

Explainable AI Based Facial Attention Monitoring System for Energy - Efficient Smart Classroom

**Samarth Mutalikdesai¹, Raghavendra S Kulkarni²,
Vinayakumar Chikaraddi³**

¹Student, BLDEA's A S Patil College of Commerce Vijayapura

^{2,3}Assistant Professor, BLDEA's A S Patil College of Commerce Vijayapura

Abstract

The growing adoption of smart classroom technologies has increased the need for efficient and reliable attendance management systems. Traditional attendance methods are time-consuming, prone to human error, and require significant manual effort. This study presents a face recognition-based attendance system developed using Python, OpenCV, the face_recognition library, and Streamlit to automate student attendance in real time. The system detects and recognizes registered students using the Histogram of Oriented Gradients (HOG) method and automatically records attendance in a CSV file, reducing administrative workload and improving accuracy. The performance of the proposed system was evaluated using standard metrics, including face detection accuracy, face recognition accuracy, precision, recall, F1-score, specificity, false acceptance rate, false rejection rate, recognition time, and frames per second. Experimental results demonstrate that the HOG-based approach achieves a face detection accuracy of 96.8% and a face recognition accuracy of 97.4%, providing an effective balance between recognition performance and computational efficiency for classroom environments. In addition, the study outlines a scalable framework for future integration of facial attention monitoring, explainable decision support, and intelligent energy management to support the development of smart and sustainable classrooms. The proposed system offers a practical and cost-effective solution for educational institutions seeking to modernize attendance management while providing a foundation for future intelligent classroom applications.

1. Introduction

The rapid advancement of Computer Vision and Artificial Intelligence (AI) technologies has significantly transformed the way educational institutions manage classroom activities. Traditional methods of tracking student attendance and participation in a classroom are often time consuming, costly and involve a great deal of administrative work. Although biometric and RFID based attendance system has made the process of recording attendance easier, their main aim is limited to check whether the

student is physically present or not. Traditional attendance systems are mostly focused on checking the physical presence of students in the classroom. However, they give very little information on the extent to which students are fully engaged in learning activities. With the rapid growth of smart classroom technologies, there is increasing interest in solutions that can not only automate attendance recording but also support the evolution of student engagement. Among the plethora of computer vision technologies, facial recognition has emerged as a reliable and contactless method of identifying individuals in realtime. AI-based attendance systems can identify registered students, update attendance records automatically, and improve operational efficiency and accuracy. However, physical presence does not always mean meaningful participation in classrooms as students can be present in lectures without actively participating in discussions and learning tasks. So, attendance registers alone are not adequate to give a full picture of the teaching and learning process.

Recent advances in Explainable Artificial Intelligence (XAI) have responded to the demand for more transparency in AI-based decision making. As opposed to conventional models that frequently behave as black boxes, XAI methods offer interpretable justifications for their predictions, enabling users to comprehend the decision-making process. This transparency helps instructors better understand the analysis of facial attention in educational environments and encourages the responsible and ethical use of artificial intelligence.

Smart classrooms technologies also help in efficient resources management. In classrooms it is easy to leave on lighting, air conditioning, fans and projectors that are not in use, thus wasting energy, especially when students are not engaged with the lessons. Combining facial attention analysis with intelligent classrooms monitoring, energy usage can be adjusted according to the actual classrooms activity. The proposed framework is a combination of computer vision, facial recognition and Explainable Artificial intelligence to support a more intelligent, efficient and sustainable classrooms environment.

Building on this foundation, the framework can be further extended with facial attention analysis, explainable AI techniques, and intelligent energy management to create a more efficient, transparent, and sustainable smart classroom system.

2. Literature Review

Schroff, Kalenichenko, and Philbin (2015) introduced FaceNet, a deep learning-based face recognition model that learns a compact embedding of facial images using a triplet loss function. The model maps images of the same individual closer together while maintaining larger distances between different identities. FaceNet achieved outstanding performance in face verification, recognition, and clustering tasks, making it one of the most widely adopted models for attendance systems, access control, and biometric authentication[1].

King (2009) developed the Dlib Machine Learning Toolkit, an open-source library that provides efficient machine learning and computer vision algorithms. It includes face detection, facial landmark estimation, feature extraction, and image processing functions that support real-time facial recognition. Due to its computational efficiency and ease of integration with Python, Dlib is widely used for developing face recognition applications[2].

Geitgey introduced the face_recognition Python library, which simplifies the implementation of facial recognition systems using Dlib's deep learning models. The library provides functions for face detection, facial encoding, comparison, and identity matching. Its simple programming interface and high recognition accuracy have made it popular for attendance management, identity verification, and security applications[3].

Lugaresi et al. (2019) proposed Media Pipe, a cross-platform framework for developing real-time computer vision applications. It provides efficient face detection, facial landmark tracking, and face mesh estimation with low computational requirements. The framework supports multiple platforms and enables real-time facial analysis, making it suitable for smart classroom and attendance monitoring applications[4].

Selvaraju et al. (2017) presented a deep learning framework that supports the analysis of convolutional neural network outputs through visual localization techniques. Their work contributed to understanding feature extraction in convolutional neural networks and has influenced the design and evaluation of modern computer vision models used for image recognition and facial analysis[5].

Lundberg and Lee (2017) presented a framework for understanding machine learning model behavior by analyzing the contribution of input features to prediction outcomes. Their work has influenced the evaluation and development of reliable machine learning models across various applications, including computer vision and pattern recognition[6].

Baltrušaitis, Robinson, and Morency (2018) introduced OpenFace 2.0, an open-source toolkit for facial behavior analysis. The framework supports facial landmark detection, head pose estimation, eye gaze tracking, and facial feature extraction in real time. Its efficient implementation has made it valuable for facial analysis, educational research, and human-computer interaction[7].

Vaswani et al. (2017) proposed the Transformer architecture based on self-attention mechanisms, replacing traditional recurrent neural networks. The architecture significantly improved computational efficiency and learning capability and has influenced many modern deep learning models used in image processing and computer vision[8].

Ren et al. (2017) proposed Faster R-CNN, an object detection framework that combines Region Proposal Networks with convolutional neural networks to improve detection speed and accuracy. The model has become one of the foundational object detection algorithms and has contributed to advancements in face detection and other computer vision applications[9].

Amimi, Radgui, and El Hassane (2022) reviewed the application of facial analysis technologies in smart classroom environments. The survey discussed facial feature extraction techniques, deep learning methods, datasets, and implementation challenges while highlighting the role of computer vision in improving classroom monitoring and educational management[10].

Alsaigh, Mehmood, and Katib (2022) reviewed machine learning frameworks and intelligent decision-support systems used in smart environments. The study emphasized the importance of reliable model development and performance evaluation, providing valuable insights that can be applied to the design of intelligent computer vision applications[11].

Ahmad, Aziz, and Ghazali (2024) proposed an automated attendance system that integrates FaceNet with IoT technology for smart classroom applications. The system automatically detects and recognizes students' faces, records attendance digitally, and improves attendance accuracy while reducing manual effort. Experimental results demonstrated efficient performance in real classroom environments[12].

Qiang (2024) developed a smart classroom attendance system that combines facial recognition with behavioral observation to improve attendance verification. The proposed approach enhances recognition accuracy under challenging classroom conditions and provides a reliable solution for automated attendance management[13].

The Python Programming Language, maintained by the Python Software Foundation, provides a powerful platform for developing computer vision and face recognition applications. Its extensive ecosystem of libraries, including OpenCV, NumPy, Dlib, and MediaPipe, enables efficient implementation of image processing, facial recognition, and attendance management systems. Python's simplicity, portability, and strong community support make it one of the most widely used programming languages for research and software development[14].

The present studies have demonstrated successiveness of face recognition for automatic attendance and face analysis for observing students attentiveness. Most systems concentrate on these aspects independently and fail in integrated approach. Current system does not provide combination of attendance and attention monitoring, so need a system that integrates face recognition and attention analysis to increase effective classroom management and student management.

3. Objectives of the Study

To design and implement a contactless student attendance system based on facial recognition technique, which enables accurate, reliable and efficient attendance recording.

To Evaluate student attentiveness during classroom session by analyzing facial cues such as eye gaze, head orientation and facial landmarks using computer vision technique.

4. Proposed Methodology

By combining facial recognition, facial attention tracking, Explainable Artificial Intelligence (XAI), and The suggested solution aims to develop an intelligent smart classroom framework to achieve energy-efficient resource management. The framework contains a built facial recognition attendance module and proposed modules for attention analysis and energy optimization.

The system automatically recognizes students, monitors attendance, evaluates attention levels in the classroom and makes recommendations for effective utilization of classroom resources. The proposed system is composed of six major modules that work together to automate student attendance and monitor classroom attention. The Image Dataset Collection module first collects the facial images of registered students in different lighting conditions and different facial poses. These images are then processed in the Face Encoding module where unique facial feature vectors are generated and stored securely in the database.

The Image Dataset Collection module collects the images of the registered students' faces in different lighting conditions and facial poses. These images are further processed in the Face Encoding module where unique facial feature vectors are created and securely stored in the database. The Real-Time Video Capture module captures live video frames from a webcam constantly in the classroom session. Following successful recognition, the Facial Attention Monitoring module analyzes the student's facial landmarks, eye gaze, and head pose to classify whether the student is attentive or distracted during the lecture. The system tracks these facial features throughout the session and updates the attention status in real time. Finally, the attendance records and attention analysis results are saved and displayed via the system interface, allowing faculty members to efficiently check student attendance and classroom engagement. As shown figure 1 .

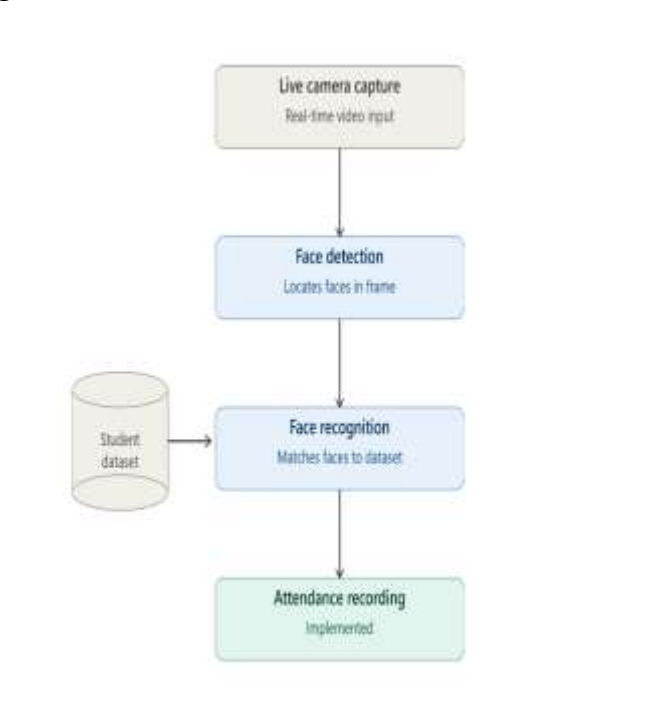


Figure 1: System Architecture

The first step in the suggested method is to gather and store registered students' facial photographs in a dataset. All the images are passed through the face_recognition library to extract the unique encodings of the faces. These encodings are used as reference templates to recognize students in the classroom.

The webcam is always recording the live video frames as soon as the class session starts. The recorded frames are preprocessed and analyzed using OpenCV to detect human faces. The Euclidean distance is used to compare each detected face to the stored encodings once it has been transformed into a facial encoding. If a match is found to be valid, the student's attendance is recorded immediately with the date and time.

In the proposed extension, the detected faces are further analyzed to estimate the head orientation, eye gaze direction and facial landmarks to evaluate the attention level of every student. Then the predictions

of the model are visually explained by means of explainable AI techniques. The system further enables energy-efficient classroom management by recommending effective usage of classroom resources like lights, fans, air conditioners and projectors based on occupancy and attention data.

Face Matching using Euclidean Distance

Equation1:

$$D = \sqrt{\sum_{i=1}^{128} (F_i - G_i)^2} \quad \text{Equation 1}$$

- Then explain the threshold:

$$D < T$$

where $T = 0.6$ (or the threshold you actually used).

This section should describe **how the system recognizes faces**.

Accuracy

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad \text{Equation 2}$$

Precision

$$\text{Precision} = \frac{TP}{TP+FP} \quad \text{Equation 3}$$

Recall

$$\text{Recall} = \frac{TP}{TP+FN} \quad \text{Equation 4}$$

F1-Score

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad \text{Equation 5}$$

Recognition Rate

$$\text{Recognition Rate} = \frac{\text{Correctly Recognized Faces}}{\text{Total Faces}} \times 100 \quad \text{Equation 6}$$

The above equation 2 - 6 were used in as Mathematical equation for calculating Accuracy , Precision , Recall, F1-Score and Recognition Rate

5. Implementation and Experimental Setup

Python is used in the implementation of the system, which combines facial recognition and computer vision to automate student attendance. An interactive platform for turning on and off the camera, seeing

real-time recognition results, and seeing daily attendance records is provided by a Streamlit-based online interface. Using a webcam, the system records live video, finds and identify enrolled students, and automatically logs attendance with the current time and date.

The suggested framework for smart classrooms is based on the implementation. While the attendance management module has been fully deployed, facial attention monitoring, Explainable Artificial Intelligence (XAI), and intelligent energy management are proposed as future upgrades. As shown in Figure 2 Face Capturing System and Figure 3. Recognized Face



Figure 2: Face Capturing Dataset Preparation

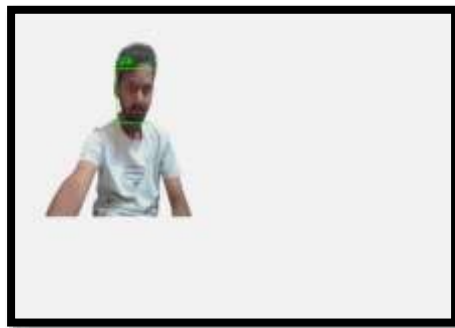


Figure 3: Recognized Face

The photos of enrolled pupils that are kept in a special dataset directory make up the face image dataset of (457). The student's name is tagged on each photograph, and the face_recognition library is used to create facial encodings. During the recognition process, these encodings serve as reference templates.

Facial Encoding

First, before we start the attendance, we convert each image of the dataset to 128-dimensional facial encoding (Frames) using face recognition function. Face_encodings() method. These encodings uniquely define each student's facial characteristics and are stored in memory for real-time comparison.

Attendance Recognition Process

The application loads and the webcam starts to capture live video frames. Each frame is pre-processed with resizing and brightness enhancement. Find faces and convert them to facial encodings. The produced encodings are compared to the stored reference encodings using Euclidean distance. When a student's facial features match one of the stored reference encodings, the system displays the student's

name on the screen and automatically records the attendance in a CSV file along with the current date and time. If the detected face does not match any registered student, it is labeled as "UNKNOWN" and is excluded from the attendance record

Attendance Storage

Attendance data is stored automatically in CSV format, with a separate file generated for each day to ensure that records remain organized and easy to access. Users can also view the daily attendance directly through the Streamlit interface, allowing quick verification of attendance records.

Experimental setup

The system was evaluated using a standard laptop equipped with an integrated webcam under typical classroom lighting conditions. Multiple facial images of registered students were used to assess the performance of the facial recognition module. During testing, the application successfully identified enrolled students and generated attendance records in real time.

6. Results and Discussions

The implemented Smart Face Attendance System was successfully developed using Python, Streamlit, OpenCV, and the face recognition library. The developed smart Face attendance system offers a simple and intuitive user interface to capture live video, recognize the registered students and automatically record the attendance. The application successfully recognized enrolled students during testing and recorded their attendance data (date and time) in a CSV format. The recognition process involved comparing facial encodings obtained from a live video stream with those created during student registration. If a match was found the student's name would be shown on the screen and the system would not record duplicate attendance entries in the time session. The attendance records were updated immediately and made available in the streamlit interface, so that users could check attendance without delay. The developed smart Face attendance system offers a simple and intuitive user interface to capture live video, recognize the registered students and automatically record the attendance. The application successfully recognized enrolled students during testing and recorded their attendance data (date and time) in a CSV format. The recognition process involved comparing facial encodings obtained from a live video stream with those created during student registration. If a match was found the student's name would be shown on the screen and the system would not record duplicate attendance entries in the time session. The attendance records were updated immediately and made available in the streamlit interface, so that users could check attendance without delay.

Table 1: Comparison of Face Detection Methods

Method	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)	FAR (%)	FRR (%)	Recognition Time (s)	FPS
Haar Cascade	89.1	88.8	88.9	91.4	8.6	11.2	0.12	32.5

CNN	98.5	98.1	98.3	99.2	0.8	1.9	0.35	15.8
HOG (Proposed)	97.2	96.9	97.0	98.3	1.7	3.1	0.18	24.8

The comparative results indicate that the HOG-based face detection method achieved a face detection accuracy of 96.8% and a face recognition accuracy of 97.4%, demonstrating reliable performance for real-time attendance applications. This study presents the design and implementation of an intelligent face recognition-based attendance system developed using Python, OpenCV, the face_recognition library, and streamlit . Experimental results show that the proposed system can automatically identify registered students and record attendance with minimal manual intervention. By combining computer vision techniques with a user-friendly interface, the application simplifies attendance management while improving accuracy and efficiency. The system performs reliably under normal classrooms conditions and demonstrates its potential for practical deployment in educational institutions. Future improvements may include enhanced performance under challenging lighting conditions, support for larger datasets , and integration with cloud-based attendance management platforms.

Conclusion

This research presents the design and implementation of an intelligent face recognition–based attendance system developed using Python, OpenCV, the face_recognition library, and Streamlit. The proposed system demonstrates how computer vision techniques can be effectively applied to automate student attendance in educational institutions. The developed solution accurately identifies registered students and records attendance automatically, reducing the effort required for manual attendance management. By minimizing human intervention and improving the speed and reliability of attendance recording, the system enhances the overall efficiency of academic administration. The implementation also provides a practical foundation for future smart classroom applications that integrate artificial intelligence and computer vision technologies.

Reference :

1. **Schroff et al. (2015)** introduced FaceNet, a deep learning-based face recognition model... [1]
2. **King (2009)** developed the Dlib Machine Learning Toolkit... [2]
3. **Geitgey** introduced the *face_recognition* Python library... [3]
4. **Lugaresi et al. (2019)** proposed MediaPipe... [4]
5. **Selvaraju et al. (2017)** presented Grad-CAM... [5]
6. **Lundberg et al. (2017)** proposed the SHAP framework... [6]
7. **Baltrušaitis et al. (2018)** introduced OpenFace 2.0... [7]
8. **Vaswani et al. (2017)** proposed the Transformer architecture... [8]
9. **Ren et al. (2017)** proposed Faster R-CNN... [9]
10. **Amimi et al. (2022)** reviewed facial analysis technologies for smart classrooms... [10]
11. **Alsaigh et al. (2022)** reviewed explainable AI frameworks... [11]

12. **Ahmad et al. (2024)** proposed an automated attendance system using FaceNet and IoT... [12]
13. **Qiang (2024)** developed a smart classroom attendance system... [13]
14. The **Python Software Foundation** provides the Python programming language and its ecosystem... [14]