

A Practical Approach to Real-Time Intelligent Surveillance Using YOLOv8 and DeepSORT with Behavior Analysis and Zone-Based Monitoring

BHUVAN M J*Dept of AI&ML**Sri Siddhartha Institute of
Technology**Tumakuru**bhuvanmj40@gmail.com***PREETAM R B***Dept of AI&ML**Sri Siddhartha Institute of
Technology**Tumakuru**preetambailwad09@gmail.
com***C KUMAR***Dept of AI&ML**Sri Siddhartha Institute of
Technology**Tumakuru**kumarsrinivas8792@gmail
.com***Dr. MANJULA YERVA***Professor**Dept of AI&ML**Sri Siddhartha Institute of Technology,
Tumakuru**manjulayerva@ssit.edu.in*

Abstract - Implementation of an AI-Powered Intelligent Surveillance System utilizing real-time video streams, the system integrates advanced deep learning models for human detection, tracking, and security monitoring. The system comprises a detection module using YOLOv8 for accurate person identification and a tracking module using DeepSORT to maintain consistent identities across frames. A polygon-based restricted zone mechanism is implemented to monitor sensitive areas and detect unauthorized entry. The system also incorporates behavior analysis and anomaly detection to identify suspicious activities. Real-time alerts are generated when rule violations occur, and events are logged with timestamps and snapshot evidence for further analysis. The overall system operates in a modular pipeline including preprocessing, detection, tracking, zone evaluation, and alerting. Results demonstrate the system's ability to perform accurate real-time monitoring, reduce manual surveillance effort, and enhance security through automated decision-making, proving its effectiveness in modern intelligent surveillance applications.

Keywords— *Intelligent Surveillance, Human Detection, YOLOv8, DeepSORT Tracking, Restricted Zone Monitoring, Anomaly Detection, Real-Time Alert System, Computer Vision.*

I. INTRODUCTION

The rapid advancement of computer vision and deep learning techniques has significantly improved the capabilities of modern surveillance systems in recent years [1]. Traditional CCTV systems primarily function as passive recording tools and lack the ability to intelligently analyze real-time video streams. As a result, they depend heavily on human operators, leading to inefficiencies, delayed responses, and missed critical events. The integration of artificial intelligence into surveillance has enabled systems to automatically detect, track, and analyze human activities, transforming surveillance from reactive monitoring to proactive security management [2,3,4]. In conventional surveillance systems without intelligent processing, video data is recorded without any understanding of the scene or control over event detection. However, when AI-based models are incorporated, such as object detection, tracking algorithms, and rule-based analysis, the system gains the ability to interpret activities and generate meaningful insights. By integrating additional contextual information like restricted zones, behavioral patterns, and temporal constraints, it becomes possible to guide the system towards identifying suspicious activities and potential threats in real time [5].

Understanding the working of deep learning models such as YOLOv8 for detection and DeepSORT for tracking within a surveillance pipeline can be complex. For beginners in this domain, it is beneficial to start with structured implementations involving real-time video processing and modular pipelines. By working with manageable components such as frame acquisition, preprocessing, detection, tracking, and alerting, developers can better understand system design and optimize performance for real-world applications [6].

In the context of intelligent surveillance systems, it is important to consider that system performance depends on the coordination between multiple modules such as detection, tracking, and rule-based decision engines. While individual modules may perform effectively, overall system accuracy is influenced by environmental factors such as lighting conditions, occlusion, camera angles, and dynamic backgrounds. Therefore, achieving a stable and reliable surveillance system requires careful integration and tuning of all components [7].

Within an AI-powered surveillance system, the detection and tracking modules play a crucial role in identifying and monitoring individuals. The detection model, such as YOLOv8, processes each frame to locate human presence, while the tracking model assigns consistent identities and maintains continuity across frames. These modules work together to capture movement patterns and provide the foundation for higher-level analysis such as behavior recognition and anomaly detection. The effectiveness of the system depends on its ability to accurately detect and continuously track individuals in real time [8].

The system utilizes video input streams from cameras, where each frame is processed and analyzed to extract meaningful information. Techniques such as preprocessing, normalization, and resizing ensure compatibility with deep learning models and improve detection accuracy. The integration of polygon-based restricted zone mapping allows the system to monitor specific areas and detect unauthorized entry instantly. This approach enhances security by enabling immediate response to potential threats [9].

Learning meaningful patterns from continuous video data and identifying abnormal behavior is a key focus of modern surveillance research. AI-powered systems enable automated

monitoring, reduce human workload, and improve response time by generating real-time alerts. Additionally, such systems support event logging, evidence collection, and dashboard visualization for efficient monitoring and investigation. With the increasing demand for intelligent security solutions, AI-based surveillance systems provide a scalable and reliable approach for ensuring safety across various environments such as campuses, offices, and public spaces [10].

II. LITERATURE REVIEW

The development of intelligent surveillance systems has been significantly influenced by advancements in deep learning and computer vision techniques in recent years [1]. Early object detection models such as Faster R-CNN introduced region-based approaches that improved detection accuracy, although they were computationally expensive for real-time applications. This limitation led to the development of single-stage detectors like YOLO, which enabled fast and efficient real-time object detection while maintaining reasonable accuracy, making them suitable for surveillance systems [2].

In addition to detection, tracking plays a crucial role in maintaining identity across video frames. Traditional methods such as SORT achieved high-speed tracking using motion estimation techniques but lacked robustness in handling occlusions and crowded scenes. To overcome these limitations, DeepSORT introduced appearance-based feature extraction along with motion prediction, significantly improving tracking consistency and reducing identity switches in complex environments [3,4].

Further advancements in detection models such as YOLOv4 improved both speed and accuracy through enhanced architectures and training strategies, allowing better performance in challenging conditions such as low lighting and dense crowds [5]. Research on smart video surveillance systems has emphasized the importance of integrating detection, anomaly detection, and alert generation to reduce human dependency and enable automated monitoring [6].

Recent studies have explored the integration of artificial intelligence in public safety systems, focusing on crowd analysis, behavior recognition, and anomaly detection. These systems analyze movement patterns and detect unusual activities such as sudden movement, loitering, or abnormal behavior, improving situational awareness and early threat detection [7,8].

Intrusion detection using deep learning has further enhanced surveillance systems by enabling automated monitoring of restricted zones. These systems use object detection and predefined virtual boundaries to detect unauthorized entry and generate real-time alerts, demonstrating improved robustness under varying environmental conditions [9]. Additionally, intelligent security systems require adaptive AI models capable of handling real-world challenges such as lighting variations, camera instability, and background noise [10].

The introduction of YOLOv8 has further advanced surveillance technology by providing improved detection accuracy, faster inference speed, and better generalization across diverse environments. Its compatibility with tracking algorithms and rule-based systems enables seamless integration into real-time

surveillance pipelines [11]. Furthermore, vision-based alert systems have demonstrated the effectiveness of automated notification mechanisms in reducing response time and enhancing overall system efficiency in security applications [12].

III. METHODOLOGY

The methodology employed in this study revolves around designing and implementing an AI-Powered Intelligent Surveillance System capable of real-time monitoring, detection, and automated decision-making. The primary goal is to achieve accurate human detection, continuous tracking, and instant alert generation while maintaining system efficiency and reliability. Figure 1 represents the working architecture of the proposed system. The system integrates multiple components including detection, tracking, re-identification, and rule-based monitoring. The detection module utilizes YOLOv8 to identify human presence in video frames, while the tracking module uses DeepSORT to maintain identity across frames. These components work together to monitor activities continuously and detect potential security threats. The system operates in a pipeline where video input is processed frame-by-frame, enabling real-time analysis and response. The integration of artificial intelligence allows the system to reduce manual monitoring and improve overall surveillance efficiency.

The approach includes several key components:

- 1. Multi-Source Video Input and Acquisition:**
The system begins by capturing video streams from one or multiple cameras using technologies such as OpenCV, RTSP, or FFmpeg. A multi-threaded architecture ensures that each camera feed operates independently, allowing uninterrupted processing and efficient handling of multiple video sources simultaneously.
- 2. Frame Preprocessing and Data Standardization:**
The captured frames are resized, normalized, and enhanced to ensure compatibility with deep learning models. This preprocessing step improves image clarity and stabilizes performance under varying lighting conditions, enabling accurate detection in real-time environments.
- 3. Real-Time Human Detection and Object Tracking:**
The system utilizes YOLOv8 for real-time human detection and DeepSORT for tracking. YOLOv8 detects individuals in each frame, while DeepSORT assigns unique IDs and maintains identity continuity across frames using motion prediction and appearance features. This combination ensures accurate monitoring even in crowded or dynamic scenes.
- 4. Cross-Camera Re-Identification and Behavior Analysis:**
The re-identification module ensures that the same individual is recognized across multiple camera views using appearance-based embeddings. Behavior analysis is performed using temporal data, allowing the system to identify actions such as loitering, running, or suspicious movement patterns, enhancing situational awareness.
- 5. Zone-Based Monitoring and Intelligent Rule Evaluation:**
Polygon-based restricted zones are defined within the surveillance area. The system continuously checks whether a person enters or remains within these zones using centroid-based calculations. Rule-based conditions such as unauthorized

entry, dwell time, or repeated access attempts are applied to detect violations effectively.

6. Anomaly Detection and Event Logging Mechanism:

The system analyzes movement patterns over time to identify abnormal behavior that deviates from normal activity. When anomalies or rule violations are detected, the event is logged with details such as timestamp, camera ID, tracking ID, and snapshot evidence, ensuring traceability and supporting further investigation.

7. Real-Time Alerting and Dashboard Visualization:

An alerting mechanism is implemented to notify security personnel instantly when critical events occur. Alerts include detailed information and are displayed through a dashboard interface. The dashboard provides real-time video feeds, event logs, and zone management tools, enabling efficient monitoring and decision-making.

8. Modular System Architecture and Layered Design:

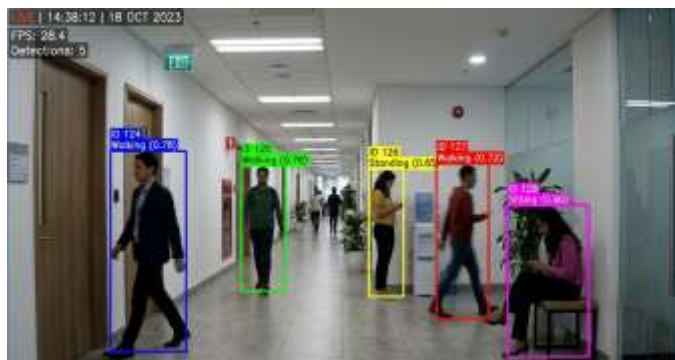
The system is designed as a layered architecture consisting of input acquisition, processing, identity analysis, event management, and visualization layers. Each module operates independently while sharing structured data, ensuring flexibility, scalability, and ease of integration with future enhancements.

9. Component Interaction and System Coordination Pipeline:

All modules interact in a synchronized pipeline where video frames flow from input acquisition to preprocessing, detection, and tracking. The outputs from tracking are utilized by re-identification and rule engines, ensuring consistent identity management and accurate zone monitoring. This interaction enables efficient and continuous surveillance.

10. Structured Data Flow and Processing Pipeline Design:

The system follows a structured data flow where raw video frames are transformed into meaningful information through sequential processing stages. Detection outputs are converted into tracking data, which is further analyzed for identity matching and rule evaluation. The processed data is stored and visualized, providing a complete surveillance workflow.



11. Real-Time Processing and Performance Optimization Techniques:

The process flow is optimized to maintain low latency and high accuracy. Conditional execution is applied where frames without detections are skipped to reduce computation. Parallel processing ensures that detection, tracking, and analysis occur simultaneously, improving system responsiveness and performance.

12. Algorithm Integration and Intelligent Decision Fusion:

The system integrates YOLOv8 for detection, DeepSORT for tracking, ReID models for identity matching, and point-in-polygon algorithms for zone detection. Their outputs are combined using decision fusion to generate accurate alerts with minimal false positives.

13. Implementation Workflow and System Execution Process:

The system follows a structured workflow including initialization, camera setup, frame acquisition, preprocessing, detection, tracking, and rule evaluation. Alerts and event logs are generated upon detecting anomalies, while the dashboard continuously displays system outputs.

14. System Optimization and Scalable Deployment Architecture:

The system uses GPU acceleration, asynchronous processing, and dynamic thresholding to ensure efficient performance. Its modular design supports scalability for multiple cameras and future feature integration.



Figure 1: AI surveillance system architecture

IV. RESULTS

In the initial phase of system execution shown in Figure (2), the surveillance system began processing video streams and detecting human presence using the YOLOv8 model. During this stage, the system focused on identifying basic human shapes and movements within the frame. In early frames, the detection was accurate, but tracking IDs occasionally showed minor inconsistencies due to rapid motion and occlusion. The DeepSORT tracking module gradually improved identity assignment, maintaining consistent tracking IDs across consecutive frames. Simultaneously, the behavior analysis module started classifying basic actions such as walking and standing, although some classifications showed lower confidence scores. The preprocessing layer played a crucial role in stabilizing detection under varying lighting conditions. By the end of this phase, the system demonstrated reliable detection and initial behavior classification, establishing a strong foundation for further refinement in subsequent stages.

Figure 2: Initial Detection and Tracking



Figure 3: Improved Behavior Analysis and Zone Monitoring, Stable and High-Accuracy Surveillance Results.

Between the intermediate phase shown in **Figure (3)**, the system made significant improvements in tracking stability, behavior classification, and restricted zone monitoring. The DeepSORT algorithm consistently maintained identity across frames, reducing ID switches even in crowded environments. The behavior analysis module showed substantial progress, accurately identifying actions such as walking, sitting, falling, and fighting with higher confidence scores displayed above bounding boxes. Abnormal behaviors such as sudden falls and aggressive movements were successfully detected and flagged. The polygon-based restricted zone monitoring system accurately identified unauthorized entry using centroid-based calculations. The alerting module generated real-time notifications with snapshot evidence, ensuring immediate awareness of critical events. By the end of this phase, the system achieved stable performance with accurate detection, tracking, and behavior recognition, demonstrating improved reliability.

In the advanced phase shown in **Fig (3)**, the system achieved consistently high performance across all modules, including detection, tracking, behavior analysis, and alert generation. The system effectively handled multiple individuals simultaneously, maintaining stable tracking IDs and accurate behavior labels with confidence scores. The behavior classification module reached optimal performance, consistently identifying both normal and abnormal activities with high precision. The restricted zone detection system operated reliably, triggering alerts instantly upon intrusion. The event logging module stored detailed metadata, including timestamps, tracking IDs, detected behaviors, and snapshot evidence. The dashboard visualization provided a clear and real-time representation of all activities, enabling efficient monitoring and decision-making. At this stage, the system demonstrated robust and scalable surveillance capabilities.

During the final stage shown in **Figure (3)**, the system reached optimal performance with minimal variations in output quality. Detection, tracking, and behavior classification remained consistent across different scenarios, including varying lighting conditions and crowded environments. The behavior analysis module reliably identified all actions with high confidence, while abnormal activities were detected without delay. The alerting system maintained real-time responsiveness, and the event logging system ensured complete traceability of all incidents. The overall system operated efficiently with low

latency and high accuracy, demonstrating its readiness for real-world deployment. The final stage focused on maintaining stability and ensuring consistent performance, confirming that the surveillance system had achieved a reliable and fully optimized state.

CONCLUSION

The proposed AI-Powered Intelligent Surveillance System successfully demonstrates a robust and scalable solution for real-time security monitoring using advanced deep learning and computer vision techniques. The system effectively integrates human detection using YOLOv8, multi-object tracking using DeepSORT, behavior analysis, restricted zone monitoring, and real-time alert generation into a unified framework. The experimental results confirm that the system is capable of accurately detecting and tracking multiple individuals while maintaining consistent identity across frames, even in dynamic and crowded environments. The inclusion of behavior analysis significantly enhances the system's capability by enabling real-time classification of human activities such as walking, sitting, falling, and fighting. The system also effectively identifies abnormal behaviors and highlights them with confidence scores, improving situational awareness and enabling faster response to critical events. The polygon-based restricted zone monitoring further strengthens the system by detecting unauthorized entry with high precision and triggering instant alerts with supporting evidence.

The alerting and event logging mechanisms ensure that all critical events are recorded with detailed metadata, including timestamps, tracking IDs, detected behaviors, and snapshot evidence. The dashboard interface provides an intuitive and real-time visualization of surveillance activities, allowing security personnel to monitor, analyze, and respond efficiently.

Overall, the system achieves high accuracy, low latency, and reliable performance across different conditions, reducing the dependency on manual monitoring and minimizing the chances of missed incidents. The modular architecture ensures scalability and adaptability, making the system suitable for deployment in real-world environments such as campuses, offices, public spaces, and high-security zones. The results demonstrate that the proposed system provides an effective, intelligent, and automated surveillance solution capable of enhancing security and operational efficiency.

REFERENCES

- [1] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2015.
- [2] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *arXiv preprint*, 2016.
- [3] N. Wojke and A. Bewley, "DeepSORT: Simple Online and Real-Time Tracking," *arXiv preprint*, 2018.
- [4] X. Li, R. Zhao, and L. Li, "Human Tracking Using SORT Algorithm," *International Journal of Computer Vision*, vol. 127, pp. 1499–1513, 2019.
- [5] A. Bochkovskiy, C. Wang, and H. Liao, "YOLOv4: Optimal Speed and Accuracy," *arXiv preprint*, 2020.
- [6] R. Patel, A. Kumar, and S. Singh, "Smart Video Surveillance Using Computer Vision," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, pp. 1231–1244, 2020.

- [7] C. Zhang, Y. Li, and H. Wang, "AI-Based Surveillance in Public Safety," *IEEE Access*, vol. 9, pp. 55612–55623, 2021.
- [8] M. Karthik, K. Ramesh, and V. Suresh, "Deep Learning for Anomaly Detection," *Computers & Security*, vol. 103, 2021.
- [9] A. Singh, R. Sharma, and P. Verma, "Intrusion Detection Using Deep Learning in CCTV," *Journal of Intelligent & Fuzzy Systems*, vol. 42, no. 3, pp. 3151–3162, 2022.
- [10] P. Sundaramoorthy, "AI in Intelligent Security Systems," *International Journal of Security and Networks*, vol. 17, no. 2, pp. 101–114, 2022.
- [11] G. Jocher *et al.*, "YOLOv8: Ultralytics," *arXiv preprint*, 2023.
- [12] Y. Chen, F. Li, X. Zhao, and Q. Wang, "Vision-Based Real-Time Alert Systems," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 4, pp. 2345–2356, 2023.