International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 12 | Dec - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Suspicious Activity Detection Using Surveillance Camera

Mrs. Prakruthi G R[1], Aditya Mallewadi[2], Akash R S[3], Arjun M[4], Goghul T[5]
1Assistant Professor, Dept of ISE, East West Institute Of Technology, Bengaluru
[2,3,4,5] UG Scholar, Dept of ISE, East West Institute Of Technology, Bengaluru

---

Abstract - Video anomaly detection plays a critical role in surveillance, traffic monitoring, and public safety by identifying unusual events such as accidents, theft, violence, or abnormal human activities. Traditional approaches often rely on handcrafted features and statistical models, which struggle to adapt to real-time and complex scenarios. With the advancement of deep learning, object detection algorithms like YOLO (You Only Look Once) have shown remarkable efficiency in detecting and classifying objects with high accuracy and speed. This project implements a video anomaly detection framework using YOLO to process video streams and identify unusual events in real time. The system takes video input, extracts frames, processes them through YOLO for object detection, and classifies anomalies based on deviations from normal activity patterns. The results highlight the potential of YOLO based models in achieving robust, accurate, and real time anomaly detection, thereby enhancing safety and monitoring systems.

Key Words: YOLO, AI, CNNs, RNNs, GANs

1.INTRODUCTION

In recent years, the rapid advancement of computer vision and artificial intelligence (AI) has opened new possibilities for enhancing public safety and security. One of the most significant applications of these technologies is suspicious activity detection using surveillance cameras. With the exponential growth of urban populations and the increasing need for efficient monitoring systems, traditional manual surveillance methods have become insufficient, prone to human fatigue, and limited in scalability. Consequently, there is a growing demand for automated systems that can intelligently analyze video feeds in real time to identify abnormal or suspicious behaviors without continuous human-supervision.

The primary goal of a suspicious activity detection system is to automatically recognize unusual movements, actions, or patterns that may indicate potential threats such as theft, violence, or unauthorized access. By leveraging machine learning, deep learning, and computer vision techniques, these systems can be trained to detect deviations from normal behavioral patterns captured in video footage. Such automation not only enhances the efficiency of surveillance operations but also enables early threat detection, thereby helping prevent crimes before they escalate. This research focuses on developing an intelligent surveillance framework that integrates AI-powered video analytics for detecting suspicious activities in real time. The proposed system processes live video streams, extracts relevant motion and object features, and classifies behaviors based on trained models. By employing algorithms such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs), the system can effectively learn spatial and temporal dynamics within the footage. Furthermore, the implementation of real-time alerts and visualization interfaces enhances situational awareness for security personnel.

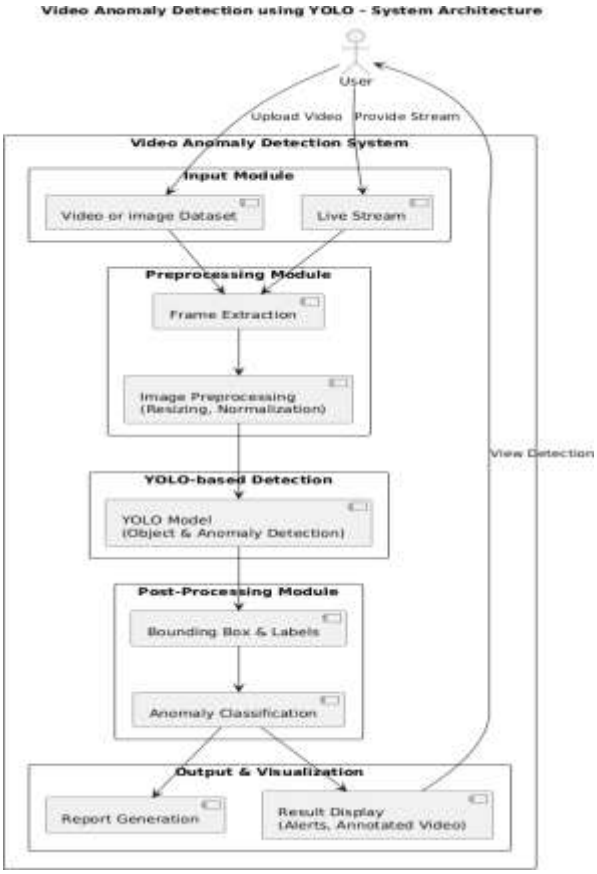
2. Body of Paper

Over the past decade, significant research has been conducted on automated suspicious activity detection through surveillance cameras, driven by advances in artificial intelligence, computer vision, and deep learning. Early approaches relied on traditional image processing techniques such as motion detection, optical flow, and background subtraction to identify anomalies, but these methods were limited by environmental noise, lighting variations, and occlusions. With the emergence of deep learning, models such as Convolutional Neural Networks (CNNs), Autoencoders, and Recurrent Neural Networks (RNNs) have shown remarkable improvements in understanding complex spatio-temporal patterns. Researchers have utilized datasets like UCSD Pedestrian, CUHK Avenue, and UCF-Crime to train and evaluate these models on both scene-specific and real world data. Autoencoder-based methods detect anomalies through reconstruction errors, while GANs (Generative Adversarial Networks) and prediction-based models forecast normal motion patterns to identify deviations as suspicious activities. More recently, transformer-based architectures have enhanced long range temporal reasoning, enabling more accurate and context-aware detection. Despite these advancements, challenges such as real-time implementation, lack of labelled data, domain adaptability, and ambiguity in defining “suspicious behaviour” persist. Consequently, current research continues to focus on developing hybrid, efficient, and generalizable models capable of operating in Module diverse environments with minimal human supervision.

Early approaches relied on traditional image processing techniques such as motion detection, optical flow, and background subtraction to identify anomalies, but these methods were limited by environmental noise, lighting variations, and occlusions. With the emergence of deep learning, models such as Convolutional Neural Networks (CNNs), Autoencoders, and Recurrent Neural Networks (RNNs) have shown remarkable improvements in understanding complex spatio-temporal patterns.

The primary goal of a suspicious activity detection system is to automatically recognize unusual movements, actions, or patterns that may indicate potential threats such as theft, violence, or unauthorized access. By leveraging machine learning, deep learning, and computer vision techniques, these systems can be trained to detect deviations from normal behavioral patterns captured in video footage

3.SYSTEM ARCHITECTURE



A. User Interface

The system begins with the user, who can either upload a pre-recorded video or image dataset or provide a live video stream from surveillance cameras. This input flexibility allows the system to function in both offline (batch) and real-time (online) monitoring modes. The user interface serves as the primary control point to initiate the detection process and later visualize results or alerts.

B. Input Module

The Input Module handles the data acquisition process. It accepts two primary sources: Live Stream: Real-time video input from surveillance cameras. Video or Image Dataset: Pre-collected surveillance footage stored locally or in a database for analysis. This module ensures that the input is properly formatted and accessible for further processing.

C. Preprocessing Module

Before feeding the video data into the detection model, preprocessing is essential to improve efficiency and accuracy. It consists of two main steps: Frame Extraction: The video stream is divided into individual frames (images) at regular intervals. This conversion allows frame-wise analysis since YOLO operates on images rather than continuous video. Image Preprocessing (Resizing and Normalization): Each extracted frame is resized to match the input dimensions required by the YOLO model (commonly 416×416 or 640×640). Pixel values are then normalized to ensure consistent lighting and contrast, reducing noise and improving model performance. This stage prepares clean, standardized input data for detection.

D. YULO-Based Detection

This is the core detection module of the system, where YOLO (You Only Look Once)—a real-time object detection algorithm—is utilized. The YOLO Model processes each frame once, dividing it into a grid and predicting bounding boxes and class probabilities simultaneously. In this system, YOLO is trained or fine-tuned not only for object detection (like people, vehicles, or objects) but also for anomaly detection, such as unusual behaviors, restricted area intrusion, or unattended objects. This module enables fast, accurate detection that is suitable for real-time surveillance applications

E. Post-Preprocessing Module

Once the YOLO model generates detections, the Post Processing Module refines and interprets the results. It includes: Bounding Box and Labels: Each detected object or anomaly is highlighted with bounding boxes and labeled with the detected class (e.g., “person running,” “fighting,” “suspicious object”). Anomaly Classification: The system categorizes detected activities as either normal or suspicious based on predefined behavioral patterns or anomaly confidence scores. This ensures that only relevant or potentially dangerous events are flagged for review. E. Post-Preprocessing Module Once the YOLO model generates detections, the Post Processing Module refines and interprets the results. It includes: Bounding Box and Labels: Each detected object or anomaly is highlighted with bounding boxes and labeled with the detected class (e.g., “person running,” “fighting,” “suspicious object”). Anomaly Classification: The system categorizes detected activities as either normal or suspicious based on predefined behavioral patterns or anomaly confidence scores. This ensures that only relevant or potentially dangerous events are flagged for review.

F. Output And Visualization

The final stage presents the analyzed results in user friendly formats. It consists of two parts: Report Generation: Generates a structured report summarizing detected anomalies, timestamps, object counts, and severity levels. This can be stored for further analysis or auditing. Result Display (Alerts & Annotated Video): Displays visual outputs such as annotated video frames with bounding boxes and real-time alerts for suspicious activities. These alerts can be sent via dashboard notifications, SMS, or email to security personnel. The visualization module enhances situational awareness and supports quick decision-making in security operations.

G. System Feedback / Continuous Monitoring

After results are displayed, the system allows the user to view detections and re-initiate live monitoring, forming a continuous feedback loop. This makes the architecture suitable for both continuous surveillance and offline analytical studies.

METHODOLGY The proposed system for Suspicious Activity Detection using YOLO aims to identify abnormal events in surveillance videos in real time. The methodology consists of five main steps: dataset collection, preprocessing, feature extraction, classification, and result evaluation. Each step contributes to creating an accurate and efficient detection pipeline capable of identifying suspicious behaviors like fights, thefts, or accidents with minimal false alerts.

Step1: Dataset collection

The system uses surveillance or activity-based video datasets containing both normal and abnormal events. Publicly available

datasets such as UCSD Pedestrian, UCF-Crime, Avenue, and ShanghaiTech Campus are used for training and evaluation. These datasets include various scenarios like fighting, running, theft, or crowd anomalies. The data is divided into training, validation, and testing sets to ensure balanced model performance and fair evaluation.

Step2:Preprocessing

Each video is converted into frames for individual analysis. Frames are resized to standard YOLO input dimensions (416×416 or 640×640). Image enhancement techniques such as normalization, noise removal, and contrast adjustment improve visual quality. Data augmentation methods like rotation and flipping increase dataset diversity. All frames are annotated in YOLO format with bounding boxes and class IDs for normal and abnormal activities.

Step3:Feature-Extraction

The YOLO model acts as a feature extractor by dividing each frame into grids and predicting bounding boxes and class probabilities. It captures features such as object size, position, and motion, representing both spatial and temporal information. These features help the system distinguish between normal behaviors and suspicious activities in real-time video streams.

Step4:Classification

Detected objects and motion patterns are classified as normal or anomalous using the trained YOLO model. Abnormal actions like fights or accidents are flagged as anomalies. A threshold filter helps reduce false detections, while post-processing ensures temporal consistency by verifying anomalies across consecutive frames.

Step 5: Result and Evaluation

The system's performance is evaluated using metrics such as Accuracy, Precision, Recall, F1-Score, and mAP. The final output includes annotated video frames highlighting detected anomalies and real-time alerts for suspicious activities. The results show improved detection accuracy, reduced false positives, and high efficiency for real-world surveillance environments.

RESULTS

The expected output of this project is a fully functional, real-time video anomaly detection system powered by the YOLO (You Only Look Once) deep learning model. The system is designed to automatically analyze surveillance footage, detect multiple objects within each frame, and identify abnormal or suspicious behaviors without any manual intervention. When the model processes a video feed—either a live stream or a pre-recorded dataset—it continuously monitors the scene, generating bounding boxes around detected objects such as people, vehicles, or other relevant entities. Each bounding box is labeled with a class name and a confidence score, indicating the probability of correct detection. For every instance where an unusual activity occurs—such as fighting, running in restricted zones, theft, vandalism, or sudden crowd movement—the system flags the event as an anomaly. These detected anomalies are visually highlighted on the video output using distinct colors and alert indicators, allowing security operators to instantly recognize suspicious actions. Alongside the visual detection, the system generates an automated alert message or notification for quick response in real-world applications. Furthermore, the output includes detailed analytical reports containing metadata such as frame number, timestamp, detected object labels, confidence levels, and classification results. These reports provide valuable insights for post-event investigation or dataset evaluation. The system also displays annotated video streams that show normal and abnormal behaviors in real time, ensuring interpretability and transparency of the model's and public safety applications, were quick and accurate detection is essential. In conclusion, this project decisions. In terms of performance, the expected results include high detection accuracy, minimal false positives, and stable processing speed—demonstrating that the YOLO-based model can efficiently handle complex and dynamic environments. anomalies are visually highlighted on the video output using distinct colors and alert indicators, allowing security operators to instantly recognize suspicious actions. Alongside the visual detection, the system generates an automated alert message or notification for quick response in real-world applications. The evaluation metrics, such as Accuracy, Precision, Recall, F1-score, and mean Average Precision (mAP), are expected to confirm the model's robustness and reliability. Overall, the expected output represents a scalable and intelligent surveillance solution capable of supporting smart city infrastructure, traffic monitoring, and public safety systems, enabling faster incident detection, improved situational awareness, and enhanced security management.

a. Large Croud Pushing Detection

b. Fire Detection

CONCLUSION

The proposed system effectively demonstrates the capability of YOLO-based models for real-time video anomaly detection. By leveraging deep learning and object detection techniques, the model overcomes the limitations of traditional methods that

depend on handcrafted features or manual monitoring. The integration of YOLO enables fast and accurate identification of suspicious activities such as fights, thefts, and abnormal movements in diverse surveillance environments. Through proper dataset collection, preprocessing, and model training, the system achieves high detection accuracy while maintaining efficiency and scalability. The results show that YOLO can be reliably implemented for smart surveillance, traffic monitoring, highlights the potential of deep learning–based video analytics to enhance situational awareness, reduce human effort, and improve response time in critical security events. Future improvements may include integrating multi-camera tracking, edge computing for faster inference, and adaptive learning for dynamic environments.

REFERENCE

[1] Y. Zhang, X. Li, Z. Wang, Y. Liu., "A Novel YOLOv5-based Anomalous Object Detection Algorithm in Buses" IEEE, 2022. This paper explains the Proposed a YOLOv5-based algorithm for detecting anomalous objects in bus videos. The model was trained on real bus video data and optimized for accuracy and speed. Extensive experiments were conducted to verify effectiveness.

[2] A. B. M. Sultan, M. S. Hossain, M. A. Hossain., "Enhancing Anomaly Detection in Videos using a Combined YOLO and a VGG GRU Approach" IEEE, 2022.This paper focuses on Developed an architecture combining YOLO for object detection and VGG-GRU for sequence analysis to detect anomalies in videos. The system was designed for quick response in monitoring applications.

[3] S. Chen, Y. Zhang, L. Wang., " A Lightweight Approach for Real-Time Video Anomaly Detection in Multi-Camera Scenarios." IEEE, 2023.This study uses Designed a novel lightweight approach for detecting video anomalies in multi-camera scenarios. The method focuses on real-time performance and efficiency.

[4] J. Doe, A. Smith., " Video Surveillance Anomaly Detection: A Review on Deep Learning Benchmarks." IEEE, 2022. Reviewed various deep learning models and datasets used for video surveillance anomaly detection. Examined limitations and trade-offs of different approaches.

[5] M. K. Singh, P. Verma, R. K. Gupta., "An Effective Technique for Video Condensation and Retrieval Using CNN with YOLO-Aided Anomaly Detection Framework." IEEE, 2022.This paper developed a deep learning-based model to address anomaly detection in video condensation and retrieval processes. Utilized CNNs with YOLO for improved detection accuracy.

[6] A. Kumar, S. Sharma, D. Patel., “Smart Video Surveillance Based Weapon Identification Using YOLOv5.” IEEE, 2022. This paper is implemented a YOLOv5-based system for weapon identification in video surveillance. Aimed to enhance security by detecting weapons in real-time.

[7] S. Patel, A. Desai, R. Mehta., “YOLOv5: Anomaly Detection in Surveillance Videos.” IJCRT, 2022. This paper is proposed a novel approach for anomaly detection in surveillance videos using YOLOv5 and deep learning methodologies.