AI BASED STUDENT ENGAGEMENT DETECTION IN ONLINE LEARNING

A Project Report

submitted to

the APJ Abdul Kalam Technological University

in partial fulfillment of the requirements for the degree of

Bachelor of Technology

by

VIVEK RAJEEV V(VML21CS183)

ABHINAV K(VML21CS008)

SWATHI KRISHNA(VML21CS171)

ANUSREE K(VML21CS061)

under the supervision of

Ms. Clara Joseph

Assistant Professor

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

VIMAL JYOTHI ENGINEERING COLLEGE CHEMPERI

CHEMPERI P.O. - 670632, KANNUR, KERALA, INDIA

March 2025

VIMAL JYOTHI
ENGINEERING COLLEGE
JYOTHI NAGAR, CHEMPERI – 670632, KANNUR, KERALA
ACCREDITED BY IEI, NBA & NAAC ◆ ISO 9001:2015 CERTIFIED
AFFILIATED TO KTU ◆ APPROVED BY AICTE

NATIONAL BOARD OF ACCREDITATION ACCREDITED

NAAC ACCREDITED

DEPT. OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the report entitled **AI BASED STUDENT ENGAGEMENT DETECTION IN ONLINE LEARNING** submitted by **VIVEK RAJEEV V** (VML21CS183), **ABHINAV K** (VML21CS008), **SWATHI KRISHNA** (VML21CS171), **ANUSREE K** (VML21CS061) to the APJ Abdul Kalam Technological University in partial fulfillment of the B.Tech. degree in Computer Science and Engineering is a bonafide record of the project carried out by them under our guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

Ms. CLARA JOSEPH
(Project Guide)
Assistant Professor
Dept.of CSE
Vimal Jyothi Engineering College
Chemperi

Ms. NEETHU MATHEW
(Project Coordinator)
Assistant Professor
Dept.of CSE
Vimal Jyothi Engineering College
Chemperi

Ms. SREELAKSHMI M
(Project Coordinator)
Assistant Professor
Dept.of CSE
Vimal Jyothi Engineering College
Chemperi

Place: VJEC Chemperi

Date:27-03-2025

Head of the department

(Office Seal)

DECLARATION

We hereby declare that the project report **AI BASED STUDENT ENGAGEMENT DETECTION IN ONLINE LEARNING**, submitted for partial fulfillment of the requirements for the award of degree of Bachelor of Technology of the APJ Abdul Kalam Technological University, Kerala is a bona fide work done by us under supervision of **Ms. Clara Joseph**.

This submission represents our ideas in our own words and where ideas or words of others have been included, we have adequately and accurately cited and referenced the original sources.

We also declare that I have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. We understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Vivek Rajeev V

Abhinav K

Swathi Krishna

Anusree K

CHEMPERI

27-03-2025

ACKNOWLEDGEMENT

We are extremely grateful to **Dr. Benny Joseph**, Principal and **Ms. Divya B**, Head of CSE Department and our project coordinators for providing all the required resources for the successful completion of our project.

We are highly indebted to the guide **Ms. Clara Joseph** for her guidance and constant supervision, as well as for providing necessary information regarding the project and also for her support.

We express my thanks to our project coordinators **Ms. Neethu Mathew**, **Ms. Sreelakshmi M**, Assistant Professors of Computer Science and Engineering Department and all staff members, and friends for all the help and co-ordination extended in bringing out this project successfully in time. Our thanks and appreciation also goes to the authors of the references and other literatures referred to in this project.

Last but not the least, We would like to express our gratitude towards our parents and friends for their kind cooperation and encouragement.

Vivek Rajeev V

Abhinav K

Swathi Krishna

Anusree K

Abstract

The student app provides a platform for students to log in using their unique user ID and password, allowing them to access links to online classes uploaded by their teachers. To ensure the authenticity of the student's presence, the app incorporates a face recognition feature that operates every 30 seconds. Once the teacher sets the class start time, the app initiates a monitoring process using an AI model. This model makes predictions on the student's engagement every second. The predictions are then averaged over a 60-second window, and the engagement state for each minute is recorded. This continuous monitoring provides a detailed analysis of the student's attentiveness, ensuring active participation and focus throughout the class session. By leveraging advanced AI techniques, the app not only facilitates seamless access to online learning resources but also ensures that students remain engaged and present during their virtual lessons.

Contents

List of Figures

List of Tables

Nomenclature

AI	Artificial Intelligence
CNN	Convolutional Neural Network
FER	Facial Emotion Recognition
DAiSEE	Dataset for Affective States in E-learning Environments
EAR	Eye Aspect Ratio
uLBP	Uniform Local Binary Pattern
VGG-16	Visual Geometry Group
SGD	Stochastic Gradient Descent
FedAvg	Federated Averaging Algorithm
SVD	Singular Value Decomposition
EI	Engagement Index
Dlib	A toolkit for machine learning and data analysis

Chapter 1

Introduction

1.1 Overview

AI-based student engagement detection systems help teachers monitor how focused students are during online classes. These systems use artificial intelligence to track students' eye movements, gaze direction, and emotions to detect engagement in online learning, generating an Engagement Index (EI) score.The system watches where students are looking, using their webcam and predict their emotion as each emotion contribute a specific index value to the engagement index to see if they are paying attention to the lesson.

The system collects all this information and creates reports for teachers. These reports show how focused each student is and how the whole class is doing in terms of engagement. Teachers can use this data to spot distracted students, keep track of participation, and decide who may need extra help or feedback. This system helps make online learning more effective by providing clear information on whether students are paying attention or getting off-task.

1.2 General Background

AI-based student engagement detection in online learning has emerged as a transformative solution to address the growing need for effective monitoring and assessment of student participation in virtual classrooms. With the global shift toward online education, particularly during the COVID-19 pandemic, traditional methods of evaluating student attentiveness have become less effective. In physical classrooms, teachers can easily observe students' non-verbal cues, such as body language, facial expressions, and eye contact, to assess their level of engagement. However, in virtual learning environments, these visual cues are either diminished or entirely lost, making it difficult for instructors to gauge student participation accurately. This challenge has led to the development of AI-driven systems capable of automatically analyzing and tracking student engagement in real time, providing valuable insights that empower educators to make informed decisions.

AI-based engagement detection systems leverage machine learning models and computer vision technologies to interpret a variety of student behaviors that are indicative of focus and attention. These behaviors include eye gaze patterns, facial expressions, head movements, and eye status (open or closed). By processing data captured from webcams and screen recordings, these systems can estimate how attentive students are during lessons, identify instances of distraction, and generate detailed engagement reports. For instance, if a student frequently looks away from the screen or exhibits signs of fatigue, the system can flag these behaviors, allowing teachers to intervene or modify their teaching strategies to re-engage the student.

Machine learning models used in these systems are trained on large datasets containing labeled examples of engaged and disengaged behaviors. As a result, they can accurately predict the level of student attention and identify patterns that may indicate potential distractions. Additionally, computer vision techniques analyze visual inputs, such as micro-expressions and head movements, to assess a student's emotional state and cognitive load. Together, these technologies provide a comprehensive understanding of student engagement, offering real-time feedback to instructors and enabling dynamic adjustments in teaching methods.

A key advantage of AI-based engagement detection systems is their ability to provide real-time analysis and feedback. During live online sessions, the system continuously processes incoming data and alerts instructors when signs of disengagement are detected. This immediate feedback enables educators to take timely action, such as introducing interactive elements, asking questions, or adjusting the pace of the lesson to recapture student attention. Moreover, these systems generate detailed reports after each session, highlighting engagement trends, identifying patterns of attentiveness, and pinpointing areas where students may need additional support.

As technology continues to advance, AI-based student engagement detection systems offer promising solutions to bridge the gap between physical and virtual classrooms. By simulating the teacher's ability to observe and respond to student behavior, these systems create a more interactive and engaging online learning experience. Future advancements in AI and deep learning are expected to further enhance the accuracy and effectiveness of these systems, incorporating additional data points such as voice analysis and physiological signals to provide a more comprehensive understanding of student engagement.

In conclusion, AI-driven engagement detection systems play a pivotal role in enhancing the effectiveness of online education by providing real-time insights, detailed reports, and personalized learning experiences. As these technologies continue to evolve, they have the potential to transform online learning environments, ensuring that students remain engaged, focused, and successful in their educational journey.

1.3 Problem statement

With the rise of online learning, teachers face significant challenges in assessing student engagement in virtual classrooms. Unlike traditional in-person settings, where educators can observe students' facial expressions, body language, and eye contact, online learning environments make it difficult to gauge attentiveness. This limitation often leads to reduced participation, disengagement, and ineffective learning outcomes.

Existing methods to measure student engagement, such as post-class surveys, quizzes, or manual monitoring through video feeds, are inefficient and do not provide real-time insights. These approaches are often subjective, time-consuming, and fail to capture subtle variations in student focus throughout a lesson.

To address this issue, a system is needed to help teachers track student engagement more accurately during online classes. Such a system would provide real-time feedback, allowing educators to identify students who may be losing focus and take timely action to keep them engaged. By offering a clearer understanding of student attentiveness, this approach ensures a more interactive and effective learning experience.

1.4 Scope of the system

- As many educational institutions are providing online classes these days even in technical degree it is necessary to check whether the students are listening to these lectures or not.
- Attendance of the students can be monitored as it is done real classroom conditions.
- Time taken to take attendance during the online class can be saved by making the process automatic by the help of machine learning.

- It protects privacy of the students.In online classes the video of the students can be seen by teachers as well as other known and unknown students in the classroom it can be prevented as machine learning model works on the student's system and only the result is send to the server.
- It reduces the work load of teacher in verifying the engagement of the student.
- Provides a place for teacher to share the class link and time

1.5 Objective

The primary objective of the Student Engagement Detection System is to provide teachers with a real-time, data-driven approach to monitoring and improving student attentiveness in online classes. By continuously tracking Engagement Index (EI) scores, the system enables educators to assess both individual and collective student engagement levels, allowing them to make timely interventions and foster a more interactive learning environment.

This project aims to bridge the gap between traditional in-person monitoring and digital learning by offering instant feedback on student attentiveness. The automated engagement tracking system ensures that teachers are no longer reliant solely on visual cues or manual participation tracking, reducing subjectivity in evaluating student focus. Additionally, storing minute-wise average EI scores in a MySQL database allows for long-term analysis of engagement trends, enabling educators to identify patterns and adjust their teaching strategies accordingly.

Another key objective is to simplify classroom management through a feature-rich PyQt6-based teacher interface. The system provides tools for student registration, class link management, and engagement data visualization, ensuring teachers can easily access and analyze student performance. The modern dark-themed UI with bold, centered text and larger windows enhances usability, making the system intuitive and visually appealing for educators.

Furthermore, the project seeks to improve student participation and motivation by allowing teachers to reward engagement with attendance marks or participation points, ensuring students remain actively involved. It also supports early intervention strategies, helping teachers identify students who may need additional support due to consistently low engagement levels.

In summary, the Student Engagement Detection System enhances the effectiveness of online learning by providing real-time engagement insights, efficient classroom management tools, and an intuitive teacher interface. By integrating technology and data analytics, the system empowers educators to create a more engaging, interactive, and supportive virtual classroom experience.

Chapter 2

Literature Review

2.1 Multimodal Engagement Recognition From Image Traits Using Deep Learning Techniques

Abstract

This study introduces a multimodal engagement recognition system that uses deep learning techniques to assess student engagement in online learning environments. The system integrates multiple image-based traits, such as facial expressions, gaze direction, and eye status, to create a comprehensive understanding of student attentiveness. A convolutional neural network (CNN)-based facial emotion recognition (FER) model, trained on the FER 2013 dataset, identifies emotions such as happiness, sadness, anger, and neutrality, which are key indicators of engagement. Additionally, the system uses eye gaze tracking and head pose estimation to determine the student's focus and monitors eye blinking rates to detect signs of fatigue or distraction. These combined modalities are fed into an 'Engagement Indicator' algorithm, which calculates an overall engagement score in real-time, providing detailed engagement analytics to help educators make informed decisions and improve student focus.

The system captures real-time video data during online sessions, extracting facial

features using Haar Cascade classifiers and Dlib's 68-point facial landmark detectors. The CNN-based FER model processes facial expressions to identify emotional states, achieving an accuracy of 73.4%, ensuring reliable emotion detection. Gaze direction and head pose estimation are used to determine attentiveness, where significant head movements indicate distraction. Eye blinking patterns are monitored to detect fatigue, with high blinking rates suggesting drowsiness and low blinking rates indicating increased focus. These visual cues collectively contribute to refining engagement predictions.

Methodology

The Engagement Indicator Algorithm aggregates the results from FER, eye gaze tracking, head pose estimation, and blinking rates to compute a dynamic engagement score. This score is validated by correlating it with student performance over three consecutive days of online sessions. The system shows a strong correlation between engagement levels and academic performance, enabling educators to adapt teaching strategies and provide timely interventions.

While the system demonstrates high accuracy and effectiveness, it faces certain limitations. The small sample size used during model training limits its ability to generalize across diverse student populations. Additionally, overfitting in facial emotion detection may occur due to the constrained variety of training data. Future improvements should focus on training the model with larger, more diverse datasets to enhance generalizability and accuracy. Despite these limitations, the proposed system enhances the quality of online education by offering real-time engagement insights, ensuring that students remain attentive and participative.

Conclusion

This study presents a multimodal engagement recognition system that leverages deep learning techniques to assess student engagement in online learning environments. By integrating facial expression analysis, gaze tracking, head pose estimation, and blinking rate monitoring, the system provides a comprehensive understanding of student attentiveness. The Engagement Indicator Algorithm effectively aggregates these visual cues to generate a real-time engagement score, which has been shown to correlate strongly with academic performance.

Despite achieving a high level of accuracy in detecting engagement, the system faces limitations related to dataset size and generalizability. Future research should focus on expanding the dataset to include a more diverse student population and refining emotion recognition models to reduce overfitting. Nonetheless, this system represents a significant step toward improving online education by offering real-time engagement insights, allowing educators to adapt teaching strategies and enhance student learning outcomes.

2.2 Automatic Detection of Students' Engagement During Online Learning: A Bagging Ensemble Deep Learning Approach

Abstract

This research introduces a bagging ensemble deep learning approach to detect student engagement in online learning environments, addressing the limitations of individual models in handling data variability and complexity. The approach integrates 1D Convolutional Neural Networks (CNNs), 1D Residual Networks (ResNet), and hybrid ensemble models to improve the robustness and accuracy of engagement detection. Bagging ensemble techniques aggregate predictions from multiple models to reduce variance and minimize overfitting, resulting in more reliable classification of engagement levels.

The system is trained and evaluated using the DAiSEE (Dataset for Affective

States in E-Environments) dataset, which consists of 10-second video clips of students recorded during online sessions. OpenFace is used to extract 709 facial feature values from each frame, including facial landmarks, eye gaze estimations, and head pose data. Dimensionality reduction is performed using Singular Value Decomposition (SVD), enhancing computational efficiency by retaining the most relevant information. The bagging ensemble approach achieves an impressive accuracy of 94.25%, demonstrating its effectiveness in maintaining prediction stability and improving classification accuracy over individual models.

Real-time feedback provided by the system empowers educators to make data-driven adjustments in their teaching strategies, enhancing student participation and focus. By integrating multiple models and averaging their predictions through soft voting, the system improves engagement detection accuracy while contributing to a more adaptive and personalized online learning experience. This approach efficiently handles data imbalance issues, making it suitable for real-world applications where engagement levels vary across student populations.

Methodology

The methodology involves a multi-stage process that combines deep learning models, feature extraction, dimensionality reduction, and ensemble learning strategies to achieve high-accuracy engagement detection. The DAiSEE dataset, a publicly available benchmark designed to assess affective states and engagement levels in e-learning environments, is used to train and validate the model. Each 10-second video clip is annotated with labels indicating different engagement levels (high, low, and neutral), enabling the system to predict engagement effectively.

Feature extraction is performed using OpenFace, a powerful facial behavior analysis toolkit. It extracts 709 feature values from each frame, including facial landmarks that identify key points on the face, eye gaze estimations that indicate where the student is looking, and head pose estimations that provide information about head orientation and tilt. To enhance computational efficiency and reduce complexity, SVD is applied to project the data onto a lower-dimensional space while preserving the most relevant information for engagement classification.

The core of the system comprises a bagging ensemble approach that integrates 1D CNNs and 1D ResNet models. CNNs process sequential feature data from video frames, capturing spatial dependencies and identifying patterns in student behavior. ResNet models address the vanishing gradient problem, enabling the training of deeper models that capture complex engagement patterns. Bagging (Bootstrap Aggregating) is used to train multiple models on different subsets of the dataset and aggregate their predictions using a soft voting mechanism, where the final prediction is obtained by averaging the probability scores of all individual models. This reduces variance and enhances generalizability.

To evaluate system performance, standard classification metrics such as accuracy, precision, recall, and F1-score are used. The bagging ensemble approach achieves an impressive accuracy of 94.25%, demonstrating superior effectiveness in detecting student engagement across diverse contexts. This high accuracy validates the robustness of the system and highlights its ability to generalize well across different engagement scenarios.

Conclusion

The proposed bagging ensemble deep learning approach significantly enhances the detection of student engagement in online learning environments by outperforming standalone models. By leveraging ensemble learning, the system effectively reduces bias and variance, resulting in more stable, accurate, and reliable predictions. The combination of 1D CNN and ResNet models, along with bagging and soft voting techniques, ensures that the system can accurately detect engagement levels even in cases of noisy or imbalanced data.

The achieved accuracy of 94.25% highlights the robustness of the approach, validating its ability to provide accurate engagement predictions. Real-time feedback on student engagement empowers educators to dynamically adapt their teaching strategies, ensuring that students remain focused and attentive during online sessions. This approach contributes to creating a more personalized and effective online learning environment, where timely interventions can be implemented to improve student engagement and learning outcomes

2.3 Privacy-Preserving On-Screen Activity Tracking and Classification in E-Learning Using Federated Learning

Abstract

This study introduces a privacy-preserving architecture for tracking and classifying on-screen activities of students in e-learning environments using Federated Learning (FL). With the growing reliance on e-learning platforms, there is an increasing need to monitor student engagement and activity without compromising individual privacy. Traditional centralized models often require transmitting sensitive user data to a central server, raising concerns about data privacy and security. Federated Learning addresses these concerns by enabling local model training on individual devices without sharing raw data, ensuring privacy while contributing to an improved global model.

The proposed method leverages Federated Averaging (FedAvg), an algorithm that aggregates locally computed model updates from multiple clients (students) to iteratively enhance the global model. Each client processes its data locally, computes gradients using Stochastic Gradient Descent (SGD), and transmits only the model updates to a central server. This ensures that raw on-screen activity data never leaves the user's device, preserving privacy while improving accuracy in classifying various

activities such as reading, browsing, and participating in interactive tasks.

To achieve highly accurate classification, the system integrates InceptionV3, a deep convolutional neural network model known for its efficiency in image classification. The model achieves an impressive test accuracy of 99.75%, ensuring reliable detection of student activities. The proposed approach not only safeguards student privacy but also enhances engagement analysis, allowing educators to personalize learning experiences based on individual learning patterns.

Additionally, this framework enables the detection of various levels of student engagement, helping instructors optimize teaching methods based on real-time insights. By reducing reliance on centralized data storage, the system significantly lowers the risk of data breaches, making it a secure and scalable solution for modern e-learning environments.

Methodology

The methodology behind the proposed privacy-preserving on-screen activity tracking and classification system involves a multi-phase process that integrates federated learning with deep learning models to improve accuracy while maintaining data privacy. The system begins by initializing a global model with a pre-trained version of InceptionV3, a CNN architecture that captures high-dimensional feature representations effectively. This pre-trained model is distributed to multiple clients (student devices), where it is fine-tuned using local datasets, ensuring that the initial model has a solid foundation for subsequent learning.

The federated learning workflow consists of multiple communication rounds between the central server and individual client devices. Each communication round follows a structured sequence. First, during local training, each client (student device) receives the global model and fine-tunes it using its local dataset, which includes on-screen activity data such as reading, typing, or browsing. The local datasets are shuffled before training to prevent sequence bias. SGD is used to compute gradients and update model weights based on the local data. After local training, clients transmit only the updated model parameters (gradients) to the central server, ensuring that no raw data is shared, thus preserving privacy.

The next step involves Federated Averaging (FedAvg), where the central server aggregates the locally computed model updates. FedAvg combines model updates by weighting contributions based on the size of local datasets, ensuring that models trained on larger datasets have a proportional influence on the global model. Finally, the aggregated model updates are used to refine the global model, which is redistributed to all client devices for the next communication round. This iterative process continues until the global model converges and achieves high accuracy in activity classification.

To reinforce privacy and security, several techniques are integrated into the system. Secure aggregation ensures that model updates are encrypted before transmission, preventing potential interception. Differential privacy adds controlled noise to local model updates, making it difficult to infer information about individual clients. Additionally, local data confidentiality ensures that since raw data never leaves the client device, the risk of data leakage is minimized. These measures collectively enhance the system's privacy and security, ensuring that sensitive data remains protected throughout the federated learning process.

Conclusion

The proposed privacy-preserving architecture effectively detects and classifies on-screen student activities while addressing privacy concerns. Using federated learning, the system achieves 99.75% accuracy with the FedInceptionV3 model, enabling real-time, privacy-conscious activity tracking. This approach ensures that sensitive data remains on local devices, contributing to a robust global model. The system is scalable and adaptable, making it deployable across various e-learning platforms while providing educators with actionable insights to enhance student engagement.

Despite its high performance, the system faces challenges such as device heterogeneity and the need for larger, more diverse datasets to improve generalization. Future enhancements could include integrating advanced cryptographic protocols for better security and incorporating multimodal data, such as audio and physiological signals, to enrich activity classification. Beyond e-learning, this system has potential applications in areas such as telemedicine and remote collaboration, contributing to the development of secure and ethical digital environments. By maintaining a balance between privacy and performance, this system paves the way for future innovations in privacy-preserving AI-based learning environments

2.4 Decoding Student Emotions: An Advanced CNN Approach for Behavior Analysis Application Using Uniform Local Binary Pattern

Abstract

This paper introduces an advanced facial emotion detection model using Uniform Local Binary Pattern (uLBP) for student behavior analysis, addressing the limitations of traditional methods in evaluating student engagement. Conventional approaches rely heavily on subjective observations, which can be inconsistent, time-consuming, and prone to human bias. To overcome these challenges, this research integrates computer vision techniques with deep learning models to automate the identification and classification of student emotions in real-time.

The proposed model extracts facial features using uLBP, a robust texture descriptor that enhances feature representation, making emotion recognition more accurate and efficient. By incorporating Convolutional Neural Networks (CNNs), the system classifies emotions such as happiness, sadness, anger, and neutrality, which serve as key indicators of student engagement. Real-time emotion tracking allows educators

to monitor student attentiveness, detect signs of disengagement, and implement timely interventions.

This approach enhances teaching effectiveness by enabling personalized learning experiences, allowing instructors to adjust teaching strategies based on student emotions. By providing deeper insights into student engagement during both online and offline classes, the system contributes to the advancement of educational technology, offering an innovative and scalable solution for improving student learning outcomes while minimizing manual effort in monitoring engagement.

Methodology

The methodology involves developing a Convolutional Neural Network (CNN) model specifically designed for real-time emotion classification. The model leverages uLBP for feature extraction, a texture descriptor that effectively captures the local patterns of facial expressions. uLBP enhances the model's ability to identify subtle changes in facial emotions, making it suitable for analyzing student behavior. To ensure robustness and generalization, data augmentation techniques such as rotation, flipping, and zooming were applied to the training dataset. These techniques prevent overfitting and improve the model's ability to adapt to different lighting conditions, facial angles, and environmental variations. The CNN architecture was fine-tuned using performance metrics, including accuracy, precision, recall, and F1-score, to validate the model's effectiveness. Multiple iterations were performed to optimize the model's parameters, ensuring high accuracy in emotion classification. Extensive testing was conducted using benchmark emotion datasets such as FER-2013 and CK+, which contain diverse facial expressions representing various emotions. The model's performance was compared with traditional emotion recognition techniques, demonstrating its superior accuracy, efficiency, and suitability for real-world educational applications. The results highlighted the model's ability to provide real-time, accurate emotion classification, making it an ideal choice for monitoring student engagement.

Conclusion

The proposed CNN-based emotion detection model, enhanced with uLBP for feature extraction, demonstrates significant potential in improving the analysis of student engagement by detecting emotional responses in real-time. The findings suggest that educators can dynamically adapt teaching methods by using insights from the model, ensuring that instructional strategies align with student emotions and engagement levels. This approach allows educators to identify disengagement early, enabling timely interventions and fostering a more inclusive and supportive learning environment. By integrating this technology into educational platforms, teachers can personalize learning experiences, address student concerns proactively, and create an environment that promotes continuous student engagement..

2.5 Student Recognition and Activity Monitoring in E-Classes Using Deep Learning in Higher Education

Abstract

This study explores the application of a deep learning model for real-time student identification and activity monitoring in online education, addressing the increasing need for effective engagement tracking in virtual learning environments. As online education becomes more prevalent, ensuring that students remain attentive and engaged is a growing challenge. The proposed system utilizes facial recognition and head posture estimation to analyze student behavior during online classes. By detecting key indicators of attentiveness, such as facial expressions, gaze direction, and head movements, the system determines whether a student is focused or distracted.

Real-time monitoring enables educators to identify disengaged students and take proactive measures to enhance engagement. The system provides instant feedback, allowing instructors to initiate personalized interactions, send reminders, or adjust

teaching strategies to re-engage students effectively. By using computer vision and deep learning, the model offers a non-intrusive, automated approach to tracking student participation.

This approach improves learning outcomes by fostering a dynamic, responsive, and supportive online learning environment. Educators can make data-driven decisions to tailor instruction based on student engagement levels, ultimately leading to higher retention rates, better comprehension, and improved student satisfaction. By integrating advanced AI-driven engagement tracking, this system significantly enhances the effectiveness of virtual education.

Methodology

The model is built using a convolutional neural network (CNN) that processes 11,342 grayscale images from the UPNA Head Pose Database, which contains various head orientations. This data allows the model to accurately predict head poses and assess student attentiveness. To improve performance and ensure robustness, the model applies various optimization techniques such as data augmentation, transfer learning, and fine-tuning. Data augmentation techniques, including rotation, flipping, zooming, and brightness adjustments, are used to artificially expand the dataset and reduce overfitting. Transfer learning leverages pre-trained CNN models, such as VGG16 or ResNet, which are fine-tuned on the target dataset to adapt to the task of head pose and facial expression recognition. Hyperparameter tuning further optimizes the learning rate, batch size, and number of epochs to enhance model efficiency.

Real-time monitoring is achieved by integrating facial expression recognition and head posture analysis to identify disengaged or inattentive students during online sessions. The system categorizes student behavior into different levels of engagement, such as fully attentive, slightly distracted, or disengaged, providing instructors with actionable insights. This enables educators to dynamically modify their teaching strategies and address student inactivity effectively.

Conclusion

The proposed system achieves an impressive accuracy of 99%, demonstrating its reliability in detecting student activity and identifying disengagement. Beyond enhancing academic performance, the system fosters an interactive and supportive learning environment by addressing student disengagement in real time. The study also emphasizes the importance of maintaining ethical standards and protecting student privacy when implementing facial recognition technologies in educational settings. Secure data storage, anonymization, and obtaining informed consent are essential to safeguard student information and uphold privacy norms. Future research could focus on expanding dataset diversity, incorporating additional behavioral cues, and enhancing real-time monitoring capabilities to further refine and personalize student engagement tracking systems.

2.6 Consolidated Table

Table 2.1: Consolidated Table

PAPER	PROBLEM STATEMENT	TECHNOLOGY USED	ADVANTAGES	DISADVANTAGES
Multimodal Engagement Recognition From Image Traits Using Deep Learning Techniques	Presents a system for automatically recognizing student engagement in online learning environments using deep learning model.	Haar Cascade Algorithm, Dlib's Shape Predictor, Eye Aspect Ratio (EAR), Engagement Indicator (EI)	Real-Time Processing Comprehensive Engagement Analysis	Overfitting Issues Inaccuracy with Occlusion.
Automatic Detection of Students' Engagement During Online Learning	Automatically detecting student engagement during online learning	1D CNN,1D ResNet,Singular Value Decomposition (SVD)	High Accuracy, Comprehensive Feature Extraction.	Facial Feature Dependence: Potential Ethical Concerns.
Privacy-Preserving On-Screen Activity Tracking and Classification in E-Learning Using Federated Learning	Classify student activities without sending data to a central server, ensuring privacy.	Federated Learning,Transfer Learning	High Accuracy Privacy-Preserving	Complex Implementation Limited to Certain Devices.
Student Recognition and Activity Monitoring in E-Classes Using Deep Learning in Higher Education	Track and monitor student behavior, mood, and engagement during online classes	CNNs, OpenCV	Real-Time Monitoring Improved Student Engagement	Privacy Concerns Computational Requirements.
Decoding Student Emotions: An Advanced CNN Approach for Behavior Analysis	To use modern techniques for feature extraction to automatically recognize student emotions in real-time	CNNs,uLBP	Real-time Feedback,More Accurate ,Better Engagement	Privacy Issues,Expensive,Limited Understanding

Chapter 3

Requirement Specification

3.1 Functional Requirements

Student Application

- Send Engagement Scores: The system must allow students to send their engagement index (EI) scores every 5 seconds.
- Submit Average Engagement Score: The application should calculate and send the average EI score every minute.
- Login and Authentication: Students must be able to log in securely using their credentials.
- Real-time Data Transmission: The app must continuously send engagement scores to the server for processing.

Teacher Application

- Real-time Engagement Monitoring: Teachers should be able to view live engagement scores of students.
- Historical Data Analysis: The system must store engagement scores and allow teachers to analyze trends over time.

- Classroom Management: Teachers should be able to register students, update class details, and manage sessions.

Server-Side Functionality

- Data Processing and Forwarding: The server must receive student engagement scores and forward them to the teacher's application in real time.
- Database Management: Engagement data should be stored in a structured format for analysis.
- Authentication and Authorization: The system should verify user credentials before allowing access to student or teacher dashboards.

Database Requirements

- Student Records: Maintain student details, including name, roll number, and engagement history.
- Class Data Storage: Store engagement scores with timestamps for each student.
- Teacher Records: Store teacher credentials and class details

3.2 Non Functional Requirements

- **Performance:** The system should process and transmit student engagement data in real-time with minimal latency.
- **Scalability:** The system should support multiple students and teachers without performance degradation.
- **Reliability:** The system must ensure consistent data transmission and storage, minimizing data loss.

- **Security:** User authentication and authorization must be implemented to protect student and teacher data.
- **Usability:** The user interface should be intuitive, easy to navigate, and maintain a dark-themed modern design.
- **Maintainability:** The system should be designed with modular components to allow for future updates and improvements.
- **Availability:** The system should be accessible with at least 99.9% uptime, ensuring minimal disruptions during online classes.
- **Data Integrity:** Engagement scores should be stored accurately without duplication or corruption.
- **Compatibility:** The system should work on multiple devices and operating systems, including Windows, macOS, and Linux.
- **Extensibility:** The architecture should allow for additional features, such as AI-based engagement detection, without major redesign.

Chapter 4

Proposed system and Design

4.1 Proposed system

The proposed system consist of two application one for the student and other for the teacher.

Student Application

User Authentication

- Students must log in using their registered credentials.
- Secure authentication mechanisms ensure data privacy.
- Session management prevents unauthorized access.

Engagement Score Submission

- The app collects the student's Engagement Index (EI) score every 5 seconds.
- The engagement score is computed based on predefined parameters (e.g., facial recognition, focus time, or manual input).

- These scores are automatically transmitted to the server for real-time processing.

Minute-Wise Engagement Average

- Every minute, the application calculates the average engagement score.
- The average score is sent to the server for long-term storage and trend analysis.
- This feature helps in tracking engagement patterns over time.

User Interface

- The app provides a dark-themed, modern, and user-friendly interface.
- It displays the current engagement score in an easy-to-read format.
- Students receive visual indicators (e.g., color-coded engagement levels) to help them track their own focus.

Teacher Application

User Authentication

- Teachers must log in using their registered credentials.
- Secure authentication mechanisms ensure data privacy.
- Role-based access control prevents unauthorized usage.

Real-Time Engagement Monitoring

- The application receives engagement index (EI) scores from students every 60 seconds.
- The EI score with the time wil be stored in the server too.

Historical Data Analysis

- The application stores engagement data for long-term trend analysis.
- Teachers can view minute-wise average engagement scores of each students.

Classroom Management

- Teachers can register new students and manage existing student profiles.
- The system allows teachers to create and manage class sessions.
- Teachers can assign class links and resources to students through the application.

User Interface

- The application follows a dark-themed, modern, and intuitive design.
- Engagement data is displayed in a structured format using tables and graphs.
- The interface is optimized for ease of use, ensuring minimal distractions.

4.1.1 Data Transmission and Storage

- The application retrieves engagement data from the server in real-time.
- Data is securely stored in a structured format for future reference.
- Encryption ensures that student engagement records remain private and secure.

4.1.2 Device Compatibility

- The teacher application is designed to run on multiple platforms, including Windows, macOS, and Linux.
- It ensures seamless interaction with the student application.

4.1.3 Benefits

- Enables teachers to track student engagement effectively in an online setting.
- Helps in identifying disengaged students early and taking corrective actions.
- Improves the overall effectiveness of virtual learning by providing real-time insights.

4.1.4 Data Transmission and Storage

- The student app securely transmits engagement data to the server using an optimized data transfer mechanism.
- It ensures minimal bandwidth usage to avoid network congestion.
- Data is encrypted to prevent unauthorized interception.

4.1.5 Device Compatibility

- The application is designed to run on multiple platforms, including Windows, macOS, and Linux.
- It supports integration with various input methods, such as webcams for facial analysis or manual input options.

4.1.6 Benefits

- Helps students stay engaged by providing real-time feedback on their focus levels.
- Allows teachers to identify disengaged students and provide timely interventions.
- Enhances the overall effectiveness of online learning environments.

Faculty Application

- Link to the online class,date,time and details of the class can be updated by the teacher in the application.
- The faculty can view the report after when the online class is over.
- Live report of the students presence in front of the camera will be reported to the faculty.

4.2 Feasibility Study

4.2.1 Technical Feasibility

- The student engagement detection system is technically feasible as it leverages existing and well-supported technologies to ensure smooth operation and scalability. The system will be developed using **FastAPI** for backend processing, **PyQt6** for the teacher's user interface, and **MySQL** for data storage. These technologies are widely used, open-source, and provide robust support for real-time data processing and user interaction.
- The system will collect Engagement Index (EI) scores from student applications every 5 seconds and transmit them to the teacher's dashboard. Real-time data

transmission will be handled efficiently using WebSockets, ensuring low-latency communication between students and teachers. The backend will process and store minute-wise average engagement scores in a structured MySQL database, allowing for easy retrieval and analysis.

- Scalability is a key consideration, and the system is designed to handle multiple students and teachers concurrently. The backend will support asynchronous processing to manage high data traffic efficiently. Cloud deployment options, such as AWS or Azure, can be integrated for further scalability and reliability.

- The system is compatible with various devices and operating systems, ensuring accessibility for users on Windows, macOS, and Linux. A secure authentication mechanism will be implemented to protect user data, with encryption techniques ensuring privacy and compliance with security standards.

- Given the availability of open-source libraries, existing frameworks, and scalable architectures, the system is technically feasible and can be implemented efficiently with minimal risk.

4.2.2 Operational Feasibility

- The student engagement detection system is operationally feasible as it seamlessly integrates into existing online learning environments with minimal disruption. The teacher application, built using **PyQt6**, provides an intuitive dashboard displaying real-time engagement scores, allowing teachers to monitor student focus effortlessly.

- For students, the application runs in the background, automatically transmitting Engagement Index (EI) scores every 5 seconds. This ensures real-time updates without requiring manual input. The system is independent of specific learning platforms, making it adaptable to various educational settings.

- Overall, the system enhances online learning by enabling teachers to track student engagement effortlessly, ensuring timely support and improved participation.

4.3 Design

4.3.1 Architecture Diagram

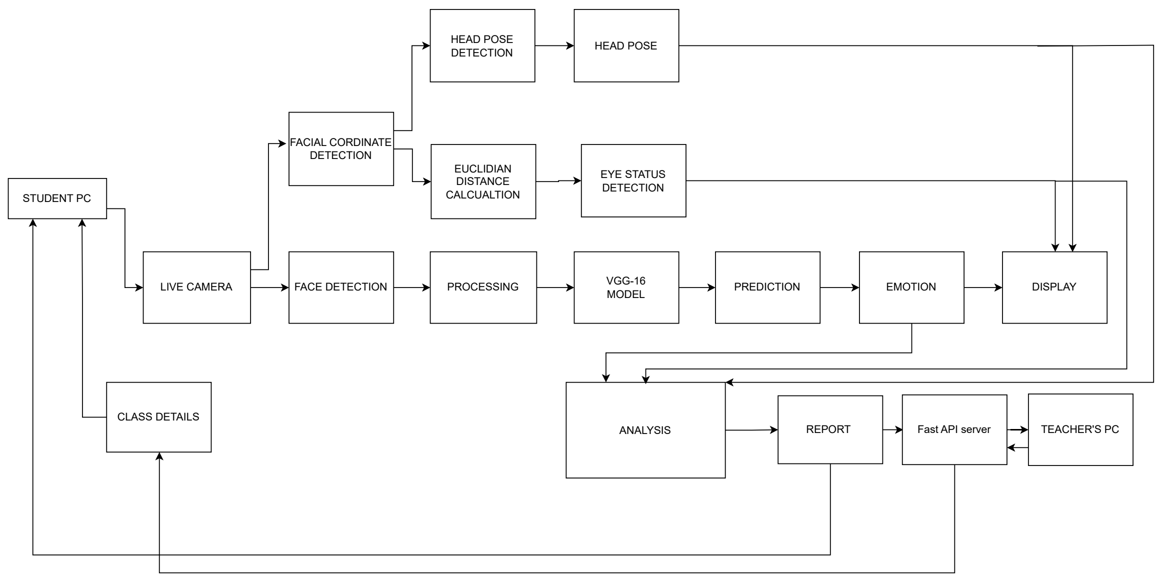

Figure 4.1: Architecture Diagram

The architecture diagram depicts engagement monitoring system designed for online education. The system begins with a Student PC that captures a live video feed via a connected camera. This feed undergoes Face Detection.

Facial Feature Analysis for Engagement

The detected facial features undergo multiple processes:

- Head Pose Detection: Determines the student's head orientation to assess attentiveness.
- Euclidean Distance Calculation: Measures facial displacement to detect changes in position or movement.
- Eye Status Detection: Monitors eye movements and blinks to evaluate focus and engagement levels.

VGG-16 Model for Prediction and Emotion Analysis

The facial data is passed through a VGG-16 Model, a pre-trained convolutional neural network (CNN), to perform Prediction and Emotion Analysis. This model identifies emotional states such as engaged, bored, sleepy, or confused by analyzing facial expressions.The Prediction module classifies student engagement based on analyzed data.The results are displayed in real-time, providing immediate feedback to both students and teachers.

EI Score Calculation

EI score calculation of the student is done according to the instruction provided in the base paper *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

Figure 4.2: EI scores

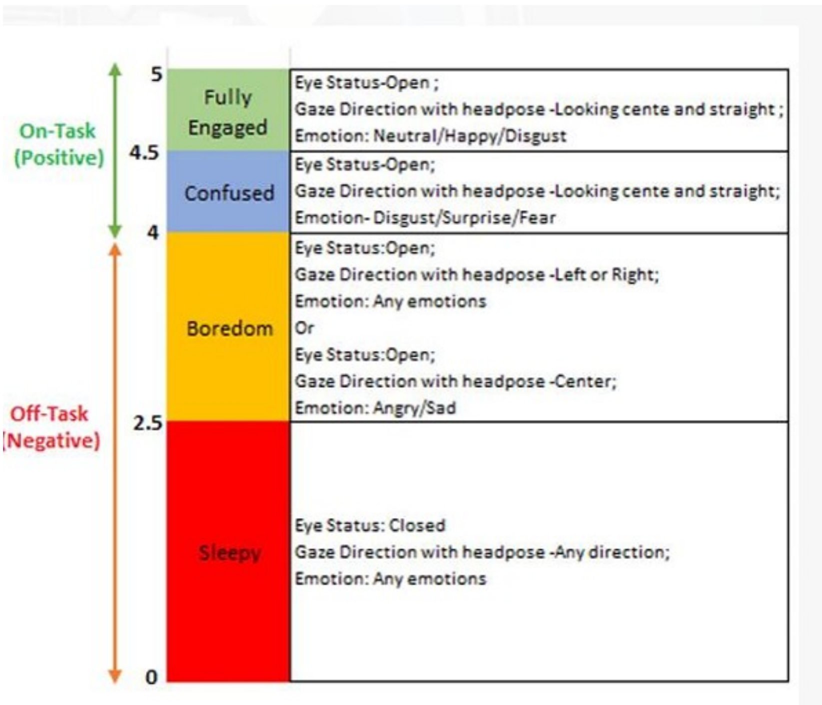

Figure 4.3: Engagement levels

Data Processing and Analysis

Once processed, the engagement data is transmitted to the Analysis module, which interprets the collected information and generates EI score on individual student activity and class-level trends. These reports highlight metrics such as presence time and engagement levels over the duration of the class.

Real-Time Communication with Teacher's PC

Processed data, along with generated reports and alerts, is forwarded to the Teacher's PC via the Django Server, ensuring real-time updates and enabling continuous monitoring. Teachers can use this data to intervene when necessary, enhancing student engagement and learning outcomes.

4.3.2 Use Case Diagram

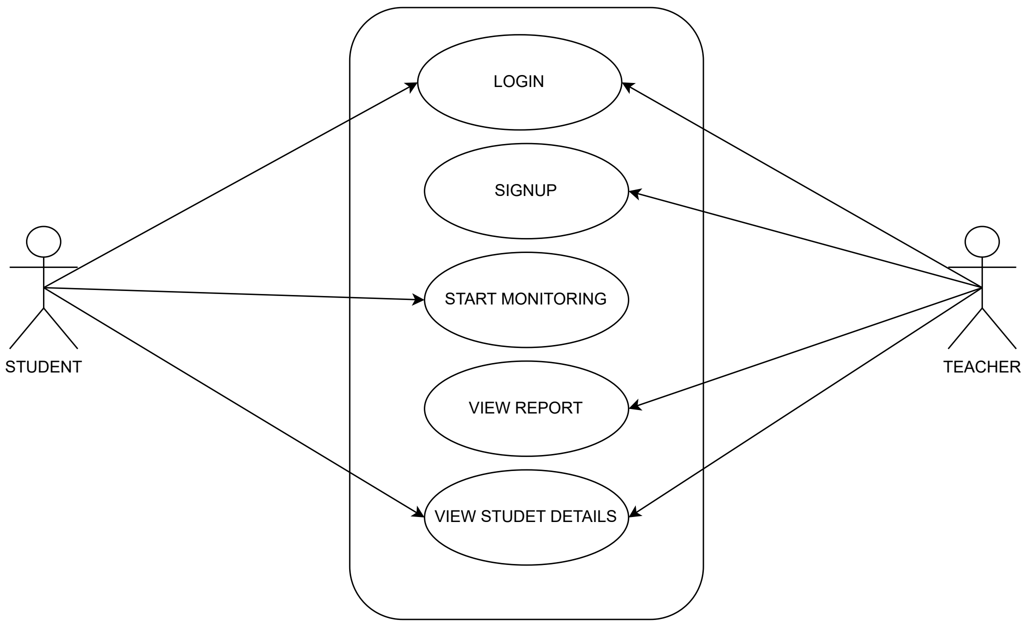

Figure 4.4: Use Case Diagram

The use case diagram depicts the interactions between two actors, Student and Teacher, with a system.

- Login: Both the Student and Teacher can log in to the system.
- Sign Up: Both actors have the capability to sign up.
- Start Monitoring: This use case is accessible by both Student and Teacher.
- View Report: Both the Student and Teacher can view reports.
- View Student Details: The Teacher can view student details.

4.3.3 Data Flow Diagram

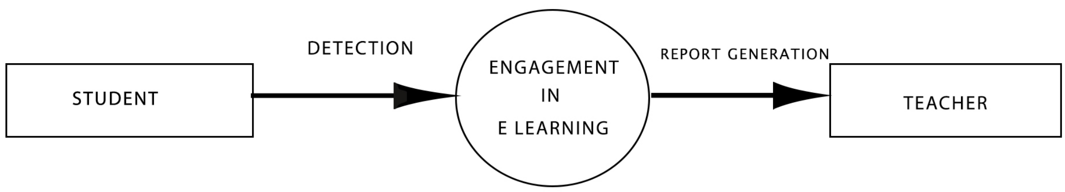

Figure 4.5: Level 0

The Level 0 Data Flow Diagram represents the system's high-level process flow between two external entities: Student and Teacher.

- Student: Provides input (Detection) to the system.
- Engagement in E-Learning: This is the central process that handles the data received from the Student. It processes this data to generate reports or results.
- Teacher: Receives output (Report Generation) from the system.

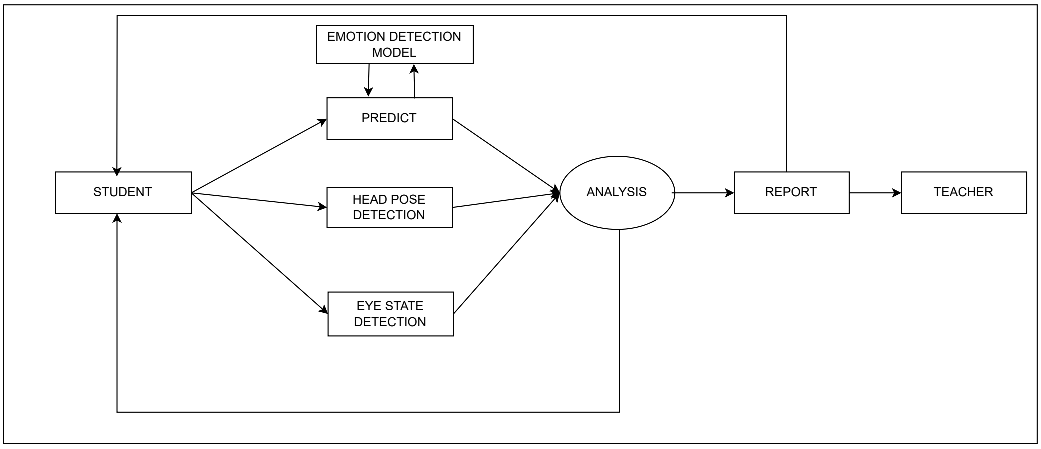

Figure 4.6: Level 1

The diagram illustrates a Level 1 Data Flow Diagram for a student engagement monitoring system, designed to assess engagement through multiple detection processes.

- Student: The system observes the student using a camera, capturing visual data.
- Face Recognition: This module identifies the student's face within the camera feed to focus on relevant features.
- Emotion Detection Model (Predict): Using the detected face, an emotion detection model predicts the student's emotional state, such as happy, sad, or bored.

- Head Pose Detection: This component evaluates the student's head orientation, which can indicate attention levels based on where they are looking.
- Eye State Detection: This module detects whether the student's eyes are open or closed, helping identify signs of drowsiness or inattentiveness.
- Analysis: All outputs from the above detection modules are fed into an analysis module to interpret the overall engagement level.
- Report: The analysis results are compiled into a report.
- Teacher: The report is sent to the teacher, providing insights into the student's engagement status.

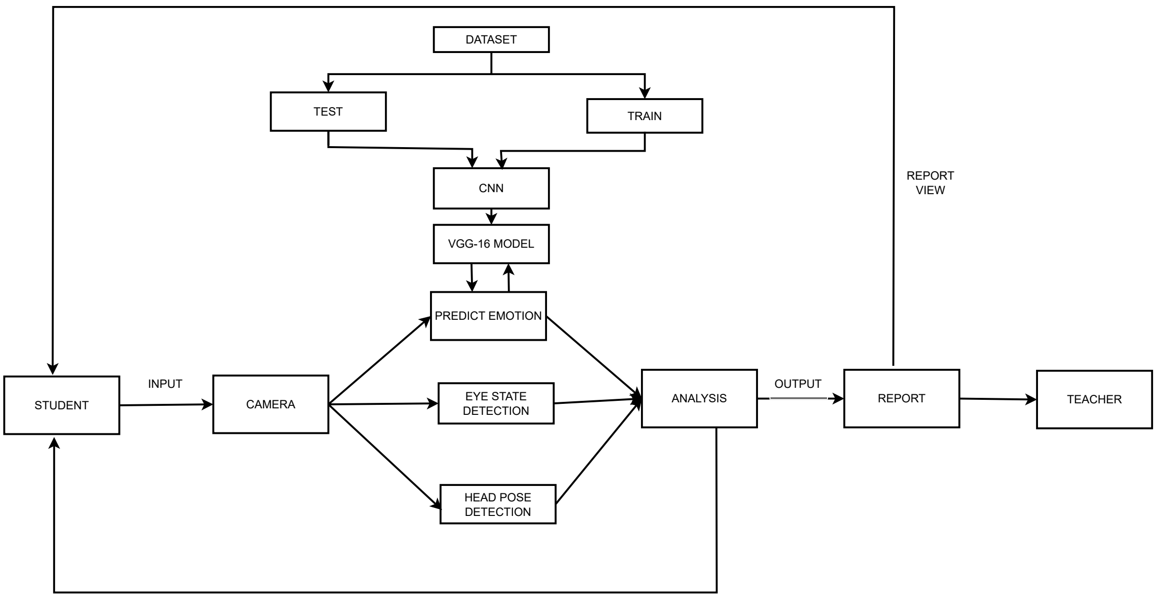

Figure 4.7: Level 2

This diagram represents the Level 2 Data Flow Diagram for a student engagement recognition system.

- Dataset: The system uses a dataset, which is split into Test and Train subsets, for training the model.

- CNN (Convolutional Neural Network): The core deep learning model, which is trained on the dataset to recognize emotions and facial features.

- VGG-16 Model: A specific CNN architecture, VGG-16, is used for its effectiveness in image recognition tasks, providing features for accurate engagement detection.

- Prediction: This module leverages the trained model to predict the student's emotional state and engagement.

- Camera: Captures real-time input from the student, including facial expressions and head movement.

- Face Recognition: Identifies the student's face within the camera feed.

- Eye State Detection and Head Pose Detection: Modules that analyze eye openness (indicating drowsiness or alertness) and head orientation (indicating attention).

- Analysis: Aggregates information from facial recognition, eye state, and head pose detection to assess engagement.

- Teacher: The final report is sent to the teacher, providing valuable insights on the student's engagement level.

4.3.4 ER Diagram

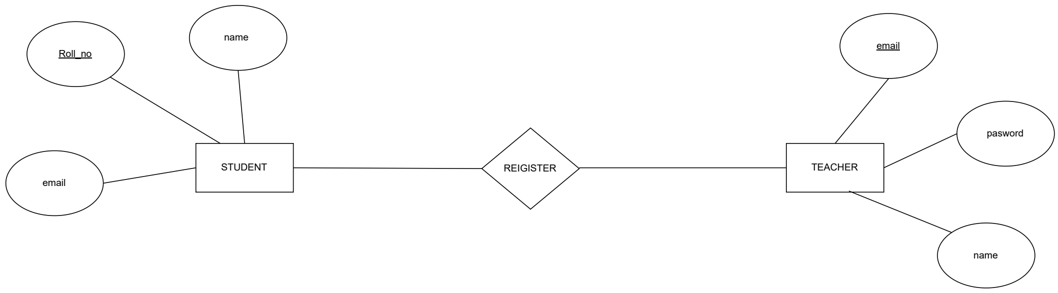

Figure 4.8: ER Diagram

The ER diagram illustrates a system where students and teachers access class details and generate/view reports after authentication. The Student entity includes attributes such as Roll No, Name, Age, Address, Email, and Class, while the Teacher entity consists of ID, Name, Email, Password, and Subject. The Register relationship connects students and teachers, allowing them to access class details that include Start Time, Total Time, and Class Link. During the class, the system monitors student engagement through facial recognition and head posture analysis, generating reports that track Presence Time and categorize engagement levels as Highly Engaged, Bored, Sleepy, or Confused. Both students and teachers can view these reports to analyze class performance. The system ensures secure authentication, maintains data privacy, and provides actionable insights to improve student engagement and learning outcomes.

Chapter 5

Implementation

5.1 Implementation

5.1.1 Module Split up

(a) Eye Status

- The Eye Aspect Ratio (EAR) is used to detects open or closed state of the eye by measuring the distances between specific eye landmarks.
- This approach calculates a scalar value based on Euclidean distances between selected points around the eye, which indicates whether the eye is open or closed.
- When the eye is open, the height of the eye (distance between upper and lower eyelid points) is relatively large compared to the width. This makes the EAR value stay above a certain threshold (often around 0.25).
- When the eye closes, the height becomes very small as the eyelids meet, causing the EAR value to drop significantly, approaching zero.
- eye status is given EI value as open eye will get a value of 2. 5 and closed eye will get a value of 0.

5.2 Tools And Techniques

- VGG-16: It is a deep learning model pre-trained on large image datasets. It is used in this system to analyze facial expressions and accurately predict student emotions such as happiness, sadness, or confusion, contributing to real-time engagement detection during online learning.

- Dlib's Shape Predictor: Detects specific facial landmarks, including the eyes, nose, and mouth, by identifying distinct points on the face. This enables precise tracking of facial features for tasks like gaze direction, emotion recognition, and head pose estimation.

- The system uses a Convolutional Neural Network (CNN) model, specifically VGG-16 This model is trained on the FER 2013 dataset, which contains labeled images of various emotions such as happiness, sadness, and anger, allowing the system to recognize facial expressions efficiently.

5.3 Login Page

The Login functionality allows students and teacher to access their accounts securely, with each session requiring valid credentials for entry. This login process is fortified with secure protocols, which helps prevent unauthorized access to sensitive data and control features. The system has two login page,one for student and other for teacher.

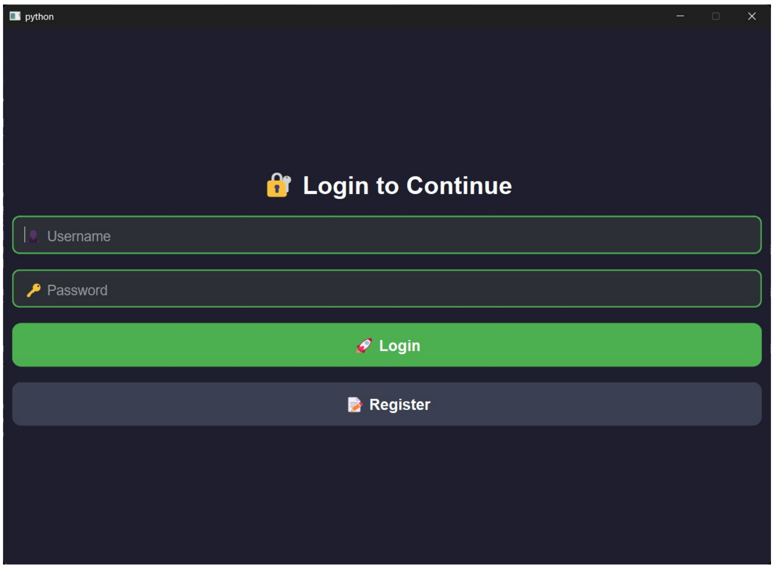

Figure 5.1: Teacher Login

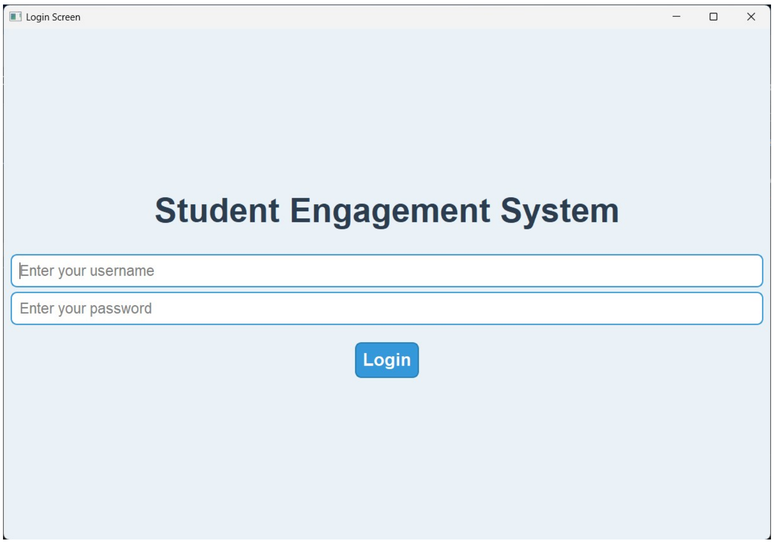

Figure 5.2: Student Login

5.4 Registration Page

The system has two separate Registration page,one for student and other for teacher.The user can input and store their information .This ensure that only registered users can login.

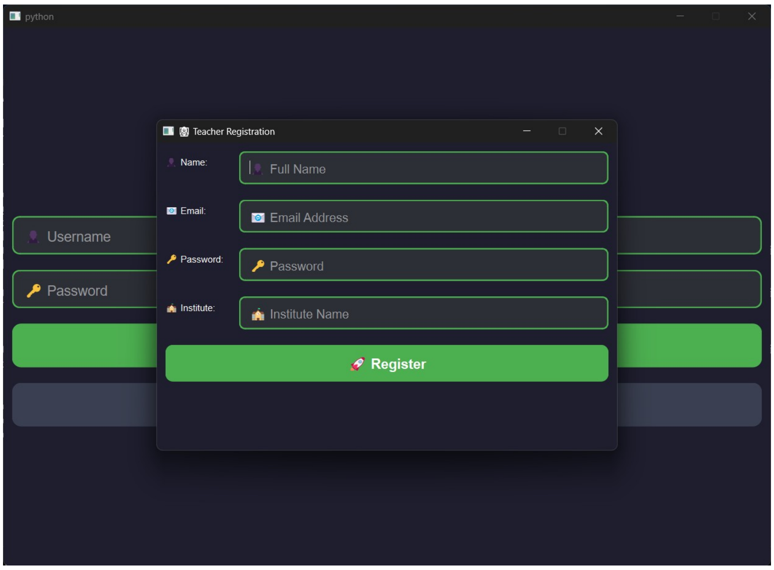

Figure 5.3: Teacher Registration

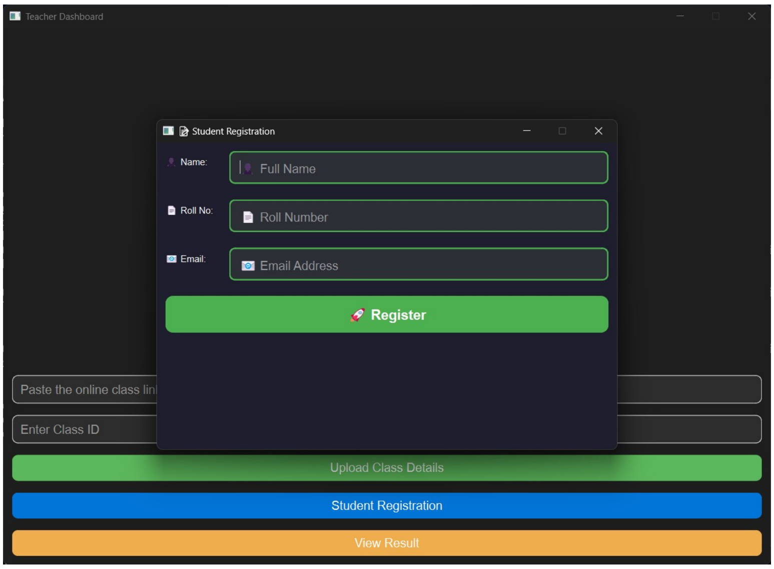

Figure 5.4: Student Registration

5.5 Fetching Student Data

The system allows users to analyse student engagement level in the class by inputting student Roll number and Class ID.The student score will be displayed for each minute.

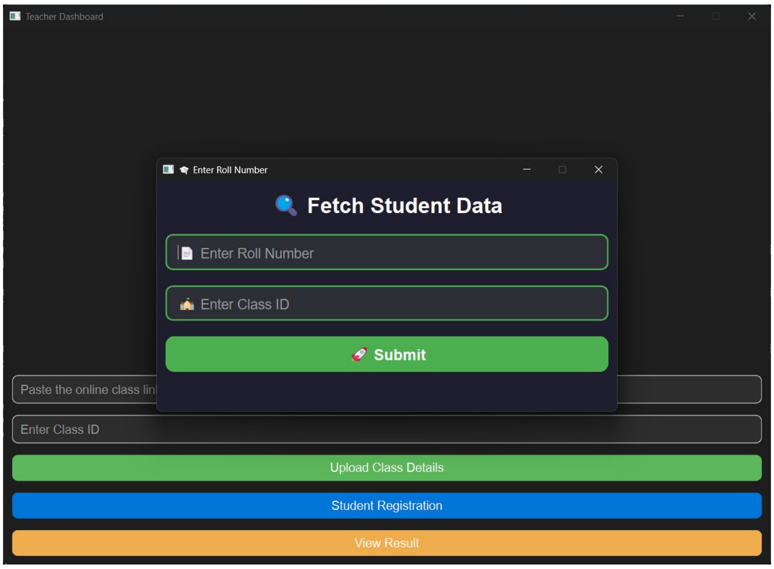

Figure 5.5: Fetching Student Data

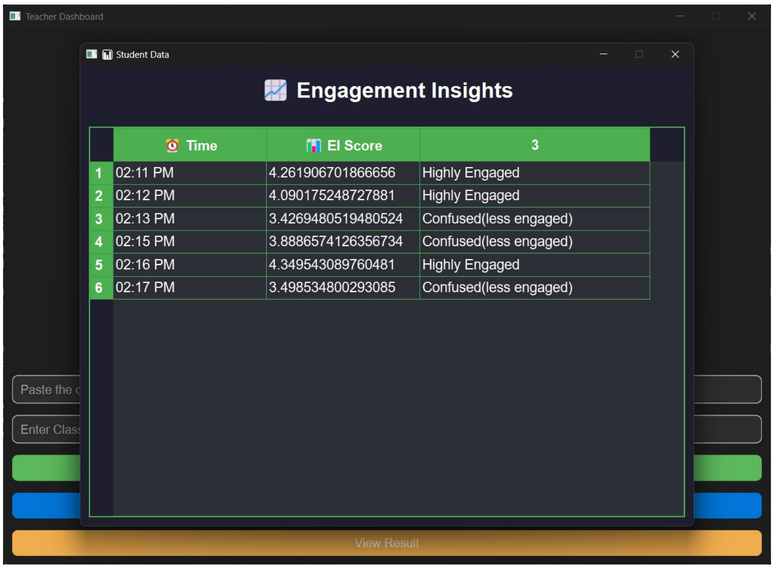

Figure 5.6: Engagement Insights

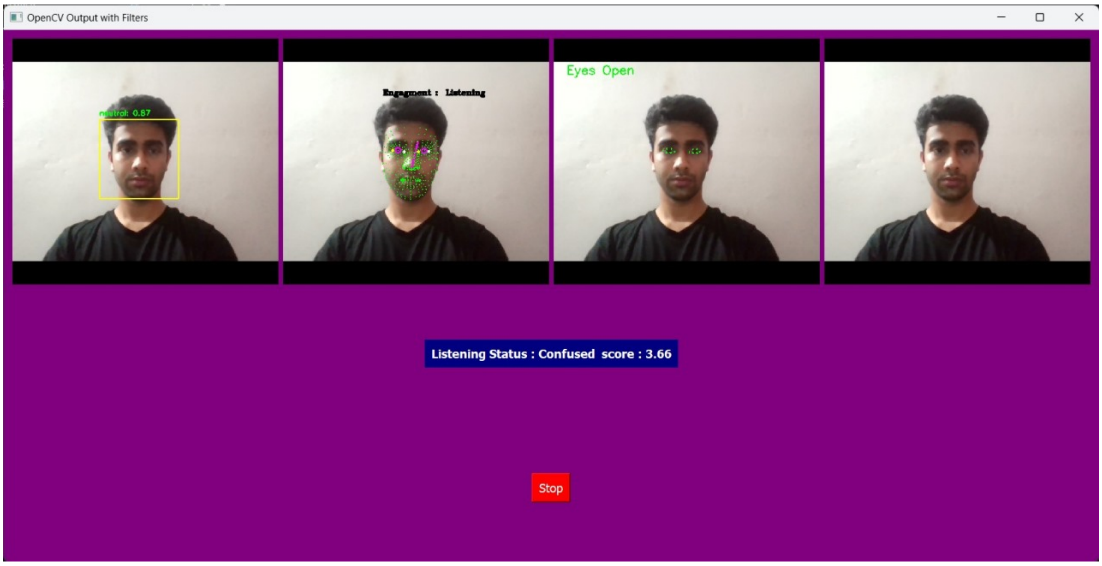

Figure 5.7: Displaying the engagement and total score

Chapter 6

Conclusion

Student engagement detection involves an AI model to detect the student engagement levels and a facial recognition system to detect the presence of the registered student. It is very much useful system in today's world as many educational institutions conduct their class online. Teachers duty to check weather the students is listening to class can be done automatically by AI models.It helps in improve the quality of the online class. The presence of AI model to detect two different aspects of the learning will provide more precise information towards the students attention in the class. Send the report of each student to the server for the teacher to view and analyze the report and give attendance to the student according to their participation in the online class.

References

[1] M. M. Santoni, T. Basaruddin, K. Junus, and O. Lawanto, "Automatic detection of students' engagement during online learning: a bagging ensemble deep learning approach," *IEEE Access*, 2024.

[2] A. Sukumaran and A. Manoharan, "Multimodal engagement recognition from image traits using deep learning techniques," *IEEE Access*, 2024.

[3] D. Mistry, M. F. Mridha, M. Safran, S. Alfarhood, A. K. Saha, and D. Che, "Privacy-preserving on-screen activity tracking and classification in e-learning using federated learning," *IEEE Access*, 2023.

[4] U. L. Sowjanya and R. Krithiga, "Decoding student emotions: An advanced cnn approach for behavior analysis application using uniform local binary pattern," *IEEE Access*, 2024.

[5] N. Alruwais and M. Zakariah, "Student recognition and activity monitoring in e-classes using deep learning in higher education," *IEEE Access*, 2024.