

e-ISSN: 2583-9241

**Volume 09 Issue 03
Sep-Dec, 2026**

***Corresponding**

Author: Miss. Snehal Phad, Department of Information Technology, Pravara Rural Engineering College, Loni, Ahilyanagar (M.S.) India

Real Time Attention Monitoring System in Classrooms

¹Miss. Snehal Phad, ²Mr. Akash Dange, ³Mr. Arjun Waman, ⁴Miss. Yukta Gaikwad, ⁵Mr. D. B. Ghorpade

¹⁻⁴Student, ⁵Assistant Professor, Department of Information Technology, Pravara Rural Engineering College, Loni, Ahilyanagar (M.S.) India

ASBSTRACT

A real-time attention monitoring system in classrooms focuses on understanding how attentive students are during lectures using modern technology. It combines computer vision and machine learning to observe student behavior without interrupting the class. Cameras capture live video, and the system analyzes facial expressions, eye movement, and posture to detect attention levels. Instead of manual observation, the system automatically identifies whether students are focused or distracted. The collected data is processed instantly, allowing teachers to get real-time feedback on classroom engagement. This helps in identifying students who may be struggling to concentrate. A dashboard provides a simple visualization of attention levels, making it easier for teachers to adjust their teaching methods. At the same time, privacy and data protection are carefully considered while designing the system. It can be implemented in different classroom sizes and supports smart classroom environments. Overall, this system helps create a more interactive and effective learning experience by bridging the gap between teaching and student engagement.

Keywords: *real-time attention monitoring, classroom engagement, and student behavior analysis, facial expression recognition, eye gaze tracking, and posture detection. real-time feedback, smart classroom systems.*

I. INTRODUCTION

A real-time attention monitoring system is an emerging solution designed to improve the quality of teaching and learning in classrooms. In traditional classroom settings, teachers often find it difficult to continuously monitor the attention level of every student. This becomes even more challenging in large classrooms where individual observation is nearly impossible. As a result, some students may lose focus without being noticed, which directly affects their learning outcomes. To address this issue, modern technologies such as computer vision and machine learning are being integrated into educational environments. These technologies make it possible to automatically analyze student behavior in real time. The system observes visual cues like facial expressions, eye movement, and head posture to determine whether a student is attentive or distracted. The main idea behind this system is to assist teachers rather than replace them. It acts as a supportive tool that provides additional insights into student engagement. By using cameras and intelligent algorithms, the system captures live classroom data and processes it instantly. This allows teachers to receive immediate feedback on how well students are paying attention during lectures. One of the key advantages of this system is its ability to provide objective and unbiased analysis. Unlike manual observation, which can be subjective, the system uses data-driven methods to evaluate attention levels. This helps in maintaining accuracy and consistency in monitoring student behavior. In addition, the system offers a dashboard interface where teachers can easily view attention statistics. These statistics may include individual attention scores, overall classroom engagement, and patterns over time. Such information helps teachers identify which topics are more engaging. Another important aspect of this system is real-time feedback. If the attention level in

the classroom drops significantly, the system can generate alerts. This allows teachers to quickly adjust their teaching methods, such as by asking questions, changing the pace, or introducing interactive activities. The system also supports data storage for future analysis. By maintaining historical records, it becomes possible to track student performance over a period of time. This helps in understanding long-term behavior patterns and improving teaching strategies accordingly. Moreover, the system can be integrated with smart classroom technologies. This makes it suitable for modern digital learning environments where automation and data analytics play a key role. It also enhances the overall classroom experience by promoting active participation among students. However, while implementing such systems, privacy and ethical considerations must be taken into account. Student data should be handled securely, and proper consent must be obtained before monitoring. Ensuring data protection is essential to maintain trust and transparency. The system is highly scalable and can be implemented in classrooms of different sizes. It can also be integrated with smart classroom technologies, such as digital boards and online learning platforms. This makes it suitable for modern education systems that rely on digital transformation. Despite these challenges, the benefits of a real-time attention monitoring system are significant. It helps improve student engagement, enhances teaching effectiveness, and supports data-driven decision-making in education. By providing continuous insights into classroom behavior, the system contributes to creating a more focused and productive learning environment. One of the main objectives of this system is to support teachers in their teaching process. It does not replace the role of a teacher but enhances their ability to understand student engagement. The system provides real-time feedback,

allowing teachers to adjust their teaching methods according to the current attention level of the class. This makes the learning process more dynamic and interactive.

In conclusion, a real-time attention monitoring system is a valuable innovation in the field of education. It helps bridge the gap between teaching and student engagement by providing real-time insights and data-driven feedback. By supporting teachers and encouraging student participation, it contributes to creating a more effective and interactive learning environment.

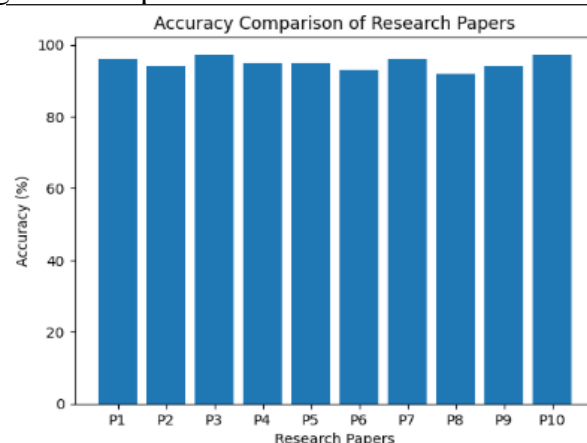
II. LITERATURE REVIEW

A study published by MDPI (2023) by Xiangyu Zhang and Yifan Liu focused on a deep learning-based approach for detecting student attention. They used models like YOLOv5 to classify attentive and inattentive behaviors based on facial expressions and body posture. Their research showed that such systems can achieve high accuracy and support teachers in understanding classroom engagement.

Another important contribution comes from Springer (2025) by Liang Chen and Min Zhao, who proposed a system combining face detection and head pose estimation. Their study introduced the concept of attention variation during lectures, highlighting that students' focus changes over time. This helped in understanding when students are most attentive and when engagement drops.

A significant contribution in this field is made by Xiaojing Sheng, along with Suqiang Li and Sixian Chan, who focused on improving the accuracy of classroom behavior detection. Their study highlights that the learning capacity of students is strongly influenced by the quality of instruction and their level of attention during lectures. With the advancement of behavior detection technologies, identifying student activities in classrooms has become more practical, but challenges still exist.

The researchers pointed out that one of the major problems is accurately detecting student behavior in complex and dynamic classroom environments. Factors such as lighting conditions, student movement, and varying seating positions make detection difficult. Another key issue is achieving real-time performance, as delays in detection reduce the usefulness of such systems in live classroom situations. The researchers pointed out that one of the major problems is accurately detecting student behavior in complex and dynamic classroom environments. Factors such as lighting conditions, student movement, and varying seating positions make detection difficult. Another key issue is achieving real-time performance, as delays in detection reduce the usefulness of such systems in live classroom situations. Overall, their work demonstrates that advanced deep learning techniques can significantly improve the performance of real-time classroom monitoring systems.



Attention plays a very important role in the learning process and directly influences how effectively students understand new concepts. Researchers like Michael I. Posner and Mary K. Rothbart (2014) emphasized that attention is a key cognitive function that supports better learning outcomes. Similarly, Richard Schwartzstein (2024) highlighted that active attention helps students engage more deeply with the subject matter.

Further, studies by Anne C. Hlas et al. (2017) show that students learn more effectively when they actively participate in classroom activities and remain focused on the lecture. When students are attentive, they are more likely to ask questions, make meaningful connections between different ideas, and develop a deeper understanding of the topic.

Another notable contribution in this field is by Liu Yan and Long Yanbin from Liaoning University of Science and Technology, China. Their work focuses on analyzing student attention and behavior within classroom environments using modern technological approaches.

Another important contribution is made by Ryan S. Baker, who has extensively worked in the field of educational data mining and student engagement analysis. His research focuses on understanding how students interact with learning environments and how their behavior reflects attention levels.

Another valuable contribution is made by Najmusher H, along with Abdul Ahad Siddique, A Shireesha, Bhavya, and Bharat Bhushan from HKBK College of Engineering, Bengaluru. Their research focuses on developing intelligent systems for monitoring student attention in classroom environments using modern computing techniques.

Another significant contribution in this area is made by Sergio Escalera, who has worked extensively on applying computer vision techniques to analyze human behavior in real-world environments,

including classrooms. His research focuses on detecting patterns such as facial expressions, gestures, and attention-related activities using advanced image processing methods.

The final important contribution is by Mohammad H. Mahoor, who has worked extensively in the field of affective computing and human behavior analysis. His research focuses on understanding human emotions and attention levels using facial expression recognition and deep learning techniques.

III. PROBLEM STATEMENT

In traditional classroom environments, it is often difficult for teachers to continuously monitor the attention level of every student. As the number of students increases, individual observation becomes limited and less effective. Many students may lose focus during lectures without being noticed, which negatively impacts their understanding and academic performance. Another major issue is the lack of objective methods to measure student engagement. Current approaches mainly rely on manual observation, which can be subjective and inconsistent. Teachers may not always be able to accurately identify which students are attentive and which are distracted. Additionally, classroom environments are dynamic, with factors like movement, noise, and varying student behavior making attention detection more challenging. There is also a lack of real-time feedback systems that can assist teachers in adjusting their teaching methods instantly. Therefore, there is a need for an automated, accurate, and real-time attention monitoring system that can analyze student behavior effectively. Such a system should provide reliable insights, improve student engagement, and support teachers in creating a more interactive and productive learning environment.

IV. METHODOLOGY

The methodology of the real-time attention monitoring system is designed to automatically detect and analyze student attention using advanced technologies like computer vision and machine learning. The system follows a structured pipeline, starting from data collection to final output generation. First, the system captures real-time video input using cameras installed in the classroom. These cameras continuously record student activities without interrupting the learning process. The captured video frames act as the primary input for further processing.

Next, the system performs face detection to identify and locate each student in the frame. This step is important because it isolates individual students for detailed analysis. After detecting faces, the system moves to feature extraction, where it analyzes facial expressions, eye gaze direction, and head posture. In the next stage, a machine learning or deep learning model (such as a YOLO-based algorithm) processes these features. The model is trained to classify whether a student is attentive or inattentive based on patterns observed in the data. Advanced modules like multi-scale feature extraction help improve accuracy even in complex classroom conditions. After classification, the system assigns an attention score to

each student. These scores represent the level of focus and engagement during the lecture. The data is then aggregated to calculate overall classroom attention. The processed information is displayed on a dashboard interface, where teachers can easily monitor attention levels in real time. If the attention level drops below a certain threshold, the system generates alerts to notify the teacher.

A. Image Input Module

The system's ability to accurately identify patterns related to affective states shows strong potential for future research and practical applications. By analyzing emotional and behavioral cues, the system can help in understanding how different factors influence students' mental and emotional conditions. This is especially useful in conducting large-scale experiments focused on stress assessment in educational environments.

The inclusion of stress detection and emotion recognition features makes the system more than just an attention monitoring tool. It allows educators to gain deeper insights into students' well-being, which is directly connected to their learning ability and academic performance. When students experience stress or negative emotions, their concentration and engagement levels often decrease.

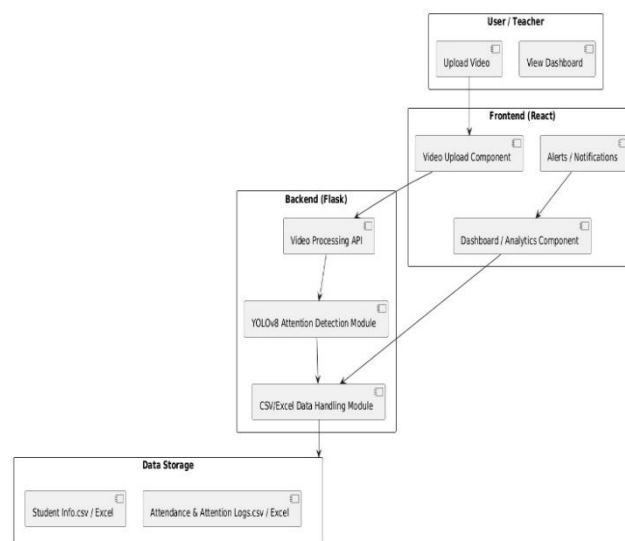


Fig.1: Smart classroom

B. Image Preprocessing Stage

The image processing stage is one of the most important parts of the system diagram, as it is responsible for analyzing the visual data captured from the classroom. In this stage, the system takes raw video input from the camera and converts it into meaningful information that can be used to detect student attention.

The process begins with frame extraction, where the continuous video stream is divided into individual frames. These frames act as images that can be processed separately. After that, the system performs pre-processing, which includes resizing the images, adjusting brightness, and removing noise to improve quality. This step ensures that the data is clean and suitable for further analysis.



C. Feature Extraction Using CNN

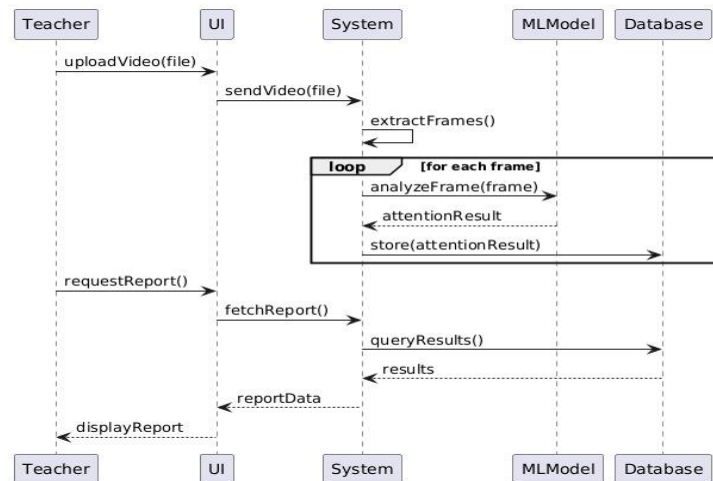
The image presents a structured visual representation of different student activities commonly observed in a classroom environment. It is divided into three main categories: hand-raising, reading, and writing, each showing multiple examples of students performing these actions.

The second section focuses on reading behavior, where students are seen concentrating on books or study materials. This activity represents cognitive engagement, where students are processing and understanding the content being taught. It highlights a focused and attentive learning state.



D. UML Sequence Diagram

The diagram represents the complete working flow of a real-time attention monitoring system, showing how data moves between the teacher, user interface, system, machine learning model, and database. It clearly explains how video input is processed step by step to generate meaningful attention reports.



The process begins when the teacher uploads a classroom video through the user interface. This video is then sent to the main system for processing. Once received, the system performs frame extraction, where the video is divided into multiple individual frames. These frames act as the basic units for analysis.

E. Database and Web Interface

The database is responsible for storing all the processed information generated by the system. After analyzing each video frame, the attention results are saved in the database in an organized manner. This includes details such as student attention scores, timestamps, and overall classroom engagement levels. Storing this data allows the system to maintain records for future analysis and comparison.

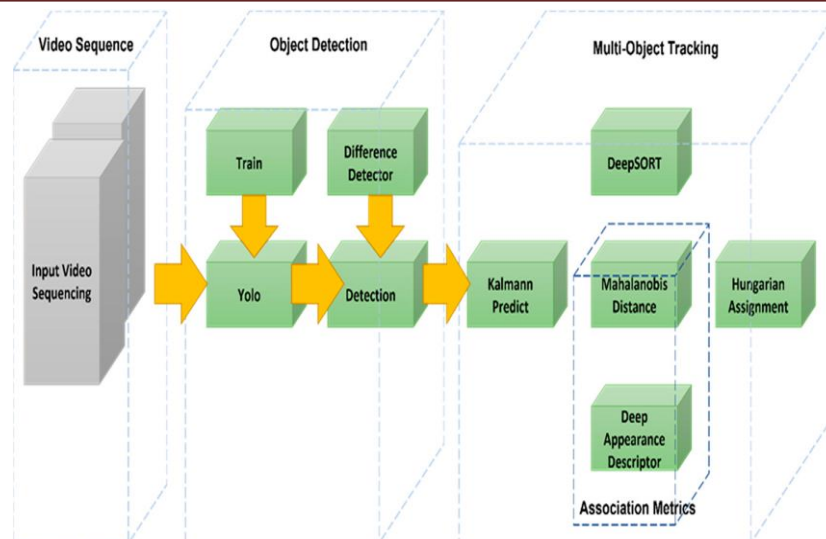
On the other hand, the web interface acts as the communication layer between the teacher and the system. It provides a user-friendly platform where teachers can upload videos, monitor real-time attention levels, and view reports. The interface is designed to be simple and interactive so that users can easily navigate through different features.

F. Teacher Verification and Decision Support

Teacher verification and decision support is an important part of the real-time attention monitoring system, as it ensures that the final output is reliable and useful for improving teaching methods. While the system automatically analyzes student attention, the teacher plays a key role in validating and interpreting the results.

The decision support aspect provides meaningful insights rather than just raw data. For example, the system may highlight students who are frequently inattentive or identify time periods where attention levels drop significantly. Based on this information, the teacher can make informed decisions to improve engagement.

G. Deep Sort Algorithm Architecture



Deep SORT is an advanced multi-object tracking algorithm used for real-time tracking in videos. It improves the SORT algorithm by adding deep learning-based appearance features, which increases tracking accuracy and reduces identity switching. The architecture consists of an object detector, Kalman filter, deep CNN-based feature extractor, and a data association module.

The object detector identifies objects in each frame, while the Kalman filter predicts their movement over time. The CNN extracts appearance features to recognize objects even after occlusion. Finally, the Hungarian algorithm is used to match detections with existing tracks based on motion and appearance similarity.

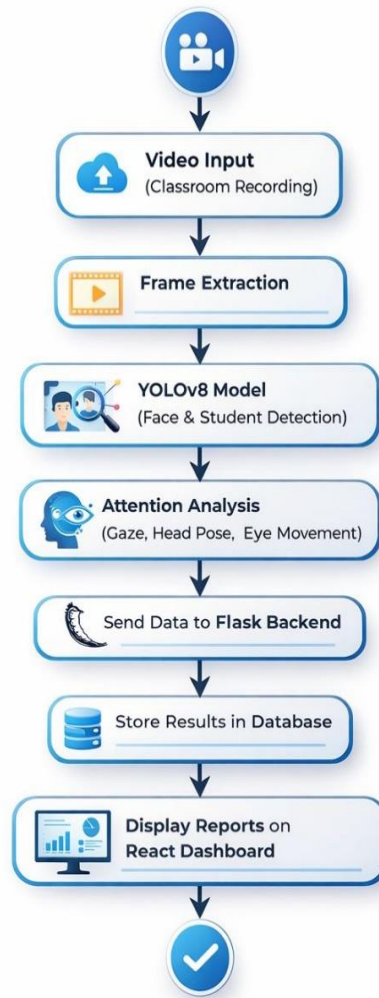
Overall, Deep SORT provides efficient and reliable real-time object tracking.

V. RESULTS

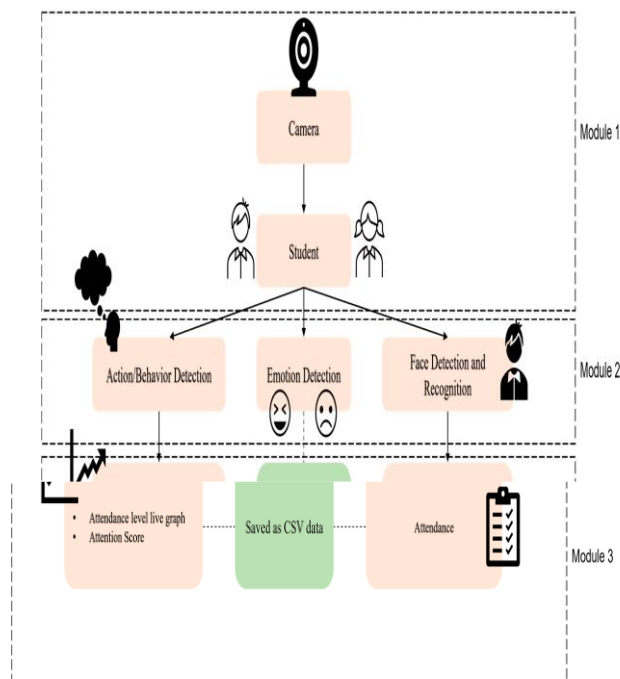
The results of the real-time attention monitoring system demonstrate its effectiveness in analyzing and improving student engagement in the classroom. After processing the video input, the system successfully identifies different student behaviors such as attentiveness, distraction, reading, writing, and participation.

Flowchart Working

It starts with a Start symbol, then moves through different process steps where actions are performed. At certain points, it includes decision blocks (Yes/No conditions) to choose the next path based on logic.



A. System Diagram



System architecture is the structured design of a system that defines how different components such as client, server, database, and network interact with each other. It provides a blueprint for organizing hardware, software, and communication processes to ensure smooth functioning of the system. It helps in building systems that are scalable, efficient, and easy to manage.

B. Evolution of Object Detection Algorithms

Object detection has improved a lot over time



In 2016, YOLO (You Only Look Once) was introduced by Redmon et al., which changed everything by making detection a single-step process instead of multiple stages. After that, YOLO kept improving with different versions like YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOv7, and YOLOv8. Each version improved speed, accuracy, and performance using better architecture and training techniques. The latest version, YOLOv8 (2023 by Ultralytics), is very advanced and flexible. It comes in different model sizes (nano to extra-large), so it can be used on both low-power devices and high-performance systems.

D. Overall Outcome

The overall outcomes of the real-time attention monitoring system show a significant improvement in understanding and enhancing student engagement in the classroom. The system successfully provides accurate and real-time analysis of student attention, helping teachers gain deeper insights into classroom behavior.

One of the key outcomes is the ability to continuously monitor attention without manual effort, reducing the workload on teachers. The system enables quick identification of inattentive students, allowing timely intervention and better classroom management.

VI. CONCLUSION

In conclusion, the real-time attention monitoring system provides an effective solution for analyzing and improving student engagement in classroom environments. By using technologies such as computer vision and machine learning, the system is able to automatically detect and evaluate student attention without interrupting the learning process.

In conclusion, the real-time attention monitoring system provides an effective solution for analyzing and improving student engagement in classroom environments. By using technologies such as computer vision and machine learning, the system is able to automatically detect and evaluate student attention without interrupting the learning process. Additionally, the system supports long-term data analysis, enabling teachers to understand patterns in student attention and improve their teaching strategies over time. It also reduces manual effort and provides a more objective way of monitoring classroom engagement. Overall, the real-time attention monitoring system plays a significant role in enhancing both teaching effectiveness and student learning outcomes. It represents a step towards smarter and more technology-driven education systems.

VII. Future Scope


The real-time attention monitoring system has strong potential for future improvements and wider applications in the field of education. One of the key areas of enhancement is increasing the accuracy of detection by using more advanced deep learning models and larger training datasets. This will help the system perform better in complex and dynamic classroom environments.

Another important scope is the integration with online learning platforms, making the system useful for virtual classrooms.

Additionally, the use of edge computing and IoT devices can enhance real-time

processing and reduce dependency on centralized systems. This will improve speed, efficiency, and data security and improve datasize.

VIII Limitations and Challenges

Even though modern object detection models like YOLO are very powerful , they still have some limitations and challenges. One major challenge is the need for large and high-quality datasets. Poor or unbalanced data can reduce

accuracy and affect model performance. Another limitation is difficulty in detecting small or overlapping objects, especially in complex backgrounds, where the model may miss or confuse objects. These models also require high computational power, particularly for large versions like YOLOv8, which may not run efficiently on low-end devices. In addition, performance can be affected in low-light or real-time noisy environments, where image quality is not clear.

IX. Classroom Experiment Result

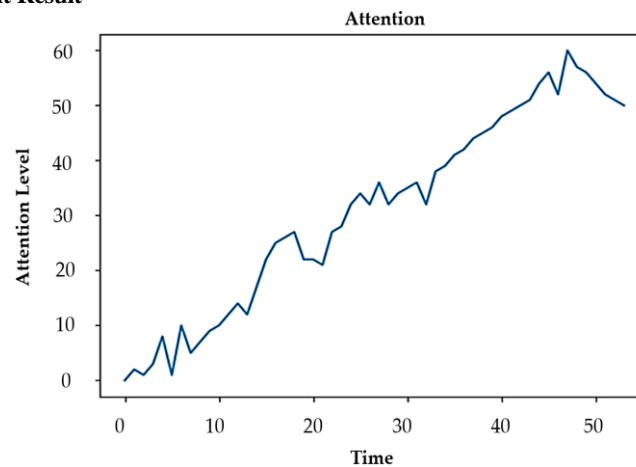


Fig 1.Attention level of student 2 is plotted on the graph

A	B	C	D	E
Student	Action/Behaviour detected	Attention State	Emotions	Time
Student 1	Focused	Attention	Happy	13:10:22
Student 2	Focused	Attention	Happy	13:10:22
Student 3	Bored	No attention	Sad	13:10:22
Student 4	Raising hand	Attention	Neutral	13:10:22
Student 5	Using Phone	No attention	Neutral	13:10:22
Student 6	Focused	Attention	Sad	13:10:22
Student 7	Using Phone	No attention	Neutral	13:10:22

Fig 2.student info saved in csv format.

A	B	C	D
Date	Time	Lecture/Code	Student
12.02.2022	13:08:02	Computer Networks(CN-0200D)	NCB002
12.02.2022	13:08:03	Computer Networks(CN-0200D)	NCB003
12.02.2022	13:08:03	Computer Networks(CN-0200D)	NCB004
12.02.2022	13:08:03	Computer Networks(CN-0200D)	NCB005
12.02.2022	13:08:04	Computer Networks(CN-0200D)	NCB006
12.02.2022	13:08:04	Computer Networks(CN-0200D)	NCB007
12.02.2022	13:08:05	Computer Networks(CN-0200D)	NCB008
12.02.2022	13:08:05	Computer Networks(CN-0200D)	NCB009

Fig 3. Student attendance saved in csv format.

X. Acknowledgment

We would like to express our sincere gratitude to our guide, D.B. Ghorpade, for providing us with continuous support, valuable guidance, and encouragement throughout the development of this

project. His insights and mentorship played a crucial role in helping us successfully complete this work.

This project has greatly contributed to enhancing our knowledge and

understanding, and it has been a valuable learning experience for all of us.

REFERENCES

1. Classroom Behavior Recognition Using Computer Vision: A Systematic Review
<https://doi.org/10.3390/s25020373>
MDPI
2. Real-Time Attention Monitoring System for Classroom: A Deep Learning Approach for Student's Behavior Recognition
<https://doi.org/10.3390/bdcc7010048>
MDPI
3. Real-time Monitoring of Students' Attention in Classroom using Transfer Learning
<https://doi.org/10.56201/ijcsmt.v10.no2.2024.pg170.178>
4. Predicting students' attention in the classroom from Kinect facial and body features Available at:
<https://jivpeurasipjournals.springeropen.com/articles/10.1186/s13640-017-0228-8>
5. Detecting Student Engagement in Classrooms for Intelligent Tutoring Systems October 2019 Available at:
<https://www.researchgate.net/publication/338942125>
6. Y. Li, G. R. Lan, Z. Li, X. Zhong and F. Han, "Real-time Gaze Tracking via Head-Eye Cues on Head-Mounted Displays", IEEE Trans. on Mobile Computing, 2024. <https://doi.org/10.1109/TMC.2024.3425928>
7. S. Maiti and A. Gupta, "LocalEyenet: Deep Attention Framework for Localization of Eyes", arXiv preprint, 2023. <https://arxiv.org/abs/2303.12728>
8. R. Daza, L. F. Gomez, A. Morales, J. Fierrez, R. Tolosana, R. Cobos and J. Ortega-Garcia, "MATT: Multimodal Attention Level Estimation for E-learning Platforms", arXiv preprint, 2023. <https://arxiv.org/abs/2301.09174>
9. J. Oh et al., "Improved Feature-Based Gaze Estimation Using Self-Attention Modules", Sensors, 2022. <https://www.mdpi.com/1424-8220/22/11/4026>
10. K. Shen et al., "Model-Based 3D Gaze Estimation Using a TOF Camera", PMC (NCBI), 2024. <https://www.ncbi.nlm.nih.gov/articles/PMC10891597>
11. E. Bozkir, "Automated Visual Attention Detection using Mobile Eye Trackers and Face Recognition", Educational Data Mining Conf., 2025. <https://educationaldatamining.org/EDM2025/proceedings/2025.EDM.long-papers.26/index.html>
12. "Classroom Attention Estimation Method Based on Mining Facial Landmarks of Students", ResearchGate, 2025. https://www.researchgate.net/publication/359223137_Classroom_Attention_Estimation_Method_Based_on_Mining_Facial_Landmarks_of_Students