Journal of Research Technology and Engineering
www.jrte.org

A Review of Deep Learning-Based Automated Biometric Attendance Monitoring Systems for Educational Institutions: Algorithms, Architectures, and Deployment Challenges

*Rashmi D. Jayasekara, Ridmi D. Jayasekara, W. S. C. Rodrigo, Udara. S. P. R. Arachchige, R. A. Prabhath Buddhika, Mohamed Sapraz

Department of Electrical, Electronic & Systems Engineering, NSBM Green University

*rdjayasekara@nsbm.ac.lk

Received: 18 February 2026; Revised: 28 February 2026; Accepted: 28 March 2026; Available online: 10 April 2026

Abstract - Automated biometric attendance monitoring has emerged as a critical research domain within educational technology, driven by the persistent limitations of conventional attendance methods such as manual roll calls, sign-in sheets, and card-based systems. Recent advances in deep learning and computer vision have substantially transformed this field, enabling the development of robust, contactless, and scalable face recognition systems. This review critically examines the evolution of automated biometric attendance monitoring systems across 34 primary studies, tracing the progression from classical handcrafted feature extraction methods, including Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), and Eigenface-based approaches, to contemporary deep convolutional neural network architectures and deep metric learning frameworks such as DeepFace, FaceNet, VGGFace, OpenFace, and ResNet-based embedding models. Particular attention is given to algorithmic complexity, high-dimensional feature representation, embedding optimization, threshold-based open-set recognition, and multimodal biometric fusion strategies. System-level deployment challenges are critically evaluated, including illumination variability, pose and occlusion sensitivity, computational latency in high-density classroom environments, privacy-preserving learning frameworks, fairness considerations, and anti-spoofing mechanisms. Emerging research directions, including edge intelligence, federated learning, adaptive threshold optimization, and learning management system integration, are analyzed as future pathways for institution-scale deployment. By synthesizing both foundational and recent literature, this review establishes a comprehensive analytical framework for understanding the algorithmic evolution, deployment feasibility, and future research potential of biometric attendance monitoring systems in smart educational infrastructures.

Index Terms - Automated attendance systems, biometric authentication, convolutional neural networks, deep learning, deep metric learning, face detection, face recognition, feature extraction, privacy-preserving biometrics, transfer learning

1. Introduction

The accurate recording of student attendance is a fundamental component of institutional administration and academic accountability in universities and higher education institutions. Attendance data directly influences grading, compliance with regulatory requirements, and assessment of student engagement [1], [2]. Attendance records are increasingly utilized as institutional indicators for academic risk assessment, student retention analysis, and evaluation of learning engagement patterns across academic semesters. Despite this criticality, most universities continue to rely on methods that are demonstrably inefficient and susceptible to manipulation. Manual roll calls require substantial educator time, disrupt lecture flow, and are readily circumvented by proxy responses. Paper-based sign-in sheets, QR code scanning, and RFID card readers similarly fail to prevent fraudulent entries, as students can sign on behalf of absent peers, share cards, or forward QR codes to individuals outside the classroom [3], [4].

Biometric attendance systems represent a significant advancement over these credential-based approaches. By anchoring identification to physiological traits intrinsically bound to the individual, biometric systems substantially reduce the possibility of impersonation or proxy attendance [5]. Early biometric deployments in educational settings included fingerprint recognition and iris scanning. While these systems improved security over card-based methods, they introduced new operational challenges: contact-based acquisition raised hygiene concerns, particularly during the COVID-19 pandemic; sequential individual scanning created bottlenecks in large lecture cohorts; and hardware sensitivity to moisture, skin damage, or environmental variation degraded recognition reliability [6], [7] .

Face recognition has since emerged as the most operationally suitable biometric modality for educational attendance monitoring, offering contactless operation, passive identification without subject cooperation, and compatibility with widely available camera infrastructure [8]. A face recognition-based attendance system captures live video or image frames, detects faces within the frame, extracts discriminative facial features, and matches them against an enrolled database to mark attendance automatically [9]. This process eliminates manual intervention and enables simultaneous identification of multiple individuals within a single frame, offering significant scalability advantages in lecture hall environments.

The field has been profoundly transformed by the advent of deep learning, particularly convolutional neural networks (CNNs), which have superseded classical handcrafted feature approaches in accuracy, robustness, and generalizability. Landmark systems, including DeepFace [10], FaceNet [11], VGGFace [12], and OpenFace have demonstrated face verification accuracies exceeding 97% on benchmark datasets, approaching human-level performance. These advances have enabled practical deployment of face recognition in resource-constrained environments using libraries such as dlib and OpenCV, which provide accessible implementations of HOG-based face detection, CNN-based verification, and deep metric learning for face embedding and comparison [13], [14]. Nevertheless, the performance advantages reported under benchmark datasets do not always directly translate into operational classroom environments, where unconstrained lighting, pose variations, partial occlusions, and heterogeneous image acquisition conditions continue to present substantial deployment challenges.

Despite these advances, significant challenges remain in translating laboratory performance to real-world deployment. Variability in illumination conditions, non-frontal facial orientations, partial occlusion by accessories such as masks or glasses, and limited training dataset size continue to degrade accuracy in practical settings [15], [16]. Furthermore, the collection, storage, and processing of biometric data raise substantial ethical and regulatory concerns, particularly regarding compliance with data protection legislation such as the General Data Protection Regulation (GDPR) and institutional privacy policies [17]. Scalability to large student cohorts, integration with learning management systems (LMS), and real-time processing requirements impose additional computational and infrastructural demands.

This review surveys existing deep learning-based face recognition approaches applied to automated attendance monitoring, organizing the literature into four principal categories: (1) handcrafted feature-based methods employing descriptors such as Histogram of Oriented Gradients (HOG) with support vector machine

(SVM) classifiers; (2) deep convolutional neural network approaches leveraging pre-trained architectures; (3) end-to-end deep metric learning systems employing facial embedding and distance-based comparison; and (4) multi-modal biometric systems combining face recognition with complementary modalities. The review further addresses deployment architectures, system-level challenges, and identifies research gaps to guide future development in this domain. In addition to summarizing existing approaches, this review provides a critical comparison of methodological strengths, practical deployment limitations, and suitability for real-world educational environments. Particular attention is given to issues of scalability, privacy compliance, and robustness under unconstrained classroom conditions.

2. Overview of Biometric Modalities for Attendance Monitoring

Attendance monitoring systems in educational institutions have evolved through several generations of technological approaches, each representing a trade-off between convenience, accuracy, cost, and susceptibility to circumvention. A systematic understanding of the biometric landscape is essential for contextualizing the emergence of face recognition as the dominant modality.

A. Non-Biometric Attendance Methods

Traditional non-biometric methods include manual roll calls, paper sign-in sheets, QR code scanning, and RFID/NFC card systems [3]. Manual roll calls and sign-in sheets rely entirely on verbal or written student responses and are therefore completely vulnerable to proxy attendance. QR code methods improve operational speed but remain susceptible to code sharing. RFID systems provide electronic records and reduced processing time, but card sharing is trivially achievable, and hardware malfunctions present reliability concerns [4], [18]. These methods share a common deficiency: they authenticate a credential or behavior rather than the individual, rendering them fundamentally inadequate for secure attendance verification.

B. Contact-Based Biometric Methods

Fingerprint recognition was among the first biometric technologies deployed in educational attendance systems. Individual fingerprint patterns provide high uniqueness and are effectively impossible to share. However, contact-based acquisition creates hygiene concerns, requires each student to individually interact with the scanner, creating sequential bottlenecks, and performance degrades with wet, dirty, or injured fingers [6]. Iris recognition offers comparable uniqueness with high accuracy, but requires controlled acquisition conditions and specialized hardware, limiting cost-effective large-scale deployment [7]. Palm vein recognition, while highly accurate and hygiene-compliant given its use of subcutaneous vascular patterns, requires dedicated proprietary hardware unavailable in most educational institutions [19].

C. Face Recognition as the Preferred Modality

Face recognition offers a unique operational profile among biometric modalities: it is passive, contactless, scalable to multiple simultaneous subjects, compatible with widely available camera hardware, and well-suited to natural interaction scenarios such as entering a lecture hall [8]. Unlike fingerprint or iris systems, face recognition can function without any explicit cooperation from the subject, enabling attendance to be recorded continuously from a fixed camera without interrupting lecture flow. The availability of open-source libraries, including OpenCV [13] and dlib [20], which provides pre-trained HOG-based detectors and CNN-based recognition pipelines, has substantially lowered the implementation barrier for educational institutions.

Face recognition-based attendance systems consist of four principal pipeline stages: (1) image acquisition from a webcam or CCTV camera; (2) face detection to localize facial regions within the captured frame; (3) face encoding to extract a compact, discriminative feature representation; and (4) face matching against an enrolled database to identify the individual and record attendance [9], [21]. Database management, session control, and web-based reporting interfaces constitute the system's administrative layer. Subsequent sections examine the algorithmic approaches employed at each of these stages, particularly the face detection and encoding components where deep learning has had the most transformative impact.

3. HANDCRAFTED FEATURE-BASED FACE RECOGNITION FOR ATTENDANCE SYSTEMS

Before the widespread adoption of deep learning, face recognition systems in attendance applications relied on handcrafted feature descriptors and classical machine learning classifiers. These approaches established the foundational pipeline that deep learning subsequently transformed, and several remain in practical use due to their computational efficiency.

The Histogram of Oriented Gradients (HOG) descriptor, introduced for pedestrian detection but rapidly extended to face detection, operates by computing the distribution of gradient orientations within localized image cells [22]. The HOG algorithm divides the input image into small overlapping cells, computes gradient magnitude and orientation for each pixel, constructs orientation histograms within each cell, and normalizes these histograms within larger blocks to achieve invariance to illumination changes. The resulting feature vector captures edge and shape information characteristic of the human face. When combined with a Linear Support Vector Machine (SVM) classifier in a sliding window detection framework, HOG provides computationally efficient face detection with acceptable accuracy under controlled conditions [23].

In the context of attendance systems, the HOG+SVM pipeline provides fast initial face detection suitable for real-time video processing. Studies utilizing dlib's pre-trained HOG face detector report detection accuracies of 75-80% in well-lit environments, with performance declining to approximately 70% under low-light conditions [9]. The computational efficiency of HOG detection, achieving processing times of approximately 50 milliseconds per frame on standard laptop hardware, makes it suitable for resource-constrained deployment without dedicated GPU acceleration.

Local Binary Patterns (LBP) represent another classical feature descriptor applied to face recognition in attendance systems. LBP encodes the local texture of an image by comparing each pixel with its neighbors and encoding the result as a binary string, yielding a histogram feature robust to monotonic illumination changes [24]. LBP-based face recognition was demonstrated in early automated attendance systems for small classroom environments, achieving identification accuracy around 85% under controlled lighting with frontal face acquisition. While LBP systems are computationally lightweight, their accuracy under real-world conditions with illumination variability and pose changes is substantially inferior to deep learning approaches.

Eigenface and Fisherface methods, based on Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), respectively, project facial images into lower-dimensional subspaces that capture the primary modes of facial variation [25]. These methods were among the earliest face recognition approaches applied to attendance systems, but have been largely superseded due to their sensitivity to illumination and pose variation and their requirement for controlled acquisition conditions. Comparative evaluations consistently demonstrate the superiority of CNN-based approaches over these classical projection methods across all operational conditions relevant to educational attendance monitoring.

4. DEEP CONVOLUTIONAL NEURAL NETWORK APPROACHES TO FACE RECOGNITION

The transition from handcrafted features to deep convolutional neural networks represents the most significant paradigm shift in face recognition for attendance monitoring. CNNs automatically learn hierarchical feature representations from training data, eliminating the need for manually designed descriptors and achieving substantially higher accuracy across diverse operating conditions [26].

DeepFace, developed by Facebook in 2014, was among the first systems to demonstrate near-human performance in face verification using a CNN architecture [10]. The DeepFace architecture consists of six convolutional layers followed by a face alignment preprocessing step using 3D face models, achieving 97.35% accuracy on the Labeled Faces in the Wild (LFW) benchmark. The key innovation of DeepFace was its explicit 3D-aligned frontalization preprocessing, which normalized facial images to a canonical pose before CNN feature extraction, substantially improving robustness to pose variation. This architectural approach demonstrated that combining geometric alignment with deep feature learning was more effective than either approach in isolation.

FaceNet, introduced by Google in 2015, proposed a fundamentally different learning objective: rather than training a classifier over a fixed set of identities, FaceNet trains a CNN to directly embed facial images into a 128-dimensional Euclidean space where distance is proportional to face dissimilarity [11]. This embedding approach, trained using a triplet loss function, enables face verification, recognition, and clustering within a unified framework. FaceNet employs an Inception-based CNN architecture and achieves 99.63% accuracy on LFW. The Euclidean distance metric in the embedding space enables straightforward addition of new identities to the database without retraining, making FaceNet particularly well-suited to attendance system applications where the enrolled population changes each academic year.

VGGFace, developed by the Visual Geometry Group at Oxford University in 2015, extends the VGG convolutional architecture to face recognition [12]. The VGGFace model uses an 18-layer convolutional network trained on a large-scale dataset of celebrity faces, achieving 98.95% accuracy on LFW. The depth of the VGGFace architecture enables the extraction of highly discriminative features but introduces significant computational cost, necessitating GPU acceleration for real-time deployment.Such computational demands may restrict its large-scale deployment in institutions with limited hardware infrastructure or low-cost attendance monitoring setups. The availability of pre-trained VGGFace weights through deep learning frameworks such as TensorFlow and Keras has enabled transfer learning approaches, where the pre-trained model is fine-tuned on domain-specific datasets of student faces, substantially reducing training data requirements.

OpenFace represents a particularly significant development for attendance system deployment in educational institutions, as it provides an open-source implementation of a FaceNet-inspired architecture optimized for training with small datasets [27]. OpenFace employs the FaceNet triplet loss framework within a modified architecture with reduced parameter count, enabling effective training on datasets too small to support the original FaceNet model. This characteristic is directly relevant to educational attendance systems where the enrolled population of one institution may be insufficient for training large-scale commercial models from scratch. OpenFace achieves 92.92% accuracy on LFW while requiring substantially less computational resources than FaceNet or VGGFace, making it a practical choice for deployment on laptop or server hardware without dedicated GPU infrastructure. The dlib face recognition library encapsulates a ResNet-based deep metric learning model trained on a dataset of approximately three million faces, producing 128-dimensional facial embeddings [20]. In attendance system implementations using Python and dlib, the face_recognition library provides a high-level API that wraps dlib's detection and embedding functions, enabling attendance system development with minimal signal processing expertise. The combination of dlib's HOG-based face detection for initial localization, 68-point facial landmark estimation for alignment, and ResNet-based embedding for recognition represents the most widely adopted technical pipeline in academic attendance system implementations [9], [28].

A comparative analysis of these architectures indicates that model selection for attendance monitoring systems should not be based solely on benchmark recognition accuracy. While DeepFace and FaceNet demonstrate superior performance on large-scale benchmark datasets, their computational complexity and dependence on extensive training data may limit direct deployment in resource-constrained educational environments. In contrast, OpenFace and dlib-based embedding frameworks offer a more practical balance between recognition accuracy, computational efficiency, and ease of integration with real-time attendance systems. This trade-off between algorithmic performance and deployment feasibility remains a critical consideration in educational biometric applications.

5. END-TO-END DEEP METRIC LEARNING AND FACE EMBEDDING ARCHITECTURES

The deep metric learning paradigm, which underpins FaceNet, OpenFace, and dlib's recognition model, has become the dominant architectural approach for attendance system face recognition due to its open-set recognition capability, the ability to identify individuals not seen during training, and its practical scalability for updating enrolled databases without model retraining [11], [27].

Deep metric learning for face recognition trains a CNN to produce facial embeddings such that embeddings of the same identity cluster closely in the embedding space while embeddings of different identities are well-separated. The choice of loss function critically determines the quality of the learned embedding space. Triplet loss, used in FaceNet, operates on triplets of anchor, positive, and negative samples, minimizing the distance between anchor and positive embeddings while maximizing the distance to negative embeddings [11]. Centre loss, used as an auxiliary objective in conjunction with softmax loss, simultaneously learns class centers and minimizes intra-class variation, improving cluster compactness in the embedding space [29]. ArcFace (Additive Angular Margin Loss) introduces an angular margin penalty in the softmax loss formulation, achieving improved class separability and state-of-the-art performance on multiple face recognition benchmarks [30].

In operational attendance systems, deep metric learning manifests as a three-phase pipeline: enrollment, where student facial images are processed through the embedding CNN to generate 128-dimensional or 512-dimensional feature vectors stored in a database; detection and embedding, where live camera frames are processed in real-time to extract facial embeddings; and matching, where Euclidean or cosine distance between the live embedding and enrolled embeddings determines identity [9]. A distance threshold governs the match decision: if the minimum distance to any enrolled embedding falls below the threshold, the corresponding identity is assigned and attendance marked; otherwise, the face is classified as unknown. The selection of this threshold represents a fundamental operating point trade-off between false acceptance rate (FAR) and false rejection rate (FRR), and must be calibrated empirically for each deployment environment. This trade-off becomes increasingly critical in high-density lecture environments, where even marginal increases in false recognition rates may lead to significant cumulative attendance errors across large student populations.

Studies implementing dlib-based attendance systems report optimal threshold values in the range of 0.6 Euclidean distance in the 128-dimensional embedding space, corresponding to approximately 60% similarity [9], [28]. This threshold achieves over 90% overall recognition accuracy under optimal lighting conditions while maintaining acceptable false acceptance rates. However, threshold optimality varies with lighting conditions, camera resolution, and population size, necessitating system-specific calibration. Face counting, implemented by tracking the number of unique face detections across frames, serves as a complementary mechanism for verifying complete class attendance without individual recognition, providing a lightweight cross-check against the recognition-based attendance log.

Real-time processing performance for deep metric learning-based attendance systems depends primarily on the face detection and embedding stages. HOG-based detection achieves frame rates of approximately 20 FPS on CPU, while CNN-based detection (dlib's MMOD detector) offers improved accuracy at approximately 1-2 FPS on CPU, necessitating GPU acceleration for real-time operation [20]. The embedding computation using dlib's ResNet model requires approximately 200ms per face on CPU, limiting throughput to approximately 5 faces per second without parallelization. Such latency may become a major bottleneck in large lecture halls with simultaneous multi-face recognition requirements, particularly in institutions operating without GPU-assisted infrastructure. For lecture hall deployments with large simultaneous student cohorts, batch processing strategies or edge computing solutions are required to meet real-time attendance recording requirements.

From a deployment perspective, deep metric learning architectures offer a substantial advantage over closed-set classification models due to their flexibility in handling open-set recognition scenarios. However, system performance is highly sensitive to threshold calibration, embedding quality, and environmental variability. In practical classroom deployments, threshold misconfiguration may significantly increase false acceptance and false rejection rates, thereby affecting attendance reliability and system trustworthiness. Consequently, adaptive thresholding and environment-specific calibration strategies remain critical research directions.

6. MULTIMODAL BIOMETRIC AND HYBRID ATTENDANCE SYSTEM ARCHITECTURES

While single-modality face recognition systems have achieved substantial deployment in educational attendance applications, research has increasingly explored multi-modal biometric architectures that combine face recognition with complementary modalities to improve robustness, accuracy, and security. Multi-modal systems exploit complementary information from independent biometric channels to overcome the limitations of individual modalities.

The combination of face recognition with fingerprint recognition represents a logical multi-modal pairing for attendance systems, as the two modalities have largely complementary failure modes: face recognition degrades with occlusion and extreme illumination variation, while fingerprint recognition is sensitive to moisture and skin condition but performs well under diverse illumination [31]. Score-level fusion of face and fingerprint recognition results, implemented through weighted sum or product rules, has demonstrated improved verification accuracy compared to either modality alone, with studies reporting verification error rates approximately 40% lower than the best single-modality result in controlled evaluation settings. However, the contact requirement of fingerprint acquisition limits the practical advantage of this combination in education contexts where non-contact operation is desirable.

Face recognition combined with iris recognition offers a fully contactless multi-modal architecture with high complementary accuracy. Iris recognition achieves extremely high uniqueness and stability across age and provides strong performance under varying illumination through active near-infrared illumination. Fusion of face and iris recognition at the score level has achieved near-zero error rates in controlled environments [32]. However, the specialized hardware requirements for iris recognition, particularly the near-infrared camera and constrained acquisition distance, present high cost and installation barriers for widespread educational deployment.

Liveness detection, or anti-spoofing, represents a critical security layer in face recognition-based attendance systems that has received increasing research attention. Standard face recognition systems are vulnerable to presentation attacks in which photographs, video replays, or 3D masks of enrolled students are presented to the camera to fraudulently mark attendance. Deep learning-based liveness detection approaches, including texture analysis using fine-grained CNNs and temporal analysis of physiological signals such as remote photoplethysmography (rPPG), provide effective discrimination between live faces and spoofing attacks without requiring additional hardware [33]. Integration of liveness detection into attendance system pipelines is increasingly considered essential for operational deployment where motivated attack scenarios are plausible. This is particularly important in university settings where proxy attendance and presentation attacks remain realistic misuse scenarios.

The integration of attendance systems with a broader Learning Management System (LMS) infrastructure represents a system-level multi-modal fusion approach, combining biometric identification with digital session management, grade records, and learning analytics. Attendance records captured by face recognition systems can be automatically synchronized with LMS databases, enabling real-time correlation of attendance patterns with academic performance metrics, early identification of at-risk students, and automated notification workflows [34]. This integration transforms the attendance system from an isolated operational tool into a component of a comprehensive learning analytics infrastructure, substantially increasing the institutional value of the biometric investment. Beyond administrative automation, such integration also creates opportunities for predictive learning analytics, early intervention frameworks, and data-driven student support mechanisms. Although multimodal biometric systems demonstrate improved robustness and security compared to single-modality face recognition frameworks, their practical adoption in educational institutions remains constrained by hardware cost, system complexity, and maintenance overhead. The integration of multiple biometric channels must therefore be evaluated not only in terms of recognition accuracy but also with respect to scalability, deployment feasibility, user acceptance, and long-term operational sustainability within academic environments.

7. Research gaps and future directions

Despite the substantial progress achieved by deep learning-based facial recognition systems for automated attendance monitoring, the existing body of literature continues to exhibit several unresolved methodological, computational, and deployment-oriented limitations that restrict large-scale real-world adoption in educational institutions.

One of the most prominent research gaps lies in robustness under unconstrained classroom environments. Although contemporary architectures such as FaceNet, ArcFace, and ResNet-based embedding frameworks report recognition accuracies exceeding 97% on benchmark datasets such as LFW and VGGFace2, these performance metrics are often obtained under highly controlled image acquisition settings. In contrast, practical classroom environments introduce significant variability in illumination intensity, directional lighting, motion blur, partial occlusions caused by masks, spectacles, or hairstyles, as well as non-frontal pose deviations. Existing studies have not sufficiently addressed domain adaptation mechanisms capable of preserving embedding discriminability under these highly heterogeneous conditions. Recent surveys continue to identify robustness under unconstrained environments as a major open problem in modern face recognition systems.

A second major gap concerns computational scalability and inference latency in high-density academic settings. Most reported systems are evaluated on small to medium student populations, whereas real institutional deployments may require simultaneous recognition of 100–300 students within a single lecture session. Such large-scale recognition pipelines impose significant demands on feature extraction throughput, embedding comparison efficiency, and database retrieval latency. Deep CNN architectures with high parameter complexity, particularly VGG-based and ResNet-derived models, may introduce unacceptable delays in CPU-only environments. This limitation becomes especially critical in institutions lacking dedicated GPU infrastructure. Consequently, lightweight inference architectures, model pruning strategies, quantization techniques, and edge-optimized deployment pipelines remain important future research directions. Recent work increasingly highlights Edge AI and edge intelligence architectures as viable solutions for real-time low-latency biometric systems.

Another critical gap lies in privacy-preserving biometric learning frameworks. Since facial embeddings and raw face images constitute highly sensitive biometric identifiers, centralized model training and cloud-based storage introduce substantial regulatory and ethical concerns, particularly in relation to GDPR-compliant data governance and institutional privacy policies. Traditional centralized training pipelines require aggregation of sensitive student data, thereby increasing the risk of privacy leakage, identity reconstruction attacks, and unauthorized biometric profiling. Recent advances in federated learning, privacy-preserving aggregation, and secure multiparty computation provide a technically promising research direction whereby model training can be distributed across local edge devices without transmitting raw biometric data. Studies in privacy-preserving federated face recognition have demonstrated that distributed learning frameworks can preserve recognition performance while significantly reducing privacy risk.

In addition to privacy, fairness and demographic bias remain underexplored in attendance monitoring literature. Recent large-scale face recognition surveys emphasize that model performance may vary significantly across demographic groups, including skin tone, gender, facial morphology, and age-based variations. Such disparities may lead to systematic false rejection of certain student populations, directly affecting attendance integrity and academic fairness. Despite its importance, the majority of attendance-system studies focus primarily on aggregate accuracy metrics while neglecting subgroup-sensitive performance evaluation. This represents a substantial methodological gap that future research must address through fairness-aware loss functions, balanced training datasets, and bias-sensitive performance auditing frameworks.

Furthermore, anti-spoofing and liveness verification mechanisms remain insufficiently integrated into many existing systems. Conventional face recognition attendance frameworks remain vulnerable to presentation attacks using printed photographs, replay videos, or digital screen images. While standalone liveness detection literature has progressed significantly, its systematic fusion with classroom-scale attendance architectures remains limited. Future research should prioritize multimodal anti-spoofing pipelines incorporating texture-based CNN classifiers, temporal physiological signal analysis, and depth-aware facial verification mechanisms.

From a systems perspective, future research should also focus on institutional-scale integration with Learning Management Systems (LMS), predictive learning analytics platforms, and academic risk monitoring frameworks. Beyond simple attendance recording, biometric attendance data can serve as a valuable temporal signal for student engagement modelling, dropout risk prediction, and early academic intervention systems. Such integration would significantly increase the strategic institutional value of attendance monitoring systems beyond administrative automation alone.

Collectively, these unresolved challenges indicate that the current literature remains predominantly algorithm-centric, with comparatively limited emphasis on deployment engineering, privacy governance, fairness auditing, and long-term institutional integration. Future advances in robust feature learning, federated edge intelligence, lightweight inference optimization, and fairness-aware biometric analytics will be essential for the development of scalable, secure, and ethically compliant next-generation attendance monitoring systems.

8. COMPARATIVE ANALYSIS OF MAJOR FACE RECOGNITION APPROACHES FOR ATTENDANCE MONITORING

Table I summarizes the major face recognition approaches reviewed in this paper across architecture, benchmark datasets, strengths, limitations, and deployment suitability.

Approach / Algorithm	Core Architecture	Typical Dataset / Benchmark	Strengths	Limitations	Deployment Suitability
HOG + SVM	Handcrafted gradient features with a linear classifier	Small institutional datasets	Fast CPU inference, lightweight implementation, suitable for real-time basic systems	Poor robustness to pose, illumination, and occlusion	Suitable for low-cost small classroom deployment
LBP	Texture-based	Controlled	Computationally	Low	Limited

	local binary encoding	frontal face datasets	efficient, illumination-tolerant to some extent	discriminative power under real-world variation	suitability
Eigenfaces / PCA	Linear dimensionality reduction	Controlled grayscale face datasets	Low computational cost, easy implementation	Highly sensitive to noise, pose, and lighting changes	Mostly academic / legacy use
DeepFace	Deep CNN with alignment-based preprocessing	LFW	High benchmark accuracy, robust feature extraction	High computational complexity, large training data requirement	Moderate, requires strong infrastructure
FaceNet	Deep metric learning with triplet loss	LFW, VGGFace	Excellent embedding quality, scalable open-set recognition	Training complexity, threshold sensitivity	High suitability
VGGFace	Deep CNN with large parameter depth	VGGFace, LFW	Strong feature discriminability	Heavy memory and GPU requirement	Suitable for server-based deployment
OpenFace	Lightweight FaceNet-inspired model	Medium-scale datasets	Good balance between performance and speed	Lower accuracy than large-scale commercial models	High practical suitability
dlib ResNet Embeddings	ResNet-based 128D embedding	Real-world institutional datasets	Practical implementation, strong open-source support	CPU latency in high-density environments	Very high suitability

Table 1 - Comparative Analysis of Major Face Recognition Approaches for Attendance Monitoring

9. Conclusion

This review has critically examined the algorithmic progression of automated biometric attendance monitoring systems, with particular emphasis on the transition from classical handcrafted feature extraction frameworks to contemporary deep learning and deep metric learning architectures. The literature clearly demonstrates that the evolution from HOG-SVM, LBP, Eigenface, and PCA-based recognition pipelines toward convolutional neural network models, embedding-based metric learning systems, and multimodal biometric frameworks has substantially improved recognition robustness, feature discriminability, and operational scalability in educational environments. The shift from handcrafted descriptors to deep learned representations has proven transformative, with CNN architectures such as DeepFace, FaceNet, VGGFace, and ResNet-based embedding models enabling hierarchical feature abstraction, while triplet loss and angular margin objectives significantly improve open-set recognition performance. Beyond algorithmic advances, institutional-scale deployment demands real-time inference optimization, privacy-preserving frameworks such as federated learning and differential privacy, and fairness-aware training to address demographic bias. Anti-spoofing integration and LMS-linked analytics remain critical open challenges for future research.In conclusion, automated biometric attendance monitoring has evolved into a highly sophisticated interdisciplinary domain characterized by complex data representations, advanced optimization algorithms, and real-time deployment challenges. The future of this field lies in integrating scalable deep learning architectures, privacy-preserving distributed intelligence, and adaptive data-driven optimization frameworks capable of supporting robust, fair, and institution-scale smart attendance ecosystems.

10. References

[1] M. A. Thalor and Omkar S. Gaikwad, "Facial Recognition Attendance Monitoring System using Deep Learning Techniques," *International Journal of Integrated Science and Technology*, vol. 1, no. 6, pp. 841–848, Jan. 2024, doi: 10.59890/ijist.v1i6.685.

[2] B. Xu, "Application of face recognition technology in academic administration in colleges and universities," in Proc. International Conference on Trends in Multimedia Computing, Information and Communication Technology (ICTMIC), Francis Academic Press, Jun. 2020, doi: 10.25236/ictmic.2020.088.

[3] S. M. Ahmed, R. Hasan, and M. T. Rahman, "Real-time facial recognition attendance system using Python and OpenCV for educational institutions," International Journal of Advanced Computer Science and Applications, vol. 15, no. 2, pp. 145–153, 2024.

[4] A. Abdelbar *et al.*, "Face Recognition Using Python and OpenCV: Review," *International Journal of Scientific and Research Publications*, vol. 14, no. 1, p. 70, 2024, doi: 10.29322/IJSRP.14.01.2024.p14508.

[5] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 1, pp. 4–20, Jan. 2004.

[6] K. Okereafor, I. Ekong, I. O. Markson, and K. Enwere, "Fingerprint biometric system hygiene and the risk of COVID-19 transmission," JMIR Biomedical Engineering, vol. 5, no. 1, p. e19623, 2020, doi: 10.2196/19623.

[7] J. R. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D. J. LoIacono, S. Mangru, M. Tinker, T. Zappia, and W. Y. Zhao, "Iris on the move: Acquisition of images for iris recognition in less constrained environments," Proceedings of the IEEE, vol. 94, no. 11, pp. 1936--1947, Nov. 2006, doi: 10.1109/JPROC.2006.884091.

[8] D. M. Prasanna and C. G. Reddy, "Development of real time face recognition system using OpenCV," International Research Journal of Engineering and Technology (IRJET), vol. 4, no. 12, pp. 791, Dec. 2017. [Online]. Available: https://www.irjet.net/archives/V4/i12/IRJET-V4I12151.pdf

[9] M. V. B. Chandrasekhar, C. Vamsi, B. Yogeswara Rao, R. Meghana, and Y. Premkumar, "Automated attendance system with AI and real-time camera access using facial recognition," Communications on Applied Nonlinear Analysis, vol. 32, no. 9s, 2025. [Online]. Available: https://internationalpubls.com/index.php/cana/article/view/4261

[10] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 1701–1708.

[11] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 815–823.

[12] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in Proceedings of the British Machine Vision Conference (BMVC), 2015, pp. 41.1–41.12.

[13] P. Udaya Varshini, R. Venkata, V. Lalitha, S. V. Lakshmi, and D. Swathi, "Enhanced facial recognition attendance system for educational institutions with OpenCV," International Journal of Progressive Research in Engineering Management and Science, 2023. [Online]. Available: www.ijprems.com

[14] D. E. King, "Dlib-ml: A machine learning toolkit," Journal of Machine Learning Research, vol. 10, pp. 1755--1758, 2009.

[15] S. Kumar, S. Sumathi, and D. Shree, "Artificial intelligence (AI) strategies for face mask recognition and facial authenticity," in Proceedings of the International Conference on Innovative Computing and Communication (ICICC), 2022, doi: 10.2139/ssrn.4143816.

[16] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2001, vol. 1, pp. I–511–I–518.

[17] O. Alniemi and H. F. Mahmood, "Class Attendance System Based on Face Recognition," *Revue d'Intelligence Artificielle*, vol. 37, no. 5, pp. 1245–1253, 2023, doi: 10.18280/ria.370517.

[18] U. R. Ughade, S. M. Gikwad, A. N. Yeole, and Dr. A. O. Mulani, "Automatic Attendance System Using Face Recognition," *Journal of Image Processing and Intelligent Remote Sensing*, no. 34, pp. 11–18, Jul. 2023, doi: 10.55529/jipirs.34.11.18.

[19] J. Daugman, "How iris recognition works," IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 1, pp. 21–30, Jan. 2004.

[20] P. K. Tiwari, Rohan, Rinku, and A. Kumar, "Implementation of face detection and recognition using OpenCV for real-time biometric authentication and attendance systems," International Journal of Research - GRANTHAALAYAH, vol. 12, no. 11, pp. 82--93, Nov. 2024, doi: 10.29121/granthaalayah.v12.i11.2024.6121.

[21] M. Turk and A. Pentland, "Eigenfaces for recognition," Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71–86, 1991.

[22] O. Déniz, G. Bueno, J. Salido, and F. De la Torre, "Face recognition using histograms of oriented gradients," in Proceedings of the International Conference on Computer Vision Theory and Applications (VISAPP), 2011, pp. 139–144.

[23] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005, vol. 1, pp. 886--893.

[24] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: Application to face recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.

[25] S. Gong, V. N. Boddeti, and A. K. Jain, "On the Capacity of Face Representation," Apr. 2019, [Online]. Available: http://arxiv.org/abs/1709.10433

[26] S. Kishore, P. H. G. Padmavathi, S. C. S. Sowmyshree, and M. K. B. Manjula, "Evaluation of deep learning methods in face recognition: Datasets, metrics, and results," International Journal on Science and Technology (IJSAT), vol. 16, no. 4, pp. 1--8, Oct.--Dec. 2025. [Online]. Available: https://www.ijsat.org/papers/2025/4/9116.pdf

[27] B. Amos, B. Ludwiczuk, and M. Satyanarayanan, "OpenFace: A general-purpose face recognition library with mobile applications," Carnegie Mellon University, CMU-CS-16-118, 2016.

[28] A. Jadhav, T. Argade, P. Darekar, N. Badhe, and S. Badhe, "Automated attendance system using face recognition," International Journal of Creative Research Thoughts (IJCRT), vol. 13, no. 4, Apr. 2025. [Online]. Available: https://www.ijcrt.org/papers/IJCRT25A4030.pdf

[29] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in European Conference on Computer Vision (ECCV), 2016, pp. 499–515.

[30] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 4690–4699.

[31] A. N. Matar, "Smart attendance system using face recognition: Robustness and accuracy evaluation under occlusions, lighting variations, and facial angle changes," M.S. thesis, KTH Royal Institute of Technology, Stockholm, Sweden, 2024. [Online]. Available: https://kth.diva-portal.org/smash/get/diva2:2046973/FULLTEXT01.pdf

[32] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, "Image understanding for iris biometrics: A survey," Computer Vision and Image Understanding, vol. 110, no. 2, pp. 281–307, May 2008.

[33] H. Xing, S. Y. Tan, F. Qamar, and Y. Jiao, "Face anti-spoofing based on deep learning: A comprehensive survey," Applied Sciences, vol. 15, no. 12, p. 6891, 2025.

[34] R. Vhatkar, A. Shaikh, Q. Hasan, and D. Deore, "Attendance management system using face recognition," International Journal for Research Trends and Innovation, vol. 8, no. 4, pp. 1430, 2023. [Online]. Available: https://www.ijrti.org/papers/IJRTI2304235.pdf