

# Surveillance System

Abhi Tyagi

*Department of AIML*

*Moradabad Institute Of Technology*

Moradabad, India

[abhityagi23002@gmail.com](mailto:abhityagi23002@gmail.com)

Dr. Neeraj Kumari

*Department of AIML*

*Moradabad Institute Of Technology*

Moradabad, India

[neerajkumari@gmail.com](mailto:neerajkumari@gmail.com)

Ashish Saini

*Department of AIML*

*Moradabad Institute Of Technology*

Moradabad, India

[ashishsaini1310@gmail.com](mailto:ashishsaini1310@gmail.com)

Nikhil Saini

*Department of AIML*

*Moradabad Institute Of Technology*

Moradabad, India

[ns7008080@gmail.com](mailto:ns7008080@gmail.com)

kartik

*Department of AIML*

*Moradabad Institute Of Technology*

Moradabad, India

[kartiksaini89419@gmail.com](mailto:kartiksaini89419@gmail.com)

## ABSTRACT

Surveillance systems have become a critical requirement for ensuring safety, real-time monitoring, and intelligent analysis in public and private environments. Traditional CCTV systems lack automated intelligence and require manual effort to identify individuals, verify attendance, or search for a person in large volumes of video footage. This paper presents a comprehensive review of modern surveillance systems capable of real-time person detection, face recognition, automated attendance marking, and rapid retrieval of a person from stored videos using AI-based techniques. The study analyzes three major categories: computer-vision-based approaches, deep-learning-based biometric recognition, and video analytics algorithms for person search and tracking. A comparative evaluation of contemporary techniques is provided considering accuracy, robustness, computational complexity, and scalability. Findings indicate that hybrid deep-learning models integrating CNN, face embedding, optical flow, and re-identification (Re-ID) networks offer superior performance. The proposed system also integrates GPS-based location mapping to track the last known position of a person, enhancing real-time surveillance and security applications.

**Index Terms**— Surveillance System, Face Recognition, Deep Learning, Automatic Attendance, Person Re-identification (Re-ID), Video Analytics.

## I. INTRODUCTION

Because the cities are getting bigger, and the routine of people has become busier, so the people need surveillance systems that not only just record videos, but also understand what is happening[1, 2]. In the creation of CCTV cameras, they only capture the videos, and someone was watching constantly on the screen, which may lead to delay and missed incidents[4, 5]. Because of these limitations, many organizations decided to move forward, where the system can monitor automatically and react on their own[5, 6].

Recent progress in computer vision machine learning has enhanced the modern surveillance system. It not

only stores video, but also analyzes live footage, identifies object faces, and tracks movement without any manual effort[16, 14]. Due to this, the workload is decreased, and it improves the response time. And because of the lightweight framework like Zask, it has made it much easier with intelligent features. People now can check live streams and time detection models and control systems remotely[13, 6].

The main goal of this project is to make a useful smart surveillance system that brings live video, automatic detection, and clean web interface. And by combining all these parts, we will be able to check live streams, get quick alerts, and reduce manual checking. And it will be easier to stay aware of what is happening in a monitor video. Overall, this will improve security and allow the user to respond faster and more confidently when something unusual happens[20].

This research paper examines literature across three major domains:

1. **Vision-Based Real-Time Person Detection**
2. **Face Recognition and Attendance Automation**
3. **Person Search, Video Retrieval, and GPS-Based Tracking**

The remaining sections follow this structure:

- Section II reviews existing research.
- Section III describes the system methodology.
- Section IV highlights research gaps.
- Section V presents future directions.
- Section VI concludes the paper.

## II. LITERATURE REVIEW

AI-powered surveillance has progressed rapidly over the past decade, shifting from traditional motion-based detection to intelligent video analytics powered by CNNs and transformer-based models. Research falls into three major categories.

### A. Vision-Based Person Detection

Early surveillance systems used classical methods such as Haar Cascades and HOG (Histogram of Oriented Gradients) for person detection [6,7]. While computationally efficient, they struggled with occlusions, illumination changes, and non-frontal faces.

The introduction of deep-learning models revolutionized detection:

- **YOLO (You Only Look Once)** enabled high-speed detection in real time with strong accuracy [13].
- **SSD (Single Shot Detector)** provided a lightweight alternative for embedded systems [6].
- **Faster R-CNN** improved detection quality but required higher computational power [16].

These models detect human figures regardless of pose or orientation, making them ideal for real-time surveillance. However, they do not identify the person; they only detect presence.

### B. Face Recognition and Automated Attendance

Face recognition is the core component for person identity verification. Modern recognition systems rely on deep feature extraction using CNNs. Studies show that embeddings produced by networks such as:

- **FaceNet,**
- **VGGFace,**
- **ArcFace,**
- **ResNet-based Siamese networks**

achieve high accuracy in embedding and matching faces across environments [1, 2, 13, 14].

Attendance automation systems leverage face recognition to automatically mark entry and exit times of individuals in organizations or campuses [1, 5, 20]. Research indicates that face-recognition-based attendance improves time efficiency, eliminates proxy attendance, and integrates well with real-time surveillance.

Challenges include:

- variation in lighting,
- partial face visibility,
- different camera angles,
- fast head movement.

Hybrid systems combining detection + recognition improve performance [4, 6].

### C. Person Re-Identification (Re-ID) and Video Search

Person Re-ID is a rapidly growing research area. It aims to identify the same person across different camera views, resolutions, and environments. Re-ID models use features such as:

- body appearance,
- color histograms,
- gait patterns,
- deep convolutional embeddings.

Studies like Zheng et al. introduced large-scale datasets (Market-1501, DukeMTMC-ReID) to train robust models [17]. Transformer networks and 3D-CNNs capture temporal data for improved video-based Re-ID [18].

Person retrieval across long videos becomes extremely efficient with Re-ID models combined with optical flow tracking [18].

### D. GPS Location Integration in Surveillance

Modern smart cities integrate camera networks with geospatial tagging. Research shows that mapping the detected person's last location significantly reduces search time in public spaces [20]. By storing camera locations, AI systems can instantly show the physical location of the identified person.

Research Area	Method / Technique	Key Findings	Challenges	References
Vision-Based Person Detection	Haar Cascades, HOG	Early detection methods; computationally efficient; detect human presence	Struggle with occlusion, lighting changes, non-frontal faces	[6, 7]
Vision-Based Person Detection	YOLO (You Only Look Once)	High-speed detection in real time; strong accuracy	Computational requirements for high-res video	[13]
Vision-Based Person Detection	SSD (Single Shot Detector)	Lightweight; suitable for embedded systems	Lower accuracy than Faster R-CNN	[6]
Vision-Based Person Detection	Faster R-CNN	Improved detection quality	High computational cost; slower for real-time	[16]
Face Recognition & Attendance	CNN-based Embeddings (FaceNet, VGGFace, ArcFace, ResNet Siamese)	High accuracy in face matching; robust across environments	Lighting variations, partial faces, camera angles	[1, 2, 13, 14]
Face Recognition & Attendance	Attendance Automation Systems	Automates entry/exit marking; reduces manual effort; prevents proxy attendance	Fast head movement, multiple faces in frame	[1, 5, 20]
Person Re-Identification (Re-ID)	Body appearance, color histograms, gait patterns, deep CNN embeddings	Identify same person across multiple cameras; works with varying resolutions	Different lighting, occlusion, background clutter	[17]
Person Re-Identification (Re-ID)	Transformer networks, 3D-CNNs	Capture temporal data for video-based Re-ID; improve retrieval in long videos	Computationally heavy; requires large datasets	[17, 18]
GPS Location Integration	Camera network geospatial tagging	Reduce search time; map physical location of detected person; useful in smart cities	Requires accurate camera/GPS calibration	[20]

**Table 1**

### III. METHODOLOGY

The proposed surveillance system works in several stages to detect, recognize, track, and identify a person in real time. The overall workflow includes video input, human detection, face recognition, attendance marking, person search, and GPS-based location mapping.

#### A. Video Input and Pre-processing

The system begins by capturing real-time video from CCTV or IP cameras. Each frame is processed to improve clarity and speed. Basic operations include resizing the frame for faster processing, converting the image to suitable color format, removing noise or blur in low-light conditions.

This ensures that later stages receive a clean and stable frame.

#### B. Face Recognition and Identity Matching

A deep-learning face recognition model (FaceNet, ArcFace, or ResNet) converts the face into a numerical feature vector.

This vector is compared with stored database vectors to identify the person.

If the distance between vectors is low, the system confirms the identity.

This step enables knowing exactly *who* the person is, recognizing multiple people in one frame, preventing mistakes through threshold-based matching.

#### C. Automated Attendance Marking

- Once a person is recognized:
- A timestamp is recorded (entry/exit).
- Duplicate entries are prevented.
- The attendance is saved in the database automatically.
- This eliminates manual attendance and avoids proxy marking.

#### D. Person Tracking Across Frames

To follow the person continuously, the system uses tracking algorithms such as SORT or DeepSORT.

Tracking helps to maintain the same ID for a person, follow them across the camera view, avoid repeated recognition for every frame.

This makes the system smoother and faster.

#### E. GPS-Based Location Mapping

Every camera is assigned a known GPS coordinate. When a person is recognized, the system automatically records: camera ID, location of that camera, time of appearance. This allows showing the last known location of the person on a map.

#### F. Alerts and Notifications

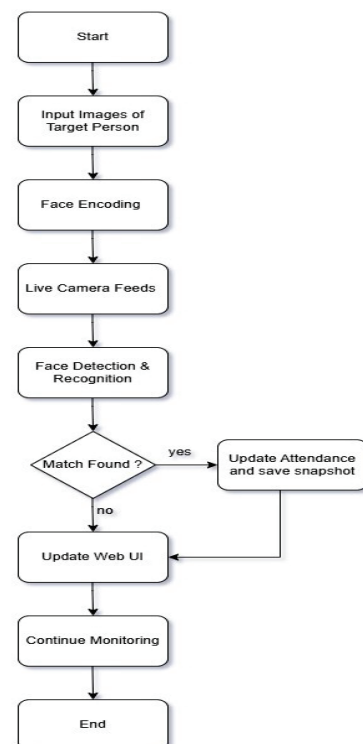
The system can generate alerts when an unknown or unauthorized person appears, a blacklisted person is detected, suspicious movement is observed.

Alerts can be sent to admins through email.

#### G. Database and Storage Management

- The system maintains:
- Face Embedding
- Attendance logs
- Video snapshots
- Person information.

This makes the system easy to manage, update, and integrate with existing security systems.



#### IV. FUTURE RESEARCH DIRECTIONS

- **Stronger night performance:** Will Use IR or thermal cameras to keep detection accurate in low light.
- **Better handling of masks/occlusions:** Use models that will also consider body features or multiple face angles.
- **Improved cross-camera matching:** Advanced Re-ID systems for handling lighting, clothing, and angle changes.
- **Lightweight edge models:** Pruning and quantization help the system to run smoothly on devices like Raspberry Pi.
- **Privacy-focused methods:** Use encrypted face data and on-device processing to keep identities secure.
- **Multi-sensor integration:** Combine CCTV with audio, motion, or IoT sensors for more reliable detection.
- **Real-time behavior analysis:** Detect actions like running, falling, or fights to improve security response.
- **Faster person search:** Use better indexing and distributed processing for large video datasets.
- **Workforce and crowd analytics:** Track attendance, crowd size, queues, and safety compliance automatically.
- **GPS-linked tracking:** Map a person's movement across multiple cameras for clearer location tracking.

#### V. CONCLUSION

The proposed system offers a smart and automated way to detect people, recognize faces, mark attendance, and quickly search for a person in recorded videos. By using modern deep-learning and real-time vision methods, it solves the main problems of old CCTV systems that depend on manual watching.

The combination of person detection, face recognition, and tracking helps the system identify people even in busy or moving scenes. Automatic attendance reduces manual work and removes chances of false entries, making it useful for offices, colleges, and secure places.

The person-search feature makes it easier to find someone in long video recordings, saving a lot of time during investigations. GPS-based camera location also helps by showing where the person was last seen.

Overall, the system improves accuracy, speed, and reliability in monitoring. Although issues like low light, face coverings, and camera quality still exist, future improvements in deep learning and multi-camera processing will make such systems even stronger and more practical for real-world use.

#### REFERENCES

1. **AttenFace (Rao, 2022)** — Real-time attendance captured directly from live camera feeds using facial recognition; useful for fast classroom/institution setups.
2. **Real-Time Video Attendance (Kavana et al., 2024)** — Uses continuous video processing to detect and recognize faces frame-by-frame; matches your “camera + automatic attendance” workflow.
3. **Attendance Capture System (IJERST, 2024)** — Covers the full flow from image capture to detection, recognition, database update; good for explaining system architecture.
4. **Face Recognition Attendance System (IJERT, 2020)** — Basic model demonstrating detection + recognition + attendance marking; good for foundational background.
5. **Face Recognition Attendance System (IJRASET, 2022)** — Discusses practical advantages and limitations of face-based attendance compared to manual or RFID systems.
6. **Smart Attendance Using Facial Recognition (IJARCCE)** — Uses LBPH + Haar cascades for simple, lightweight implementation; suitable for low-power devices.
7. **Survey of Face-based Attendance Systems (ResearchGate)** — Reviews traditional (Haar, LBPH, eigenfaces) and modern CNN-based approaches; helpful for literature comparison.
8. **LBPH vs CNN Attendance System (Budiman, 2023)** — Performance comparison between classic and deep-learning models on real university data.

9. **Review on Face Attendance Systems (JATIT, 2023)** — Systematic literature review summarizing methods, challenges, and accuracy trends.
10. **Automatic Attendance using Face Recognition (IJNRD, 2024)** — Presents a recent real-time, contactless model; good for showing modern relevance.
11. **Python + OpenCV Attendance (Quest Journals, 2020)** — Simple software-only design using Python libraries; useful for implementation section.
12. **Face Detection & Recognition Attendance System (PMU, 2019)** — Describes a video-based pipeline and database structure for attendance logging.
13. **Smart Attendance Using CNN (arXiv, 2020)** — Uses deep CNNs to improve recognition accuracy under varying lighting and angles.
14. **Contactless Attendance Review (IJCA, 2024)** — Reviews algorithms and compares real-time performance; useful for methodology justification.
15. **Attendance System with Simulation (JERR, 2024)** — Combines software with hardware simulation; helpful if your system uses embedded devices.
16. **State-of-the-Art Face Detection/Recognition (Ahmad et al., 2013)** — Classic survey explaining the evolution of face recognition algorithms.
17. **Unconstrained Face Recognition Review (Shepley, 2019)** — Focuses on challenges like pose variation, occlusion, and lighting; important for video surveillance context.
18. **I-VISE Edge Surveillance (Nikouei et al., 2020)** — Shows how surveillance tasks can run on edge devices; useful if using IP or CCTV cameras.
19. **Face Recognition Benchmark Standards (FRGC, FERET)** — Introduces standard datasets used to measure recognition accuracy; helpful if evaluating your model.
20. **Face Recognition via CCTV (Sirisha et al., 2024)** — Attendance taken using existing CCTV cameras; directly aligns with your surveillance + attendance idea.