

# STUDENT ENGAGEMENT DETECTION IN E-LEARNING

Vivek Rajeev V

Computer Science and Engineering  
Vimal Jyothi Engineering College  
Chemperi, Kerala, India  
vivekrajeev2002@gmail.com

Abhinav K

Computer Science and Engineering  
Vimal Jyothi Engineering College  
Chemperi, Kerala, India  
abhinavkizhakkepurayil@gmail.com

Swathi Krishna

Computer Science and Engineering  
Vimal Jyothi Engineering College  
Chemperi, Kerala, India  
swathikrishna2518@gmail.com

Anusree K

Computer Science and Engineering  
Vimal Jyothi Engineering College  
Chemperi, Kerala, India  
anusree8436@gmail.com

Clara Joseph

Computer Science and Engineering  
Vimal Jyothi Engineering College  
Chemperi, Kerala, India  
clarajoseph@gmail.com

**Abstract—** This paper introduces a novel e-learning engagement system using multimodal analysis of student behaviour. By monitoring eye closure, head pose, and emotional expression, the system infers student attention levels. Critically, it employs randomly timed alerts that require active student response via an "I am listening" button, mitigating passive alert habituation. Initial findings indicate the potential for improved student engagement and learning outcomes.

**Keywords—** Media pipe, Emotion Recognition, dlib, Facial landmark detection, Python, Fast API

## I. INTRODUCTION

The rapid evolution of technology has revolutionized education, with e-learning platforms becoming increasingly prevalent. While offering flexibility and accessibility, e-learning presents unique challenges in maintaining student engagement. Traditional classroom settings provide inherent social interaction and teacher presence, which naturally foster attention and participation. In contrast, online learning environments often lack these cues, leading to distractions, decreased motivation, and ultimately, compromised learning outcomes. Student engagement, defined as the learner's active involvement and investment in the learning process, is crucial for effective knowledge acquisition and retention. Engaged students are more likely to be attentive, motivated, and persistent, leading to deeper understanding and improved academic performance. Conversely, disengaged students may struggle to focus, easily become discouraged, and ultimately fail to reach their full potential. Therefore, fostering and maintaining student engagement in e-learning is a paramount concern for educators and researchers alike.

The need for effective engagement strategies in e-learning is underscored by the increasing prevalence of online courses and the growing recognition of the limitations of passive learning. Traditional methods of assessing engagement, such as surveys and self-reported feedback, often provide a delayed and subjective perspective. Furthermore, they may not capture the subtle, real-time fluctuations in a student's attention and involvement. To address these limitations, there is a growing interest in leveraging technology to develop more objective and dynamic measures of student engagement. This project focuses on developing such a system, utilizing multimodal analysis of learner behaviour to infer engagement levels in real-time.

In this system, these individual cues – eye closure, head pose, and emotional expression – are processed and combined to generate an Engagement Index (EI) score. Each cue contributes to the overall EI, providing a more holistic and nuanced measure of engagement than any single cue could offer alone. This EI score serves as a dynamic indicator of the student's current engagement level. A crucial element of this project is the implementation of a dynamic alert mechanism. Instead of relying on pre-defined thresholds or predictable patterns, the system delivers alert messages to the student at random intervals. These alerts are designed to interrupt potential lapses in attention and encourage active participation. To ensure the student is actively engaged with the alert, they are required to respond by clicking an "I am listening" button. This active response mechanism aims to minimize passive alert habituation, a common problem with static or predictable alert systems, and promote active learning. By requiring a response, the system ensures that the student is not simply passively acknowledging the alert, but actively refocusing their attention on the learning material. This project aims to demonstrate the effectiveness of this multimodal approach, combined with the dynamic alert mechanism, in improving student engagement and learning outcomes in e-learning environments.

## II. EYES CLOSURE DETECTION

### A. EAR

The Eye Aspect Ratio (EAR) is a measure of how open or closed the eye is. It is calculated using the distances between certain facial landmarks around the eye. Specifically, it's the ratio of the average of the distances between the upper and lower eyelids to the distance between the two corners of the eye. A higher EAR value indicates an open eye, while a lower EAR suggests a closed eye.

### B. EAR CALCULATION

$$EAR = \frac{|P2 - P6| + |P3 - P5|}{2 \times |P1 - P4|}$$

$$Average\ EAR = \frac{EAR_L + EAR_R}{2}$$

To detect the eye closure, we first need to detect the eye region from the given input image using Dlib's shape predictor. Six landmark points around the eye are identified,

and the Euclidean distances between these points are calculated. These distances are then used to compute the Eye Aspect Ratio (EAR) for each eye. The EAR is a measure of how open or closed the eye is, with a higher value indicating an open eye and a lower value indicating a closed eye. The EAR is calculated as the ratio of the distance between the upper and lower eyelids to the distance between the two corners of the eye.

## HEAD POSE DETECTION AND EMOTION DETECTION

### A. Head pose

To estimate the head pose, six specific facial landmarks are identified: the left and right edges of the eyes, the left and right corners of the mouth, the tip of the nose, and the chin. These 2D coordinates are then used to solve the Perspective-n-Point (PnP) problem. This mathematical procedure determines the 3D orientation and position of the head by using the corresponding 2D image points and a 3D model of the head. The PnP solution yields a rotation vector, representing the head's orientation in 3D space, and a translation vector, indicating its position. These vectors define the head pose, describing how the head is turned and where it is located in space. The rotation vector is converted into a rotation matrix using the Rodrigues' rotation formula, and Euler angles are extracted to provide a more intuitive representation of the head's orientation.

### C. Emotion Detection

The Fer (Facial Expression Recognition) module in Python is a deep learning-based tool designed for detecting emotions from facial expressions in images and videos. It leverages convolutional neural networks (CNNs) to classify emotions with high accuracy. Emotion detection is a crucial component in various applications, including e-learning, mental health monitoring, and human-computer interaction. The Fer module simplifies the implementation of emotion recognition by providing a pre-trained model that identifies emotions such as happiness, sadness, anger, surprise, fear, disgust, and neutrality. To use Fer, one must install it using pip install fer and then load an image or video stream containing human faces. The model processes the image, detects facial regions, and assigns probability scores to each emotion. A simple implementation involves using OpenCV to read an image and then applying Fer's detect\_emotions() function to analyze the facial expressions. Additionally, the top\_emotion() method extracts the most dominant emotion with its confidence score. Fer can also operate in real time with webcam feeds, allowing continuous emotion tracking by analyzing video frames sequentially.

Compared to traditional machine learning approaches that require extensive feature engineering, Fer's deep learning model provides a more robust and scalable solution with minimal preprocessing. Moreover, its compatibility with OpenCV and deep learning frameworks like TensorFlow or Keras enables seamless integration into various AI-driven applications. With growing interest in affective computing, emotion recognition technologies like Fer are gaining significance in multiple fields. Thus, the **Fer module** represents a powerful, accessible, and efficient solution for integrating emotion recognition capabilities into Python-based applications.

## III. IMPLEMENTATION

The implementation involves facial emotion recognition, gaze tracking, head pose estimation, and eye blinking analysis using deep learning techniques.

The approach involves several stages. First, facial emotion recognition is implemented using a Convolutional Neural Network (CNN) included in the python FER module. This model classifies emotions into seven categories: angry, fear, disgust, happy, sad, neutral, and surprise. The classified emotions are then mapped to academic engagement states by assigning predefined weights. Gaze tracking and head pose estimation are performed using Dlib's facial landmark detection algorithm and Perspective-n-Point (PnP) solutions, which estimate the direction of gaze based on head orientation. Additionally, eye blinking and status detection (open or closed) are determined using the Eye Aspect Ratio (EAR) method.

The engagement recognition system integrates these three modalities (facial emotions, gaze direction, and eye status) and calculates an Engagement Indicator (EI) score by aggregating the weighted results over 60-second intervals. Based on the EI score, students are categorized into engagement levels such as Highly Engaged, Confused, Bored, or Sleepy. According to the base paper [1] system was validated by comparing engagement scores with quiz performance after online classes. The results showed a positive correlation between engagement levels and academic performance. The engagement statistics were recorded and stored in a spreadsheet, providing teachers with actionable insights for monitoring student participation.

Apart from this a alert system is implemented to avoid students from cheating the system with recorded video. The alert system will display message on in the screen in random time intervals the message need to be acknowledged by the student within a given time limit if the student failed to do the same the EI score till the last acknowledged alert will be set to zero and the event will be reported to the teacher

FastAPI is used as the backend server to connect both the student and teacher applications. Student details, teacher details ,Authentication details, EI score are stored in the MySQL database connected to the server.

### A. Figures and Tables

#### a) EI score for head pose:

Look Center	Left head pose/Left gaze	Right Head pose/Right Gaze
1	0	0

#### b) EI score for eye status

Eye Close	Eye Open
0	2.5

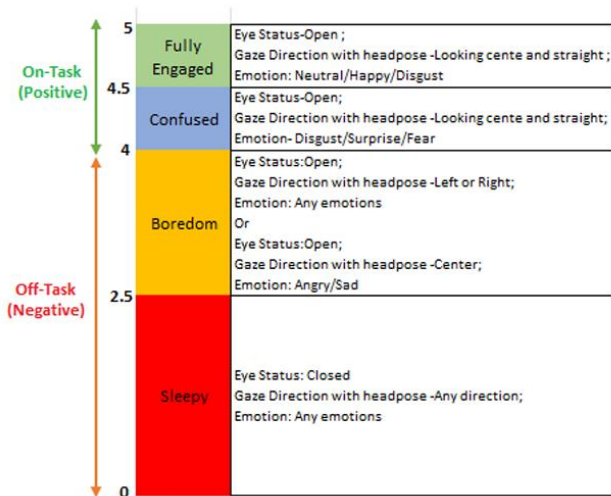
#### c) EI score for emotion

Emotions	Angry	Sad	Fear	Surprise	Disgust	Happy	Neutral
Weight of Emotions towards Engagement	0.1	0.3	0.5	0.7	0.9	1.1	1.4

#### d) Engagement classification based on EI value

Engagement status	EI
Highly Engaged	$EI \geq 4.5$
Confused	$4 \leq EI < 4.5$
Boredom	$2.5 \leq EI < 4$
Sleepy	$EI < 2.5$

e)



## REFERENCES

- [1] Sukumaran, Ajitha, and Arun Manoharan. "Multimodal engagement recognition from image traits using deep learning techniques." *IEEE Access* 12 (2024): 25228-25244.
- [2] Santhosh, Jayasankar, Akshay Palimar Pai, and Shoya Ishimaru. "Toward an interactive reading experience: Deep learning insights and visual narratives of engagement and emotion." *IEEE Access* 12 (2024): 6001-6016.
- [3] Ruiz, Nataniel, et al. "ATL-BP: a student engagement dataset and model for affect transfer learning for behavior prediction." *IEEE Transactions on Biometrics, Behavior, and Identity Science* 5.3 (2022): 411-424.
- [4] Santoni, Mayanda Mega, et al. "Automatic detection of students' engagement during online learning: a bagging ensemble deep learning approach." *IEEE Access* (2024). M. Young, *The Technical Writer's Handbook*. Mill Valley, CA: University Science, 1989.
- [5] Sowjanya, Uppanapalli Lakshmi, and R. Krithiga. "Decoding Student Emotions: An Advanced CNN Approach for Behavior Analysis Application Using Uniform Local Binary Pattern." *IEEE Access* (2024).
- [6] Alruwais, Nuha Mohammed, and Mohammed Zakariah. "Student Recognition and Activity Monitoring in E-Classes Using Deep Learning in Higher Education." *IEEE Access* 12 (2024): 66110-66128.