AI-Powered Real-Time Theft Detection System for Shopping Malls

Project Team

Muhammad Nadeem-22P-9366

Laiba Shahi-22P-9247

Sultan Mehmood Mughal-22P-9242

Session 2022-2025

Supervised by

Dr. Hafeez Anwar

Department of Computer Science

National University of Computer and Emerging Sciences
Peshawar, Pakistan

June, 2022

Student's Declaration

We declare that this project titled "AI-Powered Real-Time Theft Detection System for Shopping Malls", submitted as requirement for the award of degree of Bachelors in Computer Science, does not contain any material previously submitted for a degree in any university; and that to the best of our knowledge, it does not contain any materials previously published or written by another person except where due reference is made in the text.

We understand that the management of Department of Computer Science, National University of Computer and Emerging Sciences, has a zero tolerance policy towards plagiarism. Therefore, We, as authors of the above-mentioned thesis, solemnly declare that no portion of our thesis has been plagiarized and any material used in the thesis from other sources is properly referenced.

We further understand that if we are found guilty of any form of plagiarism in the thesis work even after graduation, the University reserves the right to revoke our BS degree.

Muhammad Nadeem Signature: ________________

Laiba Shahi Signature: ________________

Sultan Mehmood Mughal Signature: ________________

Verified by Plagiarism Cell Officer

Dated:

Certificate of Approval

The Department of Computer Science, National University of Computer and Emerging Sciences, accepts this thesis titled AI Powered Real-Time Theft Detection System For Shopping Malls”, submitted by Muhammad Nadeem(22P-9366), Laiba Shahi(22P-9247), Sultan Mehmood Mughal(22P-9242), in its current form, and it is satisfying the dissertation requirements for the award of Bachelors Degree in Computer Science.

Supervisor

Mr./ Ms./ Dr. Supervisor Name   Signature: _________________

Sir Riaz Nawab
FYP Coordinator
National University of Computer and Emerging Sciences, Peshawar

Dr. Qasim Jan
HoD of Department of Computer Science
National University of Computer and Emerging Sciences

Acknowledgements

Abstract

This project focuses on developing an AI-powered real-time theft detection system for shopping malls. The system aims to automatically detect suspicious activities using computer vision and deep learning.The proposed system will use computer vision and deep learning to detect people and products in video footage and analyze their movements. If suspicious behavior, such as hiding items, is detected, an alert is instantly generated and shown on a dashboard. This real-time detection helps reduce losses, improve surveillance efficiency, and support security teams in responding quickly to incidents. The system provides an intelligent, low-cost, and scalable solution to enhance safety in retail environments.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

The AI-powered theft detection system of shopping malls is an intelligent surveillance solution that makes use of computer vision and deep learning to analyze CCTV footage in real time. It is designed to automatically identify suspicious behaviors like hiding products, loitering near shelves, or making unusual hand movements and then generate instant alerts for mall security personnel.

The system will be designed to achieve several key improvements in mall internal security operations by reducing human dependency and enhancing response time, while continuously and reliably monitoring multiple camera feeds.

1.2 Background

Shopping malls are high-traffic public areas where many people interact with products on display every day. Classic surveillance in these places is usually performed by the human resources department, involving some security personnel who are tasked with observing up to dozens of different CCTV screens simultaneously. This most definitely leads to fatigue, distraction, and overlooking a crime incident due to that very approach.

Present developments in AI and deep learning have finally enabled automated video analysis to detect and classify human activities with a high degree of accuracy. Integrating AI into mall surveillance systems enables continuous monitoring of all areas, spotting theft-related activities, and raising immediate alerts, making the system faster, smarter, and more reliable than conventional methods.

1.3 Motivation

The increasing number of shoplifting and theft cases in shopping malls is emerging as a growing concern in the eyes of both the mall management and law enforcement agencies. Traditional monitoring systems, using CCTV cameras, cannot handle these large crowds where occlusions are quite common.

An AI-based theft detection system provides a robust solution by ensuring 24/7 automated monitoring, reducing human error, and improving response times. This project is motivated to develop a cost-effective, intelligent surveillance system in enhancing safety, reducing financial losses, and generally improving mall security management.

1.4 Problem Definition

Moreover, most of the existing surveillance systems in shopping malls have been designed around human operators who cannot effectively monitor a large number of cameras for a long period. Due to this, many incidents related to theft are either not noticed or are notified well after the incident has happened. In addition, occlusions, multiple light sources, and crowded conditions further worsen the effectiveness of observation.

Therefore, there is a strong demand for a real-time AI-powered detection system that could automatically detect and identify suspicious activities or theft-related incidents from shopping mall CCTV footage to ensure quicker alerts and timely action by the security team.

1.5 Purpose of the Study

This paper is dedicated to the design and implementation of an AI-powered real-time theft detection system for shopping malls. The system automatically analyzes surveillance footage for the identification of actions that show traces of theft and issues alerts without necessarily requiring the full-time engagement of human services. This project will enhance the safety and operational efficiency of mall security systems by providing accurate, real-time detection of theft events.

1.6 Significance of the Project

The significance of this project rests in the fact that it incorporates AI into the conventional mall surveillance infrastructure, making it smart and automated. The proposed model will:

- Improve the accuracy and reliability in the detection of theft in crowded shopping malls.
- Reduce manual workload and human dependency for monitoring multiple CCTV screens.
- Use real-time alerts for quick responses on any theft issues.
- Support evidence-based reporting for mall security and law enforcement. The adoption of such a system will afford shopping malls a far better sense of safety and operational efficiency, ensuring a secure shopping experience for both customers and retailers.

Chapter 2

Review of Literature

Theft detection is important for reducing losses and improving mall security. With the increase of CCTV cameras, AI-based systems are now used to automatically detect suspicious activities. Traditional methods often fail in crowded areas and can't handle complex situations.

Deep learning models like CNNs, 3D-CNNs, BiLSTM, and Transformers can now analyze video frames to recognize theft-related actions more accurately. This chapter reviews recent research, their methods, and results, which help build the foundation for our proposed AI-powered theft detection system.

2.1 CNN-Based Theft Detection

Convolutional Neural Networks (CNNs) are often used for image and video analysis because they can detect people, objects, and actions in video frames. In theft detection, CNNs help recognize suspicious movements quickly. However, they struggle to understand time-based patterns, which makes it hard to tell normal actions from theft behavior.

2.2 Hybrid CNN-RNN Approaches

Hybrid models combine CNNs for image features and RNNs or BiLSTMs for learning time-based patterns. These models can detect actions like hiding or moving items more accurately. Still, they need large datasets and have difficulty handling crowded scenes or multiple people at once.

2.3 3D-CNN Approaches

3D-CNNs process video as a sequence instead of single frames, allowing them to capture both motion and appearance. They work well for identifying theft movements but require powerful hardware and large, diverse datasets to perform effectively.

2.4 Transformer-Based Theft Detection

Transformers use attention mechanisms to understand long-term movement patterns. They can detect unusual body poses or actions. However, they often rely only on pose data and may not perform well in complex mall scenes with multiple people or occlusions.

Year	Title	Methodology	Results	Proposed Solution
2025	Anomaly detection in shopping centers (RLCNN)	Hybrid RLCNN (RNN + CNN) on UCF50, UCF101, UCF-YouTube datasets	96–97% accuracy; effective on benchmark dataset	Collect local store videos with traditional attire & real shopping behaviors
2025	Pose-Based Anomaly Detection (PoseLift)	YOLOv5 + activity recognition on human poses	80% accuracy; privacy-preserving	Collect videos with diverse clothing, body types & shopping styles
2024	ML-Based Theft Detection Using YOLO	YOLOv5 + activity recognition	Fast real-time detection; high-quality results	Create labeled local dataset including multiple shoppers per scene
2023	Shoplifting Detection Using	CNN + BiLSTM on	81% accuracy; multiple theft	Collect diverse videos, augment

	CNN-BiLSTM	900-video custom datase	patterns detected	data, add multi-camera setup
2025	Shopformer: Transformer-Based Detection	Transformer on skeletal keypoints with self-attention	Outperformed CNN-LSTM & 3D-CNN; captured long-range dependencies	Add object detection, multi-camera videos, include diverse cultural behaviors
2020	Suspicious Behavior Detection (3D CNN)	3D-CNN for spatiotemporal feature extraction	~75% accuracy; detected subtle movements	Collect larger dataset with varied lighting, angles & shoppers

2.1 Literature Review

2.5 Limitations of Existing Approaches

Despite notable progress, existing theft detection methods have several limitations:

1. **Dataset Diversity:** Most studies rely on benchmark datasets that lack variations in cultural attire, shopping behavior, and multi-person scenarios, reducing real-world applicability.

2. **Occlusion and Clutter:** Crowded environments with occluded hands, objects, or faces reduce detection accuracy.

3. **Camera Constraints:** Single-camera setups limit coverage and the detection of multiple simultaneous thefts.

4. **Incomplete Feature Modeling:** Pose-only or object-only methods miss comprehensive cues of suspicious behavior.

5. **Computational Complexity:** Advanced models like 3D-CNNs and transformers require high computational power, hindering real-time deployment.

2.6 Proposed Solutions

To address the limitations identified in previous studies, the proposed **Meta-Theft Detection Framework** incorporates the following solutions:

- **Hybrid Architecture:** Combines **RLCNN, CNN, and BiLSTM layers** to extract both spatial and temporal features efficiently.

- **Multi-Modal Input:** Integrates **pose estimation, object detection, and motion tracking** to capture full human-object interactions.
- **Multi-Camera Setup:** Utilizes **multiple surveillance angles** to detect simultaneous thefts and mitigate occlusion issues.
- **Diverse Dataset:** Employs a **locally collected dataset** representing cultural attire, traditional shopping behavior, and multi-person scenarios for robust training.
- **Real-Time Optimization:** Enhances model efficiency for **fast inference** without compromising detection accuracy, suitable for deployment in retail environments.

These solutions collectively aim to achieve **high accuracy, robustness, and real-time performance**, addressing the gaps in existing approaches.

Chapter 3

3.1 Problem Statement

In these crowded shopping malls, existing theft detection systems often fail due to occlusion, a limited dataset, and poor generalization capability. These systems are unable to accurately detect multiple theft incidents within a single camera frame and often generate false or missing alerts.

Therefore, a robust deep learning-based surveillance system for the detection and recognition of theft-related activities is greatly needed in real time under complex conditions with multiple people in a mall, lighting variations, and different camera angles.

3.2 Business Opportunity

Retail theft has been one of the major areas of financial losses for several shopping malls and retail businesses around the world. Most traditional manual surveillance systems are generally inefficient, labor-intensive, and expensive, as they mainly depend on continual human observation.

An AI-powered theft detection system addresses a business opportunity by automating surveillance tasks. It would be able to:

Reduce manpower requirements and operational costs.

Reduce theft incidents by alerting on time.

Improve customer safety and enhance general mall security.

This system can be implemented by retail organizations and mall management to ensure intelligent, scalable, real-time monitoring with enhanced security and improved operational efficiency.

3.3 Objectives

Primary Objective

The design and implementation of an intelligent, deep learning-based system for detecting multiple shoplifting incidents in real time within shopping malls.

Specific Objectives

The model development identifies activities related to theft; for example, hiding items under traditional clothes such as shalwar kameez.

For reliable detection in crowded areas with occlusion and clutter.

Improving Detection Accuracy and Performance with a Large, Diverse, and Locally Collected Dataset.

Generate instant alerts and store evidence (video clip or frame) for further review automatically.

3.4 Project Scope

The proposed system focuses on enhancing CCTV-based surveillance in shopping malls by using techniques of AI and computer vision for real-time theft detection.

The following activities are included within the project scope:

Application limited to shopping mall environments.

Detection limited to human activities related to theft only.

Support for multi-camera integration and real-time video processing.

Scalability to extend to other retail settings like supermarkets and hypermarkets.

3.5 Constraints

Limitations may include the following in the development and implementation of the system:

Limited availability of labeled real-world theft datasets for training and testing.

Differences in lighting conditions and angles of cameras at various malls.

Hardware limitations of low-end or edge devices that could impact real-time processing.

Privacy and ethics concerning the use of surveillance footage for training AI models.

3.6 Stakeholders Description

The main stakeholders of this project are the mall management, security personnel, developers, project supervisor, and mall customers.

The mall management is the primary end user who will be responsible for deploying and managing the operation of the system. Security personnel are system operators responsible for monitoring alerts and taking action on possible theft incidents in real time. Developers are the technical team for the design, training, and maintenance of the AI-based system. The project supervisor provides technical supervision, guidance, and ensures that the project aligns with academic and research objectives. Customers are the indirect stakeholders who benefit from improved safety and security within the shopping mall environment.

3.7 Key High-Level Goals

To develop real-time AI-based theft detection in shopping malls.

Reduce financial losses by proactive generation of alerts and faster responses. To improve public safety by automating and intelligently monitoring with AI. The goal is to enhance surveillance efficiency while reducing human workload.

Chapter 4
4 Software Requirements Specifications

This chapter will have the functional and non functional requirements of the project.

4.1 List of Features

Real-time video surveillance analysis.

Theft activity detection and classification.

Automatic alert notifications.

Evidence storage for each incident.

Admin dashboard with live feed monitoring.

4.2 Functional Requirements

The system shall detect suspicious activities in video streams.

The system shall generate an alert with camera ID and timestamp.

The system shall store event clips as evidence.

The admin shall access the dashboard to review past incidents.

4.3 Quality Attributes

Performance: Real-time analysis at ≥30 FPS.

Accuracy: ≥90% detection accuracy.

Security: Encrypted storage for video and alert data.

Scalability: Support for multiple simultaneous camera inputs.

Usability: Simple interface for security personnel.

4.4 Non-Functional Requirements

Reliability – Consistent detection across environments.

Maintainability – Easy to update models and components.

Efficiency – Optimized for GPU and edge processing.

Portability – Compatible with multiple hardware systems.

Robustness – Stable performance under noise and lighting variations.

4.5 Test Plan (Test Level, Testing Techniques)

The system will undergo four levels of testing:

- **Unit Testing:** To verify each module (YOLO detection, alert system) works correctly.
- **Integration Testing:** To ensure smooth interaction between detection, alert, and database modules.
- **System Testing:** To test the complete system performance and real-time accuracy.
- **User Acceptance Testing:** To confirm usability and reliability under real-world conditions.

Techniques Used:

- **Black Box Testing** for functionality,
- **White Box Testing** for internal logic,
- **Performance Testing** for speed and accuracy,
- **Regression Testing** after updates.

4.6 Software Development Plan

The system is developed using the Agile Model, allowing iterative improvement and testing.

Phases:

1. Requirement Analysis
2. Dataset Collection & Annotation
3. Model Training (YOLOv5 + LSTM)
4. Integration & GUI Development
5. Testing & Validation
6. Documentation & Deployment

Tools: Python, PyTorch, OpenCV, Flask, LabelImg, and Google Colab.
The final outcome is a real-time theft detection system with high accuracy and fast response.

Chapter 5

Iteration Plan

Midterm FYP 1

1. Dataset collection
2. Diagrams (Use Case,Flow Chart,Sequence) etc made.
3. 20% work done

Final FYP 1

1. Designing Poster
2. Preprocessing and fine tuning of the collected dataset.
3. Model training (initial phase)

Midterm FYP 2

1. Model training (continuation)
2. Testing of the model trained
3. Making the dashboard,Frontend

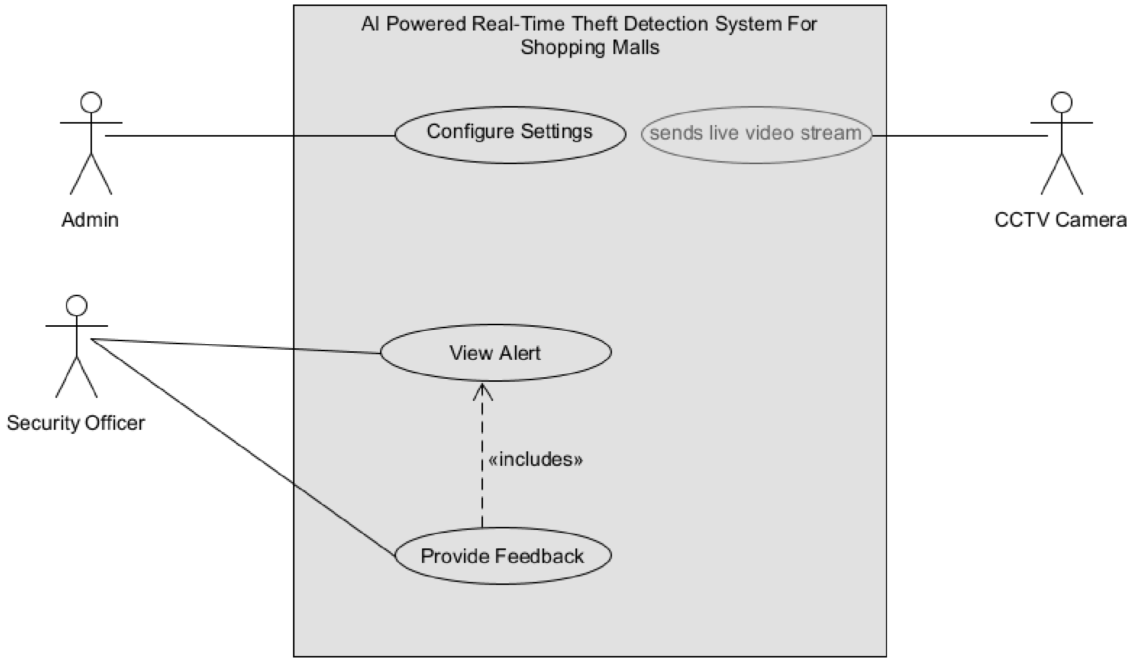
Final FYP 2

1. Integration with the system
2. Deploying the software
3. Reviewing the feedback

Chapter 6

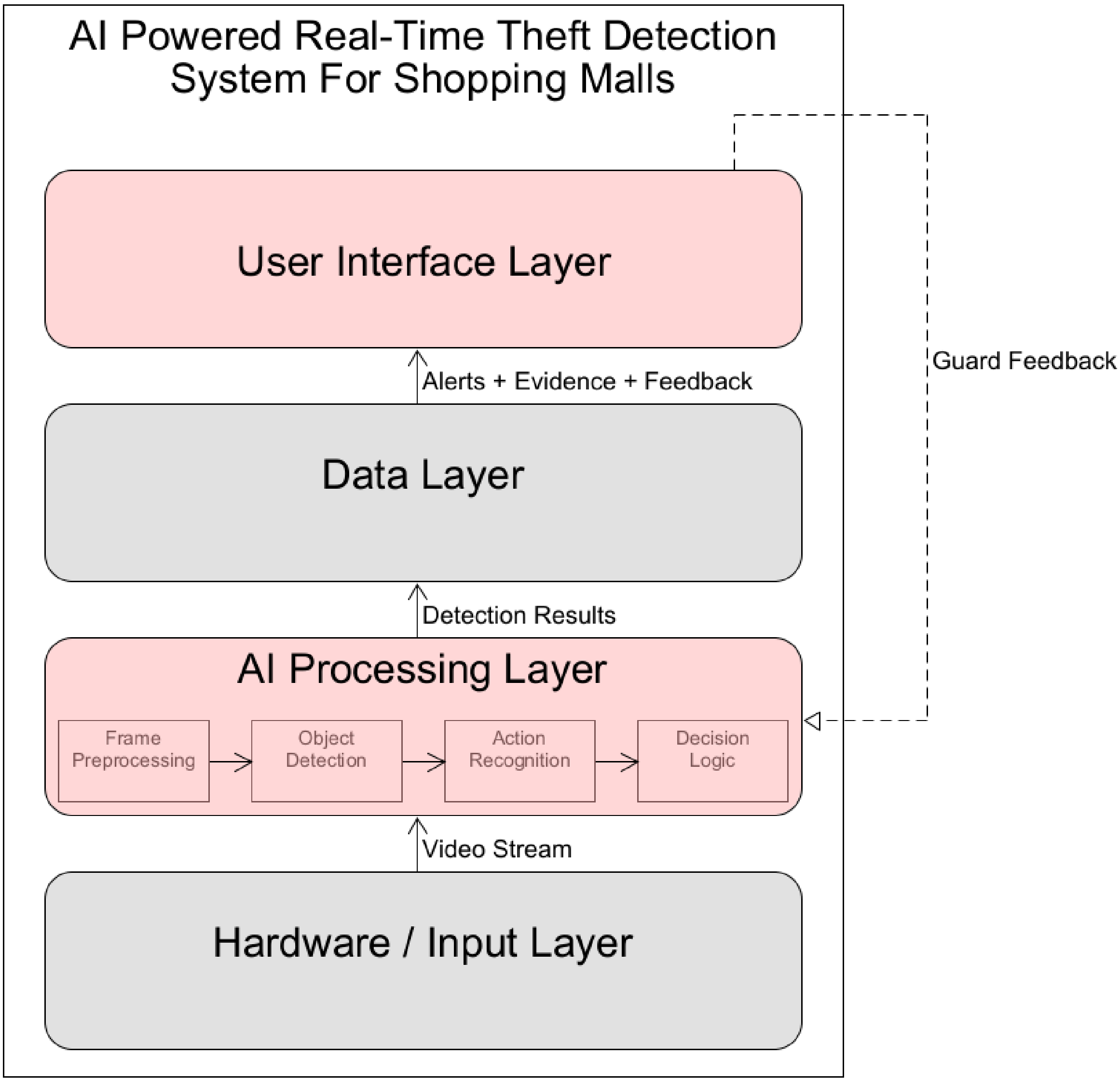
6.1 Structural design

6.1.1 Use Case Diagram



CCTV cameras continuously send live video streams to the AI model, which analyzes the footage to identify suspicious activities. The **Admin** is responsible for configuring system settings, such as camera connections and alert preferences. When the AI detects potential theft, an **alert** is sent to the **Security Officer**, who can immediately view the alert and take necessary action. After reviewing the alert, the Security Officer can **provide feedback** to the system, helping improve its accuracy over time. This feedback process allows the AI to learn from real incidents and reduce false alarms.

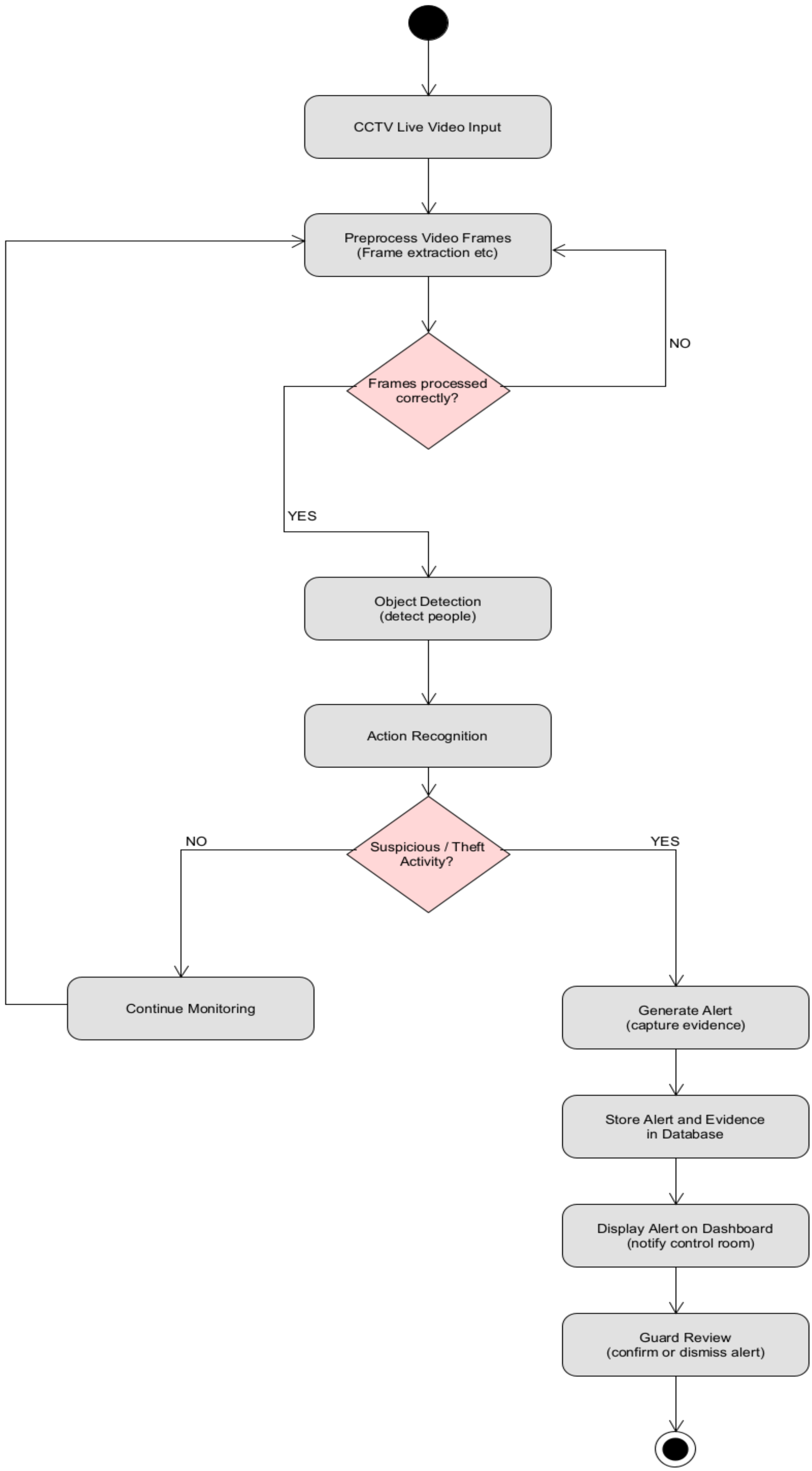
6.1.2 System Architecture Diagram



The diagram shows the overall structure and working mechanism of the AI-powered theft detection system. It is divided into multiple layers, each layer is responsible for some functions. The Input Layer consists of CCTV cameras that continuously capture real-time video from the shopping mall. The Processing Layer includes the core AI components such as the Frame Ingestor, Preprocessing Unit, Object Detection Model, and Decision-Making Module. These components work together to analyze video frames and identify suspicious activities using Deep learning and computer vision techniques. The Storage Layer contains a central Database where all alerts, logs, and video evidence are stored for future reference. On top of this, the Application Layer provides an interface (Dashboard) for the Admin and Security Officers to interact with the system, monitor activities, and manage alerts. At the end,the notification layer handles real-time alerting by notifying security staff whenever theft or unusual behavior is detected.

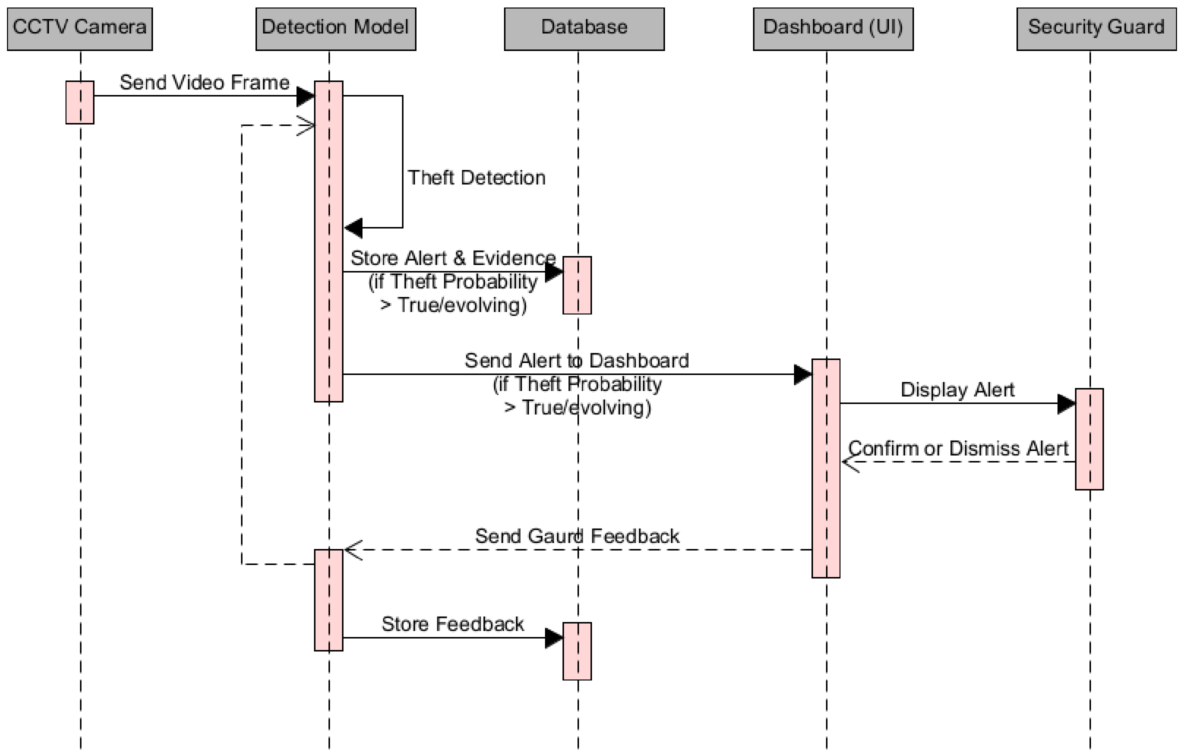
6.2 Behaviour Design:-

6.2.1 Flow Chart


CCTV Live Video Input
Preprocess Video Frames
(Frame extraction etc)
NO
Frames processed
correctly?
YES
Object Detection
(detect people)
Action Recognition
NO
Suspicious / Theft
Activity?
YES
Continue Monitoring
Generate Alert
(capture evidence)
Store Alert and Evidence
in Database
Display Alert on Dashboard
(notify control room)
Guard Review
(confirm or dismiss alert)

The system begins by receiving live video input from CCTV cameras installed throughout the shopping mall. The video feed is then preprocessed by extracting and cleaning frames to ensure that the data is suitable for analysis. Once the frames are correctly processed, the AI model performs **object detection** to identify people and other relevant objects within the video. After this, **action recognition** is carried out to analyze the behavior of individuals and determine their activities. If the AI detects suspicious or theft-related behavior, it triggers an **alert**, captures evidence such as images or short video clips, and stores them in the database. The alert is then displayed on the **dashboard** in the control room, notifying the security officers immediately. The security guard reviews the alert to confirm whether it represents an actual theft or a false alarm. Based on this feedback, the system continues monitoring and improves its accuracy over time.

6.2.2 Sequence Diagram



CCTV cameras send live video frames to the detection model, which analyzes them to identify suspicious or theft-related activities. If the model detects a potential theft, it stores the alert and evidence in the database and sends an alert notification to the dashboard (UI). The security guard views the alert on the dashboard and either confirms or dismisses it based on observation. The guard's feedback is then sent back to the detection model and saved in the database to help refine and improve the system's accuracy.

6.3 Algorithm Design

Algorithm design defines the logical steps that enable the system to detect, verify, and report theft in real time. The AI-Based Theft Detection System operates through a sequence of computational algorithms that convert video input into actionable alerts.

The following algorithms are involved:

1.Frame Processing

2.Object Detection

3.Logic and Decision Engine

4.Alert,Feedback

6.4 Unit Test

Unit testing focuses on verifying that each individual component of the system works correctly in isolation.For this project, unit tests are applied to functions below:

- Frame ingestion (checking if frames are captured correctly)
- Preprocessing (verifying image normalization and resizing)
- Object detection model (ensuring correct detection outputs)
- Alert generation (checking if alerts are triggered with proper conditions)

Test Cases

Test	Test Case Example	Expected Result
Frame Capture	Test video stream connection	Frame captured successfully
Preprocessing	Verify resizing of frames	Frame resized to model input size
Object Detection	Running model on test image	Person detected correctly
Alert System	Simulate theft behavior	Alert generated with correct message
Database	Insert and retrieve alerts	Data stored and fetched successfully
Dashboard	Display active alerts	Alerts visible to admin/security

Table 6.1 *Test Cases*

6.5 Maintainable Phase

The Maintainable Phase aims for system updates, debugging, and improvement after initial deployment.
It ensures that:

- The model can be retrained with new data.
- A security officer's feedback can be used to fine-tune the detection algorithm.
- Any new camera configurations or locations can be added easily.

Goal:
To make the system scalable, adaptable, and easy to modify in the future.