model to detect attentiveness with a significant AUROC of 92.12%, highlighting the benefit of integrating emotional and non-emotional measures via webcam [41]. A further study developed an affordable webcam-based method to effectively monitor student engagement, achieving a 94% accuracy level in virtual learning environments [42]. WiFi Channel State Information (CSI) has also been used to develop a privacy-preserving approach to estimate engagement through head movement detection, showcasing the potential for non-video-based engagement monitoring [43]. In assessing the efficacy of wearable sensors over machine-based methods, one study demonstrated the ability of a model to predict mental effort, offering insights into students' cognitive states with considerable accuracy, although the study was limited to a single participant [44]. Another high school study successfully applied Commercial Off-The-Shelf (COTS) eye-trackers in noisy classrooms to detect mind wandering, underscoring the potential for enhancing engagement through automated attention tracking [45]. Despite these promising developments, researchers acknowledge limitations, including small sample sizes, a lack of demographic information, and the challenges inherent in generalizing findings beyond the specific learning environments tested.

4.7 Miscellaneous or unique approaches

Emerging from the diverse realm of engagement detection methodologies are unique approaches tailored to the intricacies of online learning. An application focusing on the real-time analysis of students' emotions was developed to assess engagement levels during virtual lectures, successfully identifying various emotional indicators among students in a higher education setting. Despite the complexity of detecting genuine engagement from facial expressions, this system demonstrated

the potential to enhance educational outcomes by providing immediate feedback to educators [46]. Additionally, an engagement detection and intervention service were crafted to discern and bolster student participation in distance learning. Effective in maintaining engagement, this innovative service harnesses real-time analysis to foster improved interaction and participation, albeit with considerations regarding the impact of video quality on the accuracy of facial expression analysis and the system's reliance on physical manifestations of engagement [47]. These studies signify strides in educational technology, offering tools for teachers to dynamically respond to students' emotional states and engagement levels, though they acknowledge the need for broader demographic inclusivity and the nuanced nature of engagement beyond visible cues.

5 CONCLUSION

The systematic review provides a nuanced perspective on current methodologies for detecting student engagement in educational settings, emphasizing the integration of AI technology in pedagogical strategies. Most studies prefer novel datasets and primarily concentrate on college students in virtual learning contexts. A key finding, however, is the apparent omission of age-specific details, potentially obstructing the development of age-appropriate engagement detection tools and questioning the generalizability of results. The apparent scarcity of research on school-aged children suggests an area ripe for future exploration, especially given their distinct needs for engagement and expression. The prevalence of virtual over physical learning environments in the research corpus aligns with the digital shift in education but may introduce biases, underrepresenting the subtleties of traditional classroom engagement. This review addresses our second research question: "What are the emerging avenues for future research to enhance student engagement and disengagement detection?" Based on the synthesized literature, the following recommendations emerge to guide future research directions:

- a) **Detailing demographics:** Future investigations should consistently report participant demographics to enhance the transferability of engagement detection methods across educational stages, with a heightened focus on younger learners.
- b) **Holistic environment integration:** To develop effective engagement detection across diverse learning settings, research should strive for a balanced examination of both physical and virtual environments.
- c) **Universal technology accessibility:** The use of innovative datasets underscores the need for equitable access to engagement detection technologies to ensure broad applicability.
- d) **Advanced multimodal techniques:** To capture the multifaceted nature of engagement, further studies are encouraged to explore multimodal methods that integrate behavioral, emotional, and cognitive aspects.
- e) **Prioritizing privacy:** With the increasing use of biometric and sensor technologies, future research must diligently navigate privacy and ethical considerations.

Future research can address these recommendations, advance our understanding, and contribute meaningful solutions to the nuanced challenges of detecting student engagement. It is imperative that emerging technologies are adaptable, inclusive, and ethically implemented to truly enhance the educational journey for all learners.

6 REFERENCES

- [1] E. A. Skinner, T. A. Kindermann, and C. J. Furrer, "A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom," *Educ. Psychol. Meas.*, vol. 69, no. 3, pp. 493–525, 2009. <https://doi.org/10.1177/0013164408323233>
- [2] J. C. Turner, D. K. Meyer, and H. Patrick, "This issue," *Theory Into Practice*, vol. 50, no. 4, pp. 259–261, 2011. Accessed: Jan. 10, 2024. <https://doi.org/10.1080/00405841.2011.607368>
- [3] Elsevier, "Scopus: A comprehensive abstract and citation database for impact makers," 2023. Accessed: Feb. 26, 2023. [Online]. Available: <https://www.elsevier.com/en-gb/solutions/scopus>
- [4] Critical Appraisal Skills Programme, "How to use the PICO framework to aid critical appraisal," 2024. Accessed: Feb. 12, 2024. [Online]. Available: <https://casp-uk.net/pico-framework/>
- [5] J. M. M. McKeon and P. O. McKeon, "There's more than one way to skin a CAT," *Int. J. Athl. Ther. Train.*, vol. 24, no. 3, pp. 93–94, 2025. <https://doi.org/10.1123/ijatt.2019-0037>
- [6] K. A. Robinson, I. J. Saldanha, and N. A. Mckoy, "Development of a framework to identify research gaps from systematic reviews," *J. Clin. Epidemiol.*, vol. 64, no. 12, pp. 1325–1330, 2011. <https://doi.org/10.1016/j.jclinepi.2011.06.009>
- [7] L. M. Portsmouth, "What is PRISMA guideline & what's new in the 2020 guideline?" *Covidence*, 2023. Accessed: Feb. 26, 2023. [Online]. Available: <https://www.covidence.org/blog/what-is-prisma-whats-new-in-the-2020-guideline-2/>
- [8] R. Todeschini and A. Baccini, *Handbook of Bibliometric Indicators: Quantitative Tools for Studying and Evaluating Research*. Weinheim, Germany: John Wiley & Sons, 2016. <https://doi.org/10.1002/9783527681969>
- [9] H. A. Poonja, M. A. Shirazi, M. J. Khan, and K. Javed, "Engagement detection and enhancement for STEM education through computer vision, augmented reality, and haptics," *Image Vis. Comput.*, vol. 136, p. 104720, 2023. <https://doi.org/10.1016/j.imavis.2023.104720>
- [10] S. K. Vishnumolakala, V. S. Vallamkonda, C. C. Sobin, N. P. Subheesh, and J. Ali, "In-class student emotion and engagement detection system (iSEEDS): An AI-based approach for responsive teaching," in *2023 IEEE Global Engineering Education Conference (EDUCON)*, Kuwait, 2023, pp. 1–5. <https://doi.org/10.1109/EDUCON54358.2023.10125254>
- [11] H. K. T. Bamunuge, H. M. Perera, S. Kumara, P. A. P. Savindri, D. Kasthurirathna, and A. Kugathasan, "E-learn detector: Smart behaviour monitoring system to analyze student behaviours during online educational activities," in *ICAC – 2021 3rd Int. Conf. Adv. Comput.*, 2021, pp. 19–24. <https://doi.org/10.1109/ICAC54203.2021.9671073>
- [12] D. Biedermann *et al.*, "Detecting the disengaged reader – Using scrolling data to predict disengagement during reading," in *LAK23: 13th International Learning Analytics and Knowledge Conference*, 2023, pp. 585–591. <https://doi.org/10.1145/3576050.3576078>
- [13] M. U. Uçar and E. Özdemir, "Recognizing students and detecting student engagement with real-time image processing," *Electronics*, vol. 11, no. 9, p. 1500, 2022. <https://doi.org/10.3390/electronics11091500>
- [14] Z. Zhang, Z. Li, H. Liu, T. Cao, and S. Liu, "Data-driven online learning engagement detection via facial expression and mouse behavior recognition technology," *J. Educ. Comput. Res.*, vol. 58, no. 1, pp. 63–86, 2020. <https://doi.org/10.1177/0735633119825575>
- [15] J. Madake, S. Shende, S. Bhatlawande, R. Shinde, S. Govekar, and S. Shilaskar, "Vision-based monitoring of student attentiveness in an e-learning environment," in *2022 IEEE Pune Sect. Int. Conf. (PuneCon)*, 2022, pp. 1–6. <https://doi.org/10.1109/PuneCon55413.2022.10014782>

- [16] Z. Zhou, D. Pu, and F. Yang, "Engagement detection of online learners based on key frames," in *2022 IEEE Smartworld, Ubiquitous Intelligence & Computing, Scalable Computing & Communications, Digital Twin, Privacy Computing, Metaverse, Autonomous & Trusted Vehicles*, Haikou, China, 2022, pp. 1065–1070. <https://doi.org/10.1109/SmartWorld-UIC-ATC-ScalCom-DigitalTwin-PriComp-Metaverse56740.2022.00157>
- [17] B. Perumal, P. Nagaraj, T. S. Narsimha Charan, Y. V. S. SaiDeepak, C. V. Vignesh Reddy, and S. Nagendra, "Student engagement detection in classroom using deep CNN-based learning approach," in *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2023, pp. 1233–1238. <https://doi.org/10.1109/ICCES57224.2023.10192809>
- [18] T. S. Ashwin and R. M. R. Guddeti, "Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks," *Educ. Inf. Technol.*, vol. 25, no. 2, pp. 1387–1415, 2020. <https://doi.org/10.1007/s10639-019-10004-6>
- [19] M. Thiruthuvanathan, B. Krishnan, and M. Rangaswamy, "Engagement detection through facial emotional recognition using a shallow residual convolutional neural networks," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 2, pp. 236–247, 2021. <https://doi.org/10.22266/ijies2021.0430.21>
- [20] M. M. Santoni, T. Basaruddin, and K. Junus, "Convolutional neural network model based students' engagement detection in imbalanced DAiSEE dataset," *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, vol. 14, no. 3, 2023. <https://doi.org/10.14569/IJACSA.2023.0140371>
- [21] N. Xie, Z. Liu, Z. Li, W. Pang, and B. Lu, "Student engagement detection in online environment using computer vision and multi-dimensional feature fusion," *Multimed. Syst.*, vol. 29, no. 6, pp. 3559–3577, 2023. <https://doi.org/10.1007/s00530-023-01153-3>
- [22] P. Bhardwaj, P. K. Gupta, H. Panwar, M. K. Siddiqui, R. Morales-Menendez, and A. Bhaik, "Application of deep learning on student engagement in e-learning environments," *Comput. Electr. Eng.*, vol. 93, p. 107277, 2021. <https://doi.org/10.1016/j.compeleceng.2021.107277>
- [23] Q. Xu, Y. Wei, J. Gao, H. Yao, and Q. Liu, "ICAPD framework and simAM-YOLOv8n for student cognitive engagement detection in classroom," *IEEE Access*, vol. 11, pp. 136063–136076, 2023. <https://doi.org/10.1109/ACCESS.2023.3337435>
- [24] S. Mandia, F. Mushtaq, K. Singh, R. Mitharwal, and A. Panthakkan, "YOLO-v5 based single step student affect state detection system," in *2023 2nd International Conference on Paradigm Shifts in Communications Embedded Systems, Machine Learning and Signal Processing (PCEMS)*, Nagpur, India, 2023, pp. 1–6. <https://doi.org/10.1109/PCEMS58491.2023.10136090>
- [25] H. Monkaresi, N. Bosch, R. A. Calvo, and S. K. D'Mello, "Automated detection of engagement using video-based estimation of facial expressions and heart rate," *IEEE Trans. Affect. Comput.*, vol. 8, no. 1, pp. 15–28, 2017. <https://doi.org/10.1109/TAFFC.2016.2515084>
- [26] N. Alruwais and M. Zakariah, "Student-engagement detection in classroom using machine learning algorithm," *Electronics*, vol. 12, no. 3, p. 731, 2023. <https://doi.org/10.3390/electronics12030731>
- [27] F. De Moraes and P. A. Jaques, "Improving sensor-free affect detection by considering students' personality traits," *IEEE Trans. Learn. Technol.*, vol. 17, pp. 542–554, 2024. <https://doi.org/10.1109/TLT.2023.3280008>
- [28] T. GopalaKrishnan and P. Sengottuvelan, "A hybrid PSO with Naïve Bayes classifier for disengagement detection in online learning," *Program*, vol. 50, no. 2, pp. 215–224, 2016. <https://doi.org/10.1108/PROG-07-2015-0047>
- [29] A. Kaur, B. Ghosh, N. D. Singh, and A. Dhall, "Domain adaptation-based topic modeling techniques for engagement estimation in the wild," in *2019 14th IEEE Int. Conf. Autom. Face Gesture Recognit., (FG 2019)*, Lille, France, 2019, pp. 1–6. <https://doi.org/10.1109/FG.2019.8756511>

- [30] B. Zhu, X. Lan, X. Guo, K. E. Barner, and C. Boncelet, "Multi-rate attention based GRU model for engagement prediction," in *ICMI '20: Proc. Int. Conf. Multimodal Interact.*, 2020, pp. 841–848. <https://doi.org/10.1145/3382507.3417965>
- [31] S. Tu, "Engagement prediction and visualization in online learning," *IET Conference Proceedings*, vol. 2020, no. 4, pp. 57–62, 2021. <https://doi.org/10.1049/icp.2021.1428>
- [32] X. Song, J. Li, S. Sun, H. Yin, P. Dawson, and R. R. M. Doss, "SEPN: A sequential engagement based academic performance prediction model," *IEEE Intell. Syst.*, vol. 36, no. 1, pp. 46–53, 2021. <https://doi.org/10.1109/MIS.2020.3006961>
- [33] S. Dash, M. A. Akber Dewan, M. Murshed, F. Lin, M. Abdullah-Al-Wadud, and A. Das, "A two-stage algorithm for engagement detection in online learning," in *2019 Int. Conf. Sustain. Technol. Ind. 4.0 (STI)*, 2019, pp. 1–4. <https://doi.org/10.1109/STI47673.2019.9068054>
- [34] N. Alyuz *et al.*, "Semi-supervised model personalization for improved detection of learner's emotional engagement," in *ICMI – Proc. ACM Int. Conf. Multimodal Interact. (ICMI '16)*, 2016, pp. 100–107. <https://doi.org/10.1145/2993148.2993166>
- [35] S. Gupta, P. Kumar, and R. Tekchandani, "A multimodal facial cues based engagement detection system in e-learning context using deep learning approach," *Multimed. Tools Appl.*, vol. 82, no. 18, pp. 28589–28615, 2023. <https://doi.org/10.1007/s11042-023-14392-3>
- [36] L. Zhang, J.-L. Hung, X. Du, H. Li, and Z. Hu, "Multimodal fast–slow neural network for learning engagement evaluation," *Data Technol. Appl.*, vol. 57, no. 3, pp. 418–435, 2023. <https://doi.org/10.1108/DTA-05-2022-0199>
- [37] F. M. Fahid *et al.*, "Effects of modalities in detecting behavioral engagement in collaborative game-based learning," in *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK2023)*, 2023, pp. 208–218. <https://doi.org/10.1145/3576050.3576079>
- [38] Y.-Y. Li and Y.-P. Hung, "Feature fusion of face and body for engagement intensity detection," in *Proc. Int. Conf. Image Processing (ICIP)*, 2019, pp. 3312–3316. <https://doi.org/10.1109/ICIP.2019.8803488>
- [39] N. Alyuz, E. Okur, U. Genc, S. Aslan, C. Tanriover, and A. A. Esme, "An unobtrusive and multimodal approach for behavioral engagement detection of students," in *Proc. ACM SIGCHI Int. Workshop Multimodal Interact. Educ. (MIE 2017)*, 2017, pp. 26–32. <https://doi.org/10.1145/3139513.3139521>
- [40] S. Aslan *et al.*, "Learner engagement measurement and classification in 1:1 learning," in *Int. Conf. Mach. Learn. Appl. (ICMLA)*, 2014, pp. 545–552. <https://doi.org/10.1109/ICMLA.2014.111>
- [41] M. Elbawab and R. Henriques, "Machine learning applied to student attentiveness detection: Using emotional and non-emotional measures," *Educ. Inf. Technol.*, vol. 28, pp. 15717–15737, 2023. <https://doi.org/10.1007/s10639-023-11814-5>
- [42] S. Khenkar and S. K. Jarraya, "Engagement detection based on analyzing micro body gestures using 3D CNN," *Comput. Mater. Contin.*, vol. 70, no. 2, pp. 2655–2677, 2022. <https://doi.org/10.32604/cmc.2022.019152>
- [43] V. K. Singh, P. Kar, A. M. Sohini, M. Rangaiah, S. Chakraborty, and M. Maity, "Monitoring engagement in online classes through WiFi CSI," in *2023 15th International Conference on COMMunication Systems & NETWORKS (COMSNETS)*, Bangalore, India, 2023, pp. 462–465. <https://doi.org/10.1109/COMSNETS56262.2023.10041341>
- [44] N. Bosch, "Detecting student engagement: Human versus machine," in *Proc. Conf. User Model. Adapt. Personalization (UMAP '16)*, 2016, pp. 317–320. <https://doi.org/10.1145/2930238.2930371>
- [45] S. Hutt *et al.*, "Automated gaze-based mind wandering detection during computerized learning in classrooms," *User Model. User-Adapt. Interact.*, vol. 29, pp. 821–867, 2019. <https://doi.org/10.1007/s11257-019-09228-5>

- [46] M. N. Hasnine, H. T. T. Bui, T. T. Thu Tran, H. T. Nguyen, G. Akçapınar, and H. Ueda, “Students’ emotion extraction and visualization for engagement detection in online learning,” *Procedia Comput. Sci.*, vol. 192, pp. 3423–3431, 2021. <https://doi.org/10.1016/j.procs.2021.09.115>
- [47] B. Setyawan, W. Muhamad, and Suhardi, “Detection and intervention engagement service development for new normal distance learning,” in *2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, Surabaya, Indonesia, 2023, pp. 433–437. <https://doi.org/10.1109/ISITIA59021.2023.10221195>

7 AUTHORS

Shatha Radeef is with the Department of Computer Science and Software Engineering, College of Information Technology, UAE University, Al Ain, United Arab Emirates (E-mail: 202090002@uaeu.ac.ae).

Ayham Zaitouny is with the Department of Mathematical Sciences, College of Science, UAE University, Al Ain, United Arab Emirates.

Negmeldin Alsheikh is with the Department of Curriculum and Instruction, College of Education, UAE University, Al Ain, United Arab Emirates.

Shayma Alkobaisi is with the Information Systems & Security, College of Information Technology, UAE University, Al Ain, United Arab Emirates.

Nazar Zaki is with the Department of Computer Science and Software Engineering, College of Information Technology, UAE University, Al Ain, United Arab Emirates.