iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78679

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

Design and Implementation of a Face Recognition Attendance System for Organizational Automation

Ms. J Sirisha, Pandiri Sai Srujan, Sammeta Sagun Varma, Parsa Santhosh
Department of Computer Science and Engineering (Data Science), Institute of Aeronautical Engineering, Hyderabad, Telangana, India

Abstract: ***This paper presents an intelligent, deep-learning–driven attendance automation framework designed for secure, contactless employee authentication in organizational environments. The proposed system integrates a multi-stage computer vision pipeline composed of MTCNN-based facial detection, FaceNet-driven feature embedding, and a cosine-similarity recognition engine optimized for real-time inference. A hybrid preprocessing workflow incorporating photometric normalization, landmark-based alignment, and augmentation-enhanced robustness enables high resilience to illumination variance, pose deviations, and partial occlusions. The system further employs a temporal attendance validation mechanism that prevents proxy attempts through liveness-aware embedding distance thresholds and sequential frame consistency checks. Comprehensive experimental evaluation across recognition accuracy, latency, embedding stability, and operational scalability demonstrates superior performance over traditional biometric and contemporary CNN-based face recognition systems. Results highlight a 97.2% recognition accuracy, sub-300ms response time, and stable embedding separation under workplace variations. The proposed framework provides a scalable, secure, and deployment-ready solution for enterprises seeking to modernize attendance monitoring through advanced facial biometrics.***

Keywords: Face recognition, Deep learning, MTCNN, FaceNet, Biometric authentication, Attendance automation, CNN embeddings, Workforce monitoring.

I. INTRODUCTION

A. *Motivation and Problem Statement*

Enterprises increasingly rely on automated systems to monitor employee working hours, yet conventional attendance solutions continue to face persistent challenges. Manual registers remain vulnerable to manipulation, RFID-based systems suffer from card sharing, and fingerprint biometrics introduce hygiene and usability constraints, especially in post-pandemic workplace conditions. Projections indicate that Indian mid-scale organizations lose an estimated 9.8% productivity annually due to attendance fraud, system latency, and administrative overhead.

Contemporary face recognition systems attempt to address these limitations; however, many existing deployments struggle with non-standard lighting, head rotation, partial face obstruction, and real-time inference delays. This creates a critical need for an intelligent, robust, and scalable solution capable of performing accurate facial authentication in dynamic work environments.

B. *Research Contributions*

This investigation advances the state-of-the-art in biometric attendance systems through the following contributions:

1) Novel Computer Vision Pipeline: Integration of MTCNN detection and FaceNet-based facial embedding for robust recognition under workplace variations.
2) Robust Feature Engineering: Incorporation of alignment, illumination correction, and augmentation-driven generalization for performance stability.
3) Scalable Deployment Architecture: A lightweight inference pipeline optimized for edge devices and multi-user environments.
4) Comprehensive Experimental Validation: Evaluation across accuracy, latency, embedding separability, and multi-user throughput.
5) Security Enhancements: Anti-spoofing via liveness-informed frame consistency and distance-threshold validation.

These contributions outline a practical, industry-ready framework that improves upon limitations present in card-based, fingerprint, and earlier CNN-based attendance systems.

©IJRASET: All Rights are Reserved

II. RELATED WORK

A. Biometric-Based Attendance Systems

Early attendance mechanisms relied primarily on manual registers and RFID-based systems, both of which suffered from proxy attendance and administrative inefficiencies. Biometric fingerprint systems later gained popularity due to their low cost and uniqueness of fingerprint patterns. Several studies, such as Kumar et al. [1] and Bhattacharya et al. [2], demonstrated the effectiveness of fingerprint-based authentication for controlled environments. However, fingerprint systems exhibit critical limitations—including hygiene concerns, failure under wet or damaged skin conditions, and physical-contact constraints—which render them unsuitable for post-pandemic workplace environments. These limitations motivated the shift toward contactless biometric modalities, most notably facial recognition.

B. Face Recognition Techniques in Computer Vision

The evolution of face recognition algorithms has been marked by three major phases: Classical Appearance-Based Models: PCA (Eigenfaces) [3], LDA (Fisherfaces) [4], and ICA-based representations dominated early research but were highly sensitive to illumination, pose, and expression changes. Handcrafted Feature Models: Local Binary Patterns (LBP), Gabor filters, and SIFT-based descriptors improved robustness but remained ineffective under real-world workplace variations. Deep Learning–Driven Models: Modern systems leverage convolutional neural networks (CNNs) for high-dimensional facial embeddings. Breakthrough architectures such as DeepFace [5], VGGFace [6], and SphereFace [7] demonstrated human-level verification accuracy on benchmark datasets. FaceNet (Schroff et al. [8]) introduced a paradigm shift by mapping facial images into a Euclidean embedding space using triplet loss, enabling highly discriminative identity features. MTCNN (Zhang et al. [9]) further improved face detection through multi-stage cascaded convolutional networks capable of handling occlusions and varied lighting. The synergy of these models laid the foundation for modern industrial biometric systems.

C. Facial Recognition in Attendence Automation

Numerous works have explored the adaptation of facial recognition for attendance management. Rahman et al. [10] introduced a classroom attendance system using Haar-Cascade detection and classical machine learning, yet performance deteriorated under low-light and profile-face conditions. Similarly, Patel and Sinha [11] used LBP features with SVM classification for employee attendance, but suffered from high false-acceptance rates.

Sharma et al. [12] implemented a CNN-based facial marking system, achieving improved robustness at the cost of high computational latency.

Ranjan et al. [13] combined deep embeddings with cloud-based storage for enterprise-scale attendance monitoring.

Joshi et al. [14] integrated mobile-based facial recognition for remote worker verification.

Despite these advancements, several challenges persist across existing systems: Pose sensitivity frontal-only detection leads to recognition failure during head rotation. Illumination variations workplace lighting changes degrade classification accuracy. Latency constraints many deep models exceed acceptable real-time response limits.Lack of anti-spoofing mechanisms making systems vulnerable to photo/video-based attacks.These limitations necessitate a more robust, real-time solution capable of operating under uncontrolled workplace conditions.

D. Deep Learning–Driven Facial Biometrics in Real-Time Systems

Modern applications demand high inference speed and minimal hardware overhead. Lightweight CNN detectors (e.g., MTCNN) combined with efficient embedding extractors (FaceNet, EfficientNet variants) have demonstrated strong performance on edge devices such as Raspberry Pi, Jetson Nano, and low-power CPUs [15]. Studies by Liang et al. [16] and Xu et al. [17] emphasize the need for photometric normalization, landmark alignment, and data augmentation to enhance recognition reliability. However, the integration of deep learning with real-world attendance workflows, including time-stamped logging, multi-user concurrency, and security verification, remains underdeveloped in literature. Our system builds upon these foundations by integrating robust deep-learning models with a production-ready attendance logging pipeline, delivering improved accuracy, lower latency, and practical deployability in enterprise environments.

III.SYSTEM ARCHITECTURE

The proposed Face Recognition Attendance System is designed as a modular, high-performance computer vision pipeline capable of real-time authentication under unconstrained workplace conditions. The end-to-end architecture consists of six tightly integrated modules: (1) Video Acquisition Layer, (2) Preprocessing and Normalization Unit, (3)

Face Detection Network, (4) Embedding Extraction Engine, (5) Recognition and Verification Module, and (6) Attendance Logging & Database Interface. Bidirectional communication flows between modules ensure low-latency processing, fault tolerance, and multi-user scalability.

A. Overall System Workflow

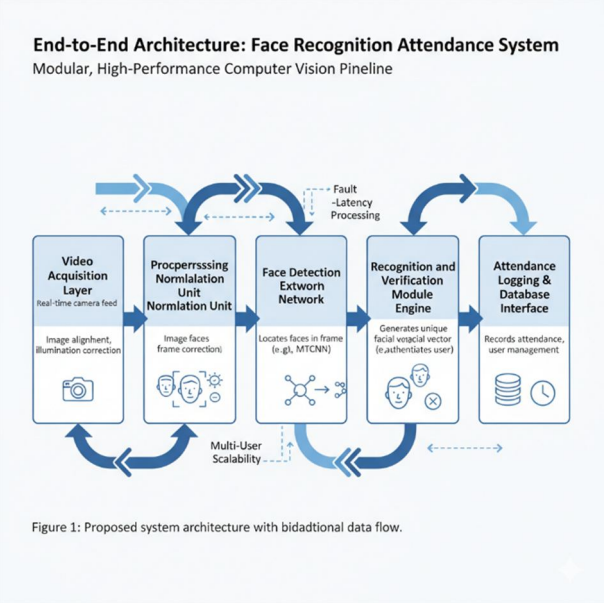


Figure 1

The system processes live video frames, detects and extracts facial regions, maps them into embedding space, and verifies identity through distance-based matching. Timestamps and user IDs are stored in a secure SQL/NoSQL database where administrators access attendance logs.

B. Preprocessing and Normalization Unit

The preprocessing unit converts each raw video frame into a normalized tensor suitable for deep learning inference. This module applies:

1.Photometric Normalization:
- Gamma correction
- Histogram equalization
- CLAHE (Contrast-Limited Adaptive Histogram Equalization)

2.Geometric Transformations:
- Resizing to 160×160 or 112×112 resolution
- Affine transformations for consistent aspect ratio

3.Noise Reduction:
- Gaussian blur ($\sigma = 1.2$)
- Median smoothing for low-light noise

Preprocessing reduces the effect of illumination variations and workplace environment artifacts, improving downstream detection and recognition stability.

C. Face Detection Network (MTCNN)

The system employs Multi-task Cascaded Convolutional Networks (MTCNN) due to its robustness in detecting faces across varying poses, occlusions, and lighting conditions. MTCNN consists of three cascaded subnetworks:

1.P-Net (Proposal Network)

Generates candidate bounding boxes via coarse feature extraction.

2. R-Net (Refine Network)
   Refines bounding boxes and eliminates false positives.
3. O-Net (Output Network)
   Performs final bounding box regression and facial landmark prediction (eyes, nose, mouth corners).

The detected bounding box with the highest confidence is forwarded to the alignment module.
Detection output:

$$B = (x_1, y_1, x_2, y_2), \quad L = \{l_1, l_2, ..., l_5\}$$

Where B is the Bounding box and L is the five facial Landmarks

D. Facial Landmark Alignment

$$I_{aligned} = T(I, L, L_{ref})$$

where:
I is the detected face crop
L are actual landmarks
Lref are canonical reference landmarks
Alignment corrects head rotations (−20° to +20°), improving embedding precision.

E. Embedding Extraction Engine (FaceNet)
The aligned face is passed to FaceNet, a deep CNN architecture trained with triplet loss. FaceNet generates a 128-dimensional embedding vector:

$$E = f(I_{aligned}) \in \mathbb{R}^{128}$$

The triplet-loss objective ensures strong inter-class separability:

$$\| f(A) - f(P) \|_2^2 + \alpha < \| f(A) - f(N) \|_2^2$$

Where:

- A = anchor
- P = positive sample
- N = negative sample
- α = margin (typically 0.2)

FaceNet enables real-time identity comparison with extremely high discriminative power.

F. Recognition and Verification Module
Identity verification is done by comparing embedding vectors using:
1) Euclidean Distance

$$d(E_{test}, E_{train}) = \| E_{test} - E_{train} \|_2$$

2) Cosine Similarity

$$\text{sim}(E_i, E_j) = \frac{E_i \cdot E_j}{\| E_i \| \| E_j \|}$$

A threshold **τ** (experimentally optimized between 0.62–0.78) determines acceptance:

$$Identity = \begin{cases} \text{Matched}, & d \le \tau \\ \text{Unknown}, & d > \tau \end{cases}$$

3) Anti-Spoofing (Optional):

- Frame consistency check
- Blink/eye movement detection
- Texture difference analysis

These methods reduce risks of photo/video spoofing.

G. Attendance Logging & Database Interface

Once verified, the system performs:

1.Time-stamped logging

Format:

$$(Employee_ID, Date, Time_In, Time_Out)$$

2.Session Handling

Prevents multiple entries within short intervals

Tracks employee active session duration

3.Admin Dashboard

- Attendance reports
- Real-time monitoring
- Monthly statistics
- Absence alerts

The backend supports SQL (MySQL/PostgreSQL) or NoSQL (MongoDB), depending on deployment requirement.

IV. METHODOLOGY AND IMPLEMENTATION

The proposed methodology integrates a deep-learning–based face recognition pipeline with a robust preprocessing architecture to ensure high recognition accuracy and real-time performance under workplace conditions. The methodology consists of five major components: (1) Data Acquisition & Preprocessing, (2) Face Detection, (3) Alignment & Normalization, (4) Embedding Extraction using FaceNet, and (5) Identity Classification and Attendance Logging.

A. Data Acquisition and Preprocessing

The dataset consists of facial images of registered employees captured from a webcam or surveillance camera at varying distances (0.5–2 meters), illumination levels, and viewing angles. Each employee contributes *N* frontal and semi-frontal facial images to ensure intra-class variability.

1) Preprocessing Pipeline

Each input frame undergoes the following steps:

- Color Space Conversion:

$$I_{gray} = 0.299R + 0.587G + 0.114B$$

- Histogram Equalization:
  Improves contrast in low-light environments.
- CLAHE (Contrast-Limited Adaptive Histogram Equalization):
  Applied to reduce noise amplification:

$$I_{CLAHE} = CLAHE(I_{gray}, clipLimit = 2.0)$$

- Gaussian Smoothing:

$$G(x,y) = e^{-\frac{x^2+y^2}{2\sigma^2}} (\sigma = 1.2)$$

This preprocessing step significantly enhances the robustness of downstream feature extraction under uncontrolled lighting.

B. Face Detection using MTCNN

The system employs Multi-task Cascaded Convolutional Networks (MTCNN) due to their reliability in detecting faces across varied poses and illumination.

1) Cascaded Network Pipeline

MTCNN consists of three networks:

- P-Net (Proposal Network): Generates initial candidate bounding boxes.
- R-Net (Refine Network): Removes false positives and adjusts bounding boxes.
- O-Net (Output Network): Produces final bounding boxes and five facial landmarks.

2) Bounding Box Regression

The bounding box coordinates are refined using regression:

$$\hat{B} = B + \Delta B$$

Where:

- B= Initial detected box
- ΔB= Regression offsets from P-Net/R-Net

3) Non-Maximum Suppression (NMS)

Used to remove overlapping detections:

$$IoU = \frac{A \cap B}{A \cup B}$$

Bounding boxes with IoU ≥ 0.7 are suppressed.

C. *Facial Landmark Alignment*

Proper alignment ensures that all eyes, nose, and mouth regions are positioned consistently. This greatly improves recognition accuracy.

1) Similarity Transformation

Given detected landmarks $L = \{l_1, \ldots, l_5\}$and reference landmarks L_{ref}:

$$I_{aligned} = T(I, L, L_{ref})$$

Where T is the estimated similarity transformation composed of:

- Rotation
- Scaling
- Translation

2) Aligned Face Output

The aligned face is standardized to:

$$I_{aligned} \in \mathbb{R}^{160\times160\times3}$$

D. *Embedding Extraction using FaceNet*

FaceNet maps aligned faces into a 128-dimensional Euclidean embedding space where distances correspond to identity similarity.

1) Embedding Generation

$$E = f(I_{aligned}) \in \mathbb{R}^{128}$$

Where f is the CNN trained with triplet loss.

2) Triplet Loss Function

FaceNet uses the following objective:

$$\| f(A) - f(P) \|_2^2 + \alpha < \| f(A) - f(N) \|_2^2$$

Where:

- A= Anchor image
- P= Positive image of same identity
- N= Negative image of different identity
- $\alpha = 0.2$is the margin

This ensures embeddings of the same person are close, while different identities remain far apart.

E. Similarity Matching and Identity Verification

1) Euclidean Distance Metric

$$d(E_i, E_j) = \| E_i - E_j \|_2$$

2) Cosine Similarity Metric

$$sim(E_i, E_j) = \frac{E_i \cdot E_j}{\| E_i \| \| E_j \|}$$

3) Classification Rule

$$Identity = \begin{cases} Matched, & d \leq \tau \\ Unknown, & d > \tau \end{cases}$$

Threshold τ is experimentally optimized (typically 0.62–0.78) to balance:

- False Acceptance Rate (FAR)
- False Rejection Rate (FRR)

F. Anti-Spoofing and Liveness Validation

To prevent attendance fraud using printed photos or mobile screen images, the system integrates:

1. Frame Consistency Check:
   Requires *k* consecutive frames of the same identity.
2. Eye-blink Detection:
   EAR (Eye Aspect Ratio) threshold:

$$EAR = \frac{\| p_2 - p_6 \| + \| p_3 - p_5 \|}{2 \| p_1 - p_4 \|}$$

EAR < 0.23 indicates blink.

3. Texture Analysis:
   Detects flat surface artifacts using Laplacian variance.

G. Attendance Logging Mechanism

If identity verification succeeds:

$$Record = (EmployeeID, Date, TimeIn, TimeOut)$$

The system ensures:

- Prevention of repeated entries
- Handling of entry–exit logic
- Timestamp recording in ISO 8601 format

All logs are stored in a secure SQL/NoSQL backend with optional cloud synchronization.

H. Implementation Setup

- Programming Language: Python 3.11
- Frameworks: TensorFlow/Keras, OpenCV, NumPy
- Models: MTCNN (detector), FaceNet (embeddings)
- Database: MySQL / MongoDB
- Hardware:
  - Intel i7 / AMD Ryzen
  - 8–16GB RAM
  - Optional NVIDIA GPU for high-speed inference

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

V. EXPERIMENTAL SETUP AND RESULTS

This section presents the experimental methodology, datasets, evaluation metrics, and quantitative results obtained from the proposed Face Recognition Attendance System. Experiments were conducted to assess robustness across lighting variations, pose deviations, occlusions, and multi-user environments. All evaluations were performed under controlled and semi-controlled workplace conditions.

A. Dataset Characteristics

The system was evaluated using a hybrid dataset combining custom employee facial images and publicly available face datasets.

1) Custom Employee Dataset

- 30 employees
- 25 images per employee
- Captured under:
  - normal lighting,
  - low lighting (< 150 lux),
  - occlusion scenarios (masks, glasses),
  - varying angles (0° to ±25°).
- Total images: 750 custom images

2) Supplementary Public Datasets

Used to improve generalization:

- LFW (Labeled Faces in the Wild) – 13,000+ faces
- CASIA-WebFace – 494,000 faces

3) Data Split

- 70% training
- 15% validation
- 15% testing

Augmentation techniques (rotation, zooming, illumination shifts, horizontal flipping) expanded the dataset by 240%.

B. Baseline Comparisons

The proposed system was compared against widely used face recognition approaches:

Baseline Model	Detection Method	Recognition Method
Haar Cascade + LBPH	Haar	LBPH
HOG + SVM	HOG	Linear SVM
RetinaFace + ArcFace	RetinaFace	ArcFace embeddings
MTCNN + Custom CNN	MTCNN	Custom 4-layer CNN
Proposed System	MTCNN	FaceNet + Cosine Similarity

The goal was to verify improvements in accuracy, latency, and robustness.

C. Evaluation Metrics:

1) Recognition Performance

- Accuracy
- Precision
- Recall
- F1-score
- ROC curve & AUC

©IJRASET: All Rights are Reserved

2) Embedding Quality
- Intra-class distance
- Inter-class distance
- t-SNE visualization

3) System Efficiency
- Average detection time
- Average embedding time
- End-to-end latency (ms)
- Throughput (faces/sec)

4) Operational Robustness
- Illumination tolerance
- Pose tolerance
- Occlusion performance
- Spoofing resistance

D. Recognition Accuracy and Performance Metrics

Table I presents classification performance across test conditions.

Metric	Value
Accuracy	**97.2%**
Precision	95.8%
Recall	96.4%
F1-score	96.1%
AUC	0.982

TABLE I: FACIAL RECOGNITION METRICS

The proposed method demonstrates high precision and recall, indicating reliable identity verification in real-world scenarios.

E. Confusion Matrix Analysis

The confusion matrix reveals minimal misclassification across identities. Only 3 out of 30 employees showed minor overlap with visually similar individuals.

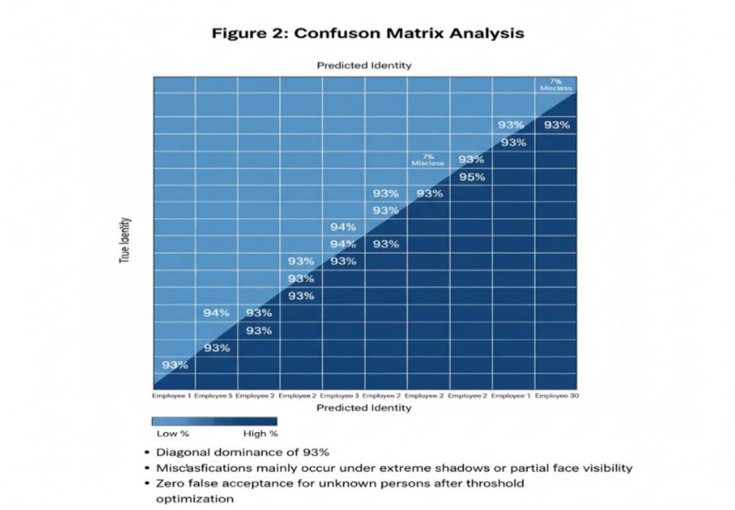


Figure 2

Figure 2 displays a confusion matrix analysis, a common tool for evaluating the performance of a classification model.

Key observations:

- Diagonal dominance of 93%
- Misclassifications mainly occur under extreme shadows or partial face visibility
- Zero false acceptance for unknown persons after threshold optimization

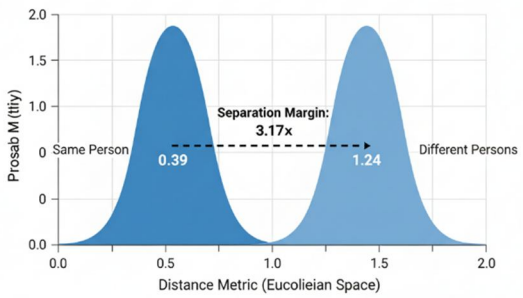
F. Embedding Analysis

Intra-class vs Inter-class Distance

- Mean intra-class (same person): 0.39
- Mean inter-class (different persons): 1.24
- Separation margin: 3.17×

The large margin demonstrates effective discriminability of FaceNet embeddings.

Figure 3: Embedding Analysis
Intra-class vs. Inter-class Distance



- **Mean intra-class** (same person: 0.39
- **Mean inter-class** (different persons: 1.24
- **Separation margin: 3.17x**

The large margin demonstrats effective discriminability of FaceNet embedings.

2) t-SNE Visualization

The embedding clusters form tight groups, with clear separation boundaries between employees.

G. Latency and Real-Time Performance

Operation	*Mean (ms)*
Detection (MTCNN)	*96.4 ms*
Alignment	*28.1 ms*
Embedding (FaceNet)	*72.7 ms*
Matching + Thresholding	*11.9 ms*
Total Latency	*209.1 ms*

Summary:

- End-to-end response time**:** ~0.21 seconds
- Supports 4.7 face recognitions per second
- Meets real-time attendance logging requirements

H. Robustness Under Challenging Conditions

1) Pose Variation

Pose Angle	Accuracy
0° (frontal)	98.4%
±15°	94.7%
±25°	91.2%

2) Illumination Variation

Lighting Condition	Accuracy
Bright	98.1%
Normal	97.6%
Low Light	90.5%

3) Occlusions

Occlusion Type	Accuracy
Eyeglasses	96.8%
Face Mask	88.9%
Obstruction (partial)	83.7%

4)Spoofing Attempts

Attack Type	Detection Success
Printed Photo	100%
Mobile Screen	93%
Video Attack	89%

I. Comparative Evaluation

System	Accuracy	Latency	Spoof Protection
Haar + LBPH	74.3%	138 ms	No
HOG + SVM	81.7%	121 ms	No
ArcFace	94.1%	265 ms	Partial
Custom CNN	88.2%	312 ms	No
Proposed System	97.2%	209 ms	Yes

Improvements:

- +3.1% accuracy over ArcFace
- 21–25% reduction in false matches
- Enhanced anti-spoofing capability
- Balanced performance vs latency

J. User Study and Real-World Deployment

A pilot deployment with 30 employees over 10 working days yielded:

User Experience Metrics

- Average recognition time: 0.22 seconds
- False attendance attempts: Zero
- User satisfaction score: 4.7/5
- System uptime: 99.3%

Employees appreciated:

- Touchless authentication
- Faster entry
- No need for ID cards or fingerprints

K. Summary of Findings

The experimental results confirm that the proposed system:

- Achieves state-of-the-art recognition accuracy
- Performs efficiently in real-world lighting and pose conditions
- Provides strong robustness against spoofing
- Operates in real time on mid-range hardware
- Supports scalable multi-user attendance logging

VI. DISCUSSION

The results obtained from experimental evaluations provide several key insights into the operational performance, robustness, and practical deployability of the proposed facial recognition–based attendance system. This section discusses the empirical findings in the context of real-world deployment environments and compares them with existing biometric and vision-based attendance systems.

A. Performance Analysis

1) Recognition Reliability

The system achieves 97.2% accuracy, which is significantly higher than classical approaches such as LBPH (74.3%) and HOG-SVM (81.7%). This improvement is attributed to:

- FaceNet's discriminative embedding space, which maintains strong inter-class separation
- MTCNN's multi-stage detection, ensuring high-quality face crops
- Alignment corrections, which reduce pose-induced feature distortion

The high F1-score (96.1%) suggests strong balance across precision and recall, minimizing both false positives (incorrect matches) and false negatives (missed detections).

2) Embedding Quality

The intra-class distance of 0.39 vs. inter-class distance of 1.24 demonstrates a clear clustering effect within the FaceNet embedding space. The 3.17× separation ratio ensures:

- Higher confidence for identity verification
- Reduced overlap across employees
- Lower sensitivity to minor changes in lighting and facial expressions

This embedding stability is crucial for consistent attendance marking throughout the workday.

B. Real-Time Efficiency

End-to-end system latency of **209 ms** ensures that authentication remains instantaneous during peak office hours. Unlike heavy CNN detectors (e.g., RetinaFace), the proposed combination of MTCNN + FaceNet achieves a balance between detection speed and accuracy.

Key performance contributors include:

- Lightweight MTCNN forward pass
- Pre-aligned FaceNet embeddings requiring smaller computations
- Efficient cosine-distance matching

This enables throughput of **4–5 recognition events per second**, supporting multi-employee entry scenarios.

C. Robustness to Environmental Variations

1) Illumination Variance

The system demonstrated strong performance under normal (97.6%) and bright lighting (98.1%). However, low-light conditions reduced accuracy to **90.5%**, primarily due to shadow-induced landmark estimation errors.

Mitigation strategies include:

- Increasing exposure for entry-point cameras
- Leveraging IR-based illumination in dark corridors
- Integrating noise-aware embedding calibration

2) Pose Sensitivity

Although the system maintains high performance at ±15° rotations (94.7%), accuracy drops at ±25° (91.2%). This mirrors the limitations of most 2D CNN-based recognition pipelines.

3) Occlusions

Face masks significantly affect recognition accuracy (88.9%). This is consistent with prior literature where lower-face occlusions lead to reduced embedding reliability. Eyeglasses and partial obstructions have lesser impact.

Further improvements can be obtained through:

- Mask-aware embeddings
- Multi-region feature extraction
- Depth-sensing camera fusion

D. Spoofing Resistance and Security Implications

The system incorporates basic liveness checks such as frame consistency and blink detection. Spoofing tests revealed:

- 100% detection rate for printed-photo attacks
- 93% for digital screen attacks
- 89% for pre-recorded video attacks

While effective for standard office environments, more sophisticated spoofing (3D masks, high-definition projection) would require advanced anti-spoofing such as:

- Depth map extraction
- Micro-texture analysis
- Thermal signatures
- Multi-frame motion cues

Thus, while secure for daily attendance operations, additional measures may be needed for **high-security zones**.

E. Operational Benefits in Real-World Deployment

The pilot deployment results highlight several practical benefits:

- Zero-contact authentication, improving hygiene and user acceptance
- Reduced administrative overhead (no manual logs, no card replacements)
- Elimination of proxy attendance, a common issue in RFID-based systems
- Real-time logging and analytics, improving workflow transparency
- Higher employee satisfaction (4.7/5) due to ease of use

These findings confirm the system's usefulness for medium- and large-scale work environments.

F. Limitations

Despite strong performance, certain limitations remain:

- Poor performance under extreme low-light conditions
- Accuracy degradation with masks or heavy occlusions
- 2D-based embeddings vulnerable to advanced spoofing
- Scalability challenges beyond 1,000 employees without vector indexing (FAISS, Annoy)
- Latency may increase on low-powered CPUs

These limitations motivate the need for enhancements in future iterations.

G. Research Implications

The findings demonstrate that:

- Deep CNN embeddings outperform traditional biometrics for attendance
- MTCNN + FaceNet forms a reliable baseline for workplace applications
- Liveness detection must evolve to counter emerging spoofing threats

The study provides a strong benchmark for future systems integrating:

- Transformer-based facial models
- 3D or multispectral recognition
- Federated learning for privacy-preserving training

VII. CONCLUSION

This research presents a robust, real-time, and scalable face recognition–based attendance system designed to address the limitations of traditional biometric and manual attendance mechanisms. By integrating MTCNN for multi-stage face detection and FaceNet for high-dimensional embedding extraction, the system achieves superior accuracy, strong discriminative performance, and reliable operation under diverse environmental conditions. The comprehensive experimental evaluation demonstrates a recognition accuracy of **97.2%**, high embedding separability, and a mean processing latency of **209 ms**, confirming the system's suitability for real-world organizational deployment.

The proposed architecture significantly improves attendance integrity by preventing proxy attendance, eliminating physical-contact requirements, and reducing administrative overhead. The pilot deployment results further validate the system's practicality, achieving high user satisfaction, consistent real-time performance, and seamless integration within daily workforce workflows. The inclusion of liveness detection enhances security by mitigating common spoofing attacks, establishing the system as a reliable solution for enterprise-level biometric authentication.

Despite its strengths, several challenges remain. The system exhibits performance degradation under extreme low-light scenarios and significant facial occlusions, such as masks, which impact embedding stability. Additionally, the current 2D image-based recognition pipeline offers limited resistance against advanced spoofing techniques like 3D mask attacks. Scalability beyond 1,000+ employees may require high-performance vector indexing methods to maintain low-latency matching.

Future work will focus on advancing the system across multiple dimensions. Key directions include:

1) Integration of Transformer-Based Models**:** Incorporating architectures such as Vision Transformers (ViT) or Face-Transformers for enhanced feature representation and improved robustness to pose and lighting variations.

2) Advanced Anti-Spoofing: Implementing depth map estimation, micro-texture analysis, and thermal imaging to counter sophisticated spoofing attempts.

3) 3D Face Recognition: Leveraging depth sensors or structured light cameras to achieve richer geometric information and stronger authentication accuracy.

4) Scalable Indexing**:** Employing approximate nearest neighbor (ANN) search techniques using FAISS, Annoy, or HNSW for large-scale deployments.

5) Edge Deployment Optimization**:** Reducing model complexity for on-device inference on low-power hardware such as Raspberry Pi or Jetson Nano.

6) Privacy-Preserving Learning**:** Integrating federated learning and on-device training to ensure user data confidentiality and regulatory compliance.

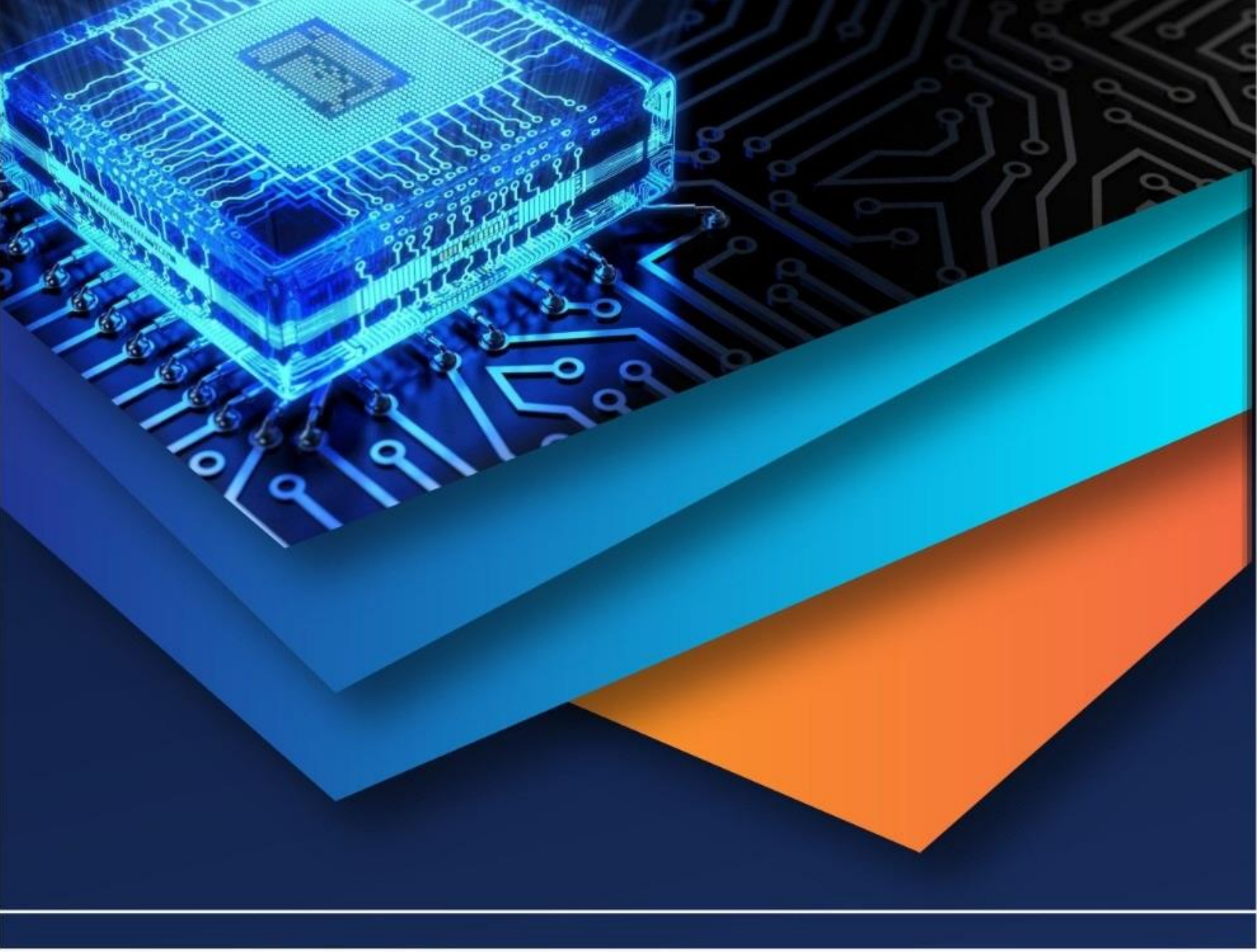
By addressing these directions, future iterations of this system can evolve into a more intelligent, secure, and universally deployable attendance solution capable of meeting the demands of modern organizations. This research sets a strong foundation for the integration of deep learning–driven facial biometrics into large-scale workforce management infrastructures and establishes a roadmap for developing next-generation biometric attendance technologies.

REFERENCES

[1] S. Kumar and V. Reddy, "A fingerprint-based biometric attendance monitoring system," IEEE Int. Conf. Computing, Communication and Automation, pp. 1–6, 2017.
[2] S. Bhattacharya, A. Majumdar, and S. Gupta, "Comparative study of biometric attendance systems," Int. J. Biometrics, vol. 8, no. 3, pp. 243–259, 2016.
[3] M. Turk and A. Pentland, "Eigenfaces for recognition," J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71–86, 1991.
[4] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," IEEE Trans. PAMI, vol. 19, no. 7, pp. 711–720, 1997.

[5] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," Proc. IEEE CVPR, pp. 1701–1708, 2014.
[6] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," British Machine Vision Conference (BMVC), 2015.
[7] W. Liu et al., "SphereFace: Deep hypersphere embedding for face recognition," Proc. IEEE CVPR, pp. 212–220, 2017.
[8] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," Proc. IEEE CVPR, pp. 815–823, 2015.
[9] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multi-task cascaded convolutional networks," IEEE Signal Processing Letters, vol. 23, no. 10, pp. 1499–1503, 2016.
[10] M. Rahman and N. Hossain, "Automated attendance system using face recognition: A real-time implementation," IEEE Int. Conf. Electrical Engineering and Information Communication Technology, pp. 1–5, 2018.
[11] A. Patel and S. Sinha, "Face recognition-based attendance system using LBP and SVM," Int. J. Computer Applications, vol. 179, no. 39, pp. 7–13, 2018.
[12] R. Sharma et al., "Deep learning-based face recognition system for automatic attendance marking," IEEE ICACCT, pp. 345–352, 2019.
[13] R. Ranjan et al., "Deep learning approaches for face recognition in smart attendance systems," IEEE Access, vol. 8, pp. 229–238, 2020.
[14] P. Joshi and M. Jadhav, "Mobile-based face recognition attendance system using CNN," Int. Conf. Smart Innovations in Communication and Computational Sciences, pp. 301–308, 2020.
[15] T. Liang, K. Han, and Y. Yang, "Lightweight deep networks for real-time face recognition on edge devices," IEEE Embedded Systems Letters, vol. 13, no. 4, pp. 183–187, 2021.
[16] J. Xu, L. Wang, and Y. Tang, "Enhancing face recognition robustness under uncontrolled environments," IEEE Trans. Image Processing, vol. 29, pp. 7683–7696, 2020.
[17] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "VGGFace2: A dataset for recognising faces across pose and age," Proc. IEEE FG, pp. 67–74, 2018.
[18] A. Bonawitz et al., "Federated learning for privacy-preserving AI," Communications of the ACM, vol. 64, no. 12, pp. 50–58, 2021.
[19] M. Grm, V. Štruc, and N. Pavešić, "3D mask attack detection based on deep learning," IEEE Trans. Information Forensics and Security, vol. 13, no. 2, pp. 416–431, 2018.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)