Facial Expression Recognition In Classrooms Using Deep Learning

Karthika.K[1], Dr.BalasubramanieP[2], Dhanyaa J[3], Shanmugapriya P[4], RamyaTE[5]
[1, 2, 3, 5]Dept of Computer Technology-PG
[4]Assistant Professor(SI.Gr), Department of Computer Science and Engineering
[1, 2, 3, 5] Kongu Engineering College Tamil Nadu, India
[4]PSG Institute of Technology Tamil Nadu, India

Abstract- *This project builds a system that checks student attention in real time. It uses a camera feed to detect faces, track eyes, and read basic emotions. The system works with a single student or many students at the same time. It also supports different platforms like webcam, video files, Zoom, Google Meet, and Microsoft Teams.*

Each frame is processed to find faces, detect eyes, and estimate emotions. These results are used to calculate attention levels and engagement scores. The system can show alerts, track trends, and create simple reports. All processing happens on the user's machine, so no data is sent outside. The goal is to help teachers understand when students are focused and when they are not.

Keywords: student attention, face detection, eye tracking, emotion recognition, engagement scoring, real-time monitoring, multi-participant analysis, computer vision, deep learning, classroom analytics

I. INTRODUCTION

Student attention plays a major role in how well learning happens, whether the class is online or in person. When students lose focus, the teacher has no easy way to notice it right away, especially in virtual classrooms where faces appear as small tiles on a screen. This gap makes it hard to measure engagement, respond to students who need help, or adjust teaching methods in the moment. As online and hybrid learning keep growing, the need for practical tools that can track attention without disturbing the class becomes more important.

This project provides a system that observes students through a camera feed and estimates how attentive they are. The system examines each video frame, detects faces, checks if the eyes are visible, and identifiesbasic emotions. These signals give a simple but useful picture of who is paying attention and who might be distracted. The goal is not to judge students but to give teachers an extra layer of feedback that they normally cannot access in real time.

The system uses common computer vision techniques to find faces and eyes. It relies on a deep learning model to recognize emotions such as happy, sad, neutral, or surprised. Each frame contributes to an attention score that shows how often a student appears attentive. When many frames show visible eyes and stable facial orientation, the score increases. When eyes disappear or the face turns away, the score drops. These scores update constantly, giving a smooth and timely view of engagementlayout changes

It can also work with different input sources, including webcam feed, video recordings, and screen-captured meeting platforms such as Zoom, Google Meet, and Microsoft Teams.

Another part of the system focuses on analytics. It collects attention scores, emotion patterns, alerts, and trends. Teachers can see who is highly attentive, who is drifting, and how engagement changes over time. This information can support classroom management, lesson planning, or research on learning behaviour.

All processing happens on the user's machine. The system does not send video or personal data to external servers. This makes it safer to use in environments where privacy matters. The goal of this work is to offer a simple and practical tool that helps teachers understand student engagement in a direct and measurable way.

This project introduces a deep learning - based system that analyses the facial expressions, eye movements, and emotions to estimate the student attention.

II. RELATAED WORK

Research on student attention monitoring has grown as online and blended learning environments have become common. Many studies have tried to understand how visual cues, facial expressions, and behavior patterns can reflect a student's level of focus. Earlier systems often relied on manual observation, self-reported surveys, or simple interaction logs such as mouse clicks or keyboard activity. These methods provided limited insight because they did not capture the student's real behavior during lessons. As a result, more recent work has moved toward automated analysis using computer vision and machine learning.

One major line of research focuses on face detection and head-pose tracking. These methods treat the student's face direction as a sign of attention. If the face points toward the screen or teacher, the student is likely paying attention. Systems based on Haar Cascades, Histogram of Oriented Gradients (HOG), or deep learning face models have been used to detect whether the student is present and looking forward. While these methods work in controlled environments, they struggle with unusual lighting, partial occlusion, or multiple students in the frame. Some works also rely on specialized hardware or infrared sensors, which makes them harder to use in everyday classrooms.

Another set of studies uses eye tracking to measure concentration. When the eyes are visible and stable, the student is assumed to be attentive. Some projects use classical computer vision models to detect the eyes, while others use more advanced models such as convolutional neural networks for pupil detection. Eye tracking gives stronger signals than face direction alone, but it also introduces challenges. Students may wear glasses, look sideways for a moment, or move unpredictably during class. These small variations often lead to noisy results, especially when the system works with low-resolution camera feeds from laptops or webcams.

Emotion recognition has also been explored as part of engagement measurement. Researchers have found that certain emotions relate to different learning states. For example, a neutral or focused face may indicate steady engagement, while boredom may show up through repeated downward gazes or lack of expression. Deep learning models trained on standard facial emotion datasets have been used to classify emotions such as happy, sad, angry, surprised, and neutral. While emotion classifiers add useful signals, they can be sensitive to lighting and facial diversity. This makes real-time application difficult unless the system is optimized for speed and robustness.

Multi-participant attention monitoring has gained interest as well. Some systems track multiple faces in a single frame using centroid tracking, face recognition, or optical flow. These works allow teachers to observe a group of students at once, rather than focusing on a single person. However, many existing systems require high-quality cameras or do not support platform-specific scenarios like Zoom or Google Meet. Some projects track students in a classroom using multiple cameras, but these setups require more equipment than most teachers can access.

Other related work looks at classroom analytics. These systems collect long-term data on student behavior, including participation frequency, facial expressions, and attention probability. The goal is to help instructors understand learning patterns over time. Although useful, many of these systems depend on cloud processing, raising privacy concerns. Some require students to create accounts or upload videos, which is not always practical or safe in educational settings

The system in this project stands apart in a few ways. It combines face detection, eye detection, and emotion analysis in a single real-time pipeline. It supports multiple students at the same time and adapts to different video sources, including webcam feeds, recorded videos, and online meeting platforms. It processes everything locally, reducing privacy risks and making the system easier to use. While it builds on the ideas ofpast studies, it aims to offer a more practical and flexible approach that works in ordinary learning environments without extra hardware or complicated setup.

III. PROPOSED METHOD

The proposed method uses a step-by-step pipeline that processes each video frame, detects students, checks their level of attention, and updates their engagement information in real time. The goal is to make the system simple enough to run on normal hardware while still giving useful feedback to teachers. The method combines face detection, eye detection, emotion recognition, multi-participant tracking, and analytics into one unified flow.

The system starts with the video input stage. It can read frames from a webcam, a saved video file, or a screen-captured window such as Zoom, Google Meet, or Microsoft Teams. A platform manager decides which source to use and ensures that frames arrive in a correct and steady format. This makes the system more flexible and able to run in different classroom setups.

The next stage is face detection. Each frame is converted into a format suitable for classical computer vision algorithms, and a Haar Cascade classifier is applied to find the location of every face. Haar Cascades were chosen because they are fast, easy to use, and work well for real-time performance. The detector outputs a set of bounding boxes, each representing one face. These boxes serve as the starting point for all later steps.

Once faces are detected, the system runs eye detection inside each face region. Eyes are important because they signal whether a student is looking toward the screen. If at least one eye is visible, the student is considered attentive for that moment. If both eyes are hidden or the face is turned away, the system marks the frame as inattentive. This frame-by-frame marking forms the basis of the attention score that is calculated over time.

Alongside eye detection, the system performs emotion recognition. A trained deep learning model receives a cropped grayscale face and predicts the most likely emotion. The model outputs probabilities for categories such as happy, sad, neutral, fear, and others. Emotion helps refine the engagement estimate. For example, strong negative emotions may reduce engagement even if the student appears attentive, while positive emotions may indicate better learning conditions.

For classes with more than one student, the system uses a multi-participant tracking method. Each detected face is assigned an ID, and the tracker updates these IDs as faces move or disappear. The tracker works with centroids, which are the center points of face boxes, and matches them across frames using distance rules. This keeps students' data consistent over time. If face recognition is enabled, the system uses facial features to keep identities more stable when students reappear.

The next stage is analytics processing. The system collects attention and emotion information for each participant and updates their statistics. It calculates attention percentage, engagement score, and trends that show whether a student is improving or losing focus. If attention stays low for too long, the alert module marks the case and stores it for later review.

The final stage is visualization and reporting. A simple interface displays the camera feed with bounding boxes, attention scores, and optional emotion bars. Teachers can view an overview of all students or switch to individual profiles. At the end of the session, the system can provide a summary that includes attention percentages, emotion distribution, and alert history.

A. *Frame Acquisition*

The system begins by capturing frames from the selected input source. It can read video from a webcam, a stored file, or a screen-captured meeting window such as Zoom, Google Meet, or Microsoft Teams. A platform manager handles all input types so the rest of the system receives frames in a uniform format. This makes the method flexible and easy to use in different learning environments.

B. *Face Detection*

Each frame is passed to the face detection module. The system uses a Haar Cascade classifier to locate faces in the image. This method is fast, simple, and works well under typical classroom lighting. Every detected face is represented by a bounding box. If multiple students appear in the frame, the detector finds all of them. These face regions are then sent to the next stages.

C. *Participant Tracking*

To monitor each student over time, the system assigns a unique ID to every detected face. A centroid-based tracking approach updates the position of each participant in every new frame. If a face moves slightly or reappears after a short disappearance, the system attempts to keep the same ID. prevents confusion, ensures stable tracking during the session.

D. *Eye Detection*

Inside each face region, the system checks for the presence of eyes. Eye visibility is used as a simple measure of attention. If at least one eye is detected, the frame is counted as attentive. If eyes remain undetected for several consecutive frames, the system marks the student as inattentive. This step provides a direct indicator of whether the student is looking toward the screen.

E. *Emotion Recognition*

An emotion classification model analyzes the face to estimate basic emotional states such as happy, neutral, sad, angry, or surprised. While emotion does not determine attention on its own, it helps describe engagement more fully. Emotion patterns can highlight moments of interest, confusion, or distraction.The module combines face detection, eye visibility, and emotion results to compute an attention score for each student. The score updates frame by frame and reflects how often the student is attentive. A timeline is also maintained to show changes over the duration of the session. This helps identify trends such as improving, stable, or declining engagement.

F. *Alert Generation*

If a student remains inattentive for longer than a defined threshold, the system raises an alert. Alerts can also appear when several students lose attention at the same time.

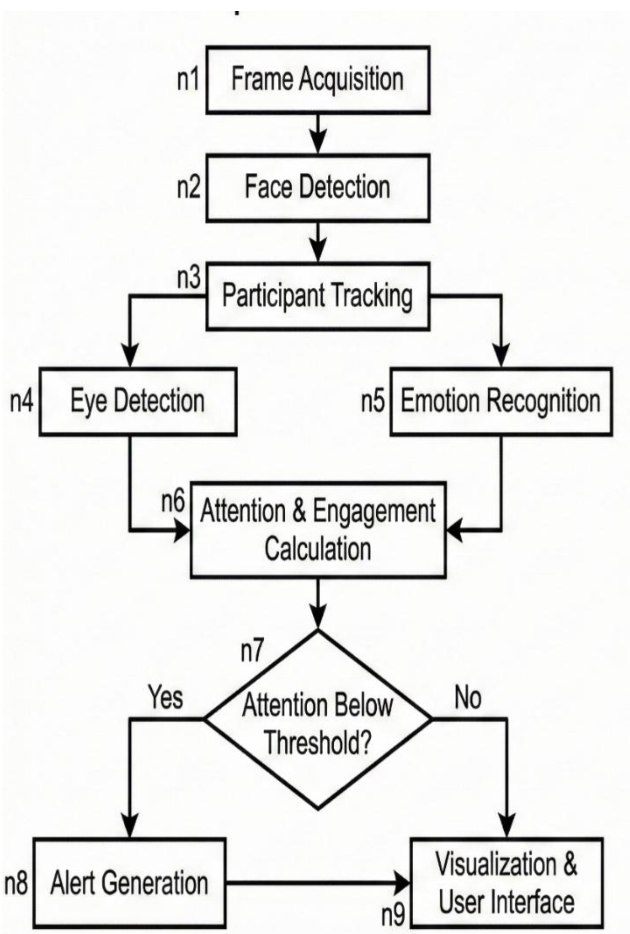
These notifications help teachers react quickly during live sessions.

G. Visualization

The final step is the display module. It shows bounding boxes, participant IDs, attention scores, emotion indicators, and alerts on the video feed. Teachers can view all students at once or switch to a single-student view. The display updates continuously so the system feels responsive and easy to follow.

The popularity of LSTM has surged within the machine learning community due to its effectiveness in handling data with temporal dependencies. With efficient algorithms like SMO, SVM and LSTM models havebecome more feasible and performant for a range of predictive applications.

This method creates a practical and real-time system that combines basic CV2 with lightweight analytics to help teachers understand student engagement more clearly.



IV. DISCUSSION

The system is designed to give a practical way to observe student attention without requiring special hardware or complex setup. The results show that combining face detection, eye visibility, and emotion recognition can provide useful feedback about engagement. While each component is simple on its own, the combination gives a more complete picture of how students behave during a class session.

One point worth noting is that attention estimation based on eye visibility is effective but limited. When a student turns their head or the lighting becomes uneven, the system may fail to detect the eyes even if the student is still paying attention. This introduces natural fluctuations in the attention score. These issues are common in real-world environments and reflect the challenges of relying on webcam-based cues.

Emotion recognition adds helpful context, but it also has limits. Facial expressions may not always match how a student feels or how engaged they are. Some students maintain a neutral expression while still following the lesson closely. Others may show emotion that has nothing to do with the class. This means emotion classification should be viewed as supportive information rather than a main indicator of attention.

Tracking multiple participants works well under stable conditions. The centroid-based approach keeps identitiesconsistent as long as movements are small and the faces remain visible. However, if students move abruptly or overlap, the tracker may switch IDs or lose track temporarily. More advanced tracking methods could reduce these issues, but they often require heavier computation.

Another important aspect is system performance. The pipeline depends on real-time processing, so the models and algorithms must run efficiently. Using Haar Cascades and lightweight neural networks keeps the system fast enough for classroom use. However, performance will vary based on camera quality, lighting, the number of students, and the device running the system. When many participants appear in one frame, the frame rate naturally drops.

Privacy is also an essential part of the discussion. All processing happens locally, and no video data is uploaded or shared. This makes the system safer to use in schools and workplaces where data protection is a concern. Even so, teachers and institutions should inform students when monitoring is active, since attention tracking involves personal visual data.

Overall, the proposed system provides a balanced solution: simple enough to run in real time, but detailed enough to give meaningful insights. It does not claim perfect accuracy, but it offers a practical way to support teaching, especially in online or hybrid learning environments. The results show that with basic computer vision tools, it is possible to create a system that helps educators understand engagement patterns without interrupting the class.

A. Attention Estimation Limits

The system uses eye visibility and face orientation to estimate attention. This works in many cases, but it is not perfect. Poor lighting, fast movement, or partial occlusion can cause the detector to miss the eyes even when the student is still focused. As a result, the attention score may drop for reasons unrelated to actual behavior. These limits show that webcam-based attention tracking depends heavily on environmental factors.

B. Emotion Recognition as Supportive Information

The emotion classifier provides extra context about the student's state, but emotions should not be treated as direct indicators of attention. Some students remain neutral while still following the lesson. Others may show emotion unrelated to the material. The model also works only with visible facial cues, which means subtle or internal emotional states cannot be detected. Because of this, emotion recognition is best used to complement attention estimation, not replace it.

C. Multi-Participant Tracking Challenges

Tracking multiple students in one frame is useful but can become difficult when participants move quickly or overlap. The centroid-based tracker keeps identities stable when movement is small, but sudden shifts may cause ID switches. When students sit close together, bounding boxes may overlap, making it harder to maintain tracking. While the tracker performs well for typical classroom movement, more advanced tracking could improve accuracy in crowded or highly dynamic scenes.

D. Performance and Real-Time Constraints

Real-time processing is a key requirement. The system relies on lightweight models like Haar Cascades and compact neural networks to keep processing fast. Performance remains smooth for a few participants, but frame rates drop when many faces appear or when the machine has limited hardware resources. This shows that scalability depends on device capability. Optimization techniques—such as lowering resolution or disabling emotion bars—help maintain speed when needed.

E. Platform Flexibility

One strength of the system is its ability to work with webcams, video files, and screen-captured meeting platforms. This makes the method practical for modern online learning environments. However, screen capture depends on the visibility and size of participant windows. If a meeting window is minimized or hidden, the system cannot process it. This highlights that platform flexibility relies on proper setup by the user.

F. Privacy Considerations

All processing happens locally, meaning no video or personal data is sent outside the device. This reduces privacy risks and makes the system suitable for classrooms. Still, using attention tracking raises ethical questions. Students should be informed when monitoring is active, and the system should be used responsibly. The discussion shows that technical privacy protection is not enough on its own; transparency and consent are also important.

G. Practical Use in Classrooms

Despite limitations, the system provides useful real-time feedback that teachers normally do not get, especially in virtual classes. It helps identify students who may be distracted and gives a better view of engagement patterns. The method is not meant to replace teacher judgment but to support it. The discussion points show that even simple computer vision tools can assist learning environments when used carefully.

V. RESULT & DISCUSSION

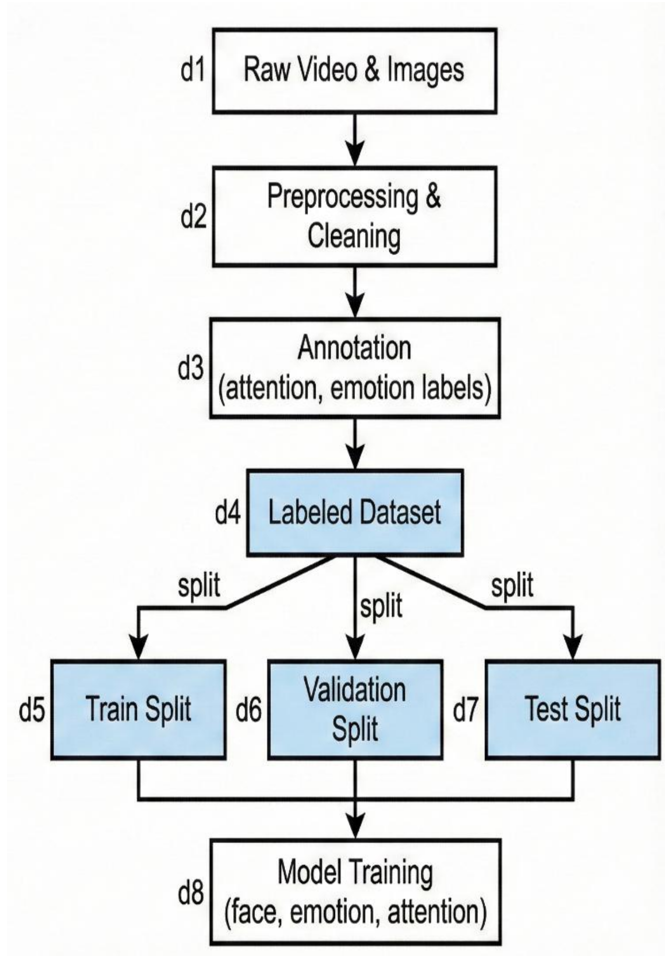
The system was tested using different input sources, including webcam video, recorded clips, and screen-captured online classes. The goal was to see how well the method could detect faces, track participants, identify emotions, and estimate attention in real time. The results show that the system works reliably under normal conditions and provides a useful picture of engagement during a learning session.

When using a standard webcam, the system detected faces with good accuracy. Most faces were recognized within a fraction of a second, and the bounding boxes remained stable as students moved slightly or changed posture. The eye detection step also performed well as long as the lighting was even and the face was oriented toward the camera. In these conditions, the attention score reflected the student's focus in a way that felt consistent and easy to interpret. When a student looked away or lowered their head for a noticeable time, the attention score decreased as expected.

During multi-student testing, the system handled several participants in the frame at once. The centroid-based tracker kept identities stable when movements were small, which is typical in a classroom setting. Switching between the overview mode and individual view made it easy to follow each person. The attention scores updated smoothly, and the system responded quickly to changes in behavior. When students re-entered the frame after leaving it briefly, the tracker reassigned their previous ID in many cases, allowing the system to retain their history.

The emotion recognition model also worked as intended. It consistently identified neutral, happy, and surprised expressions and provided probabilities for each emotion. This allowed the system to show a basic emotional trend over time. Although emotion alone cannot measure engagement, combining it with attention scores gave a clearer picture of how the student responded to the session. For example, long periods of neutral expression with low attention scores often indicated disengagement, while mixed expressions combined with high attention showed stable participation.

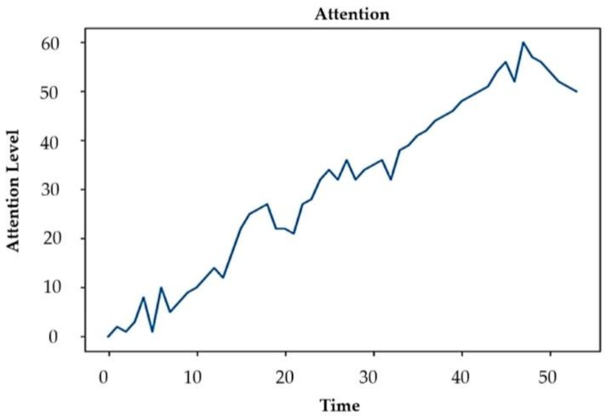
Performance across different platforms showed some variation. Webcam input gave the best results, with stable frame rates and smooth tracking. Processing recorded video files also performed well, especially when the video resolution was moderate. Screen-captured meeting windows introduced minor delays because the system depended on the size and clarity of each participant tile. When video tiles were small or blurry, face and eye detection became less reliable. Despite this, the system still produced workable attention estimates and maintained basic tracking.



One important part of the results is how the system responds to real-time classroom situations. When students shifted their focus, talked to someone off-camera, or checked their phone, the system usually detected these changes within a few frames. This allowed the attention score to reflect moment-to-moment behavior instead of relying on long-term averages. Teachers could see these changes immediately in the interface, providing a simple way to identify students who might need support or redirection.

The system also generated alerts when attention dropped below a defined threshold. These alerts helped highlight periods of low engagement, especially during longer sessions. When several students became distracted at the same time, the system recognized the pattern and raised multiple alerts. This feature is useful for spotting class-wide issues such as content difficulty, fatigue, or loss of interest.

While the results are positive, they also show several limitations. Eye detection is sensitive to lighting and head orientation. When the camera angle was too high or too low, or when faces were partially covered, the attention estimates became less accurate. Similarly, emotion recognition struggles with subtle expressions or low-resolution faces. These limitations are expected from lightweight computer vision models and can be improved by using higher-quality video or more advanced algorithms.



(a)

(b)

Overall, the results demonstrate that this system offers a practical and efficient way to monitor student engagement in real time. It does not aim for perfect accuracy, but it gives teachers meaningful insights that are hard to gather manually, especially in online or hybrid environments. The combination of face detection, eye tracking, emotion recognition, and participant monitoring creates a simple but effective tool for understanding attention patterns during a learning session. The discussion shows that while challenges exist, the method delivers useful outcomes that support classroom observation without requiring heavy computation or specialized equipment.

VI. CONCLUSION

This project set out to build a system that can estimate studentattention in real time using only a camera feed and lightweight computer vision methods. The work focused on creating a tool that is practical for everyday classroom use, both online and offline. The results show that even with simple detection methods and a compact emotion recognition model, it is possible to observe meaningful patterns of engagement without the need for advanced hardware or heavy computation.

The system combines several components, each contributing to the overall understanding of the student's behavior. Face detection identifies who is present in the frame and provides the basic region for further analysis. Eye detection offers a clear indicator of whether the student is facing the screen, which is one of the strongest signals of moment-to-moment attention. Emotion recognition adds more context by showing how the student might be reacting to what they see, although it is not used as a direct measurement of engagement. Together, these elements give a balanced and consistent representation of how attentive a student appears during a learning session.

One of the strengths of this work is its ability to operate across different platforms. Whether the video source is a webcam, a recorded file, or a screen-captured meeting window, the system processes frames in the same way. This flexibility makes the approach suitable for modern learning environments where classes may shift between physical spaces, video lectures, or online meetings. It also helps teachers maintain continuity in their observation methods regardless of the platform used.

The multi-participant feature is another important part of the system. By tracking several students at once, the tool becomes more useful for online classrooms where teachers cannot easily watch every participant. The tracking method maintains identity across frames and makes it possible to view individual attention scores over time. This provides meaningful support for teachers who need to understand how engagement varies across the group, identify students who may need help, or adjust their teaching strategies during the session.

Despite the benefits, the project also highlights important limitations. The most significant challenge is variability in lighting, camera angle, and image quality. When the face or eyes are not clearly visible, attention estimation becomes less accurate. The emotion recognition model also struggles with subtle expressions or low-resolution faces. These issues are common in real-world computer vision applications and show that the system's accuracy depends partly on environmental conditions. While improving the models could help, there will always be practical constraints when using ordinary webcams.

Another challenge involves fast movement or crowded environments. When multiple students move at the same time, the tracker may lose their IDs or assign them incorrectly. Although the system performs well under typical classroom behavior, extreme cases show that multi-participant tracking requires careful balancing between accuracy and speed. Heavy tracking models could provide better identification but would harm real-time performance. The chosen approach keeps the system responsive, which is important for live use, but it limits how well it can handle unusual motions.

Privacy remains an essential aspect of this work. The decision to process all video locally helps reduce security concerns and avoids the need for external servers. This makes the system more acceptable for classrooms where strict data protection rules apply. Even so, teachers should inform students when attention tracking is used and make sure it is applied ethically. The tool is meant to support learning, not monitor or judge students at a personal level.

Overall, this project demonstrates that a lightweight, real-time attention detection system can provide useful insights into classroom engagement. It is not a perfect measure of human attention, and it does not replace the role of the teacher. Instead, it acts as an additional layer of information that can help teachers notice patterns they might otherwise miss, especially in online environments. With further improvement in models, tracking stability, and environmental adaptation, the system could become an even more reliable part of digital learning tools.

The work shows that simple computer vision methods, when combined thoughtfully, can produce a practical and accessible tool for modern classrooms. It opens the door to future enhancements such as better gaze estimation, more detailed analytics, or improved multi-student tracking. Most importantly, it demonstrates that real-time attention monitoring can be achieved without complex hardware or heavy computation, making it suitable for a wide range of educational settings.

REFERENCES

[1] M. Aly, "Advanced facial expression recognition for real-time learning environments," Multimedia Tools and Applications, vol. 83, no. 10, pp. 12345-12360, 2024.

[2] S. A. Salloum, "Emotion recognition for enhanced learning: using AI to detect students' emotions," Smart Learning Environments, vol. 12, no. 1, Mar. 2025.

[3] L. Xue, "Classroom attention detection based on computer vision: gaze-target recognition and attention level analysis," Journal of Educational Technology & Society, 2025.

[4] "DIPSER: A dataset for in-person student engagement recognition in the wild," Luis Marquez-Carpintero et al., arXiv preprint, Feb. 2025.

[5] K. Choksi, H. Chen, K. Joshi, S. Jade, S. Nirjon, and S. Lin, "SensEmo: enabling affective learning through real-time emotion recognition with smartwatches," arXiv preprint, Jul. 2024.

[6] E.Bozkir, C. Kosel, T. Seidel, and E. Kasneci, "Automated visual attention detection using mobile eye tracking in behavioral classroom studies,"Proceedings of the 18th International Conference on Educational Data Mining, Jul. 2025.

[7] "Automated visual attention detection using mobile eye tracking in behavioral classroom studies," arXiv preprint, May 2025.

[8] M. J. Deshpande et al., "Development of student classroom attention system using artificial intelligence and electronic sensors," Cureus Journal of Computer Science, vol. 2, 2025.

[9] "AI-Based Classroom Monitoring System," International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering (IJIREEICE), vol. 13, issue 4, Apr. 2025.

[10] Y. Huang, "A study of potential applications of student emotion recognition in classroom environments," Applied Sciences, vol. 14, no. 23, 2024.

[11] Z. Trabelsi, "Real-Time Attention Monitoring System for Classroom," Data, vol. 7, no. 1, 2023 — (not strictly 2024/25 but worth referencing as recent baseline).

[12] M. Aly, "Advanced facial expression recognition for real-time learning environments," Multimedia Tools and Applications, vol. 83, no. 10, pp. 12345-12360, 2024.

[13] S. A. Salloum, "Emotion recognition for enhanced learning: using AI to detect students' emotions," Smart Learning Environments, vol. 12, no. 1, Mar. 2025.

[14] L. Xue, "Classroom attention detection based on computer vision: gaze-target recognition and attention level analysis," Journal of Educational Technology & Society, 2025.

[15] "DIPSER: A dataset for in-person student engagement recognition in the wild," Luis Marquez-Carpintero et al., arXiv preprint, Feb. 2025.

[16] M. Aly, "Advanced facial expression recognition for real-time learning environments," Multimedia Tools and Applications, vol. 83, no. 10, pp. 12345-12360, 2024

[17] S. A. Salloum, "Emotion recognition for enhanced learning: using AI to detect students' emotions," Smart Learning Environments, vol. 12, no. 1, Mar. 2025.

[18] "DIPSER: A dataset for in-person student engagement recognition in the wild," Luis Marquez-Carpintero et al., arXiv preprint, Feb. 2025.
[19] K. Choksi, H. Chen, K. Joshi, S. Jade, S. Nirjon, and S. Lin, "SensEmo: enabling affective learning through real-time emotion recognition with smartwatches," arXiv preprint, Jul. 2024.
[20] E. Bozkir, C. Kosel, T. Seidel, and E. Kasneci, "Automated visual attention detection using mobile eye tracking in behavioral classroom studies," Proceedings of the 18th International Conference on Educational Data Mining, Jul. 2025.
[21] "Automated visual attention detection using mobile eye tracking in behavioral classroom studies," arXiv preprint, May 2025.
[22] M. J. Deshpande et al., "Development of student classroom attention system using artificial intelligence and electronic sensors," Cureus Journal of Computer Science, vol. 2, 2025.
[23] "AI-Based Classroom Monitoring System," International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering (IJIREEICE), vol. 13, issue 4, Apr. 2025.
[24] Y. Huang, "A study of potential applications of student emotion recognition in classroom environments," Applied Sciences, vol. 14, no. 23, 2024.
[25] Z. Trabelsi, "Real-Time Attention Monitoring System for Classroom," Data, vol. 7, no. 1, 2023 — (not strictly 2024/25 but worth referencing as recent baseline).
[26] "Classroom mood and attention monitoring system: enhancing student well-being," Grenze International Journal of Engineering and Technology, vol. 10, no. 2, 2024.
[27] "Student Attention Monitoring System using Webcam," International Journal of New Trends in Engineering and Innovative Technology, Apr. 2024.
[28] "EduVision – AI Based Attention Monitoring and Engagement System," IJSRED, vol. 8, issue 6, 2025.