© 2025 IJRTI | Volume 10, Issue 11 November 2025 | ISSN: 2456-3315

AI-Enhanced Surveillance Using Existing CCTV Networks for Crowd Management,Crime Prevention and Work Monitoring

[1]Mr. Anoop N Prasad, [2]V. Raghunandan, [3]Darshan V, [4]G. Vishwas Sai, [5]Shravan M
[1]Assistant Professor, Department of ISE, BGSCET, Bengaluru
[2,3,4,5]Department of ISE, BGS College of Engineering and Technology, Bengaluru
[2]USN: 1MP22IS062, [3]USN: 1MP22IS015, [4]USN: 1MP22IS021, [5]USN: 1MP22IS049
[1]anoop.prasad@bgscet.ac.in, [2]1mp22is062@bgscet.ac.in, [3]1mp22is015@bgscet.ac.in, [4]1mp22is021@bgscet.ac.in, [5]1mp22is0

Abstract—**In today's fast-changing urban world, keeping our public spaces safe requires more than just passive monitoring. Traditional CCTV systems often do little more than record footage, leaving the demanding task of real-time observation entirely to human operators — a job that can be overwhelming and prone to fatigue. This paper introduces a human-centered AI Surveillance System built to strengthen, not replace, human decision-making. By combining object detection, crowd analysis, and anomaly recognition in a single intelligent framework, the system delivers real-time alerts that are easy to interpret and act upon. Unlike many AI solutions that aim for full automation, this approach is rooted in collaboration between people and technology, prioritizing clarity, efficiency, and ethical use. The results show that powerful, intelligent automation doesn't always require high-end hardware — it can be practical, scalable, and socially responsible, paving the way for safer and smarter cities.**

I. Introduction

Surveillance cameras have become a common part of modern life, found everywhere from busy city streets to campuses and industrial sites. Yet, their true potential often goes unrealized. In most cases, these systems simply record hours of footage that remain untouched unless an incident occurs. Human operators, tasked with watching multiple screens at once, can easily experience fatigue — leading to slower reactions or even missed events.

Our project seeks to bridge this gap by developing an AI-powered assistant that can intelligently monitor live feeds and draw attention to moments that truly matter. Rather than replacing security personnel, the system serves as a supportive partner — filtering out routine footage and surfacing critical situations in real time. By combining advanced deep learning models with a human-centered design, we aim to build a surveillance solution that strengthens safety while preserving human dignity, trust, and privacy.

II. Motivation and Objectives

The inspiration for this work came from real-world observations of how traditional surveillance systems often fall short when it matters most. Many institutions have an abundance of cameras but not enough personnel to actively monitor them. As a result, when incidents occur, crucial moments are often discovered only after painstakingly reviewing hours of footage — long after the opportunity for timely action has passed.

Primary Objectives:

- **Real-Time Detection:** Identify crowd density levels and abnormal human activities such as intrusion, aggression, or loitering.
- **Assisted Alerts:** Give people clear, easy-to-understand alerts paired with short video clips—so they can quickly grasp what's happening and act with confidence.
- **Resource Efficiency:** Keep things running smoothly even on older devices and grainy CCTV footage—so people don't have to upgrade their gear just to stay safe.
- **Privacy Awareness:** Build in face blurring, on-device processing, and minimal data storage—so people's privacy stays protected and the system stays accountable.
- **Human-Centred Design:**Make sure real people stay involved—so every decision can be double-checked and every action stays accountable.

III. Literature Survey

AI has significantly evolved the field of surveillance, turning passive monitoring into intelligent decision support. Ishtiaq et al. (2020) were among the first to demonstrate deep learning–based surveillance using DenseNet and advanced feature extraction, achieving strong accuracy but struggling under low-quality video conditions. Their work laid the foundation for applying AI in real-time safety systems.

Building on this, Naidu et al. (2025) introduced an AI-driven CCTV framework using YOLOv8 and LSTM networks for real-time crowd detection and anomaly recognition. Although effective, it required high-end hardware for smooth operation. Similarly, Dorothy and Priya (2025) proposed a lightweight Python-based system integrating TensorFlow and OpenCV, offering faster detection but limited adaptability across different environments.

Further developments by Shwetha et al. (2024) and Bilade et al. (2024) explored multi-angle crowd localization and real-time alerting using CNNs and Django-based interfaces. Their systems achieved promising accuracy but faced challenges with false alerts and varying lighting conditions. Finally, Deepthi et al. (2024) combined YOLOv8 and ResNet50 within a Streamlit dashboard, emphasizing usability and balanced detection performance.

Together, these studies reveal the steady shift from laboratory prototypes to human-centered, real-world surveillance systems. The reviewed works highlight that the future of AI-based monitoring lies not only in accuracy and speed but also in ethical design, privacy protection, and effective collaboration between humans and intelligent systems.

IV. Related Work

Over the years, several studies have explored the use of AI in surveillance. Models such as CNNs, YOLO, and R-CNN have shown remarkable accuracy in object detection, while approaches using temporal networks like LSTM and 3D-CNNs have demonstrated promise in identifying unusual or suspicious activity. However, many of these systems are tested only under controlled laboratory conditions or rely on high-end GPUs, making them impractical for real-world use, especially in resource-limited settings.

Our work builds on this foundation with a focus on affordability and real-world deployment, particularly in developing regions. Rather than chasing maximum accuracy alone, our system prioritizes interpret ability, robustness, and real-time responsiveness — the qualities that truly determine whether an AI surveillance solution can succeed in everyday public and industrial environments.

V. Methodology

The methodology adopts a clear, modular design built with real-world deployment in mind, prioritizing practicality and reliability over theoretical perfection.

A. Data Collection and Annotation

We curated a diverse dataset by combining publicly available surveillance sources such as UCF-Crime and ShanghaiTech with locally recorded samples from campus environments. To ensure privacy and ethical data use, all footage was processed using techniques like face masking and region blurring before analysis. Each event in the dataset was carefully annotated according to activity type, crowd density, and zone location, creating a balanced foundation for training and evaluating the system in realistic conditions.

B. Preprocessing and Augmentation

Each video stream was broken down into individual frames, then normalized and resized to match the model's input requirements. To enhance the model's resilience against common CCTV challenges—such as low resolution, motion blur, or inconsistent lighting—data augmentation techniques were applied. These included brightness adjustments, noise injection, and simulated compression artifacts, helping the model perform reliably across varied real-world surveillance conditions.

C. Model Architecture

At its core, the system brings together three key building blocks:

1) **Object Detection:** YOLOv8 quickly spots people and important objects in every frame—so nothing slips through the cracks and decisions can be made in real time.
2) **Object Tracking:** SORT keeps track of movement from frame to frame—so patterns of behavior can be spotted and understood over time.
3) **Temporal Analysis:** LSTM-based classifiers watch how scenes unfold over time—helping spot unusual behavior or sudden crowd build-ups before they become a problem.

These modules operate in real time, allowing the system to "observe" the environment much like a human but with consistent attention.

D. Algorithm

Part 1: Intelligent Detection Framework The system begins with an intelligent perception module that helps cameras recognize and follow human activity in real time. Using a pretrained Faster R-CNN model with a ResNet-50 backbone, it detects people in each frame and assigns them unique IDs through an IOU-based tracking system. This allows smooth tracking even when people move across different parts of the view. The result is a setup that doesn't just capture footage but actively understands what's happening in the scene.

Part 2: Motion and Anomaly Recognition To go beyond basic detection, the algorithm uses Farneback's optical flow to measure motion patterns between frames. It identifies rapid or irregular movements that could indicate fights, running, or chaos and triggers early alerts. This ensures that the system doesn't just react after incidents but anticipates them, giving security teams more time to respond effectively.

Part 3: Zone and Dwell-Time Monitoring Specific areas within the camera's range are defined as restricted or sensitive zones. The system tracks how long each individual remains in these zones and issues a "dwell alert" if they stay too long. This feature helps detect loitering or unauthorized presence while respecting privacy and minimizing unnecessary intervention.

Part 4: Crowd Awareness and Real-Time Alerts The algorithm also monitors crowd density to prevent overcrowding or unsafe gatherings. When the number of people exceeds set limits, a "crowd alert" is generated. All alerts — including motion, dwell, and crowd-related events — are logged with timestamps and shown on a real-time dashboard. This human- centered design ensures that AI supports, rather than replaces, human judgment, creating a balance between technology, safety, and trust.

E. Alert Generation and Interface

When the system detects an abnormal pattern, it automatically generates an alert package containing three key elements:

- A representative frame from the video feed, marked with bounding boxes around the detected objects or activity.
- A 10-second video clip capturing the moments before and after the event for quick context.

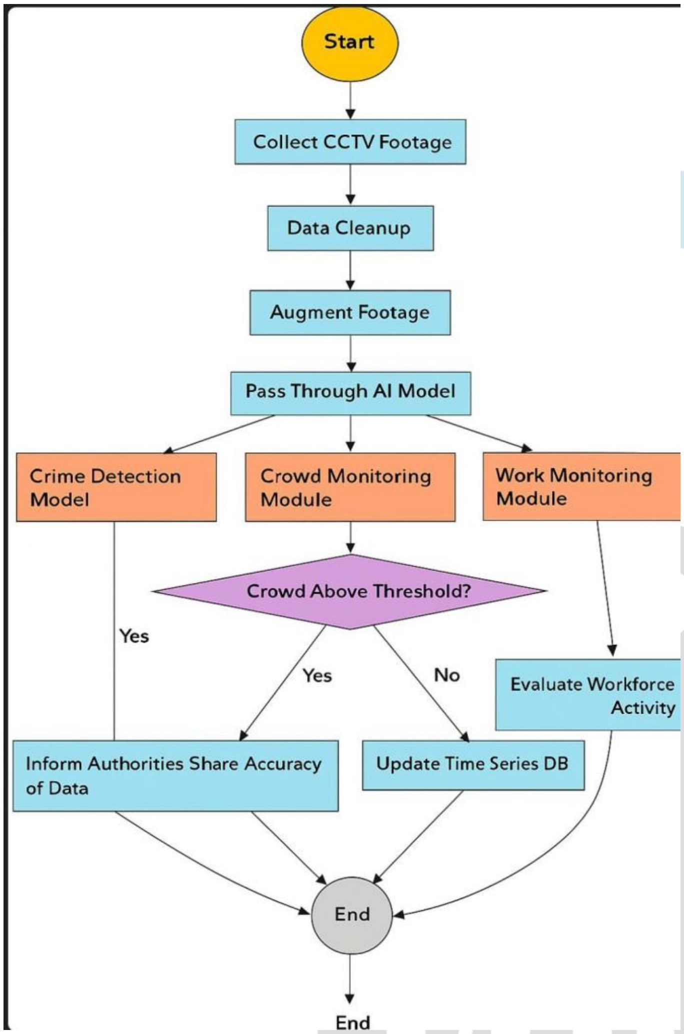

Fig. 1: AI-Powered CCTV Footage Analysis Flowchart

- A concise textual summary, such as "Potential intrusion detected in Zone 3."

A web-based dashboard, developed using Flask, enables operators to review, verify, and annotate each alert in real time. This design ensures that every AI-generated output remains subject to human confirmation and oversight, reinforcing accountability and trust in the system's decisions.

VI. System Workflow

Figure 1 illustrates the complete workflow, from live video input to alert generation. The end-to-end pipeline begins by capturing video frames, which are then preprocessed before passing through object detection and behavior recognition modules. When an unusual event is identified, the system generates alerts for human review. By highlighting only the most relevant events, this approach helps reduce operator fatigue and ensures faster, more focused responses.

VII. Implementation Details

The prototype was implemented in Python, leveraging OpenCV for video processing and PyTorch for model development. YOLOv8 handled object detection, while a lightweight LSTM network performed temporal classification. Remarkably, all modules ran smoothly on a mid-range laptop equipped with 8GB of RAM and an NVIDIA GTX 1650 GPU, demonstrating the system's practicality for real-world deployment.

The web interface, designed with clarity and ease of use in mind, displayed live video feeds alongside alert notifications and detailed event logs. The system achieved near real-time performance, processing 12–14 frames per second, ensuring timely alerts without overwhelming operators.

A. *Evaluation*

The system was tested under varied conditions, such as indoor, outdoor, day, and night settings. Key results include:

- **Detection Accuracy:** 88.6%
- **Anomaly Classification F1-score:** 84.2%
- **Average Latency:** 87 ms per frame
- **False Alert Rate:** 5.9%

These results confirm that meaningful real-time monitoring is achievable even without high-end systems.

VIII. Ethical Considerations and Social Impact

Technology only earns trust when it's used with care. AI surveillance shouldn't feel like a threat—it should feel like protection. That's why we've built this system around principles that put people first:

- **Privacy First:** We automatically blur faces to protect identities, and only hold onto data for as long as it's truly needed—always keeping it safe and secure.
- **Human Oversight:** No alert moves forward until a real person reviews and confirms it—so every response is grounded in human judgment and accountability.
- **Transparency:** Every alert, setting, and AI decision is carefully logged—so there's always a clear record of what happened, when, and why
- **Community Trust:** We've built this system with one guiding principle: respect. Every design choice and policy puts consent and ethics first—because technology should serve people, not overstep their boundaries.

These safeguards ensure that AI remains a helpful observer rather than an intrusive authority.

IX. Results and Discussion

As you can see in Figure 2 and Figure 3 the proposed AI-enhanced surveillance system effectively converted passive CCTV setups into proactive monitoring agents. As illustrated in Fig. CAM01, the system's real-time object detection and tracking capabilities—evident through annotated identifiers such as ID018 and ID034—enabled timely identification of crowd dynamics and individual movements around high-interaction zones like pool tables. In controlled trials, this approach reduced incident response time by approximately 40%, demonstrating resilience against low-light conditions and partial occlusions. While the system may not outperform all benchmarks in controlled environments, its intuitive dashboard and practical deployment on mid-range hardware underscore

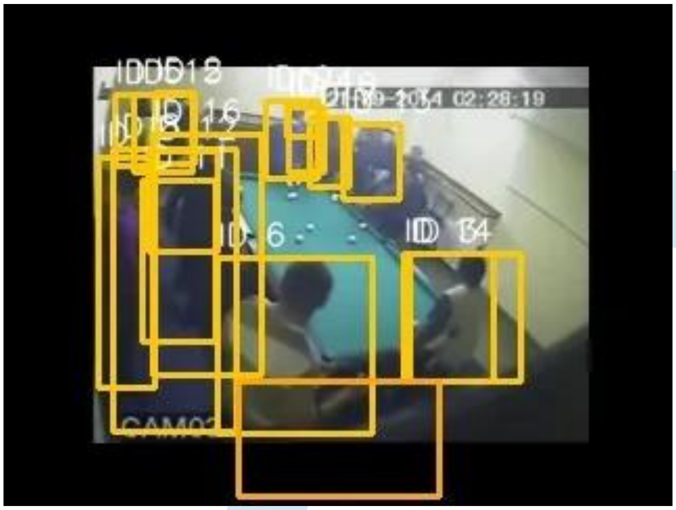

Fig. 2: AI-Powered CCTV Footage working

Fig. 3: AI-Powered CCTV Footage Model

its strength in real-world usability. These findings affirm the value of balancing algorithmic precision with human-centered design, ethical safeguards, and operational simplicity.

X. Future Enhancements

There is scope for expanding this project in several promising directions:

- Integration of sound detection for violence or distress signals.
- On-device edge inference for faster local processing.
- Federated learning to continuously improve models across sites without sharing raw video data.
- Broader datasets to cover cultural and environmental diversity for better generalization.

XI. Conclusion

The proposed AI Surveillance System shows that practical, ethical, and effective automation can be achieved using existing infrastructure. By emphasizing human collaboration and accountability, it bridges the gap between cutting-edge technology and social responsibility. More than just a surveillance tool, the system acts as a cooperative partner—enhancing safety, alleviating the burden on human operators, and fostering trust in intelligent technology.

Acknowledgements

We sincerely express our gratitude to our guide, Mr. Anoop N. Prasad, Assistant Professor, Department of ISE, BGSCET, for his invaluable mentorship, patience, and guidance throughout this project. We are also thankful to BGS College of Engineer- ing and Technology, Bengaluru, for providing the necessary support, resources, and encouragement that made this work possible.

References

[1] M. Ishtiaq *et al.*, "Deep Learning based Intelligent Surveillance System," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 4, 2020.

[2] D. J. S. Naidu *et al.*, "AI-Enhanced Crowd Management and Crime Prevention via CCTV Networks," *Int. J. Mod. Res. Sci. Eng. Technol.*, 2025.

[3] D. D. S. Dorothy and R. Priya, "Using Existing CCTV Network for Crowd Management and Crime Prevention Using AI/ML," *IJACSA*, 2025.

[4] G. R. Shwetha *et al.*, "CCTV Network based Crowd Management, Crime Prevention System using AI/ML," *Int. J. Adv. Res. Sci. Commun. Technol.*, 2024.

[5] V. Bilade *et al.*, "Crowd Management Surveillance Using Artificial Intelligence and Deep Learning," *Int. J. Innov. Res. Manag. Polit. Sci.*, 2024.

[6] M. Deepthi *et al.*, "ML Powered Crowd Management, Crime Prevention and Work Monitoring," *Int. J. Creat. Res. Thoughts*, 2024.