Theft Detection In CCTV

[1] Maaheedhar Dewaassi, [2] Sukhlal Choudhary, [3] Suraj Sharma, [4] Ansh Gurav, [5] Prof.Mohan Kumar

[1,2,3,4] Department of Data Engineering
[5] Professor, Department of Data Engineering
Universal College Of Engineering, Kaman, India

Abstract- Closed-Circuit Television (CCTV) surveillance is common in public places in many countries. The availability of CCTV footage has resulted in significant changes in policing and legal systems. While locating a person of interest is important for public safety, it is also a task of high visual complexity that necessitates sustained attention, good identity detection and recognition skills, and other cognitive resources. Also, because there is limited time and resources to take corrective action, the scale of wrongdoing is growing exponentially. It has become critical to reduce the amount of time required to identify the offender. The primary goal of our project is to expedite the process of identifying an offender from CCTV footage or other sources. The purpose of the project is to alleviate menial tasks, improve the efficiency of the identification process, and streamline the process. Metadata is used to boost the performance of identification systems. Since According to several police officers, when a crime occurs, whether it is a theft or a murder, investigators obtain CCTV footage from nearby cameras and analyse the chronology of the incident in order to identify and trace the accused persons. As a group, our aim is to create an algorithm based project employing a variety of other components to accomplish our objective to reduce the tedious work of manually watching and spending the countless hours identifying a person from the acquired footage. We'll make sure that we complete this project by the set date and break it down into more simpler modules to make it more viable as it will be a learning curve for us as we'll be exposed to unhackneyed concepts for us and expand our learning horizon.

Keywords -: - Security, CCTV, Theft, Robbery

1. INTRODUCTION

The project seeks to revolutionize the process of person identification within CCTV footage by harnessing the power of Artificial Intelligence, image processing, and machine learning. Artificial Intelligence and Machine Learning are associated with human Intelligence and natural intelligence [1]. With a focus on streamlining manual tasks, reducing time constraints, and enhancing efficiency, this endeavor aims to expedite the identification of individuals, especially in law enforcement scenarios, ultimately contributing to improved public safety and investigative processes**.** In the present scenario, the significance of Closed-Circuit Television (CCTV) video has become prominent as a crucial element in enhancing cybersecurity protocols [5].

In our project we have develop a Real-Time theft detecting system that uses basic Computer Vision and Deep Learning. The goal is to detect when a person performs the suspicious act. Most existing theft-detection papers have used older YOLO versions such as YOLOv5, YOLOV6, YOLOv7[1][3][4]. In our work, we use YOLOv8n, a lightweight and faster model, for real-time ATM theft detection. By integrating an Action Recognition model to analyse the sequence of frames, the system would classify the tracked person's action as 'concealment,' and when confirmed, it would save the incident clip and then send to security dashboard.

2. LITERATURE SURVEY

A literature survey was carried out to find various papers published in international journals such as IEEE, Springer, MDPI etc. related to tracing missing people using facial recognition to get the best algorithm for the same.

Surbhi Singla and Raman Chadha [1], In this paper, an intelligent system which will detect the crimes from the real time CCTV Feed and classify them and provides an alert system to the nearest police stations and ambulances etc. So, that system will help in reducing the crime rate in any country. This paper reviews all prior research in this area, including approaches for object recognition and finding priority frames, techniques and algorithms like Yolo used to detect crimes , various datasets used and algorithms used to analyse crime data and train the dataset .It covers the various recent trends in researches in this field and analysing the challenges faced and various research gaps and this paper also discuss how we can overcome these gaps in research so as to develop a better intelligence surveillance system in ML field.

Suryanti Awang Mohd Qhairel Rafiqi Rokei and Junaida Sulaiman [4], the paper introduces a security system known as Suspicious Activity Trigger System (SATS) that able to automatically trigger an alarm or an alert message whenever suspicious activity is detected from the CCTV video image. The algorithm will detect an object which is a person in the video and classify it as a suspicious activity or not. If the activity is classified as the suspicious activity, the system will automatically display a trigger message to alert SATS user. The user can therefore take whatever appropriate measure to prevent being a victim. Therefore, YOLOv6 can be implemented in the security system to prevent crimes in residency areas.

T sings and S. K. Das [7], In this paper, highlight that older surveillance methods (like background subtraction) were unreliable, producing too many false alarms due to things like shadows or light changes. The previous generation of accurate Deep Learning model was too slow for real-time monitoring, and there was no practical, integrated system for smart cities. Their literature review justifies a new approach that uses the YOLO algorithm for fast, accurate object detection and combines it with a custom algorithm and SORT tracking on a compact hardware platform to achieve a rapid, reliable, and contextual decision about whether an object has actually intruded into a defined restricted zone.

Saldanha, Naik, Parashtekar [8], In this paper, authors establish that traditional CCTV surveillance, which relies on manual human monitoring, is highly inefficient and often fails to prevent crimes as it's mainly for post-event analysis. They note that while machine learning offers an alternative, earlier models were too slow or inaccurate. Their review justifies adopting the YOLO (You Only Look Once) algorithm because it provides the speed necessary for realtime object detection of suspicious activities, thereby allowing the proposed system to reduce reliance on human vigilance and enhance overall security through automated, instant alerts, even distinguishing between Daytime monitoring for weapons and Nighttime detection for anonymous intruders.

Murat Koca [5], Real-Time Hand Gesture Detection and Classification Using Convolutional Neural Networks. In their work, Murat Koca. introduced a movement detector and a classifier using CNNs to identify observed motions. The researchers suggested the Levenshtein distance for evaluation, which comprehensively analyses both misclassifications and the presence/absence of detections. This offline technique for RGB images achieved an accuracy of 0.973 and a precision of 0.803. Although effective, this method was outperformed in accuracy and precision by the proposed algorithm.

N. Omkar Sainath and Malla Reddy [3], YOLO-V7 to Overcome Theft Activities:, the authors developed a YOLOv7-based framework designed to detect suspicious theft-related activities in real-time surveillance footage. The system leverages YOLOv7's object detection capabilities to monitor and identify unusual human behaviors that may indicate theft. While the framework demonstrates potential in addressing security concerns, the paper notes that performance metrics such as accuracy and FPS, along with large-scale real-world testing, are not yet fully detailed, leaving room for further research and validation.

Yang Zhou, Xianghua Xu, and Ran Wang [2], EI-YOLO: Efficiently Improved YOLO on Detection of Prohibited Items During Security Inspections. the authors introduced an improved YOLO-based model known as EI-YOLO, which aims to provide faster and more accurate detection of prohibited items during security inspections. The model enhances detection efficiency in real-time scenarios where quick response is essential for safety. Although the method shows promising results, it still requires adaptation for theft detection in CCTV surveillance and further validation in real-world environments to ensure practical usability.

Mohammed Sherooq Ali, Sreeju A, Shifa Reem, Jasna, Neethu Mathew [10] This research presents an AI-based automatic theft detection system for smart homes using CCTV footage. The system employs face recognition techniques, including HOG features and

Haar cascade, to detect and identify individuals in real-time. It compares captured faces with a criminal database and notifies the owner and police if a threat is detected, enhancing security with minimal human intervention.

3. WORKING PRINCIPLE

The proposed system performs real-time theft detection by integrating object detection, face recognition, pose estimation, object tracking, and rule-based decision logic. The overall workflow processes each video frame sequentially and generates alerts based on behavioral and spatial analysis.

Video Source and Frame Acquisition

The system accepts input from:

Live webcam

CCTV camera feed

Pre-recorded video files

Using OpenCV, video frames are captured continuously. Each frame is extracted and forwarded to the detection pipeline for further processing. This ensures real-time analysis of surveillance footage.

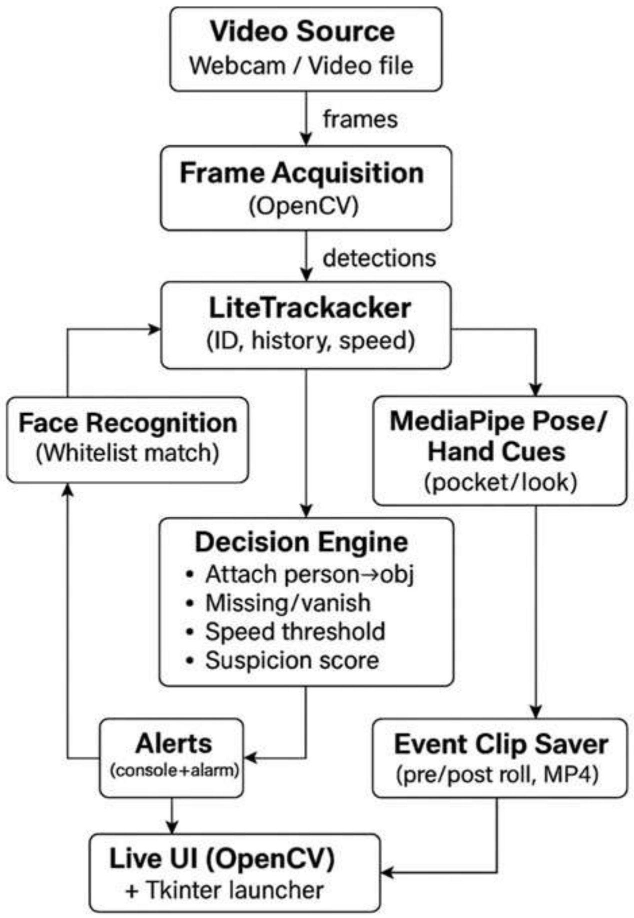


Fig 1.1 working model

3.1. OBJECT DETECTION MODULE

YOLOv8 Ultralytics The system employs the YOLOv8 Nano (yolov8n.pt) model for real-time object detection. The Nano variant is selected due to its lightweight architecture and low inference latency, making it suitable for surveillance applications.

YOLOv8 performs single-stage detection by dividing the input image into grids and predicting bounding boxes along with class probabilities in a single forward pass. The model detects the following classes:

- Person
- Backpack
- Handbag
- Suitcase
- Knife
- Bat

These objects are categorized as either potential theft targets or dangerous items.

3.2. OBJECT TRACKING MODULE

A centroid-based tracking algorithm is implemented using a custom Lite Tracker class. The tracker assigns unique IDs to detected objects and maintains their identity across consecutive frames.

For each detected bounding box, the centroid is computed as:

C = (x_center, y_center)Object association between frames is performed using Euclidean distance.

The minimum centroid distance determines object continuity. This tracking mechanism enables temporal behavior analysis such as movement speed and object disappearance detection.

3.3. FACE RECOGNITION MODULE (WHITELIST VERIFICATION)

Dlib

face_recognition

The system integrates a whitelist-based identity verification mechanism. Face detection is performed using Histogram of Oriented Gradients (HOG), which extracts gradient orientation features for robust face localization

Each detected face is encoded into a 128-dimensional feature vector using a deep convolutional neural network (ResNet-based embedding model). Identity matching is performed by computing Euclidean distance between the detected face encoding and stored whitelist encodings:

If the distance is below a predefined threshold, the individual is classified as authorized.

3.4. POSE ESTIMATION MODULE

MediaPipe

The system utilizes MediaPipe Holistic for human pose estimation and behavioral analysis. The framework detects and tracks key body landmarks including:

- Head
- Shoulders
- Hands
- Hip

By monitoring the spatial relationship and motion of these keypoints, the system identifies suspicious behaviors such as:

- Hands positioned inside pockets near valuable objects
- Frequent head rotation (indicative of scanning surroundings)
- Repeated bending or abnormal posture near objects

Pose-based cues contribute to contextual understanding beyond simple object proximity.

3.5. HEURISTIC-BASED THEFT DETECTION LOGIC

The final decision-making stage applies rule-based heuristics that combine spatial, temporal, and behavioral metrics.

3.5.1 PROXIMITY ANALYSIS

The distance between a person and a target object is computed.

3.5.2 SPEED ANALYSIS

If a detected object disappears from the frame and the associated person's displacement exceeds the defined threshold

the system flags a potential theft event.

3.5.3 LOITERING DETECTION

If an individual remains within the monitored region beyond a predefined time limit the behavior is classified as suspicious.

3.6. INTEGRATED DECISION FRAMEWORK

The system aggregates outputs from detection, tracking, face recognition, and pose analysis modules. An alert is generated only when multiple suspicious indicators are satisfied, thereby reducing false positives and improving reliability.

4. PROCEDURE

The proposed theft detection system was implemented in a structured and sequential manner. The overall development process consisted of system setup, model integration, real-time frame processing, and event evaluation.

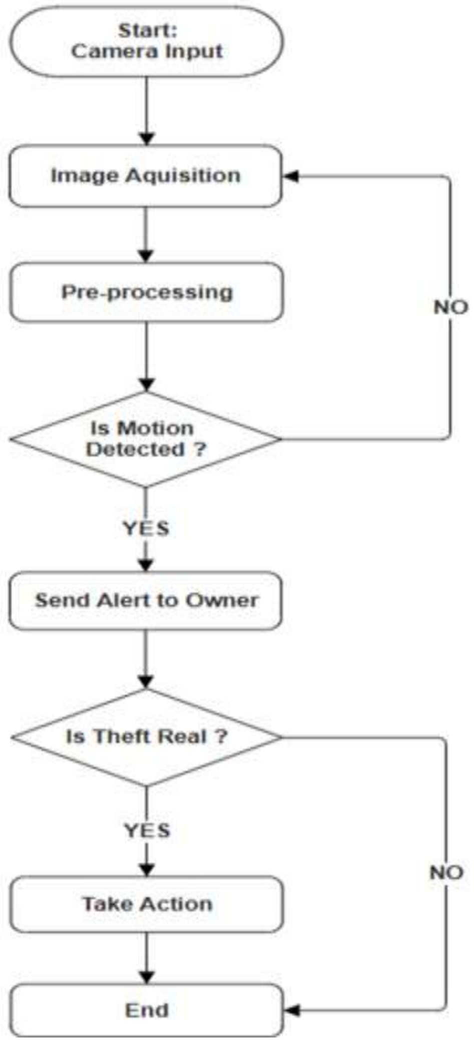


Fig 1.2 procedure

4.1. SYSTEM SETUP AND ENVIRONMENT CONFIGURATION

The system was developed using Python programming language. The required computer vision and deep learning libraries were installed and configured, including:

- Ultralytics
- dlib
- face_recognition
- MediaPipe

Pre-trained model weights (yolov8n.pt) were loaded for object detection. Face encodings of authorized individuals were pre-computed and stored in a database for whitelist verification.

4.2. VIDEO INPUT AND FRAME EXTRACTION

A live CCTV feed or recorded video file was used as input.

The system performs:

Step 1: Capture video stream

Step 2: Extract frames sequentially

Step 3: Resize frames for optimized inference

Step 4: Pass frames to detection pipeline.

4.3. DETECTION AND TRACKING EXECUTION

For each frame:

1. Object detection was performed using YOLOv8 Nano.
2. Bounding boxes and confidence scores were extracted.
3. Centroid coordinates were computed.
4. The custom centroid tracker assigned unique IDs.
5. Object displacement between consecutive frames was calculated.

This enabled speed estimation and object disappearance detection.

4.4. FACE RECOGNITION IMPLEMENTATION

When a person was detected:

- Face region was extracted.
- HOG-based face detection was applied.
- 128-dimensional embeddings were generated.
- Euclidean distance comparison was performed against stored whitelist encodings. If the distance was below threshold → Authorized. Else → Treated as unknown individual.

4.5. POSE-BASED BEHAVIORAL ANALYSIS

MediaPipe Holistic model was applied to detect body landmarks.

The system monitored:

- Head rotation frequency
- Hand proximity to detected objects
- Body posture near valuables

Landmark coordinates were tracked across frames to detect suspicious patterns.

4.6. RULE-BASED DECISIONE VALUATION

After collecting spatial and temporal data, the rule engine evaluated:

- Person-object proximity
- Running speed after object disappearance
- Loitering duration in frame

If multiple suspicious conditions were satisfied, the system generated:

- Visual alert (red bounding box)
- Warning message on dashboard
- Event logging with timestamp

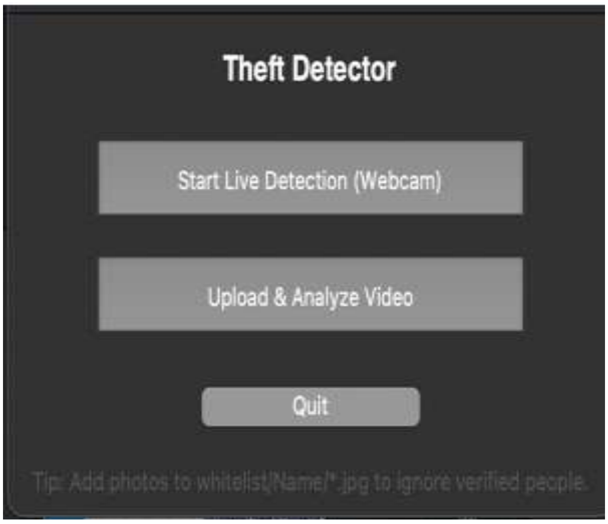
4.7. SYSTEM TESTING

The system was tested under:

- Multiple individuals in frame
- Different lighting conditions
- Occlusion scenarios
- Simulated theft behavior

Observations were recorded for detection consistency and false alert rate.

5. RESULTS



The output shows the working of a theft detection system using the YOLOv8 algorithm on CCTV footage. The system detects objects such as a person and a backpack and highlights them with bounding boxes in the video. It also analyzes the person's movement by calculating speed and a suspicion score to identify unusual behavior. If suspicious activity is detected, the system generates an alert and saves the annotated video for review. The interface provides options to start live detection, upload and analyze a video, or quit the application, and it also allows adding trusted people to a whitelist to ignore them during detection.

6. CONCLUSION

The reviewed papers collectively emphasize the growing reliance on deep learning, particularly YOLO variants (YOLOv2–YOLOv7 and EI-YOLO), for real-time crime and anomaly detection in CCTV surveillance. YOLO's real-time object detection capabilities enable quick and accurate identification of potential threats or theft-related activities, allowing for immediate response and prevention [8] While CNNs, LSTMs, and hybrid models have enhanced accuracy, speed, and contextual decision- making, challenges such as false positives, adaptability to varied environments, and large-scale real world validation remain. Overall, these studies highlight the potential of intelligent surveillance systems to move beyond manual monitoring towards proactive, automated, and reliable crime prevention solutions.

7. FUTURE SCOPE

- The system can be improved by integrating advanced AI models such as YOLOv8 to achieve more accurate and faster object detection in real-time surveillance.
- Face recognition features using tools like face_recognition can be enhanced to automatically identify authorized staff and detect unknown or suspicious individuals.
- The project can be extended to support **multiple CCTV cameras**, allowing large areas such as malls, banks, and warehouses to be monitored simultaneously.
- Cloud integration can be implemented so that alerts, images, and videos are stored online and can be accessed remotely through mobile or web applications.
- Future development may include automatic alarm systems and mobile notifications to instantly inform security personnel whenever suspicious activity or theft is detected.

8. ACKNOLEDGEMENT

The authors would like to express their sincere gratitude to their project guide for continuous guidance, valuable suggestions, and support throughout the development of this research. We also thank the faculty members of the Department of Engineering for their encouragement and assistance. We acknowledge the contributions of the open-source communities of Ultralytics for YOLOv8, dlib and face_recognition for face recognition tools, and MediaPipe for the pose estimation framework. Finally, we thank our institution for providing the necessary infrastructure and resources to complete this work.

9. REFERENCE

[1] Singla, Surbhi, and Raman Chadha. "Detecting Criminal Activities From CCTV by using Object Detection and machine Learning Algorithms." Institute of Electrical and Electronics Engineers (IEEE), 2023.

[2] Zhou, Yang, Xianghua Xu, and Ran Wang. "EI-YOLO: Efficiently Improved YOLO on Detection of Prohibited Items During Security Inspections." Springer, 2024.

[3] Sainath, N. Omkar, et al. "YOLO-V7 to Overcome Theft Activities." Institute of Electrical and Electronics Engineers (IEEE), 2025.

[4] Awang, Suryanti, Mohd Qhairel Rafiqi Rokei, and Junaida Sulaiman. "Suspicious Activity Trigger System using YOLOv6 Convolutional Neural Network (SATS)." Institute of Electrical and Electronics Engineers (IEEE), 2023.

[5] Koca, Murat. "Real-time security risk assessment from CCTV using hand gesture recognition." IEEE, 2024.

[6] Ranjith, D., Muskaan Thabassum, Kanraj, Dhanush, and Bn Bharath Kumar. "Thief guard: real-time theft detection system integrating yolo and gemini for enhanced accuracy and rapid response." International Research Journal of Modernization in Engineering Technology and Science, 2025.

[7] Das, S. K., and T. Singh. "Video-based Real-time Intrusion Detection System using Deep Learning for Smart City Applications." IEEE ANTS (Advanced Networks and Telecommunications Systems), 2020.

[8] Saldanha, V., Naik, K. D., and Parashtekar, S. S. "Intelligent Surveillance and Theft Detection Using Real Time Object Detection." ResearchGate, 2025.

*[9] Gim, U., Lee, J., Kim, J., Park, Y., & Nasridinov, A. (2020). An Automatic Shoplifting Detection from Surveillance Videos (Student Abstract). In *The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20).*

[10] Automatic Theft Detection System At Smart Home Using AI Mohammed Sherooq Ali[1,], Sreeju A[2], Shifa Reem[3], Jasna[4], Neethu Mathew[5]*