“DETECTING CRIME SCENES USING ML”

Submitted in partial fulfillment of the requirements

of the degree of

Bachelor of Computer Engineering

by

Drishti Jalgaonkar

Juilee Gund

Neha Patil

Under Guidance of

Prof. M.M.Phadke

University of Mumbai

Department of Computer Engineering

Datta Meghe College of Engineering

plot no.98 sector-3, Airoli, Navi Mumbai

Academic Year 2019-20

Datta Meghe College of Engineering

Airoli, Navi Mumbai

CERTIFICATE

This is to certify that the project entitled "**Detecting Crime Scenes Using ML**" is bona fide work of "**Drishti Jalgaonkar, Juilee Gund, Neha Patil** " submitted to the University of Mumbai in partial fulfilment of the requirement for the award of the degree of "Undergraduate" in "Computer Engineering".

Prof. M.M.Phadke	Prof. A.P.Pande	Dr.S.D.Sawarkar
Project Guide	Head of the Department	Principal

Datta Meghe College of Engineering

Airoli, Navi Mumbai

Project Approval

This project report entitled "**Detecting Crime Scenes Using ML**" of the students "**Drishti Jalgaonkar, Juilee Gund, Neha Patil**" approved for the degree of Computer Engineering.

Internal Examiner	External Examiner
Date:	Date:
Place:	Place:

Declaration

We declare that, this written submission represents our ideas in our own words and where others' ideas or words have been included; we have adequately cited and referenced the original sources. we also declare that, we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. we understand that any violation of the above will be cause for disciplinary action by the institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Name of the Students Signature

Drishti Jalgaonkar

Juilee Gund

Neha Patil

Acknowledgement

Motivation and guidance are the keys towards success. I would like to extend my thanks to all the sources of motivation.

We would like to grab this opportunity to thank D**r. S.D.Sawarkar,** principal for encouragement and support he has given for our project.

we express our deep gratitude to **Prof. A.P.Pande**, head of the department who has been the constant driving force behind the completion of this project.

We wish to express our heartfelt appreciation and deep sense of gratitude to my project guide **Prof. M.M.Phadke** for her encouragement, invaluable support, timely help, lucid suggestions and excellent guidance which helped us to understand and achieve the project goal. his/her concrete directions and critical views have greatly helped us in successful completion of this work.

We extend our sincere appreciation to all professors for their valuable inside and tip during the designing of the project. their contributions have been valuable in so many ways that we find it difficult to acknowledge of them individually.

We are also thankful to all those who helped us directly or indirectly in completion of this work.

Place:
Date:

Name of the Student
Drishti Jalgaonkar
Juilee Gund
Neha Patil

Abstract

Face detection and recognition from an image or a video is a popular topic in biometric research. Face recognition technology has widely attracted attention due to its enormous application value and market potential such as real time video surveillance system .It is widely acknowledged that the face recognition has played important role in surveillance system as it doesn't need object co-operation. Using real-time object detection to improve surveillance methods is a promising application of Convolution Neural Networks(CNNs).One particular application is a detection of hand-held weapons(such as pistols and rifles).Thus far previous work has mostly focused on weapon based detection within infrared data for concealed weapons. By contrast, we are particularly interested in the rapid detection ad identification of weapons from image and surveillance data. Also in this work we investigate the automatic detection of fire pixel regions in video imagery within real time bounds without reliance of temporal scene information.

Index

List of Figures

Chapter 1
Introduction

1.1 Motivation

The role of computers has been increased in all walks of life from the finance sector to supermarkets. In recent years police forces have been enhancing their traditional method of crime reporting with new technological advancements to increase their output by efficiently recording crimes to aid their investigation.

Human reasoning fails when presented with millions of records. Therefore, there is clearly a requirement for a tool kit to assist in analyzing the data which will make the best use of limited resources.

The aim of this study is to examine the current techniques used in crime prediction. Using real-time object detection to improve surveillance methods is a promising application of Convolutional Neural Networks.

- The objective of our project is to overcome limitations in human security.
- Normal and Abnormal Event classification in video sequence
- Automation of these tasks in video surveillance
- Once you recognize a violent situation, the assigned authority will be resolving the problems and take appropriate actions.

1.2 Problem Definition

Artificial Intelligence (AI) is proving its powers to prevent and detect everything gripping them from routine employee theft, frauds, insider trading and business risks. Many large corporations, business enterprises have been employing AI to detect and prevent money laundering and widespread frauds. Machine learning has been increasingly deployed by social media platforms to block illicit content such as child pornography and fake news. Businesses have been using AI for higher risk management and responsive fraud detection towards prevention and prediction of crimes.[3]

The earlier monitoring systems used by the industries need manual interference and are often not cent percent accurate. For instance, banks have been using transaction monitoring systems for years which are based on predefined binary rules that had to be manually checked involving time and inaccuracy. On average, only 2 percent of the transactions flagged by the software indicated true crime or malicious intent through manual checking. On the contrary, the modern machine-learning solutions use predictive rules that point out anomalies in datasets. These advanced algorithms successfully have decreased the number of false alerts by filtering out cases flagged incorrectly, while adding others missed using conventional rules.[3]

Chapter 2
2.1 Literature Review

The use of Artificial Intelligence to prevent crimes and aid in sentencing criminals may seem like a scene from a science fiction movie. However, police departments in the United Kingdom - including Durham, Kent, and South Wales - are already using facial recognition and behavioral software to prevent a crime before it occurs. Computer-driven evaluation frameworks are being used to inform custodial and sentencing decisions. The technology offers both huge promise and the prospect of dark dystopia in seemingly equal measure. [4]

2.1.1. Capturing the Relative Distribution of Features for Action Recognition Abstract:

Action recognition in natural video argue that both shape and motion information are necessary for recognition. They often combine static image features and spatiotemporal features to capture appearance and motion. It has, however, been shown in psychology experiments on Point Light Displays that it is possible to recognize hu- man actions based entirely on the dynamics of body movement. This gives motivation for our approach. Our approach is further motivated by the fact that, the appearance of subjects or background is less important to the description of actions than dynamics. In this paper, we investigate the sole use of body dynamics for automatic action recognition in complex videos. We propose a novel representation of actions in video, which captures the relative distribution of motion-based interest points by encoding their local spatiotemporal configuration in an efficient manner. This results in an action descriptor, which, in vectorized form, can be learnt using SVMs. [5]

2.1.2 Detecting emotions from everyday body movements. Abstract:

Describes a novel framework for analyzing every day or non-stylized body motion in order to detect affect. This is very different from analyzing stylized body motions. In a stylized motion the entirety of the movement encodes a particular emotion. Stylized motions normally originate from laboratory settings, where subjects are asked to freely act an emotion without any constraints. They also arise from stylized dance. This paper, however, concerns itself with the more subtle aspects of human Movement. We will examine how affect is communicated by the manner, in which everyday actions are performed. We hope that this work will open up a multitude of opportunities for intelligent human-machine interaction which is not viable with approaches that assume stereotypical, stylized body motions.[7]

2.1.3. Developing a Real-Time Gun Detection Classifier Abstract:

Today, Closed Circuit Television (CCTV) is used as monitoring and surveillance tool for fighting crimes. The aim of CCTV is to reduce the crime and social offence by monitoring the scene under surveillance. Its use varies from user to user. For example, CCTV is being used in street surveillance for monitoring various activities like finding missing person, identifying anti-social behavior, drug misuse etc1. It is being also used for capturing evidence of crime and presenting evidences to courts for prosecution. A CCTV involves the captures the video and transmits it to the television screen of base station that is monitored by use of an unmanned

and remotely mounted camera and an operator. A CCTV camera the operator to detect suspicious activity or to capture evidences. But, the detection of suspicious activity is proportional to the operator attention on each video feed of screen. So, it is not possible for a CCTV operator to effectively monitor each activity of video feeds all the time with complete attention due to low operator to screen ratio, concurrent run of multiple video feeds at same screen2 and environmental condition of operational room. There is a possibility for them to miss the detection of some abnormal activity. According to Velastinetal.16, A CCTV operator suffers from video blindness after 20 to 40 minutes of active monitoring due to which he is not able to recognize objects in video feeds. According to a study published in Security Oz Magazine17 an operator will often miss up to 45 of screen activity after 12 minutes of continuous monitoring and this miss rate increase up to 95 after. 22 minutes. So, to make the CCTV monitoring and surveillance efficient, there is a need to automate the detection of suspicious activity in video material which in turn reduces the operator overload. The automated system will raise an alarm if any abnormal activity takes place under CCTV surveillance, due to which the operator will allocate his attention on video feed and will take appropriate action.[8]

2.1.4. Recognizing Human Emotions from Body Movement and Gesture Dynamics.

Abstract:

In this paper we present an approach for the recognition of four acted emotional states (anger, joy, pleasure, sadness) based on the analysis of body movement and gesture expressivity. According to research showing that distinct emotions are often associated with different qualities of body movement, we use non-propositional movement qualities (e.g. amplitude, speed and fluidity of movement) to infer emotions, rather than trying to recognize different gesture shapes expressing specific emotions, to investigate the role of movement expressivity versus shape in gesture. We propose a method for the analysis of emotional behavior based on direct classification of time series and on a model that provides indicators describing the dynamics of expressive motion cues. Finally, we show and interpret the recognition rates for both proposals using different classification algorithms.[2][9]

2.1.5. Human activity recognition based on pose points selection.

Abstract:

A novel method for human action recognition is proposed in this paper. Traditional spatial-temporal interest point detectors are easily affected by hair, face, shadow, clothes texture or the shake of camera. Inspired by the use of points distribution information, we pro-pose a point selection method to select representative points (denoted by the pose points), which use HOG human detector and contour detector to select the points on human pose edges. The pose points carry both local gradient information and global pose information. 3D-SIFT scale selection method and novel descriptors called body scale and motion intensity feature are also studied. The descriptors calculate the width scale of different levels of human body and count motion intensity of activity in five directions. The descriptors combine spatial location with the moving intensity together and are used for further classification with SVMs.

2.1.6. A Local Feature based on Lagrangian Measures for Violent Video Classification

Abstract: In this paper we want to utilize the concept of Lagrangian measures for the violent video detection task. Following the Lagrangian framework proposed in we make use of the path line concept as the in- tegral field lines in the unsteady flow field. From the path lines we derive a Lagrangian measure, the direction Lagrangian fields. These fields are similar to the optical flow field as they contain also two- directional motion-like components that are integrated over their path lines. These fields are meant to represent the dynamic patterns in the scene related to a given time scale. To encode spatio temporal patterns, inspired by the MoSIFT algorithm, we will extend the SIFT.

2.1.7. Experimentally defined convolutional neural network architecture variants for non-temporal real time fire-detection.

Abstract:
In this work we investigate the automatic detection of fire pixel regions in video (or still) imagery within real-time bounds without reliance on temporal scene information. As an extension to prior work in the field, we consider the performance of experimentally defined, reduced complexity deep convolutional neural network architectures for this task. Contrary to contemporary trends in the field, our work illustrates maximal accuracy of 0.93 for whole image binary fire detection, with 0.89 accuracy within our super pixel localization framework can be achieved, via a network architecture of significantly reduced complexity. These reduced architectures additionally offer a 3-4 fold increase in computational performance offering up to 17 fps processing on contemporary hardware independent of temporal information. We show the relative performance achieved against prior work using benchmark datasets to illustrate maximally robust real-time fire region detection.

2.2 SUMMARIZED FINDING

- Pose estimation refers to computer vision techniques that detect human figures in images and videos.
- The analysis of emotional behavior based on direct classification of time series and on a model that provides indicators describing the dynamics of expressive motion cues.
- The appearance of subjects or background is less important to the description of actions than dynamics
- The aim of this project is to detect crime scenes and avoid crimes happening in real world, hence this surveillance system is developed that detects motion and to respond speedily by capturing an image and relaying it to an administrator device through the internet platform.

Chapter 3

Software Requirements Specification

3.1 Functional Requirements

3.1.1 Rapid Database Scanning

Although law enforcement networks are certainly more connected and streamlined than in the past, there's massive room for improvement in terms of automation.

3.1.2 Facial Recognition

Once your face is in the system, advanced algorithms can recognize you with remarkable precision. Facial key points are detected and the best facial recognition application can identify you with up to 97 accuracy.

3.2 Non-Functional Requirements

3.2.1 Performance Requirement

3.2.1.1 System can produce results faster on 4GB of RAM. It may take more time for peak loads at main node.

3.2.1.2 The system will be available 100 % of the time. Once there is a fatal error, system will provide understandable feedback to the user.

3.2.2 Safety and Security Requirement

3.2.2.1 The system is designed in modules where errors can be detected

3.2.3 Software Quality Attribute's

3.2.3.1 **Reliability:** The Client machine will change the status of data indicating successful data transmission.

3.2.3.2 **Usability:** The application should be easy to use through interactive inter- face.

3.2.3.3 **Maintainability:** The system will be developed using the standard soft- ware development conventions to help in easy review and redesigning of the system.

Software and Hardware Requirements

3.3 Software Requirements

3.3.1 **Operating System:** Windows 7 and Above

3.3.2 **Programming Language:** Python

3.3.3 **IDE:** PYCHARM, OPENCV

3.3.4 **Database:** MySQL

3.4 Hardware Requirements

3.4.1 **System:** Processor i3 and above

3.4.2 **Hard Disk:** 10 GB

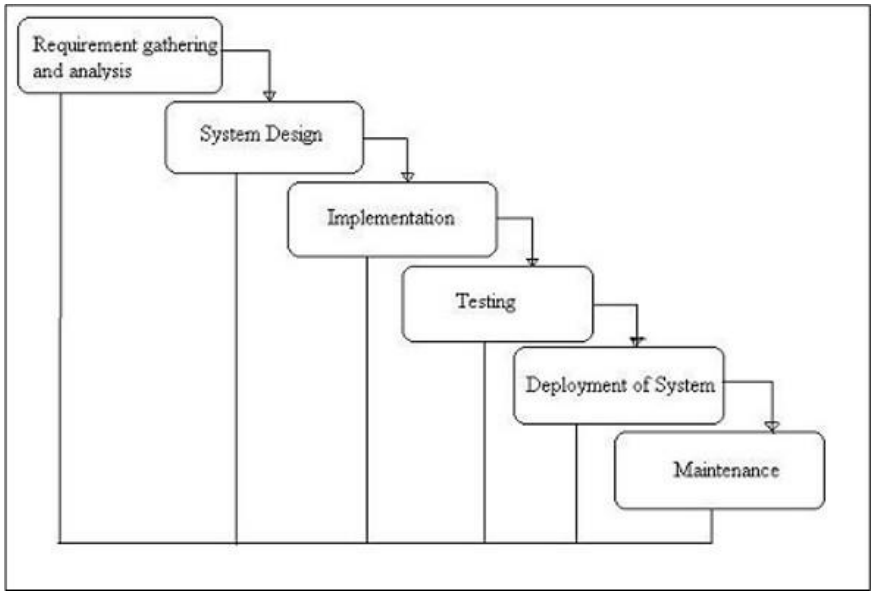
3.4.3 **RAM:** 4GB

3.4.4 **Camera:** 2 MP minimum

3.5 SDLC Model

Waterfall Model

The Waterfall Model was first Process Model to be introduced. It is also referred to as a linear-sequential life cycle model. It is very simple to understand and use. In a waterfall model, each phase must be completed fully before the next phase can begin. This type of model is basically used for the for the project which is small and there are no uncertain requirements. At the end of each phase, a review takes place to determine if the project is on the right path and whether or not to continue or discard the project. In this model the testing starts only after the development is complete. In waterfall model phases do not overlap.[12]

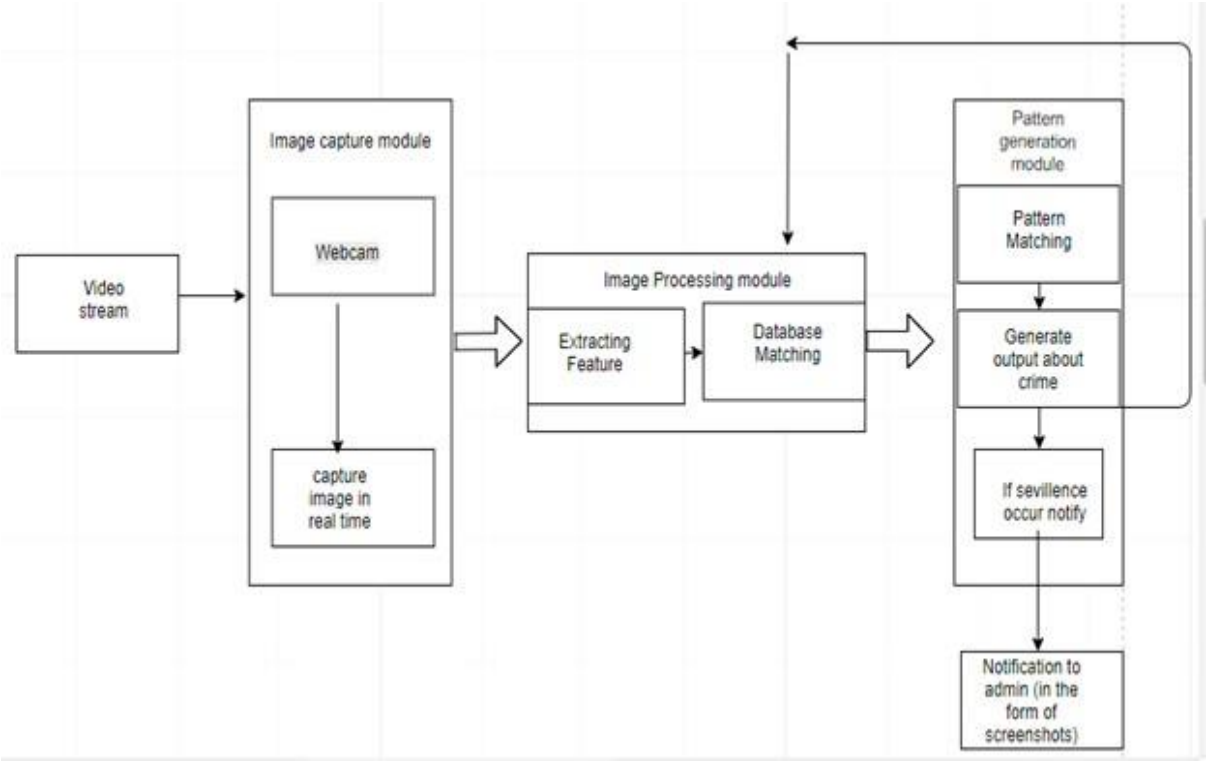


Waterfall Model

Chapter 4

Project Design

4.1 System Architecture



Firstly, Image Will be captured from webcam or any IP Cam. In database, certain images will be stored. The image captured by webcam will be matched with the image in database if match found, it will send screenshot along with the name of person that is stored in database. If the information of person is not stored in database, it will only send the screenshot of that image to the admin. Image processing mechanism is used for proccesing. It will also detect person in the video. It will divide video in number of images and will match with each different image. And lastly generate an alarm to the admin.

System Modules

4.1.1 Image Recognition:

Image recognition through photo Image recognition, in the context of machine vision, is the ability of software to identify objects, places, people, writing and actions in images. Computers can use machine vision technologies in combination with a camera and artificial intelligence software to achieve image recognition.

4.1.2 Video Recognition:

Video recognition through photo. In this section we will learn different techniques to work with videos like object tracking etc.

4.1.3 Fire and Smoke Detection:

Fire is a disaster that can strike anywhere and can be very destructive. A method to detect smoke and fires would allow the authorities to detect and put out the fires before it becomes out of control. One of the cost effective methods would be to use cameras which are already on the roads to detect the fires early in order to inform the relevant parties. This project suggests a method to use surveillance cameras in order to monitor occurrences of fire anywhere within camera range. Since cameras are already installed in most places, this method would be cost effective as there would be no need to purchase and install the hardware required. The response can also be quicker due to information being gathered early. An algorithm that can process videos in order to detect smoke is developed in this project. Using properties of smoke, such as its colour and seemingly random motion, the video is processed, and elements which match the properties are determined to be smoke. The results of this method show that visible smoke can be reliably detected in relatively bright or well lighted areas, such as indoors and on highway roads. Although the performance of the algorithm degrades on darker areas, as smoke is not so visible and difficult to trace, this can be overcome by coupling it with another fire detection algorithm.

4.1.4 Weapon Detection:

While action recognition has become an important line of research in computer vision, the recognition of particular events such as aggressive behaviors, or fights, has been relatively less studied. These tasks may be extremely useful in several video surveillance scenarios such as psychiatric wards, prisons or even in personal camera smartphones. Their potential usability has led to a surge of interest in developing fight or violence detectors. One of the key aspects in this case is efficiency, that is, these methods should be computationally fast.

Handcrafted spatiotemporal features that account for both motion and appearance information can achieve high accuracy rates, albeit the computational cost of extracting some of those features is still prohibitive for practical applications.

4.2 Use Case Diagram

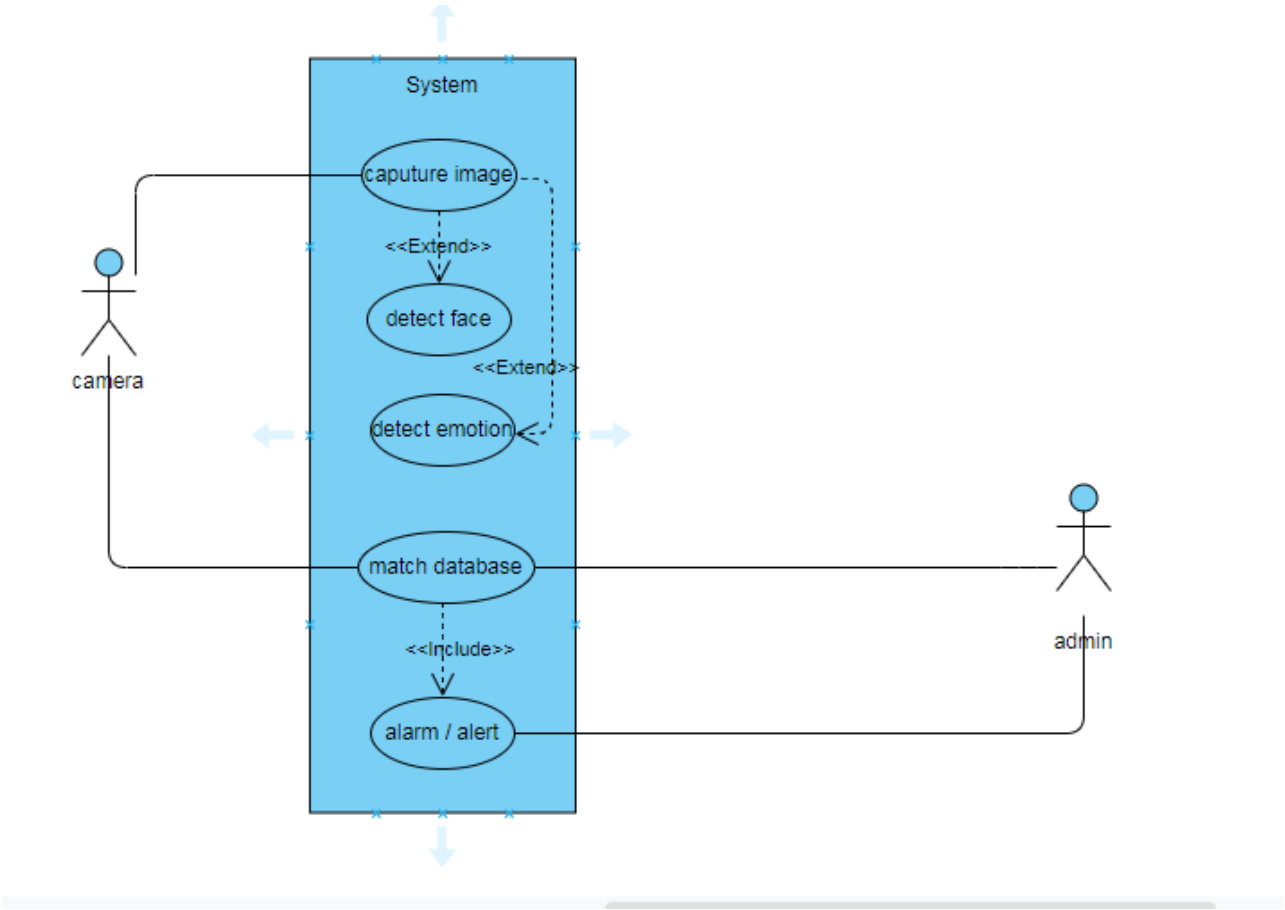


Figure 4.2: Use Case Diagram

4.3 Class Diagram

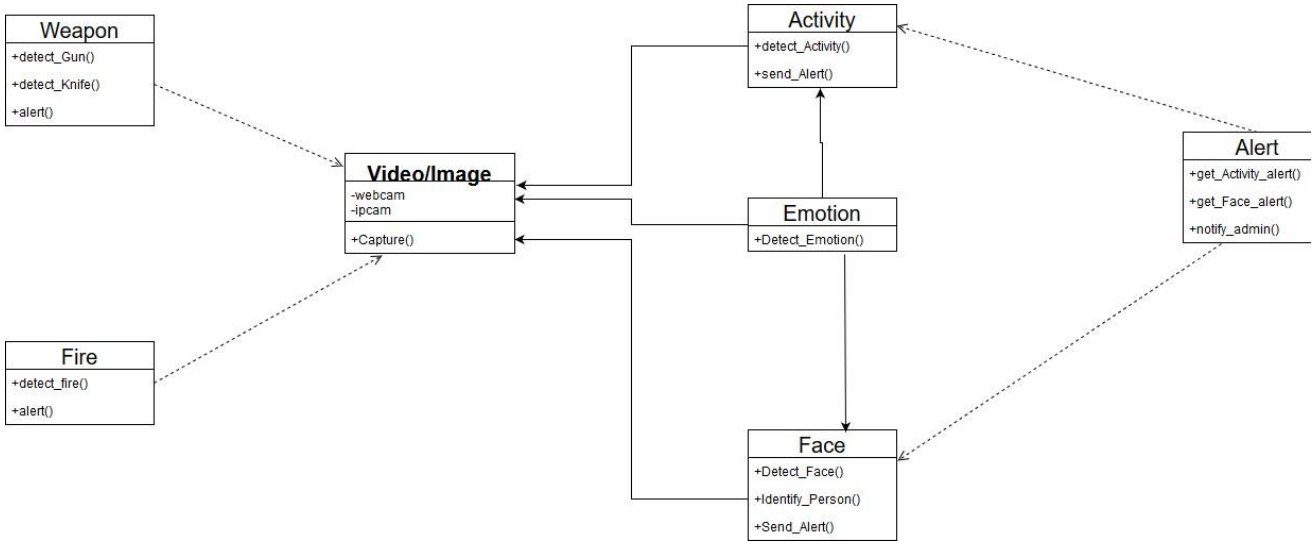


Figure 4.3: Class Diagram

4.4 Sequence Diagram

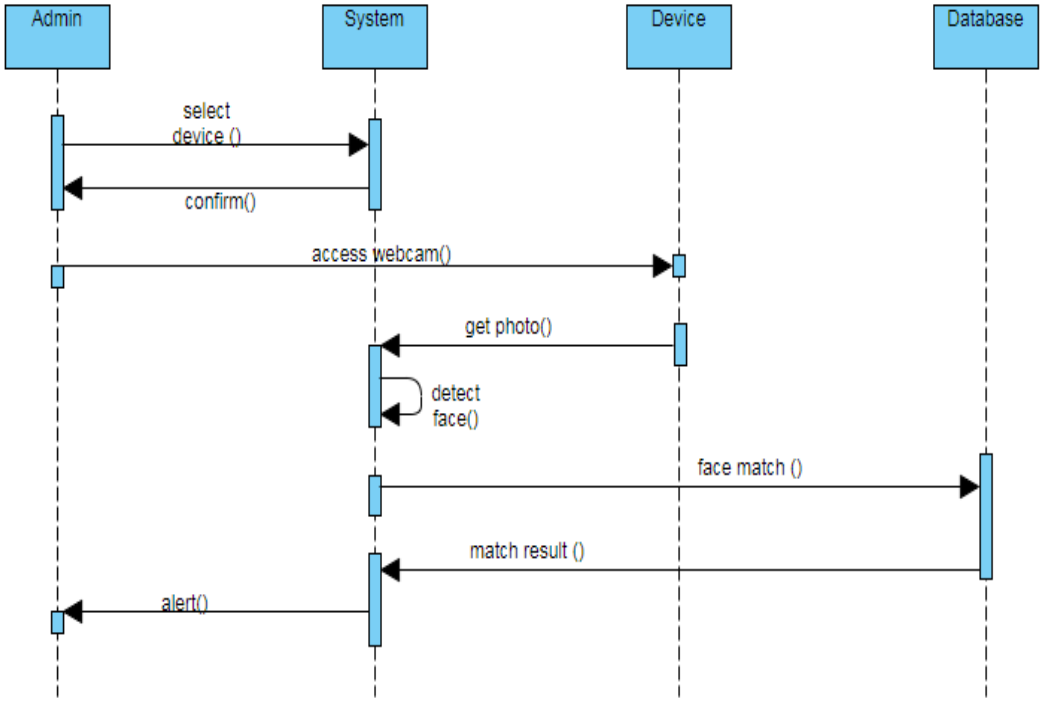


Figure 4.4: Sequence Diagram

4.5 Data Flow Diagram

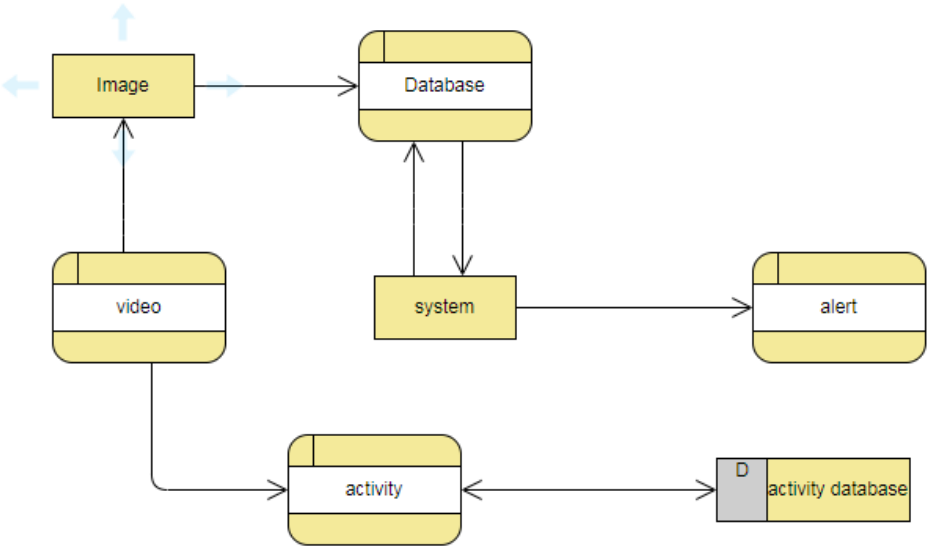


Figure 4.5.1: DFD (Level 0)

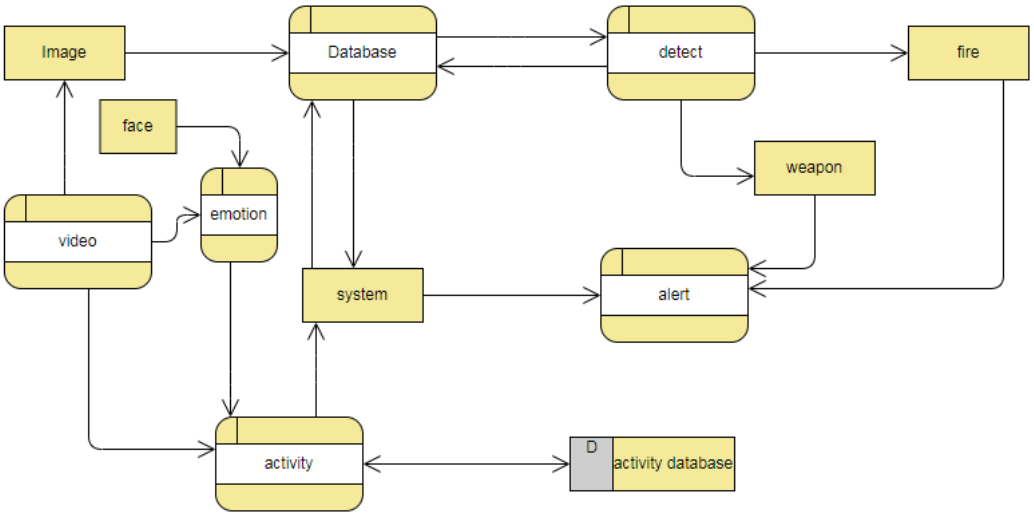


Figure 4.5.2: DFD (Level 1)

4.6 Deployment Diagram

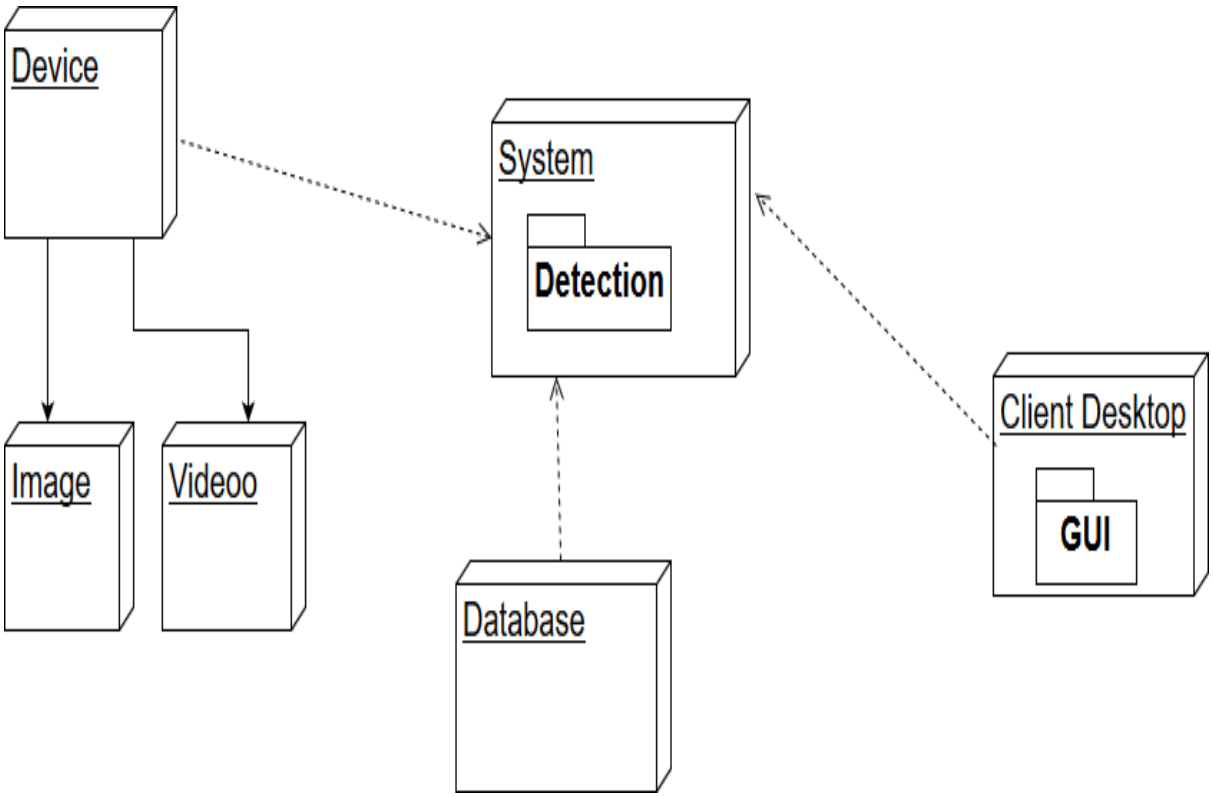


Figure 4.6: Deployment Diagram

Chapter 5

Technologies Used

5.1 Language Used:

PyCharm

PyCharm is an integrated development environment (IDE) used in computer programming, specifically for the Python language. It is developed by the Czech company JetBrains. It provides code analysis, a graphical debugger, an integrated unit tester, integration with version control systems, and supports web development with Django.

PyCharm is cross-platform, with Windows, macOS and Linux versions. The Community Edition is released under the Apache License, and there is also Professional Edition with extra features, released under a proprietary license. PyCharm provides API so that developers can write their own plugins to extend PyCharm features. Several plugins from other JetBrains IDE also work with PyCharm. There are more than 1000 plugins which are compatible with PyCharm.

PyCharm was mainly developed for emphasis on code readability, and its syntax allows programmers to express concepts in fewer lines of code. Python is a programming language that lets you work quickly and integrate systems more efficiently. We need to have an interpreter to interpret and run our programs. There are certain online interpreters like GFG-IDE, IDEONE or CodePad, etc. Running Python codes on an offline interpreter is much more compatible than using an online IDE.

PyCharm is one of the most popular Python-IDE developed by JetBrains used for performing scripting in Python language. PyCharm provides some very useful features like Code completion and inspection, Debugging process, support for various programming frameworks such as Flask and Django, Package Management, etc. PyCharm provides various tools for productive development in Python.

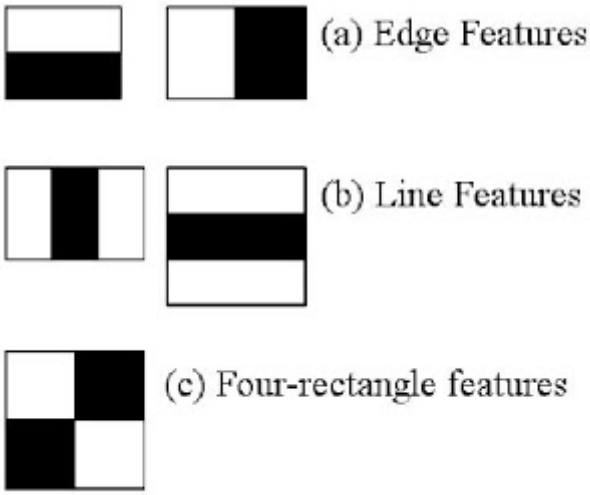
5.2 Algorithms Used:

5.2.1 Haar Cascade Algorithm:

A facial identification system is a technology capable of identifying a face of a person from a digital image or a video frame from a video source. Haar Cascade classifier is based on the Haar Wavelet technique to analyse pixels in the image into squares by function. This uses “integral image” concepts to compute the “features” detected. Haar Cascades use the AdaBoost learning algorithm which selects a small number of important features from a large set to give an efficient result of classifiers then use cascading techniques to detect face in an image. Haar cascade classifier is based on Viola Jones detection algorithm which is trained in given some input faces and non-faces and training a classifier which identifies a face.

It is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. The algorithm needs a lot of positive images of faces and negative images without faces to train the classifier. Then we need to extract features from it.

A Haar feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. Integral Images are used to make this super fast.



Adaboost selects the best features and trains the classifiers. This algorithm constructs a “strong” classifier as a linear combination of weighted simple “weak” classifiers. During the detection phase, a window of the target size is moved over the input image, and for each subsection of the image and Haar features are calculated. This difference is then compared to a learned threshold that separates non-objects from objects. Because each Haar feature is only a "weak classifier" (its detection quality is slightly better than random guessing) a large number of Haar features are necessary to describe an object with sufficient accuracy and are therefore organized into cascade classifiers to form a strong classifier.

The cascade classifier consists of a collection of stages, where each stage is an ensemble of weak learners. The weak learners are simple classifiers called decision stumps. Each stage is trained using a technique called boosting. Boosting provides the ability to train a highly accurate classifier by taking a weighted average of the decisions made by the weak learners. Each stage of the classifier labels the region defined by the current location of the sliding window as either positive or negative. Positive indicates that an object was found and negative indicates no objects were found.

5.2.2 Convolutional Neural Network Algorithm:

In neural networks, Convolutional neural network (ConvNets or CNNs) is one of the main categories to do images recognition, images classifications. Objects detections, recognition faces etc., are some of the areas where CNNs are widely used.

Convolutional neural network performs very well in the area of object classification. This network has the ability to perform feature extraction and classification within the same

architecture. Tested on real video sequences, the proposed approach achieves better classification performance as some of relevant conventional video fire detection methods and

indicates that using CNN to detect fire in videos is very promising. In conventional fire detection, much research has continuously focused on finding out the salient features of fire images. After analyzing the changes of fire using an RGB and HSI color model based on the difference between consecutive frames and proposed a rule-based approach for fire decision. In addition, Wang extracted the candidate fire area in an image using an HSI color model and calculated the dispersion of the flame color to determine the fire area. However, color-based fire detection methods are generally vulnerable to a variety of environmental factors such as lighting and shadow.

5.3.3 YOLO Algorithm:

YOLO: Real-Time Object Detection. You only look once (YOLO) is a system for detecting objects on the Pascal VOC 2012 dataset. It can detect the 20 Pascal object classes: person. Compared to other region proposal classification networks (fast RCNN) which perform detection on various region proposals and thus end up performing prediction multiple times for various regions in a image, Yolo architecture is more like FCNN (fully convolutional neural network) and passes the image (nxn) once through the FCNN and output is (mxm) prediction. This the architecture is splitting the input image in mxm grid and for each grid generation 2 bounding boxes and class probabilities for those bounding boxes. Note that bounding box is more likely to be larger than the grid itself.

A single convolutional network simultaneously predicts multiple bounding boxes and class probabilities for those boxes. YOLO trains on full images and directly optimizes detection performance. This unified model has several benefits over traditional methods of object detection. First, YOLO is extremely fast. Since we frame detection as a regression problem we don't need a complex pipeline. We simply run our neural network on a new image at test time to predict detections. Our base network runs at 45 frames per second with no batch processing on a Titan X GPU and a fast version runs at more than 150 fps. This means we can process streaming video in real-time with less than 25 milliseconds of latency.

Second, YOLO reasons globally about the image when making predictions. Unlike sliding window and region proposal-based techniques, YOLO sees the entire image during training and test time so it implicitly encodes contextual information about classes as well as their appearance. Fast R-CNN, a top detection method, mistakes background patches in an image for objects because it cant see the larger context. YOLO makes less than half the number of background errors compared to Fast R-CNN.

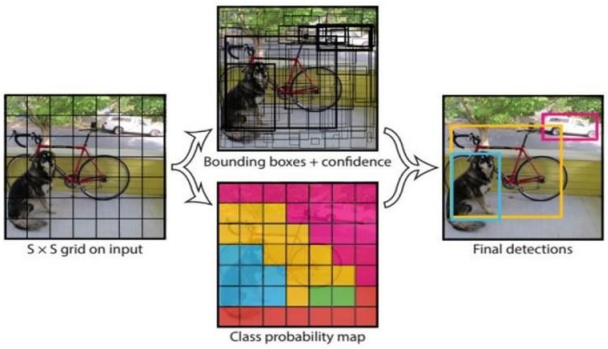


Figure 7.2: YOLO Process

Our network uses features from the entire image to predict each bounding box. It also predicts all bounding boxes across all classes for an image simultaneously. This means our network reasons globally about the full image and all the objects in the image. The YOLO design enables end-to- end training and real time speeds while maintaining high average precision. Our system divides the input image into an S S grid. If the center of an object falls into a grid cell, that grid cell is responsible for detecting that object. Each grid cell predicts B bounding boxes and confidence scores for those boxes. These confidence scores reflect how confident the model is that the box contains an object and also how accurate it thinks the box is that it predicts. Formally we define confidence as Pr(Object) IOU. If no object exists in that cell, the confidence scores should be zero. Otherwise we want the confidence score to equal the intersection over union (IOU) between the predicted box and the ground truth.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

This proposed, a technique for abnormal event detection based in a video surveillance system. The proposed method pre-trains deep spatiotemporal networks over unrelated dataset and shows promising capability in abnormal event detection of small sample problem. This project is focused on developing a surveillance system that detects motion and to respond speedily by capturing an image and relaying it to an administrator device through the internet platform. It will analyze video input and apply different image processing techniques on it. After analysis of video stream, it will use different Algorithm and techniques to detect any anomaly within surveillance system and notify to the admin.

6.2 Future Work

The task of fight recognition has attracted interest of many researchers in the last few years.3D convolutional Neural Network that processes the entire video sequence as input. Given a video, automatically classify what action is performed in the video. The task of recognition has attracted interest of many researchers in the last few years. Given the direct applications in surveillance and movie rating, the efforts are fully justised. Handcrafted features, some used for generic action recognition work, were used for this task before the advent of feature learning methods This proposed, a technique for abnormal event detection based in a video surveillance system. The proposed method pre-trains deep spatiotemporal networks over unrelated dataset and shows promising capability in abnormal event detection of small sample problem. This project is focused on developing a surveillance system that detects motion and to respond speedily by capturing an image and relaying it to an administrator device through the internet platform.

References

1. Learning spatiotemporal features with 3d convolutional networks,D. Tran,L. Bourdev, R. Fergus, L. Torresani, and M. Paluri.. In 2015 IEEE International Conference on Computer Vision (ICCV), pages 44894497. IEEE, 2015.

2.Recognising Human Emotions from Body Movement and Gesture Dynamics,] Ginevra Castellano1, Santiago D. Villalba2, University of Genoa

3. https://www.analyticsinsight.net/detecting-crime-through-artificial-intelligence/

4.https://www.forbes.com/sites/nikitamalik/2018/11/26/how-can-we-use-artificial-intelligence-to-prevent-crime/#bdbf87d498cd

5. https://books.google.co.in/books?id=xqbdxBtTY_EC&pg=PA49&lpg=PA49&dq

6.Human activity recognition based on pose points selection,Ke Xu, Xinghao Jiang, Tanfeng Sun,: 2015 IEEE International Conference on Image Processing (ICIP)

7. https://www.cl.cam.ac.uk/~pr10/publications/presenccia07.pdf

8. http://cs231n.stanford.edu/reports/2017/pdfs/716.pdf

9. https://link.springer.com/chapter/10.1007/978-3-540-74889-2_7

10. https://ieeexplore.ieee.org/document/7351339

11. https://ieeexplore.ieee.org/document/8375994

12. https://www.tutorialspoint.com/sdlc/sdlc_waterfall_model.html

13. https://en.wikipedia.org/wiki/PyCharm

14. https://link.springer.com/chapter/10.1007%2F3-540-44673-7_7