

HUMAN BEHAVIOUR AND ABNORMALITY DETECTION USING SPATIO TEMPORAL AND YOLO V5 MODEL

A PROJECT REPORT

Submitted by

AKSHAYA.R [211419104009]

HELEN ROSHNA.A [211419104099]

NEONA JOSITA.W [211419104182]

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING



PANIMALAR ENGINEERING COLLEGE

(An Autonomous Institution, Affiliated to Anna University, Chennai)

APRIL 2023

PANIMALAR ENGINEERING COLLEGE

(An Autonomous Institution, Affiliated to Anna University, Chennai)

BONAFIDE CERTIFICATE

Certified that this project report “**Human Behaviour and Abnormality Detection using Spatio Temporal and Yolo V5**” is the bonafide work of “**Akshaya.R (211419104009), Helen Roshna.A (211419104099), Neona Josita.W (211419104182)**” who carried out the project work under my supervision.

SIGNATURE

**DrL.JABASHEELA,M.E.,Ph.D.,
HEAD OF THE DEPARTMENT**

DEPARTMENT OF CSE,
PANIMALAR ENGINEERING COLLEGE,
NASARATHPETTAI,
POONAMALLEE,
CHENNAI-600 123.

SIGNATURE

**K.SANGEETHA,M.E.,
ASSISTANT PROFESSOR**

DEPARTMENT OF CSE,
PANIMALAR ENGINEERING COLLEGE,
NASARATHPETTAI,
POONAMALLEE,
CHENNAI-600 123.

Certified that the above candidate(s) was/ were examined in the End Semester

Project Viva-Voce Examination held on.....

INTERNAL EXAMINER

EXTERNAL EXAMINER

DECLARATION

We, Akshaya R (211419104009), Helen Roshna A (211419104099), Neona Josita W (211419104182) hereby declare that this project report titled **“Human Behaviour and Abnormality Detection using Spatio Temporal and Yolo V5 Model”**, under the guidance of **Mrs.K.Sangeetha,M.E.**,is the original work done by us and we have not plagiarized or submitted to any other degree in any university by us.

AKSHAYA R

HELEN ROSHNA A

NEONA JOSITA W

ACKNOWLEDGEMENT

We would like to express our deep gratitude to our respected Secretary and Correspondent **Dr.P.CHINNADURAI, M.A., Ph.D.** for his kind words and enthusiastic motivation, which inspired us a lot in completing this project.

We express our sincere thanks to our Directors **Tmt.C.VIJAYARAJESWARI, Dr.C.SAKTHI KUMAR,M.E.,Ph.D** and **Dr.SARANYASREE SAKTHI KUMAR B.E.,M.B.A.,Ph.D.,** for providing us with the necessary facilities to undertake this project.

We also express our gratitude to our Principal **Dr.K.Mani,M.E.,Ph.D.** who facilitated us in completing the project.

We thank the Head of the CSE Department, **Dr.L.JABASHEELA,M.E.,Ph.D.,**for the support extended throughout the project.

We would like to thank my **Mrs.K.SANGEETHA,M.E.,** and all the faculty members of the Department of CSE for their advice and encouragement for the successful completion of the project.

AKSHAYA R

HELEN ROSHNA A

NEONA JOSITA W

ABSTRACT

The growing urban population has brought economic prosperity while increasing the likelihood of various abnormal events. Video data captured by surveillance cameras is an important means for law enforcement officers to judge the incident. However, a large number of surveillance cameras record massive amounts of data at all times, increasing the workload of law enforcement officers and providing opportunities for criminals. This project proposes a Spatio Temporal model and a novel regularity score based on the results of the YOLO network to detect abnormal behavior in crowd scenarios. The Spatio Temporal model extracts spatial features of video frames. In the training process, a weighted loss function is proposed based on the YOLO detection results, which emphasizes the foreground part and thus overcomes the impact of complex background. In addition, a novel regularity score is put forward in the anomaly detection process. The existing system only detects violence detection which is already available in real-time CCTV abnormal detection systems. The proposed model employs Spatio Temporal and YOLO architecture. Most recent accurate models require multiple GPUs to train with a large mini-batch size, and doing so with only one GPU is extremely sluggish and inefficient. This issue is addressed in YOLO Model by creating an object detector that can be trained on a single GPU with a reduced mini-batch size. The proposed system allows a single 1080 Ti or 2080 Ti GPU to train a speedy and accurate object detector. The proposed system supports human abnormality detection with both image and video inputs. The experimental results on UCSD ped1 and ped2 dataset verify that the proposed method achieves better performance than the most of semi-supervised methods.

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	ABSTRACT	v
	LIST OF TABLES	viii
	LIST OF FIGURES	viii
	LIST OF ABBREVIATIONS	ix
1.	INTRODUCTION	1
	1.1 Overview	2
	1.2 Problem Definition	3
2.	LITERATURE SURVEY	5
3.	SYSTEM ANALYSIS	15
	3.1 Existing System	16
	3.2 Proposed system	16
		18
4.	REQUIREMENT ANALYSIS	
	4.1 Hardware Requirements	19
	4.2 Software Requirements	19
5.	SYSTEM DESIGN	21
	5.1 Data Flow Diagram	22
	5.2 UML Diagrams	23
6.	SYSTEM ARCHITECTURE	28

6.1	Module Design Specification	30
6.2	Algorithms	33
7.	PERFORMANCE ANALYSIS	36
7.1	Results and Discussion	37
7.2	Accuracy Prediction	37
8.	CONCLUSION	40
8.1	Conclusion and Future Enhancements	41
	APPENDICES	42
A.1	Sample Screens.	42
A2.	Coding	46
A3.	Dataset	54
	REFERENCES	55

LIST OF TABLES

Table No	Table Description	Page No
6.1	Dataset	30

LIST OF FIGURES

Fig No	Figure Description	Page No
5.1	Dataflow Diagram	22
5.2	Use Case Diagram	23
5.3	Class Diagram	24
5.4	Sequence Diagram	25
5.5	Activity Diagram	27
6.1	System Architecture	29
7.1	Testing accuracy of the model	38
7.2	Comparison with previous work	38
A.1	Yolo Testing	42
A.2	Detection of weapon through web camera	42
A.3	Detection of weapon in a recorded video	43
A.4	Output	43
A.5	Location of the camera	44
A.6	Email received after detecting a weapon in a live video	44
A.7	Email received after detecting weapon in a recorded video	45
A.8	SMS notifying the detection	45

LIST OF ABBREVIATIONS

S.NO	ABBREVIATIONS	EXPANSION
1.	YOLO	You Only Look Once
2.	CNN	Convolutional Neural Network
3.	RNN	Recurrent Neural Network
4.	CRNN	Convolutional Recurrent Neural Network
5.	LSTM	Long Short-Term Memory
6.	FLOPS	Floating Point Operations Per Second
7.	FPS	Frames Per Second
8.	HAD	Heat Actuated Device
9.	SOA	Service Oriented Architecture
10.	UTI DATASET	Uniform Type Identifier Dataset
11.	CRF	Conditional Random Fields
12.	CCTV	Closed-Circuit Television
13.	RGB	Red, Blue and Green
14.	DFD	Data Flow Diagram
15.	ISLD	Inter-Stimulus Level Difference

CHAPTER – 1

INTRODUCTION

CHAPTER 1

INTRODUCTION

OVERVIEW

The human behaviour and abnormality detection helps reduce life threatening acts and provide security. To detect abnormal behaviour in crowd scenarios. In this we proposed a spatio temporal model and a novel regularity score based on the results of the yolo network model. It is intended to create a system to detect abnormalities in human behaviour using the spatio temporal and yolo models. In addition, we plan to develop a smart surveillance security system that can detect weapons, specifically guns. To accomplish this, we used a combination of compute vision methods and deep learning to identify a weapon from a captured image. Recent advances in machine learning and deep learning have shown significant progress in the areas of object detection and recognition, exclusively in images. Object detection and classification are critical first steps in any video surveillance application and are required for subsequent object tracking tasks. We trained the classifier model of YOLO v5, i.e., "You Only Look Once," for this purpose. This model is a cutting-edge real-time object detection classifier. YOLO version 5 is an object detector that does not require multiple GPUs for training and has a smaller minimum batch size than previous versions. Spatio Temporal is a model that extracts spatial features from video frames. Based on the yolo detection results, a weighted loss function is proposed during the training process, which emphasises the foreground part and thus overcomes the impact of complex background. Furthermore, in the anomaly detection process, a novel regularity score is proposed. Using audio and video inputs, the proposed system is capable of identifying human abnormalities. Using the Yolov5 model and Convolution layers, we will build a Spatio Temporal Model. A pretrained model has seen a tonne of objects and is familiar with the classification rules that apply to each one. By putting the network through training on the COCO

and Imagenet datasets, the weights in the earlier pretrained model were obtained. Consequently, it is limited to detecting objects from the classes represented in the training dataset. For finding anomalies in videos, we'll build a prediction function. Making a function that can assess inputs like images, videos, and times to determine whether or not they are anomalous is the next step. Based on the trained model, we can predict the labels of the data values using the python predict method. In addition to identifying the presence of anomalous behaviour, we are also locating the incident and storing the data for later use.

PROBLEM DEFINITION

Human abnormal detection in video is critical to ensuring security in both internal and external spaces. The costs of gun violence are high in terms of the economy, the public's health, and psychological costs. Violence involving guns claims many lives each year. Children who experience high levels of violence in their neighbourhoods or through the media frequently suffer from psychological trauma. These crimes can be decreased by spotting disruptive behaviour early on and closely observing any suspicious activity so that law enforcement agencies can act right away. Human abnormality detection model give a machine the ability to recognize unsafe weapons and can also notify a human administrator when a gun or firearm is clearly visible in the vicinity. The identification of anomalous events is one of the most difficult tasks in a video surveillance system. A software programme that continuously watches for signs of offensive or disruptive behaviour is known as an anomaly recognition system. Practically speaking, video anomalies are scene- and context-dependent, so an activity that is abnormal in one scene might be normal in another. One of the main achievements of anomaly detection is the extraction of meaningful features, which can distinguish between nominal and anomalous events. To help with the quicker assessment of whether the unusual activity is abnormal or suspicious, the method also determines which frames and which portions of them

contain the abnormal event. This is also to design and develop a system that can detect guns, rifles, and fire in record time while using minimal computational resources. The most important and critical aspect of any application is having a desired and suitable dataset to train machine learning models on. As a result, we manually collected a massive amount of images from Google. The extraction of logical information from input video data during abnormal behaviour detection can be seen as a high-level image processing operation. To track the movement of moving objects in the videos, object detection algorithms are used. When there is a threat, an alert message goes off, occasionally alerting people to suspicious behaviour. Because of technological advancements, the majority of human-assisted applications are now automated and computer-based. In future, in order to detect any weapon or dangerous objects, our proposed system can also be integrated into surveillance and security robots.

CHAPTER – 2

LITERATURE SURVEY

CHAPTER 2

LITERATURE SURVEY

Abnormal Behaviour Recognition System based on improved CRNN

Model.

AUTHOR - Zheng Xu, Jingxuan Wang and Yuanyao Lu

YEAR – 2022

Abnormal behavior recognition technology based on surveillance video has been an important direction of current research, with high research value and application demand. Due to the complexity and variability of crowd movement and external environment, the recognition of abnormal behavior is quite challenging, and there is still a need for further research on the recognition technology of human abnormal behavior in surveillance video. In this paper, we proposed an adaptive video frame extraction method, and after a series of experiments and comparisons with former methods, our model achieved 86.78% accuracy on the UCF-Crime dataset and 95.2% accuracy in the subsequent system development of the algorithm model for segment detection and localization of surveillance videos. And the maximum speed of recognition can reach 543.7FPS. Data shows that the algorithm and system we designed can achieve better performance on Abnormal behavior recognition.

MERITS -This method can significantly reduce the number of parameters and FLOPs of the CRNN network without reducing accuracy.

DEMERITS-The accuracy is not mentioned properly.

Analysis of Video-based Human Activity And Detection Approaches

AUTHOR - I. Gull, A. Selwal and A. Sabha

YEAR –2022

Human action detection and recognition has become a significant topic in computer vision research over the last two decades. Intelligent techniques (such as machine learning and deep learning) have grown in popularity as a result of technological

advancements in visual data. Due to the exponential growth of video data generated by surveillance cameras, intelligent systems are in great demand for detecting specific human activities. In this work, we critically examine the state-of-the-art (SOA) methodologies for human activity detection (HAD) approaches. The study illustrates a generic classification for HAD methods and a paradigm shift is observed from traditional to modern deep learning-based techniques. We also provide an analysis of publicly available datasets for the classification of human activities. The paper outlines various open research issues as well as future directions for HAD methods. Due to problems such as a dynamic and complex background, camera motion, occlusion, and bad weather, it is evident that detecting human behaviors in surveillance videos is a challenging task. Convolution neural networks (CNNs) are used in the bulk of HAD techniques, prompting researchers to look at sequence learning models like Recurrent Neural Networks (RNNs) and Long-Short-Term Neural Networks (LSTM). Furthermore, our analysis indicates that only a few research articles on the detection of anomalous behavior have been published, with the majority of the work focusing on human action detection. Furthermore, existing HAD deep learning models could be improved by including the notion of transfer learning, which saves training time and enhances accuracy.

MERITS - Due to problems such as camera background, motion etc, CNN are used in bulk of HAD techniques to solve these problems.

DEMERITS-It needs higher software techniques

Human Behavior and Anomaly Detection using Machine Learning and Wearable Sensors.

AUTHOR - Ioana Alexandra Bozdog, Todea Daniel-Nicusor, Marcel Antal, Claudia Antal, Tudor Cioara, Ionut Anghel, Ioan Salomie

YEAR – 2021

This paper addresses the problem of detecting and analyzing human behavior using a set of non-privacy invasive wearable sensors aiming to identify potential anomalies. This may be an important tool for increasing the independence and delaying the institutionalization of older adults allowing them to live alone in their homes with little support from caregivers. We propose an experimental web-based distributed system that incorporates data from wearable sensors and machine learning-based algorithms for monitoring the person's behavior and detection of anomalies. Various configurations of feature selection techniques and features as well as manual labeling for supervised learning have been used. In case of anomalies detected in older adult behavior, the caregiver is notified. Finally, we illustrate the system implementation and functionality considering Fitbit smart band sensor and integration with Fitbit Cloud. The results obtained using a public activity dataset with different configurations of machine learning anomaly detection algorithms and features are promising, showing an accuracy of 87% and an F1-score of 0.9.

MERITS - Various configuration of feature selection techniques and features as well as manual labelling for supervised learning have been used.

DEMERITS-The accuracy is not mentioned properly.

Research on Human Abnormal Behavior Detection Based on Deep Learning.

AUTHOR - Weihu Zhang, Chang Liu

YEAR – 2020

In order to make full use of the effective information in the video and improve the recognition rate of abnormal human behavior in complex scenes, we use a mixed Gaussian model to detect clear foreground moving target contours and perform Gaussian filtering on them to remove the effects of noise in the scene . By calculating the center point of the foreground pixel, and drawing a bounding box based on this, the key area of human motion in the video is extracted. Then we use the Farneback

dense optical flow algorithm to obtain spatiotemporal information. By combining CNN and LSTM, a CNN-LSTM hybrid two-stream network model based on the Dropout mechanism is established., input the original image and the superimposed optical flow image of the key area of the video sequence motion to learn the dynamic and static features and timing information in the spatiotemporal information. The weighted fusion method is used to perform weighted calculation on the Softmax output of the two-way network to obtain results.

MERITS-A mixed Gaussian model is used to detect clear foreground moving target contours and perform Gaussian filtering on them to remove the effects of noise in the scene.

DEMERITS-It is a complex model system which may produce inaccurate results.

Detection and Recognition of Abnormal Behavior based on Multi-level Residual Network.

AUTHOR - Huifang Qian, Xuan Zhou, Mengmeng Zheng

YEAR – 2020

The ability of real-time detection and recognition of abnormal behavior in video monitoring system is a key problem in intelligent monitoring system. This paper proposes a network framework based on multi-level residual network to detect and recognize abnormal human behavior from video. The framework of multi-level residual network includes human body detection module and posture recognition module. Based on the former, this paper proposes the detection residual network (d-Res) to adopt multi-scale target detection strategy to ensure the detection speed and detection effect of human body. The latter is used to extract spatial features of abnormal behaviors, and the recognition residual network (r-Res) based on transfer learning is used to extract deep features of images, so as to classify abnormal behaviors efficiently. Experiments are carried out on UTI dataset to evaluate the

performance of the proposed algorithm. The results show that the proposed method is effective in detecting and recognizing abnormal behaviors in real-world scenes.

MERITS-Detection residual network is proposed to adopt multi-scale target detection strategy to ensure the detection of speed and detection effect of human body.

DEMERITS-It is time-consuming.

Privacy-preserving Online Human Behaviour Anomaly Detection Based on Body Movements and Objects Positions

AUTHOR - Federico Angelini; Jiawei Yan; Syed Mohsen Naqvi

YEAR – 2019

Human behaviour anomaly detection is crucial for modern artificial intelligence systems. However, privacy protection plays a great role in the realization. In this paper, an online privacy-preserving anomaly detector is presented. The proposed method is able to discriminate on human subject body movements, postures and interactions with the surrounding objects, preserving subject privacy in all the tuning, training and testing stages. ActionXPose, Single Shot MultiBox Detector and Support Vector Machine are exploited for the proposed semi-supervised anomaly detector. The method successfully detects abnormal human behaviours, including unexpected body movements and misplaced objects. A new dataset ISLD-A is also proposed¹, providing suitable benchmark for performance evaluation².

MERITS—A new dataset ISLD-A is used, which can provide more accurate results.

DEMERITS— Usage of new dataset may lead to bandwidth and memory wastage.

Detecting human abnormal behaviour through a video generated model.

AUTHOR - Thomas Gatt, Dylan Seychell, AlexieiDingli

YEAR – 2019

Detecting human abnormal activities is the process of observing rare events that deviate from normality. In this study, an automated camera-based system that is able

to detect irregular human behaviour is proposed. PoseNet and OpenPose, which are pre-trained pose estimation models are used to detect the person in the frame and extract the body keypoints. Such data is used to train two types of AutoEncoders based on LSTM and CNN units in a semi-supervised approach where the goal is to learn a general representation of the normal behaviour. Evaluated on a challenging realistic video dataset, the results show that both types of models were able to correctly distinguish between normal and abnormal data sequences, with an average F-score of 0.93. The results also show that the proposed method outperformed similar work done on the same dataset. Furthermore, it was also determined that pose estimated data compares very well with sensor data. This shows that pose estimated data can be informative enough to understand and classify human actions.

MERITS—The combination of LSTM and CNN networks can be a powerful tool for processing sequential data.

DEMERITS—They require a much larger amount of training data to achieve the same level of accuracy.

Abnormal Human Behavior Recognition Based on Image Processing Technology.

AUTHOR - Rongyong Zhao, Yan Wang, Ping Jia, Cuiling Li, Yunlong Ma, Zhishu Zhang

YEAR – 2021

In recent years, digital image processing technology and computer vision technology have been continuously developed and made considerable progress. Based on the image information, through the detection, extraction and recognition operations, the human posture is understood, and abnormal behavior can be recognized in time. In this paper, related work in the abnormal human behavior analysis is first introduced, including many advanced practical projects. Then three methods of human target detection are listed, with the advantages and disadvantages being compared. The

detection process is designed according to the inter-frame difference method. Finally, specific realization is proposed from three aspects: area identification, feature extraction, abnormality determination, and the algorithm steps are clearly summarized. Therefore, this study can provide technical support and decision-making guidance for safety management, with a wide range of applications.

MERITS—Three different aspects are considered and the algorithm is executed.

DEMERITS— The level of accuracy is not mentioned.

A video-based abnormal human behavior detection for psychiatric patient monitoring.

AUTHOR - Shih-Chung Hsu, Cheng-Hung Chuang, Chung-Lin Huang, Por-Ren Teng, Miao-Jian Lin

YEAR – 2018

This paper proposes an abnormal human behavior detection system for monitoring psychiatric patient. A normal behavior can be characterized by the spatial and temporal features of human activities. The difficulty of abnormal behavior detection is that human behavior is unpredictable and complicated. It varies in both motion and appearance. The human behavior video stream is interspersed with transition of abnormal and normal events. Here, we propose an unsupervised learning using the N-cut algorithm along with the SVM to label the video segments and then apply the Condition random field (CRF) with an adaptive threshold to distinguish the normal and abnormal events.

MERITS—An adaptive threshold is used to differentiate normal and abnormal events.

DEMERITS— Using unsupervised learning technique may lead to unpredictable or difficult to understand results.

Computer-vision-based abnormal human behavior detection and analysis in electric power plant.

AUTHOR - Yuan Cao, Hao Xu, Qiang Yang

YEAR – 2021

With the increasing demand for intelligent security in power plants, the rapid and accurate processing of massive surveillance video data is urgently needed. Researches on the detection and analysis of abnormal human behaviors in power plants still focus on traditional image processing technology, and most of them lack robustness. In this article, abnormal behavior detection and analysis system based on personnel information are proposed to solve the above problems. The proposed method using an improved YoLov3 algorithm first detects persons and extracts abnormal behavior information on this basis. In the implementation, some training tricks are introduced to improve performance. Experimental results show that the system can effectively detect abnormal human behaviors, and the improved YoLov3 algorithm can also effectively improve model performance. The proposed abnormal behavior detection and analysis system based on personnel information prove its effectiveness through experiments, which can efficiently perform in power plants with lower computing costs.

MERITS–YoloV3 outperforms YOLOv5 in terms of speed.

DEMERITS– YoloV3 may have its own disadvantages when compared to the newer versions.

CHAPTER – 3

SYSTEM ANALYSIS

CHAPTER 3

SYSTEM ANALYSIS

EXISTING SYSTEM

The current system was designed using a statistical model for the detection of human behaviour, and it assigns a distribution fitting, which reduces the efficiency of the prediction model. The sole indicator of violent behaviour that the system currently has is one that can be accessed in real time. CCTV system for detecting unusual activity is the old system which has poor accuracy and poor efficiency in terms of the amount of time needed to load data and the amount of time needed to perform changes. In addition, neither the testing nor the training are carried out using the appropriate test-to-train split ratio.

DISADVANTAGES

- It does not provide high accuracy.
- It does not provide the location of the place where abnormal activity is detected nor send notification regarding the same.

PROPOSED SYSTEM

Yolo version 5 architecture is used in the model that was suggested. To train the most recent and accurate models with a large mini-batch size requires many GPUs; doing so with one GPU is not only incredibly sluggish but also inefficient. YOLO version 5 solves this problem by introducing an object detector that doesn't require multiple GPUs for training and has a smaller minimum batch size than previous versions. The system that has been developed makes it possible for a single GPU (1080 Ti or 2080 Ti) to train a rapid and precise object detector. The suggested system is capable of detecting human abnormalities using audio and video inputs respectively. The proposed system also includes an IOT system which notifies with

both an e-mail and a SMS to the respective mail id and phone number, as soon as abnormal activity is detected.

ADVANTAGES

- It doesn't require multiple GPUs for training and has a smaller minimum batch size than previous versions.
- It identifies the location of the area where abnormal activity is found and notifies the user of the situation.
- This could be utilized for security needs and implemented in surveillance systems and security robots.

CHAPTER – 4

REQUIREMENT ANALYSIS

CHAPTER 4

REQUIREMENT ANALYSIS

These are the requirements for the doing the project. Without these tools and softwares, we cannot do the project. So we have two requirements to do the project. They are

1. Hardware Requirements
2. Software Requirements

HARDWARE REQUIREMENTS

The hardware requirements may serve as the basis for a contract for the implementation of the system and should therefore be a complete and consistent specification of the whole system. They are used by software engineers as the starting point for the system design. It shows what the system does and not how it should be implemented.

- System : Pentium i3 Processor
- Hard Disk : 500 GB
- Monitor : 15'' LED
- Input Devices : Keyboard, Mouse
- Ram : 4 GB

SOFTWARE REQUIREMENTS

The software requirements document is the specification of the system. It should include both a definition and a specification of requirements. It is a set of what the system should do rather than how it should do it. The software requirements provide a basis for creating the software requirements specification. It is useful in estimating cost, planning team activities, performing tasks and tracking the team's and tracking the team's progress throughout the development activity.

- Operating system : Windows 10
- Coding Language : Python
- IDE : Anaconda prompt

CHAPTER –5

SYSTEM DESIGN

CHAPTER 5

SYSTEM DESIGN

5.1 DATA FLOW DIAGRAM

A data flow (DFD) diagram is a graphical representation of a "flow" of data through an information system, which modifies its process characteristics. DFD is often used as the first step in creating an overview of the system without going into too much detail, which can be explained later.

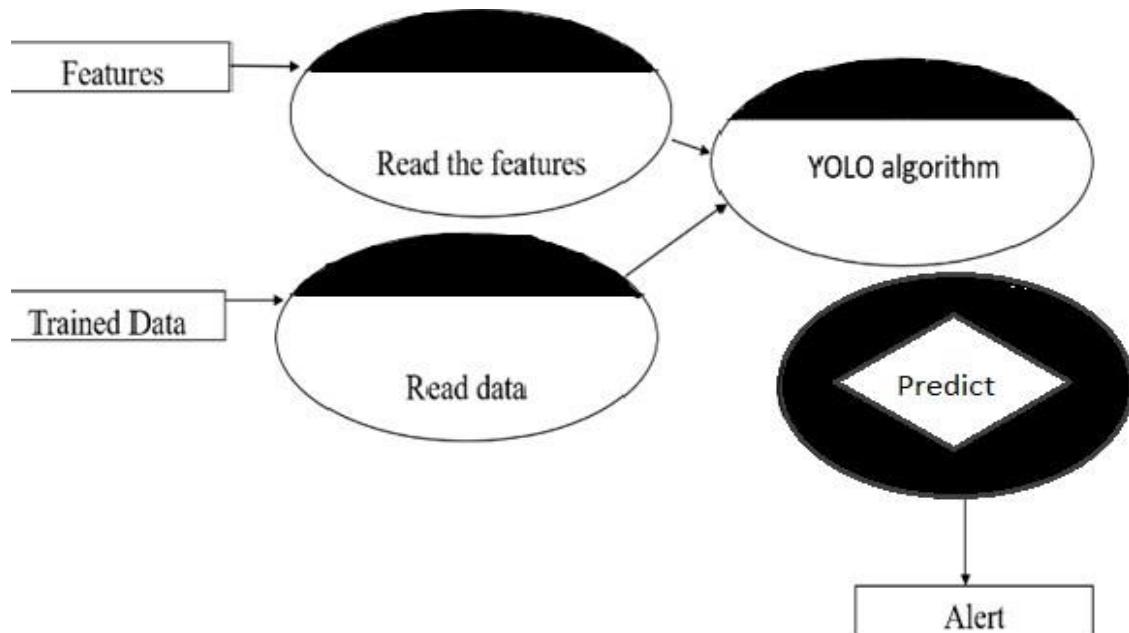


Fig. No. 5.1 Dataflow Diagram

UML DIAGRAM

USECASE DIAGRAM

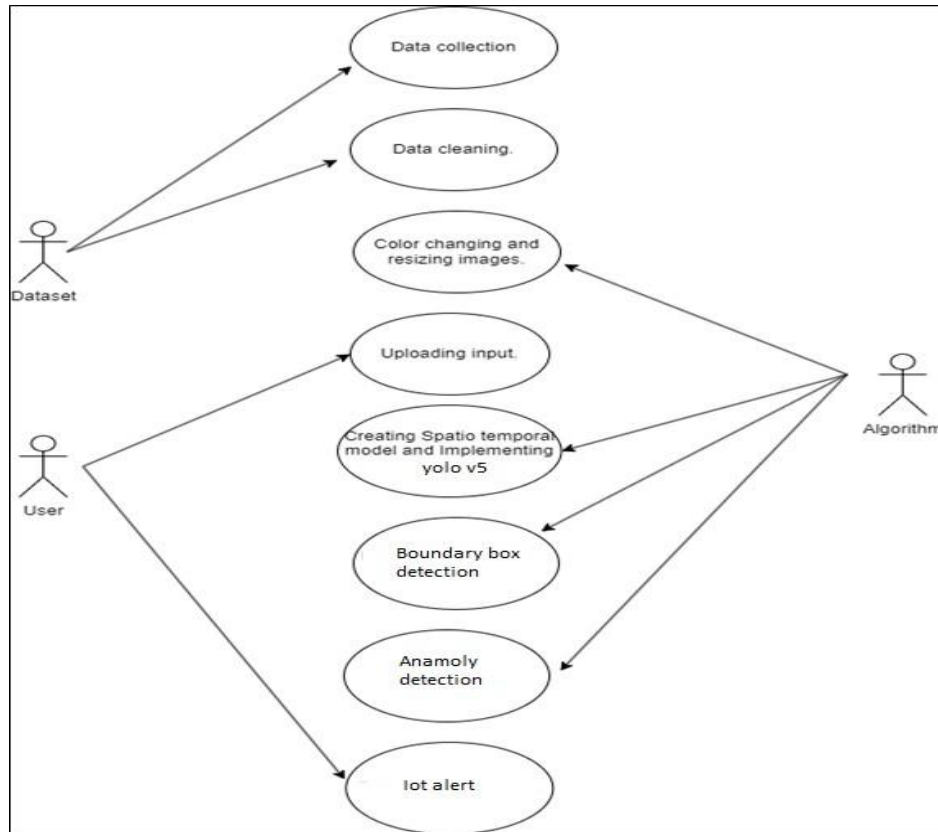


Fig. No. 5.2 Use Case Diagram

A simple user interface diagram represents the user interaction with a system that shows the relationship between the user and the different operating conditions in which the user is involved. The application case diagram can identify different types of system users and different operating conditions and will often be associated with other types of diagrams. Terms of use are represented by circles or ellipses. Three characters are present: dataset, user, and algorithm. The dataset is first gathered and cleaned. The algorithm is also in charge of altering the colour and size of images, building the spatio model, and putting Yolo V5 into practise. Algorithms are used to detect boundary boxes and anomalies. The user provides the input, and if an anomaly is found, the user is alerted anomaly behavior.

CLASS DIAGRAM

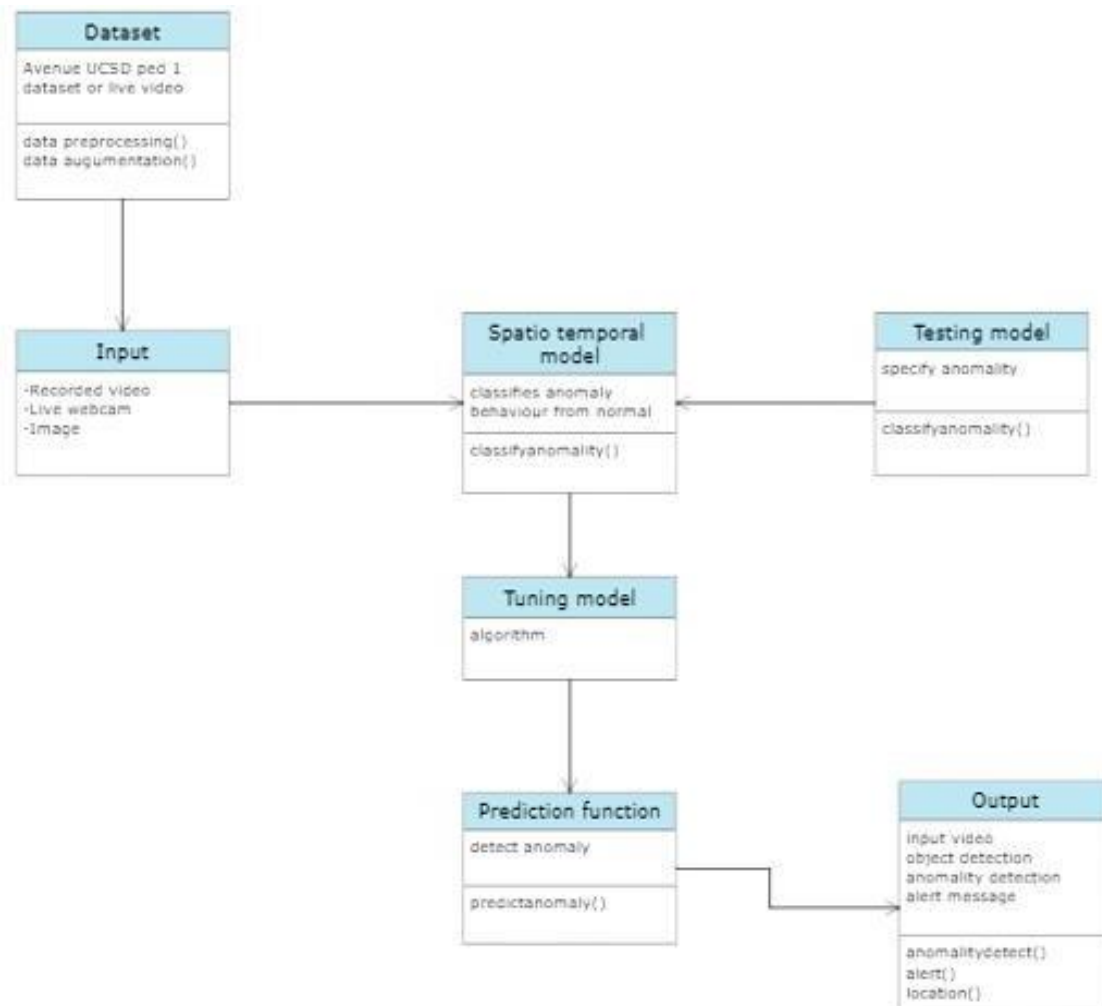


Fig. No. 5.3 Class Diagram

In the Unified Modeling Language (UML), a class diagram is a form of static structural diagram that shows the system's classes, attributes, actions (or methods), and relationships among objects. A static diagram is a class diagram. It depicts an application's static view. A class diagram is used not only for visualizing, describing, and documenting many parts of a system but also for creating executable code for a software program. A class diagram depicts a class's attributes and operations, as well as the system's limitations. The class diagram that is shown below is a detailed diagrammatic representation of the classes that are being used.

The dataset class is used for data preprocessing and data augmentation; the input information class includes recorded video, images, and live webcam; the spatiotemporal model class is used to distinguish abnormal activities from normal; the testing the model class is also used to classify the abnormality; the tuning model contains the algorithm needed for training and testing the model; the prediction function is used to detect the abnormality and finally, the output class is utilised to identify the anomaly and designate the weapon as output along with a location and alert message.

SEQUENCE DIAGRAM

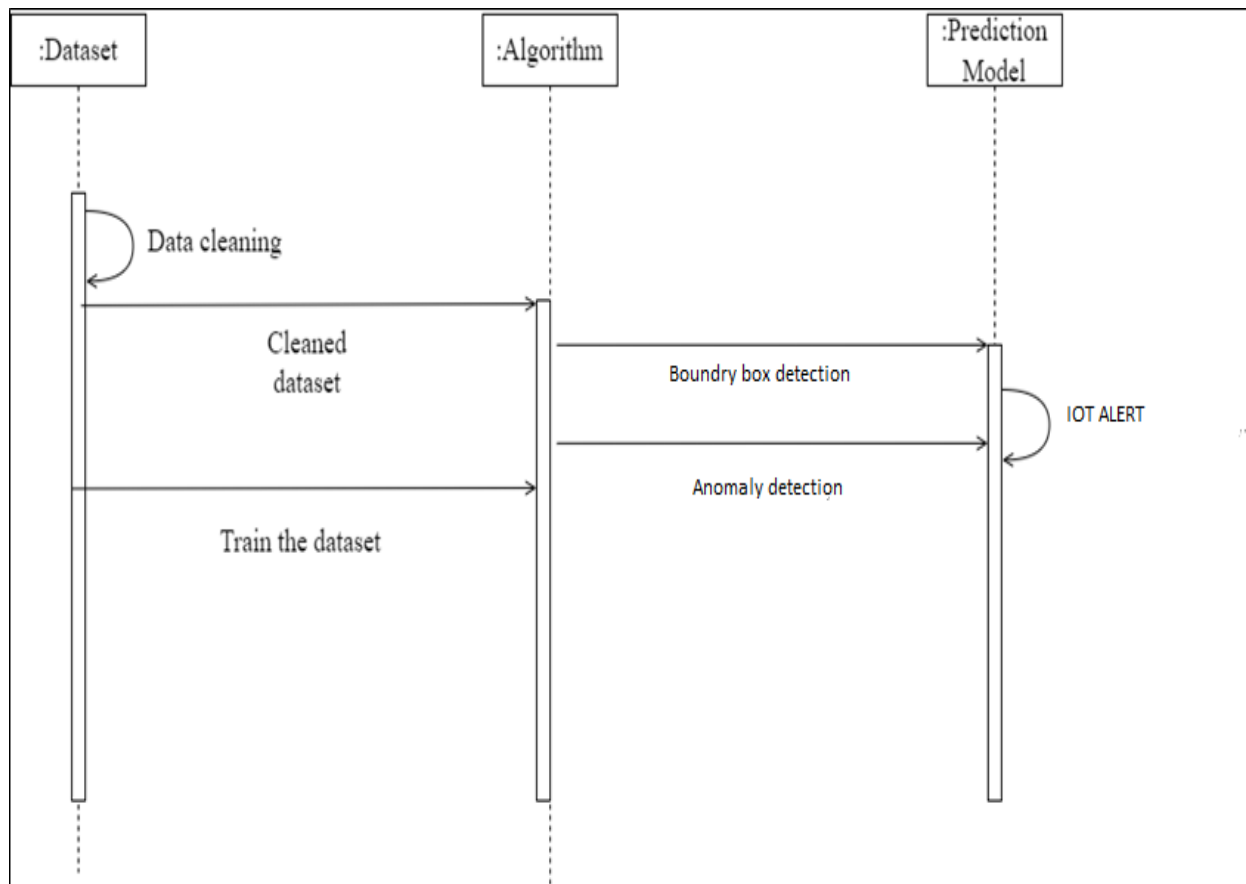


Fig. No. 5.4 Sequence Diagram

The UML Sequence Diagrams flowchart shows how the tasks are performed. They capture the interaction between objects in a shared space. Sequential diagrams focus on time and show the order of communication using the direct axis of the drawing to represent the time and messages sent and received. There are three characters in this project-dataset, algorithm and prediction. Data cleaning is done to the dataset first. The algorithm is then trained using the cleaned dataset. The prediction model is preloaded with the boundary box detection and anomaly detection algorithms. An alert message for anomalous behaviour is received after the anomaly is predicted.

ACTIVITY DIAGRAM

Activity diagrams are widely used as a flowchart showing system functions. Function diagrams are different from flowcharts in that they have additional features. Branches, parallel flow, and swimming pools are examples of these additional skills. Another important diagram in the UML to describe the dynamic features of the system is the function diagram. An activity diagram is a flow chart that shows the movement of information from one to another. The action can be defined as a system activity. From action to action, a flow of control is expressed. The activity diagram loads the dataset and performs preprocessing. Unwanted data are eliminated during preprocessing. Following that, images are resized and their colour channels are adjusted. Finally, abnormal behaviour is predicted and detected using the prediction function.

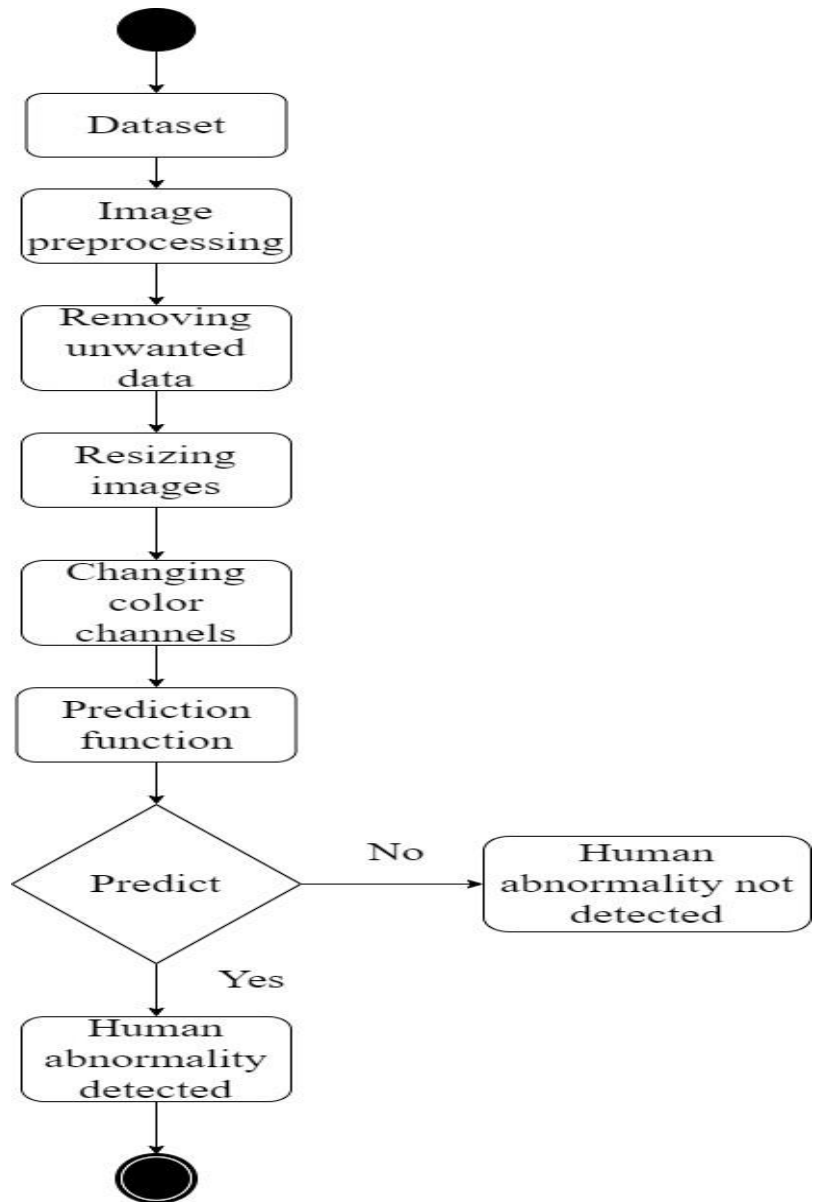


Fig. No. 4.6 Activity Diagram

CHAPTER –6

SYSTEM ARCHITECTURE

CHAPTER 6

SYSTEM ARCHITECTURE

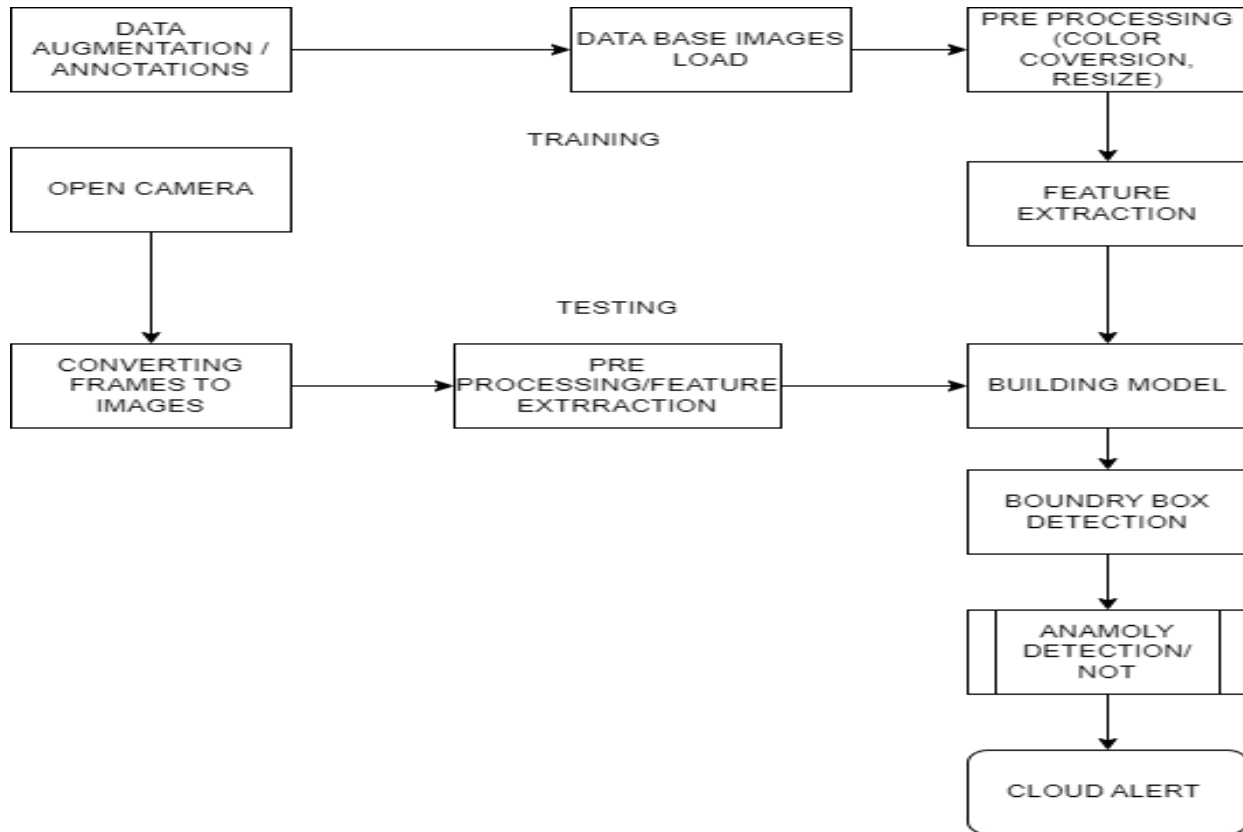


Fig. No. 6.1 System Architecture

The project's building process is shown in the architecture diagram. It starts with downloading a dataset from Kaggle and loading it for training. Preprocessing resizes the frames to 227x227px and converts them from RGB colour channel to greyscale to reduce the number of pixels. Features have been extracted. The web camera will be used in the same way for testing, and it will simply convert frames into images following conversions and resizing are completed. The YOLOv5 model will then be used to build a spatiotemporal model, and boundary box detection will be used to create a box around the object that was detected. A user's phone or email will receive

an alert message with the user's location if anomalous activity is detected after input is given and is identified.

MODULE DESIGN SPECIFICATION

MODULES

- Dataset Collection
- Data Preprocessing
- Data Augmentation
- Building model-Training
- Classification
- Prediction detection-Testing

DATASET COLLECTION

Dataset	Type	Number of images
Train	USCD Ped 1	479
	USCD Ped 2	383
	Total	862
Test	USCD Ped 1	115
	USCD Ped 2	75
	Total	190

Table 6.1Dataset

We captured every 10 frames using the Avenue UCSD Ped1 and Ped 2 dataset that we obtained from Kaggle. Avenue Dataset includes 21 video clips for assessment and 16 for training.

DATA PREPROCESSING

Data preprocessing changes the data into a format that can be processed in data mining, machine learning, and other data science tasks more quickly and efficiently. Each extracted frame is resized to 227x277. All pixel values are scaled between 0 and 1 so that to make all input image frames on same level. Normalization is performed by subtracting every frame from global mean image which is formulated as: $\text{Global Mean Image} = X/Y$

DATA AUGMENTATION

Data augmentation is a technique of artificially increasing the training set by creating modified copies of a dataset using existing data. It includes making minor changes to the dataset or using deep learning to generate new data points. Change the color channel: We'll convert the frames' RGB colour channel to Grey Scale. A grayscale image is one in which each pixel's value is a single sample carrying just information about the intensity of the light. Resizing the frame: We will resize the frames to a size of 227x227px.

BUILDING MODEL-TRAINING

We will create a Spatio Temporal Model with the Yolov5 model, using Convolution layers to build our model. YOLO v5 is natively implemented in PyTorch, eliminating the Darknet framework's limitations (based on C programming language). It is an object detection model. In this project, we make it as custom dataset. This massive change of YOLO to the PyTorch framework made it easier for the developers to modify the architecture and export to many deployment environments straightforwardly.

CLASSIFICATION

A collection of predefined classes that the algorithm was trained for are used by image classification algorithms to predict the type or class of an object in an image. Typically, input consists of an image of a singular object, like a cat. Output is a class or label that designates a specific item, frequently with a probability attached to it. Using bounding boxes, object localization algorithms identify the existence of an object in the image. They use the position, height, and width of the objects in the input picture to determine the location of one or more bounding boxes.

- Image classification: Algorithms produce a list of object categories present in the image.
- Single-object localization: Algorithms produce a list of object categories present in the image, along with an axis-aligned bounding box indicating the position and scale of one instance of each object category.
- Object detection: Algorithms produce a list of object categories present in the image along with an axis-aligned bounding box indicating the position and scale of every instance of each object category.

PREDICTION DETECTION-TESTING

Prediction methods are used to detect anomalous events by comparing them to their predicted values. We will create Prediction function to detect anomaly in a video. The next step is to create a function that can evaluate inputs such as images, videos, and times to determine if they are anomalous or not. Using the Python predict() method, we can forecast the labels of the data values based on the trained model.

ALGORITHM

YOLO V5 (OBJECT DETECTION)

An object detection algorithm is an algorithm that is capable of detecting certain objects or shapes in a given frame. YOLO, an acronym for ‘You only look once’, is an open-source software tool utilized for its efficient capability of detecting objects in a given image in real time. The YOLO algorithm uses convolutional neural network (CNN) models to find objects in images. All of the objects in the image can be detected by the algorithm with just one forward propagation through the given neural network. Through its one-forward propagation capability, the YOLO algorithm not only offers high detection speed and performance, but it also detects them very accurately and precisely. The YOLO algorithm is one of the most well-known detection algorithms to date because it has an advantage over others in terms of speed due to this. The steps listed below describe how the YOLO v5 algorithm operates.

➤ **Step 1: Residual Blocks (Dividing the Image Into Smaller, Grid-Like Boxes)**

This process entails breaking the entire frame up into smaller grids or boxes. Identical in size and shape, each grid is drawn over the original image. Each grid box in these divisions is supposed to be able to recognise the various objects that are contained within it.

➤ **Step 2: Bounding Box Regression (Identifying the Object Inside a Bounding Box)**

A bounding box is created around a specific object after it has been identified in an image. The centre point, height, width, and class are just a few of the bounding box’s attributes. (object type detected).

➤ **Step 3: Intersection Over Union**

To determine the precision of our model, we compute the intersection over union, or IOU. To do this, we measure the degree to which the real value box and the box that contains the result of our computation intersect. Assuming that our IOU value is 40%, this prediction shouldn't be taken into account if the intersection of the two boxes is less than 40%. We do this to make it easier for us to assess the precision of our forecasts.

SPATIO TEMPORAL MODEL

The term spatio-temporal modelling refers to research that records and analyses both the locations and times of the observations. The spatio-temporal intensity of incident or prevalent cases, which varies depending on the average number of incident or prevalent cases in combinations of place and time units over the geographical region and time period of interest, is the focus of spatio-temporal analysis. A rapid response team can make longer-term plans and decide where and when to focus prevention and control efforts by using real-time spatio-temporal surveillance. The Spatio Temporal model extracts spatial characteristics from video frames. For precise predictions, the spatiotemporal captures the temporal/sequential and recurrent dependencies within a specified sequence length. While anomalies frequently deviate from these dependencies and are unpredictable, normal instances usually closely adhere to them and can be accurately predicted. Therefore, the prediction error reveals the level of an instance's anomaly. In our proposed work, this model is combined with the Yolo V5 model. Additionally, it aids in pinpointing the location of anomalies and sending alert messages.

CHAPTER – 7

PERFORMANCE ANALYSIS

CHAPTER 7

PERFORMANCE ANALYSIS

RESULT AND DISCUSSION

Both recorded and live videos can be used to spot unusual activity. An alert text message with the location of the detection is sent to the user if an unusual activity or weapon is discovered. Additionally, the user receives an email with the location and an image of the abnormal activity. It can also identify people and other objects in the video, but an alert message will only be sent if abnormal activity is discovered. The project heavily relies on object detection. Receiving alert messages with location information ensures high security and quick response to threats.

ACCURACY PREDICTION

The graph demonstrates that our suggested method outperforms the majority of semi-supervised methods in terms of stability and accuracy. Additionally, our suggested method's internal group convolution structure makes it simpler to adapt our method to various data sets. To support this claim, more research and data are required. When the data set gets bigger and better in the future, we think our methodology will perform better. To show the effectiveness and precision of the suggested model, the precision of the accepted techniques in the area of perceiving anomalous behavior following the model's evaluation and testing, we applied our suggested algorithm to the Avenue UCSD-Ped 1 dataset and carried out training and testing. The results of the testing are shown below.

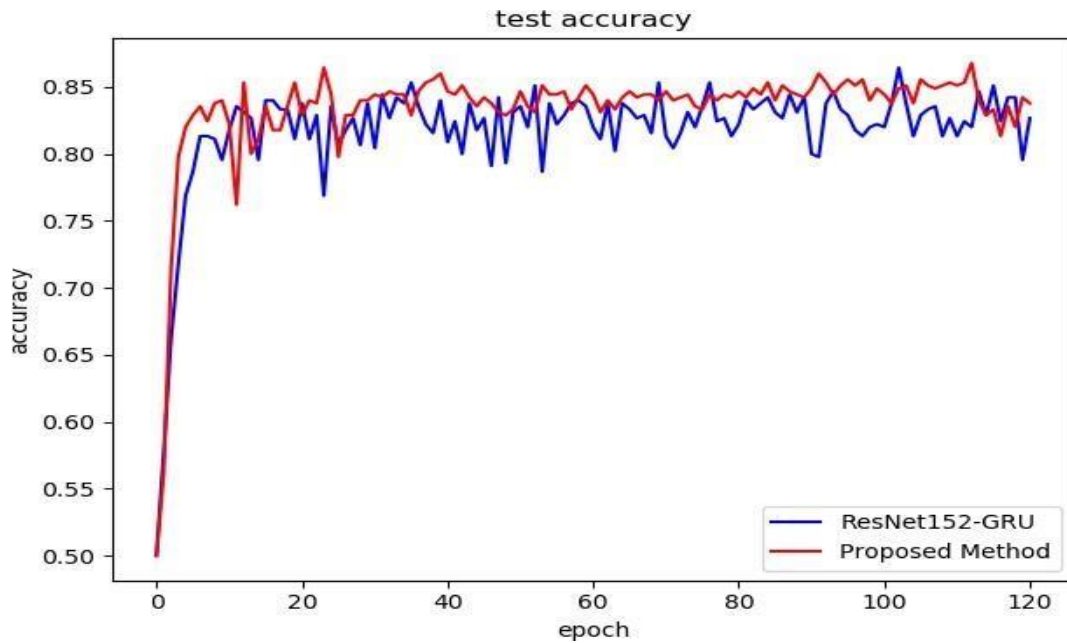


Fig No. 7.1 Testing accuracy of the model

8.3 COMPARISON OF ACCURACIES IN PROPOSED AND RELATED WORK

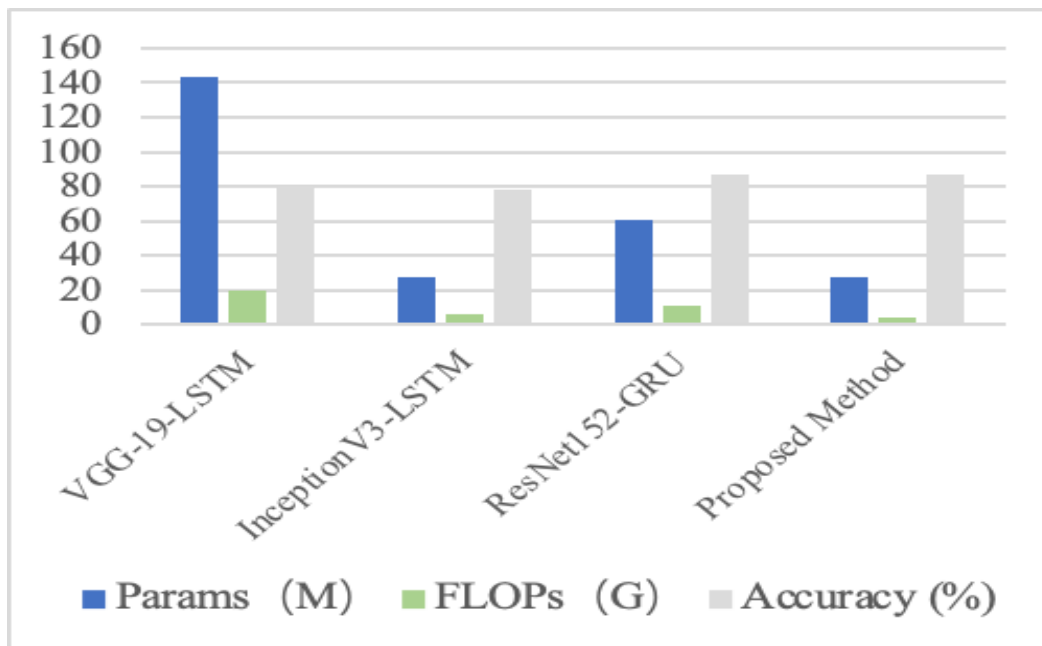


Fig No. 7.2 Comparison with previous work

The analysis above compares the number of parameters, floating-point computations per second, and accuracy of our suggested method with a number of other widely used methods. The results demonstrate that our method outperforms competing methods on the anomalous Avenue UCSD Ped 1 dataset. We chose some representative methods for comparison. Our network's accuracy is better than other methods', and it's more efficient in terms of the number of parameters and the number of operations, so it can be used with the majority of mainstream computers available today. As a result, our method will be more effective at identifying abnormal behaviour in surveillance footage that is both complex and abundant in daily life. Our approach will be more effective at identifying abnormal behaviour. Additionally, the unique structure of our model facilitates training because it allows the algorithm's overall accuracy and stability to be continuously improved as new categories and data are added.

CHAPTER –8

CONCLUSION

CHAPTER 8

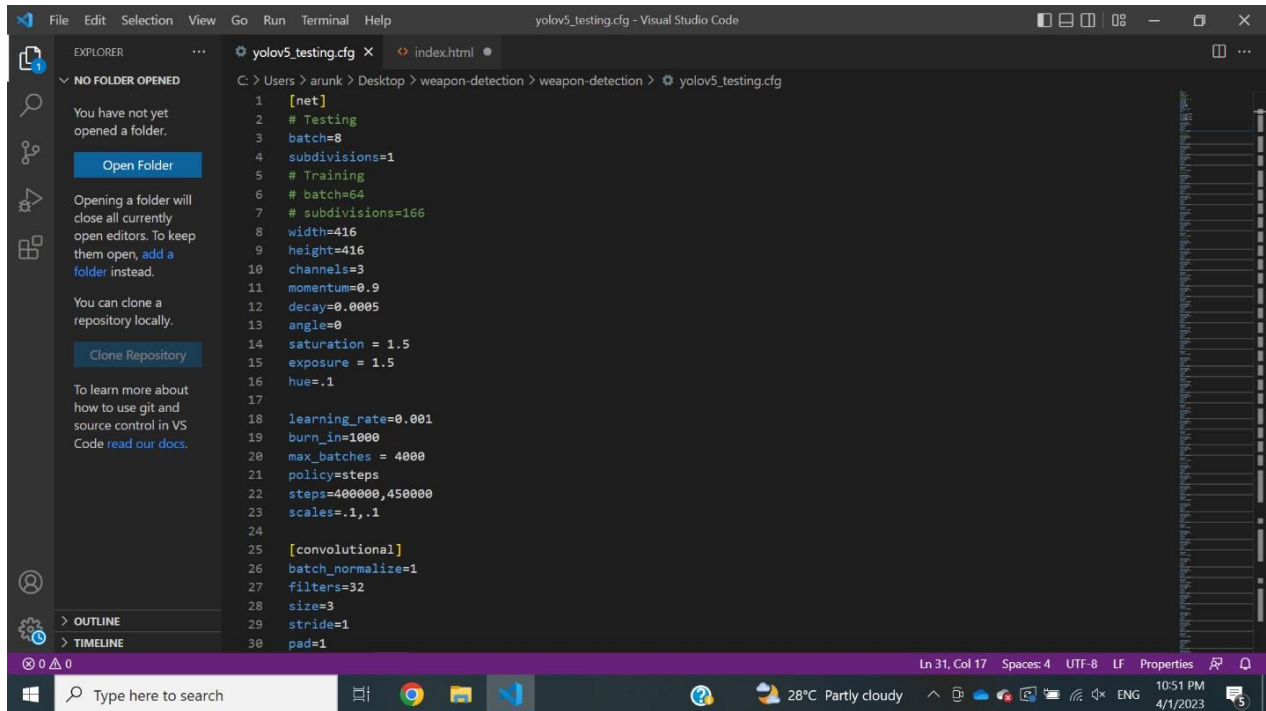
CONCLUSION

8.1 CONCLUSION & FUTURE ENHANCEMENT

In this study, the state-of-the-art YOLO V5 object detection model was implemented and trained over our collected dataset for weapon detection. We propose a model that provides a visionary sense to a machine to identify the unsafe weapon and can also alert the human administrator when a gun or a firearm is obvious in the edge. Experimental results show that the trained YOLO V5 has better performance compared to the YOLO V3 model and more or less costs the same computationally. There is an immediate need to update the current surveillance capabilities with improved resources to support monitoring the effectiveness of human operators. Smart surveillance systems would fully replace current infrastructure with the growing availability of low-cost storage, video infrastructure, and better video processing technologies. Eventually, the digital monitoring systems in terms of robots would fully replace current surveillance systems with the growing availability of cheap computing, video infrastructure, high-end technology, and better video processing.

APPENDICES

SAMPLE SCREENS



The screenshot shows the Visual Studio Code editor interface. The Explorer panel on the left indicates 'NO FOLDER OPENED'. The main editor area displays the 'yolo5_testing.cfg' file with the following configuration:

```
1 [net]
2 # Testing
3 batch=8
4 subdivisions=1
5 # Training
6 # batch=64
7 # subdivisions=166
8 width=416
9 height=416
10 channels=3
11 momentum=0.9
12 decay=0.0005
13 angle=0
14 saturation = 1.5
15 exposure = 1.5
16 hue=.1
17
18 learning_rate=0.001
19 burn_in=1000
20 max_batches = 4000
21 policy=steps
22 steps=400000,450000
23 scales=.1,.1
24
25 [convolutional]
26 batch_normalize=1
27 filters=32
28 size=3
29 stride=1
30 pad=1
```

The status bar at the bottom shows 'Ln 31, Col 17', 'Spaces: 4', 'UTF-8', 'LF', 'Properties', and the system clock '10:51 PM 4/1/2023'.

Fig No. A.1 Yolo Testing



Fig No. A.2 Detection of weapon through web camera



Fig No. A.3 Detection of weapon in a recorded video

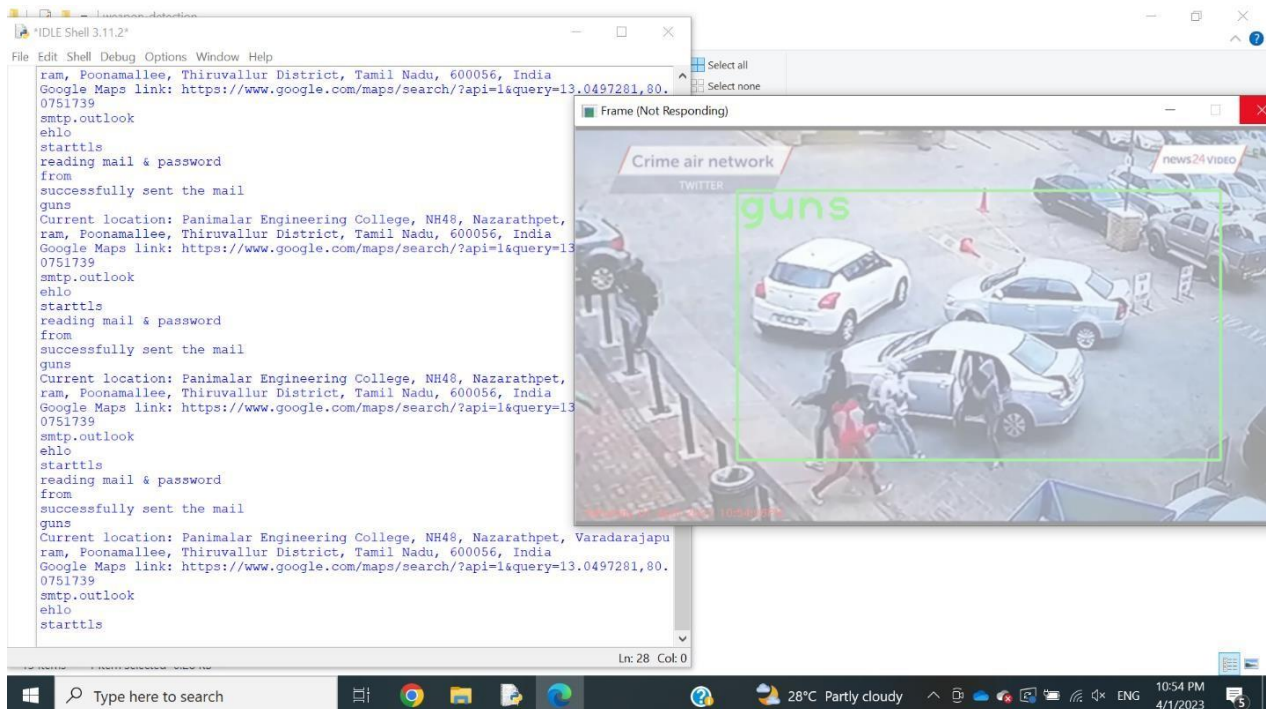


Fig No. A.4 Output

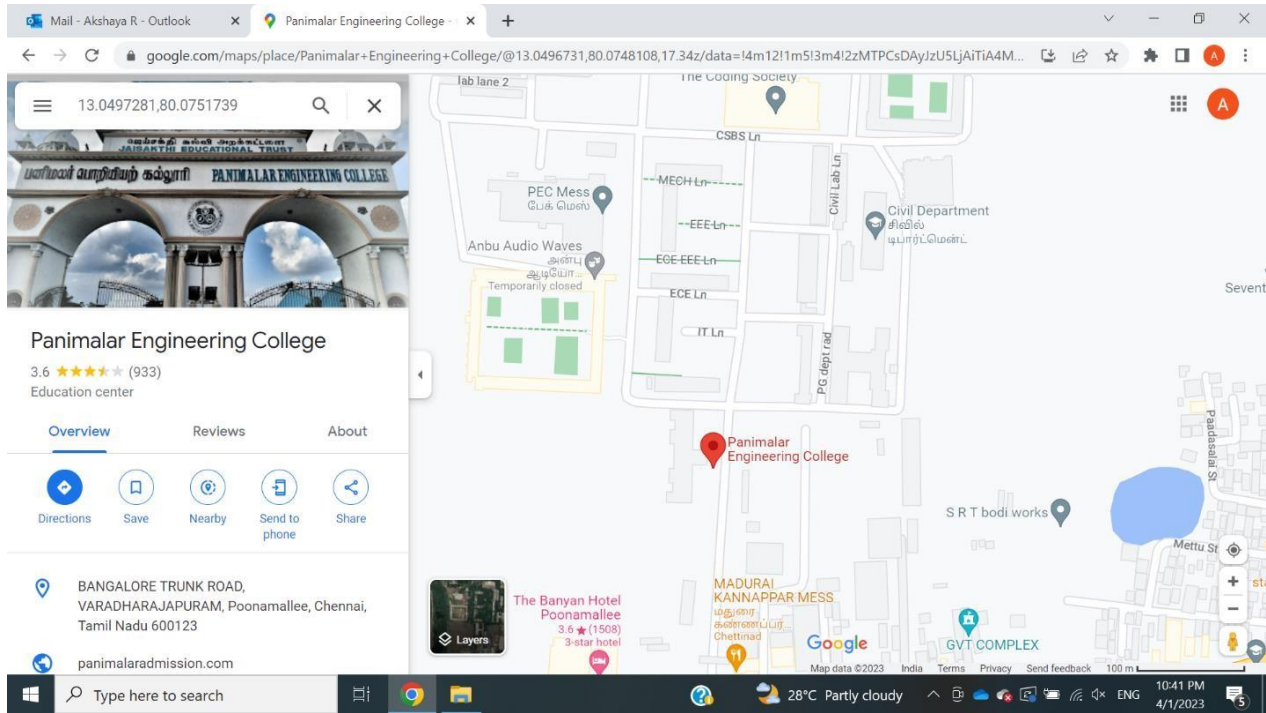


Fig No. A.5 Location of the camera

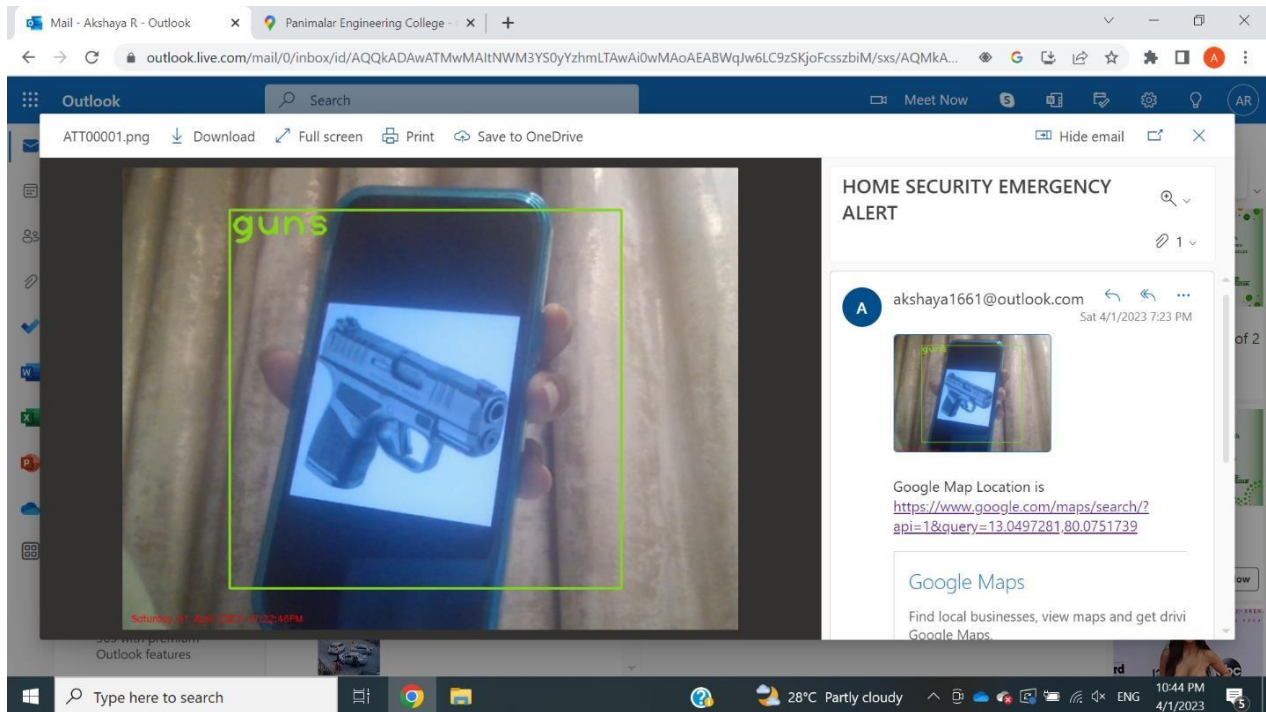


Fig No. A.6 Email received after detecting a weapon in a live video

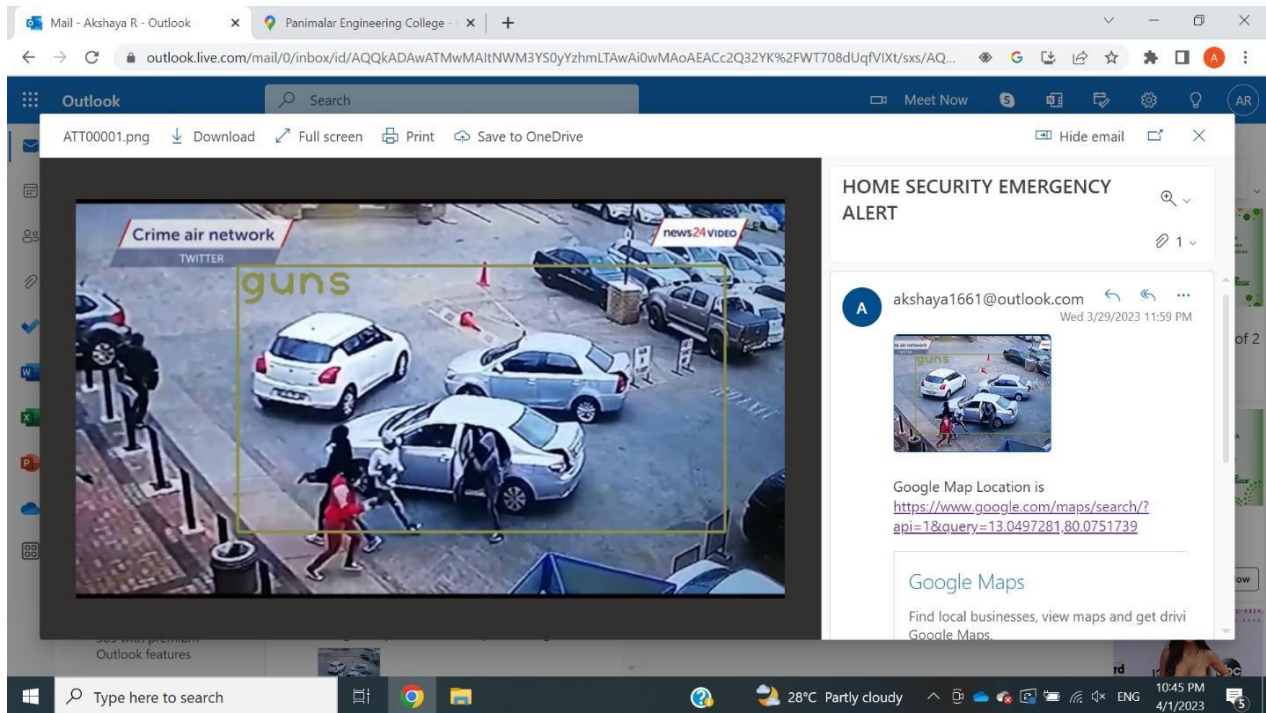


Fig No. A.7 Email received after detecting weapon in arecorded video

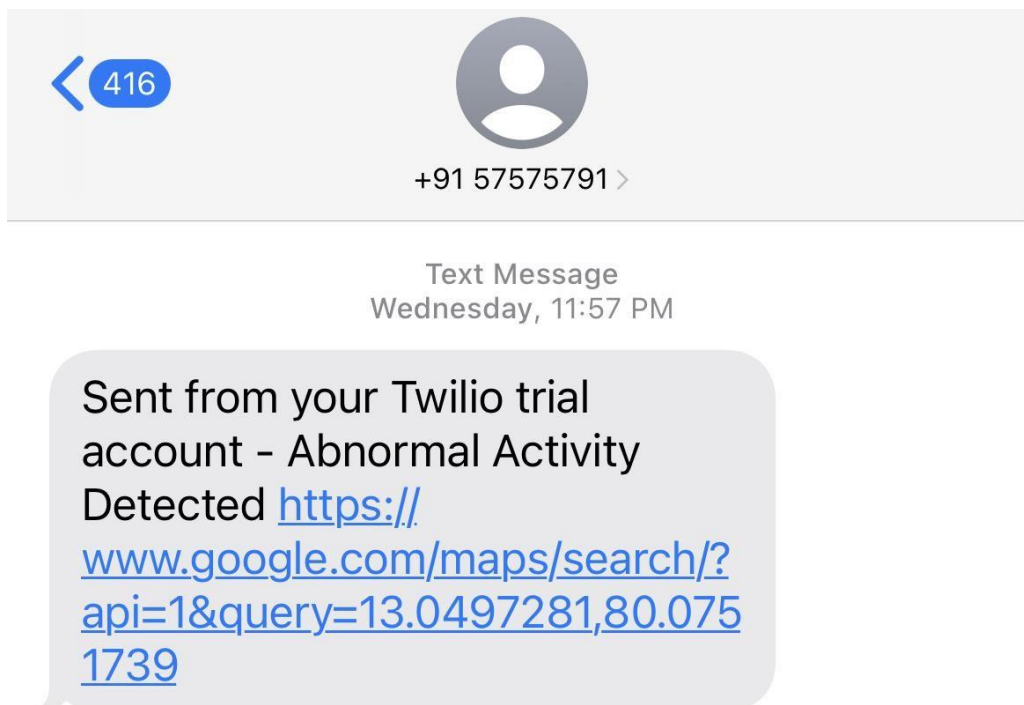


Fig No. A.8SMS notifying the detection

CODING

Yolo Object Detection

```
import cv2
import imutils
import numpy as np
#import serial
import time
import cv2
import imutils
import numpy as np
from subprocess import call
import time
import datetime
import os
import glob
import smtplib
import base64
import urllib
from email.mime.image import MIMEImage
from email.mime.multipart import MIMEMultipart
from email.mime.text import MIMEText
import random
import time
import subprocess
import webbrowser
import sys
import socket
import json
import twilio
import geopy
from urllib.request import urlopen

#https://outlook.live.com/mail/0/options/mail/accounts/popImap
yahoo_user = "akshaya1661@outlook.com"
yahoo_pwd = "akshayaram1661"
FROM = 'akshaya1661@outlook.com'
TO = ['akshaya1661@outlook.com'] #must be a list
```

```

# Load Yolo
net = cv2.dnn.readNet("yolov5_training_2000.weights", "yolov5_testing.cfg")
#weapon
#net = cv2.dnn.readNet("yolov5.weights", "yolov5.cfg")
classes = []
with open("coco.names", "r") as f:
    classes = [line.strip() for line in f.readlines()]
layer_names = net.getLayerNames()
output_layers = [layer_names[i - 1] for i in net.getUnconnectedOutLayers()]
colors = np.random.uniform(0, 255, size=(len(classes), 3))

def safe():
    #map tracking
    # Get GPS coordinates from geopy
    from geopy.geocoders import Nominatim
    geolocator = Nominatim(user_agent="geoapiExercises")

    # Open following URL to get IP address
    data = json.load(urlopen("http://ipinfo.io/json"))

    lat = '13.0497281'
    lon = '80.0751739'
    # Extract latitude and longitude
    #lat = data['loc'].split(',')[0]
    #lon = data['loc'].split(',')[1]

    # Get location address from GPS coordinates
    location = geolocator.reverse(lat + "," + lon)

# Construct Google Maps link
maps_url = f"https://www.google.com/maps/search/?api=1&query={lat},{lon}"

# Print address and Google Maps link
print("Current location: " + location.address)
print("Google Maps link: " + maps_url)

# Open Google Maps link in default browser
webbrowser.open(maps_url)

```

```

# Open Google Maps link in default browser
webbrowser.open(maps_url)

#sms alert
# Import required libraries
from twilio.rest import Client

# Define your Twilio account SID and auth token
account_sid = 'ACe30d2c962e883f99562ef210c8b441ba'
auth_token = 'f09bfb1dde1f285ecb676891b70f94b7'

# Create a Twilio client object
client = Client(account_sid, auth_token)

# Define the sender and recipient phone numbers
sender = '+14346239728' # your Twilio phone number
recipient = '+919551166631' # the recipient phone number

# Define the message to be sent
message = " Abnormal Activity Detected "+maps_url

# Send the message using the client object
message = client.messages.create(
    body=message,
    from_=sender,
    to=recipient
)

#print(message.sid)

#mail system
msg = MIMEMultipart()
time.sleep(1)
msg['Subject'] ="HOME SECURITY EMERGENCY ALERT"

#BODY with 2 argument
#variable = maps_url
#body=sys.argv[1]+sys.argv[2]
body= "Google Map Location is "+maps_url

```

```

msg.attach(MIMEText(body,'plain'))
time.sleep(1)

###IMAGE
fp = open("1.png", 'rb')
time.sleep(1)
img = MIMEImage(fp.read())
time.sleep(1)
fp.close()
time.sleep(1)
msg.attach(img)
time.sleep(1)
try:
    server = smtplib.SMTP("smtp.office365.com", 587) #or port 465 doesn't
    seem to work!
    print ("smtp.outlook")
    server.ehlo()
    print ("ehlo")
    server.starttls()
    print ("starttls")
    server.login(yahoo_user, yahoo_pwd)
    print ("reading mail & password")
    server.sendmail(FROM, TO, msg.as_string())
    print ("from")
    server.close()
    print ('successfully sent the mail')
except:
    print ("failed to send mail")
# Loading image
cap = cv2.VideoCapture(0)
while True:

    ret,img=cap.read()
    img = imutils.resize(img, width=720)
    height, width, channels = img.shape

    # Detecting objects
    blob = cv2.dnn.blobFromImage(img, 0.00392, (416, 416), (0, 0, 0), True,
    crop=False)

```



```

net.setInput(blob)
outs = net.forward(output_layers)

# Showing informations on the screen
class_ids = []
confidences = []
boxes = []
for out in outs:
    for detection in out:
        scores = detection[5:]
class_id = np.argmax(scores)
        confidence = scores[class_id]
        if confidence > 0.5:
            # Object detected
center_x = int(detection[0] * width)
center_y = int(detection[1] * height)
            w = int(detection[2] * width)
            h = int(detection[3] * height)

            # Rectangle coordinates
            x = int(center_x - w / 2)
            y = int(center_y - h / 2)
boxes.append([x, y, w, h])
confidences.append(float(confidence))
class_ids.append(class_id)

indexes = cv2.dnn.NMSBoxes(boxes, confidences, 0.5, 0.4)
font = cv2.FONT_HERSHEY_PLAIN
for i in range(len(boxes)):
    if i in indexes:
        x, y, w, h = boxes[i]
        label = str(classes[class_ids[i]])
        color = colors[i]
        cv2.rectangle(img, (x, y), (x + w, y + h), color, 2)
        cv2.putText(img, label, (x, y + 30), font, 3, color, 3)
        cv2.putText(img, datetime.datetime.now().strftime("%A %d %B %Y
%i:%M:%S%p"),
                    (10, img.shape[0] - 10), cv2.FONT_HERSHEY_SIMPLEX, 0.35, (0,
0, 255), 1)

```

```

if(label == "guns"):
    cv2.imwrite("1.png",img)
safe()
time.sleep(5)
    print(label)
elif(label == "knife"):
    cv2.imwrite("1.png",img)
safe()
time.sleep(5)
    print(label)
frametime = 100
    # show the output frame
    cv2.imshow("Frame", img)
    key = cv2.waitKey(frametime) & 0xFF

    # if the `q` key was pressed, break from the loop
    if key == ord("q"):
        break
cv2.destroyAllWindows()

```

Weapon Detection

```

import cv2
import numpy as np

# Load Yolo

net = cv2.dnn.readNet("yolov3_training_2000.weights", "yolov3_testing.cfg")
net.setPreferableBackend(cv2.dnn.DNN_BACKEND_DEFAULT)
net.setPreferableTarget(cv2.dnn.DNN_TARGET_CPU)
classes = ["Guns"]

cap = cv2.VideoCapture(0)

while True:
    _, img = cap.read()
    height, width, channels = img.shape

```

```

# width = 512
# height = 512

# Detecting objects
blob = cv2.dnn.blobFromImage(img, 0.00392, (416, 416), (0, 0, 0), True,
crop=False)
net.setInput(blob)

layer_names = net.getLayerNames()

output_layers = [layer_names[i - 1] for i in net.getUnconnectedOutLayers()]
colors = np.random.uniform(0, 255, size=(len(classes), 3))
outs = net.forward(output_layers)
# Showing information on the screen
class_ids = []
confidences = []
boxes = []
for out in outs:
    for detection in out:
        scores = detection[5:]
class_id = np.argmax(scores)
        confidence = scores[class_id]
        if confidence > 0.5:
            # Object detected
center_x = int(detection[0] * width)
center_y = int(detection[1] * height)
            w = int(detection[2] * width)
            h = int(detection[3] * height)

            # Rectangle coordinates
            x = int(center_x - w / 2)
            y = int(center_y - h / 2)

boxes.append([x, y, w, h])
confidences.append(float(confidence))
class_ids.append(class_id)

indexes = cv2.dnn.NMSBoxes(boxes, confidences, 0.5, 0.4)
print(indexes)
if indexes == 0: print("weapon detected in frame")

```

```

font = cv2.FONT_HERSHEY_PLAIN
for i in range(len(boxes)):
    if i in indexes:
        x, y, w, h = boxes[i]
        label = str(classes[class_ids[i]])
        color = colors[class_ids[i]]
        cv2.rectangle(img, (x, y), (x + w, y + h), color, 2)
        cv2.putText(img, label, (x, y + 30), font, 3, color, 3)
# frame = cv2.resize(img, (width, height), interpolation=cv2.INTER_AREA)
cv2.imshow("Image", img)
key = cv2.waitKey(1)
if key == 27:
    break
cap.release()
cv2.destroyAllWindows()

```

DATASET



REFERENCES

1. Ioana Alexandra Bozdog, Todea Daniel-Nicusor, Marcel Antal; Claudia Antal, Tudor Cioara, IonutAnghel, IoanSalomie, “Human Behavior and Anomaly Detection using Machine Learning and Wearable Sensors”, IEEE 17th International Conference on Intelligent Computer Communication and Processing (ICCP), 2022
2. Huifang Qian, Xuan Zhou, Mengmeng Zheng, “Detection and Recognition of Abnormal Behavior based on Multi-level Residual Network”, IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2020
3. Weihu Zhang, Chang Liu, “Research on Human Abnormal Behavior Detection Based on Deep Learning”, International Conference on Virtual Reality and Intelligent Systems (ICVRIS), 2021
4. Federico Angelini, Jiawei Yan, Syed Mohsen Naqvi, “Privacy-preserving Online Human Behaviour Anomaly Detection Based on Body Movements and Objects Positions”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019
5. Thomas Gatt, Dylan Seychell, AlexieiDingli, “Detecting human abnormal behaviour through a video generated model”, 11th International Symposium on Image and Signal Processing and Analysis (ISPA), 2019
6. Rongyong Zhao, Yan Wang, Ping Jia, Cuiling Li, Yunlong Ma, Zhishu Zhang, “Abnormal Human Behavior Recognition Based on Image Processing Technology”, IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2021
7. Shih-Chung Hsu, Cheng-Hung Chuang, Chung-Lin Huang, Por-Ren Teng, Miao-Jian Lin, “A video-based abnormal human behavior detection for psychiatric

- patient monitoring”, International Workshop on Advanced Image Technology (IWAIT), 2018
8. Yuan Cao, Hao Xu, Qiang Yang, “Computer-vision-based abnormal human behavior detection and analysis in electric power plant”, 33rd Chinese Control and Decision Conference (CCDC), 2021
 9. Yixue Hao, Zaiyang Tang, Bander Alzahrani, Reem Alotaibi, Reem Alharthi, Miaomiao Zhao, Arif Mahmood, “An End-to-End Human Abnormal Behavior Recognition Framework for Crowds With Mentally Disordered Individuals”, IEEE Journal of Biomedical and Health Informatics (Volume: 26, Issue: 8, August 2022), 2021
 10. Yimin DOU, Cai Fudong, Jinping LI, Cheng Wei, “Abnormal Behavior Detection Based on Optical Flow Trajectory of Human Joint Points”, Chinese Control And Decision Conference (CCDC), 2019