

A Multimodal Biometric Attendance System Using Face and Voice Recognition

Anuja Suresh Vishe

MSC Information Technology

DG Ruparel College Mumbai, Maharashtra, India

Academic Year:2025-2026

Abstract

This paper presents a multimodal biometric attendance system that integrates face recognition and voice authentication to automate attendance management in academic environments. Traditional attendance methods are often inefficient and vulnerable to manipulation such as proxy attendance. The proposed system utilizes facial feature extraction and speaker embedding techniques to ensure accurate identification of students. A role-based architecture is implemented, allowing teachers to manage subjects and record attendance, while students authenticate using biometric inputs. The system is supported by a cloud-based backend for real-time data storage and retrieval. Experimental results indicate improved accuracy, reduced manual effort, and enhanced security compared to conventional methods. The integration of multiple biometric modalities increases system robustness and reliability [1][4].

Keywords

Multimodal Biometrics, Face Recognition, Voice Recognition, Attendance System, Authentication, Machine Learning, Cloud Database

1. Introduction

Attendance tracking is an essential task in educational institutions, traditionally performed using manual methods such as roll calls or paper-based registers. These approaches are time-consuming, error-prone, and susceptible to fraudulent practices like proxy attendance. With advancements in artificial intelligence and biometric technologies, automated systems have emerged as effective alternatives [1].

Biometric systems rely on unique physiological or behavioral characteristics such as face, fingerprint, or voice for identification. Face recognition is widely used due to its non-intrusive nature, while voice recognition provides an additional authentication layer. However, unimodal systems often face limitations due to environmental factors or data variability [2][3].

To overcome these challenges, multimodal biometric systems combine multiple identification methods to improve accuracy and robustness. This paper proposes a system that integrates face and voice recognition for secure and automated attendance management using a cloud-based architecture.

2. Literature Review

Various studies have explored biometric-based attendance systems using different modalities. Face recognition systems have been widely implemented using machine learning and deep learning techniques. Methods such as Eigenfaces and deep neural networks have demonstrated effective performance in identifying individuals from images [5].

Voice recognition systems analyze speech signals to authenticate users based on vocal features. These systems utilize feature extraction and pattern matching techniques but are often sensitive to noise and environmental disturbances [4].

Recent research highlights the advantages of multimodal systems, where combining multiple biometric traits enhances reliability and reduces error rates. Such systems mitigate the weaknesses of individual modalities and provide better security against spoofing attacks [1].

Despite these advancements, many existing systems lack real-time processing capabilities and user-friendly interfaces. The proposed system addresses these limitations by integrating biometric recognition with a cloud-based platform and an interactive user interface.

3. Proposed System Architecture

3.1 System Architecture

The proposed system follows a client-server architecture consisting of two primary user roles: teacher and student. The frontend is developed using a Python-based framework, while the backend is implemented using a cloud database service [8].

The system includes the following components:

- User Authentication Module
- Face Recognition Module
- Voice Recognition Module
- Subject Management Module
- Attendance Processing Module
- Cloud Database

Teachers can create subjects, manage enrollments, and record attendance, while students can register, authenticate, and access attendance records.

3.2 Data Collection Module

The data collection module captures biometric data during user registration. Facial images are acquired through a camera interface, and voice samples are recorded using a microphone. The collected data is preprocessed to improve quality and remove noise.

Facial images are converted into feature embeddings using pre-trained models, while voice signals are processed into spectral representations for speaker identification. These embeddings are stored securely in the database and used for future matching [6][7].

3.3 Attendance Processing Module

The attendance processing module integrates outputs from face and voice recognition systems. During attendance marking, multiple images or audio inputs are analyzed to detect and identify students.

The system aggregates detection results across multiple inputs to improve reliability. Students are marked present if identified successfully in at least one input source. The final attendance records are generated with timestamps and stored in the cloud database for further analysis.

4. Methodology

4.1 Registration Phase

Students and teachers register by providing personal details and biometric data. Facial embeddings and voice embeddings are generated and stored in the database.

4.2 Face Recognition

Face recognition is implemented using feature extraction techniques that generate 128-dimensional embeddings. A Support Vector Machine (SVM) classifier is trained on these embeddings to identify students. Matching is validated using Euclidean distance thresholds [6][7].

4.3 Voice Recognition

Voice recognition is performed using speaker embedding models. Audio signals are processed to extract features, and similarity scores are computed using cosine similarity. A threshold-based approach is used to determine identity [4].

4.4 Multimodal Fusion

The system combines face and voice recognition results to enhance accuracy. If one modality fails due to environmental factors, the other can compensate, ensuring reliable attendance marking.

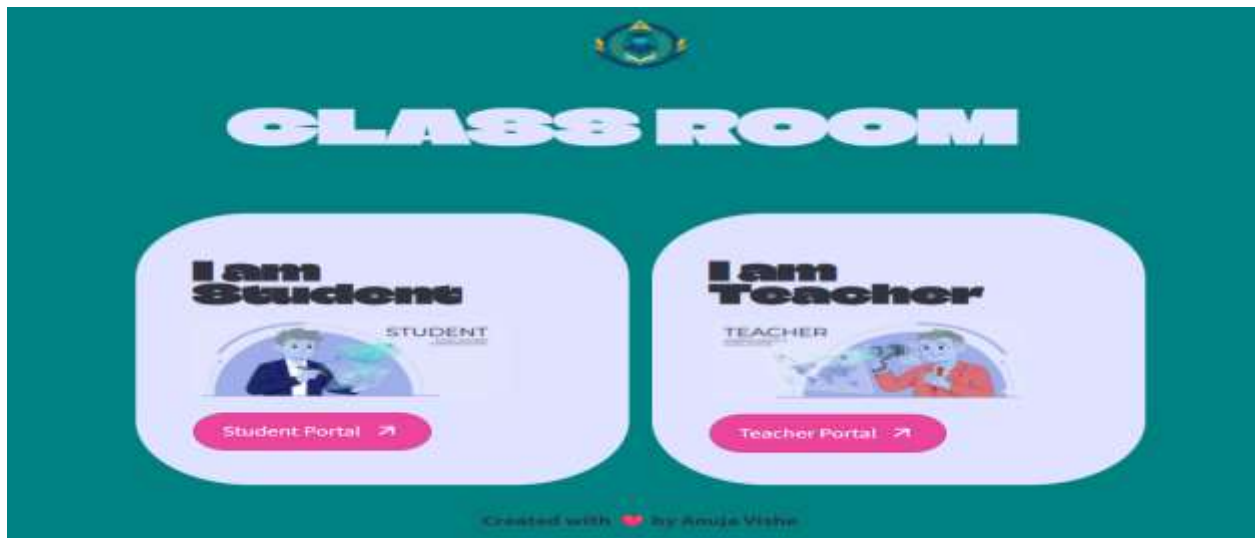
4.5 Attendance Logging

Detected student identities are mapped to enrolled subjects, and attendance records are stored with timestamps in the database.

5. System Implementation and UI

The system is developed using Python with a Streamlit interface and integrates authentication, biometric processing, and attendance management modules. Supabase is used as the backend for secure storage of user data, subjects, and attendance records. Face recognition is implemented using dlib with an SVM classifier, while voice recognition uses Resemblyzer for speaker identification. The system provides separate dashboards for teachers and students, supporting subject creation, enrollment, and attendance marking through image, camera, or voice inputs with real-time cloud synchronization.

5.1 Home Page Interface



5.2 Student dashboard



5.3 Teacher Dashboard

**Class Room**

Welcome,Anuja

Logout Ctrl + Backspace

Take attendanceManage SubjectsAttendance record

Take AI attendance

Select Subject

Computer network - 1

ADD PHOTOS

Clear all photosRun face analysisuse voice attendance

Created with ❤ by Anuja Vishe

5.4Attendance Record

Attendance reports

Please review attendance before confirming.

Name	ID	Source	Status
anuja	1	photo1	 Presnt

discred

confirm and save

6. Experimental Results and Discussion

The proposed multimodal biometric attendance system was tested using face and voice inputs under different classroom-like conditions. The face recognition module showed reliable performance in controlled lighting environments, while the voice recognition module provided additional verification in normal noise conditions [6][7]. The combination of both modalities improved overall accuracy and reduced misclassification compared to single biometric systems [1][5]. The system successfully identified students and recorded attendance in real time using

a cloud database [8]. However, performance was slightly affected in cases of poor lighting, occlusion, or background noise, indicating the need for improved robustness in uncontrolled environments [2][3].

7. Conclusion and Future Work

This paper presented a multimodal biometric attendance system using face and voice recognition for automated attendance management. The system reduces manual effort, improves accuracy, and prevents proxy attendance by using biometric authentication [1][4]. The integration of a cloud-based backend ensures real-time data storage and easy access for teachers and students [8]. Future improvements may include the use of deep learning models for higher accuracy [5], deployment as a mobile application, addition of more biometric modalities such as fingerprint recognition, and enhancement of system performance in noisy and low-light environments [2][3].

8. References

- [1] A. K. Jain, A. Ross, and S. Prabhakar, "An Introduction to Biometric Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, 2004.
- [2] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, 1991.
- [3] D. A. Reynolds, "An Overview of Automatic Speaker Recognition Technology," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2002.
- [4] R. Brunelli and T. Poggio, "Face Recognition: Features versus Templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1993.
- [5] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [6] dlib Documentation, Davis E. King, 2009.
- [7] Resemblyzer Documentation, 2019.
- [8] Supabase Official Documentation, 2020–present.