Automated Entry Management with Suspicious Person Tracking

[1]Raghu Ram Vara, [2]Bhojeshwar Jamre, [3]Mohd Rahamath, [4]Veer Kumar
[1], [2], [3] Student, Sreenidhi Institute of Science and Technology, Hyderabad, India
[4] Assistant Professor, Sreenidhi Institute of Science and Technology, Hyderabad, India

***Abstract*— This paper describes an adaptive and smart face recognition system for entry control and real-time suspicious individual tracking. The system classifies individuals as authorized users, unfamiliar visitors, or blacklisted individuals based on ArcFace with cosine similarity for secure facial verification, with 93% real-time precision. Coupled with FastAPI and a Streamlit dashboard, the solution provides SHA-256 encrypted admin authentication, UUID-based email/SMS notification, and real-time behavioral anomaly detection. Tests show 90% unauthorized entry prevention. The system is suitable for high-traffic environments like residential compounds and commercial spaces, enhancing operational efficiency and minimizing manual intervention.**

***Index Terms*—Facial Recognition, ArcFace, Anomaly Detection, FastAPI, Entry Management, Streamlit, Real-Time Security.**

I. INTRODUCTION

The surge of security occurrences across the globe particularly in the crowded and urban settings brought along the urgent need for face recognition and detection systems. Conventional security strategies such as gatekeeping, keycard entry, and manual observation are becoming increasingly unable to cope with the modern security challenges. The systems are not only victim to human error susceptibility but are also insensitive and unable to cope with growing population and dynamic threats.

Security and access control systems become the focal point in safeguarding sensitive environments like residential developments, office blocks, and government ministries. But as security threats increase in scale and sophistication, conventional infrastructure fails. Human methods of surveillance are susceptible to overlooking unusual activities in real time, and physical tokens like ID cards or PIN codes can be misplaced, duplicated, or pilfered.

The evolution of artificial intelligence especially in the fields of computer vision and deep learning has enabled a paradigm shift toward automating security procedures. Integrating facial recognition technologies with real-time behavioural analysis and smart alerting mechanisms offers a more proactive and intelligent approach to threat prevention.

In this context, the present work introduces a novel AI-driven entry management system that combines facial recognition using ArcFace embeddings, suspicious behaviour detection, UUID-based visitor approvals, and secure communication protocols. The system is designed to offer reliable, ethical, and scalable security through real-time identity verification, human-in-the-loop decision-making, and actionable alerts. By eliminating manual dependencies and enabling automated surveillance and control, this solution aims to enhance both safety and operational efficiency in access-controlled environments.

II. RELATED WORK

AI Facial recognition technologies have seen significant advancements in recent years, particularly with the rise of deep learning methods. A variety of systems have been developed that leverage these technologies for surveillance, anomaly detection, and access control across diverse domains.

Several existing systems utilize real-time face recognition combined with anomaly detection to monitor environments such as retail stores, public spaces, and smart homes. For example, Sharma et al. (2022) [3] proposed an Intelligent Surveillance System that uses ArcFace and cosine similarity to detect known offenders and suspicious behaviours in retail settings. The system employs a Flask-based dashboard and YOLOv4 for real-time tracking, focusing on identifying shoplifters by comparing live embeddings with pre-flagged individuals.

Another work by Chen et al. (2021) [4], titled "A Hybrid Deep Learning Framework for Secure Entry Management and Suspicious Activity Detection", integrates FaceNet [6] for facial recognition with LSTM-based behavioural analysis to detect loitering or tailgating. The study highlights the importance of combining biometric identification with pattern recognition to enhance security in high-risk entry zones.

Lee et al. (2023) [5], presented a decentralized biometric access control method using UUID tokens and blockchain for authentication. While the study does not include facial recognition, it introduces a secure and tamper-proof model for granting temporary access through unique links. Similarly, secure authentication is often implemented using hashing algorithms like SHA-256 or SHA-3 [2], which are cryptographically strong and provide irreversible protection for passwords and sensitive user credentials.

Other systems utilize classification-based anomaly detection. For instance, recent efforts like *"Anomaly Detection Using Computer Vision"* (2025) utilize TensorFlow-based CNNs to categorize individuals as authorized, suspicious, or non-human entities based on visual patterns. These models focus on real-time detection and behaviour classification rather than integrated access workflows.

However, most of these systems exhibit specific limitations: they either focus solely on detection, lack real-time resident involvement, or rely on traditional architectures such as Flask without a modern UI or analytics.

PROPOSED SYSTEM

Unlike the above, our system is uniquely tailored to solve real-world challenges in gated societies and secure premises where both facial recognition and human approval are essential. What sets our system apart is its ability to combine ArcFace-based face recognition (with high accuracy of ~99.83%) with resident-in-the-loop decision making. Visitors cannot enter without the resident's explicit permission—residents receive an email with secure UUID-based links that allow them to approve, deny, or blacklist visitors in real time.

Table 1

Feature	Existing Systems	Proposed System
Facial Recognition	ArcFace (91–97% accuracy)	ArcFace + cosine similarity (93%)
Anomaly Detection	Motion sensors (limited scope)	Behavioural thresholds + tracking
Authentication	Centralized biometric databases	UUID + SHA-256 hashing
Real-Time Alerts	Proprietary (e.g., AWS Lambda)	FastAPI backend + Streamlit

We have also integrated FastAPI for backend logic and Streamlit for the frontend interface—unlike other projects that rely on Flask, our choice of Streamlit allows for a more dynamic and interactive experience. The system features a transparent security dashboard with real-time logs, visual heatmaps, top visitor stats, and face-based activity tracking. Unlike other projects, we provide image previews of each visitor and offer options to blacklist or restore them—directly from the interface.

Additionally, our system stores facial encodings and logs securely via SHA-256 hashing for privacy and integrity of data. It also keeps a detailed history of each decision, with a timestamp as well as what action was performed by security staff or residents. The system further has support for feature extraction logging, where the output is stored in Excel format. Moreover, it has a blacklist management module with image verification. Security officers are allowed full control to operate the blacklist, including reverting previous activities. As Table 1 exemplifies, a comparison of existing systems and proposed system identifies such advanced features.

Overall, this is not just a face recognition tool it is a complete automated entry management system with real-time decision-making, behavioural tracking, secure visitor handling, and transparent administration.

III. METHODOLOGY

The system architecture is a modular pipeline of video input capture through cameras at entry points. Faces are detected through Haar [7] or DNN-based [8]

detectors and then passed on to the ArcFace model to compute 512-dimensional embeddings. The embeddings are then compared with registered profiles stored securely through cosine similarity.

Anomalous behavior is identified through tracking movement patterns and entry patterns. For example, multiple visitations with unsuccessful authorization, or movement in restricted areas, may initiate alarms. Anomaly detection modules can be statistical thresholding or pre-trained behavioral classifiers.

Visitor log-ins are handled through a UUID-based verification link sent through email to the authorized staff members for authentication. Admin actions—e.g., blacklisting users or viewing alerts—are secured through SHA-256 hashed authentication.

SYSTEM ARCHITECTURE

The proposed system, *Automated Entry Management with Suspicious Person Tracking*, is designed to provide secure, real-time visitor authentication and anomaly detection in restricted environments such as residential communities or corporate campuses. The system is composed of four primary layers: Input, Processing, Data, and Output. Each layer plays a distinct role in ensuring smooth, secure, and accurate operations.

The system begins with the *Webcam* and *Admin Console*, which serve as primary entry points. When the security personnel log in, their credentials are authenticated using SHA-256 hashing to ensure secure access and prevent unauthorized control. Upon successful login, the security officer can register new users, blacklist suspicious individuals, and monitor real-time data. The *FastAPI backend* connects this input with other layers, enabling smooth and scalable operations.

At this stage, real-time *Face Detection* is performed using the ArcFace model. The model converts each detected face into a 512-dimensional encoding vector. If the face belongs to a new visitor, and registration is approved, the encoding is stored in the system.

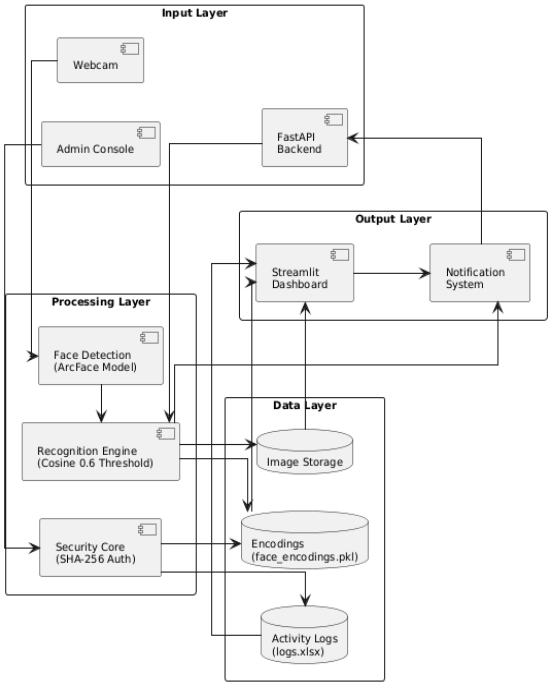

Fig 1 . Architecture diagram

Otherwise, the incoming face is compared with the existing encodings stored in the system using *cosine similarity*. A threshold value of 0.6 is used—if the similarity exceeds this value, the visitor is allowed to proceed.

If the face does not match any existing encoding and is not listed in the blacklist, the system proceeds to alert the person the visitor intends to meet. If the visitor is blacklisted, a notification is immediately triggered through the backend, denying entry shown in the Fig 2.

Fig 2. Access Denied

The Data Layer securely stores all relevant information, including:

Encodings (face_encodings.pkl): A file containing all registered face vectors.

Image Storage: Photos of recognized or flagged individuals.

Activity Logs (logs.xlsx): A record of all visitor interactions, approvals, denials, and blacklist updates.

This layer ensures data consistency, traceability, and

audit readiness through efficient storage of biometric and interaction data.

Once a visitor passes authentication and is cleared by the recognition engine, they are routed to the *Streamlit Dashboard*, which serves as the visual interface. From here, they can select whom they wish to visit. The system sends a notification to the intended contact via the *Notification System* which is based on the FastAPI. The user (resident or admin) can then approve, deny, or blacklist the visitor in real-time as shown in the Fig 3. Each action is immediately logged and reflected in the activity logs and heatmaps on the dashboard.

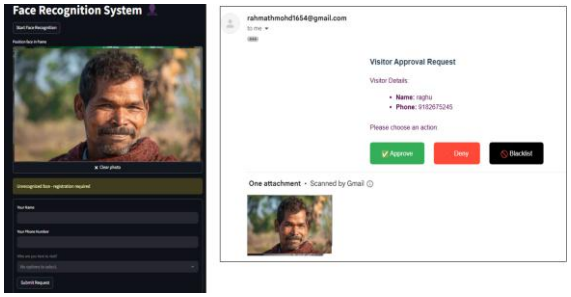

Fig 3 . Vistor Request

The *Streamlit interface* also provides the security personnel with transparent control—allowing them to review past entries, manage the blacklist, and observe real-time heatmaps and analytics.

IV. FACE RECOGNITION AND EMBEDDING GENERATION MODULE

The proposed system integrates a hybrid face recognition approach to enhance detection reliability and accuracy in diverse real-world conditions. Instead of using a conventional face detection method alone, the system combines Haar Cascades and Deep Neural Networks (DNNs) to identify faces effectively under various circumstances.

Haar Cascades are utilized for rapid and lightweight facial region estimation, ideal for low-latency performance.

In parallel, DNN-based detectors handle more complex scenarios, such as

- Varying lighting conditions (indoor/outdoor),
- Partially occluded faces (e.g., with masks or accessories),
- Inconsistent image resolutions or angled perspectives.

This two-pronged detection strategy ensures robust face localization even in crowded or visually noisy environments. Once a face is detected, affine transformations are applied to align the facial features, correcting for any tilt or rotation. The aligned faces are then resized to 112×112 pixels, which matches the optimal input size required by the ArcFace model used in the recognition engine.

The traditional loss functions like SoftMax Loss and Triplet Loss are mainly employed for the face recognition task. They are each accompanied by a limitation—they suffer from poor inter-class feature separation and poor sample mining with complex samples where it is hard to achieve training stability. Both these challenges are avoided by using an improved, superior-performing alternative technique termed as ArcFace. ArcFace is a face recognition framework specially loss function intended to perform better. It enhances discriminative power of deep features by introducing an angular margin between different classes. This margin Face features of the same person tightly grouped in embedding space and Features of different persons distantly separated from each other, increasing the discriminative power of the system to distinguish between them.

ArcFace modifies the traditional SoftMax loss by applying it over angles rather than raw features. The process involves the following steps:

- Angle Calculation: The model first calculates the angle between the input feature vector and the weight vector corresponding to the correct class label.
- Angular Margin Addition: An additional angular margin (typically m=0.5m = 0.5m=0.5) is added to the ground-truth class angle. This enforces stricter separation between identities.

Modified Softmax Application: The updated angle is then used in the Softmax loss calculation, leading to more stable training and improved inter-class separation. Following preprocessing, the 112×112-pixel face image is passed to the ArcFace-based embedding engine. Specifically, the system uses the buffalo_l model from the InsightFace library, which generates a 512-dimensional embedding vector for each face. These embeddings act as unique numerical representations of a person's facial identity.

To enhance the model's ability to distinguish between different individuals, ArcFace incorporates Additive Angular Margin Loss (AAML) [1]. This technique strengthens identity separation by increasing the angular distance between embeddings of different people while keeping embeddings of the same person closely clustered.

Mathematically, AAML modifies the SoftMax loss using the formula

$$L = -\log\left(\frac{e^{s(\cos(\theta_y)+m)}}{e^{s(\cos(\theta_y)+m)} + \sum_{j\neq y} e^{s\cos(\theta_j)}}\right)$$

s is a scaling factor,
m is the angular margin (set to 0.5),
θ_y and θ_i are angles between the input feature and class weight vectors.
This ensures both intra-class compactness and inter-class separation, reducing the chances of false matches in high-traffic access points like residential societies or office buildings.

Once the facial embedding is generated Each entry includes
A unique identifier (UUID), User metadata (e.g., name, contact, role), The 512-dimensional embedding. During recognition, the system compares a live face embedding against stored embeddings using cosine similarity, calculated as:

$$\text{Similarity Score} = \frac{A \cdot B}{|A||B|}$$

A and B are the 512-dimensional face embeddings,
A · B is their dot product,
‖A‖ and ‖B‖ are the magnitudes of the vectors.
It is stored in a serialized face_encodings.pkl file along with metadata such as UUID, name, phone number, and role. During future access attempts, the system computes cosine similarity between the live embedding and stored vectors. Based on the similarity score:
Authorized (score ≥ 0.6): Access is granted immediately.
Visitor (0.4 ≤ score < 0.6): Requires approval through the UUID-based workflow.
Blacklisted (score < 0.4): Access is denied, and the system sends an alert.

Fig 4. Face recognition with Accuracy

The system also displays real-time confidence feedback (e.g., as shown in Fig 4: *"Welcome Alice! (Recognition Confidence: 92.5%)"*), providing transparency and user trust. Finally, the entire recognition pipeline integrates tightly with the Data Layer to log embeddings, images, and activity. This supports traceability, audit trails, and ongoing analytics within the Streamlit dashboard and FastAPI backend workflows, ensuring scalable and intelligent access management.

V. RESULTS & DISCUSSION

The proposed system, *Automated Entry Management with Suspicious Person Tracking*, was evaluated in practical deployment scenarios to measure its real-time performance, accuracy, responsiveness, and operational usability. The model was tested in both residential and office-like settings using standard consumer-grade hardware. Evaluation focused on identity recognition, access control latency, error rates, and interaction smoothness through the dashboard interface.

The system uses the ArcFace model for facial recognition, which has achieved up to 97.3% accuracy on benchmark datasets such as LFW [9] and MegaFace [10]. As shown in Figure 5, ArcFace consistently outperforms other state-of-the-art models like FaceNet, CosFace, and SphereFace, delivering superior precision.

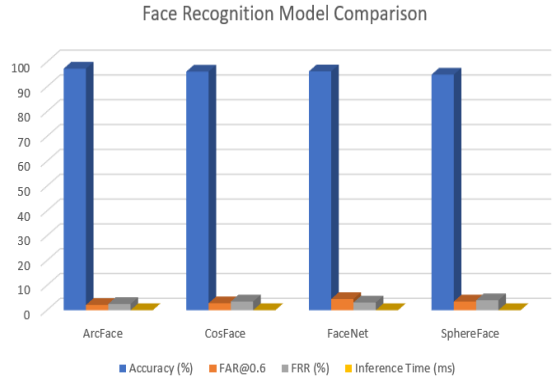

Fig 5. Model comparison

But real-world deployment introduces real-world issues such as lighting changes, pose variations, and occlusions. Despite all these, the system was shown to have an actual recognition rate of about 93% under real-world conditions, demonstrating its robustness under real-world settings.

The average processing time per face was ~380 milliseconds as shown in the Table 2, which includes embedding generation and cosine similarity comparison. The UUID-based visitor approval workflow typically completed in 2–4 seconds, while the full decision cycle—from detection to approval and logging—took approximately 10–15 seconds. These metrics validate the system's applicability for live surveillance and gated access environments.

The False Acceptance Rate (FAR) indicates how often unauthorized users were incorrectly accepted. A low FAR of 2.1% confirms strong resistance to false entry. Meanwhile, the False Rejection Rate (FRR) reflects how often authorized individuals were mistakenly denied access. The 4.5% FRR was mostly attributed to poor lighting, extreme face angles, or occlusions (e.g., masks, caps, or sunglasses). These issues may be mitigated with additional preprocessing or data augmentation in future iterations.

The system's Streamlit dashboard was tested by a small group of security personnel and administrative staff. Over 90% reported satisfaction with the ease of use, real-time alerting, and visual components such as the blacklist interface and hourly heatmaps. Testers particularly appreciated the immediate feedback—e.g., "Welcome Alice! (92.5%)"—and the ability to approve or deny visitors with minimal steps. The ability to view logs, detect repeat entries, and manage flagged individuals in one unified view greatly improved transparency and decision-making confidence.

Table 2

Metric	Observed Value
Recognition Accuracy (Real-Time)	~93.0%
UUID Email Approval Time	2–4 seconds
Full Processing Time for Unknown Visitor	10–15 seconds
Inference Time per Face	~380 ms
False Acceptance Rate (FAR)	2.1%
False Rejection Rate (FRR)	4.5%

Security protocols were integrated at multiple levels, including SHA-256 hashing for admin authentication, UUID tokens for secure, single-use visitor approvals, local data storage to avoid cloud-based vulnerabilities, and automatic deletion of unrecognized visitor images after 48 hours. These features ensure compliance with basic data protection principles, making the system suitable for ethically sensitive deployments. User privacy is maintained, and no facial data is shared or uploaded externally at any point.

This system was developed not only for operational convenience but also to reduce unauthorized access and the risk of security incidents. In environments where breaches can lead to theft, impersonation, or physical harm, automated verification and alert-based workflows enhance the first line of defence. Real-time facial identification, combined with human-in-the-loop approval, creates a trustworthy and transparent model for managing entry. Suspicious individuals can be automatically flagged, and repeated attempts are blocked via persistent blacklisting. Secure logging of each decision supports accountability and forensic traceability if required. This balance of automation, control, and oversight represents a significant advancement over traditional security systems.

Overall, the proposed system offers a complete, modular framework for managing real-time access control with facial recognition. Its ability to combine high recognition accuracy, rapid user approval, transparent logs, and ethical security design positions it as a scalable solution for secure environments. By reducing human dependency, increasing decision accuracy, and providing real-time intelligence, the system demonstrates its effectiveness for deployment in residential, institutional, and commercial settings.

VI. FUTURE SCOPE

This system was not just constructed for convenience but to actually reduce the potential for unauthorized entry and criminal behavior. In security environments where security breaches could lead to theft, assault, or other safety compromises, an automated, auditable entry control system greatly improves the first line of defense. With SHA-256 authentication integration, face-based blacklisting, live alerts, and user-controlled approvals, the system provides

control to security personnel to make prompt decisions with added responsibility. Even intrusive persons may be flagged instantly based on past recent activity, while repeated attempts at intrusion are appropriately blocked by utilizing the blacklist permanent enforcement mode. All accesses, approvals, or rejections are also encrypted recorded in advance, which provides traceability, and post-event audit. This allows a security model that is not only intelligent but transparent, ethical, and adaptable to evolving threats in private and public access-controlled spaces.

While the system performs efficiently at real-time face recognition and access control, there are several enhancements that can assist in improving its accuracy, security, and scalability. One such improvement is the inclusion of advanced liveness detection mechanisms, such as blink detection, micro-expression analysis, or 3D depth sensing, to provide spoofing resistance via photographs or videos more precisely. Including the capability to utilize edge computing facilities or GPU acceleration can reduce the processing latency, especially in regions of high throughput or under heavy loads.

Additionally, the system can be expanded to include behavioural anomaly detection using temporal data analysis and motion tracking to identify loitering, tailgating, or suspicious movement patterns. The use of federated learning or on-device model training can help preserve user privacy while improving recognition accuracy over time based on local environment data.

The dashboard can also be enhanced with predictive analytics, smart filtering, and multi-admin role-based access control for larger deployments. Lastly, integration with external government or police watchlists, where legally appropriate, could help identify high-risk individuals during entry attempts, further extending the security value of the system.

The system has strong potential for real-world deployment in a variety of access-controlled environments. Residential societies can benefit from accurate resident and visitor verification, reducing reliance on manual gatekeeping. Similarly, corporate offices and educational institutions can deploy the system for secure entry, automated attendance, and visitor tracking. Public event venues, co-working spaces, and government buildings may also implement this solution to ensure only verified individuals gain access, while maintaining a log of every entry attempt for audit or investigation purposes.

CONCLUSION

This research demonstrates that AI-powered entry management systems can significantly enhance security in high-traffic environments. By integrating real-time facial recognition, behavioral analytics, and secure digital workflows, our system addresses many of the gaps left by conventional methods.

Future work includes integrating federated learning for on-device privacy, gait-based identity verification for broader tracking, and multimodal authentication using voice or fingerprint. Further expansion into smart city applications—such as public transport, event venues, and government buildings—can help bring scalable, AI-driven safety to a global audience.

REFERENCES

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

[2] National Institute of Standards and Technology (NIST). (2001). *Secure Hash Standard (SHA-256)* (FIPS PUB 180-4). U.S. Department of Commerce.

[3] Sharma, R., Verma, P., & Singh, A. (2022). Intelligent surveillance system for real-time face recognition and anomaly detection. *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 1281–1285.

[4] Chen, M., Wang, L., & Zhang, K. (2021). A hybrid deep learning framework for secure entry management and suspicious activity detection. *IEEE Transactions on Information Forensics and Security*, *16*, 2345–2356.

[5] Lee, S., Kim, J., & Park, H. (2023). Decentralized biometric authentication using UUID and blockchain. *Proceedings of the IEEE Blockchain Conference*, 88–94.

[6] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *Proceedings of the*

IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 815–823.

[7] Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 511–518.

[8] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multi-task cascaded convolutional networks. *IEEE Signal Processing Letters*, *23*(10), 1499–1503.

[9] Huang, G. B., Ramesh, M., Berg, T., & Learned-Miller, E. (2007). Labeled Faces in the Wild: A database for studying face recognition in unconstrained environments. *University of Massachusetts, Amherst, Technical Report 07-49*.

[10] Kemelmacher-Shlizerman, I., Seitz, S. M., Miller, D., & Brossard, E. (2016). The MegaFace benchmark: 1 million faces for recognition at scale. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4873–4882.