

AI-Based Suspicious Activity Detection Using Camera with Clip Extraction and Transmission

Prof. Rajashree Shirur (M.Tech)¹

Siddhi Prashant Kshirsagar², Harshada Dilip Bhosale³, Prachi Prashant Suryawanshi⁴, Yogita Balaji karde⁵

^{1, 2, 3, 4, 5} Jaywantrao Sawant Prasarak Mandal Narhe Technical Campus, Pune

ABSTRACT

The demand for intelligence surveillance systems is increasing in public and private areas has speeded up the adoption of artificial intelligence (AI) for suspicious activity detection. Traditional closed-circuit television (CCTV) systems totally dependent on manual monitoring, which is prone to errors and inefficiency. In AI-Driven approaches, combined with real-time video analytics, enable automated detection of abnormal or suspicious behavior. This review paper discusses existing methods for suspicious activity detection, with feature extraction, machine learning, and deep learning models. In Addition, it examines the role of clip extraction to reduce the storage requirements and enable efficient review, along with transmission mechanism to ensure on time alert delivery. Current challenges, such as high false negative, privacy concerns, and issues of scalability are also highlighted, followed by discussion on future research directions in this domain.

KEYWORDS: Artificial Intelligence (AI), Suspicious Activity Detection, Deep Learning, Human Activity Recognition.

INTRODUCTION

With the Immediate growth of urbanization, public safety has become one of the most major concern Surveillance cameras are deployed widely in smart cities, banks, transportation hubs, and other crowded and sensitive areas where the crowd is major issue. However, manual monitoring of CCTV inputs is inefficient and often results in missed incidents. AI-based surveillance system offers a promising alternative by enabling real-time suspicious activity detection.

By applying advanced techniques such as deep learning and machine learning, AI systems can automatically Recent advances in machine learning (ML) and deep learning (DL) have improved accuracy and robustness of activity recognition systems. ML-based object detection algorithms, particularly YOLO (You Only Look Once), it allows real-time detection of humans and objects in video frames with high precision. Along with LSTM (Long Short-Term Memory) networks, these systems can learn patterns of suspicious and unsuspicious behavior over time, enabling the

recognize abnormal or suspicious human behaviors, including theft, violence, loitering, or intrusion. Unlike conventional methods, AI models can handle large volumes of video data more efficiently by focusing on key activity segments instead of storing bulk recordings. To address this, clip extraction feature focuses on storing only event-triggered segments, while secure transmission mechanisms ensure real-time alerts reach the concerned authorities.

identification of suspicious activities based on motion and interaction patterns.

In addition to detection, efficient data management is also important and critical in surveillance systems. Clip extraction mechanisms, which store only relevant event-triggered segments, reduce storage requirements and simplify analysis. Ability to transmit extracted clips securely and I real time ensures that monitoring personnel can respond quickly to incidents, improving overall system effectiveness.

The implementation of such systems typically involves a combination of OpenCV for video preprocessing and frame handling, TensorFlow for building and training DL models. Integrated approach allows intelligent surveillance systems to process multiple camera feeds effectively, detect suspicious activities accurately, extract relevant

clips, and transmit alerts to remote monitoring centers in real time.

This review paper explores existing methods, datasets, challenges, and future prospects of AI-based suspicious activity detection systems with a focus on clip extraction and transmission.

LITERATURE SURVEY

Ajeet Sunil [1] focus on detecting normal and abnormal human activities from videos using deep learning for security purpose. The author use a Single Shot Detector (SSD) algorithm with Convolution Neural Networks (CNN's) to identify and activities like climbing, running, punching and vandalizing. This model is trained by using transfer learning and testing real-time video frames. It can detect action and objects like guns and knives. This research is about improving safety in public areas through intelligent video surveillance. It focusses Challenges like the need of large and diverse datasets, It also needs of high computing power for real-time use. This paper includes that the future work can involve edge devices, FPGA, GNN and

quantum CNN technologies for faster and more accurate recognition.

Author [2] presents an IoT and AI-based wearable device for women's safety. In this device there is GPS tracking, voice recognition, panic button and a loud buzzer it provides quick help in critical situations. When it activates, it sends the woman's live location and emergency services through GSM module and send alert. This system provides voice commands like "Help" by using AI. This system ensures faster response and low cost with good battery life. The author compares this safety device to other existing safety devices and shows that the proposed model works more independent and faster by using smart technology in real-time situations.

In Paper [3] The author presents a smart surveillance system by using AI, motion sensors and facial expression recognition to detect unusual or dangerous activities in public places like schools, banks, malls and military areas. This system built using Raspberry Pi 4 and camera module that captures videos and analyzes videos with Python-base AI models and send alerts if suspicious behavior is found. The goal is to reduce

violence, thefts and accidents with automatic detection and send alerts without human interference. All the study reviews of past work in video surveillance and it explains how this system improves accuracy. The project is about how AI can automatically identify irregular human activities and issue alerts through speakers or messages and making surveillance more efficient.

Paper [4] This paper presents a real-time system to detect suspicious human action from CCTV video using deep learning. The researchers develop a device to analyze live video feeds without needing to send data to centralize servers. The model includes pose estimation, optical flow histograms and Multiple Instance Learning (MIL) with Early Stopping to identify and classify abnormal activities efficiently It gives us faster processing and high accuracy as compared to older devices, heavier. This device is lightweight compared to older devices. The system also includes person tracking and re-identification, it also allows us to continuously follow suspects across cameras. Overall, the aim of this project is to make surveillance cheaper, smarter and more privacy-friendly tracking and detect criminal behavior in real time

Paper [5] The author explains in this paper how artificial intelligence (AI) can improve cybersecurity by detecting and responding to threats and more accurately. In this paper it describes how old forensic methods are too slow, meanwhile AI system is faster than that it can find automatically unusual patterns, it identifies attacks and also responds in real time. Overall study is about different AI methods like deep learning, machine learning and anomaly detection and it helps in analyzing massive data from cloud systems, network and devices. It also faces challenges like data privacy and defending AI models from attacks. Finally, it suggests future directions like using behavioral biometrics and self-healing cyber systems to make digital infrastructures smarter, more secure and more resilient.

Akshay Vaja [6] explain how Artificial Intelligence (AI) can help detect suspicious activities through surveillance videos. It utilizes a surveillance video system connected to a user's device that generates an

alarm sound upon detecting illegal activities such as disputes, fighting, or any clearance of work. The main goal is to monitor places like college premises or examination halls to identify actions such as fighting, disputes, or cheating. When a suspicious activity is detected, an alarm or beep sound is generated to alert the user. The technology leverages Machine Learning and Deep Learning, specifically Convolutional Neural Networks (CNNs), for analysis and activity recognition. They conclude that while the project is currently for academic use, it has future potential to be expanded for crime detection in public and private areas to prevent incidents before they happen.

[7] This research paper focuses on creating an AI system to detect suspicious activities like fighting and disputes in surveillance videos, especially for college campuses. It reviews different deep learning techniques like CNNs, LSTMs, and transformer models that can recognize human actions and unusual activities in real-time. The authors highlight that while their current project is designed for academic settings, the technology has the potential to be expanded for use in public and private areas to prevent crimes.

[8] This research paper introduces a new AI model for automatically detecting suspicious activities like robbery, shoplifting, and fighting in surveillance videos. It combines Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to analyse

both spatial and temporal features in video footage. It uses a technique called ORB for fast and efficient feature extraction from video frames. The system was tested using the UCF Crime Dataset, focusing on crimes like robbery, fighting, and shoplifting, and achieved 99% accuracy, outperforming traditional models like VGG-16 and ResNet.

[9] The authors highlight that while their current project is designed for academic settings, the technology has the potential to be expanded for use in public and private areas to prevent crimes. They created their own dataset containing six activities — running, punching, falling, snatching, kicking, and shooting. They used Convolutional Neural Networks (CNNs) and models like Time Distributed CNN, Conv3D, Hybrid LSTM, and Keras GRU to analyse video frames.

[10] This research paper presents a system that uses deep learning to automatically detect suspicious activities from real-time surveillance videos. The approach processes video frames to identify abnormal behaviours like fighting or running by analysing motion, head movements, and contact patterns. The system aims to enhance security in public places by automatically classifying activities as normal or suspicious and generating alerts. The authors conclude that this technology can create a safer environment and has great potential for various security applications.

IMPLEMENTATION METHODOLOGY

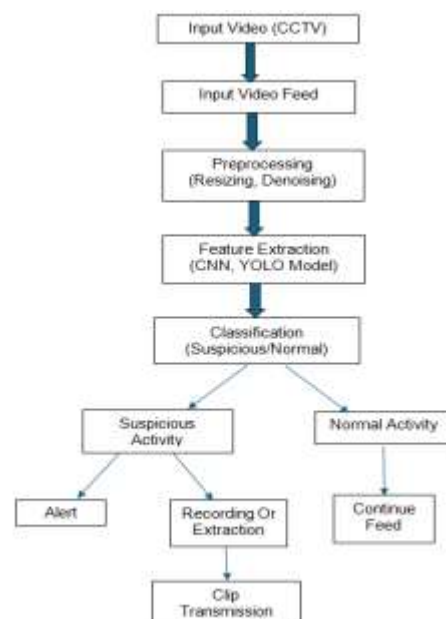


Fig1: Flow chart of design and implementation

The suspicious activity detection system begins with live video input captured from CCTV cameras installed in monitoring areas. This continuous video feed is sent to the system for analysis, where it first undergoes preprocessing steps such as resizing and denoising to ensure that the frames are clear and suitable for analysis. After preprocessing, advanced deep learning models like CNN or YOLO are used to extract important visual features, identifying people, objects, and movements within the scene. These extracted features are then classified to determine whether the detected activity is normal or suspicious. If any suspicious activity is recognized, the system automatically generates an alert

and extracts a short video clip of the event for evidence or further investigation. In the case of normal activity, the video feed continues without any interruption. This automated process helps in real-time surveillance, quick response, and efficient security monitoring.

BLOCK DIAGRAM OF A SYSTEM

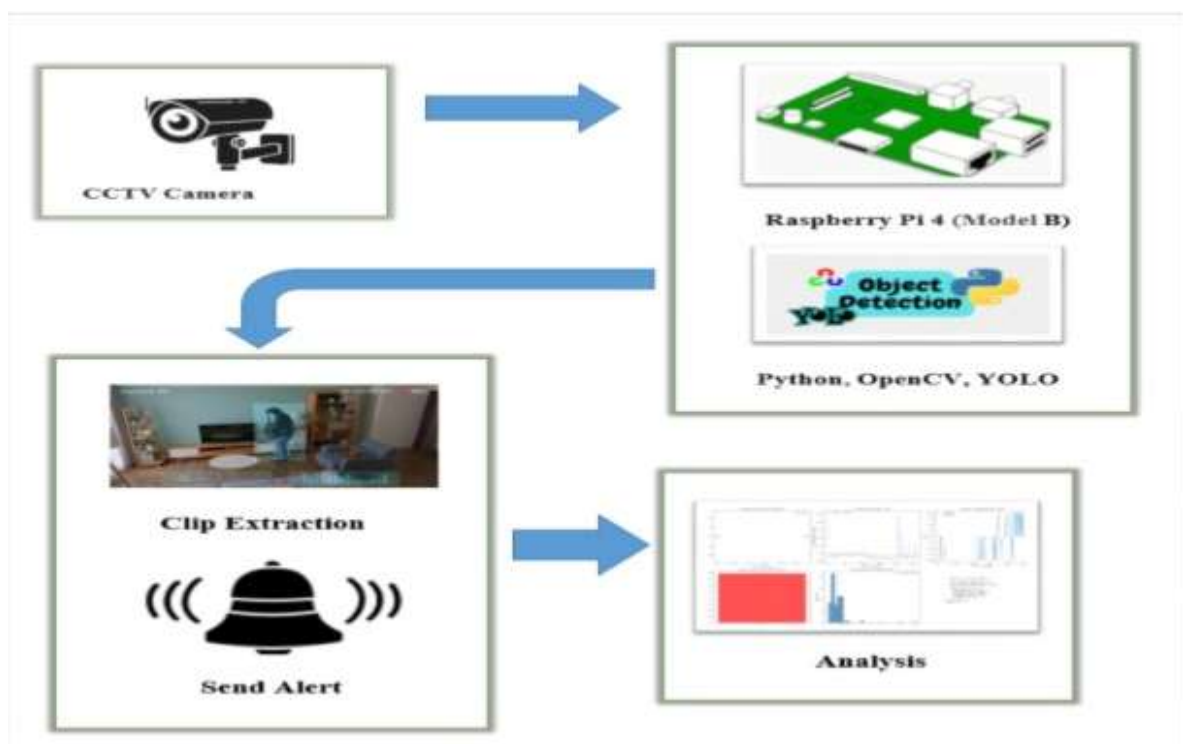


Fig2: Block Diagram of System

A CCTV camera streams video → an edge Processing Unit (Raspberry Pi 4) ingests frames, runs detection + behaviour models (YOLO + tracking), decides when a suspicious event occurs → the unit extracts and encodes a short video clip (with pre/post buffers) and sends the clip + metadata to Cloud Storage → cloud services index and (optionally) reprocess clips, populate the Dashboard → Dashboard shows live/recorded clips and metrics, and an alert (with clip link) is forwarded to the Control Room.

Processing Unit (Edge) — Raspberry Pi 5 in your diagram: runs the real-time pipeline (frame grab, pre-processing, object detection via YOLO or tiny/edge variant, tracking, behaviour analysis with LSTM/RNN), performs clip extraction and local buffering, and initiates secure upload/alert. Why

edge? low latency, reduced bandwidth (you send only clips/alerts not full continuous video), privacy

Cloud / Storage: stores clips, indexes metadata, runs heavier analytics (re-verification, analytics over time), and hosts dashboard & APIs.

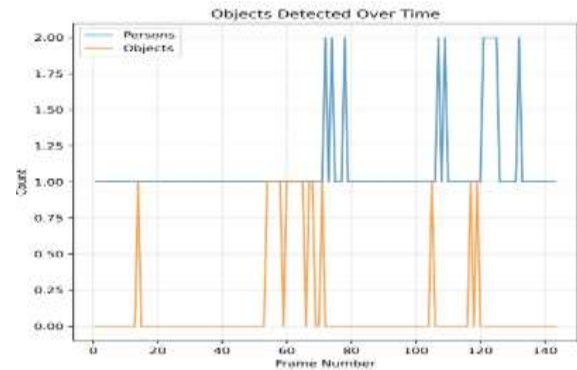
Dashboard + Control Room: UI and notification endpoints (web UI, operator consoles, SMS/Email/Push). Control room receives prioritized alerts and can play back clips or view live feed.

Control Room: Receives high-priority alerts via audible alarm + visual popup containing camera name, thumbnail, short description, confidence, play clip or jump to live feed.

This graph shows the time taken to process each video frame.

X-axis: Frame number

Y-axis: Processing time (seconds)



SIMULATION RESULT AND ANALYSIS

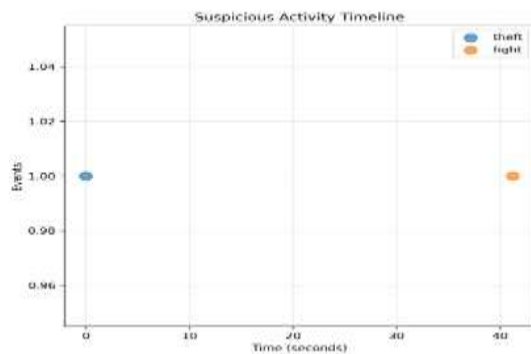


Fig.3.1 Shows Suspicious Activity Timeline
A theft event occurred at the beginning of the observation period (~0 seconds).
A fight event was detected later (~40 seconds).
his graph represents the timeline of detected suspicious events over time.

X-axis: Time (in seconds)

Y-axis: Event occurrence

Fig.3.3 Objects Detected Over Time
Shows Objects Detected Over Time This graph shows the number of persons and objects detected per frame. (The number of persons varies between 1 and 2)
X-axis: Frame number Y-axis: Count
Blue line: Number of persons
Orange line: Number of objects

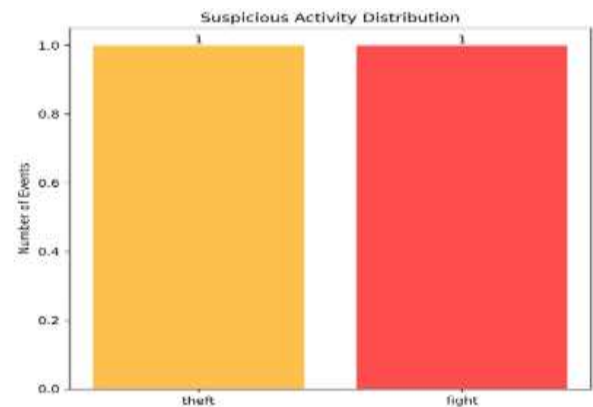


Fig.3.4 Suspicious activity distribution

This graph shows the frequency of detected events.

X-axis: Event type (theft, fight)

Y-axis: Number of events

1 theft event detected

1 fight event detected

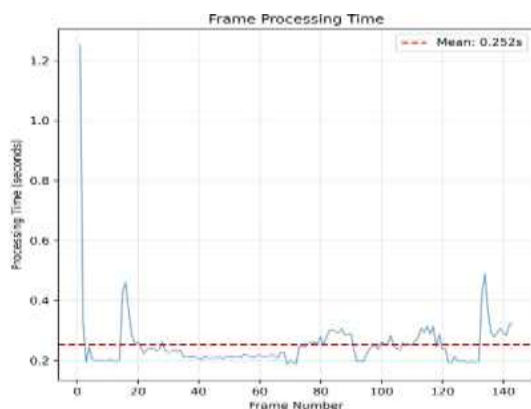


Fig.3.2 Frame processing time

Fig.3.2 Shows Frame Processing Time (Line Graph)
Red dashed line: Average processing time (≈ 0.252 seconds)

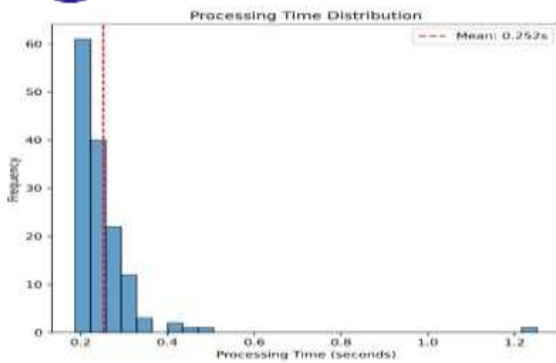


Fig.3.5 Processing time distribution

This graph shows the distribution of frame processing times.

X-axis: Processing time (seconds)

Y-axis: Frequency

Red dashed line: Mean processing time (≈ 0.252 seconds)

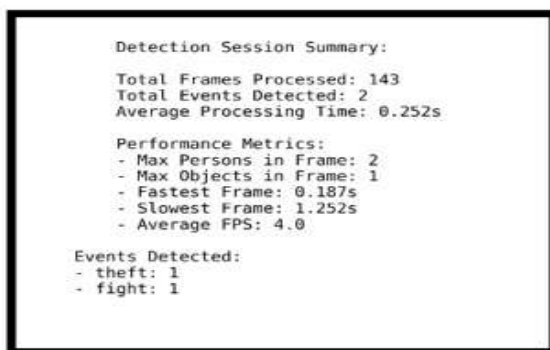


Fig.3.6 Detection Session Summary

Fig.3.6 Shows Detection Session Summary

This section provides a numerical summary of system performance:

Total Frames Processed: 143

Total Events Detected: 2

Average Processing Time: 0.252 seconds

The system achieves moderate real-time performance (4 FPS).

Detection accuracy is demonstrated by correct identification of both event types.

Max Objects in Frame: 0 No other objects detected. Fastest Frame: 0.146s Quickest processing time. Slowest Frame: 1.551s Longest processing time (possibly when fight detected). Average FPS: 4.6 Frames processed per second.



Fig. 4 Detection Output of a Fighting

Fig.4 shows the frame where two persons are involved in a physical altercation. The activity to be detected here includes aggressive motion, close physical contact, and irregular body movements, which are indicative of a fight. The detection model successfully identifies this as a suspicious activity with high confidence. The bounding boxes around the individuals indicate that the system has localized both persons participating in the conflict. Such detection helps in recognizing violent behaviour in real-time surveillance footage, allowing timely alerts for security and monitoring applications.



Fig.5 Detection output of a thefting

Fig. 5: The above image shows the detection of a suspicious activity (theft) captured by the surveillance camera. The system has successfully identified and tracked individuals within the frame using object detection and tracking algorithms, as indicated by the green bounding boxes labelled with unique IDs.

CONCLUSION

The project on Suspicious Activity Detection using Camera with Clip Extraction and Transmission successfully demonstrates an intelligent surveillance system capable of identifying abnormal human activities such as fights or thefts in real time. Using computer vision and machine learning techniques, the system detects suspicious motion patterns from live

camera footages and immediately extracts the corresponding video clip for evidence storage or further analysis. The extracted clips are automatically transmitted to a remote server or cloud platform for monitoring and alert generation. The experimental results show that the model effectively distinguishes between normal and suspicious activities with reliable accuracy and consistent processing performance.

FUTURE SCOPE

In the future, Integration with IoT and cloud computing can enable instant alerts to authorities through mobile notifications or email. Incorporating facial recognition and behaviour analysis can help identify individuals involved in the activity, thereby strengthening security and forensic investigation capabilities.

REFERENCES

- [1] Ajeet Sunil, Manav Hiren Sheth, Shreyas E, Mohana, "Usual and Unusual Human Activity Recognition in Video using Deep Learning and Artificial Intelligence for Security Applications", *2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT)* | 978-1-6654-1480-7/21/\$31.00 ©2021 IEEE | DOI: 10.1109/ICECCT52121.2021.9616791.
- [2] Bathula Prasanna Kumar, Satra Sala Jyotsnika, Vuyyuru Lakshmi Sruthi, Regalla Sumana, Shaik Hussanibi, "AI-driven security: real time women's protection via IOT-based guardian device", *International Journal of Engineering Science and Advanced Technology (IJESAT) Vol 25 Issue 03, MAR, 2025*.
- [3] Dr. Nookala Venu, Dr. A. Arun Kumar, "Suspicious Activity Tracking Artificial Intelligence Camera", *Volume 29, Issue 10, 2023*.
- [4] I. Subha, P. Narmadha, S. Nivedha, T. Sethukarasi, "Real-Time Suspicious Human Action Recognition from Surveillance Videos for Resource-Constrained Devices", *2020 Journal Computational and Theoretical Nanoscience Vol. 17, 3790–3797*.
- [5] Ogochukwu Susan Ndibe, "AI-Driven Forensic Systems for Real-Time Anomaly Detection and Threat Mitigation in Cybersecurity Infrastructures", *International Journal of Advance Research Publication and Reviews, Vol 2, no 5, pp 389-411, May 2025*.
- [6] Akshay Vaja, Dr. Dhaval Raghav Bhai Kathiriya, "Suspicious Activity Detection using AI under Surveillance Video", *Advanced International Journal of Multidisciplinary Research E-ISSN: XXXX-XXXX editor@aijmr.com Volume 1, Issue 1, July – August 2023*.
- [7] J. Arunnehr, Divya Midhunchakkaravarthy, S. Hemalatha, "Intelligent Video Surveillance Systems for Real-Time Suspicious Activity Detection Using AI and Computer Vision: A Study", *SGS Engineering & Sciences, VOL. 1 NO .1 (2025): LGPR*.
- [8] AasmaAas , Hamza Naveed , Sarah Khaleel, Junaid Asghar, Talia Noureen , Zubda Khanum , "Efficient and Sustainable Video Surveillance Using CNN-LSTM Model for Suspicious Activity Detection", *VFAST Transactions on Software Engineering VFAST Transactions on Software Engineering Volume 13, Issue 1, 2025 https://vfast.org/journals/index.php/VTSE@2025, ISSN(e): 2309-3978, ISSN(p): 2411-6246 Volume 13, Number 1, January-March 2025*.
- [9] Monji Mohamed Zaidi Ahmadalmadhor, Gabriel Avelino Sampedro, Shtwai Alsubai, Abdullah Al Hejaili, Michal Gregus, Sidra Abbas, "Suspicious Human Activity Recognition From Surveillance Videos Using Deep Learning".
- [10] Vedant Saikhede, Kiran Shende, Yuvraj Darekar, Hemant Thorat Prof. Snehal. S. Shinde, "Deep Learning Approach for Suspicious Activity Detection from Surveillance Video", *International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 10 Issue: 12 | Dec 2023*.