# FYP SRS DOCUMENT

Final Year Project

Software Requirement Specification For

AcRe: A Yolo Powered Activity Recognition Framework Learning

(BSCS)

By


| S# | Name | Registration #/Roll #/Section | Mobile # | E-Mail |
|---|---|---|---|---|
| 1. | Faizan Moin | SP-21/BSCS/001/A | 0316-0471420 | Sp21-bscs-001@lgu.edu.pk |
| **2.** | Sohail Karim | SP-21/BSCS/010/A | 0349-5074475 | Sp21-bscs-010@lgu.edu.pk |

**Supervised by:**

Mr. Omer Aftab

_____

(Signature)

Department of Computer Science

Lahore Garrison University

Lahore

# Table of Contents

# 1. Introduction

## 1.1 Purpose

This project, “AcRe: A Yolo Powered Activity Recognition Framework”, leverages YOLO models for real-time detection of anomalous activities such as smoking, harassment, fighting and gun violence in workplace environment. Traditional methods often fail to provide timely and accurate detection which lead to safety and security of many lives. In this project we are integrating advance YOLO models which not only effectively detect anomalous activities but it also timely detects. The main purpose of this project is to achieve fast, high-accuracy detection of multiple anomalous activities simultaneously that are used in different applications like security, healthcare and even in city safety cameras. In this project we are using Computer vision a branch of artificial intelligence which uses AI techniques to analyze and understand visual data. It can play a crucial role in identifying and addressing anomalous activities. To surmount these challenges the integration of Artificial Intelligence and computer vision holds to dectation of such activities within designated times frames. Activity Recognition is an essential component in various applications such as surveillance systems, security analysis and violence. This Yolo based framework is capable of accurately identifying and classifying human activities in real-time. The main purpose is to leverage speed and accuracy of Yolo to process video frames efficiently, providing real-time feedback on detected activities. The main motive is to detect the anomalous activities like harassment, gun violence, smoking at an early stage to overcome such cases coming out of the work places environment.

## 1.2 Document Conventions

The documents conventions being followed are, the font style for the SRS is times news Roman. For the headings the heading 1 style is being used for main heading with a font size of 16 while for sub-headings the heading 2 style is being used with the font no 13. For the paragraph content the font size is 10.5.

## 1.3 Intended Audience and Reading Suggestions

The intended audience for the proposed methodology “AcRe: A Yolo Powered Activity Recognition Framework” includes those people who are relevant to field of computer science or Artificial intelligence and for those clients who interested in benefiting from AI for their safety and security purpose. Technical audience like developers to understand conceptual and technical technologies requirements for putting system into operation, project managers who should coordinate on project deadline and timeframes, testers whose responsibility to test according to specified requirements. The topics discussed in this SRS with respect to the project are introduction, overall description, external interface requirements system’s featured and non-functional requirements. The introduction describes the basic purpose, document conventions, reading suggestions for audience and project scope. The Overall Description describes the product perspective, product functions, user classes, operating environment, design and implementation constraints, user documentation, assumptions and dependencies. The next is External Interface Environment that consists of user interfaces, hardware interfaces, software interfaces and communication interfaces. Further next is the System Features that includes the system features,

description and priority, stimulus sequence and functional requirements. The last one is the other non-functional requirements that further consists of performance requirements, safety requirements, software quality attributes and business rules.

### 1.4 Product Scope

- Achieving fast and high accuracy by using latest YOLO models (say YOLOv7, YOLOv8 or YOLOv9).
- Early stage detection of anomalous activities in a single framework
  Detection of multiple anomalous activities in a single framework.
- GUI interface for real-time demonstration.
- The framework aims to enhance safety and security by providing timely alerts via beep sound or notifications.

## 2. Overall Description

### 2.1 Product Perspective

Security is one of the key concerns for any person, either at home or in public places. Surveillance cameras are now being used in almost every area, be it home or public place. Security guards sit behind those CCTV screens, inspecting the behaviour of people, and monitoring their activities. But unfortunately, the human brain cannot focus on multiple screens at a single time. Several applications are being developed to automatically detect abnormal activities, therefore, the security can be ensured, and people can move without safety risks and concerns. We have developed an application that automatically detects abnormal activity in surveillance CCTV streams and generates an alarm. This project, "AcRe: A YOLO Powered Activity Recognition Framework", leverages YOLO models for real-time detection of anomalous activities such as smoking, harassment, fighting, and gun violence in workplace environments. Traditional methods often fail to provide timely and accurate detection, posing risks to safety and security. By integrating advanced YOLO models with additional technologies like deep learning-based image segmentation, this project aims to achieve fast, high-accuracy detection of multiple anomalous activities simultaneously

### 2.2 Product functions

The functions of the anomalous video detector are outlined below:

- Footage: CCTV footage is fed to the system.
- Preprocessing: Data preprocessing is applied and frames are fed
- Inferencing Model: Frames are fed to obtain probability outputs/results.
- Anomaly Detection: Anomaly is detected and key-event anomaly snippets are extracted.
- User Alert: Instant Notification is generated.

### 2.3 User Classes and Characteristics

The user class for AcRe platform includes any normal user who just have a basic knowledge of computer and CCTV footage. It can be a security guard, a peon or any normal person working in the respective environment. Additionally, researchers can also interact with the platform more deeply for the validation purposes. Design students and tech enthusiasts interested in deep learning or Yolo models can also engage with the system for educational insights. Moreover, academic researchers focused on AI-driven verification methods may study the platform for advancements in detection technology. While a diverse range of users can benefit from the system, professionals in digital security represent the primary audience, as their expertise enables them to utilize the platform effectively and gain the most accurate insights. Their educational background and experience make them best suited to utilize the system effectively.

## 2.4 Operating Environment

The respective operating environment for AcRe is designed to operate across a wide range of environments, making it accessible to media organizations, educational institutions, and offices. It can be effectively used in offices, universities and even as safe city projects, enhancing its credibility and integrity. The system is compatible with standard hardware like PCs.

## 2.5 Design and Implementation constraints

Several critical constraints govern the design and implementation of AcRe. For instance, the system needs to operate with high accuracy and real-time efficiency in order to promptly detect simple human activities as well as anomalous activities such as harassment, fighting, smoking or gun violence. The integration of Yolo ensures fast object detection but at the cost of huge computational requirements that necessitate robust hardware or optimized software to run on desktop systems. AcRe also has to balance between precision and recall, thus minimizing false positives and negatives to be reliable in security-critical scenarios. User interface design is further constrained by the need for simplicity and clarity, which targets non-technical users like security personnel. Ethical considerations such as privacy protection and reduction of bias in activity detection must be addressed. Scalability and adaptability also challenge the system; it needs to accommodate various environmental conditions, camera angles and lighting variations to make the system applicable in most real-world scenarios.

## 2.6 User Documentation

The user documentation will include a comprehensive user manual that provides all necessary information for effective use of the platform. This will guide user manually step by step through the features and functionalities. Moreover, video tutorials will be available to cater to users from non-technical backgrounds, ensuring they have a clear understanding of the platform capabilities. Further, users can access online help resources. This multi-faceted approach to user documentation aims to ensure that all users regardless of their technical expertise and can navigate the platform with confidence and ease.

## 2.7 Assumptions and Dependencies

The key element in building our AcRe, is gathering the right requirements especially the leveraging advanced technologies like deep learning and CNNs. The system needs a wide variety of authentic images and videos for training the AI models and these must be accurate and reflective of real-world scenarios. If the data is lacking or not up to standard, it can lead to misidentifications and inaccurate assessments. Ensuring that our datasets are reliable is crucial; any issues here could result in unsatisfactory user experiences and harm the platform's reputation. We also recognize that this project is entirely original, with no recycled code or borrowed components, which highlights our commitment to providing a unique and effective solution in the anomalous detection space.

# 3. External Interface Requirements

## 3.1 User Interface

This will mainly be a single user interaction framework. As it is an anomalous video detection framework and is integrated with the CCTV so only the focal persons on the duty such as security guards will have the access to see or interact with the framework. There would be probably two login options, one will be for an ordinary security guard who’s duty is to keep the check and balance on the CCTV footage. He will just have the access to see the footage. The other login will be for the security manager of that respective environment, who will have the privileges to add or remove users (security guards) from the system, have record of all the files in which the anomalous clip was detected and some more privileges likewise. The design will keep backend processes hidden, allowing users to focus on their tasks without confusion. This user-friendly approach ensures that everyone, regardless of their technical background, can navigate the platform easily and effectively, enhancing engagement and satisfaction.

## 3.2 Hardware Interfaces

The hardware devices needed for this project are laptops, PCs and tablets. The emphasis of the methodology is on the software part, these hardware devices are just needful to view the interface for the detections. Laptops and PCs will be used to view and check the saved part of the CCTV footage where the anomalous activity got detected. All detection processes will be handled within the computer’s CPU or GPU. User would simply require a device with basic computer capabilities and access to the login interface to view and check the detections.

## 3.3 Software Interfaces

The main tool under consideration for the development of the project is google colab that will be used to develop the front-end as well as back-end. For the construction of front-end gradio and flask will be used. Gradio and Flask are python web development and user interface libraries which are efficient and proven high-end interfaces which can be easily integrated with the backend

python file. The software interface is designed to be user-friendly and it requires no technical expertise. User or any security personnel just have to simply login the interface to view CCTV footage and to verify the detected portions. The simplicity of the interface guarantees that users from different backgrounds whether professionals in digital forensics or amateurs concerned about security and detections can use the system easily. The software product will run on the regular windows operating system containing PC's so there is no need of some High-Tech computers or super computers.

### 3.4 Communications Interfaces

The communication interface for AcRe serves as a seamless bridge between CCTV cameras and the desktop-based front end interface. This interface is to be intuitively operated, allowing for real-time data exchange, such that live footage or video streams from CCTV cameras are efficiently processed by YOLO model. This allows the system to pick all varieties of activity, ranging from routine behavior to anomaly incidents, and present them instant results on a user-friendly interface. With all this in place, leaves security personnel promptly identify and respond to peculiar activities without requiring technical expertise. The reliability of the communication interface is optimized for the performance and low latency, for ease of integration and making AcRe an excellent solution for most needs in modern surveillance.

## 4. System Features

The system features defined earlier will be explained in this portion as the context that how these features will work and be applied in the real time scenario

### 4.1 Footage

#### 4.1.1 Description and Priority

The main data source for the activity recognition capabilities of AcRe is footage. The system processes live video streams from CCTV cameras in real-time using YOLO v8 to identify routine and anomalous activities. These detected events are displayed on the front-end desktop-based user interface, providing immediate visibility. In case of any anomalous activity, the system sends an alert and saves the relevant portion of the footage along with timestamps for review, reporting, or evidence purposes. The priority of CCTV footage in AcRe is exceptionally high because the accuracy and timeliness of detection, alerts, and saved records depend entirely on the quality and accessibility of the video streams. This makes the effective handling of CCTV footage a critical component of the system's overall performance and reliability in enhancing security operations.

#### 4.1.2 Stimulus/Response Sequences

The stimulus/response sequences of footages in AcRe ensure real-time detection and handling of activities. The stimulus starts with the live feed from CCTV cameras, and the YOLO v8 model continuously analyzes them. When a routine activity is detected, the system updates the desktop front-end interface without any extra steps. But upon detection of anomalous behavior, the system triggers a sequence for immediate response in the following: there would be a desktop alarm where the desktop is filled with an audible beep, in addition to alerting other users that the computer

screen would then point to that particular detected anomaly and at the same time record it for possible later study. This structured stimulus/response mechanism ensures awareness of security incidents while preserving evidence for post-event review and decision-making.

### 4.1.3 Functional Requirements

**REQ1:** The functional requirements of CCTV Video is based on the following functional requirements: it must interact with CCTV cameras so that it can keep receiving and processing live video streams in real time, leveraging YOLO v8 for frame-by-frame analysis of footage to identify and classify simple and anomalous activities. Upon identification, the system will display the results clearly on the desktop front-end interface, providing a graphical representation of the current events. If there is abnormal activity, the system is required to trigger an audible beep and alarm to alert users as soon as possible. In addition, the specific segment of footage where the anomaly occurred must be retained along with a time stamp, ensuring that this information is accessible for assessment and evidentiary needs. Such provisions are paramount to establishing the dependability, responsiveness, and overall effectiveness of the system to enhance surveillance and security functions.

## 4.2 Preprocessing

### 4.2.1 Description and Priority

Preprocessing is the most important component of AcRe, as it allows for the optimization of CCTV footage to accurately and efficiently detect activities. This involves preparing video data by resizing, normalizing, and converting frames into a format compatible with the YOLO v8 model. Preprocessing reduces noise and rectifies lighting inconsistencies while extracting salient features, thus making it possible for the model to focus on relevant aspects of the footage. This step ensures that both normal and abnormal activities are detected at a high accuracy rate despite variations in environmental conditions like poor illumination or crowded frames. In AcRe, preprocessing is significantly very important as it directly influences the efficiency of the model and reliability of detection outcomes. Efficient preprocessing reduces the occurrence of false positives and guarantees real-time responsiveness.

### 4.2.2 Stimulus/Response Sequences

The preprocessing stimulus/response sequences used in AcRe are designed to ensure that the system works with accuracy and efficiency. The stimulus begins with raw video frames extracted from CCTV cameras, which can contain noise, fluctuations in lighting, and irrelevant details. As a response, the preprocessing module performs a series of operations: resizing frames to fit the input dimensions required by YOLO v8, normalizing pixel values, and removing artifacts that could interfere with detection accuracy. Once the frames are preprocessed, they are passed to the YOLO v8 model to detect activities. The preprocessing response ensures the correctness, uniformity, and refinement of input data in order to significantly enhance the capability of the model for both standard and anomalous activities identification and classification with ease. This sequence of

stimulus and response is key for maintaining the reliability and responsiveness of the system within the diversity of surveillance applications.

### 4.2.3 Functional Requirements

**REQ1:** The functional requirements of preprocessing in AcRe include activity recognition engines for precise and efficient detection of activities from CCTV footage. This preprocessing module resizes video frames to meet the required input dimensions by YOLO v8 and normalizes pixel values to provide a consistent performance model. It should also include noise reduction to remove artifacts, lighting inconsistencies, and enhance features critical for activity detection. The system should preprocess frames in real time to seamlessly integrate with live video feeds. It should also support the handling of frames from multiple cameras at the same time without introducing significant latency. All of these requirements will make the input data optimized and ready for use.

## 4.3 Inferencing Model

### 4.3.1 Description and Priority

The inferencing model is primarily responsible for analyzing CCTV footage with respect to simple and anomalous activities. This model operates based on the advanced architecture of deep learning of YOLO v8 by processing preprocessed frames of real-time video feeds that can identify objects and classifies behaviors with high precision. The inferencing model would correctly and rapidly distinguish routine behavior from anomalies to trigger alerts for specific video segments of interest promptly. Its performance determines reliability, responsiveness, and even the effectiveness of the whole system in maintaining integrity and surveillance. The inferencing model is very high priority because it constitutes the backbone of the detection pipeline, ensuring that actionable insights are generated in real time to enhance security and streamline responses to potential threats.

### 4.3.2 Stimulus/Response Sequences

The stimulus/response sequences of the inferencing model in AcRe are intended for real-time detection and action. The stimulus is initiated by preprocessed frames from CCTV footage, which are fed into the YOLO v8 inferencing model. In response, the model analyzes each frame to identify objects, classify activities, and determine whether the observed behavior is routine or anomalous. In case a routine activity is identified, it updates the desktop front-end with the classification. However, in case of the identification of an anomaly, the response sequence includes: activating an audible beep and alarm on the interface so that users are alerted towards the detected anomaly; underlining the identified anomaly graphically; and automatically saving the specific portion of video with a timestamp for the review. With a cycle of efficient stimulus response, it guarantees timely and recording of the critical event; therefore, the system makes security operations responsive and reliable.

### 4.3.3 Functional Requirements

**REQ1:** The functional requirements of the inferencing model make it capable of analyzing CCTV footage for real-time activity detection. The model has to process preprocessed video frames with accurate object detection, activity classification, and anomaly detection. It has to run in real-time

and handle continuous video streams from multiple CCTV cameras without causing any delay in detection. The model must differentiate between normal and abnormal activities with high accuracy and the lowest false positives. These requirements are essential for the inferencing model to deliver reliable, actionable insights and enhance the overall effectiveness of the surveillance system.

## 4.4 Anomaly Detection

### 4.4.1 Description and Priority

Anomaly detection is a key feature of AcRe. The system can identify unusual or potentially dangerous activities from CCTV footage. This capability is made possible by the advanced object detection and classification mechanisms available in YOLO v8 to differentiate between routine behaviors and anomalies in real-time. The system identifies anomalies and ensures an immediate response by causing an audible beep, alerting users to take swift action. The priority of anomaly detection is extremely high since it directly addresses the core objective of enhancing surveillance and security by providing timely alerts and maintaining accurate records of critical incidents.

### 4.4.2 Stimulus/Response Sequences

The stimulus/response sequences in anomaly detection of AcRe are quite essential for timely and efficient responses to unusual activities. The stimulus begins with the live video frames from the CCTV footage, which will be analyzed by the model of YOLO v8 to detect objects and classify activities. In the event of system identification of an anomalous activity, it sets up a response sequence comprising of three primary actions. They are as follows: activation of audible beep and alarm on the desktop front-end, visual highlighting of detected anomalies on the interface, and automatic saving of specific video segments of anomalous activity with timestamp for analysis or evidence purposes in future. This stimulus/response cycle ensures anomalies are addressed real time but crucial data may be kept for post-incident review to make the whole system very effective for use in surveillance and security cases.

### 4.4.3 Functional Requirements

**REQ1:** The functionality of abnormality detection will definitely be essential for ascertaining whether or not an activity captured on the CCTV can indeed be well-defined and recognized within real time. This can be so as it processes video frames for deep learning classification and categorizing activities as either normal or abnormality. The anomaly detection process should be efficient and should be able to process footage from multiple cameras in real-time with minimal latency. This is essential to ensure that AcRe can detect security threats appropriately, raise alerts in real-time, and keep an accurate record for analysis and action.

## 4.5 User Alert

### 4.5.1 Description and Priority

The user alert system of AcRe will be used to provide an alert on time to anomalies that are identified from the CCTV footage. Upon detecting an anomaly, the system triggers an immediate beep and alarm on the desktop front-end interface so that the security personnel can get informed about the issue instantly. The system will be both audio and visual in nature, where an identified anomaly is displayed and emphasized on the display. User alert priority is therefore at a high level due to its direct implication to the system's capability in alerting security personnel of possible quick reactions. A potential security threat could slip unnoticed without immediate alerts. The effectiveness of surveillance can then be compromised.

### 4.5.2 Stimulus/Response Sequences

The stimulus/response sequence aims to ensure that there are immediate notifications of anomaly events to security personnel. Once the YOLO v8 model has detected any anomalous activity while processing live CCTV footage, this stimulates the process. As soon as it detects an anomaly, the system initiates a response sequence. This sequence will ensure that the user is immediately notified, allowing for swift response to potential security threats while preserving evidence for later analysis or reporting.

### 4.5.3 Functional Requirements

**REQ1:** The functional requirements are meant to effectively support real-time alerts to the security personnel. The occurrence of any anomalous activity via the live CCTV feed, then the system should sound the audible beep and activate visual alarm at the desktop interface to inform the user instantaneously. The alert should be distinguishable and attention-grabbing for the prompt recognition of the anomaly. Furthermore, the system should indicate the identified anomaly on the front-end for easy identification. The alert system should work in real-time with minimum delay between anomaly detection and notification to ensure security personnel respond rapidly to the identified threats.

## 5. Non Functional Requirements

### 5.1 Performance Requirements

The performance requirements of AcRe are mainly to ensure operation efficiency and effectiveness under actual real-time conditions. Ideally, the system must monitor video streams from multiple cameras simultaneously without significant latency on detection of both simple and anomaly activities in real time; detection accuracy must be high, showing minimal false positives for robust anomaly identification. The user interface has to be responsive, alerts and notifications have to start flashing immediately when an anomaly is detected. The system has to be able to save video segments quickly in real time without causing lags or system crashes. Overall, it should be robust, scalable, and capable of operation continuously without degradation in performance to support seamless surveillance.

### 5.2 Safety Requirements

The safety requirements will protect both the system and its users from potential risks, thus ensuring the integrity and confidentiality of the surveillance data. The system will incorporate

secure data transmission protocols to prevent unauthorized access to live CCTV feeds and recorded footage. It would also ensure that saved footage, especially anomalous activity segments, is stored securely, with encryption mechanisms in place to safeguard sensitive information. The system must restrict user access through authentication methods such as passwords or biometric verification to prevent unauthorized users from tampering with the data or system settings. These safety requirements are crucial to the reliability and security of the system in high-stakes surveillance environments.

### 5.3 Software Quality Attributes

The software quality are very important in achieving high standards of performance, usability, and reliability. This is because the system, in this case, will be expected to run in a failure-free mode such that there is uninterrupted surveillance and detection. Another important attribute is scalability: the system should accommodate various workloads, including additional CCTV cameras or greater video resolution, without impairing performance. Efficiency is essential, especially in terms of processing high volumes of video data in real time and minimizing latency in detection and response. Maintainability is key for long-term success, with the software designed for easy updates and bug fixes. Thus, these quality attributes guarantee that AcRe provides an effective, user-friendly, and secure solution for real-time activity detection and anomaly alerts.

### 5.4 Business Rules

The business rules guarantee that the system is made to conform to organizational aims and operational standards. The 24/7 operation at high availability ensures that this system of continuous monitoring will not experience any downtime related to the CCTV footage monitored. Besides those, it should also honor policies on data retention, and video footage, especially a segment containing anomalous activity, should be stored properly for a predetermined period in order to be automatically erased. System performance benchmarks must also be met, and the detection process should be fast enough to allow for real-time alerts and to minimize response delays. These business rules are crucial to ensure that AcRe works effectively, adheres to legal and organizational standards, and meets the operational needs of its user.

## 6. Detailed Architecture and Design

### 6.1. System Architecture

- **Architecture Design Approach:**
  The system follows a modular architecture with a FastAPI-based backend, YOLO model for object detection, and a SQLite database for storing predictions. This design ensures scalability, maintainability, and efficient processing.
- **Architecture Design:**
  Key components:
  1. **Frontend:** A web interface built with Jinja2 templates for user interaction.
  2. **Backend:** FastAPI to handle HTTP requests and integrate the YOLO model.

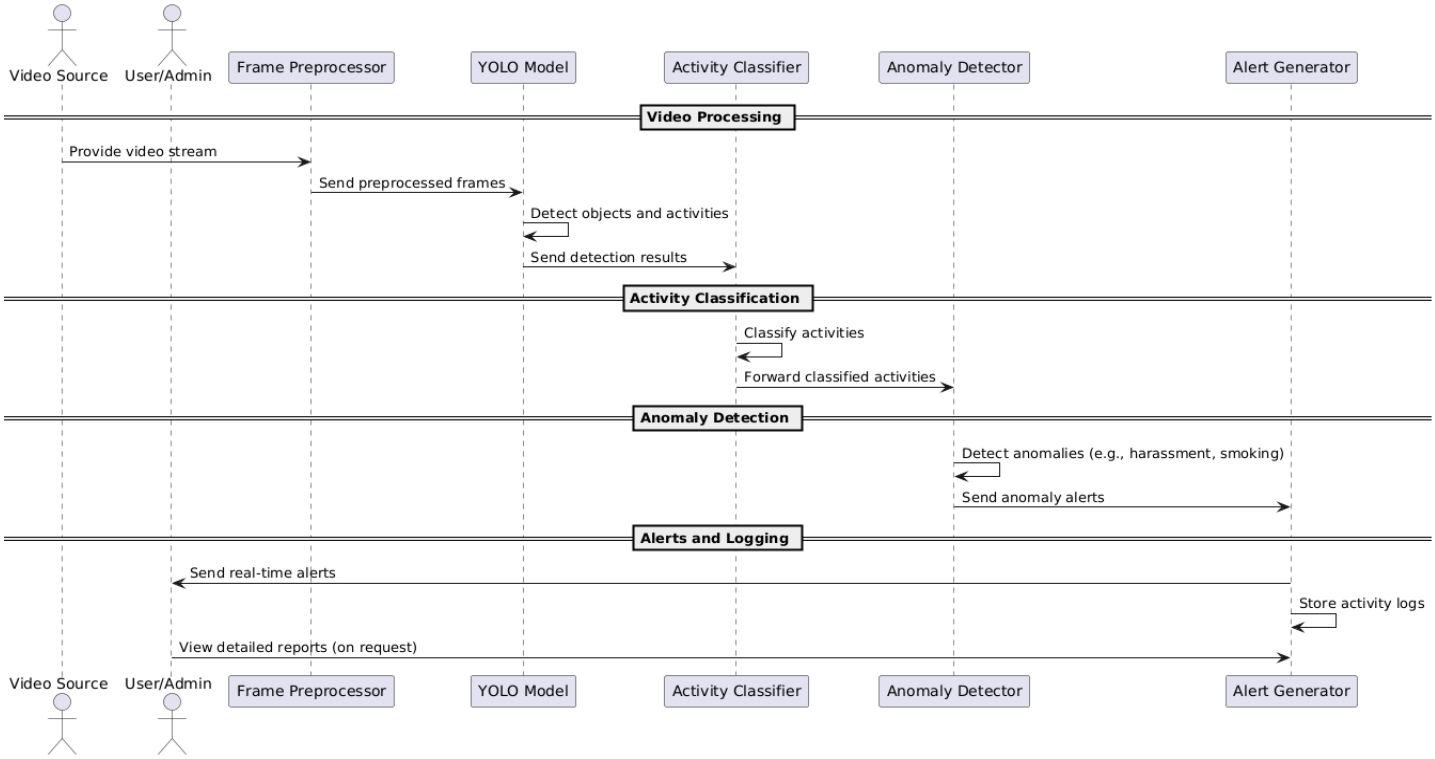
3. **YOLO Model:** A custom-trained YOLOv11 model (best5.pt) for detecting seven classes.
4. **Database:** SQLite to store prediction results and enable data analysis.
5. **Static Files:** Folder for storing uploaded images and processed results.

## 6.2. Detailed System Design

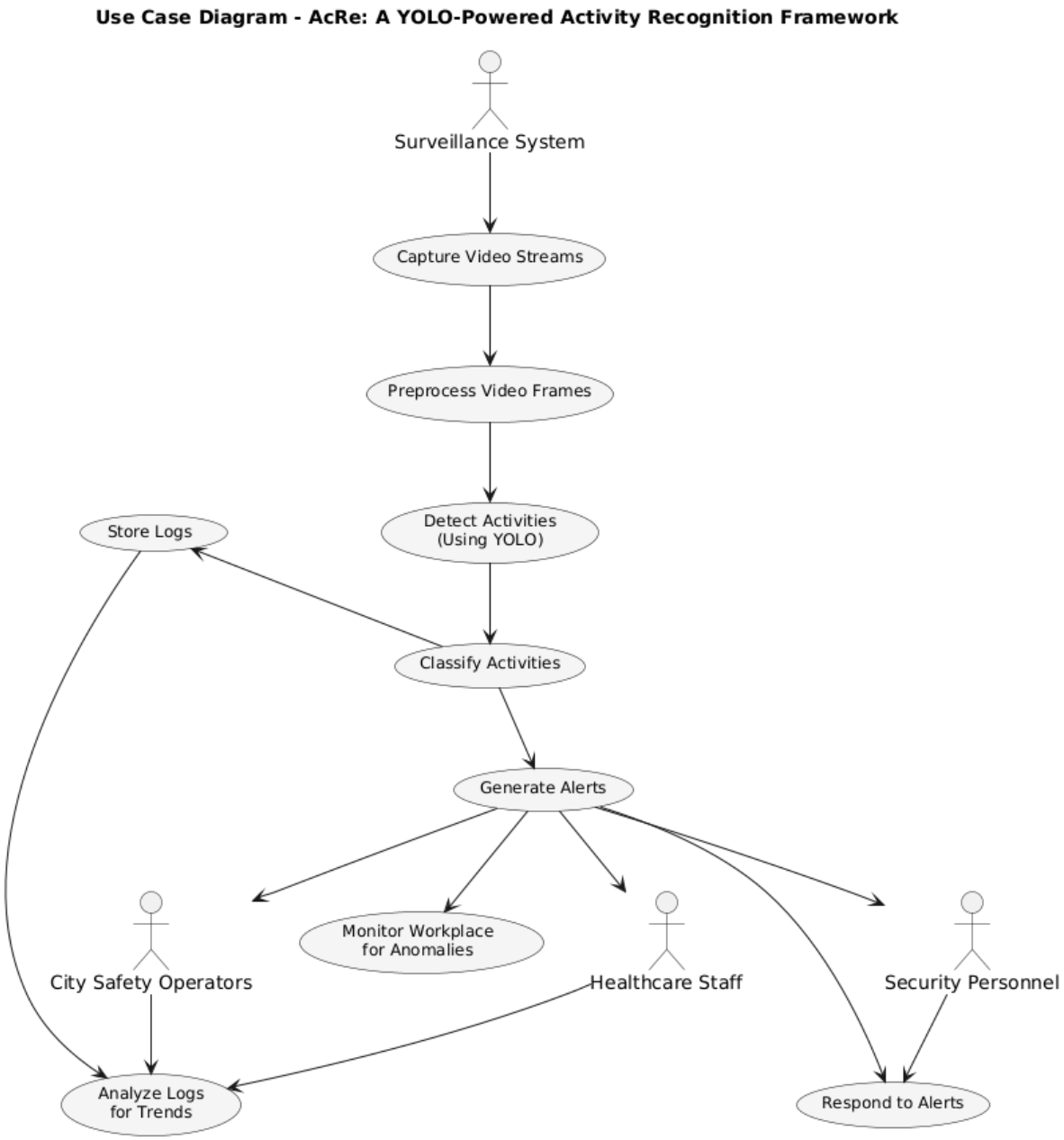
- **Classification:**
  The system consists of three main layers:
  1. **Presentation Layer:** Frontend (HTML, CSS, JavaScript).
  2. **Business Logic Layer:** Backend (FastAPI) and YOLO model processing.
  3. **Data Layer:** SQLite database.
- **Responsibilities:**
  1. **Frontend:** Collects user inputs and visualizes outputs.
  2. **Backend:** Manages API requests, processes images, and integrates with the YOLO model.
  3. **Database:** Stores prediction data for analysis.
- **Constraints:**
  1. Hardware dependency on GPUs for faster YOLO model processing.
  2. Limited to seven custom classes for detection.
- **Processing Flow:**
  1. User uploads an image via the web interface.
  2. Backend receives and preprocesses the image.
  3. YOLO model generates predictions.
  4. Results are stored in the database and returned to the frontend.

## 6.3 Key Diagrams

### 1. Sequence Diagram



## 2. Use Case Diagram:




# 7. Implementation and Testing

## 7.1. Software Development Methodology

- **Methodology Used:** Waterfall approach.
- **Reason:** Sequential development allowed focus on one phase at a time:
    1. Training YOLOv11 for high accuracy.

2. Backend development with Python 3.11 and FastAPI.
3. Frontend integration and testing.

## 7.2. Tools and Technologies

- **Programming Language:** Python 3.11.
- **Model:** YOLOv11 with 7 custom classes.
- **Backend Framework:** FastAPI.
- **Frontend:** Jinja2, Bootstrap, Chart.js.
- **Database:** SQLite with SQLAlchemy.
- **Additional Libraries:** OpenCV, NumPy, Pillow.

## 7.3. Key Capabilities

- Detects seven activities (Drinking, Eating, Violence, Sleeping, Smoking, Walking, Weapon).
- Real-time image predictions with bounding boxes.
- Stores prediction data in a database for analysis.

## 7.4. Testing Methodologies

- **Unit Testing:** Validate API endpoints and YOLO model predictions.
- **Integration Testing:** Ensure seamless interaction between backend, model, and frontend.
- **Performance Testing:** Measure API response time and resource utilization.
- **User Testing:** Gather feedback on usability and prediction reliability.

## 7.5. Evaluation and Analysis

- Model achieves high precision and recall for key classes.
- Average API response time: 1.3ms.
- Positive user feedback for the interface and prediction accuracy.

## 7.6. Comparison with Original Specifications

- All specified features implemented:
    - Real-time object detection.
    - Data storage and visualization.
- Deviations:
    - Video stream detection planned for future.

## 7.7. Libraries and Templates

- **YOLOv11:** Custom-trained model for activity detection.
- **FastAPI:** Backend API.
- **Jinja2, Bootstrap:** Frontend templates.
- **Database Libraries:** Sqlite3 for ORM.

### 7.8. Accuracy and Performance

- **YOLO Model Accuracy:** Precision: 90%, Recall: 88% (average across classes).
- **Backend Performance:** API response time: 1.3ms, max throughput: Y requests/second.
- **Database Efficiency:** Fast data retrieval and storage.

## 8. Results and Discussion

### 8.1 Results

- **Model Performance:** High accuracy for all seven classes.
- **System Performance:** API delivers predictions in real-time.
- **Screenshots:** Annotated outputs with bounding boxes and confidence scores.


### 8.2 Discussion

- **Strengths:** High detection accuracy, user-friendly interface, efficient API.

- **Limitations:** Dependent on high-quality datasets, resource-intensive for real-time processing.
- **Challenges:** Model tuning for overlapping object detection, efficient frontend-backend integration.

# 9. Conclusion

## 9.1 Future Testing

- Test with larger and more diverse datasets for improved reliability.
- Address limitations in resource optimization and occlusion handling.

## 9.2 Scalability and Interoperability

- Plan for integration with larger systems and higher traffic handling.

## 9.3 Integration of AI and Advanced Features

- Explore using generative AI for dataset expansion or preprocessing.

## 9.4 Mobile Application Development

- Extend functionality to mobile platforms for real-time detection.

## 9.5 AR and VR Integration

- Introduce AR/VR for immersive prediction visualization.

## 9.6 Personalized Features

- Expand the system for tailored use cases, such as monitoring specific activities in different environments.