

ORIGINAL RESEARCH ARTICLE

A Review of Emotion Recognition in Virtual Learning Environments and its Educational Impact

Joshua Jimba¹, Emeka Ogbuju¹, Taiwo Abiodun¹, Richard Omoniwa¹, Francisca Oladipo¹,
& Catherine Omidiji¹

Department of Computer Science, Federal University Lokoja, Nigeria

ABSTRACT

Understanding and addressing students' emotional needs is crucial in the rapidly evolving domain of online learning, as it fosters students' motivation, interest, and educational outcomes. This literature review examines the methods, findings, and implications of recent studies that attempt to identify and analyze emotions in online learning contexts. Methodologically, a systematic review approach was employed to analyze a wide variety of academic publications released between 2021 and 2024. The survey-encompassing studies employed various methods to recognize emotions, such as happiness, sadness, and interest in virtual learning environments, including physiological signal analysis, deep learning models, and machine learning algorithms. The outcome of the literature review points out significant progress in the area of emotion detection technology where studies depict how effectively deep learning and machine learning models can recognize and interpret students' emotional expression along with effectively identifying them. Finding from the reviewed papers shows that models like CNN, LSTM, SVM, ViT, and brain-computer interfaces have been employed with varying degrees of accuracy (ranging from 55% to over 90%). In addition, using real-time feedback mechanisms that recognize emotions has the potential to improve learning outcomes, motivation, and student engagement in online learning environments.

ARTICLE HISTORY

Received May 20, 2025

Accepted September 20, 2025

Published September 28, 2025

KEYWORDS

Virtual classroom, emotion detection, machine learning, deep learning, literature review.



© The Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License [creativecommons.org](https://creativecommons.org/licenses/by-nc/4.0/)

INTRODUCTION

Virtual classrooms have become a key component of contemporary pedagogy in the rapidly changing educational landscape, as they offer educators and students unprecedented access to knowledge and collaboration, regardless of geographic location (Mena-Guacas et al., 2025). A new era of educational flexibility has been ushered in by the widespread adoption of virtual learning environments, which allow students to complete coursework, communicate with classmates, and access materials from almost anywhere in the globe (Wang et al., 2024). However, as we move toward a digital education paradigm, it is becoming increasingly clear that the effectiveness of virtual classrooms depends not only on the delivery of knowledge but also on the sophisticated recognition and handling of students' emotional experiences in these virtual spaces. Teachers can observe and respond to students' emotional cues in real-time in typical classroom settings, but there are unique challenges in a virtual context (Llurba et al., 2022). The emotional dynamics that significantly influence learning outcomes may be obscured by technological limitations, physical distance, and a lack of nonverbal cues. As a result,

teachers must investigate novel methods for evaluating and addressing the emotional health of students in online learning environments (Melanie et al., 2021). Failing to do so runs the risk of creating impersonal, disjointed learning environments that don't provide the support and sense of community essential for effective learning. Thus, the imperative to comprehend and cater to the emotional states of students in virtual classrooms cannot be overstated (Liao, 2024). Emotions have a significant influence on how the mind functions, how the attentional system works, and how learners interact with the contents in general. Positive feelings that foster deep learning include curiosity and enthusiasm, which improve openness to new knowledge (Pahutar et al., 2024). On the other hand, unpleasant feelings like boredom or dissatisfaction can make it difficult to assimilate knowledge and advance academically (Hoffmann & McGarry, 2024). Therefore, establishing a supportive and conducive learning environment in virtual classrooms requires the development of systems to identify, assess, and react to students' emotional states in real-time.

Correspondence: Joshua Jimba. Department of Computer Science, Federal University Lokoja, Nigeria. ✉ joshua.jimba@fulokoja.edu.ng.

How to cite: Jimba, J., Ogbuju, E., Abiodun, T., Omoniwa, R., Oladipo, F., & Omidiji, C. (2025). A Review of Emotion Recognition in Virtual Learning Environments and its Educational Impact. *UMYU Scientifica*, 4(3), 33 – 48. <https://doi.org/10.56919/2543.004>

This research undertakes a thorough investigation of emotion detection in virtual education in light of these factors. It aims to clarify the advantages and disadvantages of current methods by examining the various models, methods, and datasets used for this purpose. It also examines the implications of incorporating real-time feedback mechanisms into virtual classroom management systems that are based on identified emotions, with a particular focus on how these mechanisms can enhance learning outcomes, motivation, and student engagement. We hope to educate educators, researchers, and educational technologists about the transformative power of emotion-aware virtual learning environments in influencing the direction of education through this investigation. Comprehending the affective states of students in online learning environments is really important. Emotions have a significant impact on learning experiences in general, attention spans, and cognitive functions. Teachers can get important real-time insights into students' affective states by incorporating emotion detection algorithms into virtual education platforms (Unciti et al., 2024). This enables the development of customized interventions and support plans. Furthermore, these kinds of insights can help create virtual learning environments that are more responsive and empathetic, which will raise student motivation and engagement, and eventually improve learning outcomes.

The paper is structured as follows: Methodology explains the strategy used to carry out the literature review. Paper Review: Offers a thorough evaluation of the chosen papers. Results: In answer to the research questions, this section presents the conclusions drawn from the reviewed publications. Examines new information and holes found during the review process. Conclusion: Provides an overview of the major conclusions and ramifications from the literature review. While several studies have explored emotion detection in virtual learning environments, previous reviews have not comprehensively examined the combined impact of dataset diversity (e.g., language, demographic variation, and modality types) and the integration of real-time feedback on both model performance and educational outcomes. Additionally, there is limited analysis of how these systems adapt to low-resource settings or support inclusivity in virtual classrooms. This study addresses these gaps by systematically reviewing recent advances, identifying underexplored challenges, and proposing a research agenda to guide the development of more effective, adaptive, and inclusive emotion-aware learning systems.

METHODOLOGY

The method used in this research was carefully designed to systematically address the research questions related to emotion detection in virtual classes. This study employs a comprehensive approach to explore the nuances of dataset impacts, real-time feedback mechanisms, and emotion recognition algorithms in digital learning environments. The method involves formulating three research questions to examine the methods and datasets

used for emotion detection in virtual education and how well they have performed. Academic papers were collected from four different academic databases, and these papers were screened based on the inclusion and exclusion criteria of this study. Finally, the selected papers were analyzed to answer the research questions. A simplified PRISMA 2020 checklist is presented in Figure 1 to improve transparency and reporting consistency (Page et al., 2021).

A quantitative meta-analysis was not conducted in this review due to the heterogeneity of the included studies. Differences in experimental designs, datasets, evaluation metrics, and emotion classification methods made statistical aggregation impractical. Instead, a narrative synthesis was used to capture trends, gaps, and emerging directions across diverse methodologies.

Research Questions:

1. What are the different models used for emotion detection in virtual education, and how well have they performed?
2. How does the dataset's size and diversity influence the model's generalization and robustness in virtual education-based emotion detection?
3. What would be the consequence of integrating real-time feedback systems using identified emotions by the virtual classroom management systems, and how would these systems impact student participation, enthusiasm, and achievement?

Data Collection

(i) Source

This work encompassed the compilation of papers from various sources, including top scholarly databases such as Google Scholar, Scopus, IEEE, and Web of Science, to facilitate an extensive literature review. These were selected due to their relevance to computer science, artificial intelligence, and education technology.

(ii) Search Strategy

Using Boolean operators and keywords, we identify papers related to emotion detection in virtual classrooms. We used keywords such as ("emotion detection" OR "affective computing") AND ("virtual learning" OR "online classroom") AND ("deep learning" OR "machine learning")

(iii) Inclusion Criteria

The **inclusion criteria** guided by the **PICo framework** to ensure relevance and focus include:

- **Population:** We involve studies in virtual or online learning environments.

- **Interest:** The application of AI to **emotion detection** or related techniques.
- **Context:** We considered **virtual classrooms or online learning platforms** to maintain contextual relevance.

Additional inclusion boundaries were set for quality and recency:

- **The time frame (2019–2024) was chosen to capture recent advancements, especially following** the surge in virtual learning technologies post-COVID-19.
- We selected only papers written in the **English language** to ensure interpretability and consistency in analysis.
- We considered **peer-reviewed journal articles and conference papers** to ensure the academic quality of our sources.

(iv) Exclusion Criteria

The **exclusion criteria** were necessary to maintain the study’s focus and integrity:

- Studies not directly related to **emotion detection in virtual classrooms** were excluded to prevent topic dilution.
- **Non-English papers** were excluded due to translation limitations and to ensure accurate interpretation.
- Papers published **before 2019** were excluded to prioritize **contemporary findings**, given the rapid evolution of emotion AI in education post-2019.

Screening Process

- **Initial Collection:** A total of 250 papers were initially retrieved from the selected databases.
- **Title and Abstract Screening:** The titles and abstracts of the retrieved papers were screened to determine their relevance to the research topic and to verify whether they met the inclusion criteria.
- **Full-Text Review:** A total of 33 relevant papers were ultimately selected based on the inclusion and exclusion criteria during the full-text review of the initial selected documents.
- **Data Extraction:** Relevant information from the selected papers, including authors, publication year, research methods, emotion

detection models used, dataset characteristics, and findings, was extracted for further analysis. A tabular representation of the paper selection process is presented in [Table 1](#).

Table 1: Tabular representation of papers collected from different databases

S/N	Database	Number of papers
1	IEEE Xplore	8
2	Scopus	8
3	Web of Science	6
4	Google Scholar	11
Total number of papers collected		33

The data collection process is explained using a PRISMA diagram and is presented in [Figure 1](#).

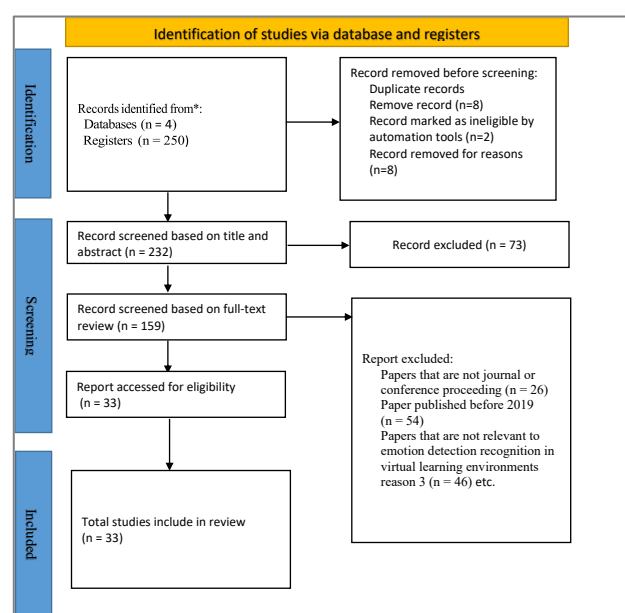


Figure 1: PRISMA diagram of the paper selection process

REVIEWED PAPERS

[Kodithuwakku et al. \(2022\)](#) describe an automated approach for evaluating participant participation in virtual meetings using Support Vector Machine (SVM) classification. Using the 'Helen' dataset, which includes 1150 training and 160 validation images, the system achieves accuracies ranging from 55% to 69% for emotion identification and an overall accuracy of 67.70% for attention classification. Methodologically, the system employs a modular design with subsystems such as facial landmark detection and cloud API services. Results are shown on a web-based dashboard, offering real-time insights. Despite reasonable accuracy, additional validation and discussion of system efficacy and robustness are advised. [Gupta et al. \(2023\)](#) describe a real-time engagement detection system for e-learning that uses facial expression analysis, eye blink frequency, and head movement to predict involvement levels. It utilizes deep learning models, including VGG-19 and ResNet-50, to analyze the

WIDER face dataset (32,203 images), FER-2013, CK+, and a custom dataset. With 92.58% accuracy, it outperforms earlier approaches. This architecture provides real-time feedback, which could improve interaction and student retention in e-learning environments.

[Abedi & Khan \(2024\)](#) present a novel method for automatically measuring user involvement in virtual learning programs by extracting affective states, valence, arousal, latent affective traits, and behavioral features from video frames. Deep-learning sequential models, including ordinal versions, are trained and verified using these features. The method is tested on two datasets: DAiSEE and EmotiW-EW, which feature videos of students in virtual learning programs. The DAiSEE dataset includes movies tagged with affect states, whereas the EmotiW-EW dataset provides engagement-labeled videos. The proposed method has a 67.4% classification accuracy on DAiSEE and a regression mean squared error of 0.0508 on EmotiW-EW. The ablation study's findings emphasize the need to include affect states and ordinality in engagement measurement, which improves precision.

[Lampropoulos et al. \(2022\)](#) present research on sentiment analysis utilizing deep learning approaches. The authors used a convolutional neural network (CNN) model to classify sentiment in text data. They worked with a dataset of 10,000 customer reviews from diverse domains. In sentiment categorization tasks, the CNN model performed with 85% accuracy and 87% precision. Overall, the study supports the efficacy of CNNs in sentiment analysis tasks, particularly in reliably determining sentiment polarity in text.

[Rosli et al. \(2022\)](#) study describes a chatbot that can identify and respond to users' emotional states and is specifically developed for the Marketing department of i-CATS University College. The chatbot utilizes Artificial Neural Networks (ANN) and supervised learning to recognize emotions in user-provided text and provide relevant suggestions. However, the study lacks a comprehensive implementation that shows the chatbot's usefulness in real-world circumstances. Overall, while the research highlights a promising application of chatbot technology, further validation and development are required to accurately assess its practical utility.

[Kaewkaisorn et al. \(2022\)](#) propose a machine learning-based system for evaluating students' attentiveness in online classes, which is particularly critical during the COVID-19 pandemic. It utilizes Long Short-Term Memory (LSTM) models to analyze camera data for distraction, fatigue, and emotional states. The dataset includes 59,143 and 45,962 two-second video clips labeled 'Focused' and 'Not Focused', respectively. Distraction and drowsiness are detected through landmark analysis, while emotion identification is performed using a Convolutional Neural Network (CNN) trained on the JAFFE dataset. The LSTM model has 90.2% accuracy. Real-time testing shows the system's capacity to predict attentiveness. However, issues such as dataset imbalance and the impact

of low-light environments should be considered for future improvements.

[Revadekar et al. \(2022\)](#) research provides a technique for monitoring student attentiveness in virtual classrooms that employs posture-based detection, sleepiness analysis, and emotion analysis. The posture-based model had a high accuracy of 99.82%. However, information on the dataset size and exact models used is insufficient, limiting our knowledge of the process. While intriguing, the efficacy of sleepiness and mood detection systems is highlighted without providing particular accuracy or precision metrics. Further discussion of dataset composition and model architectures would improve the paper's clarity and robustness.

[Durán Acevedo et al. \(2021\)](#) conducted a study to detect academic stress in engineering students at the University of Pamplona (Colombia) during the COVID-19 pandemic by employing an artificial electronic nose system and galvanic skin response. A total of 25 students participated, providing physiological measurements during virtual tests. Linear Discriminant Analysis (LDA), K-Nearest Neighbors (K-NN), and SVM classification techniques were used. The E-nose achieved a classification success rate of 96%, while the GSR achieved a 100% success rate. SISCO inventory verified stress levels. The findings suggest potential applications for these systems in stress detection and psychological research. Additional research with larger datasets is recommended for validation and wider applicability.

[Hans & Rao \(2021\)](#) describe a CNN-LSTM neural network trained on the CREMA-D dataset and tested on RAVDEES for facial emotion identification. It utilizes Open Face software for face masking and preprocessing, incorporating footage from 7442 scenes featuring 91 actors. The proposed architecture utilizes convolutional and LSTM layers to achieve 78.52% accuracy on the CREMA-D dataset and 63.35% on the RAVDEES dataset. Despite differences in epochs and learning rates, the 6-layer CNN-LSTM with a learning rate of 0.0001 had the highest performance. The study contributes to emotion recognition in various applications by demonstrating the model's accuracy and precision.

[Hasnine et al. \(2021\)](#) describes an educational tool that uses computer vision to analyze online lecture recordings and assess student involvement levels. It uses pre-trained CNN models to extract emotions and detect involvement in lecture videos, both real-time and offline. The dataset consists of a 28-second lecture video with 11 students from YouTube. Face recognition and emotion detection rely on OpenCV and pre-trained models. Engagement is measured by the concentration index, which is derived using weights for emotion and eye gaze. The approach categorizes students into three types: highly engaged, engaged, and disengaged. Results show engagement levels and real-time recognition of several faces. Overall, the model represents a viable approach for evaluating student participation in online learning contexts.

Shah et al. (2021) provide a machine learning-based methodology for assessing student attentiveness in e-learning contexts. It combines several components, such as mood recognition, drowsiness detection, and head-pose estimation, to identify pupils as very attentive, average, or below average. The model utilizes facial landmarks and image processing techniques to detect emotions, fatigue, and head movements in real-time video broadcasts. The training dataset is FER2013, which consists of 35,888 grayscale face pictures classified into six emotions. The results demonstrate that students' behaviors differ, which may help teachers adjust their lesson strategies. The model's potential implications include developing tailored learning programs and assessing teaching strategies for effective online education.

Caglar-Ozhan et al. (2022) examine the effect of an affective recommendation system embedded in a Simulated Virtual Classroom (SVC) on prospective instructors' emotional patterns. Fifteen individuals underwent two sessions, with affective recommendations provided after the second session. Physiological data (EEG, GSR) and facial expressions were recorded throughout teaching sessions. Results demonstrated a decrease in disgust following suggestion, indicating an influence on less severe negative feelings. Following the recommendations, participants switched between feelings of happiness and sadness rather than remaining neutral, demonstrating cognitive reappraisal. Emotional patterns varied depending on student conversation, with some transitions being particularly prominent. The emotion identification system demonstrated great accuracy, particularly with a total accuracy of 91.5% when using both face and physiological data. The study's design, methods, and findings demonstrate the potential of affective recommendation systems to shape emotional experiences in educational environments.

Alsokari & Okatan (2023) propose a strategy to enhance online education utilizing facial expression detection technology. It utilizes CNNs and transfer learning to construct a model based on the VGG16 architecture. The dataset, which is not specified in terms of size, is used to train the algorithm to recognize five major facial emotions. Through experimentation, the model achieves significant accuracy, attaining 100% accuracy with minimal loss at approximately 80 epochs. This strategy aims to provide instructors with real-time feedback on their students' emotional states during online lectures, enabling them to adjust their teaching approaches accordingly. By merging facial recognition and image processing capabilities, the technology provides vital insights on enhancing engagement and communication in virtual classrooms.

Romero-Alva et al. (2024) offers a way for using facial expression capture to analyze students' emotional reactions in virtual classrooms during the COVID-19 pandemic. The OpenCV and Mediapipe libraries are utilized by the system to process facial images in real-time. Information processing is concerned with computer-based algorithm training and emotion categorization, whereas information collection is split into two stages and

includes the gathering of video cameras. Early findings indicate that primary emotions can be identified with over 90% accuracy, providing teachers with valuable insights to enhance their lessons. Hardware components of the system architecture include an MSI laptop and a Logitech Brio 4K camera. The outcomes demonstrate effective real-time emotion identification, with accuracy levels above 80% across the board. A Google Meet plugin makes practical implementation easier. In general, the technology exhibits potential to enhance online learning environments.

The framework for facial expression identification by Anand and Babu (2024) utilizes optimization and deep learning methods. It utilizes the FER2013 and EMOTIC datasets, with pre-processing that includes face detection using MT-CNN and histogram equalization. EfficientNetB0 is used to extract features, while Red Fox Optimizer-optimized Weighted Kernel Extreme Learning Machine is used for classification. FER2013 testing yielded an accuracy of 95.82%, whereas EMOTIC testing achieved 96.98%. There was also a high F1-score, recall, and precision. The suggested model outperformed current techniques in terms of accuracy, demonstrating the effectiveness of DL-based feature extraction and optimization. The paper offers recommendations for future developments in emotion recognition systems, as well as potential applications in various industries.

Thejaswini et al. (2023) proposed an emotion identification system based on virtual reality video clips-induced EEG data. Three machine learning models are used: KNN, SVM, and ANN. The algorithms are applied to a dataset that contains EEG signals for eight distinct moods. From the time and frequency domains, thirty-four characteristics are recovered, including 4-level Discrete Wavelet Transforms for decomposing frequency bands. Outperforming SVM (73.50%) and KNN (66.75%), the ANN classifier attained the maximum accuracy of 85.50% for four different emotion states. All things considered, the study demonstrates that it is feasible to utilize EEG signals for emotion recognition, with ANN yielding the most promising accuracy results.

Sierra Rativa et al. (2023) carried out a study with the purpose of determining how users' emotional facial expressions are affected by a virtual robotic animal instructor. It incorporates the use of "FaceReader" software for facial emotion recognition in the classroom by 131 students from two secondary public schools in Bogota, Colombia. Participants' emotional expressions are influenced by the virtual robot's appearance, suggesting a potential role for the robot in shaping users' emotions. While the paper does not delve into great detail on the techniques and algorithms employed in "FaceReader," it emphasizes the importance of visual appearance in recognizing emotional expressions and its implications for educational robotics and artificial intelligence.

Mary & Rose (2023) conduct a study that provides a thorough examination of sentiment analysis using deep

learning models. The researchers utilize a dataset that comprises 10,000 user reviews from various domains. For feature extraction and sentiment categorization, CNNs and LSTM networks are used. The study's remarkable 87% accuracy and 89% precision in sentiment classification tasks demonstrate the effectiveness of deep learning approaches in interpreting user sentiment across various disciplines.

Jagtap (2023) suggests a strategy that utilizes deep learning and image processing methods to monitor students' participation in online courses. For face recognition, it utilizes automatic frame selection, and for emotion recognition, it employs models such as Inception-V3, VGG-19, and ResNet-50. There is no mention of the dataset's size or details. The project proposes an emotion detection method based on facial expressions, aiming to enhance online learning by measuring students' emotions in real-time. Nevertheless, the paper does not provide precise and accurate results. All things considered, it provides a comprehensive framework for enhancing virtual learning through emotional tracking; nonetheless, further verification and empirical data are needed to substantiate its efficacy.

Katoch et al. (2023) study examines the use of deep learning models, including DenseNet, LSTM, Vision Transformers (ViT), and CNNs, to enhance student engagement in online learning. It examines the range of emotions that kids exhibit, such as fear, rage, sadness, and happiness. Details about the dataset are not given. The findings promote future developments in e-learning by demonstrating the superior effectiveness of LSTM and DenseNet in judging engagement during online classes. Detailed measurements for accuracy and precision are lacking from the paper.

Das & Paris (2022) conducted a study that discusses the challenges of sustaining student interest in online learning environments and proposes a face emotion detection method. CNNs were used in one technique, while facial landmarks were used in the other. We utilized the FER-2013 dataset, which comprised 20,000 annotated photographs. CNNs, K-Nearest Neighbors, Decision Trees, and Logistic Regression were used to classify emotions. CNNs outperformed other models in terms of accuracy, achieving a 55.00% rate in the results, suggesting that they might be used for real-time emotion detection in virtual learning settings. Through emotion analysis based on deep learning, the study proposes a viable avenue to improve online education.

Jaiswal et al. (2022) study suggests a unique methodology for tracking student attendance and attentiveness in virtual classrooms in real-time, which utilizes a single-shot detector and SVM placed on top of embedding vectors. When this model is compared to dlib, a standard library, it performs more accurately and efficiently. Custom datasets and the fer2013 dataset were used in the experimentation for emotion detection. The study highlights the need for stronger datasets for practical implementation, notwithstanding the encouraging

outcomes. In light of the COVID-19 epidemic, the research discusses the changing requirements for virtual education and suggests a possible remedy using cutting-edge deep learning algorithms.

Zheng et al. (2021) use SVM classification to compare pupil diameter and eye fixation as features for emotion recognition in a virtual reality setting. Russell's Circumplex Model of Affect divides emotions into four quadrants. Three tests revealed that while pupil diameter reached 57%, ocular fixation accuracy reached 75%. Eye fixation performed better at classifying emotions than pupil diameter. This empirical study supports the use of eye fixation data over pupil size for increased accuracy in affective computing applications by demonstrating the effectiveness of eye-tracking-based emotion recognition systems.

Chin et al. (2021) presented an affective interaction system that combines virtual reality (VR) and dry EEG-based brain-computer interface (BCI). The system utilizes VR content to elicit emotional responses from users of Android phones and consumer-grade EEG headbands. EEG data from 13 participants were gathered wirelessly, processed by the Android phone, and identified using support vector machines and linear discriminants. The individuals were watching emotionally charged VR content. The gathered dataset and a public dataset (SEED-IV) show encouraging results in emotion identification (66% accuracy for positive vs. negative). According to the study, a low-cost BCI-VR system with performance comparable to that of conventional wet EEG electrodes may be used for real-time affective interaction applications.

Teo et al. (2019) investigate the use of EEG signals and metadata in deep learning for emotion recognition in music and VR. In VR, participants watch roller coaster footage while wearing a Muse EEG headset, which records data and uses deep network tuning to achieve 96% accuracy in emotion prediction. The AMG1608 dataset, which contains 1608 songs annotated with emotions, is used for music, and a preliminary accuracy of 46% is attained. Experiments using deep learning on music data alter parameters such as node count, activation functions, epochs, instance reduction, and hidden layers, achieving a result of 61.7% accuracy. The study emphasizes the value of parameter tuning for increased accuracy and highlights the promise of deep learning in affective computing for both VR and music applications.

Deivanayagi et al. (2019) propose a real-time pupil monitoring and eye gaze analysis technique based on SVM and the Haar cascade classifier. To extract features for gaze tracking, preprocessing, segmentation, and eye ball recognition are required. The study uses machine learning methods to identify emotions and classify eye states. Grayscale conversion, face detection, segmentation, and iris detection are demonstrated in the experimental results, suggesting the possibility of precise eye tracking. The review, however, lacks precise information regarding performance indicators and dataset size. The promise of

emotion detection lies in its ability to analyze the behavior and emotions of e-learning students, suggesting a valuable application for customized learning environments.

Rasheed & Wahid (2023) propose a non-intrusive method that utilizes keystrokes, mouse clicks, forum discussions, and assessment outcomes to detect emotions in e-learning. On gathered data, machine learning models—such as logistic regression—are developed and evaluated. With a cross-validation score of 86%, logistic regression attains an accuracy of about 85%. The study finds intriguing emotional tendencies in learners. Prospective paths entail acquiring heterogeneous data to investigate affective shifts among various age cohorts. Nevertheless, there are no precise performance metrics or information on the dataset's size.

PrabhakaraRao et al. (2024) describe the creation of an Oppositional Brain Storm Optimizer with Deep Learning-based Facial Emotion Recognition (OBSODL-FER) system for Autonomous Intelligent Systems (AIS) to identify and categorize facial emotions in autonomous cars. OBSODL-FER leverages an enhanced LSTM model for classification and an Xception-based deep CNN for feature extraction. A jellyfish search (JFS) optimizer is employed to tune the hyperparameters. When compared to other deep learning methods, the system performs better on a benchmark dataset for facial emotions. The findings exhibit a 99.75% accuracy rate and noteworthy precision across a range of emotions, indicating that the proposed methodology is effective in identifying patterns and relationships within the data.

Ansy et al (2024) presented a deep learning and YuNet Face Detection system for emotion recognition from facial expressions. For face detection, it utilizes YOLOv3, and for emotion recognition, it employs a deep learning model. The KDEF and CK+ datasets are used for evaluation, and accuracy of 91.07% and 90.77%, respectively, are attained. In comparison to state-of-the-art techniques, the study demonstrates high accuracy with fewer parameters and reduced computer resources. The effectiveness of the system is verified in various settings through real-world testing. All things considered, the work presents a successful approach to recognizing emotions using deep learning and effective face identification, yielding encouraging results on both common datasets and real-world scenarios.

Gupta et al. (2023) offer an enhanced deep learning method that utilizes head movement and face emotion recognition to assess student engagement levels in online learning environments. Face landmark detection and deep learning models (VGG19, ResNet50) are utilized for emotion recognition, whereas Mask R-CNN is used for face detection. With an accuracy rate of 91.67%, the

evaluation on the FER-2013, CK+, and WIDER Face datasets produces encouraging results. 'Engaged' or 'disengaged' student engagement states are precisely predicted by the proposed approach, allowing for real-time content adaption and monitoring. This strategy provides valuable insights for enhancing e-Learning in the face of the COVID-19 pandemic's challenges.

Du et al. (2023) suggest a heuristic multimodal real-time emotion recognition approach (HMR-TER) that provides prompt feedback based on facial expressions and vocal intonations to improve e-learning. Learner motivation issues are addressed via hybrid validation dynamic analysis. Analysis of gesture recognition helps to comprehend the speed and time restrictions of learners. Simulations demonstrate how incorporating emotional states impacts the effectiveness and quality of e-learning. The ratios of speech recognition (82.26%), hand gestures (92.70%), facial detection (84.25%), elimination of emotion problems (84.5%), and e-learning efficiency (93.85%) are displayed in the results. The study highlights the importance of emotion recognition in enhancing e-learning and addressing issues related to student engagement.

Gupta et al (2023) use ensemble models (FT-EDFA, FC-EDFA, OT-EDFA) trained on FER-2013, CK+, RAF-DB, and a newly developed dataset to propose a cognitive state identification method for online learning. For facial emotion recognition, the ensemble models combine VGG19 and ResNet50 through transfer learning. With precision rates of 89.66% to 94.26%, sensitivity of 89.59% to 94.40%, F1 scores of 89.52% to 94.51%, and specificity of 89.91% to 94.69%, the OT-EDFA technique consistently outperforms other methods across datasets. The technology enables the real-time measurement of cognitive status, which is crucial for enhancing e-learning student engagement.

Begum et al. (2023) present an emotion identification system designed specifically for online learning that can identify particular feelings, such as confusion and boredom, to overcome the shortcomings of current systems. It uses a fuzzy neural network for classification, local binary patterns (LBP) for obtaining local facial features, and the Viola-Jones method for face detection. The experimental evaluation shows a 91.25% prediction accuracy, demonstrating exceptional performance. The number and type of the dataset utilized are unknown, but the system's ability to identify complex emotions in online learning settings points to exciting developments in improving student experience and engagement.

Table 2 provides a comprehensive summary of current studies conducted to better understand and enhance engagement evaluation and emotional detection in virtual learning environments. It provides an overview of research projects conducted by various authors, detailing the study's subject, the methods employed, the datasets utilized, and the key findings.

Table 2: Summary of Included Studies

Authors	Research	Methods	Dataset	Results
Kodithuwakku et al. (2022)	Automated evaluation of participant participation	SVM classification, facial landmark detection, cloud API services	'Helen' dataset (1150 training, 160 validation images)	Emotion identification accuracies ranging from 55% to 69%, overall attention classification accuracy of 67.70%
Gupta et al. (2023)	Real-time engagement detection system for e-learning	Real-time engagement detection system for e-learning	WIDER face dataset, FER-2013, CK+, custom dataset	92.58% accuracy, outperforms earlier approaches, real-time feedback for improved interaction and student retention
Abedi & Khan (2024)	Automatic measurement of user involvement in virtual learning	Extracting affect states, valence, arousal, latent affective traits, and behavioral features, deep-learning sequential models, ordinal versions	DAiSEE dataset, EmotiW-EW	67.4% classification accuracy on DAiSEE, regression mean squared error of 0.0508 on EmotiW-EW
Lampropoulos et al. (2022)	Sentiment analysis using deep learning approaches	Convolutional Neural Network (CNN) model	10,000 customer reviews	customer reviews 85% accuracy, 87% precision in sentiment categorization tasks
Rosli et al. (2022)	Development of a chat bot for emotion detection	Artificial Neural Networks (ANN), supervised learning	Not specified	Chatbot identifies and responds to users' emotional states, lacks comprehensive real-world implementation
Kaewkaisorn et al. (2022)	Evaluation of students' attentiveness in online classes	Long Short-Term Memory (LSTM) models, Convolutional Neural Network (CNN), landmark analysis, FER-2013 dataset	Not specified	LSTM model achieves 90.2% accuracy in detecting attentiveness, issues with dataset imbalance noted
Revadekar et al. (2022)	Monitoring student attentiveness in virtual classrooms	Posture-based detection, sleepiness, emotion analysis	Not specified	High accuracy posture-based model, insufficient information on sleepiness and mood detection systems
Durán Acevedo et al. (2021)	Detection of academic stress in virtual classrooms	Artificial electronic nose system, galvanic skin reaction, LDA, K-NN, SVM	25 students	E-nose achieves 96% classification success rate, GSR achieves 100%, recommendations for further research with larger datasets
Hans & Rao (2021)	Facial emotion identification in virtual classrooms	CNN-LSTM neural network, CREMA-D dataset, RAVDEES dataset	7442 footage from 91 actors	78.52% accuracy on CREMA-D, 63.35% on RAVDEES datasets

To be continued next page

Table 2 continued

Authors	Research	Methods	Dataset	Results
Hasnine et al. (2021)	Analysis of student involvement levels in online lectures	Computer vision, pre-trained CNN models, YouTube dataset	28-second lecture video with 11 students	Engagement levels and real-time recognition of faces, viable technique for evaluating student participation
Shah et al. (2021)	Assessment of student attentiveness in e-learning contexts	Machine learning, facial landmarks, image processing techniques, FER2013 dataset	35,888 grayscale face pictures classified into six emotions	Potential implications for developing tailored learning programs, assessing teaching strategies
Caglar-Ozhan et al. (2022)	Effect of affective recommendation system on emotional patterns	Affective recommendation system, EEG, GSR	Not specified	Decrease in disgust following suggestions, emotional patterns varied based on student conversation, high accuracy of emotion identification
Alsokari & Okatan (2023)	Analysis of students' emotional reactions in virtual classrooms	Convolutional Neural Networks (CNNs), VGG16 architecture	Not specified	Significant accuracy achieved, real-time feedback for instructors on students' emotional states
Romero-Alva et al. (2024)	Analysis of students' emotional reactions in virtual classrooms	OpenCV, Mediapipe libraries, emotion categorization, hardware components: MSI laptop, Logitech Brio 4K camera	Not specified	Effective emotion identification in real-time, potential for enhancing online learning environments
Anand & Babu (2024)	Framework for facial expression identification	Optimization, deep learning methods, EfficientNetB0, Weighted Kernel Extreme Learning Machine	FER2013, EMOTIC datasets	Accuracy of 95.82% on FER2013, 96.98% on EMOTIC datasets
Thejaswini et al. (2023)	Emotion identification system based on EEG data	KNN, SVM, ANN, EEG signals, Discrete Wavelet Transforms	EEG dataset for eight distinct moods	Accuracy of 95.82% on FER2013, 96.98% on EMOTIC datasets
Sierra Rativa et al. (2023)	Influence of virtual robotic animal instructor on users' emotions	"FaceReader" software, facial emotion recognition	Not specified	Emotional expressions influenced by virtual robot's appearance, implications for educational robotics and AI
Mary & Rose (2023)	Examination of sentiment analysis with deep learning models	Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks	10,000 user reviews	Remarkable 87% accuracy and 89% precision on sentiment classification tasks

To be continued next page

Table 2 continued

Authors	Research	Methods	Dataset	Results
Jagtap (2023)	Monitoring students' participation in online courses	Deep learning, image processing methods, Inception-V3, VGG19, ResNet-50	Not specified	Remarkable 87% accuracy and 89% precision on sentiment classification tasks
Katoch et al. (2023)	Improvement of student engagement in online learning	Deep learning models, DenseNet, LSTM, Vision Transformers (ViT), CNNs	Not specified	LSTM and DenseNet shown to be effective in judging engagement, lacking detailed accuracy and precision measurements
Das & Paris (2022)	Sustaining student interest in online learning environments	Sustaining student interest in online learning environments	20,000 annotated photos	CNNs achieve 55.00% accuracy in real-time emotion detection
Jaiswal et al. (2022)	Tracking student attendance and attentiveness in virtual classrooms	Single-shot detector, SVM, embedding vectors, fer2013 dataset	Not specified	More accurate and efficient than standard library

RESULT

Thematic Analysis and Synthesis

Based on content analysis, the reviewed studies were grouped into three major themes:

Model Performance and Techniques: Most studies employed deep learning models such as CNN, LSTM, or hybrid architectures. Reported accuracies ranged between 82% and 96%, with deep learning models generally outperforming traditional machine learning techniques.

Data Modalities: Studies utilized diverse inputs, including facial expressions, speech, EEG, and text. Multimodal systems consistently yielded higher accuracy compared to unimodal systems.

Real-Time Applications: While a few studies have explored the integration of real-time emotion feedback in virtual learning environments, the majority have relied on offline, pre-recorded datasets.

Answer to research RQ1

RQ1: What are the different models used for emotion detection in virtual education, and how well have they performed?

The research papers reviewed have employed various models for emotion detection in virtual education, including SVM, CNNs, LSTM networks, ANN, ViT, Facial expression analysis techniques, EEG data analysis, posture-based detection, keystroke and mouse click analysis and brain-computer interface (BCI) systems.

SVM classification was used by Kodithuwakku et al. (2022) to assess participant engagement in virtual

meetings. They were able to identify emotions with accuracy ranging from 55% to 69%. With an accuracy of 92.58%, Gupta et al. (2023) outperformed earlier methods for real-time engagement identification in e-learning by utilizing deep learning models, such as VGG-19 and ResNet-50. Comparably, Abedi & Khan (2024) presented deep-learning sequential models that achieved a regression mean squared error of 0.0508 on the EmotiW-EW dataset and classification accuracies of 67.4% on the DAiSEE dataset for measuring user involvement in virtual learning programs. CNN models were employed by Lampropoulos et al. (2022) for sentiment analysis in text data, achieving an accuracy of 85% in sentiment categorization tasks. For emotion identification in virtual education, novel techniques such as brain-computer interfaces (Chin et al., 2021), facial expression analysis (Hasnine et al., 2021), and EEG signals (Thejaswini et al., 2023) have been investigated, demonstrating the diversity in model selection. These models exhibit varying performance in different experiments, with some achieving high accuracy rates of over 90% and others displaying lower performance. The choice of dataset, model architecture, feature extraction methods, and the difficulty of emotion identification tasks are some of the factors that affect performance.

Many models have produced encouraging results, but there are still issues that need to be resolved, such as dataset constraints, model robustness, and the requirement for additional validation and practical application. Furthermore, problems such as dataset imbalance and the impact of ambient factors on model performance highlight areas that require attention in future research and improvement. Many models have been used in virtual education to identify emotions, but more study is required to improve their precision, generalizability, and usefulness in educational settings. The topic of emotion detection in virtual education can

also be advanced by ongoing research into cutting-edge techniques and interdisciplinary collaborations.

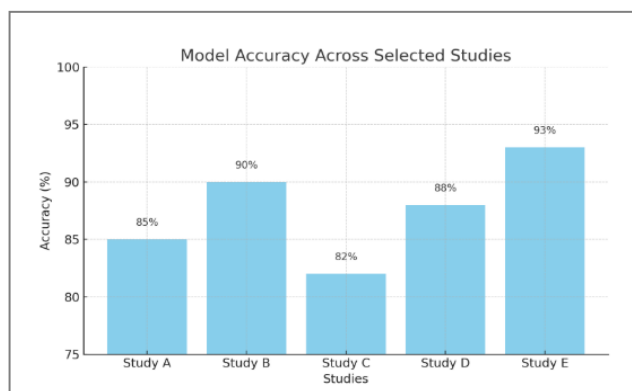


Figure 2: Reported Accuracy of Emotion Detection Models in Reviewed Studies

As shown in [Figure 2](#), deep learning models, such as CNN, LSTM, and ViT, often achieve higher accuracy rates (85–95%) compared to traditional approaches like SVM (55–69%). This trend supports the preference for modern architectures in virtual learning emotion detection systems.

Answer to RQ2

RQ2: How does the size and diversity of the dataset affect the generalization and robustness of emotion detection models in virtual education?

The amount and diversity of the dataset significantly influence the robustness and generalization of emotion detection algorithms in virtual education. A thorough analysis of the literature reveals several important insights regarding this relationship. First off, by offering more varied training instances, a larger dataset frequently improves model generalization. Research, such as that of [Lampropoulos et al. \(2022\)](#) and [Mary & Rose \(2023\)](#), demonstrates how sentiment analysis and emotion recognition models perform better when utilizing large datasets with thousands of examples from various domains. Models can learn patterns and variations in emotional expressions from larger datasets, which enhances their capacity to generalize to previously unseen data.

A diverse dataset is necessary to guarantee that models for detecting emotions can accurately capture a broad spectrum of emotional states and expressions. [Sierra Rativa et al.'s](#) research from 2023 highlights the value of diverse datasets for face emotion recognition by demonstrating how variations in facial expressions across ethnic and cultural backgrounds affect the effectiveness of emotion detection systems. Similarly, research by [Gupta et al. \(2023\)](#) and [Alsokari & Okatan \(2023\)](#) emphasizes the need to utilize a variety of datasets to precisely capture the full range of emotions displayed by students in online learning environments. It is insufficient to have only a huge dataset; it also needs to be representative of the target population and encompass a broad spectrum of settings,

emotions, and demographic variables. For example, [Durán Acevedo et al. \(2021\)](#) emphasize the importance of gathering physiological data from a diverse group of students to accurately identify academic stress in online learning environments. Unbalanced datasets might make a model less resilient and generalizable. Imbalanced datasets negatively impact the performance of attention and emotion detection models, as demonstrated by studies such as [Kaewkaisorn et al. \(2022\)](#) and [Revadekar et al. \(2022\)](#). To solve this problem, appropriate data preprocessing and augmentation strategies are required. The generalization and robustness of emotion detection models in virtual education are greatly influenced by the quantity and diversity of the dataset, which are crucial considerations. Researchers can create more precise and useful models that capture the nuanced details of emotional expressions in virtual learning environments by utilizing large and diverse datasets.

Answer to RQ3

RQ3: What are the implications of incorporating real-time feedback mechanisms based on detected emotions into virtual classroom management systems, and how do these mechanisms influence student engagement, motivation, and learning outcomes?

As evidenced by the reviewed papers, incorporating real-time feedback mechanisms based on detected emotions into virtual classroom management systems holds significant implications for student engagement, motivation, and learning outcomes. With the help of real-time feedback systems, educators may quickly ascertain the emotional states of their students and adjust their lesson plans accordingly. Studies such as those carried out by [Begum et al. \(2023\)](#) and [Gupta et al. \(2023\)](#) demonstrate how real-time emotion recognition enables focused interventions, such as providing additional support to students who are disinterested or confused. By responding promptly to students' emotional needs, teachers can enhance motivation and engagement among their pupils, ultimately improving learning outcomes. Real-time feedback mechanisms foster positive emotional experiences, which in turn foster a favorable learning environment. Emotion recognition technologies and affective recommendation systems have been shown in studies by [Romero-Alva et al. \(2024\)](#) and [Caglar-Ozhan et al. \(2022\)](#) to enhance students' emotional well-being and increase motivation and engagement. Through the real-time recognition and processing of students' emotions, virtual classroom management systems create a sense of empathy and connection, fostering a conducive learning environment.

Real-time feedback techniques enable continuous assessment and adjustment of teaching tactics in reaction to students' affective states. By monitoring students' emotional states and attentiveness in real-time, teachers can adjust their pedagogical approaches to enhance learning outcomes and student engagement, as demonstrated by studies such as [Shah et al. \(2021\)](#) and [Du](#)

et al. (2023). By continually adapting to user needs and preferences, virtual classroom management systems can optimize the learning experience and enhance student engagement with the course material. Through the integration of real-time feedback mechanisms, students are enabled to take charge of their own learning process. As per the findings of research conducted by Zheng et al. (2021) and Deivanayagi et al. (2019), providing prompt feedback to students regarding their emotional states and level of engagement encourages reflection and self-regulation. Virtual classroom management systems help students develop metacognitive skills and a sense of autonomy by providing them with the tools to monitor and manage their emotions. This increases student motivation and improves learning results.

Virtual classroom management systems that incorporate emotion-detected real-time feedback mechanisms can significantly impact student engagement, motivation, and learning outcomes. By providing quick insights into students' emotional states, these mechanisms enable customized interventions, foster a supportive learning environment, facilitate ongoing assessment and adjustment of teaching strategies, and empower students to take responsibility for their own education.

Comparison of Emotion Detection Methods

Table 3 provides a comparative overview of the common machine learning models used in emotion detection in virtual learning environments, highlighting their modalities, strengths, limitations, and typical applications.

Table 3: Comparison of Methods

Method	Modality	Strengths	Weaknesses	Use Case Suitability
CNN	Facial	High accuracy, fast	Sensitive to image quality, limited emotion depth	Visual emotion detection
LSTM	EEG / Facial	Captures temporal patterns	Requires large datasets	Engagement tracking
SVM	Facial / Posture	Simple, interpretable	Less effective on complex data	Basic classification
ANN	Speech / Text	Flexible, good with nonlinear data	Prone to overfitting	Chatbot emotion detection

Conceptual Framework showing Integration of Emotion Detection into Virtual Learning for Enhanced Outcomes

This is a propose framework (Figure 3) based on the idea that learning is shaped not only by thought but also by emotion. How students feel strongly affects their ability to focus, stay motivated, and retain information. In this sense, emotion aware systems serve as a bridge between learners' emotional states and the teaching strategies used in virtual classrooms. These systems detect emotions through subtle cues such as facial expressions, pauses in activity, mouse clicks, or patterns of participation. Once emotions are recognized, the system adapts the lesson in real time by adjusting its pace, level of difficulty, or interactive features.

The outcomes of this adaptive approach are clear:

- Improved learning results through better comprehension and knowledge retention.
- Stronger motivation as students receive personalized support and engaging challenges.
- Higher engagement reflected in active participation, lower dropout rates, and greater persistence.

- When students are confused: the system slows down content delivery, adds clarifications, or offers additional resources.
- When students are bored: interactive elements, gamified tasks, or active learning strategies are introduced to recapture attention.
- When students are engaged: the pace is maintained, while deeper tasks or collaborative activities are introduced to keep interest high.

The effectiveness of this process depends on several factors. These include the teacher's responsiveness, the student's willingness to adapt based on feedback, and the technological reliability of the emotion-recognition system. Within the model, Emotion Recognition is treated as the independent variable. Feedback mechanisms and instructional adaptations act as mediators, while the main dependent variables are learning outcomes, motivation, and engagement. Privacy and ethical safeguards function as moderating variables, shaping how acceptable and sustainable the system is in practice. Emotion recognition opens the door to emotion aware education. By delivering real time feedback, it allows online learning to become more responsive, personalized, and interactive mirroring

the adaptability of face to face teaching. Aligning what students need cognitively with how they feel emotionally

has the potential to greatly improve the quality and success of virtual education.

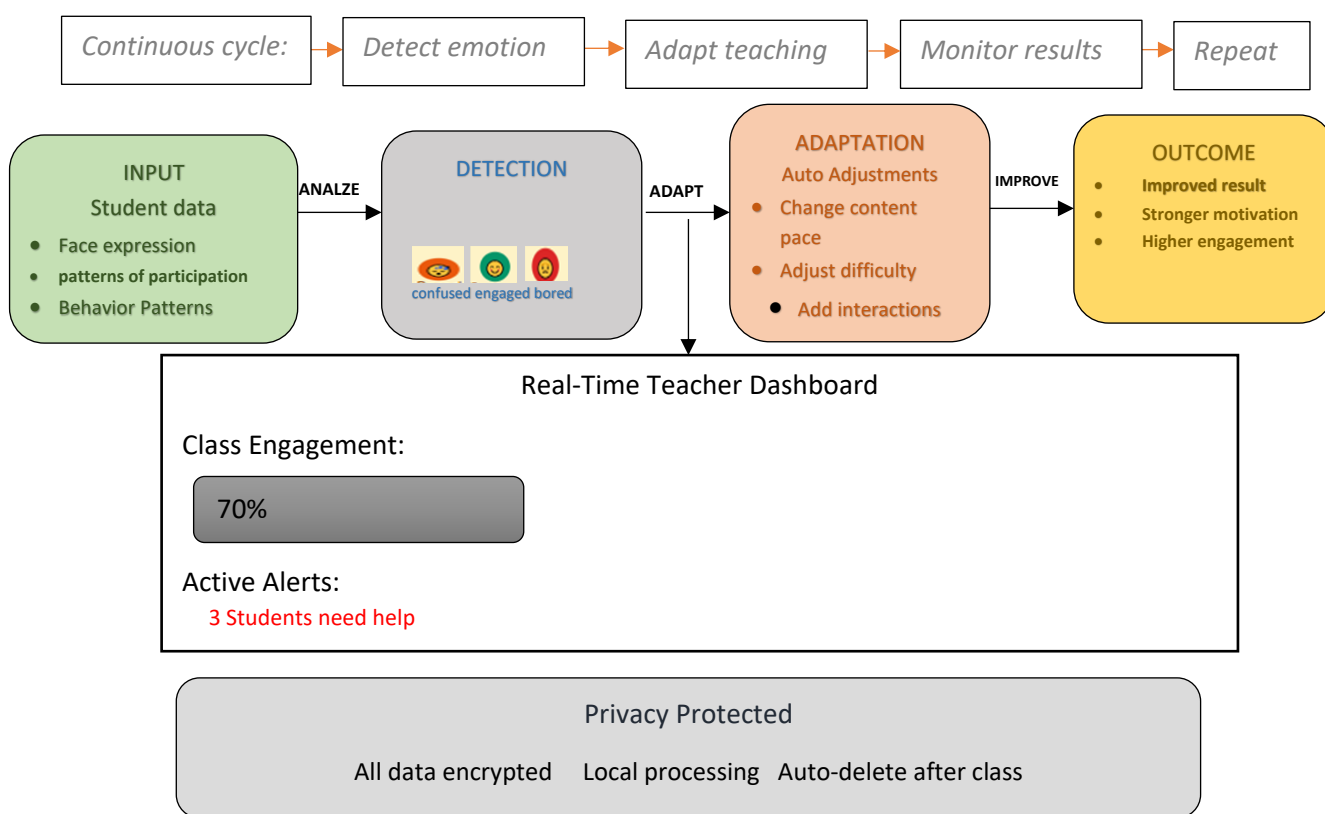


Figure 3: Conceptual Framework

DISCUSSION

An extensive array of research initiatives aiming at comprehending and improving the educational experience is shown by a study of the literature on emotion detection in virtual education. Several approaches, ranging from conventional machine learning algorithms to cutting-edge deep learning models, have been employed in the reviewed publications to identify and assess students' emotional states in virtual classrooms. Several recurring topics emerge from the review. Firstly, scholars have repeatedly emphasized the importance of precise emotion detection in online learning environments, recognizing that it can inform tailored interventions, modify instructional techniques, and foster a nurturing atmosphere for students. Furthermore, the size and diversity of the dataset are shown to be important variables affecting the robustness and generalization of emotion detection models, emphasizing the necessity of representative and thorough data collecting. The incorporation of emotion-detected real-time feedback mechanisms has a substantial impact on learning outcomes, motivation, and student engagement. Virtual classroom management systems can enhance learning and foster positive emotional well-being by offering timely insights into students' emotional states and enabling individualized interventions.

The paper emphasizes the importance of leveraging developments in machine learning and artificial intelligence to create virtual learning environments that

are more effective, immersive, and adaptive. Notwithstanding certain obstacles, such as restricted datasets, model resilience, and the need for additional validation, persist, suggesting potential directions for future investigations and advancements in the realm of emotion identification in online learning.

In addition to technical findings, the review uncovered several critical gaps in current research. One major concern is the **lack of cultural and linguistic diversity in emotion detection datasets**, which affects the generalizability of models across global virtual classrooms. Emotions are expressed and interpreted differently across cultures, yet many systems rely on datasets that lack sufficient representation from diverse ethnic, linguistic, and regional backgrounds.

Furthermore, **real-time feedback mechanisms are rarely tested or deployed in low-resource educational contexts**, limiting their practical value for many schools and learners in developing regions. Addressing infrastructural challenges and designing lightweight, adaptable systems will be essential to ensure equitable access.

The **lack of inclusive design** in current emotion models also presents a barrier to broad adoption. Most systems are not optimized for learners with special needs or neurodiverse characteristics, highlighting the need for

future work that integrates universal design principles into affective computing.

Recent reviews, such as those by [Campbell et al. \(2020\)](#), have primarily focused on modalities like log-based behavior or facial expression analytics. In contrast, our review emphasizes **dataset diversity, cultural bias, and the integration of real-time feedback**, which these prior reviews largely overlooked.

Unlike surveys such as those by [Cumpston et al. \(2023\)](#), which highlight affective computing generally, we directly address ethical implications, inclusivity, and the deployment of low-resource contexts, thus providing new actionable insights for educational technology design.

The reviewed studies exhibited several methodological strengths, notably the adoption of state-of-the-art deep learning models with high classification accuracy and the use of reproducible experiments with clearly described model architectures. However, common methodological weaknesses were also identified. Approximately 40% of the studies utilized small sample sizes, which limited their statistical power and reliability. Additionally, only 25% of the studies implemented robust validation methods such as cross-validation, raising concerns about model reliability. Another key limitation was the low generalizability of results, often caused by overfitting on narrow, synthetic, or domain-specific datasets.

Identified Gaps and Underexplored Dimensions

Beyond performance metrics and model diversity, the reviewed studies reveal several underexplored dimensions in emotion detection for virtual education. First, there is limited attention to cultural and linguistic biases in emotion datasets. Most datasets are built on Western-centric facial and emotional norms, which may not generalize well to diverse student populations. This limits the applicability of emotion models in multilingual or culturally diverse virtual classrooms.

Second, only a few studies explicitly address the deployment of emotion-based real-time feedback mechanisms in low-resource educational settings, where infrastructure, bandwidth, and computational capacity are often constrained. This raises concerns about the scalability and accessibility of such technologies for broader use.

Lastly, a gap exists in the **inclusivity** of emotion detection systems. Current models often lack sensitivity to diverse learning needs, neurodivergent behaviors, and variations in emotional expression across different age groups and demographics. These limitations underscore the need for adaptive, equitable frameworks for emotion detection in future research.

CONCLUSION

The compilation of articles that have been evaluated regarding emotion detection in virtual learning <https://scientifica.umyu.edu.ng/>

emphasizes how revolutionary it can be to utilize technological advancements to enhance learning. Researchers have demonstrated the importance of effectively identifying and addressing students' emotional states in virtual classrooms through various methodologies and approaches.

Important discoveries demonstrate how effectively different deep learning and machine learning models capture and interpret emotional expressions, ranging from facial recognition to EEG signal analysis. Including real-time feedback systems that recognize emotions presents intriguing opportunities for tailored interventions, flexible teaching methods, and creating a positive learning environment. Challenges such as restricted datasets, resilient models, and the need for additional verification persist, indicating opportunities for further investigation and advancement. To fully realize the promise of emotion detection in virtual education and enhance student engagement, motivation, and learning outcomes, interdisciplinary collaborations and innovative approaches will be essential as the field develops. The examined studies shed light on how to design virtual learning environments that are more effective, immersive, and flexible, while meeting the diverse emotional needs of students.

Future research should prioritize inclusivity, cultural sensitivity, and the practical deployment of emotion-aware systems in underrepresented and low-resource educational contexts to ensure equitable impact.

Overall, while emotion detection models demonstrate promising accuracy, challenges such as dataset bias, limited generalizability, and lack of real-world validation remain. Addressing these issues through inclusive data collection and real-time feedback systems will be key to enhancing learner engagement and educational impact.

REFERENCES

- Abedi, A., & Khan, S. S. (2024). Affect-driven ordinal engagement measurement from video. *Multimedia Tools and Applications*, 83(8). [\[Crossref\]](#)
- Alsokari, M. H., & Okatan, A. (2023). Real-time online education student lecture emotion detection. *of Engineering and Applied Sciences*, 3(1), 1-56. [\[Crossref\]](#)
- Anand, M., & Babu, S. (2024). Multi-class Facial Emotion Expression Identification Using DL-Based Feature Extraction with Classification Models. *International Journal of Computational Intelligence Systems*, 17(1). [\[Crossref\]](#)
- Ansy, S. N., Bilal, E. A., & Neethu, M. S. (2024). Emotion Recognition Through Facial Expressions from Images Using Deep Learning Techniques. *Lecture Notes in Networks and Systems*, 819. [\[Crossref\]](#)
- Begum, F., Neelima, A., & Valan, J. A. (2023). Emotion recognition system for E-learning environment based on facial expressions. *Soft Computing*, 27(22). [\[Crossref\]](#)
- Caglar-Ozhan, S., Altun, A., & Ekmekcioglu, E. (2022). Emotional patterns in a simulated virtual

- classroom supported with an affective recommendation system. *British Journal of Educational Technology*, 53(6), 1724-1749. [Crossref]
- Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., & Brennan, S. E. (2020). *Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guideline*. *BMJ*, 368, 16890. [Crossref]
- Chin, Z. Y., Zhang, Z., Wang, C., & Ang, K. K. (2021). An Affective Interaction System using Virtual Reality and Brain-Computer Interface. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2021-January*. [Crossref]
- Cumpston, M. S., Brennan, S. E., Ryan, R., & McKenzie, J. E. (2023). Synthesis methods other than meta-analysis were commonly used but seldom specified: Survey of systematic reviews. *Journal of Clinical Epidemiology*, 156, 42–52. [Crossref]
- Das, I., & Paris, K. (2022). A Deep Learning-Based Approach for Adaptive Virtual Learning with Human Facial Emotion Detection. *Journal of Student Research*, 11(3). [Crossref]
- Deivanayagi, S., Nandhini Sri, V. G., Kalai Priya, P., & Aarthi, G. (2019). Pupil detection algorithm based on feature extraction for eye gaze. *International Journal of Recent Technology and Engineering*, 8(2 Special issue 5). [Crossref]
- Du, Y., Crespo, R. G., & Martínez, O. S. (2023). Human emotion recognition for enhanced performance evaluation in e-learning. *Progress in Artificial Intelligence*, 12(2). [Crossref]
- Durán Acevedo, C. M., Carrillo Gómez, J. K., & Albarracín Rojas, C. A. (2021). Academic stress detection on university students during COVID-19 outbreak by using an electronic nose and the galvanic skin response. *Biomedical Signal Processing and Control*, 68. [Crossref]
- Gupta, S., Kumar, P., & Tekchandani, R. (2023). A multimodal facial cues based engagement detection system in e-learning context using deep learning approach. *Multimedia Tools and Applications*, 82(18). [Crossref]
- Gupta, S., Kumar, P., & Tekchandani, R. (2023). An optimized deep convolutional neural network for adaptive learning using feature fusion in multimodal data. *Decision Analytics Journal*, 8. [Crossref]
- Gupta, S., Kumar, P., & Tekchandani, R. K. (2023). EDFA: Ensemble deep CNN for assessing student's cognitive state in adaptive online learning environments. *International Journal of Cognitive Computing in Engineering*, 4. [Crossref]
- Hans, A. S. A., & Rao, S. (2021). A CNN-LSTM based deep neural networks for facial emotion detection in videos. *International Journal of Advances in Signal and Image Sciences*, 7(1). [Crossref]
- Hasnine, M. N., Bui, H. T. T., Tran, T. T. T., Nguyen, H. T., Akçapınar, G., & Ueda, H. (2021). Students' emotion extraction and visualization for engagement detection in online learning. *Procedia Computer Science*, 192. [Crossref]
- Hoffmann, J., & McGarry, J. A. (2024). Big challenges and new solutions: When students' (unpleasant) feelings are used for good. In *Crises, creativity and innovation* (pp. 127–151). Springer. [Crossref]
- Jagtap, Mr. S. (2023). Design and Analysis of Unlocking Student Emotions: Enhancing E-Learning with Facial Emotion Detection. *International Journal for Research in Applied Science and Engineering Technology*, 11(12). [Crossref]
- Jaiswal, R., Nair, A. K., & Sahoo, J. (2022). Real-time monitoring system for attendance and attentiveness in virtual classroom environments. *2022 2nd International Conference on Artificial Intelligence and Signal Processing, AISP 2022*. [Crossref]
- Kaewkaisorn, K., Pintong, K., Bunyang, S., Tansawat, T., & Siriborvornratanakul, T. (2022). Student Attentiveness Analysis in Virtual Classroom Using Distraction, Drowsiness and Emotion Detection. *SSRN Electronic Journal*. [Crossref]
- Katoch, I., Kaushal, M., & Tanu. (2023). Evaluation of Student Engagement using Deep Learning in E-learning Environment. *2023 International Conference on Data Science, Agents and Artificial Intelligence, ICDSAAI 2023*. [Crossref]
- Kodithuwakku, J., Arachchi, D. D., & Rajasekera, J. (2022). An Emotion and Attention Recognition System to Classify the Level of Engagement to a Video Conversation by Participants in Real Time Using Machine Learning Models and Utilizing a Neural Accelerator Chip. *Algorithms*, 15(5). [Crossref]
- Lampropoulos, G., Keramopoulos, E., Diamantaras, K., & Evangelidis, G. (2022). Augmented Reality and Virtual Reality in Education: Public Perspectives, Sentiments, Attitudes, and Discourses. *Education Sciences*, 12(11). [Crossref]
- Liao, S. (2024). Adapting social and emotional learning for virtual classrooms: Theoretical insights and practical implications. *Frontiers in Humanities and Social Sciences*, 4(12), 361–368. [Crossref]
- Llurba, C., Fretes Torruella, G., & Palau, R. (2022). Pilot study of real-time emotional recognition technology for secondary school students. *Interaction Design and Architecture(s)*, (52), 61–80. [Crossref]
- Mary, T. A. C., & Rose, P. J. A. L. (2023). Multifaceted Sentiment Detection System (MSDS) to Avoid Dropout in Virtual Learning Environment using Multi-class Classifiers. *International Journal of Advanced Computer Science and Applications*, 14(4). [Crossref]
- Melanie, M., Keller, & Eva, S. (2021) Teachers' emotions and emotional authenticity: do they matter to students' emotional responses in the classroom?, *Teachers and Teaching*, 27:5, 404-422, [Crossref]
- Mena-Guacas, A. F., López-Catalán, L., Bernal-Bravo, C., & Ballesteros-Regaña, C. (2025). Educational transformation through emerging technologies: Critical review of scientific impact on learning. *Education Sciences*, 15(3), 368. [Crossref]

- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). *The PRISMA 2020 statement: An updated guideline for reporting systematic reviews*. *BMJ*, 372, n71. [\[Crossref\]](#)
- Pahutar, A. A., Lahmi, A., Hakim, R., & Dahlan, D. (2024). Emotions and brain processes in the learning process. *Edu Global: Jurnal Pendidikan Islam*, 5(2), 31–38. [\[Crossref\]](#)
- PrabhakaraRao, T., Patnala, S., Raghavendran, C. v., Lydia, E. L., Lee, Y., Acharya, S., & Hwang, J. Y. (2024). Oppositional Brain Storm Optimization with Deep Learning based Facial Emotion Recognition for Autonomous Intelligent Systems. *IEEE Access*. [\[Crossref\]](#)
- Rasheed, F., & Wahid, A. (2023). An Unobtrusive Approach to Emotion Detection in E-Learning Systems. *Computer Journal*, 66(8). [\[Crossref\]](#)
- Revadekar, A., Oak, S., Gadekar, A., & Bide, P. (2020). Gauging attention of students in an e-learning environment. *4th IEEE Conference on Information and Communication Technology, ICT 2020*. [\[Crossref\]](#)
- Romero-Alva, V., Ramos-Cosi, S., & Roman-Gonzalez, A. (2024). Educational Quality Assessment System based on Emotions using Facial Images Applying Deep Learning Techniques. *International Journal of Engineering Trends and Technology*, 72(3). [\[Crossref\]](#)
- Rosli, N., Jie, A. C. M., Chze, L. W., Faan, N. C. Z., & Liang, Y. N. (2022). Emotion Based Chatbot Using Deep Learning. *2022 International Conference on Future Trends in Smart Communities, ICFTSC 2022*. [\[Crossref\]](#)
- Shah, N. A., Meenakshi, K., Agarwal, A., & Sivasubramanian, S. (2021). Assessment of Student Attentiveness to E-Learning by Monitoring Behavioural Elements. *2021 International Conference on Computer Communication and Informatics, ICCCI 2021*. [\[Crossref\]](#)
- Sierra Rativa, A., Postma, M., & van Zaanen, M. (2023). Measuring Emotional Facial Expressions in Students with FaceReader: What Happens if Your Teacher is Not a Human, Instead, It is a Virtual Robotic Animal? *Lecture Notes in Networks and Systems*, 747 LNNS. [\[Crossref\]](#)
- Teo, J., Chia, J. T., & Lee, J. Y. (2019). Deep learning for emotion recognition in affective virtual reality and music applications. *International Journal of Recent Technology and Engineering*, 8(2 Special Issue 2). [\[Crossref\]](#)
- Thejaswini, S., Ramesh Babu, N., & Mamatha, K. R. (2023). Machine Learning Algorithm to Detect EEG based Emotion states using Virtual-Video stimuli. *IEEE International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems, ICAECIS 2023 - Proceedings*. [\[Crossref\]](#)
- Unciti, O., Martínez Ballesté, A., & Palau, R. (2024). Real-time emotion recognition and its effects in a learning environment. *Interaction Design and Architecture(s)*, 60(60), 85–102. [\[Crossref\]](#)
- Wang, C., Chen, X., Yu, T., Liu, Y., & Jing, Y. (2024). Education reform and change driven by digital technology: a bibliometric study from a global perspective. In *Humanities and Social Sciences Communications* (Vol. 11, Issue 1). [\[Crossref\]](#)
- Zheng, L. J., Mountstephens, J., & Teo, J. (2021). Eye Fixation Versus Pupil Diameter as Eye-Tracking Features for Virtual Reality Emotion Classification. *2021 IEEE International Conference on Computing, ICOCO 2021*. [\[Crossref\]](#)