12th IEEE International Conference on Sustainable Technologies and Engineering (IEEE iCOSTE 2026)
1–3 November 2026, Cox's Bazar, Bangladesh

A Hybrid Attendance Monitoring Framework Using Fingerprint Authentication and FaceNet-SVM Recognition

Md. Mehedi Hasan Naeem
Dept. of Computer Science & Engineering
Jatiya Kabi Kazi Nazrul Islam University
Mymensingh, Bangladesh
mehedinaeem00@gmail.com

Mst. Kamrunnahar Ruma
Dept. of Computer Science & Engineering
Jatiya Kabi Kazi Nazrul Islam University
Mymensingh, Bangladesh
kamrunnaharruma242@gmail.com

Nafiza Anjum
Dept. of Computer Science & Engineering
Jatiya Kabi Kazi Nazrul Islam University
Mymensingh, Bangladesh
nafizaanjum.2002@gmail.com

Shakila Sultana
Dept. of Computer Science & Engineering
Jatiya Kabi Kazi Nazrul Islam University
Mymensingh, Bangladesh
shakila161470@gmail.com

Md. Sujan Ali
Dept. of Computer Science & Engineering
Jatiya Kabi Kazi Nazrul Islam University
Mymensingh, Bangladesh
sujan_cse@jkkniu.edu.bd

***Abstract*—Accurate attendance management is a very important requirement in an educational institute. But the existing methods have many practical problems. Traditional roll calls take up class time and are subject to proxy attendance. A single biometric scan, however, only verifies that the student is who they say they are at the time of scanning, but does not indicate whether or not the student remains in the classroom for the duration of the class. This paper proposes a hybrid attendance monitoring framework combining fingerprint based identity verification with CCTV assisted FaceNet-SVM face recognition. The fingerprint module authenticates the student at the entrance of the classroom and video frames sampled during the class are used to identify the students and estimate their overall presence. Face images are converted into FaceNet embeddings, which are classified using a Support Vector Machine (SVM). Attendance is considered valid when the fingerprint authentication is successful and the student's recorded presence reaches a 75% threshold. The system was evaluated on a classroom dataset of 10 students comprising 531 processed face images, together with a recorded classroom video simulating a CCTV stream. It achieved a recognition accuracy of 98.13% on the held-out test set and separated students with high classroom presence from those with low or undetected presence. The main contribution is a session-level attendance decision rule that links entry-time fingerprint verification with repeated visual evidence collected during the class.**

***Index Terms*—Attendance Monitoring, CCTV, Face Recognition, FaceNet, Fingerprint Authentication, Hybrid Attendance System, Presence-Based Attendance, Support Vector Machine (SVM).**

I. Introduction

Monitoring students' attendance is an important responsibility of universities and colleges. It helps institutions to keep track of students' involvement and to maintain proper academic records. In many institutions attendance is still taken in the traditional way by roll calls. These techniques are easy to implement, but they consume class time, increase manual work, and make proxy attendance more likely. These limitations have led to the research of automated attendance management systems. Recent research has revealed that machine-learning-based facial recognition can enable more efficient attendance recording in college environments [1].

Biometric methods can improve the reliability of attendance systems because records are directly linked to the student's identity. Fingerprint based systems are quick and convenient for identity verification during registration [2], but they work only at the time of scanning and hence cannot establish if the student is still in the classroom. Camera-based systems compensate for this by allowing contactless, repeated observation of presence throughout a class [3] but their performance depends on visual information and can be affected by illumination, pose, image quality and partial occlusion [4].

Deep learning has also improved visual recognition. FaceNet maps a face image to a compact embedding space suitable for identity comparison [5], MTCNN can detect and align faces before recognition, which is especially useful in less controlled environments [6], and SVM remains useful when deep embeddings are already discriminative and the number of identities is moderate [7].

Motivated by these observations, this work proposes a hybrid attendance monitoring framework that combines fingerprint authentication with FaceNet-SVM-based CCTV recognition. Fingerprint verification establishes the student's identity at the classroom entry point, while classroom video is sampled to obtain repeated visual evidence of presence. The final decision is based on aggregated observations, and attendance is granted only when the fingerprint verification succeeds and the recorded presence satisfies a 75% threshold.

The main contribution of this work is not a new face-recognition model, but the integration of complementary bio-

979-8-3195-2804-9/26/$31.00 © 2026 IEEE

metric and computer-vision mechanisms into a single session-level attendance decision. Specifically, this paper makes the following contributions:

- A hybrid attendance framework in which entry-time fingerprint verification and frame-wise face recognition jointly determine the attendance status.
- A presence-percentage decision rule that converts repeated frame-level detections into an interpretable Present/Absent outcome using a 75% threshold.
- A prototype evaluation on a 10-student dataset of 531 processed face images and a recorded classroom video, with the code made publicly available.

The remainder of this paper is organized as follows. Section II reviews related work and positions the proposed framework against it. Section III describes the datasets and implementation environment. Section IV presents the system architecture and decision rule. Section V reports and discusses the experimental results. Section VI concludes the paper, and Section VII provides data and code availability information.

II. Related Work

Automated attendance systems increasingly use biometric and computer-vision methods to reduce manual work and prevent proxy attendance. This section groups existing studies by the type of evidence they use and identifies the gap addressed by the proposed framework.

A. Fingerprint-Based Attendance

Fingerprint-based systems reliably verify identity using registered biometric templates and are often integrated with IoT and web platforms for automated record management [2]. Deep learning has also been explored for fingerprint-based attendance monitoring [8], and practical fingerprint systems are easy to deploy in institutional environments [9]. However, these systems mainly confirm a student's identity at the time of scanning and do not measure continued classroom attendance.

B. Camera-Based Attendance

Camera-based systems address this limitation by enabling contactless observation of students in the classroom. Son et al. developed a CCTV-based attendance support system using deep face recognition [3]. Web-based face-recognition attendance systems have also been proposed for university classroom management [10], and recent CCTV-based studies show that surveillance-camera footage can support automated attendance monitoring [11]. These approaches demonstrate the feasibility of using classroom cameras, but their performance can be affected by illumination, pose, image resolution, and classroom variation.

C. Embedding-Based and Alternative Recognition Pipelines

Deep facial representations have further improved recognition-based attendance. FaceNet maps facial images into a compact embedding space [5], and MTCNN supports face detection and alignment before recognition [6]. FaceNet-based attendance systems show that deep embeddings can support student identification [12], and RetinaFace and FaceNet have been combined for automated attendance recognition [13]. Other studies have considered alternative strategies: HOG features with SVM [14], anti-spoofing and alert mechanisms to improve security [15], face segmentation and recognition [16], and multi-task convolutional networks [17]. These works contribute useful recognition or security components, but most focus on identifying a face, and the final attendance decision often remains tied to a single recognition event or a short observation.

D. Session-Level Attendance and Research Gap

TABLE I
Comparison of existing attendance approaches with the proposed system.

Approach	Main outcome	Limitation addressed by the proposed system
Fingerprint-based attendance [2], [8], [9]	Strong identity verification at scan time	Does not show whether the student stays in class
CCTV face recognition [3], [10], [11]	Contactless classroom observation	No entry-time biometric identity check
FaceNet-based recognition [12], [13]	Compact embeddings for student identification	Focus on identifying a face rather than defining an attendance rule
Session-level monitoring [18]–[21]	Attendance from repeated observations	Not combined with fingerprint-verified entry
Proposed system	**Fingerprint-verified identity with FaceNet-SVM session monitoring and a 75% presence decision**	**–**

Session-level approaches indicate that attendance should be calculated from repeated observations during class. AttenFace used face recognition for real-time attendance monitoring and supports the idea of periodic visual checking [18], and recent work highlights the value of automated attendance recording from surveillance streams [19]. Recent 2026 studies have placed cameras at the classroom entrance and inside the room to record entry, exit, and periods of presence [20], [22], and have added temporal analysis and adaptive thresholding to CCTV-based recognition to reduce false positives [21]. Video-based attendance with RetinaFace and FaceNet has likewise been performed on frames captured at regular intervals [23]. These systems rely on face recognition at every stage, whereas the proposed framework pairs fingerprint-verified entry with frame-wise FaceNet-SVM presence estimation. Table I summarizes how these groups of studies differ from the proposed framework.

Taken together, these studies show that attendance automation has progressed from simple biometric verification to continuous vision-based monitoring. However, many existing

systems still treat identity verification and attendance estimation as separate tasks. A more reliable approach should combine secure entry verification with continuous evidence collected throughout the class session. This motivates the proposed framework, in which fingerprint verification establishes identity and frame-level face recognition estimates actual classroom presence.

III. Datasets and Implementation Environment

A. Dataset

For the experimental evaluation, we constructed a multimodal dataset of 10 students with enrolled roll numbers 22102001, 22102013, 22102017, 22102023, 22102024, 22102025, 22102026, 22102027, 22102029, and 22102042. Facial data were captured from different angles and expressions so that the recognition pipeline could learn from realistic intra-class variation.



Fig. 1. Sample processed face images from the constructed multimodal dataset.

After face detection and preprocessing, the dataset contained 531 processed face images; samples captured from various angles are shown in Fig. 1.

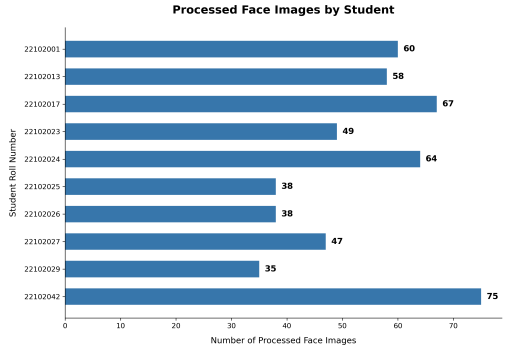


Fig. 2. Number of processed face images per enrolled student (identified by roll number). The counts range from 35 to 75 images per student.

For the fingerprint component, we selected 10 distinct fingerprint identities from the publicly available SOCOFing dataset, which is intended for academic fingerprint-recognition research [24], and assigned each one to the facial image set of one enrolled student. This pairing created a complete hybrid biometric profile for each student, allowing both fingerprint-based entry verification and continuous classroom monitoring to be represented in the evaluation. The distribution of the processed face-image is summarized in Fig. 2. The dataset is not perfectly balanced as the final count depends on successful face detection and filtering: it is between 35 and 75 images for each student, with mean 53.1 . We use this distribution consistently in the paper.

TABLE II
Dataset and Experimental Configuration Used in This Study.

Item	Value
Number of enrolled students	10
Processed face images	531
Images per student	35–75
Average images per student	53.1
Embedding dimension	128
Recognition classifier	SVM
Reported test samples	107
Attendance threshold	75%
Train/test split	80/20
SVM kernel	RBF
Regularization parameter C	10.0
Frame-sampling interval (recorded video)	3 s
Video frames for evaluation	20 sampled frames

Table II summarizes the dataset and experimental configuration.

B. Implementation Environment

The prototype was implemented in Python. OpenCV was used for image and video handling, FaceNet for embedding extraction, scikit-learn for SVM training and evaluation, and pandas/NumPy for tabular processing. The trained SVM model, label encoder, and embedding files were saved for later video-based attendance evaluation. The video test was performed by sampling frames from a recorded classroom video that simulates a CCTV stream.

C. Training and Sampling Parameters

The face embeddings were split 80/20 into training and test sets (test size 0.20), giving the 107 held-out test samples. The SVM uses a Radial Basis Function (RBF) kernel with regularization parameter $C = 10.0$. Frames were sampled at 3-second intervals for the recorded-video experiments, producing the 20 frames used for attendance evaluation, and at 180-second (3-minute) intervals for real-time CCTV deployment to limit the computational load. With 20 sampled frames, the 75% threshold corresponds to detection in at least 15 frames.

IV. Proposed System Architecture

The proposed system consists of four main stages: fingerprint verification, video frame sampling, FaceNet-SVM recognition, and attendance decision generation (Fig. 3). The fingerprint scanner first verifies the student's identity at the

classroom entrance and records a valid session entry. The CCTV/IP camera then captures the classroom video, and instead of processing every frame, the system selects frames at regular intervals (3 seconds for recorded video, 180 seconds for real-time CCTV deployment) to reduce the computational workload.

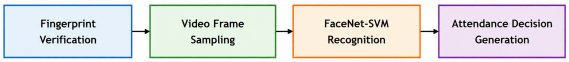


Fig. 3. Block diagram of the four main stages of the proposed system.

Each selected frame passes through face detection, preprocessing, embedding extraction, and classification. The framework is divided into modules. The authentication module verifies the fingerprint and retrieves the corresponding roll number from the registered student database. The video acquisition module captures the classroom video stream and selects the frames to be processed. The recognition module employs the trained FaceNet-SVM model to identify students from the detected faces. The attendance module integrates the recognition results from multiple frames to establish the final attendance status. The modular architecture makes testing, maintenance and extension of the system easier.

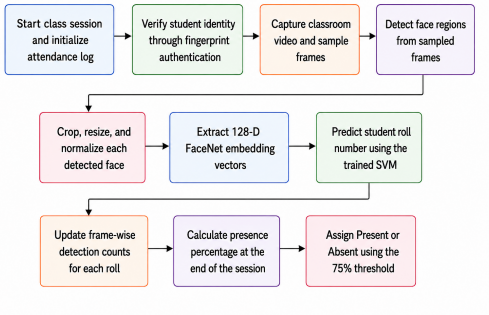


Fig. 4. Step-by-step procedure of the proposed system.

The whole procedure is summarized in Fig. 4. In each step, an identity record, recognition result, or attendance decision is generated. This will enable a teacher to check the output if the result is ambiguous.

A. Face Image Preprocessing

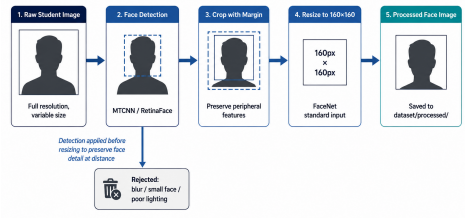


Fig. 5. Face image preprocessing pipeline applied before FaceNet embedding extraction.

The preprocessing step normalizes the input to the FaceNet encoder. The raw student images are processed by face detection, cropping with margin, resizing to 160×160 pixels and quality filtering. The detection before resizing helps to retain facial details in classroom images where the face may occupy a small part of the entire frame. Figure 5 illustrates our preprocessing pipeline.

B. FaceNet-SVM Recognition

Given the face image I detected, the FaceNet encoder produces an embedding vector:

$$\mathbf{e} = f(I), \quad \mathbf{e} \in \mathbb{R}^{128}, \tag{1}$$

where $f(\cdot)$ is the pre-trained FaceNet feature extractor. The embeddings are then used to train an SVM classifier for the 10 enrolled student identities. The trained classifier predicts the roll number of each detected face in a sampled video frame.This design separates feature extraction and classification, which is suitable for a compact dataset where training a deep recognition model from scratch would be unnecessary [5].

C. Attendance Decision Rule

Attendance is calculated from multiple detections, not a single frame. For a student s, the attendance percentage is given by

$$P_s = \frac{D_s}{N} \times 100, \tag{2}$$

where D_s is the number of detected frames with student s and N is the total number of processed frames. The final decision is

$$A_s = \begin{cases} \text{Present}, & P_s \geq 75\%, \\ \text{Absent}, & P_s < 75\%. \end{cases} \tag{3}$$

This threshold-based method derives classroom presence from aggregated evidence rather than from a single recognition result and gives an interpretable percentage for each student.

V. Experimental Results and Discussion

In this section, we first analyze the quality of the learned embeddings, then report the classification performance, and finally evaluate the full attendance decision on the classroom video.

A. Embedding Separability

The 128-dimensional FaceNet embeddings were visualized using t-SNE to inspect whether identity clusters were distinguishable in two dimensions (Fig. 6). As shown in Fig. 6, most clusters show visible separation, although some local overlap exists. Because a two-dimensional projection discards information, this supports applying the SVM to the original 128-dimensional embeddings, where more discriminative information is available.

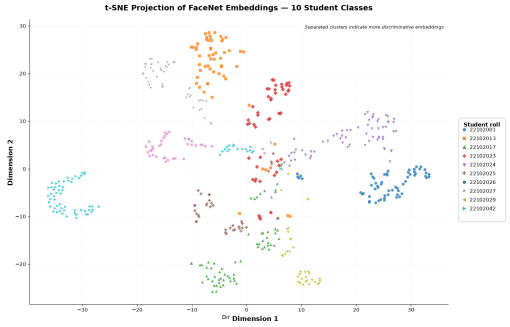


Fig. 6. t-SNE projection of the 128-dimensional FaceNet embeddings of the 10 enrolled students, colored by roll number.

B. Classification Performance

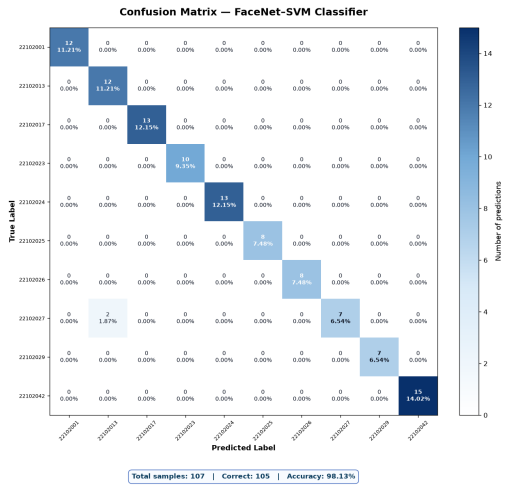


Fig. 7. Confusion matrix of the FaceNet-SVM classifier on the 107 held-out test samples.

The FaceNet-SVM classifier was tested on held-out test embeddings and obtained 98.13% accuracy, which means the embeddings are discriminative enough for the 10 enrolled identities.

TABLE III
CLASSIFICATION RESULTS OF THE FACENET-SVM CLASSIFIER ON THE HELD-OUT TEST SET (107 SAMPLES).

Metric	Value
Test samples	107
Correct predictions	105
Incorrect predictions	2
Accuracy	98.13%
Macro precision	0.99
Macro recall	0.98
Macro F1-score	0.98
Weighted precision	0.98
Weighted recall	0.98
Weighted F1-score	0.98

Fig. 7 shows the confusion matrix for test samples, with only 2 of the 107 samples misclassified. Most of the predictions are on the diagonal, showing little confusion between visually similar embeddings. The close macro and weighted averages in Table III indicate that the performance is balanced, although the number of images per student is not equal.These figures are to show feasibility given the small test set and controlled collection conditions, not general performance.

C. Video-Based Attendance Evaluation

The system processed 20 sampled frames of classroom video for attendance evaluation. Student 22102013 was detected in 18 frames (90.00% presence), student 22102017 in 17 frames (85.00%), and student 22102029 in 2 frames (10.00%). The remaining enrolled students were not detected in the sampled frames. Using the 75% threshold, the system marked 22102013 and 22102017 as Present, and marked 22102029 and the remaining students as Absent.

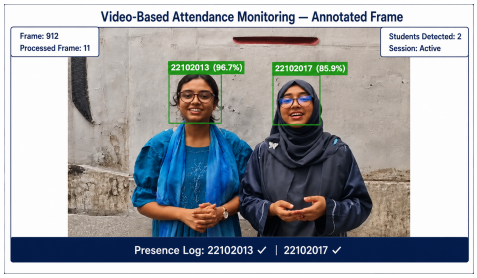


Fig. 8. Representative annotated frame from the video evaluation, showing face-detection boxes, roll-number labels with recognition confidence, and the running presence log.

Fig. 8 depicts a representative frame; the displayed confidence is for that single frame, whereas the final decision uses the number of frames in which each student is detected. Fig. 9 shows the per-student presence percentages against the 75% threshold: two students are classified as Present and eight as Absent according to their frame detection rates. Student 22102029, detected in only 2 frames, illustrates the purpose of the threshold, since a brief appearance does not qualify as session-level presence.

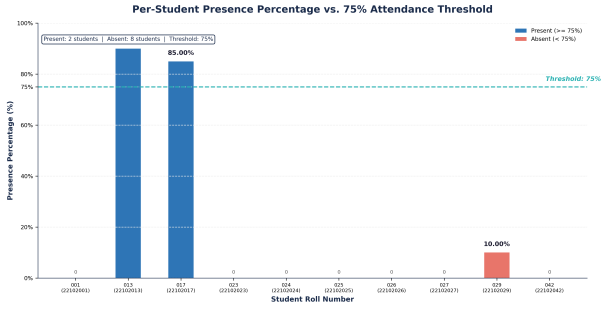


Fig. 9. Per-student presence percentage over the 20 sampled frames, compared with the 75% attendance threshold (dashed line). Students above the line are marked Present.

D. Discussion

The results show that fingerprint verification at entry and repeated visual recognition during the session provide complementary attendance evidence. The high accuracy suggests

that FaceNet embeddings with a simple SVM are sufficient for a small enrolled population, so recognition is not the limiting stage. Basing the decision on multiple frames makes it less sensitive to an isolated missed detection and gives teachers an interpretable presence percentage. The study is limited by its 10-student dataset, 20 sampled frames, and the lack of a separate evaluation of the fingerprint component.

VI. Conclusion

This paper presented a hybrid attendance monitoring framework that combines fingerprint-based authentication with FaceNet-SVM-based classroom monitoring. The system verifies student identity at the classroom entry point and evaluates continued presence from sampled video frames, with the final attendance determined using a 75% presence threshold. Evaluation on a 10-student prototype dataset containing 531 processed face images achieved 98.13% face-recognition accuracy and demonstrated the feasibility of distinguishing students with high session presence from those with low or no detected presence. Compared with single-event attendance approaches, the proposed system provides a more automated and presence-based attendance decision.

The current evaluation is limited by the small dataset, controlled classroom setting and use of sampled frames which requires further evaluation for real-time deployment.Future research will test and enhance the system in larger and more diverse classroom settings.Other improvements include multi-camera testing, real-time processing, secure storage of biometric information and access-controlled attendance management that will support the evaluation of the framework for large-scale institutional deployment.

VII. Data and Code Availability

The source code, dataset and experimental scripts for face-image preprocessing, FaceNet embedding extraction, SVM-based recognition and video-based attendance calculation are publicly available at https://github.com/mehedinaeem/Hybrid_Attendance_System.

Acknowledgment—The authors gratefully acknowledge the Research and Extension Center, Jatiya Kabi Kazi Nazrul Islam University (JKKNIU), Trishal, Mymensingh, Bangladesh, for providing financial support for this research. The authors sincerely appreciate the Center's valuable support and continued encouragement of research and academic excellence.

References

[1] A. Z. Lari, F. Khan, A. Ahmad, A. A. Raza, and M. Suaib, "An efficient attendance management system for college environments using machine learning facial recognition technology," *International Journal of Innovative Research in Computer Science and Technology*, vol. 13, no. 3, pp. 8–12, 2025.

[2] E. T. Olawole, O. F. Adebayo, J. A. Adeshina, F. O. Ariba, and R. A. Ajayi, "Development of an internet of things-based fingerprint biometric attendance system," *LAUTECH Journal of Engineering and Technology*, vol. 19, no. 3, pp. 180–187, 2025.

[3] N. T. Son *et al.*, "Implementing cctv-based attendance taking support system using deep face recognition: A case study at fpt polytechnic college," *Symmetry*, vol. 12, no. 2, p. 307, 2020.

[4] C. Sun, P. Wen, S. Zhang, X. Wu, J. Zhang, and H. Gong, "A face detector with adaptive feature fusion in classroom environment," *Electronics*, vol. 12, no. 7, p. 1738, 2023.

[5] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 815–823.

[6] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[7] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.

[8] K. B. S. Saul, J. B. Saul, and K. T. Soberano, "Utilizing convolutional neural networks for fingerprint-based attendance monitoring," *International Journal of Multidisciplinary Research and Analysis*, vol. 6, no. 4, pp. 1343–1351, 2023.

[9] M. P. Gaikar, M. S. Patil, M. Y. Bhoir, and K. Kapde, "Fingerprint based attendance system," *International Journal of Advanced Research in Science, Communication and Technology*, vol. 5, no. 7, pp. 210–215, 2025.

[10] N. A. Ismail *et al.*, "Web-based university classroom attendance system based on deep learning face recognition," *KSII Transactions on Internet and Information Systems*, vol. 16, no. 2, pp. 503–523, 2022.

[11] J. Patel, S. Gandhi, V. Katheriya, P. Pataliya, and A. Majumdar, "Enhancing classroom attendance systems with face recognition through cctv using deep learning," *Procedia Computer Science*, vol. 258, pp. 3031–3041, 2025.

[12] R. Bankar, N. Bargat, I. Hanmante, and H. Dakore, "Face recognition using facenet deep learning network for attendance system," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 2022.

[13] U. Hengaju, N. Adhikari, A. Aryal, O. K. Raut, and S. Dahal, "Automated face recognition based attendance system using retinaface and facenet," *Advancement in Image Processing and Pattern Recognition*, vol. 5, no. 3, 2022.

[14] V. Khadatkar, S. Bagadkar, and N. Dhande, "Face recognition based attendance management system using hog feature extraction and svm classifier," *International Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 3, pp. 2079–2083, 2022.

[15] M. A. Hosen, S. H. Moz, M. M. H. Khalid, S. S. Kabir, and S. M. Galib, "Face recognition-based attendance system with anti-spoofing, system alert, and email automation," *Radioelectronic and Computer Systems*, no. 2, pp. 115–127, 2023.

[16] V. Rohini, M. Sobhana, and C. S. Chowdary, "Attendance monitoring system design based on face segmentation and recognition," *Recent Patents on Engineering*, 2023.

[17] L. A. Rollon, A. Caye, J. Lusing, and B. Aquino, "Attendance system based on facial recognition using multi-task convolutional neural network," *World Journal of Advanced Research and Reviews*, vol. 18, no. 3, pp. 1–6, 2023.

[18] A. Rao, "Attenface: A real time attendance system using face recognition," in *Proc. 2022 IEEE 6th Conf. Information and Communication Technology (CICT)*, 2022, pp. 1–5.

[19] J. D. D. C. Pires, B. F. S. Neto, R. S. M. Fragoso, M. C. Oliveira, and M. M. Ribeiro, "End-to-end automated student attendance recording system using surveillance camera," *Revista Brasileira de Informatica na Educacao*, vol. 33, pp. 371–393, 2025.

[20] K. Sangeetha *et al.*, "Real-time classroom student monitoring framework using ESP32-Cam and CNN-based face recognition," in *2026 5th Int. Conf. Communication, Computing and Electronics Systems (ICCCES)*, 2026.

[21] B. Kasi *et al.*, "CCTV-based automated attendance system using deep learning for real-time monitoring," in *2026 6th Int. Conf. Inventive Computation and Information Technologies (ICICIT)*, 2026.

[22] B. Lavanya *et al.*, "Smart attendance and emotion monitoring system using face detection with CCTV integration," in *2026 7th Int. Conf. Mobile Computing and Sustainable Informatics (ICMCSI)*, 2026.

[23] G. A. Swamy *et al.*, "Multi-face recognition-based attendance system using image and video inputs," in *2026 IEEE 18th Int. Conf. Computational Intelligence and Communication Networks (CICN)*, 2026.

[24] Y. I. Shehu, A. Ruiz-Garcia, V. Palade, and A. James, "Sokoto coventry fingerprint dataset," https://arxiv.org/abs/1807.10609, 2018.