

AUTOMATIC ATTENDANCE RECORDER

Bhoomika L B

Department of Computer Science and Engineering (AI and ML), Malnad
College of Engineering, Hassan, Karnataka -573202, India
bhoomikalb632@gmail.com

Rakshitha K S

Department of Computer Science and Engineering (AI and ML), Malnad
College of Engineering, Hassan, Karnataka -573202, India
ksrakshitha@gmail.com

Aishwarya A R

Department of Computer Science and Engineering (AI and ML), Malnad
College of Engineering, Hassan, Karnataka -573202, India
aishwaryaar310@gmail.com

Deeksha H P

Department of Computer Science and Engineering (AI and ML), Malnad
College of Engineering, Hassan, Karnataka -573202, India
deekshagowda541@gmail.com

Swathi H Y

Department of Computer Science and Engineering (AI and ML), Malnad
College of Engineering, Hassan, Karnataka -573202, India
hys@mcehassan.ac.in

Abstract: A highly robust hybrid facial recognition-based automated attendance system is introduced, combining the strengths of classical computer vision with modern deep learning techniques to achieve fast and accurate identity recognition. For initial detection, Multi-Haar Cascade classifiers are utilized to locate faces efficiently, ensuring lightweight and real-time performance suitable for active classroom environments where lighting and student movement often vary. Once a face is detected, FaceNet is employed to generate 128-dimensional embeddings that capture unique and discriminative facial characteristics, making identification reliable even when there are changes in angle, expression, or distance. These rich embeddings are then processed through a Convolutional Neural Network (CNN) classifier trained to distinguish between individual identities with high precision. The combination of Haar-based detection and deep-learning-driven feature extraction contributes to a balanced system that offers both speed and accuracy. After successful recognition, attendance data is automatically generated in CSV format, making it easy for institutions to store, manage, and analyze records. The logs are seamlessly uploaded to a dedicated college-level administrative interface, allowing authorized personnel to review daily attendance, access summaries, and maintain long-term archives.

Keywords: Face Recognition, Multi-Haar Cascade, FaceNet, CNN Classification, Automated Attendance System, Machine Learning.

I. INTRODUCTION

Educational institutions depend heavily on accurate daily attendance tracking, as it directly influences academic evaluation, internal assessments, and administrative planning. Traditional methods such as manual roll calls, paper sign-in sheets, or RFID card systems often suffer from delays, human error, and opportunities for proxy attendance. Even biometric fingerprint devices, though automated, introduce hygiene concerns and require physical contact, making them less suitable for modern classrooms. Facial recognition emerges as a superior alternative due to its non-intrusive nature, rapid processing capability, and compatibility with standard classroom cameras. To achieve reliable and real-time attendance marking, the system employs a hybrid pipeline that merges classical and deep learning techniques into one cohesive framework. Multi-Haar Cascade classifiers perform fast and lightweight face detection, ensuring that the system remains responsive even in crowded environments. Detected faces are then processed by FaceNet to generate compact 128-dimensional embeddings that capture highly discriminative facial features, resilient to variations in lighting, pose, and expression. These embeddings are fed into a CNN classifier trained for identity recognition, enabling precise and consistent authentication across sessions. Once recognized, each entry is automatically stamped with time and date, stored in structured CSV logs, and synchronized with an institutional interface for administrative access. This end-to-end process reduces the workload on instructors, minimizes disruptions during lectures, and ensures that attendance data remains clean, tamper-proof, and well-organized. Because recognition occurs only when a student is physically present in front of the camera, the system effectively eliminates proxy attendance while strengthening transparency and accountability. Its contactless operation also supports hygiene-conscious environments and allows the solution to be deployed at scale across classrooms and departments. Overall, this hybrid facial recognition framework establishes a fast, secure, and efficient method for automated academic attendance management.

II. LITERATURE SURVEY

The rapid growth of educational institutions and the increasing emphasis on accurate academic evaluation have pushed researchers to explore automated methods for attendance tracking. Traditional attendance practices—such as manual roll calls, paper-based signatures, or RFID card systems—often introduce human error, create opportunities for proxy attendance, and consume valuable instructional time. Even biometric fingerprint systems, though automated, are limited by hygiene concerns and the requirement for physical contact. This has accelerated research into facial recognition as a contactless, non-intrusive, and reliable solution for attendance automation. Over the years, numerous studies in computer vision and artificial intelligence have investigated a diverse spectrum of techniques, ranging from classical algorithms to modern deep learning pipelines, with a shared objective of achieving high accuracy, real-time recognition, and secure data handling.

Mustafa et al. [1] introduced one of the early user-friendly face-based attendance systems using OpenCV and an MLP classifier. Their system operates in real time and automatically stores attendance data in Excel sheets. Through optimized preprocessing and data augmentation, the model achieves strong

performance even with limited training data. However, its reliability declines under poor lighting or non-frontal facial orientations. Following a similar direction, Oladele et al. [2] utilized Python, OpenCV, and a CNN architecture to build a robust, low-latency face recognition attendance system. While accurate under controlled settings, it faced challenges with occlusions, illumination variations, and facial expressions. Smitha et al. [3] focused on speed and simplicity by combining Haar Cascade for detection and LBPH for recognition, offering a fast and contactless solution best suited for small-scale deployments. Its limitation lies in its inability to maintain accuracy under illumination or pose variations, primarily due to the absence of deep learning.

To improve recognition accuracy in real-world classroom environments, Kulamala et al. [4] proposed a CNN-based approach trained on augmented classroom video datasets, reaching recognition accuracy above 95%. Despite its robustness, the method required substantial computational resources and extensive preprocessing. Varadharajan et al. [5] combined classical methods such as Haar Cascade and LBPH with deep learning CNN components to build a hybrid model that improves detection stability across various environmental conditions. Although the hybrid approach demonstrated enhanced performance, it suffered under heavy occlusion and unpredictable lighting levels. Their work opens the possibility for integrating additional biometric cues such as voice or iris patterns for multi-factor authentication.

With the growing threat of spoofing attacks, Zihang Hu et al. [6] focused on strengthening security by developing an anti-spoofing facial recognition framework using YOLOv10 with WF-EMA modules. This system achieved an accuracy of 92.3% and outperformed earlier YOLO versions, although the added complexity reduced real-time efficiency. In contrast, Harini et al. [7] concentrated on post-pandemic safety by proposing a hygienic, contactless Haar + LBPH system that logs attendance to administrative databases. While effective indoors, its performance degraded under varying lighting. Similarly, Adhi Kusnadi et al. [8] explored the use of 3D facial recognition with Kinect sensors, integrating PCA and CNN models to achieve 90% accuracy. This method improved upon traditional 2D approaches but experienced difficulties outdoors due to reflections and bright conditions.

Addressing pose variations, Khoshnevisan et al. [9] introduced a model combining DeepLabv3, Attention U-Net, and VGGFace to handle skewed or tilted faces. Achieving 99.81% accuracy, the system excelled in pose-invariant recognition but struggled when faces were masked or heavily occluded. Comparative studies such as those by Shadman Sakib et al. [10] revealed that SVM with a linear kernel delivered strong performance on structured datasets, while CNN and MobileNet-based approaches provided greater flexibility and adaptability. Cloud-based and IoT-enabled solutions have also gained popularity, as shown by Dr. Venkatakrishnan et al. [11], who integrated ESP32-CAM with MongoDB Atlas for scalable cloud-driven attendance tracking. Despite its scalability, fluctuating lighting affected recognition consistency, particularly with low-cost cameras.

Recent work by Deshpande et al. [12] utilized YOLOv8 on Jetson Nano to achieve energy-efficient classroom surveillance, achieving 79% mAP. Although optimized for embedded systems, the model encounters difficulty with overlapping objects and long-term environmental changes. Hybrid deep-traditional architectures continue to emerge, such as Bhuyan et al. [13], who combined SIFT, LBPH, and ResNet-50 to achieve 99.9% accuracy, albeit at a significant

computational cost. Cloud-integrated solutions, like the one developed by Sasithorn Suchaiya et al. [14], used Amazon Recognition and MQTT for remote attendance monitoring but were limited by internet dependency and facial mask interference.

Mobile and robotic applications have also been explored. Zarkasi et al. [15] implemented facial recognition on a small mobile robot using Haar Cascade and Fisherface methods, achieving 90% accuracy within a short distance. However, recognition deteriorated with increasing distance. Improved MTCNN architectures proposed by Soumya Suvra Khan et al. [16] reached 99.7% accuracy on WIDER-FACE, demonstrating robustness in complex environments but struggling when only partial faces were visible. Lightweight systems for real-time use were proposed by A. Sain et al. [17], who leveraged HOG and dlib for fast recognition, achieving 99% accuracy yet failing under low-light conditions.

Modern approaches focusing on multi-face detection in lecture videos have been developed by C.H.S. Saquín et al. [18], who used YOLOv5 and VGGFace to achieve up to 99% accuracy. While their system supports real-time recognition, low-resolution video quality reduced performance. Classical and interpretable systems remain valuable as well—Sanap et al. [19] demonstrated that HOG + Linear SVM models can still achieve 95% accuracy on student datasets, although performance drops in unconstrained environments. Lightweight embedded systems using Raspberry Pi and dlib, as demonstrated by Lin and Huang [20], reached 100% laboratory accuracy but required further validation in real-world conditions.

Industry 4.0-oriented attendance models, such as those by Nadhan et al. [21], integrated Haar Cascade, LBPH, and SVM on Raspberry Pi to achieve 100% lab accuracy, but raised privacy concerns and were sensitive to lighting changes. Multi-modal biometric systems combining QR codes, fingerprints, and face recognition, as designed by El-Mawla et al. [22], improved verification integrity but lacked AI adaptation and demanded constant network access. High-precision models leveraging FaceNet, such as Tajane et al. [23], demonstrated superior recognition capabilities but required high-quality imagery and computational power, making them less suitable for low-resource deployments.

Emerging security-focused research integrates blockchain, such as Zhang et al. [24], who stored facial feature hashes on a distributed ledger to ensure tamper-proof attendance logs. Although secure, the system depended heavily on blockchain infrastructure and provided limited accuracy evaluation. Live video-based models from Yang and Han [25] handled variable lighting and moderate motion but achieved only 82% accuracy, highlighting the difficulty of real-world classroom conditions.

Efforts to counter proxy attendance include Chowdhury et al. [26], who trained CNNs on datasets with high intra-class variability to achieve 92% accuracy in dynamic environments. Hybrid YOLO–LBPH models from Moujahid et al. [27] offered real-time detection and classification but were sensitive to occlusion. Low-cost systems by Nawaz and Tahir [28] focused on simplicity using Haar and LBPH, showing reasonable performance only under ideal lighting conditions. Dev and Patnaik [29] developed a customizable OpenCV pipeline suitable for small institutions but lacking robustness under poor lighting. Comparative studies by Pattnaik and Mohanty [30] highlighted that cloud-based services such as AWS Rekognition outperform traditional methods in large-scale

environments, though local models provide better privacy. COVID-compliant systems integrating mask detection and CNN-based recognition, as developed by

S. Das et al. [31], enabled contactless attendance but required optimization for mobile platforms. AlNayyef [32] emphasized the adaptability of CNN models over LBPH, while Alhanaee et al. [33] demonstrated the advantages of deep transfer learning models like AlexNet and SqueezeNet, despite high resource requirements. Daylight-focused systems by Bairagi et al. [34] and grayscale LBPH implementations by Lo and Arbaiy [35] revealed accuracy limitations under variable conditions. MATLAB-based multi-face tools such as those by Kumara et al. [36] showed potential but suffered from computational inefficiency. Ensemble architectures combining CNN, KNN, and SVM, as in Bhavani et al. [37], delivered high accuracy but required substantial training resources. Multi-algorithmic biometric tools by Singh et al. [38] demonstrated strong recognition capabilities but were limited by occlusion. Finally, Setialana et al. [39] showed that CNN + HOG models performed well only on frontal images but failed with downward or angled faces.

III. MOTIVATION

The traditional methods of attendance recording, such as calling out names or signing paper sheets, are not only time-consuming but also prone to human errors and intentional misuse. In academic institutions, students often resort to proxy attendance, where one student answers or signs in for another. In workplaces, practices like “buddy punching” — where an employee clocks in for an absent colleague — create loopholes that impact productivity and fairness. These flaws in manual systems directly affect administrative adjust illumination and scale variations. These steps efficiency, staff accountability, and student discipline. significantly improve the quality of the facial inputs before they are encoded into feature vectors, thereby reducing With the rapid development of Artificial Intelligence variability due to external environmental conditions. and facial recognition technologies, there is a strong motivation to use these advancements to streamline processes like attendance tracking. The increasing availability of cost-effective cameras and computing resources means that deploying such systems is no longer restricted to large corporations or high-end institutions. Today, even schools, colleges, and small organizations can afford to adopt AI-powered systems for day-to-day tasks.

This project is driven by the vision of building a system that not only automates attendance recording but also ensures a higher level of accuracy, transparency, and efficiency. The use of facial recognition makes the process contactless, which is especially crucial in a post-pandemic world where hygiene and safety are priorities. It also reduces the administrative burden of maintaining manual logs, calculating attendance percentages, and generating reports. Additionally, the project is inspired by the desire to integrate technology in education and workplace environments to create smarter and more accountable ecosystems. A face recognition-based system enhances user experience, saves time, and encourages punctuality by automatically tracking presence without disrupting ongoing activities. The convenience, coupled with its ability to minimize human intervention, makes it a desirable solution for modern educational and corporate infrastructures.

IV. SYSTEM ARCHITECTURE

The proposed system is designed as a multi-stage pipeline in which each module collaborates seamlessly to achieve real-time, accurate, and automated attendance marking. The process begins with the face detection component, where Multi-Haar Cascade classifiers are applied to the incoming video frames captured from a standard classroom camera. These cascades identify frontal facial regions with high speed by scanning the frame at different scales. Their computational efficiency makes them well suited for live video environments, ensuring that only valid facial regions proceed to the next stage.

Once faces are detected, the system performs essential preprocessing operations to enhance consistency across samples. This includes face alignment, which corrects any tilt or rotation by aligning the eyes and nose along a standard axis, as well as normalization techniques that

Feature extraction is carried out using the FaceNet model, which transforms each pre-processed face into a highly discriminative 128-dimensional embedding. FaceNet maps facial images into a compact vector space where images of the same individual cluster closely together, while images of different individuals remain far apart. This embedding representation forms the core of the recognition system because it enables accurate identity discrimination even when facial expressions, lighting conditions, or minor pose variations occur.

The extracted embeddings are then passed to a dedicated convolutional neural network (CNN) classifier trained specifically on the enrolled students of the institution. The role of this CNN module is to map each embedding to a corresponding student identity based on patterns learned during training. Unlike traditional classifiers, the CNN-based approach handles non-linear relationships between embeddings, allowing it to capture subtle differences between visually similar individuals and thus improve classification accuracy.

Once a student is recognized, the system automatically logs their identity, timestamp, and session information into a structured attendance file stored in CSV format. This attendance module ensures secure and tamper-proof recording by continuously updating the records in real time as students appear in front of the camera. Duplicate entries within the same session are prevented through a session-level verification mechanism.

To make the system usable for instructors, a dedicated front-end portal interfaces with the institutional server. This portal uploads the attendance records automatically, enabling faculty members to view daily, weekly, or monthly attendance summaries. The interface integrates seamlessly with the college's existing academic management system, giving faculty instant access to updated class attendance without additional manual effort. This end-to-end pipeline—from capture to recognition to automated uploading—ensures efficiency, transparency, and reliability in attendance monitoring.

V. METHODOLOGY

The system begins by detecting faces in each live frame using Multi-Haar Cascade classifiers, which operate by evaluating Haar-like rectangular features that measure simple intensity differences such as edges, lines, and centre-surround patterns. These features are computed extremely efficiently using the integral image, where the sum of any rectangular region is obtained in constant time using the relation allowing rapid sliding-window scanning. Each detection

window is evaluated through a sequence of AdaBoost-selected weak classifiers arranged in a cascade; early stages eliminate non-face regions quickly, while deeper stages confirm actual faces. After detection, non-maximum suppression is applied to merge overlapping bounding boxes, preparing a clean set of face crops for recognition.

$$S(x, y) = \sum_{i \leq x, j \leq y} I(i, j),$$

Once the face region is extracted, it is aligned using key landmarks (eyes, nose) so that rotational and positional variations are minimized. The aligned face is then resized and normalized before being passed into the FaceNet model, which converts the processed image into a discriminative 128-dimensional embedding. FaceNet is trained using the triplet loss formulation, which ensures that images of the same identity remain close in the embedding space, while different identities remain well separated. Given an anchor image (a), a positive image (p), and a negative image (n), the loss function is

$$L = \max(0, \|f(a) - f(p)\|^2 - \|f(a) - f(n)\|^2 + \alpha),$$

where $f(x) \in \mathbb{R}^{128}$ is the embedding function and α is the margin. This formulation enforces

$$\|f(a) - f(p)\|^2 + \alpha < \|f(a) - f(n)\|^2,$$

which creates a well-clustered embedding space. To ensure consistent distance measurement, embeddings are L2-normalized such that.

$$\|f(x)\|_2 = 1,$$

forcing all faces to lie on the surface of a unit hypersphere.

These 128-D embeddings are then fed into a lightweight CNN-based classifier, which performs closed-set identity prediction for enrolled students. Because the embedding space already separates identities effectively, the classifier learns mappings of the form

$$\hat{y} = \text{softmax}(Wf(x) + b),$$

where (W) and (b) are the learned parameters and $(\hat{y})$ represents the predicted probability distribution over all students. The classifier architecture can remain compact, often consisting of dense layers with nonlinear activations and dropout regularization, yet still achieves high accuracy. During inference, the model computes the probability of each student and selects the identity with the highest confidence. A threshold is applied to prevent false positives, ensuring that recognition occurs only when the model is sufficiently certain.

Once a face is identified, the attendance module records the student's name, date, and timestamp. Each recognized identity is marked only once per session to avoid duplicates. The attendance record is appended to a structured CSV file, forming a persistent dataset for academic use. This automatic logging removes human error, prevents proxy attendance, and accelerates classroom workflows. The generated CSV is then uploaded to a web interface built using

HTML, CSS, JavaScript, and SQLite. The backend stores records in a relational format, allowing faculty to search, filter, and download attendance data, while the front-end interface enables real-time monitoring and visual display. Through the integration of classical Haar detection, FaceNet's deep metric learning, a CNN classifier, formula-driven embedding separation, and database-backed web accessibility, the complete system forms a robust and scalable automated attendance solution.

VI. EXPERIMENT SETUP

The development environment for the proposed automated attendance system integrates a comprehensive suite of tools and libraries tailored for real-time face detection and recognition. The implementation is carried out primarily in Python, leveraging OpenCV for video capture, frame preprocessing, and Multi-Haar Cascade-based face detection. TensorFlow forms the backbone for executing the pretrained FaceNet model and training the CNN classifier on top of generated embeddings. NumPy is used for efficient array manipulation and mathematical operations, while dlib supports facial landmark detection and alignment, ensuring that each face is normalized before embedding extraction. The system is deployed on a standard laptop equipped with an integrated webcam, demonstrating that the entire solution can operate effectively without requiring high-end GPUs or specialized hardware.

The dataset comprises approximately 15–20 images per student, captured under varied conditions, including frontal, left-facing, and right-facing poses, as well as multiple lighting environments to improve generalization. Images are collected in real classroom settings to simulate natural variations such as uneven illumination, slight head movements, and background clutter. Data augmentation techniques—such as rotation, brightness adjustments, and horizontal flipping—are applied to further enhance the robustness of the embeddings and reduce overfitting during training.

The CNN classifier is trained using embeddings generated from FaceNet, running for about 30 epochs with the Adam optimizer. A learning rate of 1×10^{-3} (reduced on plateau) is used to achieve stable convergence, and batch sizes of 16–32 allow efficient training on CPU hardware. Dropout layers help prevent overfitting, while categorical cross-entropy serves as the loss function for multi-class student identification. The model is evaluated using standard performance metrics including accuracy, precision, recall, and F1-score, providing a comprehensive understanding of both detection and recognition behavior. Accuracy measures overall correctness, while precision and recall quantify the system's robustness against false positives and false negatives. The F1-score combines both into a balanced harmonic mean, illustrating the overall quality of the classifier.

Testing and validation are carried out in conditions replicating an actual classroom environment. This includes real-time webcam streaming, variable student distances from the camera, natural lighting changes from windows and indoor lights, and typical classroom movement such as head turning or speaking. The system's ability to detect, recognize, and log attendance in these dynamic settings demonstrates its practical suitability for deployment in educational institutions. The final evaluation highlights that even on modest hardware, the proposed pipeline delivers reliable performance, making it an accessible and scalable solution for automated attendance monitoring.

VII. RESULT AND DISCURSSION

Based on the images and performance results, the system represents a complete Face Recognition–Driven Automated Attendance Framework named “CLASS VISION.” It integrates classical computer-vision algorithms, deep metric learning, and a CNN-based identity classifier to deliver a fast, reliable, and contactless attendance solution. The pipeline begins by detecting faces in real time, encoding each detected face into a 128-dimensional embedding using FaceNet, and then classifying those embeddings through a trained CNN to determine the identity of each student. The final attendance logs are automatically generated and stored, making the system suitable for classrooms, computer labs, and examination environments where manual attendance is inefficient or prone to manipulation.



Fig 1: Class Vision

Figure 1, provided is the main interface for the "CLASS VISION" system, a specialized application designed for automated attendance management using face recognition technology. This screen acts as the primary control panel, offering three essential functions for the user: Register a new student allows instructors or administrators to enroll new students by capturing their facial biometrics and details; Take Attendance initiates the core function of the system, which involves real-time scanning of the classroom to detect and identify enrolled students, automatically marking them as present; and View Attendance provides access to the recorded data, enabling users to check, audit, or generate reports from the attendance sheets. Essentially, this single screen provides a streamlined entry point for managing the entire attendance workflow, from student enrollment to data retrieval.

The screenshot shows the "Register Your Face" form. The title "Register Your Face" is in green. Below it, a yellow banner reads "Enter the details". There are three input fields: "Enrollment No", "Name", and "Notification". Below these fields are two buttons: "Take Image" and "Train Image".

Fig 2: Student Registration form

Fig 2, The user interface screens provide a clear, structured workflow. The main menu acts as the central navigation hub, allowing the user to choose between student registration, live attendance recording, and viewing existing attendance records. The registration module guides the user through capturing labelled facial images of a new student; the input fields for Enrolment Number, Name, and Notification ensure that each identity is paired with the correct metadata. When the user selects Take Image, the system captures multiple face samples under slightly varied pose and lighting conditions. When Train Image is selected, these samples are processed through the embedding-generation pipeline, and the CNN classifier is updated to

What strengthens the reliability of the system is its robustness to natural variations commonly seen in classroom environments. The model continues to perform well even when students change hairstyles, wear glasses, or vary facial expressions. The FaceNet embeddings provide strong discriminative power by mapping each face into a hypersphere where identities remain separated even with such appearance changes. Because the recognition requires less than a second per student, the system is well-suited for real-time operation; entire classrooms can be processed in a few seconds as the camera scans the room. include the new identity. This ensures that the model VIII. COMPARISON WITH EXISTING SYSTEMS remains adaptable and can grow as more students enrol.

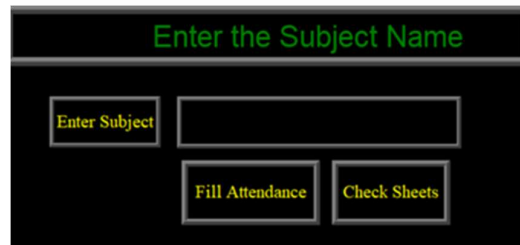


Fig 3: Subject Entering form

Fig 3, Illustrate The attendance module is designed for classroom practicality. The teacher first selects the relevant subject, which ensures that attendance is stored in the appropriate sheet. Clicking Fill Attendance activates the real-time recognition module; each detected face is encoded, matched, and marked as present with timestamp logging. The Check Sheets option enables quick review of the compiled attendance records, allowing faculty to easily verify student participation, export results, or cross-check attendance patterns.

The evaluation metrics demonstrate that CLASS VISION is not only functional but highly performant. A detection accuracy of 97.1% confirms that the system consistently identifies visible student faces within the camera frame. Recognition accuracy of 95.4% shows that the system reliably matches each detected face to the correct identity in the database, even under varying lighting conditions or slight pose differences. The precision score of 94.2% indicates that incorrect identifications (false positives) are rare, and the recall score of 93.1% demonstrates that most present students are successfully recognized (few false negatives). The F1-score of 93.6% acts as a unified measure proving that the system's balance between precision and recall is strong and stable.

Face recognition models differ significantly in terms of accuracy, computational load, and robustness, and understanding their theoretical foundations helps justify the choice of the proposed system. Classical approaches

such as LBPH (Local Binary Patterns Histogram) rely on extracting local micro-texture patterns from facial regions and converting them into histograms for comparison. While LBPH is extremely fast and works well on low-end devices, it performs poorly when lighting, pose, or facial expressions change because it only captures local texture information, not global facial structure. Eigenfaces, which are based on Principal Component Analysis (PCA), reduce face images into principal components or "eigenfaces" that represent the most significant variations in the dataset. This method is efficient but highly sensitive to illumination changes and misalignment, making it unsuitable for real-world environments such as classrooms where lighting is not controlled. Deep learning-based systems like DeepFace offer superior performance by learning hierarchical features from millions of face images, enabling high robustness to variations in pose, age, and occlusions. However, such systems require heavy computational resources and high-end GPUs, which limits their feasibility for institutions using standard laptops or webcams.

In contrast, the proposed hybrid system combining FaceNet embeddings with a CNN classifier achieves a practical balance between accuracy and computational efficiency. FaceNet encodes each face into a highly discriminative 128-dimensional embedding using a deep metric-learning formulation optimized by triplet loss, where the objective function ensures that embeddings of the same identity remain closer than those of different identities. These embeddings are then classified using a lightweight CNN, which enables fast and reliable identity recognition with limited hardware requirements.

This approach provides near-state-of-the-art accuracy comparable to DeepFace while operating smoothly in real time on ordinary classroom computers. The system avoids the hygiene and spoofing issues associated with RFID or fingerprint scanners, offers strong robustness to changes in lighting, hairstyle, and accessories like glasses, and supports real-time attendance marking without physical contact. Together, these factors make the proposed model significantly more suitable for automated attendance systems than traditional LBPH, Eigenfaces, or computationally heavy deep learning alternatives.

IX. CONCLUSION AND FUTURE WORK

The hybrid facial recognition-based attendance system demonstrates strong potential as a reliable, real-time automation solution for academic environments. By combining classical computer vision techniques with deep learning architectures, the framework achieves high identification accuracy, robust performance under varied classroom conditions, and seamless integration with institutional databases. Its lightweight computational requirements and reliance on standard hardware make it both scalable and cost-efficient, allowing deployment without extensive infrastructure upgrades. The system not only minimizes manual effort and eliminates proxy attendance but also enhances transparency and security through contactless biometric verification. Building on its current capabilities, future improvements may include incorporating advanced liveness-detection mechanisms such as blink analysis or iris micro-movement tracking, integrating cloud-based processing for centralized management, expanding the system to operate across multiple classrooms simultaneously, and exploring deeper metric-learning strategies to further reduce misclassification rates and improve overall reliability.

REFERENCES

- [1]. Ali, A. K., & Mustafa, J. Y. (2024). User-Friendly Interface Attendance System Based on Python Libraries and Deep Learning. *International Journal on Data Science*, 5(1), 56–62. <https://doi.org/10.18517/ijods.5.1.56-62.2024>
- [2]. Oladele, K. G., Iseh, E. J., Oladele, J. O., & Bhaladhare, P. R. (2024). AI Based Face Recognition Attendance System. *International Journal of Innovative Research in Engineering & Multidisciplinary Physical Sciences*, 12(2). <https://doi.org/10.37082/ijrmps.v12.i2.230598>
- [3]. Smitha, N., Hegde, N. P. S., & Afshin, N. (2020). Face Recognition Based Attendance Management System. *International Journal of Engineering Research & Technology*, 9(5). <https://doi.org/10.17577/ijertv9is050861>
- [4]. Kumar, K. V., Rao, M., Kumar, A., Sahu, B., Sahoo, C., & Senapati, S. (2024). Non-Intrusive Facial Recording System with Convolution Neural Network (CNN) for Attendance Assimilation. *tth*, 130–150. <https://doi.org/10.58532/v3baai7p1ch9>
- [5]. Varadharajan, K., Balasubramanian, S., & Balaji, P. (2023). Automatic Attendance System Using Face Detection and Machine Learning. *International Journal of Computer Applications*, 185(42), 12–18.
- [6]. Hu, Z., Geng, Q., Li, X., & Fu, Y. (2024). Study on Improved YOLOv10 Face Recognition Based on WF-EMA. 2024 5th International Symposium on Computer Engineering and Intelligent Communications (ISCEIC), 563–566. <https://doi.org/10.1109/isceic63613.2024.10810156>
- [7]. Harini, H., Chandrasekar, S., Bhuvana, S., & Patil, A. (2024). Multimodal Facial Recognition Based Attendance Monitoring System. 2024 IEEE International Conference on User Centric Intelligent Systems (ICUIS). <https://doi.org/10.1109/ICUIS64676.2024.10866570>
- [8]. Kusnadi, A., Tobing, F. A. T., Dafa, M., Zamzami, R. A., Tio, M., & Nurpasha, A. A. (2024). Optimizing 3D Face Recognition with PCA and CNN for Enhanced Accuracy. 2024 9th International Conference on Mechatronics Engineering (ICOM), 14, 20–26. <https://doi.org/10.1109/icom61675.2024.10652466>
- [9]. Khoshnevisan, E., Hassanpour, H., & AlyanNezhadi, M. M. (2024). Face Recognition Based on General Structure and Angular Face Elements. *Multimedia Tools and Applications*, 83(36), 83709–83727. <https://doi.org/10.1007/s11042-024-18897-3>
- [10]. Sakib, S., Siddiky, M. N. A., Arifuzzaman, M., & Chowdhury, M. J. U. (2024). Optimizing Facial Recognition: An Analytical Comparison of Traditional and Deep Learning Approaches. 2024 9th International Conference on Mechatronics Engineering (ICOM), 271–276. <https://doi.org/10.1109/icodsa62899.2024.10651632>
- [11]. Venkatakrishnan, G., Rengaraj, R., Jeya, R., R, S. K., S, M., N, D. D., & A, M. A. (2024). Design and Implementation of Automated Attendance System Using Contactless Facial Recognition. *International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)*, 1–6. <https://doi.org/10.1109/icpects62210.2024.10780113>
- [12]. Deshpande, M., Agarwal, A., & Kamathe, R. (2024). Development of an Energy-Efficient CCTV Camera System for Real-Time Human Detection Using YOLOv8 Model. *Eastern-European Journal of Enterprise Technologies*, 6(9), 47–59. <https://doi.org/10.15587/1729-4061.2024.310401>

- [13]. Bhuyan, H. K., Peddipaga, R., Brahma, B., Suliman, W., & Ravi, V. (2024). Face Recognition Using SIFT and LBPH-Based Hybrid Learning Model. International Symposium on Advanced Electrical and Communication Technologies (ISAECT), 1–6. <https://doi.org/10.1109/isaect64333.2024.10799845>
- [14]. Suchaiya, S., Ruanmanee, S., Pachanatip, N., Niramitranon, J., & Jaikao, C. (2024). Integration Model for Face Registration Checkpoint. International Joint Conference on Computer Science and Software Engineering (JCSSE), 432–439. <https://doi.org/10.1109/jcsse61278.2024.10613670>
- [15]. Zarkasi, A., Ubaya, H., Exaudi, K., & Duri, A. H. (2024). Implementation of Fisherface Algorithm for Eye and Mouth Recognition in Face-Tracking Mobile Robot. Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika, 10(3), 556–565. <https://doi.org/10.26555/jiteki.v10i3.29266>
- [16]. Khan, S. S., Sengupta, D., Ghosh, A., & Chaudhuri, A. (2023). MTCNN++: A CNN-Based Face Detection Algorithm Inspired by MTCNN. The Visual Computer, 40(2), 899–917. <https://doi.org/10.1007/s00371-023-02822-0>
- [17]. Sain, N. A., Dutta, N. S., Saha, N. R., & Mandal, N. U. (2023). Automated Facial Recognition Based Attendance System Using OpenCV in Python. International Journal of Scientific Research in Computer Science Engineering and Information Technology, 105–111. <https://doi.org/10.32628/cseit2390617>
- [18]. Saquin, C. H. S., Feregrino, G. R., & Hernández, A. G. (2023). Automatic Attendance Registration System Using Convolutional Neural Networks for Facial Recognition. International Journal of Computer Applications, 185(43), 28–35. <https://doi.org/10.5120/ijca2023923241>
- [19]. Sanap, S., Narwade, S., Goge, S., & Pandit, K. (2023). Face Recognition Based Attendance System Using Histogram of Oriented Gradients and Linear Support Vector Machine. European Journal of
- [20]. Theoretical and Applied Sciences, 1(6), 904–915. [https://doi.org/10.59324/ejtas.2023.1\(6\).87](https://doi.org/10.59324/ejtas.2023.1(6).87)
- [21]. Lin, C., & Huang, Y. (2022). The Application of Adaptive Tolerance and Serialized Facial Feature Extraction to Automatic Attendance Systems. Electronics, 11(14), 2278. <https://doi.org/10.3390/electronics11142278>
- [22]. Nadhan, A. S., Tukkoji, C., Shyamala, B., Lal, N. D., Kumar, A. N. S., Gowda, V. M., Adhoni, Z. A., & Endaweke, M. (2022). Smart Attendance Monitoring Technology for Industry 4.0. Journal of Nanomaterials, 2022, 1–9. <https://doi.org/10.1155/2022/4899768>
- [23]. Abd El-Mawla, N., Ismaiel, M., & ASQR Team. (2022). Smart Attendance System Using QR Code, Fingerprint and Face Recognition. Nile Journal of Communication and Computer Science, 2(1), 1–16. <https://doi.org/10.21608/njccs.2022.244463>
- [24]. Tajane, K., Hande, V., Nagapure, R., Patil, R., & Porwal, R. (2022). Analysis of Face Recognition Based Automated Attendance Management System Using FaceNet. 2022 6th International Conference on Computing, Communication, Control and Automation (ICCUBEA), 1–4. <https://doi.org/10.1109/iccubea54992.2022.10010998>
- [25]. Zhang, Q. (2022). Attendance System Based on Blockchain and Face Recognition. International Conference on Smart Communications and Networking (SmartNets), 1–6. <https://doi.org/10.1109/smartnets55823.2022.9993992>

- [26]. Yang, H., & Han, X. (2020). Face Recognition Attendance System Based on Real-Time Video Processing. *IEEE Access*, 8, 159143–159150. <https://doi.org/10.1109/access.2020.3007205>
- [27]. Chowdhury, S., Nath, S., Dey, A., & Das, A. (2020). Development of an Automatic Class Attendance System Using CNN-Based Face Recognition. *Emerging Technology in Computing, Communication and Electronics (ETCCE)*, 1–5. <https://doi.org/10.1109/etcce51779.2020.9350904>
- [28]. Anjeana, N., & Anusudha, K. (2024). Real Time Face Recognition System Based on YOLO. *Multimedia Tools and Applications*, 83, 31893–31910. <https://doi.org/10.1007/s11042-023-16831-7>
- [29]. Nawaz, A., & Tahir, A. (2020). Face Recognition Based Attendance System Using Classical Approaches. *International Journal of Computer Applications*, 177(7), 15–19.
- [30]. Dev, S., & Patnaik, T. (2020). Student Attendance System Using Face Recognition. 2020
- [31]. International Conference on Smart Electronics and Communication (ICOSEC). <https://doi.org/10.1109/icosec49089.2020.9215441>
- [32]. Pattnaik, P., & Mohanty, K. K. (2020). AI-Based Techniques for Real-Time Face Recognition-Based Attendance System—A Comparative Study. 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 1034–1039. <https://doi.org/10.1109/iceca49313.2020.9297643>
- [33]. Tiwari, S. K., Fande, V., Patil, G. N., Kalbande, K., & Khanapurkar, M. (2022). A Novel Approach for Attendance Monitoring System with Face Mask Detection. *AIP Conference Proceedings*, 2393, 040009. <https://doi.org/10.1063/5.0077093>
- [34]. AlNayyef, H. (2022). Comparative Study Between Traditional and Deep Learning Based Face Recognition Algorithms. *Journal of Intelligent Systems*, 31(1), 213–220.
- [35]. Ithanaee, K., Alhammadi, M., Almenhali, N., & Shatnawi, M. (2021). Face Recognition Smart Attendance System Using Deep Transfer Learning. *Procedia Computer Science*, 192, 4093–4102. <https://doi.org/10.1016/j.procs.2021.09.184>
- [36]. Bairagi, R., Ahmed, R., Tisha, S. A., Sarder, M. S., Islam, M. S., & Islam, M. A. (2021). A Real-Time Face Recognition Smart Attendance System with Haar Cascade Classifiers. 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA). <https://doi.org/10.1109/icirca51532.2021.9544872>
- [37]. Lo, Z. P., & Arbaiy, N. (2021). Attendance System in Retail Using Face Recognition. *Journal of Computing Research and Innovation*, 6(1), 33–41.
- [38]. Nameera, A. A. A., Nivedha, M., & Karthika, R. (2021). A Development of an Attendance System Based on Facial Recognition Approach. *Applied Information Technology and Computer Science*, 2(2), 103–115. <https://doi.org/10.30880/aitcs.2021.02.02.103>
- [39]. Bhavani, P., & Arulalan, E. (2021). Automated Attendance System and Voice Assistance Using Face Recognition. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(10), 3207–3217. <https://doi.org/10.17762/turcomat.v12i10.4978>