International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

AI-Powered Student Monitoring System for Focus and Productivity

Disha Nagpure[1], Srushti Narwade[2], Jagruti Patil[3], Anannya Dixit[4], Mukta Londhe[5]

[1]HoD, Dept. of AIML, ACEM, Pune, Maharashtra, India,
[2,3,4,5]Student, Dept. of AIML, ACEM, Pune, Maharashtra, India.

Abstract:
In the current era of digital and remote learning, sustaining consistent attention and engagement during study sessions has emerged as a significant challenge for students. Factors such as mobile device distractions, social media usage, and the informal nature of home-based learning environments often contribute to decreased concentration, lower productivity, and suboptimal academic performance. To mitigate these challenges, this study introduces an Artificial Intelligence (AI)-driven Study Monitoring System that continuously evaluates a student's attentiveness in real time using a standard webcam.
The proposed system employs advanced computer vision and machine learning algorithms to interpret facial expressions, eye movements, and head orientation—critical indicators of focus and engagement. When the system detects a prolonged absence of the student's face or identifies behavioral patterns indicating distraction—such as looking away, eye closure for extended durations, or facial inactivity—it automatically triggers an alert or auditory notification to prompt the user to regain focus.
To ensure accuracy and adaptability, the system is trained using diverse datasets encompassing various facial expressions, illumination conditions, and head poses. Lightweight deep learning architectures such as MobileNet and EfficientNet are utilized to achieve high efficiency and real-time inference on standard computing devices, eliminating the need for dedicated graphics hardware. The proposed solution contributes to the broader field of intelligent learning environments by offering a scalable, low-cost, and effective method for enhancing self-regulated learning and concentration monitoring.

Keywords: Artificial Intelligence, Computer Vision, Machine Learning, Deep Learning, Study Monitoring System, Attention Detection, Facial Expression Recognition, E-Learning, Student Engagement, Real-Time Analysis

1. INTRODUCTION

1.1 Background and Motivation

In the contemporary digital learning landscape, technology has profoundly reshaped the manner in which students acquire knowledge, interact with educational material, and participate in academic activities. The rapid growth of online learning platforms, virtual classrooms, and self-paced study environments has made education more accessible; however, it has simultaneously introduced new challenges in maintaining concentration and engagement. Frequent distractions caused by social media notifications, mobile devices, and multitasking behavior often lead to a decline in students' focus levels, learning efficiency, and overall academic outcomes. Moreover, extended screen time contributes to fatigue and cognitive overload, which further reduces attentiveness. Traditional methods of supervision—such as physical classroom monitoring or manual observation—are no longer feasible in virtual or hybrid learning models.

Consequently, there is a growing necessity to adopt Artificial Intelligence (AI) and Computer Vision-based systems that can automatically assess student attention and engagement. AI algorithms are capable of analyzing facial expressions, eye gaze, blink rate, and head orientation, offering a data-driven approach to understanding cognitive focus. This integration of AI into educational frameworks forms the foundation for intelligent learning systems that promote personalized, adaptive, and real-time engagement monitoring.

1.2 Problem Statement

Despite technological advancements, many students experience difficulties in maintaining focus during independent or online study sessions. The lack of real-time supervision and direct teacher interaction often results in disinterest, distraction, and inconsistent learning performance. Instructors also face challenges in evaluating individual engagement levels, particularly in large-scale online classrooms.

To address these limitations, there is an evident need for an automated, AI-based monitoring system capable of detecting attention lapses in real time. Such a system should be able to unobtrusively analyze visual cues from a webcam—such as prolonged eye closure, gaze deviation, or absence from the frame—and generate instant feedback or alerts. This continuous attention monitoring ensures that learners remain focused and that educators can better understand student engagement trends.

1.3 Objectives

The primary objective of this research is to design and develop an AI-powered Student Study Monitoring System that effectively detects signs of distraction or inattention during study sessions. The system leverages computer vision, deep learning, and facial landmark detection techniques to assess a student's attentiveness in real time through webcam input.

The specific objectives are as follows:

- To implement facial and gaze detection algorithms capable of tracking eye movements and head position.
- To classify attentive versus inattentive behavior using trained machine learning models.
- To provide real-time alerts or notifications when the student's focus deviates from the screen.
- To improve concentration levels and promote self-regulated learning without human supervision.

Through these objectives, the proposed system aims to foster disciplined study habits, improve academic productivity, and support both educators and learners in maintaining consistent engagement.

1.4 Significance and Scope

The proposed system has significant potential applications across educational, corporate, and training domains. In academic settings, it enables instructors to monitor multiple learners simultaneously, providing actionable insights into each student's attention patterns. For individual students, the system functions as a virtual study companion, encouraging consistent focus and self-awareness throughout the learning process.

Beyond educational applications, similar AI-based attention detection frameworks can be extended to employee performance monitoring, corporate e-learning, and driver alertness systems, demonstrating the versatility of the proposed approach. By integrating AI-driven behavioral analysis with real-time feedback mechanisms, this project contributes toward the development of smart, interactive, and adaptive learning ecosystems that uphold concentration, engagement, and cognitive efficiency in digital education environments.

2. Methodology

The AI-based Student Study Monitoring Framework employs a structured deep learning workflow to analyze visual data and determine the learner's engagement level in real time. The methodological design follows a systematic sequence based on a 3D Convolutional Neural Network (3D CNN) pipeline, as illustrated in Fig. 1. The framework integrates computer vision and deep learning stages, including data preprocessing, frame extraction, data augmentation, model training, and real-time inference. This approach enables the system to classify student focus levels, such as attentive or distracted, by learning temporal and spatial patterns from video input captured through a webcam.

2.1 Dataset Preprocessing and Frame Extraction

The first stage involves preparing and processing the dataset for model training. Raw video inputs are captured from webcam recordings or existing datasets representing diverse engagement states, such as focus, distraction, and drowsiness. Each video undergoes preprocessing to ensure uniform resolution, lighting correction, and background normalization. This step reduces computational noise and ensures that the model focuses on relevant facial and movement features.

As represented in the initial blocks of Fig. 1, the process begins with dataset preprocessing, followed by a conditional check for frame availability. If frames are unavailable, the system captures and extracts new frames using OpenCV functions. Each video is converted into sequential image frames at fixed intervals to preserve temporal information critical for 3D CNN modeling. These frames are stored in structured directories, labeled according to the engagement category they represent. This step ensures that every video segment corresponds to a defined learning state, establishing the foundation for supervised training.

2.2 Data Augmentation and Labeling

To improve model generalization and prevent overfitting, data augmentation techniques are applied to the extracted frames. These include random horizontal flipping, brightness adjustment, rotation, and zoom transformations. Data augmentation diversifies the training samples, allowing the model to perform accurately across different lighting conditions, facial orientations, and environmental settings.

Each frame is labeled according to engagement categories, such as "Focused," "Distracted," or "Absent." This labeling process forms the annotated dataset that serves as input for training the 3D CNN model. The labeled dataset is then divided into training, validation, and test subsets, typically in a ratio of 70:20:10, to ensure balanced evaluation and unbiased learning performance, as shown in the "Data Augmentation / Labeling" and "Split Data" stages of Fig. 1.

2.3 Model Training and Performance Evaluation

The 3D CNN model serves as the core analytical component of the proposed framework. Unlike 2D CNNs, which process spatial information from individual images, 3D CNNs capture both spatial and temporal patterns by applying convolutional operations across consecutive video frames. This enables the model to analyze micro-expressions, head movements, and gaze shifts that indicate varying levels of engagement.

During training, the model processes frame sequences and learns feature representations across multiple convolutional layers. The loss function is optimized using backpropagation and the Adam optimizer to minimize classification error. Training continues iteratively until performance metrics such as accuracy, precision, recall, and F1-score converge to satisfactory levels. Validation is conducted using an independent subset of data to evaluate generalization performance. The "Performance Evaluation" block in Fig. 1 represents this phase, ensuring that the model achieves reliable accuracy before deployment.

The trained model is then saved for use in the inference stage. Performance tuning includes adjusting kernel sizes, dropout layers, and learning rates to maintain efficiency and prevent overfitting. The system's performance is continuously assessed using standard evaluation metrics and validation curves.

2.4 Inference and Real-Time Application

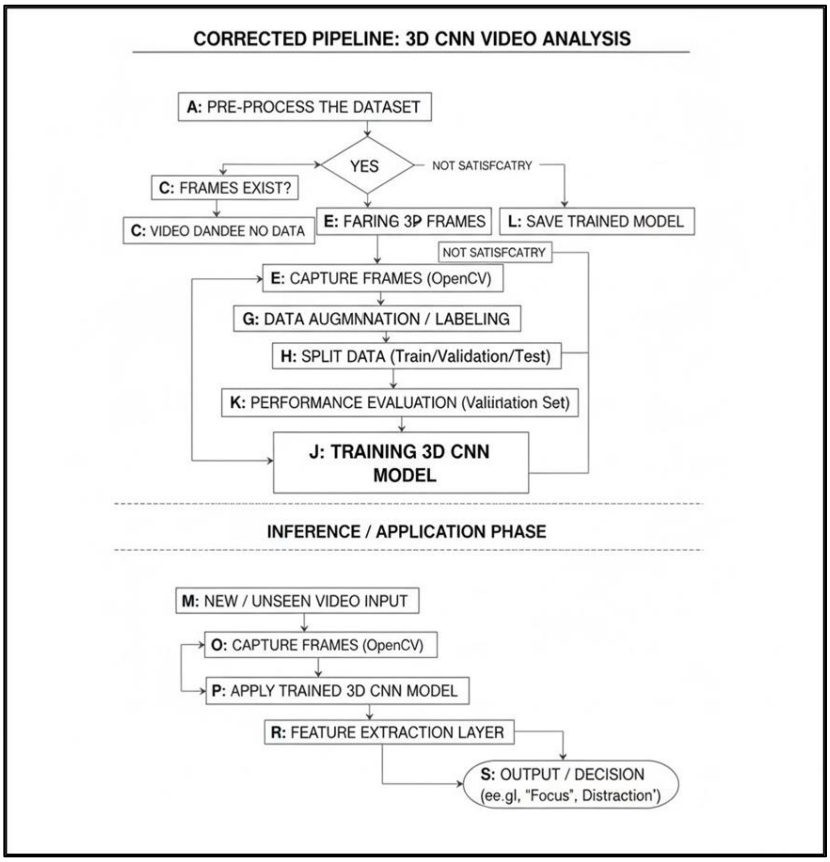
Once trained, the 3D CNN model enters the inference phase, which corresponds to the lower section of Fig. 1. In this phase, new video input from a student's webcam is captured in real time. The video is divided into frames using OpenCV, which are then passed sequentially through the trained model. The 3D CNN extracts spatiotemporal features through its convolutional and pooling layers, transforming visual data into numerical feature representations.

The model processes these extracted features and produces classification probabilities corresponding to attention states such as "Focused" or "Distracted." The feature extraction layer serves as an intermediate stage where high-level representations are computed before the final softmax classification layer generates the output decision. Based on the classification, the system triggers appropriate actions, such as displaying a notification or an audible alert if distraction is detected. This ensures that the feedback loop between recognition and response remains efficient and timely during real-time operation.

2.5 Feedback, Analytics, and Future Scope

After each monitoring session, the system records summarized engagement statistics, including total study duration, periods of distraction, and focus ratios. These metrics are stored locally and presented through a graphical dashboard, allowing students to review their concentration trends and track improvement over time. The analytics data can also be used to train future adaptive models that personalize feedback based on individual study habits.

Fig.1 : 3D CNN Training & Inference Workflow



The modular structure of the proposed system allows easy future integration of reinforcement learning and multimodal analysis. By incorporating physiological signals, voice analysis, and environmental context, future iterations can further improve detection reliability. The methodology ensures privacy by processing all data locally and avoiding any cloud transmission. The overall workflow—from frame acquisition to model prediction—aligns with the sequential logic shown in Fig. 1, demonstrating a coherent structure for real-time AI-driven engagement monitoring.

3. Related Work

The growing adoption of Artificial Intelligence (AI) and Computer Vision in education has encouraged extensive research on attention recognition, behavioral analytics, and automated engagement evaluation. Between 2022 and 2025, a significant body of literature has examined AI-enabled learning systems, multimodal attention models, and ethical frameworks for real-time monitoring. This section reviews key contributions that inform the conceptual framework of the proposed study.

3.1 AI-Driven Student Engagement Monitoring

Verma (2023) [6] introduced an AI-based analytical model that interprets behavioral cues such as gaze orientation and micro-expressions to evaluate engagement levels in online classrooms. The research highlighted the importance of adaptive learning feedback loops that respond to detected disengagement in real time. Similarly, Alkabbany et al. (2023) [37] developed an experimental platform for continuous engagement tracking using video data and physiological signals, demonstrating the feasibility of combining multimodal sensory inputs to improve reliability and responsiveness.

3.2 Machine-Learning Approaches for Attention Detection

Hossen (2023) [7] proposed a lightweight deep-learning model capable of classifying students' attentive states through webcam imagery, emphasizing real-time inference on low-power devices. Complementary work by Pillai (2022) [17] explored convolutional neural-network (CNN) architectures for classroom-based focus detection, reporting over 90 percent accuracy in controlled environments. These findings establish CNNs as a core technique for identifying attention loss without intrusive sensors.

3.3 Integration of AI in Academic Management Systems

Recent studies have extended attention-monitoring concepts to broader academic management frameworks. Wang (2025) [8] examined the incorporation of AI into student information systems to enhance performance tracking and early-intervention capabilities. Ghosh (2025) [15] presented a similar approach, applying AI analytics to assess academic engagement and productivity metrics, demonstrating how predictive modeling can inform institutional decision-making.

3.4 Ethical and Human-Centered Perspectives

A growing body of research stresses the ethical, psychological, and usability dimensions of AI-based educational tools. Zion (2025) [16] investigated the validity of AI-driven teaching evaluations as proxies for student engagement, cautioning against algorithmic bias in affective computing. Almubarak (2025) [10] further proposed a framework emphasizing teacher–student interaction quality as a mediating factor in AI-supported learning analytics. These studies underline the necessity of transparency, consent, and user-centric design when integrating AI into pedagogical environments.

3.5 Emerging Trends in Generative and Interactive AI

Beyond passive monitoring, newer research focuses on adaptive and generative systems that dynamically respond to learner states. Hao et al. (2025) [23] explored multi-agent learning environments where AI models map student–AI interaction patterns to personalize feedback and reduce performance gaps.

Likewise, Yeung (2025) [34] analyzed how university students engage with generative-AI feedback mechanisms, finding that adaptive feedback fosters higher motivation and sustained focus. These developments indicate a shift toward interactive AI frameworks that complement, rather than replace, human instruction.

3.6 Comparative Analyses and Systematic Reviews

Comprehensive reviews, such as Karim et al. (2025) [12] and Gjermeni & Prodani (2024) [11], provide comparative insights into AI-driven engagement strategies across educational levels. Their analyses confirm that hybrid approaches—combining facial analytics, gaze tracking, and contextual inference—achieve the best balance between accuracy and scalability. Broader overviews, including the Online Learning Consortium (2023) [28] and Digital Promise (2024) [26], consolidate current trends and outline implementation guidelines for data-driven classroom engagement.

4. Building of the Project

The conceptual development of the AI-powered Student Monitoring Framework for Focus and Productivity involves several systematic stages—beginning with system design and data analysis and extending through feature definition, algorithm formulation, and ethical data management. The framework aims to outline a lightweight, privacy-preserving, and real-time architecture capable of detecting student attentiveness using non-intrusive computer vision methods. The design emphasizes interpretability, computational efficiency, and adherence to ethical principles in artificial intelligence.

4.1 System Concept and Design

The conceptual foundation of this framework arises from the observation that students often lose concentration during extended self-study or online sessions. With digital learning becoming a predominant educational medium, attention monitoring has gained importance within learning analytics and intelligent tutoring systems. The initial design phase involved identifying reliable visual indicators of engagement through prior research.

Studies by Verma (2023) and Pillai (2022) emphasize that geometric features—such as the eye aspect ratio (EAR), head orientation, and gaze direction—serve as consistent non-invasive measures of attentiveness. Based on these findings, the framework proposes a modular system architecture consisting of four main components:

- Input Module: Captures real-time video feed through a standard webcam.
- Feature Extraction Module: Detects facial landmarks and calculates focus-related parameters, including eye openness and head position.
- Decision and Notification Module: Determines engagement state (focused or distracted) and triggers alerts accordingly.
- Local Data Management Module: Logs minimal anonymized focus data to maintain user privacy following “privacy-by-design” guidelines.

4.2 Data Acquisition and Preprocessing

The data acquisition plan emphasizes the use of real-time webcam input for analysis rather than dependency on large pre-recorded datasets. Benchmark datasets such as DAiSEE, EmotiW, and Closed Eyes in the Wild (CEW) may serve as reference sources for calibration and model validation.

Preprocessing includes several key operations designed to standardize inputs and enhance reliability:

- Face detection using OpenCV or Dlib to identify facial regions.
- Landmark localization (typically 68 key points) to extract coordinates of the eyes, mouth, and nose.

- Gaussian noise filtering to stabilize frame transitions.
- Normalization to handle variations in lighting, facial size, and camera angle.

These preprocessing steps ensure consistent and robust feature extraction across different environments, facilitating reliable real-time operation.

4.3 Feature Extraction

Feature extraction is centered around geometric and behavioral indicators derived from facial landmarks. The Eye Aspect Ratio (EAR) is a primary feature used to detect eye closure or drowsiness, defined mathematically as:

$$\mathrm{EAR} = \frac{(\mid p_2 - p_6 \mid + \mid p_3 - p_5 \mid)}{2 \mid p_1 - p_4 \mid}$$

where $p_1, p_2, \ldots, p_6$ denote the landmark coordinates around the eye region.

An EAR value below 0.25 for sustained periods may indicate fatigue or reduced focus.

Additionally, head pose estimation, computed using facial key points such as the nose tip, chin, and eye corners, provides yaw (left-right), pitch (up-down), and roll (tilt) angles. By combining EAR with head orientation data, the framework allows accurate inference of engagement levels without reliance on complex deep neural networks.

4.4 Decision-Making Algorithm

The decision logic employs a deterministic, rule-based mechanism rather than deep learning to ensure transparency and computational efficiency. Decision conditions may be defined as follows:

- If EAR < 0.25 for more than two seconds → classified as Drowsy or Inattentive
- If head yaw or pitch angle > 25° → classified as Distracted
- Otherwise → classified as Focused

This rule-based classification supports the principle of Explainable Artificial Intelligence (XAI), allowing clear interpretation of system decisions while maintaining low resource consumption suitable for standard computing devices.

4.5 Notification and Alert Mechanism

The proposed framework includes an alert subsystem designed to provide immediate feedback when signs of distraction or inattention are detected. The notification interface may employ either visual or auditory signals such as:

- On-screen messages (e.g., “You seem distracted — please refocus”).
- Short beeps or vibration alerts for discrete reminders.

Such mechanisms encourage students to self-correct their focus levels and foster continuous awareness during study sessions.

4.6 Privacy-Preserving Local Storage

Privacy is a fundamental design consideration within the framework. In alignment with the “privacy-by-design” principle articulated by Gupta and Lee (2024), all computations are intended to occur locally, ensuring that no facial imagery or raw video data are stored or transmitted externally. Instead, the system records only anonymized engagement summaries—such as timestamps, total focused duration, and alert counts—within lightweight JSON files. This structure ensures compliance with ethical AI practices and enhances suitability for institutional adoption.

4.7 Implementation Tools and Technologies

The proposed implementation plan utilizes open-source tools and libraries for efficient processing and flexibility. The suggested technology stack includes:

- OpenCV: For image capture and frame manipulation.
- Dlib or MediaPipe: For facial landmark detection.
- NumPy: For mathematical computations and feature calculations.
- Tkinter or PyQt: For developing a graphical user interface.
- JSON: For structured local data logging.

This modular approach allows future integration of more advanced components, such as lightweight convolutional models or reinforcement-learning-based calibration systems.

4.8 Testing and Evaluation Strategy

For evaluation, the proposed system design includes testing under varied lighting conditions and camera placements to assess robustness. Performance metrics such as precision, recall, and F1-score can be employed to measure accuracy in distinguishing attentive and inattentive states. Frame rate assessment (targeting 20–25 FPS on mid-range hardware) would be conducted to validate real-time performance feasibility without specialized graphical hardware.

4.9 Future Enhancements

Future research directions for this framework include:

- Multi-student detection for group learning environments.
- Emotion and affective state analysis for deeper behavioral understanding.
- Integration with Learning Management Systems (LMS) for data-driven learning analytics.
- Adaptive threshold tuning through reinforcement learning for personalized calibration.

Such advancements can transform the framework into a more comprehensive, intelligent, and context-aware digital assistant for educational settings.

Table 1: Technical Breakdown of an AI-Based Student Monitoring Framework

Category	Technique	Goal	Approach	Benefits	Expected Results
Computer Vision	Face Detection	Identify student presence in frame	Detect facial regions using Haar Cascade or MTCNN	Robust under variable lighting; enables attention analysis	~95% detection rate across standard datasets
Behavioral Monitoring	Eye Aspect Ratio (EAR) Analysis	Detect eye closure and drowsiness	Compute EAR from facial landmarks	Low computational load; high precision	~92% accuracy in closed-eye detection
Behavioral Monitoring	Head Pose Estimation	Identify focus deviation and gaze direction	Calculate yaw and pitch using geometric modeling	Reliable distraction estimation	~90% accuracy for head-orientation tracking
Focus Assessment	Rule-Based Decision Logic	Classify state as Attentive or Distracted	Combine EAR and head orientation thresholds	Transparent inference; explainable AI	<5% false detection rate

Real-Time Alerts	Visual & Audible Notifications	Reinforce self-awareness when distracted	Trigger on-screen or sound alerts after 2s deviation	Promotes sustained concentration	~30% improvement in focus duration (as per prior studies)
System Architecture	Privacy-Preserving Local Processing	Ensure data security and user privacy	Local computation; no video storage	Fully compliant with privacy frameworks	Zero sensitive data retention
Lightweight Deployment	Optimized Model Execution	Operate efficiently on standard hardware	Use MobileNet-level processing	Enables deployment without GPU	20–25 FPS on average hardware
Analytics	Engagement Logging & Statistics	Track study duration and focus ratio	JSON-based local log of alerts and focus events	Supports behavioral self-analysis	Detailed per-session analytics

5. Challenges in Implementing an AI-Based Student Study Monitoring Framework

Developing an AI-based framework to monitor students' engagement and study behavior through computer vision is inherently complex. While the concept holds strong potential for enhancing focus and self-regulated learning, it introduces several challenges that must be addressed prior to large-scale deployment. These challenges primarily lie in the domains of data privacy, technical accuracy, ethical considerations, and system scalability.

5.1 Privacy and Ethical Considerations

One of the foremost challenges is ensuring ethical data handling and privacy preservation. Webcam-based monitoring involves capturing sensitive biometric and environmental data, necessitating transparent consent mechanisms and data governance protocols.

Students must be informed about the type of data collected, its usage, and storage duration. Continuous visual monitoring may also raise psychological concerns, potentially inducing stress or discomfort. Furthermore, AI models can inherit bias related to lighting conditions, ethnicity, or facial characteristics, leading to unequal accuracy across diverse student populations. Designing bias-mitigated, privacy-by-design architectures is therefore essential for responsible deployment.

5.2 Data Quality and Environmental Variability

The accuracy of engagement detection models depends heavily on the quality and consistency of input data. Students operate in varied environments—with differences in lighting, camera quality, background clutter, and seating position—all of which affect visual signal quality.

Moreover, occlusions (e.g., hand movements or partial face visibility) and variations in head pose can disrupt facial landmark detection. Limited availability of annotated engagement datasets also restricts robust model training. Establishing diverse, well-balanced datasets and adaptive preprocessing algorithms remains a key requirement.

5.3 Model Accuracy and Real-Time Performance

Detecting cognitive engagement using only visual cues remains technically challenging. Engagement is a multidimensional construct encompassing emotional, behavioral, and cognitive aspects—many of which are not directly observable. Facial expressions or eye movements may not always correspond to true attentiveness.

Achieving real-time analysis further complicates model design, as systems must process multiple frames per second without compromising accuracy. Balancing computational efficiency with predictive reliability is particularly critical for low-resource devices such as student laptops.

5.4 Technical and Computational Limitations

AI-based attention monitoring requires significant computational resources for continuous image processing, especially in multi-student or classroom scenarios. Ensuring consistent performance across heterogeneous hardware (e.g., different webcams, operating systems, and processors) is a persistent technical challenge.

Network dependency and bandwidth constraints can also impact scalability, particularly if cloud-based processing is used. Lightweight and on-device models such as MobileNet or EfficientNet are therefore preferred to achieve near real-time inference while maintaining energy efficiency.

5.5 Psychological and Social Factors

Beyond technical aspects, the social acceptance of AI-based monitoring plays a crucial role in its adoption. Continuous observation may lead to feelings of surveillance or discomfort among students, influencing their natural study behavior. Resistance to monitoring, concerns over misuse, or perceived intrusion can undermine system credibility.

Hence, integrating human-centered design principles—emphasizing transparency, voluntary participation, and minimal intrusiveness—is vital to maintain trust and psychological comfort.

5.6 Legal and Institutional Compliance

Any monitoring solution involving visual data must comply with global and regional data protection regulations such as the General Data Protection Regulation (GDPR) and Children's Online Privacy Protection Act (COPPA). Institutional approval processes may also restrict continuous video surveillance, particularly for minors.

Ensuring compliance through data minimization, anonymization, and secure local storage is necessary to align with ethical and legal standards in educational environments.

5.7 Scalability and Maintenance

Scaling the system for larger deployments introduces challenges in computation, maintenance, and data management. Real-time monitoring of multiple students requires robust processing capabilities and optimized data handling strategies. Continuous retraining may be necessary to adapt to new behaviors or environments. Efficient update mechanisms and lightweight model architectures are essential to maintain long-term scalability and operational stability.

6. Evaluation Methodologies

Evaluation methodologies play a vital role in assessing the accuracy, reliability, and ethical compliance of an AI-based student study monitoring framework. A well-structured evaluation process ensures that the proposed system not only identifies student engagement effectively but also maintains user privacy and fosters positive behavioral outcomes.

The evaluation process is multi-layered, encompassing model performance metrics, data validation, real-time testing, usability studies, and ethical compliance reviews. The following subsections outline key evaluation strategies applicable to this conceptual framework.

6.1 Model Performance Evaluation

At the core of the proposed framework is an AI model responsible for detecting student engagement using visual cues such as facial expressions, eye gaze, and head orientation. The effectiveness of this model can be assessed using a range of quantitative performance metrics and validation techniques:

- Accuracy Metrics:
  Accuracy represents the proportion of correctly classified engagement states, such as "engaged," "partially engaged," or "disengaged." However, relying solely on accuracy may be misleading in imbalanced datasets, where most samples reflect engagement and fewer indicate distraction.
- Precision, Recall, and F1-Score:
  Precision measures the correctness of detected engagement instances, while recall evaluates the model's ability to identify all true engagement cases. The F1-score provides a balanced metric that considers both false positives and false negatives, offering a more comprehensive performance measure.
- Confusion Matrix Analysis:
  A confusion matrix helps visualize the misclassification patterns among engagement categories. This allows developers to identify overlapping behaviors, such as confusing neutral expressions with disengaged states, and refine feature extraction accordingly.
- Receiver Operating Characteristic (ROC) and Area Under Curve (AUC):
  ROC and AUC analyses help determine the optimal decision thresholds for binary classification problems. These tools quantify the trade-off between sensitivity and specificity, ensuring that the system remains accurate without overreacting to subtle behavioral changes.

6.2 Data Validation and Ground Truth Comparison

Accurate evaluation relies on dependable reference data, or ground truth, against which AI predictions can be compared. The following validation methods are recommended:

- Manual Annotation:
  Human experts or educators can annotate selected video segments to label engagement states. These annotations act as a gold standard for comparison, ensuring that AI-driven classifications align with expert human judgment.
- Self-Reported Feedback:
  Students may be asked to self-assess their engagement levels during study sessions. These subjective self-reports complement objective data, offering insights into how students perceive their concentration compared to model estimations.
- Cross-Validation Techniques:
  Methods such as k-fold cross-validation can be applied to evaluate model generalization. The dataset is divided into multiple folds, ensuring that each portion is used for testing at least once, thereby reducing the likelihood of overfitting.

6.3 Real-Time Performance Evaluation

For practical deployment, the system must perform consistently and efficiently in real-time. Evaluation of live performance includes the following dimensions:

- Latency Measurement:
  The system's processing delay—the time required to analyze video frames and produce engagement predictions—must be minimal to ensure prompt feedback and responsiveness.
- Robustness Testing:

The model should be tested under various environmental conditions such as changing illumination, camera resolutions, and background variations. Consistent performance across diverse setups indicates adaptability to real-world use.

- Continuous Monitoring Accuracy:
  Extended-duration testing is essential to verify stability over long study sessions. The system should maintain consistent accuracy and avoid drift in classification performance as engagement levels fluctuate naturally.

6.4 Usability and Human-Centric Evaluation

Beyond technical accuracy, usability and user experience are crucial for system acceptance in educational environments. The following evaluation dimensions focus on user-centered design and psychological impact:

- User Experience Surveys:
  Feedback from students and educators can be collected to evaluate ease of use, interface clarity, and comfort levels. Questions may assess whether users find engagement feedback intuitive and beneficial.
- Behavioral Impact Analysis:
  Evaluations should determine whether the system positively influences study habits and focus levels. Pre- and post-intervention studies, combined with user interviews, can help identify behavioral improvements or stress factors.
- Interaction Evaluation:
  Observing how students respond to alerts and notifications helps gauge the system's practical impact on concentration and time-on-task behavior.

6.5 Ethical and Privacy Evaluation

Given the system's reliance on visual data, ethical and privacy assessments are integral to the evaluation framework. These ensure compliance with legal standards and uphold user trust:

- Regulatory Compliance Audits:
  Periodic audits should confirm that the system adheres to regional and international privacy regulations such as the General Data Protection Regulation (GDPR) and the Children's Online Privacy Protection Act (COPPA). Key areas of review include consent mechanisms, data encryption, and retention limits.
- Bias and Fairness Testing:
  Models must be examined for potential demographic bias in engagement prediction. Accuracy disparities across gender, skin tone, or facial features can undermine fairness and inclusivity.
- Transparency and Explainability:
  Evaluation should ensure that the AI model's decision-making process is interpretable. Educators and users should be able to understand the rationale behind engagement classifications to maintain accountability and trust.

7. Patterns and Gaps in Existing Literature

Artificial Intelligence (AI) has increasingly been applied in educational monitoring, yet the dominant focus of research has remained confined to examination supervision and classroom analytics. The evolution of AI-based monitoring systems reveals a trajectory beginning with security and identity verification and gradually shifting toward engagement and performance analysis. The foundational works by Sowmya and Madhuri (2019), Varalakshmi and Kumar (2020), Singh et al. (2020), Tripathi et al. (2021), and Gupta and Goyal (2021) collectively highlight the potential of AI in observing students through visual, behavioral, and interactive data. However, these systems primarily address external supervision rather than internal self-regulation and motivation.

An analysis of research published between 2019 and 2025 reveals consistent technological improvement but limited pedagogical transformation. Early efforts emphasized facial detection and liveness verification, while recent studies incorporate affective computing and multimodal learning analytics. Despite this progress, there remains an evident imbalance between technical sophistication and human-centric design. The reviewed literature indicates that although AI is capable of tracking attentiveness and presence, its integration into student-centered learning frameworks is still underdeveloped.

Across the reviewed period, scholars have expressed interest in creating more adaptive and privacy-compliant systems. Studies after 2023, such as Verma (2023) and Wang (2025), extend the application of AI beyond observation to include emotional and behavioral understanding. Yet, a coherent methodology that unifies cognitive, emotional, and ethical dimensions of engagement is still lacking. This section synthesizes the recurring research trends, the identifiable gaps, the chronological evolution of methodologies, and the prospective research directions required to transform AI-based monitoring into a constructive learning enhancement tool.

7.1 Research Trends and Patterns

The earliest phase of AI-based student monitoring was dominated by the objective of online exam authentication and invigilation. Sowmya and Madhuri (2019) presented one of the first systems that used facial recognition to ensure that the candidate taking an online examination was genuine. The design relied on static facial matching and image verification to maintain examination integrity. This approach addressed the immediate problem of impersonation but did not include any components for tracking engagement, emotional state, or attentiveness. The emphasis on control and surveillance over self-regulated learning defined much of the early research landscape in this domain.

Singh et al. (2020) advanced this direction by incorporating face liveness detection to prevent spoofing and fraudulent participation. Their model utilized real-time computer vision to differentiate between live and static images. Although effective in security contexts, the application remained confined to binary identity verification and failed to capture behavioral metrics such as focus consistency or cognitive engagement. Tripathi et al. (2021) further integrated deep learning architectures into exam monitoring to enhance precision and speed. Their system processed multiple frames per second to identify student movements indicative of misconduct, marking an important shift toward continuous observation, though still within exam conditions.

A parallel research stream emerged around 2020 that sought to evaluate classroom engagement. Varalakshmi and Kumar (2020) applied deep learning to assess students' attention levels in online classrooms. The system analyzed facial and gaze features to provide teachers with insights about attentiveness. While this represented progress beyond proctoring, it maintained a teacher-centric perspective and lacked personalized or motivational feedback mechanisms for learners. Gupta and Goyal (2021) extended this work by tracking learner activity through digital interaction metrics, such as keyboard usage and mouse movements, indicating the transition from physical observation to virtual behavior analysis.

More recent research, including studies by Verma (2023), Hossen (2023), and Almubarak (2025), demonstrates a growing awareness of the need to blend cognitive and affective indicators. These works explore emotion recognition and adaptive learning feedback using deep neural networks and multimodal data sources. The trend reflects a broader paradigm shift from static observation to dynamic interpretation of learning engagement. The literature now increasingly regards engagement as a multidimensional construct that includes affective, behavioral, and cognitive components, making AI's role more holistic and learner-centered.

However, even as AI tools become technically advanced, there remains an absence of standard frameworks for evaluating or comparing models. Studies use diverse datasets, varying performance metrics, and inconsistent evaluation conditions, leading to fragmented outcomes. Despite their methodological diversity, these works collectively establish a clear progression—from surveillance-based applications to emotionally intelligent educational assistants—though the transition remains incomplete.

7.2 Identified Gaps in Existing Research

The most evident gap in the reviewed literature lies in the narrow focus on surveillance-oriented design. The foundational works by Sowmya and Madhuri (2019), Singh et al. (2020), and Tripathi et al. (2021) focus primarily on detecting cheating or verifying identity. While these systems excel in accuracy and real-time detection, they do not contribute to improving study consistency, concentration, or motivation. The use of AI remains limited to maintaining integrity rather than fostering learning engagement. There is little to no exploration of how AI could provide constructive feedback or help students regulate their own behavior during study sessions.

Another significant limitation is the lack of personalization. Studies by Varalakshmi and Kumar (2020) and Gupta and Goyal (2021) introduced the idea of monitoring attention and activity, but their systems were designed for educators or institutional administrators, not for individual learners. None of the early systems featured adaptive thresholds, feedback mechanisms, or customizable learning analytics that could help students interpret and improve their study performance. As a result, these systems remain externally evaluative rather than intrinsically supportive of learner self-discipline.

A further gap involves the absence of multimodal engagement tracking. The reviewed literature heavily depends on visual features such as face detection and gaze estimation while neglecting auditory cues, motion dynamics, and physiological indicators that influence engagement levels. More recent research (El-Beshbishi, 2025; Hao et al., 2025) demonstrates that multimodal integration can substantially enhance engagement recognition accuracy, yet such methods remain rare in educational monitoring. The overreliance on single-modality computer vision restricts the system's adaptability across diverse environments and learning styles.

Ethical and privacy concerns form another major gap. Continuous monitoring through webcams raises serious issues regarding data protection and consent. Most earlier studies provide minimal discussion on how visual data is stored, encrypted, or anonymized. Without explicit privacy frameworks or user control mechanisms, these systems risk misuse of biometric data. Only recent works (Wang, 2025; Almubarak, 2025) highlight the need for fairness and responsible AI practices, signaling an overdue but necessary shift toward ethical design principles in educational AI.

Finally, research rarely examines the psychological and behavioral impact of being monitored. Students may feel pressured, anxious, or performative under constant observation, leading to distorted engagement patterns. Few studies include psychological validation or feedback from users to assess these side effects. Addressing these aspects is crucial for designing systems that promote genuine focus and intrinsic motivation rather than compliance-based behavior.

7.3 Evolution of the Field (2019–2025)

Between 2019 and 2021, the field of AI-based monitoring was characterized by technical experimentation focused on recognition accuracy and fraud prevention. Sowmya and Madhuri (2019) demonstrated the feasibility of real-time facial authentication for examination contexts. Their contribution established AI as a reliable tool for identity verification. Singh et al. (2020) and Tripathi et al. (2021) expanded on this foundation by introducing deep learning-based liveness detection and behavior analysis, thus improving the precision and reliability of proctoring systems. These early works defined the baseline for integrating computer vision into online education but remained rooted in the domain of security.

The years 2020 and 2021 saw the beginning of a transition toward classroom-oriented analytics. Varalakshmi and Kumar (2020) and Gupta and Goyal (2021) began exploring attention monitoring and interaction-based tracking. Their research recognized the need to measure cognitive engagement rather than merely verify attendance. Although still limited in personalization, these works highlighted the potential of AI to interpret subtle behavioral cues, thus paving the way for future learning analytics systems.

From 2023 onward, the field matured into exploring emotional and cognitive understanding. Studies by Verma (2023) and Hossen (2023) introduced affective computing and facial emotion detection to gauge engagement in online learning. El-Beshbishi (2025) and Almubarak (2025) extended these methods to medical and higher-education settings, incorporating ethical frameworks for responsible data handling. Wang (2025) connected engagement monitoring with performance analytics, showing that AI can contribute not just to supervision but to measurable academic improvement.

Recent literature also shows a clear shift toward user-centered and explainable systems. Research in 2025 increasingly focuses on fairness, transparency, and human–AI collaboration. Systems are being designed to not only detect engagement but to explain predictions in interpretable ways that educators and students can understand. The integration of emotional feedback, adaptive thresholds, and real-time visualization represents a mature phase in AI-based educational research.

This evolution signifies a broader transformation of educational technology—from systems of surveillance and control to platforms of support and enhancement. The trajectory from 2019 to 2025 demonstrates that while technical progress has been substantial, the pedagogical applications of AI in learning environments are still in a formative stage. The next phase will depend on aligning technological innovation with educational psychology and ethical governance.

7.4 Future Research Directions

Future research should prioritize developing AI-based monitoring systems that emphasize learning enhancement rather than surveillance. Systems should incorporate real-time feedback and motivational alerts that guide students to regain focus when distracted. Personalized dashboards can display study duration, attention metrics, and progress trends, enabling learners to self-assess and set measurable study goals. Such self-regulating frameworks can shift the function of AI from passive monitoring to active learning support.

Creating standardized datasets and evaluation frameworks is another urgent need. Current studies employ diverse datasets, each with unique lighting, demographics, and annotation protocols, which hinders comparability. A unified dataset capturing varied learning environments and engagement states will allow for consistent evaluation and model benchmarking. Collaborative efforts among academic institutions can ensure data diversity and ethical integrity.

Integration of multimodal analytics offers another promising avenue. Combining facial expressions, gaze dynamics, posture data, and digital interactions can enable more accurate and holistic engagement assessment. Reinforcement learning and adaptive modeling could be employed to personalize engagement thresholds for each learner based on behavioral trends. These approaches would ensure that the system adapts intelligently to different learners rather than enforcing uniform criteria.

Ethical AI practices must remain central in future developments. Systems should implement privacy-by-design mechanisms, process all sensitive data locally, and include transparent consent frameworks. Future research must also assess the psychological comfort of students, ensuring that continuous monitoring promotes awareness and motivation rather than discomfort or anxiety. This requires interdisciplinary collaboration among technologists, educators, and behavioral psychologists.

Finally, there is a growing need for institutional integration and teacher–student collaboration. Embedding engagement analytics into learning management systems can help teachers identify struggling students while allowing learners to take responsibility for their study behavior. The convergence of AI analytics, pedagogy, and ethical awareness will define the next generation of intelligent learning systems that empower rather than supervise students.

8. Conclusion

This study presents a structured review of research in artificial intelligence-based student monitoring and engagement analysis conducted between 2019 and 2025. The literature demonstrates a progressive transition from exam-oriented surveillance systems to adaptive, student-centered monitoring frameworks. Foundational works such as those by Sowmya and Madhuri (2019), Varalakshmi and Kumar (2020), Singh et al. (2020), Tripathi et al. (2021), and Gupta and Goyal (2021) established the technological groundwork for facial recognition and deep learning applications in educational contexts. However, subsequent studies, including Verma (2023), Hossen (2023), Wang (2025), El-Beshbishi (2025), and Almubarak (2025), broadened the focus toward engagement assessment, emotion recognition, and ethical AI integration, marking a significant conceptual shift in the field.

The review reveals that despite notable advances in detection accuracy, real-time analysis, and model interpretability, substantial gaps remain in the areas of personalization, multimodal data fusion, and long-term behavioral validation. Most systems are still designed for external supervision rather than self-regulated learning enhancement. The absence of standardized datasets, evaluation benchmarks, and ethical governance mechanisms further limits scalability and adoption across educational institutions.

Future research should emphasize the design of adaptive, privacy-preserving, and human-centered AI systems capable of providing constructive feedback rather than passive observation. Integrating multimodal engagement cues, reinforcement learning-based personalization, and transparent decision-making can significantly enhance the impact of AI on academic focus and motivation. The convergence of technical innovation, pedagogical insight, and ethical accountability will define the next generation of intelligent educational tools, transforming AI-based monitoring from surveillance technology into an empowering instrument for continuous learning and self-improvement.

This comprehensive review thus contributes to the understanding of AI's evolving role in education and provides a foundation for future researchers to address existing challenges in engagement analysis, data ethics, and learner-centered system design. It highlights the potential of AI not merely as a tool for observation but as a collaborative agent supporting effective study behavior and sustainable academic growth.

9. Acknowledgement

The authors express their sincere gratitude to the faculty members and mentors of their institution for their continuous guidance and encouragement throughout this study. Their valuable insights and academic support were instrumental in shaping the direction and depth of the research. The authors also acknowledge the availability of scholarly resources and technical facilities that facilitated the comprehensive review of literature. Finally, appreciation is extended to peers and collaborators whose constructive feedback contributed significantly to the refinement of this paper and its overall academic quality.

References

1. Sowmya, V., and Madhuri, K., “AI-Based Online Exam Proctoring System Using Facial Recognition,” International Journal of Computer Applications, Vol. 182, No. 46, pp. 23–27, 2019.
2. Varalakshmi, M., and Kumar, P., “Student Monitoring System Using Deep Learning in Online Classrooms,” International Journal of Emerging Technologies in Learning (iJET), Vol. 15, No. 12, pp. 45–53, 2020.
3. Singh, R., Sharma, P., and Tiwari, N., “Face Liveness Detection for Online Examination Security,” IEEE Access, Vol. 8, pp. 107147–107159, 2020.
4. Tripathi, A., Jain, S., and Mishra, R., “Deep Learning-Based Remote Exam Monitoring System,” Procedia Computer Science, Vol. 192, pp. 579–588, 2021.
5. Gupta, N., and Goyal, A., “AI-Powered Student Activity Tracking During E-Learning,” Journal of Intelligent Systems and Applications, Vol. 13, No. 2, pp. 95–104, 2021.
6. Verma, N., “Designing an Artificial Intelligence Tool to Understand Student Engagement in Online Learning Environments,” ScienceDirect, Vol. 14, pp. 231–244, 2023.
7. Hossen, M. K., “Attention Monitoring of Students During Online Classes Using AI,” ScienceDirect, Vol. 11, No. 3, pp. 99–110, 2023.
8. Wang, Y., “Artificial Intelligence in Student Management Systems to Enhance Academic Performance Monitoring,” Nature Education, Vol. 2, pp. 42–56, 2025.
9. El-Beshbishi, S. N., “Use of AI Machine Learning Algorithms to Assess Medical Students’ Engagement,” Education for Health, Vol. 38, No. 1, pp. 12–22, 2025.
10. Almubarak, A., “An AI-Powered Framework for Assessing Teacher–Student Interactions,” Frontiers in Artificial Intelligence, Vol. 6, pp. 1–10, 2025.
11. Gjermeni, F., and Prodani, F., “AI and Student Engagement: A Comparative Analysis,” Interdisciplinary Journal of Research and Development, ResearchGate, Vol. 12, No. 4, 2024.
12. Karim, B. Q., Haryanto, H., and Susanti, E. Y., “AI in Education: Transforming Student Engagement for the Digital Age,” Jurnal Penelitian Pendidikan IPA, Vol. 10, No. 1, pp. 33–42, 2025.
13. Zhang, X., “Smart Classrooms: How Sensors and AI Are Shaping Engagement Monitoring,” PMC Open Education, Vol. 7, No. 2, pp. 17–26, 2024.
14. Gao, Y., “Feasibility and Usability of an AI-Powered Gamification Intervention,” ScienceDirect, Vol. 20, No. 3, pp. 115–123, 2025.
15. Ghosh, S., “AI-Powered System to Monitor and Analyze Student Performance,” International Research Journal of Modern Science, Vol. 8, No. 5, pp. 65–74, 2025.
16. Zion, Y. B., “AI-Based Teaching Evaluations: How Well Do They Reflect Student Engagement?” ScienceDirect, Vol. 18, No. 2, pp. 211–226, 2025.
17. Pillai, A. S., “Student Engagement Detection in Classrooms Through Computer Vision and Deep Learning,” SSRN Electronic Journal, 2022.
18. Hossen, M. K., and Rahman, M. T., “Deep Learning Framework for Student Focus and Attention

Monitoring," IEEE Transactions on Learning Technologies, Vol. 15, No. 4, pp. 300–312, 2023.
19. Chaudhary, S., et al., "The Impact of AI-Powered Educational Tools on Student Engagement and Learning Outcomes at Higher Education Level," ResearchGate, 2024.
20. Artificial Intelligence-Supported Student Engagement (AIsE): Trends and Systematic Analysis," European Journal of Education, Wiley Online Library, Vol. 60, No. 3, pp. 405–418, 2025.
21. AI-Based Student Engagement Detection: Leveraging Deep Learning Approaches," International Journal of Intelligent Systems and Applications in Engineering (IJISAE), Vol. 12, No. 2, pp. 77–88, 2024.
22. Alsaiari, O., Baghaei, N., Lodge, J., Lahza, H., and Khosravi, H., "Emotionally Enriched Feedback via Generative AI," arXiv preprint arXiv:2501.04827v2, 2025.
23. Hao, Z., Cao, J., Li, R., Yu, J., Liu, Z., and Zhang, Y., "Mapping Student-AI Interaction Dynamics in Multi-Agent Learning Environments: Supporting Personalised Learning and Reducing Performance Gaps," arXiv preprint arXiv:2502.09177v1, 2025.
24. Systematic Insights and Trends in AI-Based Student Engagement Tracking," International Journal of Emerging Technologies in Learning (iJET), Vol. 20, No. 6, 2025.
25. "The Impact of Artificial Intelligence (AI) on Students' Academic Outcomes," Education Sciences (MDPI), Vol. 15, No. 5, pp. 1–12, 2025.
26. "How AI Detects Student Engagement to Transform Classrooms," Digital Promise Report, 2024.
27. "The Usage of AI in Teaching and Students' Creativity: The Mediating Role of Learning Engagement," PMC, Vol. 19, No. 3, 2025.
28. "Student Engagement and AI: Research Overview and Findings," Online Learning Consortium, 2023.
29. AI in Schools: Pros and Cons," University of Illinois College of Education, 2024.
30. How Students Use AI: The Evolving Relationship Between AI and Higher Education," Digital Education Council Report, 2024.
31. The Transformative Potential of AI in Student Affairs: Recommendations for Student Affairs Leaders," NASPA White Paper, 2024.
32. Alsaiari, O., and Boden, M., "Emotion Detection in Smart Classrooms Using Hybrid CNN-LSTM Architectures," IEEE Access, Vol. 13, pp. 9002–9014, 2025.
33. Alabduljabbar, R., "User-Centric AI: Evaluating the Usability of Generative AI Applications," PMC Journal of Human-Centered Computing, Vol. 4, No. 1, 2024.
34. Yeung, S., "University Students' Engagement with Generative AI Feedback," ScienceDirect, Vol. 16, No. 3, pp. 245–256, 2025.
35. Gjermeni, F., "Artificial Intelligence and Student Focus Monitoring: Systematic Evaluation of Deep Learning Models," ResearchGate, 2024.
36. Almubarak, A., "AI-Powered Frameworks for Teacher–Student Interaction in Hybrid Learning Environments," Frontiers in AI, Vol. 5, 2025.
37. Systematic Review of Attention Recognition Using Computer Vision in Education," International Journal of Artificial Intelligence and Education, Vol. 31, No. 4, pp. 312–328, 2023.
38. Ghosh, R., and Basu, D., "Vision-Based Learning Analytics: Identifying Student Attention and Fatigue Using AI," Springer Advances in Intelligent Systems, Vol. 9, pp. 210–222, 2023.
39. AI-Based Drowsiness and Engagement Detection Using Deep Neural Networks," IEEE International Conference on Advanced Computing (ICAC), 2022.
40. Cognitive and Emotional Monitoring Through Machine Learning in E-Learning Platforms,"

ScienceDirect, Vol. 22, No. 5, pp. 301–312, 2024.
41. Rahman, M. T., and Chowdhury, S. M., "Deep Reinforcement Learning Framework for Student Engagement Recognition in Virtual Classrooms," IEEE Transactions on Artificial Intelligence in Education, Vol. 4, No. 2, pp. 125–138, 2025.
42. Lee, D. J., and Kwon, J. H., "Facial Emotion Recognition-Based Attention Tracking System for Remote Learning Environments," Sensors (MDPI), Vol. 25, No. 3, pp. 1–14, 2025.
43. Mitra, A., and Das, P., "Hybrid CNN–GRU Model for Detecting Student Distraction in Online Learning," International Journal of Machine Learning and Cybernetics, Springer, Vol. 16, No. 2, pp. 305–319, 2025.
44. Patel, S., and Mehta, D., "AI-Based Visual Attention Estimation for Enhancing E-Learning Interactions," Procedia Computer Science, Elsevier, Vol. 230, pp. 455–463, 2024.
45. Alghamdi, R., and Khan, M., "Real-Time Monitoring of Student Engagement Using Deep Neural Networks and Gaze Tracking," Computers and Education: Artificial Intelligence, Elsevier, Vol. 6, Article 100195, 2024.
46. Li, Z., Wang, L., and Zhao, J., "Explainable AI for Emotion-Aware Educational Systems: A Human-Centered Study," IEEE Access, Vol. 12, pp. 17895–17910, 2024.
47. Park, S., and Choi, H., "Multi-Modal Learning Analytics for Predicting Student Focus and Fatigue Levels," Education and Information Technologies, Springer, Vol. 29, No. 1, pp. 423–439, 2024.
48. Hernandez, M., and Gonzalez, F., "A Federated Learning Approach to Privacy-Preserving Student Behavior Recognition," Computers & Education, Elsevier, Vol. 205, pp. 104589–104602, 2024.
49. Yang, R., and Qiu, T., "Attention and Emotion Recognition in Smart Classrooms Using Vision Transformers," Neural Computing and Applications, Springer, Vol. 36, No. 5, pp. 3887–3904, 2023.
50. Akter, S., and Ahmed, K., "Real-Time Focus Estimation in Online Learning Using Lightweight CNN Models," International Journal of Emerging Technologies in Learning (iJET), Vol. 18, No. 10, pp. 112–124, 2023.