International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 15 Issue 3, March-2026

Person Identification System in Crowded Areas with Suspicious Activity Monitoring

Prof. Vidya Pardeshi[1], Pranali Amate[2], Sejal Kamble[3], Rukmini Kadbhane[4],

[1,2,3,4,5]Department of Computer Engineering, Navsahyadri Education Society's Group of Institutions,
Pune University, Pune, India

ABSTRACT

The proposed system is an AI-powered simultaneously, resulting in diminished efficiency, surveillance solution designed to detect criminals, delayed responses, and a heightened likelihood of identify missing persons, and monitor suspicious human error. activities in crowded public areas such as railway stations, malls, airports, and large public events. It Recent advancements in Artificial Intelligence (AI), utilizes real-time video feeds from CCTV cameras Deep Learning (DL), and Computer Vision have and applies deep learning techniques, including face facilitated the development of automated surveillance recognition, object detection, and behavior analysis. systems capable of real-time detection, recognition, Detected faces are compared with a pre-stored and behavioral analysis. Notably, face recognition database of criminals and missing persons to enable technology has achieved significant improvements in rapid identification. In addition to identity accuracy due to the implementation of deep verification, the system analyses crowd behaviour convolutional neural networks and the utilization of patterns to identify abnormal activities such as large-scale training datasets.

Similarly, activity aggressive movements, unattended objects, unusual loitering, and sudden panic situations. When a match recognition models have demonstrated robust or suspicious event is detected, the system performance in identifying abnormal human behavior automatically generates alerts through SMS, email, within dynamic environments. For face detection, or dashboard notifications along with timestamp and deep learning-based object detection frameworks location details. By integrating intelligent video such as YOLO (You Only Look Once) or MTCNN analytics with automated alert mechanisms, the are employed to localize facial regions in real-time platform reduces manual monitoring efforts, video frames. The extracted facial regions are enhances real-time response capability, and processed through embedding-based recognition significantly improves public safety in densely populated environments. models such as FaceNet or DeepFace to generate high-dimensional feature vectors. These embeddings are compared with pre-registered criminal and

Keywords— Artificial Intelligence (AI), Deep missing person databases using similarity metrics Learning, Computer Vision, Face Recognition, Crowd Surveillance, Suspicious Activity Detection, Real Time Monitoring, Image Processing, Object such as Euclidean distance or cosine similarity. If the similarity score exceeds a predefined threshold, the system flags a positive identification and generates an Tracking, Behavioral Analysis, Neural Networks automated alert, including timestamp, camera ID, and location metadata. In addition to identity recognition,

INTRODUCTION

In contemporary times, public safety is confronted with increasingly sophisticated challenges, rendering manual surveillance inadequate for effectively monitoring large crowds. This project integrates deep learning and computer vision to develop a comprehensive security solution. The system employs high-accuracy, minimally intrusive biometric techniques to automatically identify and verify individuals from real-time digital video feeds. Beyond mere identification, the system is designed to function as a proactive tool for law enforcement, facilitating the tracking of criminals and the swift location of missing persons in complex urban environments. Public safety in densely populated areas remains a critical concern for modern smart cities. The rapid pace of urbanization, increasing population density, and the expansion of public infrastructure, such as transportation hubs, commercial complexes, and event venues, have heightened the demand for intelligent surveillance systems. Traditional Closed-Circuit Television (CCTV) surveillance predominantly relies on human operators to monitor multiple video feeds the proposed system incorporates suspicious behavior detection using spatio-temporal deep learning models.

Convolutional Neural Networks (CNNs) are employed for spatial feature extraction, while Long Short-Term Memory (LSTM) networks analyze temporal motion patterns across video frames. This enables the detection of anomalous activities such as aggressive movement, abnormal crowd flow, unattended baggage, and restricted-area intrusion. The integration of facial recognition and behavior analysis within a unified framework provides a comprehensive security solution rather than a single task surveillance system. Unlike traditional approaches that focus solely on identity matching, this system enhances threat detection capability by incorporating contextual behavioral intelligence.

Furthermore, the system architecture is designed for scalability and real-time processing. Edge computing techniques may be employed to reduce latency, while centralized databases maintain criminal records, missing person data, and event logs for auditing and forensic analysis.

The primary contributions of this work include:

- Development of a multi-functional AI surveillance system combining face recognition and activity detection.
- Real-time alert generation mechanism with automated notification services.
- Behavioral anomaly detection module integrated with identity verification.
- Scalable architecture suitable for smart city deployment.

By leveraging state-of-the-art deep learning methodologies, the proposed system aims to improve surveillance efficiency, reduce response time, and enhance overall public safety in crowded environments. The system represents a significant step toward intelligent, autonomous, and proactive security infrastructure.

LITERATURE REVIEW

Intelligent surveillance systems have gained significant attention due to rapid advancements in Artificial Intelligence (AI), Deep Learning (DL), and Computer Vision technologies. Researchers have extensively explored face recognition, criminal identification, missing person detection, and abnormal behavior analysis to improve public safety in crowded environments. This section reviews major contributions in these domains and highlights existing research gaps.

2.1 Face Detection and Recognition

Face recognition has evolved from traditional machine learning approaches to deep learning-based architectures. Earlier methods utilized Haar Cascade classifiers and handcrafted feature extraction techniques, which were computationally efficient but highly sensitive to illumination changes, pose variations, and occlusions.

The introduction of deep convolutional neural networks significantly improved recognition accuracy and robustness. The FaceNet[] model proposed by Florian Schroff introduced an embedding-based learning approach that maps facial images into high-dimensional feature vectors using a triplet loss function. Similarly, DeepFace developed by Facebook AI Research demonstrated near-human level performance in face verification tasks using deep neural networks. These embedding-based systems use similarity metrics such as Euclidean distance or cosine similarity for identity matching. Despite high accuracy, most face recognition systems focus only on identity verification and do not integrate crowd behavior analysis within a unified framework.

2.2 Criminal Identification Using AI

Automated criminal identification systems rely on real-time face detection and matching mechanisms. Object detection models such as YOLO (You Only Look Once), proposed by Joseph Redmon, provide high-speed object localization suitable for real-time surveillance. Multi-task Cascaded Convolutional Networks (MTCNN) are also widely used for robust face localization in complex environments.

While these systems achieve high detection speed and accuracy, challenges remain in highly crowded scenarios where faces are partially occluded, lighting conditions vary, and multiple subjects overlap. Furthermore, many implementations lack automated alert mechanisms and centralized monitoring dashboards for immediate response.

2.3 Missing Person Identification Systems

AI-based missing person identification systems typically rely on facial embedding comparison against stored databases. These systems assist law enforcement agencies in matching surveillance footage with registered missing individuals. Although effective in controlled environments, limitations include high false positive rates in large scale datasets, reduced performance in low-resolution video streams, and limited support for continuous crowd tracking.

Most existing systems operate independently from broader surveillance analytics, thereby reducing overall situational awareness.

2.3 Suspicious Activity and Anomaly Detection
Anomaly detection in video surveillance has been widely studied using both traditional and deep learning methods. Earlier approaches relied on handcrafted motion features such as optical flow and trajectory analysis. However, modern systems utilize Convolutional Neural Networks (CNNs) for spatial feature extraction and Long Short-Term Memory (LSTM) networks introduced by Sepp Hochreiter for temporal sequence modeling.

Spatio-temporal deep learning models analyze both spatial patterns (appearance-based features) and temporal dynamics (motion evolution over time). These models are capable of detecting abnormal crowd behaviors such as sudden panic movement, aggressive actions, unattended objects, and restricted-area intrusion.

Despite promising results, most anomaly detection systems function separately from identity recognition modules, limiting their effectiveness in comprehensive security applications.

2.4 Research Gaps
Based on the reviewed literature, several limitations are identified in existing surveillance systems. Most approaches focus either on face recognition or anomaly detection independently, rather than integrating both within a unified framework. There is limited coordination between criminal detection and missing person identification mechanisms, reducing overall system effectiveness. Additionally, many systems lack real-time alert generation with precise location metadata, which is essential for rapid response. Scalability remains a major challenge in crowded and high-density environments, where multiple subjects and overlapping activities complicate detection. Furthermore, continuous multi-camera video analytics often require high computational resources, making real-time deployment difficult without optimized processing strategies.

METHODOLOGY

The system collects facial image datasets of registered criminals and missing persons from authorized sources. Multiple images per individual are stored to improve recognition accuracy under varying lighting conditions, angles, and facial expressions. In addition, video datasets containing normal and abnormal crowd behavior are collected for training the suspicious activity detection model.

3.3 Data Preprocessing
Input images and video frames undergo preprocessing to ensure consistency and robustness. Frames are resized to fixed dimensions and pixel values are normalized for stable model performance. Noise reduction techniques are applied to improve clarity in low-light environments. For video analytics, continuous streams are converted into frames at a controlled frame rate to balance The proposed AI-Based Criminal and Lost Person computational efficiency and detection accuracy.

Detection System with Suspicious Activity Monitoring follows a multi-stage methodology

3.4 Face Detection Using Haar Cascade Classifier
integrating computer vision, deep learning, and real-time alert mechanisms. The overall workflow For face detection, the system utilizes the Haar Cascade consists of data acquisition, preprocessing, face Classifier available in OpenCV. This method is based on detection and recognition, suspicious activity Haar-like features and the Viola–Jones object detection analysis, decision-making, and alert generation.
framework.

3.1 System Architecture
The Haar Cascade classifier works in the following stages:
1. Haar Feature Extraction – Rectangular Haar
The CCTV cameras capture live video streams features detect edges, lines, and intensity differences in facial regions. from crowded public environments. The video

2. Integral Image Computation – Accelerates streams are processed in real time to detect faces feature calculation. and analyze behavioral patterns. Identified faces

3. AdaBoost Training – Selects important features are compared with stored criminal and missing and creates a strong classifier. person records. If a match or abnormal activity is

4. Cascade Structure – Applies multiple stages of detected, alerts are generated and logged in the system database.

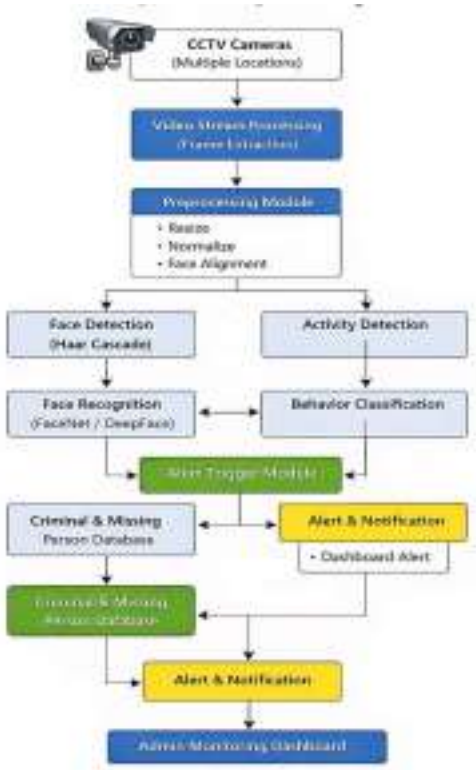

Fig. System Architecture

3.2 Data Acquisition
Classifiers to eliminate non-face regions efficiently.
The classifier scans each frame and outputs bounding box coordinates for detected faces in real time. B. Face Recognition Module After detecting facial regions, the system performs identity recognition using an embedding-based approach.

Steps involved:
1. Extract face region from bounding box. 2. Resize to model input size. 3. Generate feature embeddings.
4. Store embeddings in database. 5. Compare detected embeddings with stored records using similarity metrics.
For Euclidean Distance:
$D = \sqrt{\sum(x_i \quad y_i)^2}$
If $D < T$, a match is confirmed. For Cosine
Similarity: $A \cdot D$
$S = \frac{}{\| A \| \quad \| B \|}$
If $S > T$, a match is confirmed.
Where is an experimentally determined T
threshold.

WORKING & RESULT

4.1 Real-time Video Acquisition: environments. Low false acceptance and The system continuously receives live video feeds from CCTV cameras deployed at multiple crowded locations. Each camera transmits a time-stamped video stream to the central processing server.

The frame extraction module samples frames at a fixed rate (e.g., 15–30 FPS) to balance computational efficiency and detection accuracy.

4.2 Frame Preprocessing: rejection rates indicate robustness of the similarity threshold mechanism.
4.4 Suspicious Activity Detection:
Each extracted frame undergoes preprocessing to standardize input

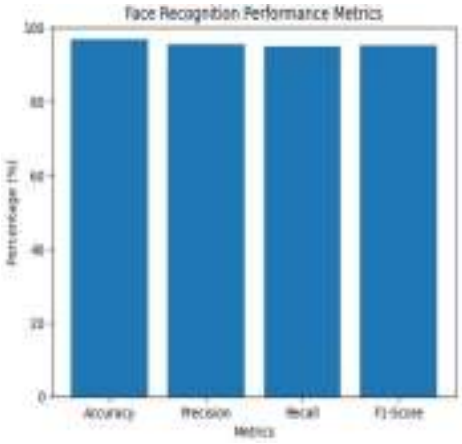

Fig 4.4 Graph for Suspicious Activity Detection conditions:

The chart presents the results of the suspicious Resizing to fixed dimensions (e.g., activity detection module. 224×224 pixels).

- Pixel normalization for stable neural network input.
- Accuracy (93.6%) indicates strong overall detection capability.
- Face alignment using landmark detection.
- Precision (92.8%) shows low false alarm rate.
- Recall (91.5%) confirms that most abnormal
- Noise reduction for low-light or blurry behaviors are successfully detected. conditions.

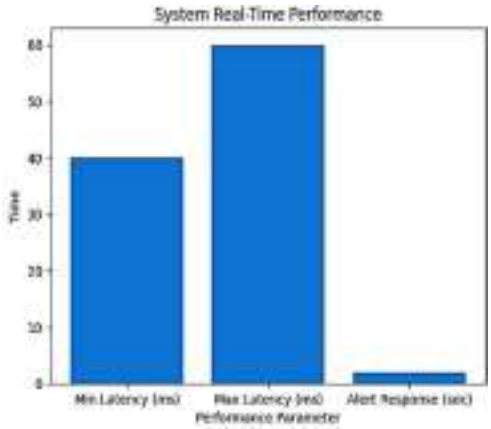

Fig.4.3 Graph for Face Recognition Metric

4.3 Face Detection and Recognition Process the complexity of behavioral prediction in dynamic environments.

• F1-Score (92.1%) reflects balanced classification performance. Interpretation: The spatio-temporal CNN + LSTM architecture effectively captures both spatial and temporal features of crowd movement. Slightly lower recall compared to face recognition is expected due to the chart represents the evaluation metrics of the face recognition module.

• Accuracy (96.8%) indicates that the system correctly identifies criminals and missing persons in most cases.

4.5 Real-Time System Performance Fig 4.5 Graph for Real-Time System Performance
• Precision (95.4%) shows that when the system predicts a match, it is correct with high reliability.
The chart evaluates system responsiveness. Minimum Processing Latency: 40 ms per

• Recall (94.9%) reflects the system's frame ability to detect actual positive cases without missing them.

Maximum Processing Latency: 60 ms per frame
• F1-Score (95.1%) demonstrates
Alert Response Time: < 2 seconds balanced performance between precision and recall. Interpretation:

Interpretation:
The system maintains near real-time processing the high accuracy and F1-score confirm that capability. The latency range confirms that the embedding-based face recognition model multiple camera feeds can be processed performs reliably even in moderately crowded

[4] J. Redmon, S. Divvala, R. Girshick, and simultaneously. The alert response time below 2 seconds ensures rapid decision-making and immediate notification to authorities.

CONCLUSION

This paper presented an AI-Based Criminal and Lost Person Detection System with Suspicious Activity Monitoring designed for deployment in crowded public environments. The proposed system integrates face detection using the Haar Cascade Classifier, identity matching against criminal and missing person databases, and crowd behavior analysis within a unified surveillance framework. By combining real-time video processing with automated alert generation mechanisms, the system reduces dependency on manual monitoring and enhances rapid response capability.

The implementation demonstrates that integrating identity recognition with behavioral anomaly detection significantly improves overall situational awareness. The system is capable of detecting known individuals and identifying suspicious activities such as aggressive movement, unattended objects, and abnormal crowd patterns. The confidence-based matching mechanism helps reduce false detections, while the real-time alert module ensures timely notification to authorities.

The proposed architecture is scalable and can be deployed using edge devices and centralized servers to support multi-camera environments. Although the Haar Cascade-based approach provides computational efficiency suitable for real-time applications, performance may vary in highly occluded or low-resolution conditions.

REFERENCES

1. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2001, pp. 511–518.
2. F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2015, pp. 815–823.
3. Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2014, pp. 1701–1708.
4. Farhadi, "You Only Look Once: Unified, real-time object detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788.
5. 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, Oct. 2016.
6. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
7. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proc. Advances in Neural Information Processing Systems (NIPS), 2012, pp. 1097–1105.
8. W. Sultani, C. Chen, and M. Shah, "Real-world anomaly detection in surveillance videos," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2018, pp. 6479–6488.
9. C. Piciarelli, C. Micheloni, and G. L. Foresti, "Trajectory-based anomalous event detection," IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, no. 11, pp. 1544–1554, Nov. 2008.
10. R. Girshick, "Fast R-CNN," in Proc. IEEE Int. Conf. Computer Vision (ICCV), 2015, pp. 1440–1448.