Multimodal Video Summarization For Crime Scene Analysis Using Yolo Based Object Detection

Karwin Vikash Tr[1], Jean Benedict M[2], Iyappank[3], Subasree[4]
[1, 2, 3, 4]Dept of Electronics and Communication Engineering
[1, 2, 3, 4] CARE College of Engineering Approved by AICTE | Affiliated To Anna University

Abstract- *The rapid growth of surveillance cameras in public places has led to the generation of large volumes of video data, making manual monitoring difficult and inefficient. Detecting violent activities such as fights, weapon usage, and suspicious behaviour in real time is essential for maintaining public safety. This paper proposes a deep learning-based system for crime scene analysis using multimodal video summarization and YOLO-based object detection. The system processes surveillance videos by extracting frames and applying preprocessing techniques, followed by object detection using the YOLO algorithm. Convolutional Neural Networks (CNN) are used for feature extraction to identify objects such as people and weapons. The system analyses detected objects to identify suspicious activities and generates real-time alerts with relevant information such as timestamps and confidence scores. This approach improves the efficiency of surveillance monitoring and helps security personnel respond quickly to potential threats.*

Keywords: Video Surveillance, YOLO Object Detection, Deep Learning, Crime Detection, Convolutional Neural Networks (CNN), Video Summarization.

I. INTRODUCTION

1.1.PROJECT OVERVIEW

This project focuses on developing an intelligent surveillance system that can automatically detect violent or suspicious activities in video footage using deep learning techniques. The system uses YOLO-based object detection to identify objects such as people and weapons in real time.

- Detects suspicious activities from surveillance videos.
- Uses YOLO algorithm for real-time object detection.
- Generates alerts when abnormal activities are detected.

1.2. PROBLEM DESCRIPTION

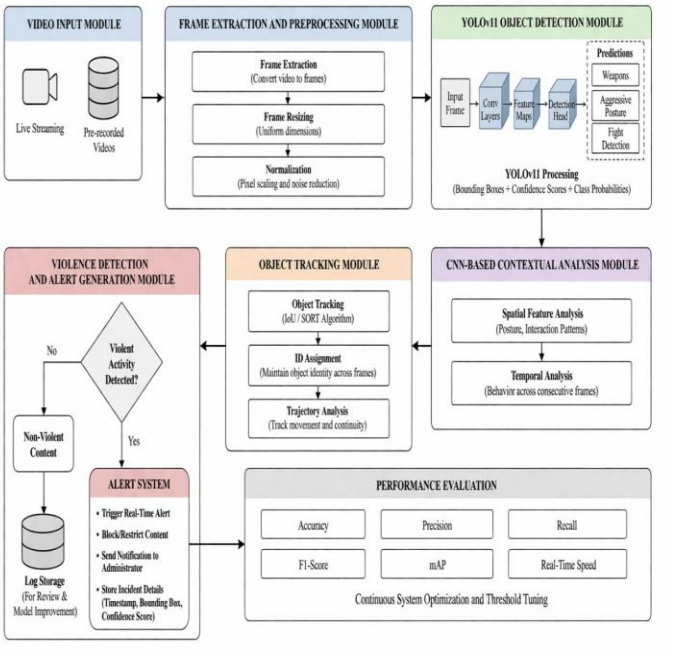
- Manual monitoring of surveillance videos is time-consuming and inefficient.
- Security personnel may miss critical events due to human limitations.
- Traditional systems lack automated detection of violent activities.

II. LITERATURE SURVEY

Many studies have focused on developing intelligent surveillance systems using deep learning techniques. Convolutional Neural Networks (CNN) are widely used for feature extraction and object recognition. The YOLO algorithm provides fast and accurate real-time object detection. These methods improve the detection of suspicious activities in surveillance videos.

III. PROPOSED SYSTEM

3.1.SYSTEM ARCHITECTURE



The platform comprises:

- Input Layer: Receives video streams from surveillance cameras.
- Processing Layer: Performs frame extraction and preprocessing.
- Detection Layer: Uses YOLO algorithm to detect objects.
- Alert Layer: Generates alerts when suspicious activity is detected.

3.2.KEY FEATURES

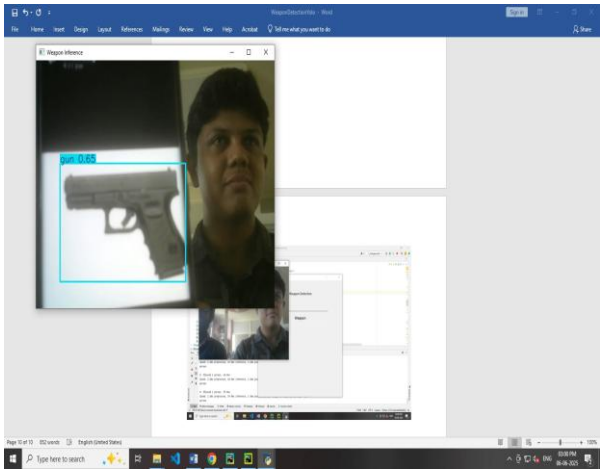
- Real-time detection of suspicious activities in surveillance videos.
- YOLO algorithm enables fast and accurate object detection.
- Automatic alerts generated when abnormal activities are detected.

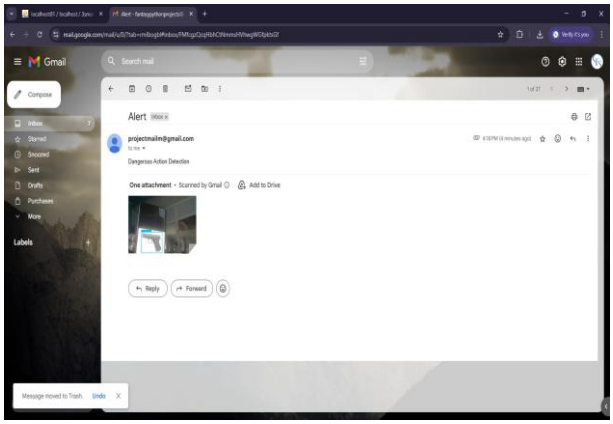
IV. METHODOLOGY

The system follows an deeplearning-based approach:

- Collect video input from surveillance cameras.
- Extract frames from the input video stream.
- Apply preprocessing to improve frame quality.
- Use CNN for feature extraction from frames.
- Detect objects using YOLO algorithm.
- Generate alerts when suspicious activity is detected.

SCREENSHOTS





V. DISCUSSION

STRENGTHS

- Provides real-time detection of suspicious activities.
- Improves surveillance efficiency using deep learning.
- Reduces manual monitoring by security personnel.

LIMITATIONS

- Requires large datasets for accurate training.

- Performance may decrease in low-light conditions.

FUTURE WORK

- Improve detection accuracy using advanced AI models.
- Integrate system with smart city surveillance networks.
- Develop mobile alert systems for faster response.

VI. CONCLUSION

The proposed system provides an intelligent solution for detecting suspicious activities in surveillance videos. It uses the YOLO-based object detection model to identify objects such as people and weapons in real time. The system improves the efficiency of surveillance monitoring by automatically analysing video frames. This helps security personnel respond quickly and enhances public safety.

REFERENCES

[1] J. Redmon and A. Farhadi, "YOLO: You Only Look Once – Unified Real-Time Object Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
[2] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015.
[3] W. Sultani, C. Chen, and M. Shah, "Real-World Anomaly Detection in Surveillance Videos," IEEE Conference on Computer Vision and Pattern Recognition, 2018.
[4] R. Girshick, "Fast R-CNN," IEEE International Conference on Computer Vision, 2015.