ASFS - Auto Surveillance for Security

Type: Development Project

FYP Group (F22-028-D-ASFS)

Adeel Subhan   (19I-1869)

Fahad Baig   (19I-1883)

Fateh Muhammad (19I-1750)

Supervised by

Dr. Hasan Mujtaba

National University of Computer and Emerging Sciences, Islamabad, Pakistan

Anti-Plagiarism Declaration

This is to declare that this Final year project report produced under the following title:

Title: ASFS - Auto Surveillance for Security

This is our own contribution and no part of this FYP report has been reproduced on as it is basis, which can be considered as plagiarism. All reused parts have been referenced to argue the idea, have been cited properly, and are reused in order to avoid reinventing the wheel. The work presented in this report is our own and it has not been submitted or used by any other institution or organization previously. We will accept the consequences if violation of this declaration is determined.

Date: 22/10/2022

Group member 1:

Name: Fateh Muhammad (19I-1750)

Signature: _____

Group member 2:

Name: Adeel Subhan (19I-1869)

Signature: _____

Group member 3:

Name: Fahad Baig (19I-1883)

Signature: _____

Supervised by

Dr. Hasan Mujtaba

Professor, School of Computing,
Director, Office of Industrial Liaison and Digital Media,
Founding Member AIM Lab,
National University of Computer and Emerging Sciences,
Islamabad.

Signature: _____

Abstract:

Our web-based application is useful in detecting threat on publicly accessible places. This system will serve the purpose of auto surveillance using CCTV cameras. Threat could be a detected gun, shooting, or explosion at some public place. Most of the time, threats will be detected even before an offence. Shooting might be after loitering in front of some private property or explosion might be caused by some unattended luggage. Our system will detect such suspicious activities as well and report to authorities. Depending on the user needs and sensitivity of activities around the place, user will be able to customize this web-based application features. User can also define, who should be reported in case of any mishap or suspicious activity, this approach allows us to target a broader audience. Given time constraints and for demo purposes we have narrow down the scope of this system and we will only be concentrating on auto surveillance of private properties during our final year project. This report will give a brief overview of our approach to achieve this milestone during this development based FYP (final year project).

Executive Summary:

This FYP report is an overview of our final year project which we started working on from August 2022. This report covers brief overview of this development project.

We are creating a web-application that can provide services of auto surveillance using CCTV cameras. This web application will be reporting authorities on any suspicious activities in front of private properties or potential threats. Due to time constraints, we are just considering two classes for suspicious activity detection, loitering and unattended luggage. This application will also be detecting threats like shooting, weapon detection, or explosion.

This application is customizable and user can change it according to once own needs, and depending on the sensitivity of area where the system will be installed. This system will be storing any detected threats or suspicious activities into a database so that user can review them later.

Once, this system is installed and customized according to sensitivity of area and needs of user, User would be asked to define authorities in case of any mishap. This system will automatically be reporting those authorities in case of any threat detection like shooting or explosion. One more feature is to check if the person causing suspicious activity warning is the admin user and owner of system or not. We'll using deep learning approach for face detection if that would match to our database of recognized people like security guards, or family members, we'll not report authorities and just send a notification to owner using progressive web apps.

Table of Contents

1. Introduction:

Our Project is about threat detection before it happens and informing user about that suspicious activity that can cause threat. Furthermore, we want to detect and report any potential threats like weapon detection or shooting and report authorities. So, we'll be looking into both threat detection like detecting some weapon in publicly accessible places or any suspicious activities like unattended luggage left.

The project is about having a customizable application that can detect suspicious activity and threats by auto surveillance using CCTV cameras. We want to make this application customizable for all kind of users. This auto surveillance system will detect following threats and suspicious activities and warn users about it before any mishap. This system would be detecting such activities on common areas or public areas that could in front of some private property like houses as well. We are considering following activities as suspicious, loitering in front of house and unattended luggage. We'll be considering detection of weapons or fighting as a threat.

Furthermore, there could be a fire at pubic place followed by any suspicious activity like someone left unattended luggage and then it causes a blast. We'll try to warn authorities about this luggage before any incident which is our strategy about detecting threats and trying to avoid them before they happen at all. User of application will be able to decide and customize who is the authority to be warned in case of any suspicious activity or threat detection.

2. Project Vision:

Auto Surveillance of private properties like houses using CCTV cameras to detect and report threats or any suspicious activities that can cause damage to authorities.

2.1 Problem Statement:

Manual surveillance using CCTV cameras need to utilize much more resources like hiring a guard, who is supposed to monitor camera footage than auto surveillance. Even after having multiple CCTV cameras and hiring guards, criminals are able to snatch bags from public or shoot people sometimes due to negligence of guards. On the other hand, there could be potential threats like shooting, and explosions which can be detected before incident happens by looking at the suspicious activities before such incidents. We want to automate surveillance of publicly accessible places for example in front of any private property like house using CCTV cameras. It'll detect suspicious activities like unattended luggage, or loitering and report to user about such activities. This system will also report authorities like police about any threats or incidents like shooting, Fire, fighting or explosions. Which will significantly decrease the crime rate in our country and make it more peaceful place.

2.2　Business Opportunity

"According to the research report published by Polaris Market Research, the global physical security market size is expected to reach USD 137.99 Billion by 2028. At a CAGR of 6.1% during the forecast period." [(1)](1)

Providing security solutions in countries like Pakistan will not only be helpful for the security of the country, but it'll create new doors for export services that can solve economic problems of this country as well. There has been one recent opening of company in Pakistan whom are providing auto surveillance solutions using Computer Vision. Pakistan has potential for this industry. If we can provide our product as services to people, we can generate good revenue from that along with exporting these software services to other countries as well globally. Our goal is to leverage this opportunity, and we'll try to fill this market gap by providing these services in Pakistan.

2.3　Objectives:

Here are some objectives that we'll be working on which we think are crucial for the safety of people. These objectives are as the following:

1. **Threat Detection**: Using visual threats and object detection techniques, we aim to detect any apparent threat like a person holding a gun, a fire nearby or individuals fighting.
2. **Suspicious Activity Detection**: We'll be detecting any suspicious activity that can result into incidents like loitering can result into shooting or explosion by any unattended luggage.
3. **Notifying User and Authorities**: Once threat or suspicious activity has been detected, our web-based application needs to notify authorities, so that appropriate action can be taken timely. We want these notifications to be customizable according to preference of user i.e., user will define what authority to call when a specific threat/suspicious activity is detected.
4. **Simple and Customizable Web Application:** This is one of the most important objectives for us, as we want to target a large audience with security needs. Building a simple application means that even a layman will be able to use our application without any issue. Customizability will help us target a larger audience, and meet unique needs of customers depending on the sensitivity of location where system is being installed.

2.4   Project Scope**:**

Our long-term goal is to build a surveillance-based security applications which can be used by individuals, organizations or any institutions for the security of publicly accessible areas. Given the time constraints for our FYP and demo purposes, we have narrow down the scope to private properties only like houses. We'll be detecting any suspicious activity or threat by auto surveillance using CCTV cameras in front of houses or private properties and will notify authorities about such activities.

2.5   Constraints

We have some limitations for this project. Starting with time constraints of our Final Year Project, we are considering limited scope for now that is, we'll be giving demo and testing only for private properties yet. Some other limitations could be limited hardware resources. We do not have quality data available for the training of deep learning models. And the data which is available on internet doesn't have good video quality. Angle of camera may also make difference while assessing the accuracy of models. We can face other issues as well, like picture of a person may not be clear enough in CCTV videos, or storing streams of data for previous mentioned time period. Due, to low quality we'll be getting low pixels which give small number of features to understand. The quality of video from camera may be good at day but might not be good at night. Another factor can be the angle of CCTV camera, because that'll also make difference when it comes to getting quality data.

1. Angle of Camera
2. Defined Scope is just for private properties like Houses
3. Lighting can affect quality of video
4. CCTV Camera quality or in general low quality video form Cameras
5. In crowd, if too many people in the frame model may not perform well
6. Limited hardware and Software Resource

2.6   Stakeholders Description:

2.6.1 Stakeholder's Summary:

Stakeholder	Responsibilities	Description
User	Proper Installation of hardware, to make sure that uninterrupted video being sent to the server.	We are sending CCTV video to server to detect if there is any suspicious activity or threat in it or not.
Authorities	Authorities to be reported in case of any threat detection should respond as fast as possible to stop the crime or catch the criminal.	After a threat have been detected by the system, authorities will be notified about this. The institution or notified organization should respond quickly to stop the crime before any mishap if possible else they should make sure to arrest criminal.

2.6.2 Stakeholder's Goals and Problems:

Stakeholder	Problem	Goal
User	When it comes to Machine Learning or Deep Learning models, we say "garbage in, garbage out". User will have to make sure the quality of video coming from CCTV camera is good. Because low quality video will give us low pixels that means there will be small number so features to detect threat or suspicious activity.	The user should behave responsibly after being notified about any suspicious activity like loitering, or unattended luggage. He can share details to authorities as well like police.
Authorities	Authorities need to react as soon as possible. We are expecting from our system to notify authorities like police as soon as possible, but they should react within time to catch the suspicious person or criminal with proof.	Using our system, authorities can trace criminals very easily with solid proof of their committed crime and will be able to decrease crime rate in the society. One thing on the end of Authorities is to respond quickly to make sure criminals do not get away with crimes.

3. Software Requirement Specifications:

3.1   List of Features:

3.1.1 Threat Detection:
Our web-based application will be able to detect threats on in real time. We will be detecting threats of following categories, fighting, shooting, and explosions.

3.1.2 Notify Authorities:
If any threat is detected, the system is supposed to report authorities on the incident, User of the system is supposed to define the authorities in case of any mishap.

3.1.3 Suspicious Activity Detection:
Our system will also be detecting the suspicious activities that can cause any damage or can become threat. For example, there could be shooting incident after loitering. Our system should report loitering to User so they can avoid threat. System will be detecting suspicious activities on run time.

3.1.4 Log of any suspicious activity or threat:
This system will keep record of previous suspicious activities or threats detected by the system. User will be able to view log of threat or suspicious activities using web application.

3.2   Functional Requirements:
- User shall be able view video stream from CCTV camera on web application
- User shall be able to see detected suspicious activities.
- User shall be able to see all threats detected by the system.
- User shall be able to delete threat records form history
- System shall be able to inform authorities urgently
- User shall be able to define, who have the authority in case of specific incident
- System shall maintain threat log and suspicious activity log

3.3   Quality Attributes:

Attributes	Brief Details
Usability	User should be able to see CCTV cameras at dashboard. And should be able to add CCTV cameras. User should be able to view the video stream of CCTV cameras on Dashboard.
Reliability	When system detects any suspicious activity or threats, results should be correct and reliable. So that user should be able call authorities for help.
Maintainable, Scalable, and Customizable	The system is supposed to be used by a diverse audience, so we'll have to customize it according to needs of customers. Hence, we should be able to add or remove functionality easily without major changes. That means system should be scalable, maintainable, and Customizable.

3.4   Non-Functional Requirements:

➢ System shall be able to detect any suspicious activity that can cause threats like shooting or explosion. So that authorities can react before criminal make its move.

➢ If crime have already taken place, system should inform authorities, so that they can respond quickly and catch the criminal.

➢ The web application of our system should provide an easy-to-use interface for users.

➢ System should be able to support multiple cameras.

4. High Level Use Cases:

Use Case (UC01)	Setup User Profile
Actor:	User
Type:	Primary
Description:	The user of our system should be able to register on our website by providing required information like no of cameras, their address, and User should define authorities in their area to be reported in case of any incident, along with other credentials.

Use Case (UC02)	View Video Stream
Actor:	User
Type:	Primary
Description:	The user of our system should be able to view video stream from CCTV cameras on Dashboard of website after logging into system successfully. Any threat, or suspicious activity detected also would be reflected as well.

Use Case (UC03)	View Threat Log
Actor:	User
Type:	Primary
Description:	The user of our system should be able to view all threats detected previously. We'll be maintaining threat log to save those video clips separately that had any potential threat.

Use Case (UC04)	View Suspicious Activity Log
Actor:	User
Type:	Primary
Description:	The user of our system should be able to view all suspicious activities detected previously. We'll be maintaining log of suspicious activities to save those video clips separately.

| Use Case (UC05)	Redefining Authorities
Actor:	User
Type:	Primary
Description:	The user of our system should be able to redefine authorities to be reported after logging into the system. For example, User wants our system to report to their private security instead of police in case of any incident.

Use Case (UC06)	Check threat or suspicious activity by Uploading a Video
Actor:	User
Type:	Primary
Description:	The user of our system should be able to upload a video to check any kind of suspicious activity or threat in the uploaded video.

Use Case (UC07)	Delete Suspicious Activity Log
Actor:	User
Type:	Primary
Description:	The user of our system should be able to delete all suspicious activities detected previously. We'll be maintaining log of suspicious activities to save those video clips separately. So that user can delete them. (either all of them or specific ones)

Use Case (UC08)	Delete Threat Log
Actor:	User
Type:	Primary
Description:	The user of our system should be able to delete all threats detected previously or specific ones from the list. We'll be maintaining log of threats to save those video clips separately. So that user should be able to delete log using web application.

Use Case (UC09)	Report Authorities
Actor:	Alert Generator
Type:	Primary
Description:	The Alert Generator of our system should be able to report authorities as quickly as possible about any threat detected or suspicious activity detected in front of CCTV cameras.

4.1   Use Case Diagram:
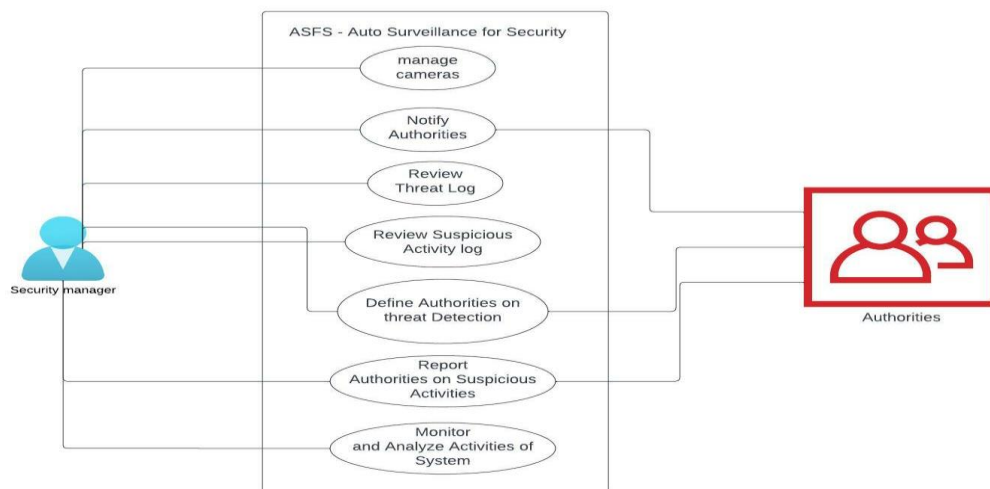
5. Iteration Plan

We'll completing our Final Year Project according to following plan. We have divided work load on into iterations. First two iterations are for Final Year Project 1. Last two iterations will be completed in next semester.

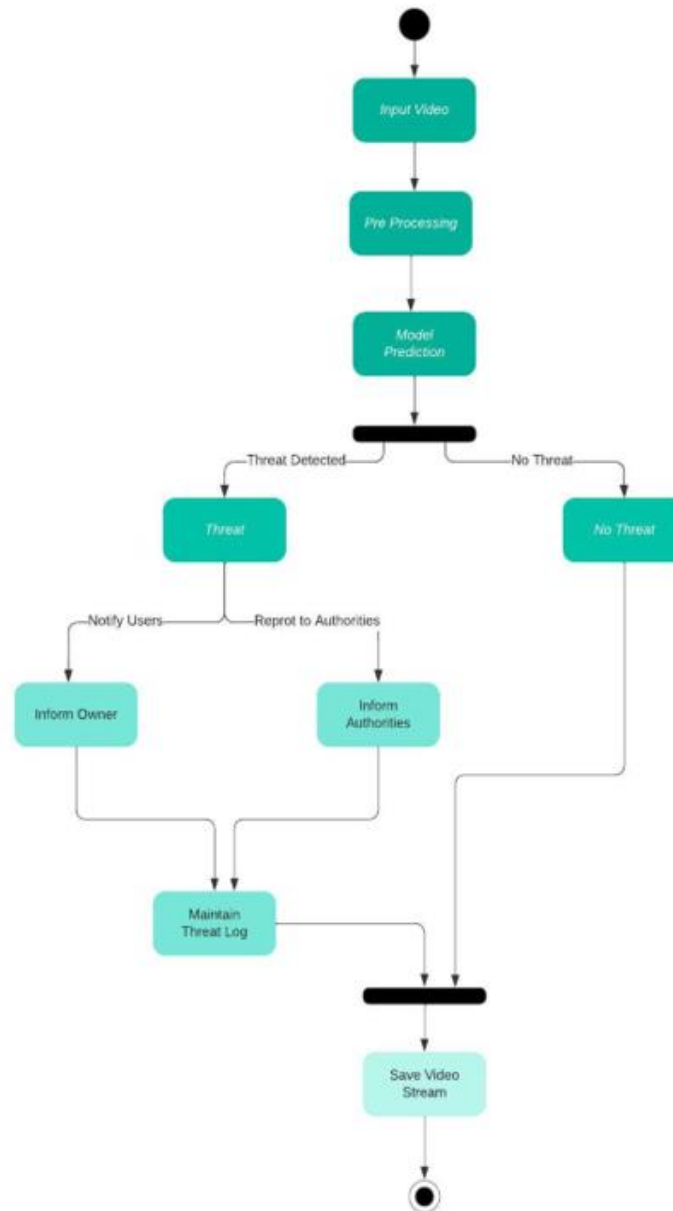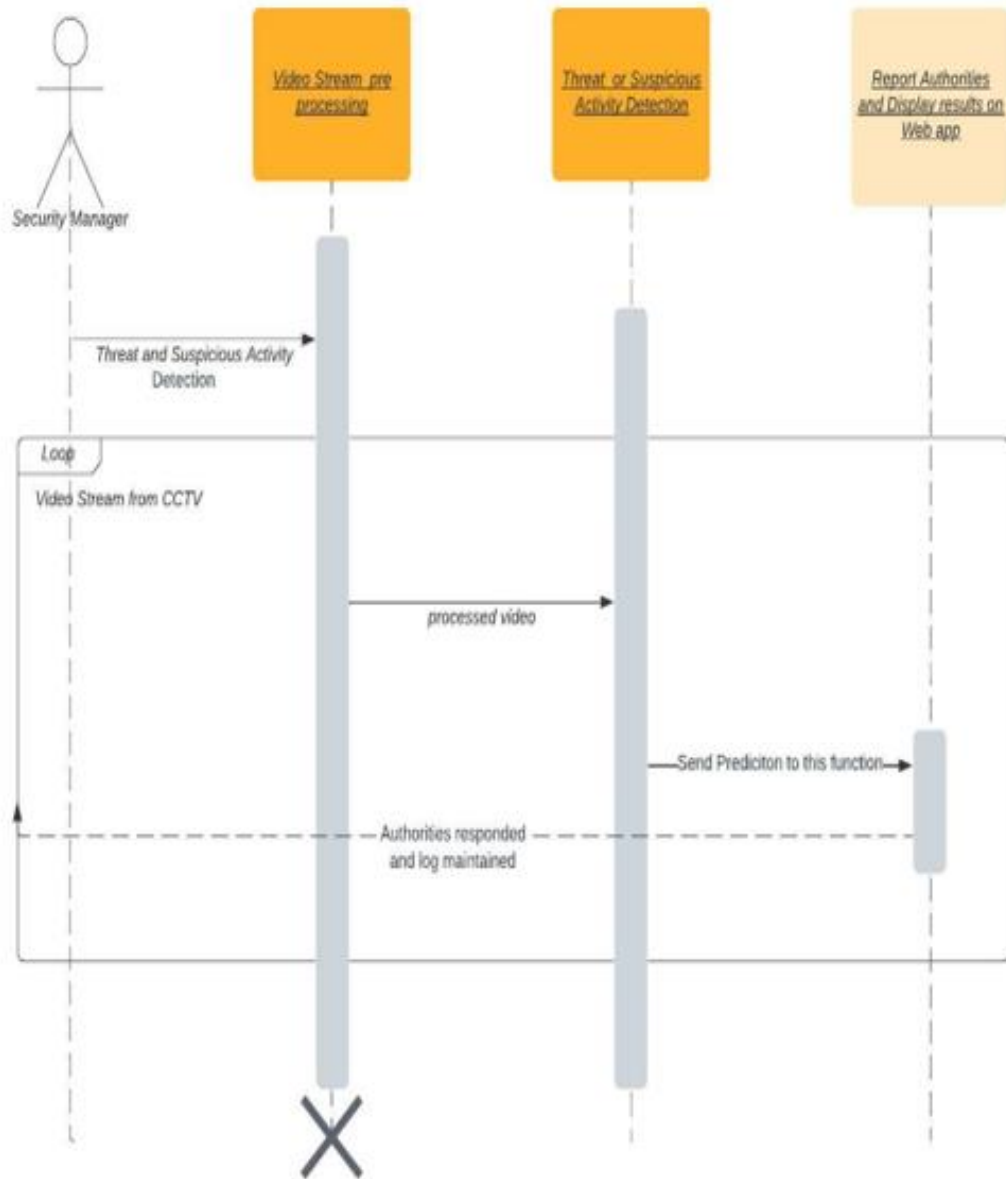| Time Planned	Description of Task to be Completed	Evaluations
Sep-Oct	• Suspicious Activity Detection (Loitering Detection) • Static Web Page for presenting results • Database and System Design (Complete Design phase)	FYP1 (Mid-Evaluation)
Nov-Dec	• Suspicious Activity Detection (Unattended luggage Detection) • Dynamic Web Page, deployment with models deep learning models • Extended Features added into Web application	FYP1 (Final-Evaluation)
Jan-Feb	• Threat Detection (Fire and Arson) • Demo Functionality for new Users will be added into web application • Alert Generation for reporting authorities and sending notification to user of system.	FYP2 (Mid-Evaluation)
Mar-April	• Threat Detection (Shooting) • Web application will be made customizable to scalable • Testing & Deployment	FYP2 (Final-Evaluation)

6. Iteration #01:

6.1   Activity Diagram:

Find Threat
Fateh Muhammad | October 23, 2022

Input Video

Pre Processing

Model Prediction

Threat Detected — No Threat

Threat

No Threat

Notify Users — Reprot to Authorities

Inform Owner

Inform Authorities

Maintain Threat Log

Save Video Stream

Find Threat

Fateh Muhammad | October 23, 2022

6.2   Sequence Diagram:

6.3   System Sequence Diagram:

Sequence diagram

I191869 Adeel Subhan | October 23, 2022
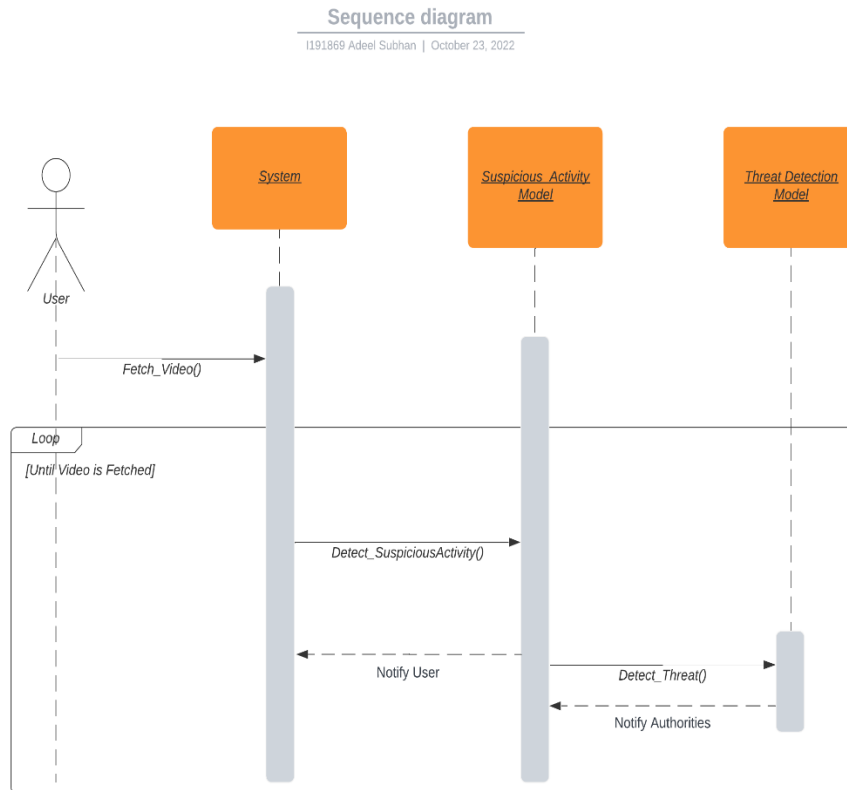
7. Implementation Details

7.1   Object Detection

Video from CCTV or Demo Page is fetched by the system and preprocessing is done. Preprocessing involves motion compensation and resizing of each frame to 448x448x3. Once Pre-Processing is done each frame goes to Object Detector, YOLO V5 (State of the art object detector), which detects each person in the frame and returns proposed bounding box coordinates per frame for each person. Object detection model is trained on COCO dataset with the following architecture.
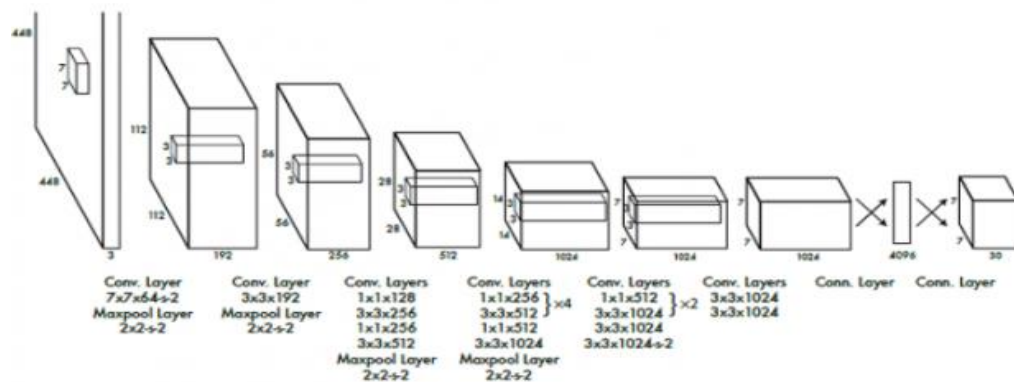

7.2   Deep SORT Tracking

Once we have bounding box coordinates, we then pass the coordinates to Deep SORT tracking algorithm. The algorithm returns a tracking ID associated with each bounding box coordinate that is passed through it. As a result, we're able to assign unique id to each person.

Deep SORT is a tracking algorithm which not only tracks object based on velocity and motion but also appearance of the object. Deep SORT using the following techniques:

1. **Kalman Filter:** are used to optimally estimate the variables of interests when they can't be measured directly, but an indirect measurement is available.
2. **Optical Flow:** is a vector field between two images, showing how the pixels of an object in the first image can be moved to form the same object in the second image
3. **Mean Shift:** is a data clustering algorithm commonly used in computer vision and image processing.
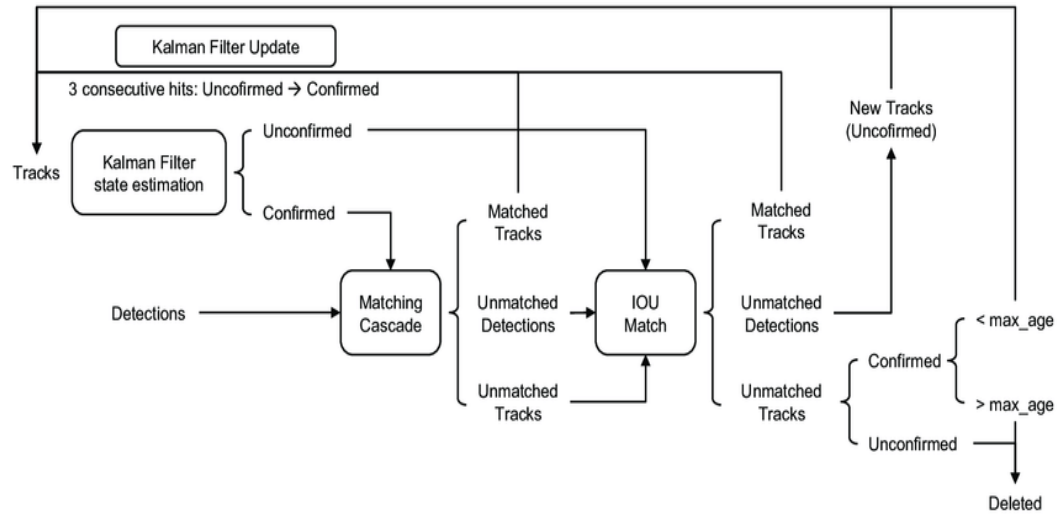
7.3   Suspicious Activity Classification

Having tracking ID for each object allows us to track activity of each object. We maintain a state dictionary to keep track of each person that enters the frame. State dictionary contains bounding box coordinates (x,y,w,h), first occurrence of the person in frame and last occurrence of the person in frame. This allows us to compute the time each person was in frame, if a person just walks by, their state will be deleted, else if person stays in the frames more than the time specified by the user, it'll be classified as Loitering (Suspicious Activity)

8. Results

8.1   Evaluation

This detector was trained with YOLO loss function with SoftMax function to convert scores into probabilities that sum up to one. This means smaller bounding boxes are penalized more as we need better prediction on smaller objects than on bigger objects. The smaller values are punished more if we follow "square-root" method. This SoftMax loss helped us to normalize probability as the activation function in the output layer of neural network models that predict a multinomial probability distribution. As you may notice, the accuracy we were able to achieve was approximately 80% on test dataset (we aim to improve this in upcoming iterations).
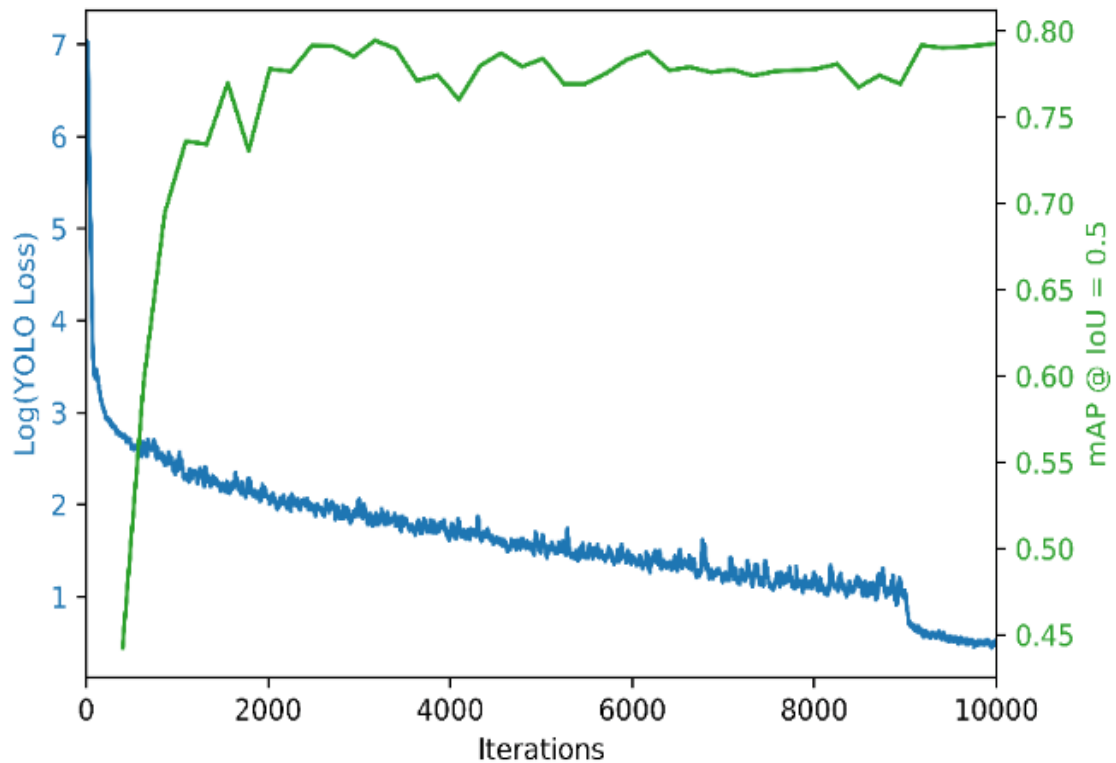
8.2   Output

This detector makes a box around the person who is suspected for Loitering. This is done using the state dictionary which was generated from DeepSORT results. Using results and OpenCV, we generate output that'll identify the user of Loitering. Later, this will be sent as notification to user's devices.

References:

1. By Polaris Market Research published, October 20, 2022

   https://www.digitaljournal.com/pr/physical-security-market-expected-to-reach-us-137-99-billion-and-cagr-6-1-by-2028-polaris-market-research#:~:text=WORLD-,Physical%20Security%20Market%20Expected%20to%20Reach%20US%20~%24137.99%20Billion,October%202020%2C%202022

2. T. Subetha and S. Chitrakala, "A survey on human activity recognition from videos," 2016 International Conference on Information Communication and Embedded Systems (ICICES), 2016, pp. 1-7, doi: 10.1109/ICICES.2016.7518920.

3. A. W. M. Smeulders, D. M. Chu, R. Cucchiara, S. Calderara, A. Dehghan and M. Shah, "Visual Tracking: An Experimental Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 7, pp. 1442-1468, July 2014, doi: 10.1109/TPAMI.2013.230.

4. arXiv:1801.04264 [cs.CV], https://doi.org/10.48550/arXiv.1801.04264